Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 35, 1–14.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP301

Wasserstein-2 bounds in normal approximation under local dependence*

Xiao Fang†

Abstract

We obtain a general bound for the Wasserstein-2 distance in normal approximation for sums of locally dependent random variables. The proof is based on an asymptotic expansion for expectations of second-order differentiable functions of the sum. We apply the main result to obtain Wasserstein-2 bounds in normal approximation for sums of m-dependent random variables, U-statistics and subgraph counts in the Erdős-Rényi random graph. We state a conjecture on Wasserstein-p bounds for any positive integer p and provide supporting arguments for the conjecture.

Keywords: central limit theorem; local dependence; Erdős-Rényi random graph; Stein's method; U-statistics; Wasserstein-2 distance.
AMS MSC 2010: 60F05.
Submitted to EJP on February 1, 2019, final version accepted on March 24, 2019.

1 Introduction

For two probability measures μ and ν on $\mathbb{R}^d$, the Wasserstein-p distance, $p \geq 1$, is defined as

$$\mathcal{W}_p(\mu, \nu) = \Big(\inf_{\pi \in \Gamma(\mu,\nu)} \int |x-y|^p d\pi(x,y) \Big)^{\frac{1}{p}},$$

where $\Gamma(\mu, \nu)$ is the space of all probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with μ and ν as marginals and $|\cdot|$ denotes the Euclidean norm. Note that $\mathcal{W}_p(\mu,\nu) \leq \mathcal{W}_q(\mu,\nu)$ if $p \leq q$. For a random vector W whose distribution is close to ν, it is of interest to provide an explicit upper bound on their Wasserstein-p distance. See, for example, [10], [3], [15], [4] and [8] for a recent wave of research in this direction.

We consider the central limit theorem in dimension one where μ is the distribution of a random variable W of interest, $\nu = N(0,1)$ and $d = 1$ in the above setting. A large class of random variables that can be approximated by a normal distribution exhibits a *local dependence* structure. Roughly speaking, with details deferred to Section 2.1,

*Supported by Hong Kong RGC ECS 24301617, a CUHK direct grant and a CUHK start-up grant.
†The Chinese University of Hong Kong, Hong Kong. E-mail: xfang@sta.cuhk.edu.hk

we assume that the random variable W is a sum of a large number of random variables $\{X_i : i \in I\}$ and that each X_i is independent of $\{X_j : j \notin A_i\}$ for a relatively small index set A_i. Barbour, Karoński and Ruciński [2] obtained a Wasserstein-1 bound in the central limit theorem for such W and Chen and Shao [6] obtained a bound for the Kolmogorov distance. We refer to these two papers for a number of interesting applications.

To prove their Wasserstein-1 bound, Barbour, Karoński and Ruciński [2] used Stein's method and the following equivalent definition of the Wasserstein-1 distance:

$$\mathcal{W}_1(\mu, \nu) = \sup_{h \in \mathrm{Lip}_1(\mathbb{R})} \Big| \int_{\mathbb{R}} h d\mu - \int_{\mathbb{R}} h d\nu \Big|,$$

where $\mathrm{Lip}_1(\mathbb{R})$ denotes the class of Lipschitz functions with Lipschitz constant 1. There seems to be no such expression for $\mathcal{W}_p$ for general p. The optimal Wasserstein-p bound in normal approximation for sums of independent random variables (cf. Lemma 3.4) was only recently obtained by Bobkov [3] using characteristic functions. Our main result, Theorem 2.1, provides a Wasserstein-2 bound in normal approximation under local dependence, which is a generalization of independence. We also state a conjecture on Wasserstein-p bounds for any positive integer p.

To prove our main result, we follow the approach of Rio [12], who used the asymptotic expansion of Barbour [1] and a Poisson-like approximation to obtain a Wasserstein-2 bound in normal approximation for sums of independent random variables. We first use Stein's method to obtain an asymptotic expansion for expectations of second-order differentiable functions of the sum of locally dependent random variables W. We then use this expansion and the upper bound for the Wasserstein-2 distance in terms of Zolotarev's ideal distance of order 2 to control the Wasserstein-2 distance between the distributions of W and a sum of independent and identically distributed (i.i.d.) random variables. Finally, we use the triangle inequality and known Wasserstein-2 bounds in normal approximation for sums of i.i.d. random variables to prove our main result. This approach enables us to potentially bound the Wasserstein-p distance for any positive integer p.

We apply our main result to the central limit theorem for sums of m-dependent random variables, U-statistics and subgraph counts in the Erdős-Rényi random graph.

The paper is organized as follows. Section 2 contains the Wasserstein-2 bound in normal approximation under local dependence, the applications and the conjecture on Wasserstein-p bounds. Section 3 contains some related literature, the proofs of the results in Section 2 and supporting arguments for the conjecture. In the following, we use C to denote positive constants independent of all other parameters, possibly different from line to line.

2 Main results

In this section, we provide a general Wasserstein-2 bound in normal approximation under local dependence and apply it to the central limit theorem for sums of m-dependent random variables, U-statistics and subgraph counts in the Erdős-Rényi random graph. We also state a conjecture on Wasserstein-p bounds.

2.1 A Wasserstein-2 bound under local dependence

Let $W = \sum_{i \in I} X_i$ for an index set I with $\mathbb{E}X_i = 0, \mathbb{E}W^2 = 1$ and satisfies the following *local dependence* structure:

(LD1): For each $i \in I$, there exists $A_i \subset I$ such that X_i is independent of $\{X_j : j \notin A_i\}$.

(LD2): For each $i \in I$ and $j \in A_i$, there exists $A_{ij} \supset A_i$ such that $\{X_i, X_j\}$ is independent of $\{X_k : k \notin A_{ij}\}$.

(LD3): For each $i \in I$, $j \in A_i$ and $k \in A_{ij}$, there exists $A_{ijk} \supset A_{ij}$ such that $\{X_i, X_j, X_k\}$ is independent of $\{X_l : l \notin A_{ijk}\}$.

Assume that $\beta := \mathbb{E}W^3$ exists.

Theorem 2.1. *Under the above setting, we have*

$$\mathcal{W}_2(\mathcal{L}(W), N(0,1)) \le C\big[|\beta| + (\gamma_1 + \gamma_2 + \gamma_3)^{\frac{1}{2}}\big], \tag{2.1}$$

where

$$\beta = \sum_{i\in I}\sum_{j,k\in A_i} \mathbb{E}X_iX_jX_k + 2\sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}\setminus A_i} \mathbb{E}X_iX_jX_k,$$

$$\gamma_1 = \sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}}\sum_{l\in A_{ijk}} \mathbb{E}|X_iX_jX_kX_l|,$$

$$\gamma_2 = \sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}}\sum_{l\in A_{ijk}} \mathbb{E}|X_iX_j|\mathbb{E}|X_kX_l|,$$

$$\gamma_3 = \sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}}\sum_{l\in A_{ijk}} \mathbb{E}|X_iX_jX_k|\mathbb{E}|X_l|.$$

Remark 2.2. The conditions (LD1)–(LD3) and the bound (2.1) represent a natural extension of (2.1)–(2.5) and (2.7) of [2]. The sizes of *neighborhoods* A_{ij} and A_{ijk} are typically smaller than those used in [6]. It would be interesting to prove a bound for the Kolmogorov distance under the above setting.

2.2 Applications

2.2.1 m-dependence

Let $X_1, \ldots, X_n$ be a sequence of m-dependent random variables, namely, $\{X_i : i \le j\}$ is independent of $\{X_i : i \ge j+m+1\}$ for any $j = 1, \ldots, n-m-1$. Let $W = \sum_{i=1}^n X_i$. Assume that $\mathbb{E}X_i = 0$ and $\mathbb{E}W^2 = 1$. We have the following corollary of Theorem 2.1.

Corollary 2.3. *For sums of m-dependent random variables as above, we have*

$$\mathcal{W}_2(\mathcal{L}(W), N(0,1)) \le C\Big\{m^2\sum_{i=1}^n \mathbb{E}|X_i|^3 + m^{3/2}(\sum_{i=1}^n \mathbb{E}X_i^4)^{1/2}\Big\}.$$

2.2.2 U-statistics

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables from a fixed distribution. Let $m \ge 2$ be a fixed integer. Let $h : \mathbb{R}^m \to \mathbb{R}$ be a fixed, symmetric, Borel-measurable function. We consider the Hoeffding [9] U-statistic

$$\sum_{1\le i_1<\cdots<i_m\le n} h(X_{i_1}, \ldots, X_{i_m}).$$

Assume that

$$\mathbb{E}h(X_1, \ldots, X_m) = 0,\ \mathbb{E}h^4(X_1, \ldots, X_m) < \infty,$$

and the U-statistic is non-degenerate, namely,

$$\mathbb{E}g^2(X_1) > 0,$$

where

$$g(x) := \mathbb{E}(h(X_1, \ldots, X_m)|X_1 = x).$$

Applying Theorem 2.1 to the U-statistic above yields the following result:

Theorem 2.4. *Under the above setting, let*

$$W_n = \frac{1}{\sigma_n} \sum_{1 \le i_1 < \cdots < i_m \le n} h(X_{i_1}, \ldots, X_{i_m}),$$

where

$$\sigma_n^2 = \mathrm{Var}\big[\sum_{1 \le i_1 < \cdots < i_m \le n} h(X_{i_1}, \ldots, X_{i_m})\big].$$

We have

$$\mathcal{W}_2(\mathcal{L}(W_n), N(0,1)) \le \frac{C}{\sqrt{n}}.$$

Remark 2.5. Chen and Shao [7] obtained a bound on the Kolmogorov distance in normal approximation for non-degenerate U-statistics. We refer to the references therein for a large literature on the rate of convergence in normal approximation for U-statistics. For simplicity, we assumed above that $\mathcal{L}(X_1)$, m and $h(\cdot)$ are fixed. They may be taken into account explicitly in the Wasserstein-2 bound. We omit the details.

2.2.3 Subgraph counts in the Erdős-Rényi random graph

Let $K(n,p)$ be the Erdős-Rényi random graph with n vertices. Each pair of vertices is connected with probability p and remain disconnected with probability $1-p$, independent of all else. Let G be a given fixed graph. For any graph H, let $v(H)$ and $e(H)$ denote the number of its vertices and edges, respectively. Theorem 2.1 leads to the following result.

Theorem 2.6. *Let S be the number of copies (not necessarily induced) of G in $K(n,p)$, and let $W = (S - \mathbb{E}S)/\sqrt{\mathrm{Var}(S)}$ be the standardized version. Then*

$$\mathcal{W}_2(\mathcal{L}(W), N(0,1)) \le C(G) \begin{cases} \psi^{-\frac{1}{2}} & \text{if } 0 < p \le \frac{1}{2} \\ n^{-1}(1-p)^{-\frac{1}{2}} & \text{if } \frac{1}{2} < p < 1, \end{cases} \tag{2.2}$$

where $C(G)$ is a constant only depending on G and

$$\psi = \min_{H \subset G, e(H) > 0} \{n^{v(H)} p^{e(H)}\}.$$

Remark 2.7. Barbour, Karoński and Ruciński [2] proved the same bound as in (2.2) for the weaker Wasserstein-1 distance. In the special case where G is a triangle, the bound in (2.2) reduces to

$$C \begin{cases} n^{-\frac{3}{2}} p^{-\frac{3}{2}} & \text{if } 0 < p \le n^{-\frac{1}{2}} \\ n^{-1} p^{-\frac{1}{2}} & \text{if } n^{-\frac{1}{2}} < p \le \frac{1}{2} \\ n^{-1}(1-p)^{-\frac{1}{2}} & \text{if } \frac{1}{2} < p < 1. \end{cases}$$

Röllin [13] proved the same bound for the Kolmogorov distance in this special case.

2.3 Conjecture on Wasserstein-p bounds

Here we state a conjecture on Wasserstein-p bounds for any positive integer p. We provide supporting arguments, including a complete proof for $p = 3$, for the conjecture at the end of the next section. Let $W = \sum_{i \in I} X_i$ for an index set I with $\mathbb{E}X_i = 0, \mathbb{E}W^2 = 1$ and satisfies (LD1)–(LD$(p+1)$) where

(LDm): For each $i_1 \in I, i_2 \in A_{i_1} \ldots, i_m \in A_{i_1 \ldots i_{m-1}}$, there exists $A_{i_1 \ldots i_m} \supset A_{i_1 \ldots i_{m-1}}$ such that $\{X_{i_1}, \ldots, X_{i_m}\}$ is independent of $\{X_j : j \notin A_{i_1 \ldots i_m}\}$.

Conjecture 2.8. *Under the above setting, we have*

$$\mathcal{W}_p(\mathcal{L}(W), N(0,1)) \le C_p \sum_{m=1}^{p} (R_m)^{\frac{1}{m}}, \tag{2.3}$$

where C_p is a constant only depending on p,

$$R_m = \sum_{i_1 \in I} \sum_{i_2 \in A_{i_1}} \cdots \sum_{i_{m+2} \in A_{i_1 \dots i_{m+1}}} \sum_{(\mathbb{E})} \mathbb{E}|X_{i_1} X_{i_2}|(\mathbb{E})|X_{i_3}| \cdots (\mathbb{E})|X_{i_{m+2}}|,$$

and $\sum_{(\mathbb{E})}$ denotes the sum over a possible $\mathbb{E}$ in front of each X_i with the constraint that any pair of $\mathbb{E}'s$ must be separated by at least two $X_i's$.

Remark 2.9. The case $p = 1$ was proved by Barbour, Karoński and Ruciński [2]. For the case $p = 2$, we have $R_2 = \gamma_1 + \gamma_2 + \gamma_3$ where γ_1–γ_3 are defined as in Theorem 2.1. In this case, the bound in (2.3) is clearly an upper bound for the bound in (2.1).

3 Proofs

3.1 Preliminaries

To prepare for the proof of Theorem 2.1, we need the following lemmas. The first lemma relates Wasserstein-p distances to Zolotarev's ideal metrics.

Definition 3.1. *For $p > 1$, let $l = \lceil p \rceil - 1$ be the largest integer that is smaller than p and Λ_p be the class of l-times continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$ such that $|f^{(l)}(x) - f^{(l)}(y)| \le |x - y|^{p-l}$ for any $(x, y) \in \mathbb{R}^2$. The ideal distance Z_p of Zolotarev between two probability distributions μ and ν is defined by*

$$Z_p(\mu, \nu) = \sup_{f \in \Lambda_p} \Big\{ \int_{\mathbb{R}} f d\mu - \int_{\mathbb{R}} f d\nu \Big\}.$$

Lemma 3.2 (Theorem 3.1 of [12]). *For any $p > 1$ there exists a positive constant C_p, such that for any pair (μ, ν) of laws on the real line with finite absolute moments of order p,*

$$\mathcal{W}_p(\mu, \nu) \le C_p \big[Z_p(\mu, \nu) \big]^{\frac{1}{p}}.$$

We use Stein's method to obtain the asymptotic expansion (3.5) in the proof of Theorem 2.1. Stein's method was introduced by Stein [14] to prove central limit theorems. The method has been generalized to other limit theorems and drawn considerable interest recently. We refer to the book by Chen, Goldstein and Shao [5] for an introduction to Stein's method. Barbour [1] used Stein's method to obtain an asymptotic expansion for expectations of smooth functions of sums of independent random variables. Rinott and Rotar [11] considered a related expansion for dependency-neighborhoods chain structures. See Remark 3.6 below for more details.

For a function h, denote $\mathcal{N}h := \mathbb{E}h(Z)$, where $Z \sim N(0, 1)$, provided that the expectation exists. Consider the Stein equation

$$f'(w) - wf(w) = h(w) - \mathcal{N}h. \tag{3.1}$$

Let

$$\begin{aligned} f_h(w) &= \int_{-\infty}^{w} e^{\frac{1}{2}(w^2 - t^2)} \big\{ h(t) - \mathcal{N}h \big\} dt \\ &= - \int_{w}^{\infty} e^{\frac{1}{2}(w^2 - t^2)} \big\{ h(t) - \mathcal{N}h \big\} dt. \end{aligned} \tag{3.2}$$

We will use the following lemma.

Lemma 3.3 (Special case of Lemma 6 of [1])**.** *For any positive integer $p > 1$, let $h \in \Lambda_p$ where Λ_p is defined in Definition 3.1. Then f_h in* (3.2) *is a solution to* (3.1)*. Moreover, f_h is p times differentiable, and satisfies*

$$|f_h^{(p)}(x) - f_h^{(p)}(y)| \le C_p|x-y|, \quad \forall\, x, y \in \mathbb{R},$$

where C_p is a constant only depending on p.

In the final step of the proof of Theorem 2.1, we will invoke the known Wasserstein-2 bounds in the central limit theorem for sums of i.i.d. random variables. The following result was recently proved by Bobkov [3].

Lemma 3.4 (Theorem 1.1 of [3])**.** *Let $V_n = \sum_{i=1}^n \xi_i$ where $\{\xi_1, \dots, \xi_n\}$ are independent, with $\mathbb{E}\xi_i = 0$ and $\mathbb{E}V_n^2 = 1$. Then for any real $p \ge 1$,*

$$\mathcal{W}_p(\mathcal{L}(V_n), N(0,1)) \le C_p \big[\sum_{i=1}^n \mathbb{E}|\xi_i|^{p+2}\big]^{\frac{1}{p}}, \tag{3.3}$$

where C_p continuously depends on p.

The results for $p \in (1, 2]$ and for $p > 1$ but i.i.d. case were first proved by Rio [12], who also showed that the bound in (3.3) is optimal.

3.2 Proof of Theorem 2.1

As noted in the Introduction, the proof consists of three steps. We first obtain an asymptotic expansion for $\mathbb{E}h(W)$ for $h \in \Lambda_2$. We then use the expansion and Lemma 3.2 to control the Wasserstein-2 distance between the distributions of W and a sum of i.i.d. random variables. Finally, we use the triangle inequality and known Wasserstein-2 bounds in Lemma 3.4 for sums of i.i.d. random variables to prove our main result. Without loss of generality, we assume that the right-hand side of (2.1) is finite.

3.2.1 Asymptotic expansion for $\mathbb{E}h(W)$

In this step, we prove the following proposition.

Proposition 3.5. *Let W be as in Theorem 2.1, let $h \in \Lambda_2$ and let f_h be the solution* (3.2) *to the Stein equation*

$$f'(w) - wf(w) = h(w) - \mathcal{N}h. \tag{3.4}$$

We have

$$\begin{aligned} &\Big|\mathbb{E}h(W) - \mathcal{N}h + \frac{\beta}{2}\mathcal{N}f_h''\Big| \\ \le& C\Big[|\beta|\mathcal{W}_2(\mathcal{L}(W), N(0,1)) + \gamma_1 + \gamma_2 + \gamma_3\Big], \end{aligned} \tag{3.5}$$

where β, γ_1–γ_3 are as in Theorem 2.1.

Remark 3.6. Rinott and Rotar [11] obtained an asymptotic expansion for $\mathbb{E}h(W) - \mathcal{N}h$ under a different set of conditions, which allows certain weak global dependence. It may be possible to obtain a Wasserstein-2 bound for their W. We leave it for future research.

Proof of Proposition 3.5. In the proof, we denote $f := f_h$. From $h \in \Lambda_2$ and Lemma 3.3, we have

$$|f''(x) - f''(y)| \le C|x-y| \tag{3.6}$$

for any $x, y \in \mathbb{R}$. From (3.4), we have

$$\mathbb{E}h(W) - \mathcal{N}h = \mathbb{E}f'(W) - \mathbb{E}Wf(W). \tag{3.7}$$

For each index $i \in I$, let
$$W^{(i)} = W - \sum_{j \in A_i} X_j.$$

By (LD1), X_i is independent of $W^{(i)}$. From $\mathbb{E}X_i = 0$, Taylor's expansion and (3.6), we have

$$\begin{aligned}\mathbb{E}Wf(W) &= \sum_{i \in I} \mathbb{E}X_i f(W) = \sum_{i \in I} \mathbb{E}X_i [f(W) - f(W^{(i)})] \\ =& \sum_{i \in I} \sum_{j \in A_i} \mathbb{E}X_i X_j f'(W^{(i)}) + \frac{1}{2} \sum_{i \in I} \sum_{j,k \in A_i} \mathbb{E}X_i X_j X_k f''(W^{(i)}) + O(\gamma_1),\end{aligned} \tag{3.8}$$

We begin by dealing with the first term on the right-hand side of (3.8). The second term will be dealt with similarly. In (LD2), let

$$W^{(ij)} = W - \sum_{k \in A_{ij}} X_k.$$

By the independence of $\{X_i, X_j\}$ and $W^{(ij)}$ and (3.6), we have

$$\begin{aligned}&\mathbb{E}X_i X_j f'(W^{(i)}) = \mathbb{E}X_i X_j \mathbb{E}f'(W^{(ij)}) + \mathbb{E}X_i X_j \big[f'(W^{(i)}) - f'(W^{(ij)})\big] \\ =&\mathbb{E}X_i X_j \mathbb{E}f'(W) + \mathbb{E}X_i X_j \Big\{\mathbb{E}\big[f'(W^{(ij)}) - f'(W)\big] + \big[f'(W^{(i)}) - f'(W^{(ij)})\big]\Big\} \\ =&\mathbb{E}X_i X_j \mathbb{E}f'(W) + \mathbb{E}X_i X_j \mathbb{E}\big[-\sum_{k \in A_{ij}} X_k f''(W^{(ij)}) + O\big(\sum_{k \in A_{ij}} |X_k|\big)^2\big] \\ &+ \mathbb{E}X_i X_j \big[\sum_{k \in A_{ij} \setminus A_i} X_k f''(W^{(ij)}) + O\big(\sum_{k \in A_{ij}} |X_k|\big)^2\big].\end{aligned}$$

By the assumption that $\mathbb{E}W^2 = \sum_{i \in I} \sum_{j \in A_i} \mathbb{E}X_i X_j = 1$, we have

$$\sum_{i \in I} \sum_{j \in A_i} \mathbb{E}X_i X_j \mathbb{E}f'(W) = \mathbb{E}f'(W).$$

Therefore,

$$\begin{aligned}&\sum_{i \in I} \sum_{j \in A_i} \mathbb{E}X_i X_j f'(W^{(i)}) \\ =&\mathbb{E}f'(W) - \sum_{i \in I} \sum_{j \in A_i} \sum_{k \in A_{ij}} \mathbb{E}X_i X_j \mathbb{E}X_k f''(W^{ij}) \\ &+ \sum_{i \in I} \sum_{j \in A_i} \sum_{k \in A_{ij} \setminus A_i} \mathbb{E}X_i X_j X_k f''(W^{ij}) + O(\gamma_1 + \gamma_2).\end{aligned} \tag{3.9}$$

In (LD3), let
$$W^{(ijk)} = W - \sum_{l \in A_{ijk}} X_l.$$

By the independence of $\{X_i, X_j, X_k\}$ and $W^{(ijk)}$, $\mathbb{E}X_k = 0$ and (3.6), we have

$$\begin{aligned}&\sum_{i \in I} \sum_{j \in A_i} \sum_{k \in A_{ij}} \mathbb{E}X_i X_j \mathbb{E}X_k f''(W^{ij}) \\ =&\sum_{i \in I} \sum_{j \in A_i} \sum_{k \in A_{ij}} \mathbb{E}X_i X_j \mathbb{E}X_k \big[f''(W^{ij}) - f''(W^{(ijk)})\big] \\ =&O(\gamma_2).\end{aligned} \tag{3.10}$$

Similarly,

$$
\begin{aligned}
&\sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}\setminus A_i}\mathbb{E}X_iX_jX_kf''(W^{ij})\\
=&\sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}\setminus A_i}\mathbb{E}X_iX_jX_k\mathbb{E}f''(W^{(ijk)})\\
&+\sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}\setminus A_i}\mathbb{E}X_iX_jX_k\big[f''(W^{ij})-f''(W^{(ijk)})\big]\\
=&\sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}\setminus A_i}\mathbb{E}X_iX_jX_k\mathbb{E}f''(W)+O(\gamma_1+\gamma_3)
\end{aligned}
\tag{3.11}
$$

Combining (3.9), (3.10) and (3.11), we have

$$
\begin{aligned}
&\sum_{i\in I}\sum_{j\in A_i}\mathbb{E}X_iX_jf'(W^{(i)})\\
=&\mathbb{E}f'(W)+\sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}\setminus A_i}\mathbb{E}X_iX_jX_k\mathbb{E}f''(W)+O(\gamma_1+\gamma_2+\gamma_3).
\end{aligned}
\tag{3.12}
$$

Similar arguments applied to the second term on the right-hand side of (3.8) yield

$$
\begin{aligned}
&\frac{1}{2}\sum_{i\in I}\sum_{j,k\in A_i}\mathbb{E}X_iX_jX_kf''(W^{(i)})\\
=&\frac{1}{2}\sum_{i\in I}\sum_{j,k\in A_i}\mathbb{E}X_iX_jX_k\mathbb{E}f''(W)\\
&+\frac{1}{2}\sum_{i\in I}\sum_{j,k\in A_i}\mathbb{E}X_iX_jX_k\Big\{\mathbb{E}\big[f''(W^{ijk})-f''(W)\big]+\big[f''(W^{(i)})-f''(W^{(ijk)})\big]\Big\}\\
=&\frac{1}{2}\sum_{i\in I}\sum_{j,k\in A_i}\mathbb{E}X_iX_jX_k\mathbb{E}f''(W)+O(\gamma_1+\gamma_3).
\end{aligned}
\tag{3.13}
$$

From (3.7), (3.8), (3.12) and (3.13), we have

$$
\begin{aligned}
&\mathbb{E}h(W)-\mathcal{N}h=\mathbb{E}f'(W)-\mathbb{E}Wf(W)\\
=&-\sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}\setminus A_i}\mathbb{E}X_iX_jX_k\mathbb{E}f''(W)-\frac{1}{2}\sum_{i\in I}\sum_{j,k\in A_i}\mathbb{E}X_iX_jX_k\mathbb{E}f''(W)\\
&+O(\gamma_1+\gamma_2+\gamma_3)\\
=&-\frac{\beta}{2}\mathbb{E}f''(W)+O(\gamma_1+\gamma_2+\gamma_3).
\end{aligned}
\tag{3.14}
$$

From (3.6) and the equivalent definition of the Wasserstein-1 distance

$$
\mathcal{W}_1(\mu,\nu)=\sup_{g\in\mathrm{Lip}_1(\mathbb{R})}\Big|\int gd\mu-\int gd\nu\Big|,
$$

we have

$$
\big|\mathbb{E}f''(W)-\mathcal{N}f''\big|\le C\mathcal{W}_1(\mathcal{L}(W),N(0,1))\le C\mathcal{W}_2(\mathcal{L}(W),N(0,1)).
$$

This proves (3.5). □

3.2.2 $\mathcal{W}_2$ bound for approximating $\mathcal{L}(W)$ by the distribution of a sum of i.i.d. random variables

Note that in proving Theorem 2.1, we can assume that $|\beta|$ is smaller than an arbitrarily chosen constant $c_1 > 0$. If $\beta \neq 0$, let $n = \lfloor c_2\beta^{-2} \rfloor$ for a constant $c_2 > 0$ to be chosen. Let $\{\xi_i : i = 1, \dots, n\}$ be i.i.d. such that

$$\begin{aligned}
\mathbb{P}(\xi_1 = -\frac{3}{2}) &= \frac{3}{16} - \frac{\sqrt{n}\beta}{6},\\
\mathbb{P}(\xi_1 = -\frac{1}{2}) &= \frac{5}{16} + \frac{\sqrt{n}\beta}{2},\\
\mathbb{P}(\xi_1 = \frac{1}{2}) &= \frac{5}{16} - \frac{\sqrt{n}\beta}{2},\\
\mathbb{P}(\xi_1 = \frac{3}{2}) &= \frac{3}{16} + \frac{\sqrt{n}\beta}{6},
\end{aligned}$$

where we choose c_2 to be small enough so that the above is indeed a probability distribution, and then choose c_1 to be small enough so that $n \geq 1$. By straightforward computation, we have

$$\mathbb{E}\xi_i = 0,\ \mathbb{E}\xi_i^2 = 1,\ \mathbb{E}\xi_i^3 = \sqrt{n}\beta,\ \mathbb{E}\xi_i^4 \leq C.$$

Let $V_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_i$. Note that $\kappa_3(V_n) = \beta$, where κ_r denotes the rth cumulant, and $\sum_{i=1}^n \frac{\mathbb{E}\xi_i^4}{n^2} \leq \frac{C}{n} \leq C\beta^2$. The expansion in Theorem 1 of [1] implies

$$\Big|\mathbb{E}h(V_n) - \mathcal{N}h + \frac{\beta}{2}\mathcal{N}f_h''\Big| \leq C\beta^2. \tag{3.15}$$

If $\beta = 0$, let $V_n \sim N(0,1)$ and (3.15) automatically holds. From Lemma 3.2 and the expansions (3.5) and (3.15), we have

$$\begin{aligned}
&\mathcal{W}_2(\mathcal{L}(W), \mathcal{L}(V_n))\\
\leq& C\Big\{\sup_{h\in\Lambda_2}\big[\mathbb{E}h(W) - \mathbb{E}h(V_n)\big]\Big\}^{\frac{1}{2}}\\
\leq& C\Big\{|\beta| + \big[|\beta|\mathcal{W}_2(\mathcal{L}(W), N(0,1))\big]^{\frac{1}{2}} + (\gamma_1+\gamma_2+\gamma_3)^{\frac{1}{2}}\Big\}.
\end{aligned} \tag{3.16}$$

We remark that Rio [12] used a Poisson-like approximation for $\mathcal{L}(W)$. Approximating by sums of i.i.d. random variables enables us to potentially bound the Wasserstein-p distance for any positive integer p.

3.2.3 Triangle inequality and the final bound

By Lemma 3.4,

$$\mathcal{W}_2(\mathcal{L}(V_n), N(0,1)) \leq C\Big\{\sum_{i=1}^n \frac{\mathbb{E}\xi_i^4}{n^2}\Big\}^{\frac{1}{2}} \leq C|\beta|. \tag{3.17}$$

Using the triangle inequality, (3.16) and (3.17), we obtain

$$\begin{aligned}
&\mathcal{W}_2(\mathcal{L}(W), N(0,1))\\
\leq& \mathcal{W}_2(\mathcal{L}(W), \mathcal{L}(V_n)) + \mathcal{W}_2(\mathcal{L}(V_n), N(0,1))\\
\leq& C\Big\{|\beta| + \big[|\beta|\mathcal{W}_2(\mathcal{L}(W), N(0,1))\big]^{\frac{1}{2}} + (\gamma_1+\gamma_2+\gamma_3)^{\frac{1}{2}}\Big\}.
\end{aligned}$$

Finally, we use the inequality $\sqrt{ab} \leq \frac{1}{2\epsilon}a + \frac{\epsilon}{2}b$ with $a = |\beta|$ and $b = \mathcal{W}_2(\mathcal{L}(W), N(0,1))$, choose a sufficiently small ϵ and solve the recursive inequality for $\mathcal{W}_2(\mathcal{L}(W), N(0,1))$ to obtain the bound (2.1).

3.3 Proof of Corollary 2.3

For each $i = 1, \dots, n$, let $A_i = \{j : |j - i| \le m\}$. For each $i = 1, \dots, n$ and $j \in A_i$, let $A_{ij} = \{k : \min\{|k - j|, |k - i|\} \le m\}$. For each $i = 1, \dots, n$, $j \in A_i$ and $k \in A_{ij}$, let $A_{ijk} = \{l : \min\{|l - i|, |l - j|, |l - k|\} \le m\}$. By the m-dependence assumption, they satisfy the assumptions (LD1)–(LD3) for Theorem 2.1. For the first term in the definition of β of Theorem 2.1, we have

$$\begin{aligned}&|\sum_{i=1}^n \sum_{j,k\in A_i} \mathbb{E} X_i X_j X_k| \\ \le& C \sum_{i=1}^n \sum_{j,k\in A_i} (\mathbb{E}|X_i|^3 + \mathbb{E}|X_j|^3 + \mathbb{E}|X_k|^3) \\ \le& Cm^2 \sum_{i=1}^n \mathbb{E}|X_i|^3,\end{aligned}$$

where the last inequality is from the fact that each i is counted at most Cm^2 times in the previous expression. The second term of β has the same upper bound. Similarly, for γ_1, we have

$$\begin{aligned}&\sum_{i\in I} \sum_{j\in A_i} \sum_{k\in A_{ij}} \sum_{l\in A_{ijk}} \mathbb{E}|X_i X_j X_k X_l| \\ \le& C \sum_{i\in I} \sum_{j\in A_i} \sum_{k\in A_{ij}} \sum_{l\in A_{ijk}} (\mathbb{E}|X_i|^4 + \mathbb{E}|X_j|^4 + \mathbb{E}|X_k|^4 + \mathbb{E}|X_l|^4) \\ \le& Cm^3 \sum_{i=1}^n \mathbb{E}|X_i|^4,\end{aligned}$$

and γ_2 and γ_3 have the same upper bound. This proves the corollary.

3.4 Proof of Theorem 2.4

Consider the index set

$$I = \{i = (i_1, \dots, i_m) : 1 \le i_1 < \dots < i_m \le n\}.$$

For each $i \in I$, let $\xi_i = \sigma_n^{-1} h(X_{i_1}, \dots, X_{i_m})$. Then $W_n = \sum_{i\in I} \xi_i$. For each $i \in I$, let

$$A_i = \{j \in I : i \cap j \neq \emptyset\}.$$

For each $i \in I$ and $j \in A_i$, let

$$A_{ij} = \{k \in I : k \cap (i \cup j) \neq \emptyset\}.$$

For each $i \in I$, $j \in A_i$ and $k \in A_{ij}$, let

$$A_{ijk} = \{l \in I : l \cap (i \cup j \cup k) \neq \emptyset\}.$$

Then they satisfy the conditions (LD1)–(LD3) of Theorem 2.1. Moreover, the sizes of the neighborhoods are all bounded by Cn^{m-1}. Note that by the non-degeneracy condition, $\sigma_n^2 \asymp n^{2m-1}$. By Theorem 2.1, we have

$$\begin{aligned}&\mathcal{W}_2(\mathcal{L}(W_n), N(0,1)) \\ \le& C\Big\{ n^m (n^{m-1})^2 \frac{\mathbb{E}|h(X_1, \dots, X_m)|^3}{\sigma_n^3} + \big[n^m (n^{m-1})^3 \frac{\mathbb{E}(h(X_1, \dots, X_m))^4}{\sigma_n^4} \big]^{1/2} \Big\} \\ \le& C/\sqrt{n}.\end{aligned}$$

3.5 Proof of Theorem 2.6

In this subsection, the constants C are allowed to depend on the given fixed graph G. Let the potential edges of $K(n,p)$ be denoted by $(e_1, \dots, e_{\binom{n}{2}})$. Let $v = v(G), e = e(G)$. In applying Theorem 2.1, let $W = \sum_{i\in I} X_i$, where the index set is

$$I = \Big\{ i = (i_1, \dots, i_e) : 1 \le i_1 < \dots < i_e \le \binom{n}{2}, G_i := (e_{i_1}, \dots, e_{i_e}) \text{ is a copy of } G \Big\},$$

$$X_i = \sigma^{-1}\big(Y_i - p^e\big), \quad \sigma^2 := \mathrm{Var}(S), \quad Y_i = \Pi_{l=1}^e E_{i_l},$$

and E_{i_l} is the indicator of the event that the edge e_{i_l} is connected in $K(n,p)$. It is known that (cf. (3.7) of [2])

$$\sigma^2 \ge C(1-p)n^{2v}p^{2e}\psi^{-1}.$$

For each $i \in I$, let

$$A_i = \{j \in I : e(G_j \cap G_i) \ge 1\}.$$

For each $i \in I$ and $j \in A_i$, let

$$A_{ij} = \{k \in I : e(G_k \cap (G_i \cup G_j)) \ge 1\}.$$

For each $i \in I$, $j \in A_i$ and $k \in A_{ij}$, let

$$A_{ijk} = \{l \in I : e(G_l \cap (G_i \cup G_j \cup G_k)) \ge 1\},$$

Then they satisfy (LD1)–(LD3) of Section 2.1. Note that the Y's are all increasing functions of the E's. By the arguments leading to (3.8) of [2], we have

$$\begin{aligned}
&\gamma := \gamma_1 + \gamma_2 + \gamma_3 \\
\le &\Big\{ \frac{C}{\sigma^4} \sum_{i\in I} \sum_{j\in A_i} \sum_{k\in A_{ij}} \sum_{l\in A_{ijk}} \mathbb{E}(Y_iY_jY_kY_l) \Big\} \wedge \Big\{ \frac{C}{\sigma^4} \sum_{i\in I} \sum_{j\in A_i} \sum_{k\in A_{ij}} \sum_{l\in A_{ijk}} \mathbb{E}(1-Y_i) \Big\}.
\end{aligned}$$

For $\frac{1}{2} < p < 1$, the latter term directly yields the estimate

$$\begin{aligned}
\gamma \le & C\sigma^{-4} n^v n^{3(v-2)}(1-p) \\
\le & Cn^{4v-6}(1-p)[n^{2v-2}(1-p)]^{-2} \\
\le & Cn^{-2}(1-p)^{-1}.
\end{aligned}$$

Let $\cong$ denote graph homomorphism. For $0 < p \le \frac{1}{2}$, the former term gives

$$\begin{aligned}
\gamma \le & C\sigma^{-4} \sum_{\substack{H\subset G \\ e(H)\ge 1}} \sum_{\substack{i,j\in I \\ G_i\cap G_j \cong H}} \sum_{\substack{K\subset (G_i\cup G_j) \\ e(K)\ge 1}} \sum_{\substack{k\in I \\ G_k\cap(G_i\cup G_j)=K}} \\
& \Big\{ \sum_{\substack{L\subset (G_i\cup G_j\cup G_k) \\ e(L)\ge 1}} \sum_{\substack{l\in I \\ G_l\cap(G_i\cup G_j\cup G_k)=L}} p^{4e-e(H)-e(K)-e(L)} \Big\} \\
\le & C\sigma^{-4} \sum_{\substack{H\subset G \\ e(H)\ge 1}} \sum_{\substack{i,j\in I \\ G_i\cap G_j \cong H}} \sum_{\substack{K\subset (G_i\cup G_j) \\ e(K)\ge 1}} \sum_{\substack{k\in I \\ G_k\cap(G_i\cup G_j)=K}} \\
& \Big\{ \sum_{\substack{L\subset (G_i\cup G_j\cup G_k) \\ L\subset G_m \text{ for some } m, e(L)\ge 1}} n^{v-v(L)} p^{4e-e(H)-e(K)-e(L)} \Big\} \\
\le & C\sigma^{-4}\psi^{-1}n^v p^e \sum_{\substack{H\subset G \\ e(H)\ge 1}} \sum_{\substack{i,j\in I \\ G_i\cap G_j \cong H}} \sum_{\substack{K\subset (G_i\cup G_j) \\ e(K)\ge 1}} \sum_{\substack{k\in I \\ G_k\cap(G_i\cup G_j)=K}} p^{3e-e(H)-e(K)} \\
\le & C\sigma^{-2}(\psi^{-1}n^v p^e)^2,
\end{aligned}$$

where in the last step, we used (3.10) of [2]. This gives

$$\gamma \le C\psi^{-1}.$$

In summary, we have proved that $\gamma^{1/2}$ is bounded by the right-hand side of (2.2). By a similar and simpler argument which is essentially the same as (3.10) of [2], we also have that $|\beta|$ is bounded by the right-hand side of (2.2). Theorem 2.6 is now proved by invoking Theorem 2.1.

3.6 Supporting arguments for Conjecture 2.8

We follow the proof of Theorem 2.1, obtain higher-order expansions and choose appropriate sums of i.i.d. random variables for the intermediate approximation.

We first give a complete proof for the case $p = 3$. Without loss of generality, we assume that the right-hand side of (2.3) is finite. Let $h \in \Lambda_3$. Let $f := f_h$ in (3.2) be the solution to the Stein equation

$$f'(w) - wf(w) = h(w) - \mathcal{N}h.$$

From $h \in \Lambda_3$ and Lemma 3.3,

$$|f^{(3)}(x) - f^{(3)}(x)| \le C|x-y|. \tag{3.18}$$

We further let $g := g_{f''}$, defined by replacing h by f'' on the right-hand side of (3.2), be the solution to

$$g'(w) - wg(w) = f''(w) - \mathcal{N}f''.$$

From $\frac{1}{C}f'' \in \Lambda_2$ and Lemma 3.3, we have

$$|g''(x) - g''(y)| \le C|x-y|.$$

Denote the third cumulant of W by

$$\kappa_3 := \kappa_3(W) = \sum_{i\in I}\sum_{j,k\in A_i} \mathbb{E}X_iX_jX_k + 2\sum_{i\in I}\sum_{j\in A_i}\sum_{k\in A_{ij}\setminus A_i} \mathbb{E}X_iX_jX_k,$$

which we denoted by β before. Denote the fourth cumulant of W by $\kappa_4 := \kappa_4(W)$. A tedious but similar expansion as for (3.14) yields

$$\begin{aligned} &\mathbb{E}h(W) - \mathcal{N}h = \mathbb{E}f'(W) - \mathbb{E}Wf(W) \\ =& -\frac{\kappa_3}{2}\mathbb{E}f''(W) - \frac{\kappa_4}{6}\mathbb{E}f^{(3)}(W) + O(R_3). \end{aligned} \tag{3.19}$$

Since $\frac{1}{C}f'' \in \Lambda_2$, from (3.5), we have

$$|\mathbb{E}f''(W) - \mathcal{N}f'' + \frac{\kappa_3}{2}\mathcal{N}g''| \le C\big[|\kappa_3|\mathcal{W}_3(\mathcal{L}(W), N(0,1)) + R_2\big]. \tag{3.20}$$

From (3.18), we have

$$\mathbb{E}f^{(3)}(W) - \mathcal{N}f^{(3)} = O(\mathcal{W}_3(\mathcal{L}(W), N(0,1))). \tag{3.21}$$

From (3.19)–(3.21) and $|\kappa_3| \le CR_1$, $|\kappa_4| \le CR_2$, we have

$$\begin{aligned} &\Big|\mathbb{E}h(W) - \mathcal{N}h + \frac{\kappa_3}{2}\mathcal{N}f'' + \frac{\kappa_4}{6}\mathcal{N}f^{(3)} - \frac{\kappa_3^2}{4}\mathcal{N}g''\Big| \\ \le& C\Big[(R_1^2 + R_2)\mathcal{W}_3(\mathcal{L}(W), N(0,1)) + R_1R_2 + R_3\Big]. \end{aligned} \tag{3.22}$$

Without loss of generality, assume that R_1 and R_2, hence $|\kappa_3|$ and $|\kappa_4|$ are smaller than an arbitrarily chosen constant $c_1 > 0$. Otherwise, the bound (2.3) is trivial for $p = 3$ by choosing a large enough C_3. If $\kappa_3 \neq 0$ or $\kappa_4 \neq 0$, let

$$n = \lfloor c_2 \kappa_3^{-2} \rfloor \wedge \lfloor c_2 |\kappa_4|^{-1} \rfloor$$

for a constant $c_2 > 0$ to be chosen. Let $\{\xi_i : i = 1, \dots, n\}$ be i.i.d. such that

$$\begin{aligned}
\mathbb{P}(\xi_1 = -2) &= \frac{1}{12} + \frac{-2\sqrt{n}\kappa_3 + n\kappa_4}{24}, \\
\mathbb{P}(\xi_1 = -1) &= \frac{1}{6} + \frac{\sqrt{n}\kappa_3 - n\kappa_4}{6}, \\
\mathbb{P}(\xi_1 = 0) &= \frac{1}{2} + \frac{n\kappa_4}{4}, \\
\mathbb{P}(\xi_1 = 1) &= \frac{1}{6} - \frac{\sqrt{n}\kappa_3 + n\kappa_4}{6}, \\
\mathbb{P}(\xi_1 = 2) &= \frac{1}{12} + \frac{2\sqrt{n}\kappa_3 + n\kappa_4}{24},
\end{aligned}$$

where we choose c_2 to be small enough so that the above is indeed a probability distribution, and then choose c_1 to be small enough so that $n \geq 1$. By straightforward computation, we have

$$\mathbb{E}\xi_1 = 0,\ \mathbb{E}\xi_2^2 = 1,\ \kappa_3(\xi_1) = \sqrt{n}\kappa_3,\ \kappa_4(\xi_1) = n\kappa_4,\ \mathbb{E}|\xi_1|^5 \leq C.$$

Let $V_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \xi_i$. The expansion in Theorem 1 of [1] implies

$$\left| \mathbb{E}h(V_n) - \mathcal{N}h + \frac{\kappa_3}{2}\mathcal{N}f'' + \frac{\kappa_4}{6}\mathcal{N}f^{(3)} - \frac{\kappa_3^2}{4}\mathcal{N}g'' \right| \leq \frac{C}{n^{3/2}} \leq C(R_1^3 + R_2^{3/2}). \tag{3.23}$$

If $\kappa_3 = \kappa_4 = 0$, let $V_n \sim N(0,1)$ and (3.23) automatically holds. The expansions (3.22) and (3.23) imply

$$|\mathbb{E}h(W) - \mathbb{E}h(V_n)| \leq C\big[(R_1^2 + R_2)\mathcal{W}_3(\mathcal{L}(W), N(0,1)) + R_1^3 + R_2^{3/2} + R_3\big],$$

where we used Young's inequality $|ab| \leq C(|a|^3 + |b|^{3/2})$. As in the proof of Theorem 2.1, we have

$$\begin{aligned}
&\mathcal{W}_3(\mathcal{L}(W), N(0,1)) \\
\leq &\mathcal{W}_3(\mathcal{L}(W), \mathcal{L}(V_n)) + C(R_1 + R_2^{1/2}) \\
\leq &C(R_1 + R_2^{1/2} + R_3^{1/3}) + C(R_1 + R_2^{1/2})^{2/3}(\mathcal{W}_3(\mathcal{L}(W), N(0,1)))^{1/3} \\
\leq &\frac{1}{2}\mathcal{W}_3(\mathcal{L}(W), N(0,1)) + C(R_1 + R_2^{1/2} + R_3^{1/3}).
\end{aligned}$$

This implies the conjectured result for $p = 3$.

For the case $p \geq 4$ and $h \in \Lambda_p$, we start with the expansion

$$\begin{aligned}
\mathbb{E}h(W) - \mathcal{N}h &= \mathbb{E}f'(W) - \mathbb{E}Wf(W) \\
&= -\sum_{m=1}^{p-1} \frac{\kappa_{m+2}}{(m+1)!}\mathbb{E}f^{(m)}(W) + O(R_p),
\end{aligned}$$

where $f = f_h$ in (3.2) is the solution to (3.1) and $\kappa_{m+2} := \kappa_{m+2}(W)$ is the $(m+2)$th cumulant of W. To see that the coefficients must be of the given form of the cumulants, take $f(w) = w^2, w^3, \dots$ in the expansion. The constraint that any pair of $\mathbb{E}$'s must be separated by at least two X_i's is from the assumption that $\mathbb{E}X_i = 0$ for any $i \in I$. The conjectured result should then follow by similar arguments as for the case $p = 3$.

References

[1] Barbour, A. D. (1986). Asymptotic expansions based on smooth functions in the central limit theorem. *Probab. Theory Relat. Fields* **72**, no. 2, 289–303. MR-0836279

[2] Barbour, A. D., Karoński, M. and Ruciński, A. (1989). A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B* **47**, no. 2, 125–145. MR-1047781

[3] Bobkov, S. G. (2018). Berry-Esseen bounds and Edgeworth expansions in the central limit theorem for transport distances. *Probab. Theory Related Fields* **170**, no. 1-2, 229–262. MR-3748324

[4] Bonis, T. (2018). Rate in the central limit theorem and diffusion approximation via Stein's method. *Preprint.* Available at https://arxiv.org/abs/1506.06966

[5] Chen, L.H.Y., Goldstein, L. and Shao, Q.M. (2011). *Normal approximation by Stein's method.* Probability and its Applications (New York). Springer, Heidelberg, 2011. xii+405 pp. MR-2732624

[6] Chen, L.H.Y. and Shao, Q.M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32**, no. 3A, 1985–2028. MR-2073183

[7] Chen, L.H.Y. and Shao, Q.M. (2007). Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli* **13**, 581–599. MR-2331265

[8] Courtade, T.A., Fathi, M. and Pananjady, A. (2018). Existence of Stein kernels under a spectral gap, and discrepancy bound. *Preprint.* Available at https://arxiv.org/abs/1703.07707

[9] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics* **19**, 293–325. MR-0026294

[10] Ledoux, M., Nourdin, I. and Peccati, G. (2015). Stein's method, logarithmic Sobolev and transport inequalities. *Geom. Funct. Anal.* **25**, no. 1, 256–306. MR-3320893

[11] Rinott, Y. and Rotar, V. (2003). On Edgeworth expansions for dependency-neighborhoods chain structures and Stein's method. *Probab. Theory Related Fields* **126**, no. 4, 528–570. MR-2001197

[12] Rio, E. (2009). Upper bounds for minimal distances in the central limit theorem. *Ann. Inst. Henri Poincaré Probab. Stat.* **45**, no. 3, 802–817. MR-2548505

[13] Röllin, A. (2017). Kolmogorov bounds for the normal approximation of the number of triangles in the Erdös-Rényi random graph. *Preprint.* Available at https://arxiv.org/abs/1704.00410

[14] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Stat. Prob.* **2**, Univ. California Press. Berkeley, Calif., 583–602. MR-0402873

[15] Zhai, A. (2018). A high-dimensional CLT in $\mathcal{W}_2$ distance with near optimal convergence rate. *Probab. Theory Related Fields* **170**, no. 3-4, 821–845. MR-3773801

Acknowledgments. The author would like to thank Michel Ledoux for introducing the problem, Jia-An Yan for comments on an earlier version of this paper, and an anonymous referee for insightful suggestions.