## OPTIMAL BOUNDS IN NORMAL APPROXIMATION FOR MANY INTERACTING WORLDS

## By Louis H. Y. Chen<sup>1,a</sup> and Lê Văn Thành<sup>2,b</sup>

<sup>1</sup>Department of Mathematics, National University of Singapore, <sup>a</sup>matchyl@nus.edu.sg <sup>2</sup>Department of Mathematics, Vinh University, <sup>b</sup>levt@vinhuni.edu.vn

In this paper, we use Stein's method to obtain optimal bounds, both in Kolmogorov and in Wasserstein distance, in the normal approximation for the empirical distribution of the ground state of a many-interacting-worlds harmonic oscillator proposed by Hall, Deckert and Wiseman (*Phys. Rev. X* **4** (2014) 041013). Our bounds on the Wasserstein distance solve a conjecture of McKeague and Levin (*Ann. Appl. Probab.* **26** (2016) 2540–2555).

**1. Introduction.** In [10], Hall, Deckert and Wiseman proposed a many-interactingworlds (MIW) theory for interpreting quantum mechanics. In this theory, quantum theory can be understood as the continuum limit of a deterministic theory in which there is a large, but finite, number of interacting classical "worlds". Here, a world means an entire universe with well-defined properties, determined by the classical configuration of its particles and fields.

Hall, Deckert and Wiseman [10] proposed a MIW harmonic oscillator model for N onedimensional worlds, where the Hamiltonian for the MIW harmonic oscillator is:

$$H(\mathbf{x}, \mathbf{p}) = E(\mathbf{p}) + V(\mathbf{x}) + U(\mathbf{x}).$$

Here  $\mathbf{p} = (p_1, \dots, p_N)$  are the momenta and  $\mathbf{x} = (x_1, \dots, x_N)$ ,  $x_1 > x_2 > \dots > x_N$ , the locations of the *N* worlds, each regarded as a particle with unit mass,

$$E(\mathbf{p}) = \sum_{n=1}^{N} p_n^2 / 2$$

is the kinetic energy,

$$V(\mathbf{x}) = \sum_{n=1}^{N} x_n^2$$

is the potential energy, and

$$U(\mathbf{x}) = \sum_{n=1}^{N} \left( \frac{1}{x_{n-1} - x_n} - \frac{1}{x_n - x_{n+1}} \right)^2$$

is the "interworld" potential, where  $x_0 = \infty$  and  $x_{N+1} = -\infty$ , which is a discretization of Bohm's quantum potential (see Bohm [1, 2]).

Hall, Deckert and Wiseman [10] showed that in the ground state, where the Hamiltonian is minimized, all the momenta  $p_n$  vanish and the locations of the N particles satisfy this recursion equation:

(1.1) 
$$x_{n+1} = x_n - \frac{1}{x_1 + \dots + x_n}, \quad 1 \le n \le N - 1,$$

subject to the constraints  $x_1 + \cdots + x_N = 0$  and  $x_1^2 + \cdots + x_N^2 = N - 1$ .

Received June 2020; revised April 2021.

MSC2020 subject classifications. 60F05, 81Q65.

Key words and phrases. Stein's method, normal approximation, many interacting worlds, harmonic oscillator.

Their numerical calculations suggest that the empirical distribution  $\mathbb{P}_N$  of the locations converges to the standard Gaussian distribution  $\gamma$  as  $N \to \infty$ , which agrees with the ground state probability distribution of a quantum harmonic oscillator. Here the empirical distribution  $\mathbb{P}_N$  is defined by

(1.2) 
$$\mathbb{P}_N(A) := \frac{\#\{n : x_n \in A\}}{N} \quad \text{for } A \in \mathcal{B}(\mathbb{R}).$$

In [12], McKeague and Levin proved that the recursion equation (1.1) has a unique solution if the solution is monotonic, zero-median, namely  $x_m = 0$  for N odd and  $x_m = -x_{m+1}$  for N even, where m = (N + 1)/2 if N is odd, and m = N/2 if N is even. They proved that  $\mathbb{P}_N$ converges to  $\gamma$  as  $N \to \infty$ .

The Wasserstein distance between  $\mathbb{P}_N$  and  $\gamma$ , denoted by  $d_W(\mathbb{P}_N, \gamma)$ , is defined by

$$d_{\mathbf{W}}(\mathbb{P}_N,\gamma) := \sup_{|h(x) - h(y)| \le |x - y|} \left| \int_{\mathbb{R}} h \, d\mathbb{P}_N - \int_{\mathbb{R}} h \, d\gamma \right|$$

Using Stein's method and zero-bias coupling, McKeague and Levin [12] further proved that

$$d_{\mathrm{W}}(\mathbb{P}_N,\gamma) \leq \frac{4}{\sqrt{\log N}},$$

and conjectured that the correct order of the bound on  $d_W(\mathbb{P}_N, \gamma)$  should be  $\sqrt{\log N}/N$ .

The Kolmogorov distance between  $\mathbb{P}_N$  and  $\gamma$ , denoted by  $d_{\mathrm{K}}(\mathbb{P}_N, \gamma)$ , is defined by

$$d_{\mathrm{K}}(\mathbb{P}_{N},\gamma) := \sup_{z \in \mathbb{R}} |\mathbb{P}_{N}((-\infty,z]) - \gamma((-\infty,z])|.$$

In this paper, we prove that the upper and lower bounds on the Wasserstein distance between  $\mathbb{P}_N$  and  $\gamma$  are both of the order  $\sqrt{\log N}/N$ , thereby proving the conjecture of McKeague and Levin [12] while at the same time showing that  $\sqrt{\log N}/N$  is optimal. We also prove that the upper and lower bounds on the Kolmogorov distance between  $\mathbb{P}_N$  and  $\gamma$  are both of the order 1/N, showing that 1/N is also optimal. This is a surprising outcome as optimal bounds on the Kolmogorov distance are usually of no smaller order than those on the Wasserstein distance for a particular problem.

Our proof of the upper bound on the Wasserstein distance turned out to be an easy consequence of Theorem 1.1 of Goldstein [8] using the zero-bias coupling of McKeague and Levin [12] and an upper bound on  $x_1$ . Our approach is also the same as that of the proof of the upper bound on the Wasserstein distance for two-sided Maxwell approximation in McKeague, Peköz and Swan [13], page 109.

REMARK 1.1. Stein's method is applicable to Wasserstein distance of order p > 1 (see, e.g., [3, 6, 7, 11] for recent results in this direction). We have not succeeded in obtaining optimal bounds on the Wasserstein distance of order p > 1 between  $\mathbb{P}_N$  and  $\gamma$ . It remains an open problem for future research.

Throughout this paper,  $\gamma$  denotes the standard normal distribution. Let  $N \ge 2$  be an integer number, we will denote m = (N + 1)/2 if N is odd, and = N/2 if N is even. For a positive number x,  $\lfloor x \rfloor$  denotes the greatest integer number which is less than or equal to x, and  $\log x$ denotes the natural (base e = 2.7182...) logarithm of x. The symbol C denotes a positive constant which does not depend on N, and its value may be different at each appearance. We also denote by  $\{x_1, \ldots, x_N\}$  the unique zero-median and strictly decreasing solution to the recursion equation (1.1) and by  $\mathbb{P}_N$  the empirical distribution as in (1.2), and let

$$S_n = \sum_{i=1}^n x_i, \quad 1 \le n \le N.$$

The rest of the paper is organized as follows. In Section 2, we reproduce the construction of the zero-bias coupling of McKeague and Levin [12] and state the main result. Section 3 focuses on the proof of the bounds on the Kolmogorov distance. The bounds on the Wasserstein distance are presented in Section 4. Finally, some technical results are proved in the Appendix.

**2. Main results.** Before proving the main result, we will need some preliminary lemmas. The first lemma establishes the existence of a solution to the recursion relation. This lemma is proved in McKeague and Levin [12].

LEMMA 2.1. Every zero-median solution  $\{x_1, \ldots, x_N\}$  of (1.1) satisfies the following properties:

- (P1) *Zero-mean*:  $x_1 + \cdots + x_N = 0$ .
- (P2) *Variance-bound*:  $x_1^2 + \dots + x_N^2 = N 1$ .

(P3) Symmetry:  $x_n = -x_{N+1-n}$  for n = 1, ..., N.

Further, there is a unique solution  $\{x_1, \ldots, x_N\}$  of (1.1) such that (P1) holds and

(P4) *Strictly decreasing*:  $x_1 > \cdots > x_N$ .

This solution has the zero-median property, and thus also satisfies (P2) and (P3).

It was shown by Goldstein and Reinert [9] that for any mean zero random variable W with finite variance  $\sigma^2$ , there exists a random variable  $W^*$  which satisfies

(2.1) 
$$\mathbb{E}Wf(W) = \sigma^2 \mathbb{E}f'(W^*)$$

for all absolutely continuous f with  $\mathbb{E}|Wf(W)| < \infty$ . We say such a  $W^*$  has the W-zerobiased distribution. The following result is due to Goldstein and Reinert [9] (see also in Chen, Goldstein and Shao [4], Proposition 2.1).

LEMMA 2.2. Let W be a random variable with mean zero and finite positive variance  $\sigma^2$ , and let W\* have the W-zero-biased distribution. Then the distribution of W\* is absolutely continuous with density given by

$$p^*(x) = \mathbb{E}(W\mathbf{1}(W > x))/\sigma^2 = -\mathbb{E}(W\mathbf{1}(W \le x))/\sigma^2.$$

We recall that  $\mathbb{P}_N$  denotes the empirical probability measure as in (1.2). McKeague and Levin ([12], page 9), constructed a zero bias coupling  $(W, W^*)$ , where W has the probability distribution  $\mathbb{P}_N$ . For completeness, we describe the construction in this paper. From Lemma 2.1, we have Var(W) = (N - 1)/N. By Lemma 2.2, the density of  $W^*$  is given by

$$p^*(x) = \frac{S_n}{N-1} = \frac{1}{(N-1)(x_n - x_{n+1})}$$
 if  $x_{n+1} \le x < x_n, 1 \le n \le N-1$ .

This implies that  $W^*$  is uniformly distributed on each interval  $[x_{n+1}, x_n)$  with mass 1/(N - 1),  $1 \le n \le N - 1$ . Let  $x_{n+1} < y_n < x_n$  such that for  $1 \le n \le N - 1$ , the area under  $p^*$  on the interval  $[y_n, x_n)$  is  $L_n = (N - n)/(N(N - 1))$ , and for  $2 \le n \le N$ , the area under  $p^*$  on the interval  $[x_n, y_{n-1})$  is  $R_n = (n - 1)/(N(N - 1))$ . Then the area under  $p^*$  on  $[y_1, x_1)$  and on  $[x_N, y_{N-1})$  is 1/N, and on  $[y_n, y_{n-1})$  is  $L_n + R_n = 1/N$  for  $2 \le n \le N - 1$ . See Figure 1 below.



FIG. 1. Density  $p^*$  of  $W^*$ .

Now we couple W to W<sup>\*</sup> by defining W and W<sup>\*</sup> on the same probability space with  $\Omega = [x_N, x_1)$  as the sample space and  $p^*$  as the probability measure as follows:  $W^*(w) = \omega$  for all  $\omega \in \Omega$ , and

$$W(w) = \begin{cases} x_1 & \text{if } \omega \in [y_1, x_1), \\ x_n & \text{if } \omega \in [y_n, y_{n-1}), 2 \le n \le N-1, \\ x_N & \text{if } \omega \in [x_N, y_{N-1}). \end{cases}$$

The following theorem is the main result of the paper.

THEOREM 2.3. Let  $\{x_n, 1 \le n \le N\}$  be the unique monotonic zero-mean solution of the recursion relation (1.1) and  $\mathbb{P}_N$  the empirical distribution. Then

(2.2) 
$$\frac{1}{2N} \le d_K(\mathbb{P}_N, \gamma) \le \frac{58}{N},$$

and

(2.3) 
$$\frac{\sqrt{\log(N/2)}}{2N} - \frac{C}{N} \le d_W(\mathbb{P}_N, \gamma) \le \frac{16\sqrt{\log N}}{N}.$$

**3.** Bounds on the Kolmogorov distance. In this section, we will prove the Kolmogorov bounds in (2.2). Here and thereafter, we denote  $\Phi(z) = \gamma((-\infty, z])$  for  $z \in \mathbb{R}$ .

PROOF OF THE UPPER BOUND IN (2.2). To prove the upper bound in (2.2), it suffices to consider z > 0 since W is symmetric. For z > 0, we have

$$(3.1) \qquad \qquad \left|\mathbb{P}(W \le z) - \Phi(z)\right| \le \max\left\{\mathbb{P}(W > z), 1 - \Phi(z)\right\} \le 0.5.$$

If  $N \le 100$ , then the upper bound in (2.2) holds by (3.1). Therefore we only need to consider N > 100. For the case where  $0 < z < x_1$ , there exists  $1 \le n \le m$  such that  $z \in [x_{n+1}, x_n)$ . In this case, we have

$$\mathbb{P}(W > z) = \frac{n}{N}$$
 and  $\mathbb{P}(W^* > z) = \frac{n-1}{N-1} + \varepsilon$ ,

where  $\varepsilon \leq 1/(N-1)$ , and therefore

$$|\mathbb{P}(W \le z) - \mathbb{P}(W^* \le z)| = \left|\frac{n}{N} - \frac{n-1}{N-1} - \varepsilon\right|$$
$$\le \frac{n}{N} - \frac{n-1}{N-1} + \varepsilon$$
$$\le \frac{1}{N} + \frac{1}{N-1} \le \frac{2.02}{N}$$

where we have applied N > 100 in the last inequality. For the case where  $z \ge x_1$ , we have

$$\mathbb{P}(W > z) = \mathbb{P}(W^* > z) = 0.$$

Therefore,

(3.2) 
$$\sup_{z>0} \left| \mathbb{P}(W \le z) - \mathbb{P}(W^* \le z) \right| \le \frac{2.02}{N}$$

Now we bound the Kolmogorov distance between distribution of  $W^*$  and  $\gamma$ . Let z > 0, and let  $f_z$  be the unique bounded solution of the Stein equation

$$f'(w) - wf(w) = 1(w \le z) - \Phi(z),$$

and

 $g_z(w) = \left(w f_z(w)\right)'.$ 

Then (see Chen and Shao [5], page 248)

(3.3) 
$$g_z(w) = \begin{cases} (\sqrt{2\pi}(1+w^2)e^{w^2/2}(1-\Phi(w))-w)\Phi(z) & \text{if } w > z, \\ (\sqrt{2\pi}(1+w^2)e^{w^2/2}\Phi(w)+w)(1-\Phi(z)) & \text{if } w \le z. \end{cases}$$

From Lemma 2.3 in Chen, Goldstein and Shao [4], we have  $0 < f_z(w) \le \sqrt{2\pi}/4$ ,  $|f'_z(w)| \le 1$ , and therefore

(3.4) 
$$|g_z(w)| \le |wf'_z(w)| + |f_z(w)| \le |w| + \sqrt{2\pi}/4.$$

Chen and Shao [5], page 249, proved that  $g_z \ge 0$ ,  $g_z(w) \le 2(1 - \Phi(z))$  for  $w \le 0$ .

We also have the following lemma whose proof is given in the Appendix.

LEMMA 3.1. We have the following properties of  $g_z$ :

(3.5) 
$$g_z(w)$$
 increases when  $w \le z$  and decreases when  $w > z$ ,

(3.6) 
$$g_z(w) \le \frac{3}{2(1-w)^3}$$
 for  $w < 0$ ,

and

(3.7) 
$$g_z(w) \le \frac{3}{(1+w)^3} \quad \text{for } w > z.$$

Chen and Shao [5] proved that

$$g_z(w) \le \frac{2}{1+w^3} \quad \text{for } w > z.$$

For large w, this bound is of the same order as (3.7) but with a better constant. We use (3.7) because we need some technical estimates such as  $(1 + w)^3 \ge x_{j-1}^3$  for  $w \ge x_{j+1}$ ,  $2 \le j \le m-1$  (see, e.g., (3.20) and (3.21) below).

The Kolmogorov distance between the distribution of  $W^*$  and  $\gamma$  can be bounded as follows:

$$|\mathbb{P}(W^* \le z) - \Phi(z)| = |\mathbb{E}f_z'(W^*) - \mathbb{E}W^* f_z(W^*)|$$
  
(3.8)
$$= \left|\frac{N}{N-1}\mathbb{E}Wf_z(W) - \mathbb{E}W^* f_z(W^*)\right|$$
$$\le |\mathbb{E}Wf_z(W) - \mathbb{E}W^* f_z(W^*)| + \frac{1}{N-1}\mathbb{E}|Wf_z(W)|.$$

Since  $0 < f_z(w) \le \sqrt{2\pi}/4$ , the last term in (3.8) is bounded by

(3.9) 
$$\frac{1}{N-1}\mathbb{E}|Wf_{z}(W)| \leq \frac{\sqrt{2\pi \mathbb{E}W^{2}}}{4(N-1)} = \frac{\sqrt{2\pi}}{4\sqrt{N(N-1)}} \leq \frac{1}{N},$$

where we have used N > 100 in the last inequality. By the definition of  $g_z$ , we have

(3.10) 
$$\mathbb{E}Wf_z(W) - \mathbb{E}W^*f_z(W^*) = -\mathbb{E}\int_0^{W^*-W} g_z(W+t) dt.$$

For  $n \le m$ , from (3.15) of Lemma 3.2 below, we see that  $x_n$  is bounded by an absolute constant as long as  $\log(m/n)$  is bounded. To bound the right-hand side of (3.10), we separate W around  $m/e^3$  (we can replace  $e^3$  by any constant c > 1, the number  $e^3$  is only for convenience when we compute the explicit constant in (2.2)) as follows:

(3.11) 
$$R_1 = \mathbb{E} \left| \int_0^{W^* - W} g_z(W + t) \mathbf{1}(W \le -x_{\lfloor m/e^3 \rfloor}) dt \right|,$$

(3.12) 
$$R_2 = \mathbb{E} \left| \int_0^{W^* - W} g_z(W + t) \mathbf{1} \left( |W| < x_{\lfloor m/e^3 \rfloor} \right) dt \right|,$$

and

(3.13) 
$$R_3 = \mathbb{E}\left|\int_0^{W^* - W} g_z(W + t) \mathbf{1}(W \ge x_{\lfloor m/e^3 \rfloor}) dt\right|.$$

We need the following lemma whose proof will be presented in the Appendix.

LEMMA 3.2. Let N > 100. The following statements hold.

$$(3.14) 0 \le x_m \le \frac{1}{m}$$

For  $1 \le n \le m - 1$ , we have

(3.15) 
$$x_n \le \sqrt{2(1 + \log(m/n))},$$

and

(3.16) 
$$\left(\frac{n(n+1)}{2}\right)^{1/2} \le S_n \le \frac{3n}{2}\sqrt{2(1+\log(m/n))}.$$

For  $1 \le n \le \lfloor m/e^3 \rfloor$ , we have

(3.17) 
$$x_n \ge \frac{1}{3}\sqrt{2(1 + \log(m/n))} \quad and \quad S_n \ge \frac{n}{3}\sqrt{2(1 + \log(m/n))}.$$

For  $1 \le i < j \le \lfloor m/e^3 \rfloor$ , we have

(3.18) 
$$x_i^2 - x_j^2 \ge \frac{4\log(j/i)}{9}.$$

REMARK 3.3. By using the first half of (3.17), we have  $x_{\lfloor m/e^3 \rfloor} \ge \sqrt{8}/3$ . We note also that m > 50 and  $S_j \ge jx_j$ ,  $1 \le j \le m$ . These simple inequalities will be used in many places later without further mention. For  $x_1$ , McKeague and Levin [12] proved the following lower bound:

$$(3.19) S_1 = x_1 \ge \sqrt{\log(m)},$$

which is of the same order as ours but with a better constant, but their method seems not to work with  $x_n$  for n > 1.

We will now bound  $R_1$ ,  $R_2$ , and  $R_3$ . From (1.1) and the first half of (3.16), we have  $x_1 - x_2 = 1/S_1 \le 1$ , and

$$x_{n-1} - x_{n+1} = \frac{1}{S_{n-1}} + \frac{1}{S_n} \le \left(\frac{2}{(n-1)n}\right)^{1/2} + \left(\frac{2}{n(n+1)}\right)^{1/2} \le 1$$

for  $3 \le n \le m$ , or equivalently,

(3.20) 
$$x_1 \le x_2 + 1$$
 and  $x_{n-1} \le x_{n+1} + 1$  for  $3 \le n \le m$ .

Let t be a real number lying between 0 and  $W^* - W$ . From the definitions of W and  $W^*$ , we observe the following facts.

- FACTS 3.4.
  - (i) If  $W = x_1$ , then

$$|W^* - W| \le x_1 - x_2$$
 and  $x_2 \le W + t \le x_1;$ 

(ii) If  $W = x_n, 2 \le n \le m$ , then

$$|W^* - W| \le x_{n-1} - x_n$$
 and  $x_{n+1} \le W + t \le x_{n-1}$ ;

and, by symmetry, we have

(iii) If 
$$W = x_N = -x_1$$
, then  
 $|W^* - W| \le x_1 - x_2$  and  $-x_1 \le W + t \le -x_2$ ;  
(iv) If  $W = x_{N+1-n} = -x_n$ ,  $2 \le n \le m$ , then  
 $|W^* - W| \le x_{n-1} - x_n$  and  $-x_{n-1} \le W + t \le -x_{n+1}$ .

Keeping Facts 3.4 and the properties of  $g_z$  in mind, we have

$$R_{1} \leq \frac{1}{N} \left( (x_{1} - x_{2})g_{z}(-x_{2}) + \sum_{j=2}^{\lfloor m/e^{3} \rfloor} (x_{j-1} - x_{j})g_{z}(-x_{j+1}) \right)$$

$$\leq \frac{3}{2N} \left( \frac{x_{1} - x_{2}}{(1 + x_{2})^{3}} + \sum_{j=2}^{\lfloor m/e^{3} \rfloor} \frac{x_{j-1} - x_{j}}{(1 + x_{j+1})^{3}} \right)$$

$$= \frac{3}{2N} \left( \frac{1}{x_{1}(1 + x_{2})^{3}} + \sum_{j=2}^{\lfloor m/e^{3} \rfloor} \frac{1}{(1 + x_{j+1})^{3}S_{j-1}} \right)$$

$$\leq \frac{3}{2N} \left( \frac{1}{x_{1}^{4}} + \sum_{j=2}^{\lfloor m/e^{3} \rfloor} \frac{1}{(j-1)x_{j-1}^{4}} \right)$$

$$\leq \frac{3}{2N} \left( \frac{2}{x_{1}^{4}} + \sum_{j=2}^{\lfloor m/e^{3} \rfloor - 1} \frac{81}{4j(1 + \log(m/j))^{2}} \right)$$

$$\leq \frac{3}{2N} \left( \frac{2}{\log^{2}(m)} + \int_{1}^{m/e^{3}} \frac{81 \, dx}{4x(1 + \log(m/x))^{2}} \right)$$

$$= \frac{3}{2N} \left( \frac{2}{\log^{2}(m)} + \frac{81}{16} - \frac{81}{4(1 + \log(m))} \right) \leq \frac{243}{32N},$$

where we have applied (3.5) in the first inequality, and (3.6) in the second inequality, (3.20) and simple bounds  $S_j \ge jx_j$   $(1 \le j \le m)$  in the third inequality, the first half of (3.17) and (3.19) in the fourth inequality. The final bound in (3.21) follows from m > 50 and the elementary inequality  $2/\log^2(m) \le 81/(4(1 + \log(m))))$ . For  $R_2$ , we have

$$R_{2} \leq \frac{2}{N} \sum_{0 \leq x_{j} < \lfloor m/e^{3} \rfloor} (x_{j-1} - x_{j})(x_{j-1} + \sqrt{2\pi}/4)$$

$$= \frac{2}{N} \sum_{\lfloor m/e^{3} \rfloor < j \leq m} \left(\frac{x_{j-1}}{S_{j-1}} + \frac{\sqrt{2\pi}}{4S_{j-1}}\right)$$

$$\leq \frac{2}{N} \sum_{\lfloor m/e^{3} \rfloor < j \leq m} \left(\frac{1}{j-1} + \frac{\sqrt{\pi}}{2(j-1)}\right)$$

$$\leq \frac{2 + \sqrt{\pi}}{N} \left(\frac{1}{\lfloor m/e^{3} \rfloor} + \frac{1}{\lfloor m/e^{3} \rfloor + 1} + \int_{m/e^{3}}^{m} \frac{dx}{x}\right)$$

$$\leq \frac{2 + \sqrt{\pi}}{N} \left(\frac{1}{2} + \frac{1}{3} + 3\right) \leq \frac{15}{N},$$

where we have applied (3.4) and Facts 3.4 in the first inequality, the first half of (3.16) and simple bounds  $x_j/S_j \le 1/j$  ( $1 \le j \le m$ ) in the second inequality, and used m > 50 in the fourth inequality.

To bound  $R_3$ , we set  $\ell = \min\{j : x_j \le z\}$  and consider the following two cases.

*Case 1*:  $x_{\ell} < x_{\lfloor m/e^3 \rfloor + 1}$ . In this case, we have  $z < x_{\lfloor m/e^3 \rfloor + 1}$ . If  $W = x_j \ge x_{\lfloor m/e^3 \rfloor}$ , then, by Facts 3.4,

(3.23)  $W + t > x_{j+1} > z$  for all t lying between 0 and  $W^* - W$ .

Therefore,

(3.24)  

$$R_{3} = \mathbb{E} \left| \int_{0}^{W^{*}-W} g_{z}(W+t) \mathbf{1}(W \ge x_{\lfloor m/e^{3} \rfloor}) dt \right|$$

$$\leq \frac{1}{N} \left( (x_{1} - x_{2})g_{z}(x_{2}) + \sum_{j=2}^{\lfloor m/e^{3} \rfloor} (x_{j-1} - x_{j})g_{z}(x_{j+1}) \right)$$

$$\leq \frac{3}{N} \left( \frac{x_{1} - x_{2}}{(1+x_{2})^{3}} + \sum_{j=2}^{\lfloor m/e^{3} \rfloor} \frac{x_{j-1} - x_{j}}{(1+x_{j+1})^{3}} \right) \leq \frac{243}{16N},$$

where we have applied (3.5) and (3.23) in the first inequality, and (3.7) and (3.23) in the second inequality. The last inequality in (3.24) follows by using the same calculations as in (3.21).

*Case 2*:  $x_{\ell} \ge x_{\lfloor m/e^3 \rfloor + 1}$ . In this case, we have  $z \ge x_{\lfloor m/e^3 \rfloor + 1}$ , and  $R_3$  is bounded by  $R_{31} + R_{32} + R_{33}$ , where

(3.25) 
$$R_{31} = \mathbb{E} \left| \int_0^{W^* - W} g_z(W + t) \mathbf{1}(x_{\lfloor m/e^3 \rfloor} \le W < x_\ell) \, dt \right|,$$

(3.26) 
$$R_{32} = \mathbb{E} \left| \int_0^{W^* - W} g_z(W + t) \mathbf{1} (x_\ell \le W \le x_{\ell-1}) \, dt \right|,$$

and

(3.27) 
$$R_{33} = \mathbb{E} \left| \int_0^{W^* - W} g_z(W + t) \mathbf{1}(W > x_{\ell-1}) dt \right|$$

Here and thereafter, we denote  $x_0 = x_1$  and  $S_0 = S_1$ . We of course only need to bound  $R_{31}$  when  $x_{\lfloor m/e^3 \rfloor} < x_\ell$ . Since  $x_\ell \le z$ , we conclude that if  $x_{\lfloor m/e^3 \rfloor} \le W = x_{j+1} < x_\ell$ , then by Facts 3.4, we have

(3.28) 
$$|W^* - W| \le x_j - x_{j+1}$$
 and  $0 < W + t < x_j \le x_\ell \le z$ 

for all t lying between 0 and  $W^* - W$ . Applying (3.3), (3.28), and Facts 3.4, we have

(3.29) 
$$R_{31} \leq \frac{1 - \Phi(x_{\ell})}{N} \left( \sum_{j=\ell}^{\lfloor m/e^3 \rfloor - 1} (x_j - x_{j+1}) \left( \sqrt{2\pi} \left( 1 + x_j^2 \right) e^{x_j^2/2} + x_j \right) \right).$$

Since  $1 - \Phi(x_\ell) \le e^{-x_\ell^2/2}/(x_\ell\sqrt{2\pi})$ , it follows from (3.29) that

(3.30) 
$$R_{31} \leq \frac{1}{N} \sum_{j=\ell}^{\lfloor m/e^3 \rfloor - 1} \left( \frac{1 + x_j^2}{S_j x_\ell} e^{(x_j^2 - x_\ell^2)/2} + \frac{x_j e^{-x_\ell^2/2}}{\sqrt{2\pi} S_j x_\ell} \right).$$

For all  $\ell \leq j \leq \lfloor m/e^3 \rfloor$ , by applying (3.17), we have

$$(3.31) S_j x_\ell \ge S_j x_j \ge 8j/9.$$

By using (3.17), (3.18), (3.31), and simple inequalities  $x_{\ell} \ge x_{\lfloor m/e^3 \rfloor} \ge \sqrt{8}/3$  and  $x_j/S_j \le 1/j$ ( $1 \le j \le m$ ), we have

$$(3.32) \qquad \frac{1+x_{j}^{2}}{S_{j}x_{\ell}}e^{(x_{j}^{2}-x_{\ell}^{2})/2} + \frac{x_{j}e^{-x_{\ell}^{2}/2}}{\sqrt{2\pi}S_{j}x_{\ell}} \\ \leq \left(\frac{1}{S_{j}x_{j}} + \frac{x_{j}}{S_{j}}\right)\left(\frac{\ell}{j}\right)^{2/9} + \frac{3x_{j}}{4\sqrt{\pi}S_{j}}\exp\left(-\frac{1}{2}\left(\frac{\sqrt{2\log(m/\ell)}}{3}\right)^{2}\right) \\ \leq \left(\frac{9}{8j} + \frac{1}{j}\right)\left(\frac{\ell}{j}\right)^{2/9} + \frac{3}{4\sqrt{\pi}j}\left(\frac{\ell}{m}\right)^{1/9}$$

for all  $\ell \le j \le \lfloor m/e^3 \rfloor$ . It follows from (3.30) and (3.32) that

$$R_{31} \leq \frac{1}{N} \left( \frac{17}{8} \sum_{j=\ell}^{\lfloor m/e^3 \rfloor - 1} \frac{\ell^{2/9}}{j^{11/9}} + \frac{3e^{-1/3}}{4\sqrt{\pi}} \sum_{j=\ell}^{\lfloor m/e^3 \rfloor - 1} \frac{\ell^{1/9}}{j^{10/9}} \right)$$

$$\leq \frac{1}{N} \left( \frac{17}{8} \left( \frac{1}{\ell} + \int_{\ell}^{m/e^3} \frac{\ell^{2/9} dx}{x^{11/9}} \right) + \frac{3e^{-1/3}}{4\sqrt{\pi}} \left( \frac{1}{\ell} + \int_{\ell}^{m/e^3} \frac{\ell^{1/9} dx}{x^{10/9}} \right) \right)$$

$$\leq \frac{1}{N} \left( \frac{17}{8} \left( 1 + \frac{9}{2} \right) + \frac{30e^{-1/3}}{4\sqrt{\pi}} \right) \leq \frac{15}{N}.$$

Now, we bound  $R_{32}$  and  $R_{33}$ . If  $x_{\ell} \ge x_3$ , then

$$R_{32} + R_{33} \leq \sum_{i=1}^{3} \mathbb{E} \left| \int_{0}^{W^{*}-W} g_{z}(W+t) \mathbf{1}(W=x_{i}) dt \right|$$

$$\leq \frac{1}{N} \left( 2(x_{1}-x_{2})\left(x_{1}+\frac{\sqrt{2\pi}}{4}\right) + (x_{2}-x_{3})\left(x_{2}+\frac{\sqrt{2\pi}}{4}\right) \right)$$

$$= \frac{1}{N} \left( \frac{2(x_{1}+\sqrt{2\pi}/4)}{S_{1}} + \frac{x_{2}+\sqrt{2\pi}/4}{S_{2}} \right) \leq \frac{6}{N},$$

where we have applied (3.4) and Facts 3.4 in the second inequality, and the first half of (3.16) and a simple bound  $x_2/S_2 \le 1/2$  in the last inequality. If  $x_\ell \le x_4$ , then, similar to (3.34), we have

(3.35)  

$$R_{32} = \mathbb{E} \left| \int_{0}^{W^{*}-W} g_{z}(W+t) \mathbf{1} (x_{\ell} \le W \le x_{\ell-1}) dt \right|$$

$$\leq \frac{1}{N} \left( \frac{x_{\ell-1} + \sqrt{2\pi}/4}{S_{\ell-1}} + \frac{x_{\ell-2} + \sqrt{2\pi}/4}{S_{\ell-2}} \right)$$

$$\leq \frac{1}{N} \left( \frac{x_{3}}{S_{3}} + \frac{\sqrt{2\pi}/4}{S_{3}} + \frac{x_{2}}{S_{2}} + \frac{\sqrt{2\pi}/4}{S_{2}} \right)$$

$$\leq \frac{1}{N} \left( \frac{1}{3} + \frac{\sqrt{2\pi}/4}{S_{3}} + \frac{1}{2} + \frac{\sqrt{2\pi}/4}{S_{2}} \right) \le \frac{2}{N},$$

and

(3.36)  

$$R_{33} = \mathbb{E} \left| \int_{0}^{W^{*}-W} g_{z}(W+t) \mathbf{1}(W > x_{\ell-1}) dt \right|$$

$$\leq \frac{3}{N} \left( \frac{x_{1}-x_{2}}{(1+x_{2})^{3}} + \sum_{j=2}^{\ell-2} \frac{x_{j-1}-x_{j}}{(1+x_{j+1})^{3}} \right)$$

$$\leq \frac{243}{16N},$$

where, in (3.36), we have applied Facts 3.4 and (3.7) in the first inequality, and used the same calculations as in (3.21) in the second inequality. Combining (3.33)–(3.36), we have

$$(3.37) R_3 \le \frac{515}{16N}.$$

It follows from (3.21), (3.22), and (3.37) that

(3.38) 
$$\left| \mathbb{E}Wf_{z}(W) - \mathbb{E}W^{*}f_{z}(W^{*}) \right| \leq R_{1} + R_{2} + R_{3} \leq \frac{1753}{32N}.$$

Combining (3.8), (3.9), and (3.38), we have

(3.39) 
$$\left|\mathbb{P}(W^* \le z) - \Phi(z)\right| \le \frac{1}{N} + \frac{1753}{32N} = \frac{1785}{32N}.$$

The upper bound of (2.2) follows from (3.2) and (3.39).  $\Box$ 

PROOF OF THE LOWER BOUND IN (2.2). The proof of the lower bound in (2.2) follows from the fact that the Kolmogorov distance between  $\gamma$  and the probability distribution of a discrete random variable is always greater than half of the minimum of the jumps. In our case, the jumps are all equal to 1/N. Therefore,

$$d_{\mathbf{K}}(\mathbb{P}_N,\gamma) \ge \frac{1}{2N}.$$

**4.** Bounds on the Wasserstein distance. PROOF OF (2.3). By Theorem 1.1 of Goldstein [8], we have

$$d_{W}(\mathbb{P}_{N}, \gamma) \leq 2\mathbb{E}|W - W^{*}|$$

$$\leq \frac{2}{N-1} \sum_{n=1}^{N-1} (x_{n} - x_{n+1})$$

$$= \frac{4x_{1}}{N-1} \leq \frac{4\sqrt{2(1 + \log(m))}}{N-1}$$

$$\leq \frac{8\sqrt{\log N}}{N-1} \leq \frac{16\sqrt{\log N}}{N},$$

where we have applied (3.15) in the third inequality, and definition of *m* in the fourth inequality. The proof of the upper bound of (2.3) is complete.

For the lower bound of (2.3), we can assume N > 100 since it is trivial when  $N \le 100$  and  $C \ge 3$ . We note that  $d_W(\mathbb{P}_N, \gamma) \ge |\mathbb{E}h(W) - \mathbb{E}h(Z)|$  for any 1-Lipschitz function h. We will use the "sawtooth" piecewise linear function considered in McKeague and Levin [12], which is defined as follows:

$$h(w) = \begin{cases} 0 & \text{if } w > x_1 \text{ or } w < x_N, \\ w - x_{n+1} & \text{if } x_{n+1} \le w < m_n, 1 \le n < N, \\ x_n - w & \text{if } m_n \le w < x_n, 1 \le n < N, \end{cases}$$

where  $m_n = (x_n + x_{n+1})/2$ ,  $1 \le n < N$ . Clearly *h* is 1-Lipschitz and  $\mathbb{E}h(W) = 0$ . Simple calculation gives

(4.2) 
$$\mathbb{E}h(W^*) = \frac{x_1}{2(N-1)} \ge \frac{\sqrt{\log(N/2)}}{2(N-1)} \ge \frac{\sqrt{\log(N/2)}}{2N},$$

where we have applied (3.19) in the first inequality.

For this "sawtooth" function h, let  $f_h$  be the unique bounded solution of the Stein equation

$$f'(w) - wf(w) = h(w) - \mathbb{E}h(Z),$$

and let

$$g_h(w) = (wf_h(w))'.$$

By Lemma 2.4 in [4], we have  $||f_h|| \le 2$  and  $||f'_h|| \le \sqrt{2/\pi}$ . Therefore,

(4.3) 
$$|g_h(w)| \le |f_h(w)| + |wf'(w)| \le 2(1+|w|).$$

It thus follows from (4.2) that

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| = |\mathbb{E}f_h'(W) - \mathbb{E}Wf_h(W)|$$

$$= \left|\mathbb{E}f_h'(W) - \left(1 - \frac{1}{N}\right)\mathbb{E}f_h'(W^*)\right|$$

$$= \left|\mathbb{E}Wf_h(W) - \mathbb{E}W^*f_h(W^*) - \mathbb{E}h(W^*) + \frac{1}{N}\mathbb{E}f_h'(W^*)\right|$$

$$\geq |\mathbb{E}h(W^*)| - |\mathbb{E}Wf_h(W) - \mathbb{E}W^*f_h(W^*)| - \left|\frac{1}{N}\mathbb{E}f_h'(W^*)\right|$$

$$\geq \frac{\sqrt{\log(N/2)}}{2N} - |\mathbb{E}Wf_h(W) - \mathbb{E}W^*f_h(W^*)| - \frac{\sqrt{2/\pi}}{N}.$$

The lower bound of (2.3) will follow if we can show that

$$\left|\mathbb{E}Wf_h(W) - \mathbb{E}W^*f_h(W^*)\right| \leq \frac{C}{N}.$$

Letting  $T = W^* - W$ , we have

(4.5) 
$$\mathbb{E} |Wf_h(W) - W^* f_h(W^*)| = \mathbb{E} (|Tg_h(W+\xi)|),$$

where  $\xi$  is a random variable lying between 0 and *T*. We will obtain an upper bound on  $\mathbb{E}(|Tg_h(W + \xi)|)$  using the same truncation for *W* as in the proof for the Kolmogorov bound. So we let

(4.6) 
$$R_1 = \mathbb{E}(|Tg_h(W+\xi)|\mathbf{1}(|W| \le x_{\lfloor m/e^3 \rfloor})),$$

and

(4.8)

(4.7) 
$$R_2 = \mathbb{E}(|Tg_h(W+\xi)|\mathbf{1}(|W| > x_{\lfloor m/e^3 \rfloor})).$$

For  $R_1$ , we have

$$R_{1} \leq 2\mathbb{E}(|T|(1+|W+\xi|)\mathbf{1}(|W| \leq x_{\lfloor m/e^{3} \rfloor}))$$
  
$$\leq \frac{C}{N} \sum_{i=\lfloor m/e^{3} \rfloor}^{m} (x_{i-1}-x_{i})(1+x_{i-1})$$
  
$$= \frac{C}{N} \sum_{i=\lfloor m/e^{3} \rfloor}^{m} \left(\frac{1}{S_{i-1}} + \frac{x_{i-1}}{S_{i-1}}\right)$$
  
$$\leq \frac{C}{N} \sum_{i=\lfloor m/e^{3} \rfloor}^{m} \left(\frac{1}{S_{i-1}} + \frac{1}{i-1}\right),$$

where we have applied (4.3) in the first inequality, Facts 3.4 in the second inequality, and simple bounds  $x_j/S_j \le 1/j$   $(1 \le j \le m)$  in the third inequality. If  $\lfloor m/e^3 \rfloor \le i \le m$ , then by using the second half of (3.17), we have  $S_{i-1} \ge S_{\lfloor m/e^3 \rfloor - 1} \ge m/C$ . Therefore, (4.8) implies

(4.9) 
$$R_1 \leq \frac{C}{N} \sum_{i=\lfloor m/e^3 \rfloor}^m \left(\frac{1}{m} + \frac{1}{i-1}\right) \leq \frac{C}{N}.$$

To bound  $R_2$ , we need the following lemma.

LEMMA 4.1. If either 
$$x_{\lfloor m/e^3 \rfloor} \le x_{n+1} < w \le x_n \text{ or } -x_n \le w < -x_{n+1} \le -x_{\lfloor m/e^3 \rfloor}$$
, then  
(4.10)  $|g_h(w)| \le C \left(\frac{1}{\log(m/n)} + \frac{1}{n^{2/9}}\right).$ 

We postpone the proof of Lemma 4.1 to the Appendix. Now, we bound  $R_2$  as follows:

$$R_{2} \leq \frac{C}{N} \sum_{n=2}^{\lfloor m/e^{3} \rfloor - 1} (x_{n-1} - x_{n}) \left( \frac{1}{\log(m/n)} + \frac{1}{n^{2/9}} \right)$$

$$= \frac{C}{N} \sum_{n=2}^{\lfloor m/e^{3} \rfloor - 1} \left( \frac{1}{\log(m/n)S_{n-1}} + \frac{1}{n^{2/9}S_{n-1}} \right)$$

$$\leq \frac{C}{N} \sum_{n=1}^{\lfloor m/e^{3} \rfloor - 2} \left( \frac{1}{n \log^{3/2}(m/n)} + \frac{1}{n^{11/9} \log^{1/2}(m/n)} \right)$$

$$\leq \frac{C}{N},$$
(4.11)

where we have applied Facts 3.4 and Lemma 4.1 in the first inequality, and the second half of (3.17) in the second inequality. It follows from (4.9) and (4.11) that

(4.12) 
$$\mathbb{E} |Wf_h(W) - W^* f_h(W^*)| \leq \frac{C}{N}.$$

Combining (4.4) and (4.12), we have

$$d_{\mathrm{W}}(\mathbb{P}_N, \gamma) \ge \frac{\sqrt{\log(N/2)}}{2N} - \frac{C}{N}$$

thereby proving the lower bound of (2.3).  $\Box$ 

## APPENDIX

In this Appendix, we will prove those lemmas in the previous sections, whose proofs have been deferred to the Appendix.

PROOF OF LEMMA 3.1. We first prove that

(A.1) 
$$\sqrt{2\pi} w e^{w^2/2} (1 - \Phi(w)) \le \frac{w^2 + 2}{w^2 + 3}$$

for w > 0. This inequality is equivalent to

(A.2) 
$$\int_{w}^{\infty} e^{-x^{2}/2} dx \le \frac{w^{2}+2}{w^{3}+3w} e^{-w^{2}/2}$$

Let

$$h(w) = \frac{w^2 + 2}{w^3 + 3w} e^{-w^2/2}.$$

Then

$$h(w) = \int_{w}^{\infty} \left(-h'(x)\right) dx.$$

For x > 0, we have

$$-h'(x) = \left(1 + \frac{6}{(x^3 + 3x)^2}\right)e^{-x^2/2} \ge e^{-x^2/2}$$

thereby proving (A.2). The proof of (A.1) is complete.

From (3.3) and (A.1), we have for w > z,

$$g'_{z}(w) = \left(\sqrt{2\pi} \left(w^{2} + 3\right) w e^{w^{2}/2} \left(1 - \Phi(w)\right) - 2 - w^{2}\right) \Phi(z) \le 0,$$

which proves  $g_z(w)$  is decreasing for w > z. Similarly,  $g_z(w)$  is increasing for  $w \le z$  since in this case,

$$g'_{z}(w) = (2 + w^{2} - \sqrt{2\pi}(w^{2} + 3)(-w)e^{w^{2}/2}(1 - \Phi(-w)))(1 - \Phi(z)) \ge 0.$$

The proof of (3.5) is complete.

Now we prove that for w > 0,

(A.3) 
$$0 < \sqrt{2\pi} (1+w^2) e^{w^2/2} (1-\Phi(w)) - w \le \frac{3}{(1+w)^3}.$$

The first inequality in (A.3) is proved by Chen and Shao [5]. The second inequality in (A.3) is equivalent to

(A.4) 
$$\int_{w}^{\infty} e^{-x^{2}/2} dx \leq \left(\frac{w}{1+w^{2}} + \frac{3}{(1+w^{2})(1+w)^{3}}\right) e^{-w^{2}/2}.$$

838

Let

$$k(w) = \left(\frac{w}{1+w^2} + \frac{3}{(1+w^2)(1+w)^3}\right)e^{-w^2/2}, \quad w > 0.$$

Then

$$k(w) = \int_{w}^{\infty} \left(-k'(x)\right) dx$$

For x > 0, we have

$$-k'(x) = \left(1 + \frac{x^4 - 5x^3 + 6x^2 + x + 7}{(1 + x^2)^2(1 + x)^4}\right)e^{-x^2/2} \ge e^{-x^2/2}$$

thereby proving (A.4). Combining (3.3) and (A.3), and noting that  $1 - \Phi(z) \le 1/2$ , we have

$$0 \le g_z(w) \le \frac{3}{(1+w)^3}$$
 for  $w > z$ ,

and

$$0 \le g_z(w) \le \frac{3}{2(1-w)^3}$$
 for  $w < 0$ .

This ends the proof of (3.6) and (3.7).

PROOF OF LEMMA 3.2. From (P3) and (P4) of Lemma 2.1, it is easy to see that  $x_m = 0$  if N is odd, and  $2x_m = x_m - x_{m+1} > 0$  if N is even. So we only need to prove the second inequality in (3.14) for the case where N is even. By (1.1), we have

$$x_j = \sum_{i=j}^{m-1} (x_i - x_{i+1}) + x_m \ge \sum_{i=j}^{m-1} \frac{1}{S_i}, \quad 1 \le j \le m-1.$$

Therefore,  $S_1 \ge 1$  and

(A.5)  
$$S_{j} \geq \frac{1}{S_{1}} + \frac{2}{S_{2}} + \dots + \frac{j-1}{S_{j-1}} + j \sum_{i=j}^{m-1} \frac{1}{S_{i}}$$
$$\geq \frac{1}{S_{j}} + \frac{2}{S_{j}} + \dots + \frac{j}{S_{j}} = \frac{j(j+1)}{2S_{j}},$$

for  $2 \le j \le m - 1$ . This implies that

(A.6) 
$$S_j \ge \left(\frac{j(j+1)}{2}\right)^{1/2}, \quad 1 \le j \le m-1.$$

Since  $x_m \ge 0$ ,

$$S_m \ge S_{m-1} \ge \left(\frac{m(m-1)}{2}\right)^{1/2}.$$

If *N* is even, then  $x_m = -x_{m+1}$ . So

$$2x_m = x_m - x_{m+1} = \frac{1}{S_m} \le \left(\frac{2}{m(m-1)}\right)^{1/2}.$$

Therefore,

$$0 \le x_m \le \left(\frac{1}{2m(m-1)}\right)^{1/2} \le \frac{1}{m}.$$

This proves (3.14).

Since  $\{x_1, \ldots, x_N\}$  is decreasing,  $jx_j \le S_j \le jx_1$ . For  $1 \le n \le m - 1$ , by using the telescoping sum and (3.14), we have

(A.7)  

$$x_n^2 = \sum_{j=n}^{m-1} (x_j^2 - x_{j+1}^2) + x_m^2 \le 2 \sum_{j=n}^{m-1} x_j (x_j - x_{j+1}) + x_m^2$$

$$= 2 \sum_{j=n}^{m-1} \frac{x_j}{S_j} + x_m^2 \le 2 \sum_{j=n}^m \frac{1}{j}$$

$$\le 2(1/n + \log(m/n))$$

$$\le 2(1 + \log(m/n)).$$

This proves (3.15). For  $2 \le n \le m - 1$ ,

(A.8)  

$$S_n = x_1 + \dots + x_n$$

$$\leq \sqrt{2} \left( \sqrt{1 + \log(m)} + \sqrt{1 + \log(m/2)} + \dots + \sqrt{1 + \log(m/n)} \right)$$

$$\leq \sqrt{2} \left( \sqrt{1 + \log(m)} + \int_1^n \sqrt{1 + \log(m/x)} \, dx \right).$$

Set

$$I_n = \int_1^n \sqrt{1 + \log(m/x)} \, dx.$$

By using integration by parts, we have

(A.9)  
$$I_n = n\sqrt{1 + \log(m/n)} - \sqrt{1 + \log(m)} + \frac{1}{2} \int_1^n \frac{dx}{\sqrt{1 + \log(m/x)}}$$
$$\leq n\sqrt{1 + \log(m/n)} - \sqrt{1 + \log(m)} + \frac{n - 1}{2\sqrt{1 + \log(m/n)}}.$$

Combining (A.8) and (A.9),

(A.10)  
$$S_n \le \sqrt{2} \left( n \sqrt{1 + \log(m/n)} + \frac{n-1}{2\sqrt{1 + \log(m/n)}} \right) \le \frac{3n}{2} \sqrt{2(1 + \log(m/n))},$$

which proves the second half of (3.16). The first half of (3.16) coincides with (A.6).

Using (A.10), we calculate lower bounds of  $x_n$  and  $S_n$  for  $1 \le n \le \lfloor m/e^3 \rfloor$  as follows:

(A.11)  
$$x_{n} = \sum_{j=n}^{m-1} (x_{j} - x_{j+1}) + x_{m} \ge \sum_{j=n}^{m-1} \frac{1}{S_{j}}$$
$$\ge \frac{\sqrt{2}}{3} \sum_{j=n}^{m-1} \frac{1}{j\sqrt{1 + \log(m/j)}}$$
$$\ge \frac{\sqrt{2}}{3} \int_{n}^{m} \frac{dx}{x\sqrt{1 + \log(m/x)}}$$
$$\ge \frac{1}{3}\sqrt{2(1 + \log(m/n))}.$$

For the lower bound of  $S_n$ , by (A.11), we have

$$S_n \ge nx_n \ge \frac{n}{3}\sqrt{2(1+\log(m/n))}.$$

The proof of (3.17) is complete.

By the second half of (3.16) and the first half of (3.17), we have

$$\begin{aligned} x_i^2 - x_j^2 &= (x_i + x_j)(x_i - x_{i+1} + \dots + x_{j-1} - x_j) \\ &= (x_i + x_j) \left(\frac{1}{S_i} + \dots + \frac{1}{S_{j-1}}\right) \\ &\ge \frac{2(\sqrt{1 + \log(m/i)} + \sqrt{1 + \log(m/j)})}{9} \left(\sum_{k=i}^{j-1} \frac{1}{k\sqrt{1 + \log(m/k)}}\right) \\ &\ge \frac{2(\sqrt{1 + \log(m/i)} + \sqrt{1 + \log(m/j)})}{9} \int_i^j \frac{dx}{x\sqrt{1 + \log(m/x)}} \\ &= \frac{4\log(j/i)}{9}. \end{aligned}$$

This proves (3.18).  $\Box$ 

PROOF OF LEMMA 4.1. We only prove (4.10) for the case where  $x_{\lfloor m/e^3 \rfloor} \le x_{n+1} < w \le x_n$ . The proof for the other case is similar and therefore omitted. We have (see [4], page 40)

(A.12)  

$$g_{h}(w) = \left(w - \sqrt{2\pi} \left(1 + w^{2}\right) e^{w^{2}/2} \left(1 - \Phi(w)\right)\right) \int_{-\infty}^{w} h'(t) \Phi(t) dt$$

$$= \left(w + \sqrt{2\pi} \left(1 + w^{2}\right) e^{w^{2}/2} \Phi(w)\right) \int_{w}^{\infty} h'(t) \left(1 - \Phi(t)\right) dt$$

$$= I_{1} + I_{2}.$$

Applying the identity

$$\int_{-\infty}^{w} \Phi(t) dt = w \Phi(w) + \frac{e^{-w^2/2}}{\sqrt{2\pi}}$$

and noting that ||h'|| = 1, we have

(A.13) 
$$\left| \int_{-\infty}^{w} h'(t) \Phi(t) \, dt \right| \le w \Phi(w) + \frac{e^{-w^2/2}}{\sqrt{2\pi}} \le 1 + w.$$

Therefore, applying the first half of (3.17) and (A.3), and noting that  $x_{\lfloor m/e^3 \rfloor} \le x_{n+1} < w \le x_n$ , we obtain

(A.14) 
$$|I_1| \le \frac{3}{(1+w)^2} \le \frac{C}{\log(m/n)}$$

For all  $1 \le i \le N-1$ , we have h'(t) = 1 on  $(x_{i+1}, m_i)$  and h'(t) = -1 on  $(m_i, x_i)$ . Therefore, for all  $1 \le i \le \lfloor m/e^3 \rfloor - 1$ , we have

(A.15)  
$$\left| \int_{x_{i+1}}^{x_i} h'(t) (1 - \Phi(t)) dt \right| = \left| \int_{x_{i+1}}^{x_i} h'(t) \Phi(t) dt \right|$$
$$= \left| \int_{x_{i+1}}^{m_i} \Phi(t) dt - \int_{m_i}^{x_i} \Phi(t) dt \right|$$
$$= \int_{m_i}^{x_i} \Phi(t) dt - \int_{x_{i+1}}^{m_i} \Phi(t) dt$$
$$\leq \frac{x_i - x_{i+1}}{2} (\Phi(x_i) - \Phi(x_{i+1}))$$
$$\leq \frac{(x_i - x_{i+1})^2 e^{-x_{i+1}^2/2}}{2\sqrt{2\pi}}$$
$$= \frac{e^{-x_{i+1}^2/2}}{2\sqrt{2\pi}S_i^2},$$

where we have applied the monotonicity of function  $\Phi(\cdot)$  in the first inequality, the mean value theorem in the second inequality, and (1.1) in the last equality. It follows from (A.15) that

$$\begin{aligned} \left| \int_{w}^{\infty} h'(t) (1 - \Phi(t)) dt \right| &= \left| \int_{w}^{x_{1}} h'(t) (1 - \Phi(t)) dt \right| \\ &\leq \int_{w}^{x_{n}} (1 - \Phi(t)) dt + \sum_{i=1}^{n-1} \left| \int_{x_{i+1}}^{x_{i}} h'(t) (1 - \Phi(t)) dt \right| \\ &\leq (x_{n} - w) (1 - \Phi(w)) + \sum_{i=1}^{n-1} \frac{e^{-x_{i+1}^{2}/2}}{2\sqrt{2\pi}S_{i}^{2}} \\ &\leq \frac{(x_{n} - x_{n+1})e^{-w^{2}/2}}{w\sqrt{2\pi}} + \sum_{i=1}^{n-1} \frac{e^{-x_{i+1}^{2}/2}}{2\sqrt{2\pi}S_{i}^{2}}. \end{aligned}$$

Therefore

$$|I_{2}| = \left(w + \sqrt{2\pi} (1 + w^{2}) e^{w^{2}/2} \Phi(w)\right) \left| \int_{w}^{\infty} h'(t) (1 - \Phi(t)) dt \right|$$

$$\leq C w^{2} e^{w^{2}/2} \left( \frac{(x_{n} - x_{n+1}) e^{-w^{2}/2}}{w\sqrt{2\pi}} + \sum_{i=1}^{n-1} \frac{e^{-x_{i+1}^{2}/2}}{2\sqrt{2\pi} S_{i}^{2}} \right)$$

$$\leq C \left( \frac{x_{n}}{S_{n}} + \sum_{i=1}^{n-1} \frac{x_{i}^{2}}{S_{i}^{2}} e^{(x_{n}^{2} - x_{i+1}^{2})/2} \right)$$

$$\leq C \left( \frac{x_{n}}{S_{n}} + \sum_{i=1}^{n-1} \frac{x_{i}^{2}}{S_{i}^{2}} \left( \frac{i+1}{n} \right)^{2/9} \right)$$

$$\leq C \left( \frac{1}{n} + \frac{1}{n^{2/9}} \sum_{i=1}^{n-1} \frac{(i+1)^{2/9}}{i^{2}} \right)$$

$$\leq \frac{C}{n^{2/9}},$$

where we have applied (A.16) and the fact that  $w \ge x_{\lfloor m/e^3 \rfloor} \ge \sqrt{8}/3$  in the first inequality, (1.1) and the fact that  $w^2 \le x_n^2$  in the second inequality, (3.18) in the third inequality, and the fact that  $S_n \ge nx_n$  in the fourth inequality. The proof of (4.10) follows from (A.12), (A.14), and (A.17).  $\Box$ 

Acknowledgements. A substantial part of this paper was written when the second author was at the Vietnam Institute for Advanced Study in Mathematics (VIASM) and the Department of Mathematics, National University of Singapore (NUS). He would like to thank VIASM and the Department of Mathematics at NUS for their hospitality. We would also like to thank Adrian Röllin for helping us draw Figure 1 and the referees for their valuable comments.

**Funding.** This research was partially supported by Grant R-146-000-230-114 from the National University of Singapore.

## REFERENCES

- BOHM, D. (1952). A suggested interpretation of the quantum theory in terms of "hidden" variables. I. *Phys. Rev.* (2) 85 166–179. MR0046287
- BOHM, D. (1952). A suggested interpretation of the quantum theory in terms of "hidden" variables. II. *Phys. Rev.* (2) 85 180–193. MR0046288
- [3] BONIS, T. (2020). Stein's method for normal approximation in Wasserstein distances with application to the multivariate central limit theorem. *Probab. Theory Related Fields* 178 827–860. MR4168389 https://doi.org/10.1007/s00440-020-00989-4
- [4] 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. MR2732624 https://doi.org/10. 1007/978-3-642-15007-4
- [5] CHEN, L. H. Y. and SHAO, Q.-M. (2001). A non-uniform Berry–Esseen bound via Stein's method. Probab. Theory Related Fields 120 236–254. MR1841329 https://doi.org/10.1007/PL00008782
- [6] COURTADE, T. A., FATHI, M. and PANANJADY, A. (2019). Existence of Stein kernels under a spectral gap, and discrepancy bounds. Ann. Inst. Henri Poincaré Probab. Stat. 55 777–790. MR3949953 https://doi.org/10.1214/18-aihp898
- [7] FANG, X. (2019). Wasserstein-2 bounds in normal approximation under local dependence. *Electron*. J. Probab. 24 Paper No. 35, 14 pp. MR3940765 https://doi.org/10.1214/19-EJP301
- [8] GOLDSTEIN, L. (2007). L<sup>1</sup> bounds in normal approximation. Ann. Probab. 35 1888–1930. MR2349578 https://doi.org/10.1214/009117906000001123
- [9] GOLDSTEIN, L. and REINERT, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* 7 935–952. MR1484792 https://doi.org/10.1214/aoap/ 1043862419
- [10] HALL, M. J. W., DECKERT, D. A. and WISEMAN, H. M. (2014). Quantum phenomena modeled by interactions between many classical worlds. *Phys. Rev. X* 4 041013.
- [11] LEDOUX, M., NOURDIN, I. and PECCATI, G. (2015). Stein's method, logarithmic Sobolev and transport inequalities. Geom. Funct. Anal. 25 256–306. MR3320893 https://doi.org/10.1007/s00039-015-0312-0
- [12] MCKEAGUE, I. W. and LEVIN, B. (2016). Convergence of empirical distributions in an interpretation of quantum mechanics. Ann. Appl. Probab. 26 2540–2555. MR3543904 https://doi.org/10.1214/ 15-AAP1154
- [13] MCKEAGUE, I. W., PEKÖZ, E. A. and SWAN, Y. (2019). Stein's method and approximating the quantum harmonic oscillator. *Bernoulli* 25 89–111. MR3892313 https://doi.org/10.3150/17-bej960