

# A GENERAL METHOD FOR LOWER BOUNDS ON FLUCTUATIONS OF RANDOM VARIABLES<sup>1</sup>

BY SOURAV CHATTERJEE

*Stanford University*

There are many ways of establishing upper bounds on fluctuations of random variables, but there is no systematic approach for lower bounds. As a result, lower bounds are unknown in many important problems. This paper introduces a general method for lower bounds on fluctuations. The method is used to obtain new results for the stochastic traveling salesman problem, the stochastic minimal matching problem, the random assignment problem, the Sherrington–Kirkpatrick model of spin glasses, first-passage percolation and random matrices. A long list of open problems is provided at the end.

## 1. Theory.

1.1. *The problem of lower bounds.* The problem of establishing upper bounds on fluctuations of random variables is widely studied in the literature on concentration inequalities [14, 34]. The theory for lower bounds, however, is not so well developed. In fact, the only available methods for computing lower bounds are the following:

(1) Prove a distributional limit theorem. This is possible only in problems where classical tools are applicable, and in a small number of modern problems that admit exact calculations. For most of the contemporary ‘hard’ problems that do not have a miraculous exactly solvable structure, we are very far from proving distributional limit theorems.

(2) Prove a lower bound on some central moment (such as the variance) and a matching upper bound on a higher central moment. The Paley–Zygmund second moment method would then give a lower bound on the order of fluctuations. There is a general method for obtaining lower bounds on variances, due to Wehr and Aizenman [66]. However, matching upper bounds are rarely available. First-passage percolation is one example where the best known upper and lower bounds on fluctuations do not match [46, 52].

---

Received August 2017; revised July 2018.

<sup>1</sup>Supported in part by NSF Grant DMS-1608249.

*MSC2010 subject classifications.* 60E15, 60C05, 60K35, 60B20.

*Key words and phrases.* Variance lower bound, first-passage percolation, random assignment problem, stochastic minimal matching problem, stochastic traveling salesman problem, spin glass, Sherrington–Kirkpatrick model, random matrix, determinant.

(3) A coupling technique invented by Janson [31] and used by Bollobás and Janson [12] to obtain the first nontrivial lower bound on the fluctuations of the longest increasing subsequence in a random permutation. The main lemma of [31] is closely related to the method of this paper.

(4) Problem-specific techniques, as in [24, 29, 30, 35, 55]. Some of these are also related to the method proposed here.

Besides the above, there is a recent work of Janson and Warnke [32] on a general lower bound for lower tails of sums of weakly dependent binary random variables. This bound, however, is for the large deviation regime; it is not meant to be used for understanding typical fluctuations. In the examples where the Janson–Warnke lower bound applies, typical fluctuations can be understood more comprehensively by proving central limit theorems using existing technology for sums of weakly dependent random variables.

What is implicit in the above discussion is that while an upper bound on the variance gives an upper bound on the order of fluctuations by Chebychev’s inequality, a lower bound on the variance cannot be used on its own for demonstrating a lower bound on the order of fluctuations. For example, one can easily construct a sequence of random variables which converge in probability to a deterministic constant, but whose variances stay bounded away from zero. In such an example, it is unreasonable to say that the fluctuations do not tend to zero. In the absence of a simple numerical measure for lower bounds on fluctuations, the following definition looks reasonable.

**DEFINITION 1.1.** Let  $\{X_n\}_{n \geq 1}$  be a sequence of random variables and let  $\{\delta_n\}_{n \geq 1}$  be a sequence of positive real numbers. We will say that  $X_n$  has fluctuations of order at least  $\delta_n$  if there are positive constants  $c_1$  and  $c_2$  such that for all large  $n$ , and for all  $-\infty < a \leq b < \infty$  with  $b - a \leq c_1 \delta_n$ ,  $\mathbb{P}(a \leq X_n \leq b) \leq 1 - c_2$ .

In other words,  $X_n$  has fluctuations of order at least  $\delta_n$  if it is impossible to construct a sequence of intervals  $I_n$ , such that  $I_n$  has length of order  $\delta_n$  and  $\mathbb{P}(X_n \in I_n) \rightarrow 1$  as  $n \rightarrow \infty$ . It is easy to see that if  $\delta_n^{-1} X_n$  tends to a nondegenerate limit in distribution, then  $X_n$  has fluctuations of order at least  $\delta_n$  according to the above definition. It is also easy to see that if  $X_n$  has fluctuations of order  $\delta_n$  according to the above definition, then  $\text{Var}(X_n)$  is at least of order  $\delta_n^2$  (but not vice versa).

Definition 1.1 is closely related to the notion of concentration functions introduced by Lévy [37]. The concentration function  $f$  of a random variable  $X$  is defined as

$$f(l) := \sup_{x \in \mathbb{R}} \mathbb{P}(x \leq X \leq x + l).$$

In the language of concentration functions, Definition 1.1 can be restated as follows: A sequence of random variables  $\{X_n\}_{n \geq 1}$ , with concentration functions

$\{f_n\}_{n \geq 1}$ , has fluctuations of order at least  $\delta_n$  as  $n \rightarrow \infty$  if for some  $c > 0$ ,

$$\limsup_{n \rightarrow \infty} f_n(c\delta_n) < 1.$$

The theory of Lévy concentration functions is well developed for sums of independent random variables [53] and more generally in situations where a central limit theorem can be proved [19], but not for more complicated objects, especially where distributional limits cannot be established by existing methods.

*1.2. Lower bounds via coupling.* The following simple lemma gives a coupling technique for uniform upper bounds on probabilities of intervals of a given length. The idea is that given a random variable  $X$  and a number  $\delta$ , we construct another random variable  $Y$  on the same probability space, such that the law of  $X$  and the law of  $Y$  are close in total variation distance, and yet there is a substantial chance of  $|X - Y| > \delta$ . Under these circumstances, the following lemma shows that the probability of  $X$  belonging to an interval of length  $\leq \delta$  is bounded away from one. This is the main tool of this paper.

LEMMA 1.2. *Let  $X$  and  $Y$  be two random variables defined on the same probability space. Then for any  $-\infty < a \leq b < \infty$ ,*

$$\mathbb{P}(a \leq X \leq b) \leq \frac{1}{2}(1 + \mathbb{P}(|X - Y| \leq b - a) + d_{\text{TV}}(\mathcal{L}_X, \mathcal{L}_Y)),$$

where  $\mathcal{L}_X$  is the law of  $X$ ,  $\mathcal{L}_Y$  is the law of  $Y$ , and  $d_{\text{TV}}$  is total variation distance.

PROOF. Let  $I$  denote the interval  $[a, b]$ . Then note that

$$\begin{aligned} 1 &\geq \mathbb{P}(\{X \in I\} \cup \{Y \in I\}) \\ &= \mathbb{P}(X \in I) + \mathbb{P}(Y \in I) - \mathbb{P}(\{X \in I\} \cap \{Y \in I\}). \end{aligned}$$

But

$$\mathbb{P}(Y \in I) \geq \mathbb{P}(X \in I) - d_{\text{TV}}(\mathcal{L}_X, \mathcal{L}_Y),$$

and

$$\mathbb{P}(\{X \in I\} \cap \{Y \in I\}) \leq \mathbb{P}(|X - Y| \leq b - a).$$

The proof is completed by combining the above inequalities.  $\square$

Some variants of this coupling approach for lower bounds on fluctuations are already present in [24, 29, 30, 35, 55] for the specific problems handled in those papers; but the potential generality of the idea was not recognized in earlier works. Janson [31] proved a similar lemma, but where  $Y$  was assumed to have the same law as  $X$ . As far as I know, the exact statement of Lemma 1.2 is actually a new result. We will see later how this nearly trivial lemma can be used to obtain optimal

lower bounds on the orders of fluctuations of some highly complicated random variables.

Incidentally, one can have a version of Lemma 1.2 with twice the Kolmogorov distance instead of the total variation distance on the right. The proof makes it clear that this stronger statement is valid. However, bounding the total variation distance is often more manageable in problems of interest (as we will see in all of the examples in this paper), which is why the lemma is presented in the above form.

1.3. *A simple example.* For an elementary application of Lemma 1.2 that uses nothing more than Chebychev’s inequality, consider the following example. Let  $X_1, \dots, X_n$  be i.i.d. Bernoulli(1/2) random variables, and let

$$S_n = X_1 + \dots + X_n.$$

By the central limit theorem, we know that  $S_n$  has fluctuations of order  $n^{1/2}$ . Can we prove a lower bound of order  $n^{1/2}$  using Lemma 1.2? To do this, we first define a suitable perturbation  $(X'_1, \dots, X'_n)$  of the vector  $(X_1, \dots, X_n)$ , coupled with  $(X_1, \dots, X_n)$  on the same probability space. Let  $\alpha \in (0, 1)$  be a constant, to be chosen later, and let  $\varepsilon := \alpha n^{-1/2}$ . For each  $i$ , let

$$X'_i := \begin{cases} X_i & \text{with probability } 1 - \varepsilon, \\ 1 & \text{with probability } \varepsilon. \end{cases}$$

Then note that  $X'_1, \dots, X'_n$  are i.i.d. Bernoulli $((1 + \varepsilon)/2)$  random variables. Recall that if  $\mu$  and  $\nu$  are probability measures on a set  $\Omega$ , and  $\nu$  has density  $f$  with respect to  $\mu$ , then

$$d_{TV}(\mu, \nu) = \int_{\Omega} (1 - f)_+ d\mu,$$

where  $x_+$  denotes the positive part of a real number  $x$ . Using this representation, a simple calculation shows that

$$\begin{aligned} d_{TV}(\mathcal{L}_{(X_1, \dots, X_n)}, \mathcal{L}_{(X'_1, \dots, X'_n)}) &= \mathbb{E}(1 - (1 + \varepsilon)^{S_n} (1 - \varepsilon)^{n - S_n})_+ \\ &= \mathbb{E}\left(1 - (1 - \varepsilon^2)^{n/2} \left(\frac{1 + \varepsilon}{1 - \varepsilon}\right)^{S_n - n/2}\right)_+. \end{aligned}$$

The quantity within the expectation is always between 0 and 1. Moreover, by Chebychev’s inequality, for any  $\beta > 0$  we have

$$\mathbb{P}(|S_n - n/2| \geq \beta n^{1/2}) \leq \frac{1}{4\beta^2}.$$

Choosing  $\beta = \alpha^{-1/2}$ , we get

$$\begin{aligned} d_{\text{TV}}(\mathcal{L}_{(X_1, \dots, X_n)}, \mathcal{L}_{(X'_1, \dots, X'_n)}) &\leq \frac{\alpha}{4} + 1 - \left(1 - \frac{\alpha^2}{n}\right)^{n/2} \left(\frac{1 + \alpha n^{-1/2}}{1 - \alpha n^{-1/2}}\right)^{-\alpha^{-1/2} n^{1/2}} \\ &\leq C\sqrt{\alpha}, \end{aligned}$$

where  $C$  does not depend on  $n$  or the choice of  $\alpha$ . Thus, we may choose  $\alpha$  so small that the total variation distance is  $\leq 1/2$  for all  $n$ . Next, let

$$S'_n := X'_1 + \dots + X'_n.$$

Notice that for each  $i$ ,  $X'_i = X_i$  with probability  $1 - \varepsilon/2$  and  $X'_i = X_i + 1$  with probability  $\varepsilon/2$ . From this observation, it is not difficult to prove using Chebyshev's inequality that

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(S'_n \geq S_n + \frac{\alpha n^{1/2}}{3}\right) = 1.$$

Choose  $n$  so large that the above probability is at least  $2/3$ . Since

$$d_{\text{TV}}(\mathcal{L}_{S_n}, \mathcal{L}_{S'_n}) \leq d_{\text{TV}}(\mathcal{L}_{(X_1, \dots, X_n)}, \mathcal{L}_{(X'_1, \dots, X'_n)}) \leq \frac{1}{2},$$

Lemma 1.2 now implies that for any interval  $I$  of length less than  $\alpha n^{1/2}/3$ ,

$$\mathbb{P}(S_n \in I) \leq \frac{1}{2} \left(1 + \frac{1}{2} + \frac{1}{3}\right) = \frac{11}{12}.$$

This shows that in the sense of Definition 1.1,  $S_n$  has fluctuations of order at least  $n^{1/2}$ .

1.4. *Total variation distance between product measures.* The above example may appear to be a little ad hoc, because we used an explicit formula for the total variation distance to do our calculations. This, however, is not the case. There is a systematic way of upper bounding the total variation distances between product measures using a measure of similarity between probability measures called the Hellinger affinity. Although this method is well known and discussed at length in various texts (e.g., in [33] and [36]), I will now give a quick summary for the sake of completeness and to save the reader the trouble of looking up references.

Let  $(\Omega, \mathcal{F})$  be a measurable space, and let  $\nu$  be a probability measure on this space. Let  $f$  and  $g$  be two probability densities with respect to  $\nu$ , and let  $\mu$  and  $\mu'$  denote the corresponding measures. The Hellinger affinity or Hellinger integral between  $\mu$  and  $\mu'$  is defined as

$$\rho(\mu, \mu') := \int_{\Omega} \sqrt{fg} \, d\nu.$$

It is easy to see that this does not depend on the choice of  $\nu$ ; the result is the same for any  $\nu$  such that  $\mu$  and  $\mu'$  are both absolutely continuous with respect to  $\nu$ . For any  $\mu$  and  $\mu'$ , there is always at least one such  $\nu$ , for example,  $(\mu + \mu')/2$ . The Hellinger affinity gives the following upper bound on the total variation distance. This is a classical result. The proof is reproduced for completeness.

LEMMA 1.3. *For any two probability measures  $\mu$  and  $\mu'$  defined on the same measurable space,*

$$d_{TV}(\mu, \mu') \leq \sqrt{1 - \rho(\mu, \mu')^2}.$$

PROOF. By the Cauchy–Schwarz inequality,

$$\begin{aligned} d_{TV}(\mu, \mu') &= \frac{1}{2} \int |f - g| \, d\nu \\ &= \frac{1}{2} \int |(\sqrt{f} - \sqrt{g})(\sqrt{f} + \sqrt{g})| \, d\nu \\ &\leq \frac{1}{2} \left( \int (\sqrt{f} - \sqrt{g})^2 \, d\nu \int (\sqrt{f} + \sqrt{g})^2 \, d\nu \right)^{1/2} \\ &= (1 - \rho(\mu, \mu'))^{1/2} (1 + \rho(\mu, \mu'))^{1/2}, \end{aligned}$$

where the last step follows because  $\int f \, d\nu = \int g \, d\nu = 1$ .  $\square$

The main advantage of using the Hellinger affinity is that it is easy to evaluate for product measures. For  $i = 1, \dots, n$ , let  $\mu_i$  and  $\mu'_i$  be probability measures on some measurable space  $(\Omega_i, \mathcal{F}_i)$ . Let  $\mu = \mu_1 \times \dots \times \mu_n$  and  $\mu' = \mu'_1 \times \dots \times \mu'_n$  be the corresponding product measures on  $\Omega_1 \times \dots \times \Omega_n$ . Then from the definition of the Hellinger affinity it is clear that

$$\rho(\mu, \mu') = \prod_{i=1}^n \rho(\mu_i, \mu'_i).$$

Combining this with Lemma 1.3, we get the following upper bound for the total variation distance between product measures.

LEMMA 1.4. *Let  $\mu$  and  $\mu'$  be as in the above paragraph. Then*

$$d_{TV}(\mu, \mu') \leq \left( 1 - \prod_{i=1}^n \rho(\mu_i, \mu'_i)^2 \right)^{1/2}.$$

Let us see what this lemma gives for the example from Section 1.3. In that example, a simple calculation shows that

$$\rho(\mathcal{L}_{X_i}, \mathcal{L}_{X'_i}) = \frac{\sqrt{1 + \varepsilon} + \sqrt{1 - \varepsilon}}{2} \geq 1 - C\varepsilon^2 = 1 - \frac{C\alpha^2}{n},$$

where  $C$  is a positive constant that does not depend on  $n$  or  $\alpha$ . Thus,

$$d_{TV}(\mathcal{L}_{(X_1, \dots, X_n)}, \mathcal{L}_{(X'_1, \dots, X'_n)}) \leq \left(1 - \left(1 - \frac{C\alpha^2}{n}\right)^{2n}\right)^{1/2} \leq C'\alpha,$$

where  $C'$  is another constant that does not depend on  $n$  or  $\alpha$ . This recovers a stronger version of the result that we previously derived using the explicit formula for the total variation distance and Chebychev's inequality.

1.5. *Perturbative coupling.* To apply Lemma 1.2, we need to get upper bounds on the total variation distance between the law of a random variable and the law of a small perturbation of the variable. We have dealt with perturbations of independent Bernoulli random variables in the previous section. This section gives a similar bound for continuous random vectors, where the perturbation is of a different nature. The bound is derived using Lemma 1.4. We will deal with the following classes of probability measures on  $\mathbb{R}^d$ ,  $d \geq 1$ .

DEFINITION 1.5. Let  $\mathcal{P}(d)$  denote the set of all probability densities of the form  $e^{-V}$  on  $\mathbb{R}^d$ , where  $V$  is a  $C^\infty$  function,  $V$  and all its derivatives have at most polynomial growth at infinity, and  $e^V$  increases faster than any polynomial at infinity.

DEFINITION 1.6. Let  $\mathcal{P}^+(d)$  denote the set of all probability densities of the form  $e^{-V}$  on  $[0, \infty)^d$ , such that  $V$  is a  $C^\infty$  function in  $(0, \infty)^d$ ,  $V$  and all its derivatives have at most polynomial growth at infinity and extend continuously to the boundary of  $[0, \infty)^d$ , and  $e^V$  increases faster than any polynomial at infinity.

Many of the familiar probability distributions, such as Gaussian distributions and exponential distributions, are of the above type. Some other commonly used distributions, such as uniform distributions on compact sets, do not belong to the above classes.

The following theorem gives a lower bound on the Hellinger affinity between a probability measure in  $\mathcal{P}(d)$  or  $\mathcal{P}^+(d)$  and a certain kind of perturbation of that measure. The proof is based on an application of the divergence theorem. This result is frequently used in later sections.

THEOREM 1.7. *Let  $X$  be a  $d$ -dimensional random vector whose law belongs to either  $\mathcal{P}(d)$  or  $\mathcal{P}^+(d)$ . Take any  $\varepsilon \in (-1/2, 1/2)$ , and let  $X' = X/(1 + \varepsilon)$ . Then*

$$\rho(\mathcal{L}_X, \mathcal{L}_{X'}) \geq 1 - C\varepsilon^2$$

where  $C$  depends only on  $\mathcal{L}_X$  and  $d$ .

PROOF. Let  $e^{-V}$  be the probability density function of  $X$ . In the following, all integrals are over  $\mathbb{R}^d$  if  $e^{-V} \in \mathcal{P}(d)$ , and over  $[0, \infty)^d$  if  $e^{-V} \in \mathcal{P}^+(d)$ . Let

$$g(\varepsilon) := \int \sqrt{(1 + \varepsilon)^d e^{-V(x+\varepsilon x)} e^{-V(x)}} dx = \rho(\mathcal{L}_X, \mathcal{L}_{X'})$$

By the inequality

$$|e^x - e^y| \leq \frac{1}{2}|x - y|(e^x + e^y),$$

we have that for any  $\varepsilon_1$  and  $\varepsilon_2$ ,

$$\begin{aligned} &|e^{-V(x+\varepsilon_1 x)} - e^{-V(x+\varepsilon_2 x)}| \\ &\leq \frac{1}{2}|V(x + \varepsilon_1 x) - V(x + \varepsilon_2 x)|(e^{-V(x+\varepsilon_1 x)} + e^{-V(x+\varepsilon_2 x)}). \end{aligned}$$

Using this inequality and the assumptions on  $V$ , it is not difficult to verify that  $g$  is a  $C^\infty$  function on  $(-1/2, 1/2)$ , and the derivatives can be computed by differentiating under the integral. Note that  $g(0) = 1$  and

$$\begin{aligned} g'(\varepsilon) &= \frac{d}{2}(1 + \varepsilon)^{-1}g(\varepsilon) \\ &\quad - \frac{1}{2}(1 + \varepsilon)^{d/2} \int x \cdot \nabla V(x + \varepsilon x)e^{-(V(x+\varepsilon x)+V(x))/2} dx, \end{aligned}$$

which gives

$$\begin{aligned} g'(0) &= \frac{d}{2} - \frac{1}{2} \int x \cdot \nabla V(x)e^{-V(x)} dx \\ &= \frac{1}{2} \int (d - x \cdot \nabla V(x))e^{-V(x)} dx. \end{aligned}$$

Let  $h = (h_1, \dots, h_d)$  be the function defined as

$$h_i(x) = x_i e^{-V(x)}.$$

Then

$$\frac{\partial h_i}{\partial x_i} = \left(1 - x_i \frac{\partial V}{\partial x_i}\right) e^{-V(x)},$$

and hence

$$\operatorname{div} h(x) = \sum_{i=1}^d \frac{\partial h_i}{\partial x_i} = (d - x \cdot \nabla V(x))e^{-V(x)}.$$

If  $e^{-V} \in \mathcal{P}(d)$ , then using the growth assumptions about  $V$ , a simple application of the divergence theorem now shows that  $g'(0) = 0$ . If  $e^{-V} \in \mathcal{P}^+(d)$ , then also the divergence theorem implies that  $g'(0) = 0$ , because  $h(x) \cdot n(x) = 0$  on the

boundary of  $[0, \infty)^d$ , where  $n(x)$  denotes the unit normal vector to the boundary at the point  $x$ . Finally, note that

$$\begin{aligned} g''(\varepsilon) &= -\frac{d}{2}(1 + \varepsilon)^{-2}g(\varepsilon) + \frac{d}{2}(1 + \varepsilon)^{-1}g'(\varepsilon) \\ &\quad - \frac{d}{4}(1 + \varepsilon)^{(d-2)/2} \int x \cdot \nabla V(x + \varepsilon x)e^{-(V(x+\varepsilon x)+V(x))/2} dx \\ &\quad - \frac{1}{2}(1 + \varepsilon)^{d/2} \int \left(x \cdot \text{Hess } V(x + \varepsilon x)x \right. \\ &\quad \left. - \frac{1}{2}(x \cdot \nabla V(x + \varepsilon x))^2\right)e^{-(V(x+\varepsilon x)+V(x))/2} dx, \end{aligned}$$

where  $\text{Hess } V$  is the Hessian matrix of  $V$ . By the assumed properties of  $V$ , the above formulas show that  $|g|$ ,  $|g'|$  and  $|g''|$  are uniformly bounded in the interval  $(-1/2, 1/2)$ . Moreover, we have already deduced that  $g(0) = 1$  and  $g'(0) = 0$ . Thus, for any  $\varepsilon \in (-1/2, 1/2)$ ,

$$g(\varepsilon) \geq 1 - C\varepsilon^2,$$

where  $C$  depends only on  $V$  and  $d$ .  $\square$

Combining this theorem with Lemma 1.4 yields the following corollary.

**COROLLARY 1.8.** *Let  $X_1, \dots, X_n$  be i.i.d.  $d$ -dimensional random vectors with probability density belonging to either  $\mathcal{P}(d)$  or  $\mathcal{P}^+(d)$ . Take any  $\varepsilon_1, \dots, \varepsilon_n \in (-1/2, 1/2)$  and let  $X'_i = X_i/(1 + \varepsilon_i)$  for  $i = 1, \dots, n$ . Then*

$$d_{\text{TV}}(\mathcal{L}_{(X_1, \dots, X_n)}, \mathcal{L}_{(X'_1, \dots, X'_n)}) \leq C \left( \sum_{i=1}^n \varepsilon_i^2 \right)^{1/2},$$

where  $C$  depends only of the law of the  $X_i$ 's and the dimension  $d$ .

**PROOF.** By Theorem 1.7 and Lemma 1.4,

$$d_{\text{TV}}(\mathcal{L}_{(X_1, \dots, X_n)}, \mathcal{L}_{(X'_1, \dots, X'_n)}) \leq \left( 1 - \prod_{i=1}^n (1 - C_0 \varepsilon_i^2) \right)^{1/2},$$

where  $C_0$  depends only on the law of the  $X_i$ 's and the dimension  $d$ . Since the total variation distance is bounded by one, we may assume without loss of generality (by suitably increasing the value of  $C$  in the statement of the corollary, if necessary) that

$$\sum_{i=1}^n \varepsilon_i^2 \leq \frac{1}{4C_0}.$$

Then by repeated applications of the inequality  $(1 - x)(1 - y) \geq 1 - x - y$ , which holds for  $x, y \in [0, 1]$ , we get the desired result.  $\square$

1.6. *Connection with the Mermin–Wagner theorem.* The technique of using Lemma 1.2 in conjunction with Lemma 1.4 and Theorem 1.7 has strong similarities with the celebrated Mermin–Wagner theorem of statistical physics [40, 41]. Roughly speaking, the Mermin–Wagner theorem shows that continuous symmetries cannot be spontaneously broken in dimensions  $\leq 2$ . The physics proof of the Mermin–Wagner theorem involves the introduction of a slowly rotating perturbation known as a “spin wave,” akin to the perturbation used in Theorem 1.7. The first rigorous proof of the Mermin–Wagner theorem, due to McBryan and Spencer [39], used complex analytic techniques. A later proof, due to Pfister [54], used a more transparent argument that resembles the method of this paper in various aspects. For a modern exposition in probabilistic language, see the lecture notes of Peled and Spinka [51].

## 2. Applications.

2.1. *Traveling salesman and minimal matching.* Let  $f$  be a measurable real-valued function on  $(\mathbb{R}^d)^n$  such that there is some  $r > 0$  so that for any  $\lambda \geq 0$  and any  $x_1, \dots, x_n \in \mathbb{R}^d$ ,

$$(2.1) \quad f(\lambda x_1, \dots, \lambda x_n) = \lambda^r f(x_1, \dots, x_n).$$

Such functions arise in many geometric combinatorial optimization problems. The length of the optimal traveling salesman path through  $x_1, \dots, x_n$ , the length of the minimal spanning tree and the length of the minimal matching (if  $n$  is even) are all examples of functions having the above property, with  $r = 1$ . The volume of the convex hull of  $x_1, \dots, x_n$  is an example that satisfies (2.1) with  $r = d$ .

If the points  $x_1, \dots, x_n$  are replaced by i.i.d. random points  $X_1, \dots, X_n$  drawn from some probability measure on  $\mathbb{R}^d$ , the resulting problem is called the stochastic version of the original optimization problem. The stochastic problems have been extensively studied by probabilists. Laws of large numbers and concentration inequalities for upper bounds on fluctuations are well understood due to the works of many authors, and beautifully expounded in the classic monograph of Steele [57]. Yet a lot remains to be understood. For example, the distribution theories for the stochastic traveling salesman and the stochastic minimal matching problems remain out of the reach of available technology.

Not much is known about lower bounds on fluctuations in the problems where the distribution theory is not understood. For the stochastic traveling salesman problem, the only result on lower bounds that I am aware of is a result of Rhee [55], who proved a fluctuation lower bound of the correct order in dimension two when the points are uniformly distributed in the unit square. Nothing is known about lower bounds in the stochastic minimal matching problem. The following theorem gives a general lower bound for fluctuations when the points are distributed according to some probability density in  $\mathcal{P}(d)$  or  $\mathcal{P}^+(d)$ .

**THEOREM 2.1.** *Take any  $d \geq 1$ . Let  $X_1, X_2, \dots$  be i.i.d. random vectors with probability density in either  $\mathcal{P}(d)$  or  $\mathcal{P}^+(d)$ . For each  $n$ , let  $f_n : (\mathbb{R}^d)^n \rightarrow \mathbb{R}$  be a function satisfying (2.1) for some fixed  $r > 0$ , and let  $L_n := f_n(X_1, \dots, X_n)$ . Let  $(t_n)_{n \geq 1}$  be a sequence of positive real numbers such that*

$$\liminf_{n \rightarrow \infty} \mathbb{P}(L_n \geq t_n) > 0.$$

*Then  $L_n$  has fluctuations of order at least  $n^{-1/2}t_n$ , in the sense of Definition 1.1.*

**PROOF.** Without loss of generality, let  $a$  be a positive constant such that

$$\mathbb{P}(L_n \geq t_n) \geq a$$

for all  $n$ . Take any  $n$ . For  $i = 1, \dots, n$ , let

$$X'_i := \frac{X_i}{1 + \alpha n^{-1/2}},$$

where  $\alpha$  will be determined later. Let  $L'_n := f_n(X'_1, \dots, X'_n)$ . Then by Corollary 1.8,

$$d_{\text{TV}}(\mathcal{L}_{L_n}, \mathcal{L}_{L'_n}) \leq d_{\text{TV}}(\mathcal{L}_{(X_1, \dots, X_n)}, \mathcal{L}_{(X'_1, \dots, X'_n)}) \leq C\alpha,$$

where  $C$  depends only on  $\mathcal{L}_{X_1}$  and  $d$ . On the other hand, by (2.1),

$$L'_n = \frac{L_n}{(1 + \alpha n^{-1/2})^r}.$$

Consequently, for any  $\beta$  such that

$$0 < \beta < \frac{(1 + \alpha n^{-1/2})^r - 1}{n^{-1/2}(1 + \alpha n^{-1/2})^r} = r\alpha + o(1)$$

[where  $o(1)$  denotes a quantity that tends to zero as  $n \rightarrow \infty$ ], we have

$$\begin{aligned} \mathbb{P}(L'_n \geq L_n - \beta n^{-1/2}t_n) &= \mathbb{P}\left(L_n \leq \frac{(1 + \alpha n^{-1/2})^r \beta n^{-1/2}t_n}{(1 + \alpha n^{-1/2})^r - 1}\right) \\ &\leq \mathbb{P}(L_n < t_n) \leq 1 - a. \end{aligned}$$

Thus, choosing  $\alpha$  sufficiently small, and then choosing  $\beta$  depending on  $\alpha$ , Lemma 1.2 completes the proof.  $\square$

Concretely, Theorem 2.1 gives the following lower bound in the traveling salesman and minimal matching problems.

**COROLLARY 2.2.** *Take any  $d \geq 2$ . In the stochastic traveling salesman and stochastic minimal matching problems with points drawn according to some probability density in  $\mathcal{P}(d)$  or  $\mathcal{P}^+(d)$ , the optimal lengths have fluctuations of order at least  $n^{(d-2)/2d}$ , in the sense of Definition 1.1.*

PROOF. In the stochastic traveling salesman problem and the stochastic minimal matching problems in  $d \geq 2$ , it is not difficult to see that the optimal values are bounded below by some constant multiple of the sum of nearest-neighbor distances. From this, it follows by a simple mean and variance calculation that  $t_n$  can be chosen to be  $n^{1-1/d}$ , since there are  $n$  points and the nearest neighbor distances are of order at least  $n^{-1/d}$  [because a density in  $\mathcal{P}(d)$  or  $\mathcal{P}^+(d)$  is necessarily bounded]. With this choice of  $t_n$ , Theorem 2.1 implies that the order of fluctuations is at least  $n^{-1/2}n^{1-1/d} = n^{(d-2)/2d}$ .  $\square$

Interestingly, the lower bound in the above corollary matches the known order of upper bounds on fluctuations in these problems [57]. The known upper bounds, however, are for points drawn from the uniform distribution on  $[0, 1]^d$ . I do not know if upper bounds are known for unbounded distributions, such as the ones in  $\mathcal{P}(d)$  and  $\mathcal{P}^+(d)$ .

The next theorem gives the optimal lower bound for the traveling salesman problem when the points are distributed uniformly in  $[0, 1]^d$ . This is a straightforward generalization of the lower bound obtained by Rhee [55] in  $d = 2$ . The coupling used in the proof of this theorem is borrowed from Rhee’s paper. I do not know if a simpler coupling can be made to work.

**THEOREM 2.3.** *Take any  $d \geq 2$ . Let  $L_n$  be the length of the optimal tour in the stochastic traveling salesman problem with  $n$  points when the points are distributed independently and uniformly in  $[0, 1]^d$ . Then  $L_n$  has fluctuations of order at least  $n^{(d-2)/2d}$ , in the sense of Definition 1.1.*

PROOF. Throughout this proof,  $C$  will denote any positive universal constant, whose value may change from line to line.

Let  $X_1, \dots, X_n$  be i.i.d. uniform points from  $[0, 1]^d$ . Assume that  $n \geq 4$  and let  $m = \lfloor n/2 \rfloor$ . Given  $X_1, \dots, X_m$ , let  $D$  be the set of all points in  $[0, 1]^d$  that are within distance  $\alpha n^{-1/d}$  from the set  $\{X_1, \dots, X_m\}$ , where  $\alpha \in (0, 1)$  will be chosen later. Generate  $Y_{m+1}, \dots, Y_n$  independently and uniformly from the set  $D$ . For each  $m + 1 \leq i \leq n$ , let

$$X'_i := \begin{cases} X_i & \text{with probability } 1 - \beta n^{-1/2}, \\ Y_i & \text{with probability } \beta n^{-1/2}, \end{cases}$$

where  $\beta \in (0, 1)$  will be chosen later. For  $1 \leq i \leq m$ , let  $X'_i := X_i$ . Let  $L_n$  be the length of the optimal tour through  $X_1, \dots, X_n$  and let  $L'_n$  be the length of the optimal tour through  $X'_1, \dots, X'_n$ .

Given  $X_1, \dots, X_m$ , the random variables  $X'_{m+1}, \dots, X'_n$  are i.i.d. with probability density function

$$1 - \beta n^{-1/2} + \frac{\beta n^{-1/2}}{\text{Vol}(D)} 1_{\{x \in D\}}$$

at  $x \in [0, 1]^d$ . From this formula and the inequality  $\sqrt{1+a} \geq 1 + a/2 - Ca^2$  that holds for  $a \geq -1/2$ , an easy calculation shows that the Hellinger affinity between this conditional law and the uniform distribution on  $[0, 1]^d$  is bounded below by

$$1 - \frac{C\beta^2}{\text{Vol}(D)_n}.$$

Therefore, by Lemma 1.4, the total variation distance between the conditional law of  $(X'_{m+1}, \dots, X'_n)$  and that of  $(X_{m+1}, \dots, X_n)$  is bounded above by

$$\frac{C\beta}{\sqrt{\text{Vol}(D)}}.$$

It is not difficult to show (e.g., as in the proof of [55], Lemma 7) that with probability tending to one as  $n \rightarrow \infty$ ,  $\text{Vol}(D) \geq C\alpha^d$ . Combining this with the above bound on the conditional laws, it follows that

$$(2.2) \quad d_{\text{TV}}(\mathcal{L}_{(X_1, \dots, X_n)}, \mathcal{L}_{(X'_1, \dots, X'_n)}) \leq \frac{C\beta}{\alpha^{d/2}}.$$

Consider the optimal tour through  $X_1, \dots, X_n$ . Each  $X_i$  has two ‘neighbors’ in this tour, one which comes before it and one that comes after. Call these points  $U_i$  and  $V_i$ . If the point  $X_i$  is erased, then the length of the optimal tour must decrease by at least

$$\|X_i - U_i\| + \|X_i - V_i\| - \|U_i - V_i\|.$$

Let  $K$  be a positive real number, to be chosen later. It is known that with probability tending to one, the optimal tour has length  $\leq Cn^{1-1/d}$  [57], Chapter 2. If this happens, then the average length of an edge in the tour is bounded above by  $Cn^{-1/d}$ , and hence the fraction of all  $i$  for which

$$(2.3) \quad \max\{\|X_i - U_i\|, \|X_i - V_i\|\} \leq Kn^{-1/d}$$

is bounded below by  $1 - CK^{-1}$ .

For each  $i$ , let  $N_i$  be the set of all  $j \neq i$  such that  $\|X_i - X_j\| \leq Kn^{-1/d}$ . Let

$$R_i := \min\{\|X_i - X_j\| + \|X_i - X_k\| - \|X_j - X_k\| : j, k \in N_i\},$$

where we follow the usual convention that the minimum of an empty set is infinity. Let  $\gamma$  be a positive real number, to be chosen later. From the local structure of a set of uniformly distributed points, it is not difficult to prove that with probability tending to one as  $n \rightarrow \infty$ , the fraction of all  $i$  for which  $R_i \geq \gamma n^{-1/d}$  is at least  $c(\gamma, K)$ , where  $c(\gamma, K) \rightarrow 1$  as  $\gamma \rightarrow 0$  for any fixed  $K$ .

Finally, note that if the point  $X_i$  is dropped, and (2.3) holds, then the length of the optimal tour decreases by at least  $R_i$ .

Combining all of the above observations, we see that if  $K$  is chosen large enough, and then  $\gamma$  is chosen small enough depending on  $K$ , then with probability

tending to one as  $n \rightarrow \infty$ , there are at least  $3n/4$  points  $X_i$  such that (2.3) holds and  $R_i \geq \gamma n^{-1/d}$ . Let  $A$  be the set of all  $i$  such that  $X_i$  has these two properties.

Let  $B$  be the set of all  $m + 1 \leq i \leq n$  such that  $X'_i = Y_i$ . Then from the conclusion of the previous paragraph, it follows that with probability tending to one,  $|A \cap B| \geq C\beta n^{1/2}$ . Also, the expected number of pairs of points in  $B$  that are neighbors of each other in the optimal tour is of order 1 as  $n \rightarrow \infty$ . To see this, just note that  $B$  is a randomly chosen subset of size  $O(\sqrt{n})$  and compute a straightforward bound on the conditional expectation of the number of such pairs given the  $X_i$ 's. From these two observations, it follows that with probability tending to one, deleting  $\{X_i : i \in B\}$  results in a decrease of at least  $C\beta\gamma n^{1/2-1/d}$  in the length of the optimal tour.

After dropping  $\{X_i : i \in B\}$ , let us now replace these points by  $\{Y_i : i \in B\}$ . Note that the resulting point set is exactly  $\{X'_1, \dots, X'_n\}$ . Each  $Y_i$  is within distance  $\alpha n^{-1/d}$  of some  $X_j$  and, therefore, adds at most  $2\alpha n^{-1/d}$  to the length of the optimal tour. Consequently, with probability tending to one, the total increase after inserting all the  $Y_i$ 's is at most  $C\alpha\beta n^{1/2-1/d}$ . Thus, choosing  $\alpha$  sufficiently small (depending only on  $\gamma$ ), we can ensure that with probability tending to one,

$$L_n - L'_n \geq C\beta n^{1/2-1/d}.$$

Finally, choose  $\beta$  small enough (depending on  $\alpha$ ) so that the right-hand side of (2.2) is less than  $1/2$ . Lemma 1.2 now completes the proof.  $\square$

2.2. *Free energy of the Sherrington–Kirkpatrick model.* Let  $n$  be any positive integer. Let  $(g_{ij})_{1 \leq i < j \leq n}$  be i.i.d. standard Gaussian random variables. The Sherrington–Kirkpatrick (S-K) model of spin glasses [56] with  $n$  spins at inverse temperature  $\beta \geq 0$  and external field  $h \in \mathbb{R}$  defines a random probability measure on  $\{-1, 1\}^n$ , which puts mass proportional to  $e^{\beta H_n(\sigma)}$  at each point  $\sigma = (\sigma_1, \dots, \sigma_n) \in \{-1, 1\}^n$ , where

$$H_n(\sigma) := \frac{1}{\sqrt{n}} \sum_{1 \leq i < j \leq n} g_{ij} \sigma_i \sigma_j + h \sum_{i=1}^n \sigma_i.$$

The S-K model has inspired a large body of work in probability theory, extensively surveyed in [49, 59–61]. One of the key quantities of interest is the free energy of the model, defined as

$$F_n(\beta, h) := \log \sum_{\sigma \in \{-1, 1\}^n} e^{\beta H_n(\sigma)}.$$

For  $\beta < 1$  and  $h = 0$ , the fluctuations of the free energy are well understood due to the work of Aizenman, Lebowitz and Ruelle [1]. In this case, the free energy has fluctuations of order 1 as  $n \rightarrow \infty$ , and satisfies a central limit theorem after centering. When  $\beta > 1$  and  $h = 0$ , the best known upper bound on the order of fluctuations is  $\sqrt{n/\log n}$  [16, 18]. I do not know of a definite conjecture about

the true order of fluctuations. For  $h \neq 0$ , Chen, Dey and Panchenko [20] have recently proved a central limit theorem for  $F_n(\beta, h)$  for any  $\beta$ , showing that it has fluctuations of order  $\sqrt{n}$ . When  $h = 0$ , I have heard it said that the fluctuations may be of order 1, but I have also heard it said that the fluctuations are of order  $n^\rho$  for some small  $\rho$ . The following result shows that in the absence of an external field, the fluctuations are at least of order 1.

**THEOREM 2.4.** *The free energy of the S-K model at zero external field and any inverse temperature  $\beta$ , has fluctuations of order at least 1, in the sense of Definition 1.1.*

**PROOF.** Since  $h = 0$ , we will write  $F_n(\beta)$  instead of  $F_n(\beta, h)$ . Let  $\alpha$  be a positive constant, to be determined later. Let

$$\tilde{g}_{ij} := \frac{g_{ij}}{1 - \alpha n^{-1}},$$

and let  $\tilde{F}_n(\beta)$  be the free energy of the model where  $g_{ij}$  is replaced by  $\tilde{g}_{ij}$ . Then by Corollary 1.8,

$$d_{\text{TV}}(\mathcal{L}_{F_n(\beta)}, \mathcal{L}_{\tilde{F}_n(\beta)}) \leq d_{\text{TV}}(\mathcal{L}_{(g_{ij})_{1 \leq i < j \leq n}}, \mathcal{L}_{(\tilde{g}_{ij})_{1 \leq i < j \leq n}}) \leq C\alpha,$$

where  $C$  does not depend on  $n$ . A simple computation shows that

$$\tilde{F}_n(\beta) - F_n(\beta) = \log \left\langle \exp \left( \frac{\beta \alpha H_n(\sigma)}{n(1 - \alpha n^{-1})} \right) \right\rangle_\beta,$$

where  $\langle \cdot \rangle_\beta$  denotes expectation under the probability measure defined by the S-K model at inverse temperature  $\beta$ . By Jensen's inequality, this gives

$$(2.4) \quad \tilde{F}_n(\beta) - F_n(\beta) \geq \left\langle \frac{\beta \alpha H_n(\sigma)}{n(1 - \alpha n^{-1})} \right\rangle_\beta.$$

On the other hand, another simple computation gives

$$(2.5) \quad F'_n(\beta) = \langle H_n(\sigma) \rangle_\beta,$$

where  $F'_n$  is the derivative of  $F_n$ . Now, the following facts are well known (see, e.g., in [59]):  $n^{-1} F_n(\beta)$  is a convex function of  $\beta$ , and converges to a deterministic limit  $P(\beta)$  as  $n \rightarrow \infty$ . Therefore, by the properties of convex functions,  $n^{-1} F'_n(\beta) \rightarrow P'(\beta)$  for every  $\beta > 0$  where  $P$  is differentiable. Moreover, it is also known that  $P'(\beta) > 0$  for every  $\beta > 0$  where  $P$  is differentiable. Thus, at every  $\beta > 0$  where  $P$  is differentiable,  $n^{-1} F'_n(\beta)$  converges to a positive limit.

Now,  $P$  is convex, and hence almost everywhere differentiable in  $(0, \infty)$ . In particular, for any  $\beta > 0$  there exists  $\beta' \in (0, \beta)$  where  $P$  is differentiable. Thus,  $n^{-1} F'_n(\beta')$  converges to a positive limit. But the convexity of  $F_n$  implies that

$F'_n(\beta') \leq F'_n(\beta)$ . Thus, there exists a positive constant  $a$  (depending on  $\beta$ ) such that

$$\lim_{n \rightarrow \infty} \mathbb{P}(n^{-1} F'_n(\beta) \geq a) = 1.$$

Thus, by (2.4) and (2.5), we see that there is some  $b > 0$  such that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\tilde{F}_n(\beta) - F_n(\beta) \geq b) = 1.$$

Choosing  $\alpha$  small enough, Lemma 1.2 completes the proof.  $\square$

Another important quantity related to the S-K model is its ground state energy, namely,

$$(2.6) \quad \max_{\sigma \in \{-1, 1\}^n} \frac{1}{\sqrt{n}} \sum_{1 \leq i < j \leq n} g_{ij} \sigma_i \sigma_j.$$

The best known upper bound on the order of fluctuations of the ground state energy is  $o(\sqrt{n})$ , proved in a recent manuscript of Chen, Handschy and Lerman [21]. No lower bound is known. It is believed that the correct order of fluctuations is  $n^\rho$ , where  $\rho$  is either  $1/6$  or  $1/4$  (see, e.g., the discussion in [48]). The following theorem proves a lower bound of order 1 on the fluctuations of the ground state energy.

**THEOREM 2.5.** *The ground state energy of the S-K model, as defined in (2.6), has fluctuations of order at least 1, in the sense of Definition 1.1.*

**PROOF.** Let  $\tilde{g}_{ij}$  be as in the proof of Theorem 2.4. Let  $G_n$  and  $\tilde{G}_n$  be the ground state energies in the two systems. Then

$$\tilde{G}_n = \frac{G_n}{1 - \alpha n^{-1}}.$$

Using the well-known fact that  $G_n/n$  converges in probability to a deterministic positive limit as  $n \rightarrow \infty$ , it is now easy to complete the proof using Corollary 1.8 and Lemma 1.2.  $\square$

**2.3. First-passage percolation.** Take any  $d \geq 2$ . Let  $E$  be the set of nearest-neighbor edges of  $\mathbb{Z}^d$ , and let  $(\omega_e)_{e \in E}$  be a collection of i.i.d. nonnegative random variables, called “edge weights.” Define the weight of a path in  $\mathbb{Z}^d$  to be the sum of the weights of the edges in the path. Define the first-passage time  $T(x, y)$  from a point  $x$  to a point  $y$  to be the minimum over the weights of all paths from  $x$  to  $y$ . First-passage percolation is the study of the behavior of these first-passage times. For a recent survey of the large mathematical literature on this model, see [9].

A lot of energy has been spent on the study of fluctuations of first-passage times, but the known results are far from optimal. If  $x$  and  $y$  are points at distance  $n$  from

each other, the best known upper bound on the fluctuations of  $T(x, y)$  is of order  $\sqrt{n/\log n}$  under mild assumptions on the law of the edge weights [10, 11, 23], although it is conjectured that at least in  $d = 2$ , the correct order should be  $n^{1/3}$ . The situation with lower bounds is even worse. The best known lower bound on the order of fluctuations in  $d = 2$ , due to Pemantle and Peres [52], is of order  $\sqrt{\log n}$  when the edge weight distribution is exponential. The proof depends crucially on the memoryless property of the exponential distribution. Newman and Piza [46] showed that for a fairly general class of edge weight distributions, the variance of  $T(x, y)$  is lower bounded by a constant multiple of  $\log n$ . The Newman–Piza lower bound is based on a technique pioneered by Wehr and Aizenman [66]. However, on its own, this lower bound on the variance does not give a lower bound on the order of fluctuations since there is no matching upper bound on any higher moment (or even on the variance itself). The following theorem proves the  $\sqrt{\log n}$  lower bound on the order of fluctuations for a large class of edge weight distributions.

**THEOREM 2.6.** *Consider the first-passage percolation model with i.i.d. non-negative edge weights in  $d = 2$ . Suppose that the edge weight distribution belongs to the class  $\mathcal{P}^+(1)$ . Let  $x_n$  and  $y_n$  be two sequences of points such that the distance between  $x_n$  and  $y_n$  grows like a constant multiple of  $n$ . Then the fluctuations of  $T(x_n, y_n)$  are at least of order  $\sqrt{\log n}$  in the sense of Definition 1.1.*

**PROOF.** Throughout this proof,  $C$  will denote any positive constant that may depend only on the edge weight distribution, and nothing else. The value of  $C$  may change from line to line or even within a line.

Let  $n$  be a positive integer greater than four, and let  $y$  be a point at distance  $n$  from 0. Let  $T = T(0, y)$ . It suffices to prove that  $T$  has fluctuations of order at least  $\sqrt{\log n}$ .

For each edge  $e$ , let  $k(e)$  denote the distance of  $e$  from the origin, where ‘distance’ means the graph distance between the origin and the endpoint of  $e$  that is closer to the origin. For all edges  $e$  with  $k(e) \leq n/2$ , let

$$\varepsilon_e := \frac{\alpha}{(k(e) + 1)\sqrt{\log n}},$$

where the constant  $\alpha$  will be chosen later. Since there are  $\leq Cr$  edges at distance  $r$ ,

$$(2.7) \quad \sum_e \varepsilon_e^2 \leq \sum_{r \leq n/2} Cr \frac{\alpha^2}{(r + 1)^2 \log n} \leq C\alpha^2.$$

Let  $T'$  be the first-passage time from 0 to  $y$  when the original edge weights  $\omega_e$  are replaced by  $\omega_e/(1 + \varepsilon_e)$  for all edges with  $k(e) \leq n/2$ . By Corollary 1.8 and the inequality (2.7),

$$(2.8) \quad d_{\text{TV}}(\mathcal{L}_T, \mathcal{L}_{T'}) \leq C\alpha.$$

Let  $m = \lfloor n/2 \rfloor$ . Note that if  $e$  and  $e'$  are two edges that share a vertex, then  $|k(e) - k(e')| \leq 1$ . Thus, considering the first  $m$  edges  $e_1, \dots, e_m$  in the optimal path, we see that

$$(2.9) \quad T - T' \geq \sum_{i=1}^m \frac{\varepsilon_{e_i} \omega_{e_i}}{1 + \varepsilon_{e_i}} \geq \frac{C\alpha}{\sqrt{\log n}} \sum_{i=1}^m \frac{\omega_{e_i}}{i}.$$

Notice that the optimal path must necessarily be a self-avoiding path. Let  $P(x, r)$  be the set of all self-avoiding paths of length  $r$  starting at a vertex  $x$ . For  $\theta \geq 0$ , let

$$\phi(\theta) := \mathbb{E}(e^{-\theta \omega_e}),$$

where  $e$  denotes a generic edge. Since the density of the edge weight distribution is uniformly bounded,

$$\phi(\theta) \leq C \int_0^\infty e^{-\theta y} dy = \frac{C}{\theta}.$$

Take any  $x$  and  $r$ , and any path in  $P(x, r)$ . Let  $e_1, \dots, e_r$  be the sequence of edges in the path. Then for any  $\theta > 0$  and  $b > 0$ ,

$$\mathbb{P}\left(\sum_{i=1}^r \omega_{e_i} \leq br\right) \leq e^{\theta br} \phi(\theta)^r \leq C^r \frac{e^{\theta br}}{\theta^r}.$$

Choosing  $\theta = 1/b$ , we get

$$\mathbb{P}\left(\sum_{i=1}^r \omega_{e_i} \leq br\right) \leq (Cb)^r.$$

Fix some  $b$ . Let  $E_r$  be the event that there is a self-avoiding path, starting at some point at distance  $\leq r$  from the origin, of length  $r$  and weight  $\leq br$ . Since there are  $\leq Cr^2$  points at distance  $\leq r$  from the origin, and  $\leq C^r$  paths of length  $r$  from any given starting point, the above inequality shows that

$$(2.10) \quad \mathbb{P}(E_r) \leq C^r r^2 (Cb)^r \leq C^r b^r.$$

Let  $E_r^c$  denote the complement of  $E_r$ , and define

$$F := \bigcap_{k=1}^{\lfloor \log_2 m \rfloor} E_{2^k}^c.$$

Then by the previous (2.10),

$$(2.11) \quad \mathbb{P}(F) \geq 1 - \sum_{k=1}^{\lfloor \log_2 m \rfloor} C^{2^k} b^{2^k}.$$

Suppose that  $F$  happens. Then for any path of length  $m$ , starting at 0, with edges  $e_1, \dots, e_m$ ,

$$\begin{aligned}
 \sum_{i=1}^m \frac{\omega_{e_i}}{i} &\geq \sum_{k=1}^{\lfloor \log_2 m \rfloor} \sum_{i=2^{k-1}}^{2^k-1} \frac{\omega_{e_i}}{i} \\
 (2.12) \qquad &\geq \sum_{k=1}^{\lfloor \log_2 m \rfloor} \frac{1}{2^k} \sum_{i=2^{k-1}}^{2^k-1} \omega_{e_i} \\
 &> \sum_{k=1}^{\lfloor \log_2 m \rfloor} \frac{b2^{k-1}}{2^k} = \frac{1}{2} b \lfloor \log_2 m \rfloor.
 \end{aligned}$$

By (2.9), (2.11) and (2.12), choosing  $\alpha$  and  $b$  small enough, it follows that there exists a positive constant  $a$ , depending only on the edge weight distribution, such that

$$\mathbb{P}(T - T' \geq a\sqrt{\log n}) \geq \frac{1}{2}.$$

By (2.8) and the above inequality, Lemma 1.2 completes the proof.  $\square$

For  $x \in \mathbb{Z}^d$ , let  $T_x := T(0, x)$ . Extend the definition of  $T_x$  to  $x \in \mathbb{R}^d$  in some reasonable way such that the map  $x \mapsto T_x$  is continuous. For example, one can define  $T_x$  in the interior of every  $k$ -cell of  $\mathbb{Z}^d$  as the unique harmonic function that takes specified values on the boundary of the cell, starting with  $k = 1$  (because  $T_x$  is a priori defined on every 0-cell), and inductively going up to  $k = d$ . The reason for insisting on the continuity of the extension is that it will make it technically easier for us to carry out a certain step in the proof of the next theorem.

For  $t \geq 0$ , define the random set

$$B(t) := \{x \in \mathbb{R}^d : T_x \leq t\}.$$

The shape theorem of Cox and Durrett [22] says that under mild conditions on the edge weight distribution, there exists a deterministic compact symmetric convex set  $B_0$  with nonempty interior, such that almost surely, for all  $\varepsilon > 0$ ,

$$(1 - \varepsilon)B_0 \subseteq \frac{1}{t}B(t) \subseteq (1 + \varepsilon)B_0 \qquad \text{for all large } t.$$

The set  $B_0$  is called the ‘limit shape’ of first-passage percolation with the given edge weight distribution.

A unit vector  $x \in \mathbb{R}^d$  is called a ‘direction of curvature’ if in a neighborhood of the boundary point of  $B_0$  in the direction  $x$ ,  $B_0$  is “at least as curved as an Euclidean sphere.” Formally, this can be defined as in [46], as follows. Take any unit vector  $x$ . Let  $z$  be the boundary point of  $B_0$  in the direction  $x$ . We say that  $x$  is a direction of curvature if there is an Euclidean ball  $D$ , with any center, such that

$D \supseteq B_0$  and  $z \in \partial D$ . Simple geometric considerations imply that  $B_0$  has at least one direction of curvature.

Suppose that  $d = 2$  and  $x$  is a direction of curvature. Then, under mild conditions on the edge weight distribution, Newman and Piza [46] showed that for all  $\varepsilon > 0$ ,  $\text{Var}(T_{nx}) \geq Cn^{1/4-\varepsilon}$  for all  $n$ , where  $C$  does not depend on  $n$  (but may depend on the edge weight distribution and  $\varepsilon$ ). However, as we have observed before, this does not actually prove anything about the true order of fluctuations of  $T_{nx}$  since we do not have a matching upper bound. The following theorem fills this gap.

**THEOREM 2.7.** *Suppose that  $d = 2$  and the edge weight distribution belongs to the class  $\mathcal{P}^+(1)$ . Take any  $\varepsilon > 0$ . If  $x$  is a direction of curvature, then the first-passage time  $T_{nx}$  has fluctuations of order at least  $n^{1/8-\varepsilon}$  in the sense of Definition 1.1.*

**PROOF.** Fix some  $n$  and  $\varepsilon$ . Let  $V_n$  be the set of all points in  $\mathbb{Z}^2$  that are within Euclidean distance  $n^{3/4+2\varepsilon}$  of the straight line joining 0 and  $nx$ . It follows by a standard argument (as outlined, e.g., in [17], Section 6) using the concentration properties of first-passage times, Alexander’s rate of convergence theorem [4, 5], and the curvature of  $B_0$  in the direction  $x$ , that the geodesic from 0 to any lattice point near  $nx$  lies entirely in  $V_n$  with probability tending to one as  $n \rightarrow \infty$ .

Replace the edge weights  $\omega_e$  in  $V_n$  by  $\omega'_e := \omega_e/(1 + \alpha n^{-7/8-\varepsilon})$ , where the constant  $\alpha$  will be chosen later. Keep all other edge weights the same as before. Let  $T'_{nx}$  be the first-passage time from 0 to  $nx$  in this new environment. Then by Corollary 1.8, it follows that

$$d_{\text{TV}}(\mathcal{L}_{T_{nx}}, \mathcal{L}_{T'_{nx}}) \leq C\alpha,$$

where  $C$  does not depend on  $n$ . On the other hand, if the original geodesic from 0 to  $[nx]$  is wholly contained in  $V_n$ , then

$$T'_{nx} \leq \frac{T_{nx}}{1 + \alpha n^{-7/8-\varepsilon}}.$$

Since  $T_{nx}/n$  converges to a deterministic positive limit [e.g., by the shape theorem, which applies to edge weight distribution in  $\mathcal{P}^+(1)$ ], this shows that there is some positive constant  $c$  such that

$$\lim_{n \rightarrow \infty} \mathbb{P}(T_{nx} - T'_{nx} \geq cn^{1/8-\varepsilon}) = 1.$$

Choosing  $\alpha$  small enough, Lemma 1.2 completes the proof.  $\square$

The behavior of the set  $B(t)$  is of interest in the theory of first-passage percolation. Newman and Piza [46] defined the following exponent to measure the discrepancy of  $t^{-1}B(t)$  from the limit shape  $B_0$ :

$$(2.13) \quad \chi' := \inf\{\kappa : (t - t^\kappa)B_0 \subseteq B(t) \subseteq (t + t^\kappa)B_0 \text{ for all large } t \text{ a.s.}\}.$$

As far as I know, it has not been proved that  $\chi' > 0$  under any conditions. Newman and Piza [46] showed that in  $d = 2$ ,  $\max\{\chi', \chi\} \geq 1/5$ , where  $\chi$  is an exponent defined in terms of lower bounds on variances of first-passage times. To be precise,  $\chi = \sup \chi_x$ , where the supremum is taken over all unit vectors  $x$ , and

$$\chi_x := \sup\{\gamma \geq 0 : \text{for some } C > 0, \text{Var}(T_{nx}) \geq Cn^{2\gamma} \text{ for all } n\}.$$

Unfortunately, the exponent  $\chi$  does not contain any information about the fluctuations of  $B(t)$  in the absence of matching upper bounds. The only result that I know, that gives a true lower bound on the order of fluctuations of  $B(t)$ , is a theorem of Zhang [67]. Zhang showed that in any dimension,  $B(t)$  has fluctuations of order at least  $\log t$  in a certain sense, if the edge weights are Bernoulli random variables. Very recently, this result has been extended to a general class of edge weight distributions by Nakajima [45]. Another relevant result, due to Auffinger, Damron and Hanson [8], gives a lower bound on the discrepancy between the expected first-passage times and their limiting values—but this does not say anything about fluctuations. The following theorem is the first result that shows  $\chi' > 0$  in two-dimensional first-passage percolation.

**THEOREM 2.8.** *Suppose that  $d = 2$  and the edge weight distribution belongs to the class  $\mathcal{P}^+(1)$ . Let  $\chi'$  be the Newman–Piza shape fluctuation exponent defined in (2.13). Then  $\chi' \geq 1/8$ .*

**PROOF.** Take any  $\kappa$  such that almost surely, for all large  $t$ ,

$$(t - t^\kappa)B_0 \subseteq B(t) \subseteq (t + t^\kappa)B_0.$$

If  $\chi' \geq 1$ , there is nothing to prove. So assume that  $\chi' < 1$ . Then we can take  $\kappa < 1$ . Let  $E_t$  denote the event in the above display, and let

$$F_t := \bigcap_{s \geq t} E_s.$$

From the above characterization of  $\kappa$ , we have

$$(2.14) \quad \lim_{t \rightarrow \infty} \mathbb{P}(F_t) = 1.$$

Take any  $t > 0$ . Let  $x$  be a direction of curvature, and let  $z$  be the unique point on  $\partial B_0$  in the direction  $x$ . Let  $u$  and  $v$  solve  $u + u^\kappa = t$  and  $v - v^\kappa = t$ . Suppose that  $E_v$  happens. Then

$$tB_0 = (v - v^\kappa)B_0 \subseteq B(v),$$

which implies that  $T_{tz} \leq v$ . Again, if  $E_u$  happens, then

$$tB_0 = (u + u^\kappa)B_0 \supseteq B(u),$$

which implies that  $T_{tz} \geq u$ , since if  $T_{tz} < u$ , then we can find  $t' > t$  such that  $T_{t'z} < u$  due to the continuity of the map  $y \mapsto T_y$  (but  $t'z \notin tB_0$ , which gives

a contradiction). If  $F_u$  happens, then both  $E_u$  and  $E_v$  happen, and so the above argument shows that  $u \leq T_{tz} \leq v$ .

Since  $\kappa < 1$ , a simple calculation shows that  $u = t - t^\kappa + o(t^\kappa)$  and  $v = t + t^\kappa + o(t^\kappa)$  as  $t \rightarrow \infty$ . Thus, by (2.14) and the conclusion of the previous paragraph, we get

$$(2.15) \quad \lim_{t \rightarrow \infty} \mathbb{P}(|T_{tz} - t| \leq 2t^\kappa) = 1.$$

Recall that  $z$  is a scalar multiple of  $x$ . Therefore, (2.15) implies that there is some constant  $c$  such that

$$\lim_{n \rightarrow \infty} \mathbb{P}(|T_{nx} - cn| \leq 2(cn)^\kappa) = 1.$$

By Theorem 2.7, this is impossible unless  $\kappa \geq 1/8$ .  $\square$

2.4. *The random assignment problem.* Suppose that we have to assign  $n$  tasks to  $n$  workers, and  $a_{ij}$  is the cost of assigning task  $j$  to worker  $i$ . Suppose that the  $a_{ij}$ 's are i.i.d. nonnegative random variables. Let  $S_n$  be the group of all permutations of  $\{1, \dots, n\}$  and let

$$C_n := \min_{\pi \in S_n} \sum_{i=1}^n a_{i\pi(i)}.$$

The problem of computing  $C_n$  is known as the random assignment problem. Aldous [2] proved that if the law of the costs has a density  $f$  that is nonzero and finite in a neighborhood of zero, then  $C_n$  converges to a deterministic limit as  $n \rightarrow \infty$ . Moreover, the limit depends only on the value of  $f(0)$  (assuming that  $f$  is continuous at 0). In [3], Aldous proved that if  $f(0) = 1$ , then the limit is  $\zeta(2) = \pi^2/6$ , confirming a conjecture of Mézard and Parisi [42, 43]. Later, an exact formula for  $\mathbb{E}(C_n)$ , when the costs are exponentially distributed, was obtained by Linusson and Wästlund [38] and Nair, Prabhakar and Sharma [44]. When the density of the cost distribution at zero is either blowing up to infinity or converging to zero, the situation becomes more complicated. These cases have been investigated by Wästlund [65].

The fluctuations of  $C_n$ , however, are not as well understood. Talagrand [58] proved using his general machinery that if the costs are uniformly distributed in  $[0, 1]$ , then the fluctuations of  $C_n$  are at most of order

$$\frac{(\log n)^2}{\sqrt{n} \log \log n}.$$

A number of exact calculations for fluctuations are possible when the costs are exponentially distributed with mean one. Under this assumption, Alm and Sorkin [6] proved that the variance of  $C_n$  is at least of order  $1/n$ , and then Wästlund [63, 64] derived the following asymptotic formula for the variance:

$$\text{Var}(C_n) = \frac{4\zeta(2) - 4\zeta(3)}{n} + O\left(\frac{1}{n^2}\right).$$

Wästlund [63, 64] also gave formulas for higher moments when the costs are exponentially distributed, but the asymptotics of these formulas are hard to understand. There is, however, a more direct way to extract lower bounds on fluctuations when the costs are exponentially distributed, using the memoryless property of the exponential distribution. This has been worked out in a related model by Hessler and Wästlund [27], also appearing in the Ph.D. thesis of Hessler [26].

When the costs are not exponential, nothing is known about lower bounds on the fluctuations of  $C_n$ . The following theorem gives a general lower bound of order  $n^{-1/2}$ , which, in the face of the evidence presented above, appears to be the correct order (recall that an upper bound of order  $n^{-1/2}$  is not yet known for nonexponential costs). The proof uses Lemma 1.2, but the simple multiplicative perturbation of Section 1.5 does not give the correct answer for this problem. Instead, a more complicated coupling is used.

**THEOREM 2.9.** *Suppose that the cost distribution belongs to the class  $\mathcal{P}^+(1)$ . Then the optimal cost  $C_n$  in the random assignment problem has fluctuations of order at least  $n^{-1/2}$ , in the sense of Definition 1.1.*

**PROOF.** Throughout this proof,  $C$  will denote any positive constant that may depend only on the cost distribution and nothing else. The value of  $C$  may change from line to line or even within a line.

Define a function  $\phi : [0, \infty) \rightarrow [0, \infty)$  as

$$\phi(x) = \begin{cases} \sqrt{n}x & \text{if } 0 \leq x \leq 1/n, \\ x + 1/\sqrt{n} - 1/n & \text{if } x > 1/n. \end{cases}$$

Then  $\phi$  is absolutely continuous and strictly increasing, and  $\phi(0) = 0$ . Moreover,  $\phi(x) \leq x + 1$  for all  $x$ , and

$$(2.16) \quad \phi'(x) = \begin{cases} \sqrt{n} & \text{if } 0 < x < 1/n, \\ 1 & \text{if } x > 1/n. \end{cases}$$

Let  $e^{-V}$  be the density function of the cost distribution. Let  $X$  be a nonnegative random variable with probability density  $e^{-V}$ . Let  $Y$  solve

$$Y + \alpha n^{-1} \phi(Y) = X,$$

where  $\alpha$  is a positive constant that will be chosen later. Since the map  $x \mapsto x + \alpha n^{-1} \phi(x)$  is strictly increasing and continuous on  $[0, \infty)$  and sends 0 to 0, it is a bijection of  $[0, \infty)$  onto itself. Therefore,  $Y$  is uniquely defined. The probability density function of  $Y$  is

$$(1 + \alpha n^{-1} \phi'(x)) e^{-V(x + \alpha n^{-1} \phi(x))}.$$

For  $\varepsilon \in (-1/2, 1/2)$ , define

$$g(\varepsilon) := \int_0^\infty \sqrt{(1 + \varepsilon\phi'(x))e^{-V(x+\varepsilon\phi(x))}e^{-V(x)}} dx,$$

so that

$$\rho(\mathcal{L}_X, \mathcal{L}_Y) = g(\alpha n^{-1}).$$

As in the proof of Theorem 1.7, it is easy to verify using the assumptions on  $V$  and  $\phi$  that  $g$  is a  $C^\infty$  function of  $\varepsilon$  and the derivatives with respect to  $\varepsilon$  can be taken inside the integral. Note that  $g(0) = 1$ , and

$$g'(\varepsilon) = \int_0^\infty \frac{\phi'(x)}{2\sqrt{1 + \varepsilon\phi'(x)}} e^{-(V(x+\varepsilon\phi(x))+V(x))/2} dx - \frac{1}{2} \int_0^\infty \sqrt{1 + \varepsilon\phi'(x)} V'(x + \varepsilon\phi(x)) \phi(x) e^{-(V(x+\varepsilon\phi(x))+V(x))/2} dx.$$

Thus,

$$g'(0) = \frac{1}{2} \int_0^\infty (\phi'(x) - V'(x)\phi(x)) e^{-V(x)} dx,$$

which equals zero by integration by parts, since  $\phi(0) = 0$ . Next, note that

$$g''(\varepsilon) = - \int_0^\infty \frac{\phi'(x)^2}{4(1 + \varepsilon\phi'(x))^{3/2}} e^{-(V(x+\varepsilon\phi(x))+V(x))/2} dx - \frac{1}{2} \int_0^\infty \frac{\phi'(x)}{\sqrt{1 + \varepsilon\phi'(x)}} V'(x + \varepsilon\phi(x)) \phi(x) e^{-(V(x+\varepsilon\phi(x))+V(x))/2} dx - \frac{1}{2} \int_0^\infty \sqrt{1 + \varepsilon\phi'(x)} V''(x + \varepsilon\phi(x)) \phi(x)^2 e^{-(V(x+\varepsilon\phi(x))+V(x))/2} dx + \frac{1}{4} \int_0^\infty \sqrt{1 + \varepsilon\phi'(x)} V'(x + \varepsilon\phi(x))^2 \phi(x)^2 e^{-(V(x+\varepsilon\phi(x))+V(x))/2} dx.$$

Because of (2.16), it is convenient to write each of the four terms in the above display as a sum of two integrals—one from 0 to  $1/n$  and another from  $1/n$  to  $\infty$ . Due to (2.16) and the boundedness of  $V$  near zero, the magnitude of the first part is bounded by  $C$  in all four cases. In the second part,  $\phi'(x)$  is bounded by 1 and  $\phi(x)$  is bounded by  $x + 1$ . Therefore, using the properties of  $V$  and a few applications of the Cauchy–Schwarz inequality and change-of-variables, we see that the magnitude of the second part is also bounded by  $C$  in all four cases. Thus,

$$\sup_{-1/2 < \varepsilon < 1/2} |g''(\varepsilon)| \leq C.$$

As a consequence,

$$(2.17) \quad \rho(\mathcal{L}_X, \mathcal{L}_Y) = g(\alpha n^{-1}) \geq 1 - C\alpha^2 n^{-2}.$$

Now for each  $i, j$ , let  $a'_{ij}$  solve

$$a'_{ij} + \alpha n^{-1} \phi(a'_{ij}) = a_{ij}.$$

Let  $C'_n$  be the optimal assignment cost with these new costs. By (2.17) and Lemma 1.4, it follows that

$$(2.18) \quad d_{\text{TV}}(\mathcal{L}_{C_n}, \mathcal{L}_{C'_n}) \leq C\alpha.$$

For  $1 \leq i \leq n$ , let

$$b_i := \min_{1 \leq j \leq n} a_{ij}.$$

Since  $e^{-V}$  is a bounded density,

$$(2.19) \quad \mathbb{P}(b_i \geq 1/n) \geq (1 - C/n)^n \geq K,$$

where  $K$  is some positive constant that does not depend on  $n$ . Let

$$A := \{i : b_i \geq 1/n\}.$$

Since  $b_1, \dots, b_n$  are i.i.d. random variables, (2.19) shows that

$$(2.20) \quad \lim_{n \rightarrow \infty} \mathbb{P}(|A| \geq Kn/2) = 1.$$

Now notice that  $a'_{ij} \leq a_{ij}$  for all  $i$  and  $j$ . Moreover, if  $i \in A$ , then for any  $j$ ,  $a_{ij} \geq 1/n$ . Since  $x \mapsto x + n^{-1} \phi(x)$  is an increasing map, this implies that  $a'_{ij} \geq x_n$ , where  $x_n$  is the unique solution of

$$x_n + \alpha n^{-1} \phi(x_n) = n^{-1}.$$

A simple calculation shows that

$$x_n = \frac{1}{n + \alpha \sqrt{n}}.$$

Thus, if  $i \in A$ , then for any  $j$ ,

$$a'_{ij} \geq \frac{1}{n + \alpha \sqrt{n}},$$

and, therefore,

$$\begin{aligned} a_{ij} - a'_{ij} &= \alpha n^{-1} \phi(a'_{ij}) \\ &\geq \alpha n^{-1} \phi\left(\frac{1}{n + \alpha \sqrt{n}}\right) = \frac{\alpha}{n^{3/2} + \alpha n}. \end{aligned}$$

Combining all observations, we get

$$C_n - C'_n \geq \frac{\alpha |A|}{n^{3/2} + \alpha n}.$$

By (2.18), (2.20) and Lemma 1.2, this completes the proof.  $\square$

2.5. *Determinants of random matrices.* Let  $n$  and  $N$  be two positive integers, and let  $f$  be a measurable function from  $\mathbb{R}^n$  into  $\mathbb{R}^{N \times N}$ , the set of all  $N \times N$  real matrices. Suppose that there is some  $r > 0$  such that for all  $x_1, \dots, x_n \in \mathbb{R}$  and all  $\lambda \geq 0$ ,

$$(2.21) \quad f(\lambda x_1, \dots, \lambda x_n) = \lambda^r f(x_1, \dots, x_n).$$

Let  $X_1, \dots, X_n$  be i.i.d. random variables. Then  $M = f(X_1, \dots, X_n)$  is an  $N \times N$  random matrix. Many common families of random matrices, such as Wigner matrices, sample covariance matrices, random Toeplitz and Hankel matrices, and random band matrices, can be obtained in the above manner as functions of independent random variables where the function satisfies (2.21) for some  $r$ . For example, for a Wigner matrix of order  $N$ ,  $r = 1$  and  $n = N(N + 1)/2$ . The following theorem gives a lower bound on the order of fluctuations of  $\log |\det M|$ .

**THEOREM 2.10.** *Let  $X_1, X_2, \dots$  be a sequence of i.i.d. random variables with probability density in either  $\mathcal{P}(1)$  or  $\mathcal{P}^+(1)$ . For each  $n$ , let  $N_n$  be a positive integer and let  $f_n$  be a measurable function from  $\mathbb{R}^n$  into  $\mathbb{R}^{N_n \times N_n}$  satisfying (2.21) for some fixed  $r > 0$ . Let  $M_n := f_n(X_1, \dots, X_n)$ . Then  $\log |\det M_n|$  has fluctuations of order at least  $n^{-1/2}N_n$ , in the sense of Definition 1.1.*

**PROOF.** Throughout this proof,  $C$  will denote any constant that may depend only on the distribution of the  $X_i$ 's, but not on  $n$ .

Fix  $n$ . Let  $X'_i := X_i/(1 + \alpha n^{-1/2})$ , where  $\alpha$  will be determined later. Let  $M'_n := f_n(X'_1, \dots, X'_n)$ . Let  $L_n := \log |\det M_n|$  and  $L'_n := \log |\det M'_n|$ . By Corollary 1.8,

$$d_{TV}(\mathcal{L}_{L_n}, \mathcal{L}_{L'_n}) \leq d_{TV}(\mathcal{L}_{(X_1, \dots, X_n)}, \mathcal{L}_{(X'_1, \dots, X'_n)}) \leq C\alpha.$$

On the other hand, by (2.21),

$$\det M'_n = (1 + \alpha n^{-1/2})^{rN_n} \det M_n,$$

and hence

$$L'_n = L_n + rN_n \log(1 + \alpha n^{-1/2}).$$

An application of Lemma 1.2 completes the proof.  $\square$

Consider now the special case of sample covariance matrices. Let  $n$  and  $p$  be two positive integers, and let  $Y_1, \dots, Y_n$  be i.i.d.  $p$ -dimensional random vectors, whose components are i.i.d. random variables. The sample covariance matrix for this data is defined as

$$(2.22) \quad W := \frac{1}{n} \sum_{i=1}^p (Y_i - \bar{Y})(Y_i - \bar{Y})^T,$$

where  $x^T$  denotes the transpose of a column vector  $x$ , and

$$\bar{Y} := \frac{1}{n} \sum_{i=1}^n Y_i.$$

Note that this is slightly different than the Wishart matrices usually considered in random matrix theory, which have the same definition, except that  $\bar{Y}$  is not subtracted from  $Y_i$  in (2.22).

A sample covariance matrix is positive semi-definite and so its determinant is always nonnegative. The logarithm of the determinant of a sample covariance matrix is an important object in statistics, where it is used to perform tests of hypotheses [7]. A central limit theorem for  $\log \det W$ , when  $p$  is fixed,  $n \rightarrow \infty$  and the  $Y_i$ 's are complex Gaussian random vectors, was established in 1963 by Goodman [25]. The high dimensional case, where  $p$  and  $n$  both tend to infinity, was solved by Cai, Liang and Zhou [15] in 2015. (The corresponding result for Wishart matrices, however, is standard fare in random matrix theory; see, e.g., [50], Chapter 7.) The main result of [15] is a central limit theorem for  $\log \det W$  when the  $Y_i$ 's are real Gaussian random vectors (possibly with correlated coordinates),  $n \rightarrow \infty$  and  $p$  is allowed to vary arbitrarily but under the constraint that  $p \leq n$  (so that  $\det W \neq 0$ ). In this scenario, Cai, Liang and Zhou [15] show that  $\log \det W$  has Gaussian fluctuations of order  $\sqrt{p/n}$  if  $p/n \rightarrow r \in [0, 1)$ , and of order  $\sqrt{\log n}$  if  $p/n \rightarrow 1$ .

The non-Gaussian case is open. In particular, a central limit theorem for  $\log \det W$  has not been proved in the setting described above, that is,  $Y_i$ 's having i.i.d. but not necessarily Gaussian coordinates. The following corollary of Theorem 2.10 provides a lower bound on the fluctuation of  $\log \det W$  which appears to be of the correct order if  $p/n \rightarrow r \in [0, 1)$ , in view of the result of Cai, Liang and Zhou [15].

**COROLLARY 2.11.** *Let  $Y_1, \dots, Y_n$  be i.i.d.  $p$ -dimensional random vectors and let  $W$  be defined as in (2.22). Suppose that the coordinates of  $Y_i$ 's are i.i.d. with probability density in  $\mathcal{P}(1)$  (which remains fixed as  $n$  and  $p$  vary). Then as  $n \rightarrow \infty$  and  $p$  varies arbitrarily as a function of  $n$  (with the constraint that  $p \leq n$ ),  $\log \det W$  has fluctuations of order at least  $\sqrt{p/n}$ , in the sense of Definition 1.1.*

**PROOF.** Note that the  $p \times p$  matrix  $W$  is a function of  $np$  i.i.d. random variables, and this function satisfies (2.21) with  $r = 2$ . Thus, the  $n$  in Theorem 2.10 should be replaced by  $np$  and  $N_n$  should be replaced by  $p$ . With these replacements, the lower bound on the order of fluctuations turns out to be  $(np)^{-1/2} p = \sqrt{p/n}$ , proving the claim.  $\square$

It is surprising to me that a soft technique based on Lemma 1.2 and Corollary 1.8 can actually yield the correct lower bound in Corollary 2.11. Indeed, the method does not yield the correct bound for Wigner matrices. Improving on an earlier work

of Tao and Vu [62], Nguyen and Vu [47] proved the central limit theorem for log-determinants of Wigner matrices with non-Gaussian entries in 2014. According to this result, the log-determinant has fluctuations of order  $\sqrt{\log n}$ . A straightforward application of Theorem 2.10, however, only gives a lower bound of order 1. Possibly a better coupling than the one provided by Corollary 1.8 is needed to achieve the  $\sqrt{\log n}$  lower bound.

**3. Open problems.** There are many open questions about lower bounds for fluctuations of random variables. Here is a list of questions that are closely associated with the examples worked out in this paper:

(1) Extend the results of this paper beyond the distribution classes  $\mathcal{P}(d)$  and  $\mathcal{P}^+(d)$ . In particular, proofs under minimal assumptions would be very desirable.

(2) In the traveling salesman problem for uniformly distributed points on  $[0, 1]^2$ , prove that the variance of the length of the optimal tour converges to a constant as the number of points tends to infinity, and identify this constant if possible. This conjecture is due to Mike Steele, who told me about it in a personal communication.

(3) Prove a tight lower bound for the fluctuations of the length of the minimal matching when the points are uniformly distributed in  $[0, 1]^d$ .

(4) Prove a tight lower bound for fluctuations in the longest common subsequence problem for random words. Considerable progress on this problem has been made in [24, 29, 30, 35], but the most important case of uniformly distributed letters is open. A solution of this problem would complete the proof of the central limit theorem for longest common subsequences, as shown by Houdré and Işlak [28].

(5) Improve the lower bound for the fluctuations of the first-passage time in two-dimensional first-passage percolation.

(6) Extend the lower bound result for first-passage percolation to the case of discrete edge weights.

(7) Prove any nontrivial lower bound for the fluctuations of the first-passage time in higher dimensions (a lower bound of order 1 is easy using the method of this paper).

(8) Improve the lower bound for the Newman–Piza exponent  $\chi'$  for two-dimensional first-passage percolation.

(9) Show that  $\chi' > 0$  in higher dimensions.

(10) Improve the lower bound on the order of fluctuations of the free energy of the S-K model or show that it is optimal. Same for the ground state energy.

(11) Prove the optimality of the lower bound in the random assignment problem for cost distributions in  $\mathcal{P}^+(1)$  or any other general class of cost distributions.

(12) Prove tight lower bounds for fluctuations of functionals of random matrices other than the determinant, such as linear statistics of eigenvalues and the maximum and minimum eigenvalues, in ensembles where such results are not known.

This includes a variety of patterned random matrix ensembles, such as random Toeplitz and Hankel matrices and random band matrices. For a survey of results on patterned random matrices, see [13].

(13) Prove a distributional limit theorem in any of the examples discussed in this paper.

**Acknowledgments.** I am grateful to David Aldous, Louis-Pierre Arguin, Erik Bates, Wei-Kuo Chen, Michael Damron, Peter Forrester, Christian Houdré, Svante Janson, Ron Peled, Yuval Peres, Mike Steele, Johan Wästlund, Harry Zhou and the anonymous referees for many helpful comments and references.

## REFERENCES

- [1] AIZENMAN, M., LEBOWITZ, J. L. and RUELLE, D. (1987). Some rigorous results on the Sherrington–Kirkpatrick spin glass model. *Comm. Math. Phys.* **112** 3–20. [MR0904135](#)
- [2] ALDOUS, D. (1992). Asymptotics in the random assignment problem. *Probab. Theory Related Fields* **93** 507–534. [MR1183889](#)
- [3] ALDOUS, D. J. (2001). The  $\zeta(2)$  limit in the random assignment problem. *Random Structures Algorithms* **18** 381–418. [MR1839499](#)
- [4] ALEXANDER, K. S. (1993). A note on some rates of convergence in first-passage percolation. *Ann. Appl. Probab.* **3** 81–90. [MR1202516](#)
- [5] ALEXANDER, K. S. (1997). Approximation of subadditive functions and convergence rates in limiting-shape results. *Ann. Probab.* **25** 30–55. [MR1428498](#)
- [6] ALM, S. E. and SORKIN, G. B. (2002). Exact expectations and distributions for the random assignment problem. *Combin. Probab. Comput.* **11** 217–248. [MR1909500](#)
- [7] ANDERSON, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Wiley-Interscience, Hoboken, NJ. [MR1990662](#)
- [8] AUFFINGER, A., DAMRON, M. and HANSON, J. (2015). Rate of convergence of the mean for sub-additive ergodic sequences. *Adv. Math.* **285** 138–181. [MR3406498](#)
- [9] AUFFINGER, A., DAMRON, M. and HANSON, J. (2017). *50 Years of First-Passage Percolation. University Lecture Series* **68**. Amer. Math. Soc., Providence, RI. [MR3729447](#)
- [10] BENAÏ M, M. and ROSSIGNOL, R. (2008). Exponential concentration for first passage percolation through modified Poincaré inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 544–573. [MR2451057](#)
- [11] BENJAMINI, I., KALAI, G. and SCHRAMM, O. (2003). First passage percolation has sublinear distance variance. *Ann. Probab.* **31** 1970–1978. [MR2016607](#)
- [12] BOLLOBÁS, B. and JANSON, S. (1997). On the length of the longest increasing subsequence in a random permutation. In *Combinatorics, Geometry and Probability* (Cambridge, 1993) 121–128. Cambridge Univ. Press, Cambridge. [MR1476438](#)
- [13] BOSE, A., SUBHRA HAZRA, R. and SAHA, K. (2010). Patterned random matrices and method of moments. In *Proceedings of the International Congress of Mathematicians. Volume IV* 2203–2231. Hindustan Book Agency, New Delhi. [MR2827968](#)
- [14] BOUCHERON, S., LUGOSI, G. and MASSART, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford Univ. Press, Oxford. [MR3185193](#)
- [15] CAI, T. T., LIANG, T. and ZHOU, H. H. (2015). Law of log determinant of sample covariance matrix and optimal estimation of differential entropy for high-dimensional Gaussian distributions. *J. Multivariate Anal.* **137** 161–172. [MR3332804](#)
- [16] CHATTERJEE, S. (2009). Disorder chaos and multiple valleys in spin glasses Preprint. Available at <https://arxiv.org/abs/0907.3381>.

- [17] CHATTERJEE, S. (2013). The universal relation between scaling exponents in first-passage percolation. *Ann. of Math. (2)* **177** 663–697. [MR3010809](#)
- [18] CHATTERJEE, S. (2014). *Superconcentration and Related Topics*. Springer, Cham. [MR3157205](#)
- [19] CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q.-M. (2011). *Normal Approximation by Stein's Method*. Springer, Heidelberg. [MR2732624](#)
- [20] CHEN, W.-K., DEY, P. and PANCHENKO, D. (2017). Fluctuations of the free energy in the mixed  $p$ -spin models with external field. *Probab. Theory Related Fields* **168** 41–53. [MR3651048](#)
- [21] CHEN, W.-K., HANDSCHY, M. and LERMAN, G. (2018). On the energy landscape of the mixed even  $p$ -spin model. *Probab. Theory Related Fields* **171** 53–95. [MR3800830](#)
- [22] COX, J. T. and DURRETT, R. (1981). Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Probab.* **9** 583–603. [MR0624685](#)
- [23] DAMRON, M., HANSON, J. and SOSOE, P. (2015). Sublinear variance in first-passage percolation for general distributions. *Probab. Theory Related Fields* **163** 223–258. [MR3405617](#)
- [24] GONG, R., HOUDRÉ, C. and LEMBER, J. (2018). Lower bounds on the generalized central moments of the optimal alignments score of random sequences. *J. Theoret. Probab.* **31** 643–683. [MR3803910](#)
- [25] GOODMAN, N. R. (1963). The distribution of the determinant of a complex Wishart distributed matrix. *Ann. Math. Stat.* **34** 178–180. [MR0145619](#)
- [26] HESSLER, M. (2009). Optimization, matroids and error-correcting codes. Ph.D. thesis, Linköping Univ.
- [27] HESSLER, M. and WÄSTLUND, J. (2008). Concentration of the cost of a random matching problem. Preprint. Available at <http://www.math.chalmers.se/~wastlund/martingale.pdf>.
- [28] HOUDRÉ, C. and IŞLAK, U. (2014). A central limit theorem for the length of the longest common subsequences in random words. Preprint. Available at <https://arxiv.org/abs/1408.1559>.
- [29] HOUDRÉ, C. and MA, J. (2016b). On the order of the central moments of the length of the longest common subsequences in random words. In *High Dimensional Probability VII. Progress in Probability* **71** 105–136. Springer, Cham. [MR3565261](#)
- [30] HOUDRÉ, C. and MATZINGER, H. (2016). On the variance of the optimal alignments score for binary random words and an asymmetric scoring function. *J. Stat. Phys.* **164** 693–734. [MR3519215](#)
- [31] JANSON, S. (1994). Self-couplings and the concentration function. *Acta Appl. Math.* **34** 5–6. [MR1273842](#)
- [32] JANSON, S. and WARNKE, L. (2016). The lower tail: Poisson approximation revisited. *Random Structures Algorithms* **48** 219–246. [MR3449596](#)
- [33] LE CAM, L. and YANG, G. L. (2000). *Asymptotics in Statistics: Some Basic Concepts*, 2nd ed. Springer, New York. [MR1784901](#)
- [34] LEDOUX, M. (2001). *The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs* **89**. Amer. Math. Soc., Providence, RI. [MR1849347](#)
- [35] LEMBER, J. and MATZINGER, H. (2009). Standard deviation of the longest common subsequence. *Ann. Probab.* **37** 1192–1235. [MR2537552](#)
- [36] LEVIN, D. A., PERES, Y. and WILMER, E. L. (2009). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. [MR2466937](#)
- [37] LÉVY, P. (1937). *Théorie de L'addition des Variables Aléatoires*. Gauthier-Villars, Paris.
- [38] LINUSSON, S. and WÄSTLUND, J. (2004). A proof of Parisi's conjecture on the random assignment problem. *Probab. Theory Related Fields* **128** 419–440. [MR2036492](#)
- [39] MCBRYAN, O. A. and SPENCER, T. (1977). On the decay of correlations in  $SO(n)$ -symmetric ferromagnets. *Comm. Math. Phys.* **53** 299–302. [MR0441179](#)

- [40] MERMIN, N. D. (1967). Absence of ordering in certain classical systems. *J. Math. Phys.* **8** 1061–1064.
- [41] MERMIN, N. D. and WAGNER, H. (1966). Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models. *Phys. Rev. Lett.* **17** 1133–1136.
- [42] MÉZARD, M. and PARISI, G. (1985). Replicas and optimization. *J. Phys. Lett.* **46** 771–778.
- [43] MÉZARD, M. and PARISI, G. (1987). On the solution of the random link matching problem. *J. Physique* **48** 1451–1459.
- [44] NAIR, C., PRABHAKAR, B. and SHARMA, M. (2005). Proofs of the Parisi and Coppersmith–Sorkin random assignment conjectures. *Random Structures Algorithms* **27** 413–444. [MR2178256](#)
- [45] NAKAJIMA, S. (2017). Divergence of shape fluctuation in first passage percolation. Preprint. Available at <https://arxiv.org/abs/1706.03493>.
- [46] NEWMAN, C. M. and PIZA, M. S. T. (1995). Divergence of shape fluctuations in two dimensions. *Ann. Probab.* **23** 977–1005. [MR1349159](#)
- [47] NGUYEN, H. H. and VU, V. (2014). Random matrices: Law of the determinant. *Ann. Probab.* **42** 146–167. [MR3161483](#)
- [48] PALASSINI, M. (2008). Ground-state energy fluctuations in the Sherrington–Kirkpatrick model. *J. Stat. Mech.* **2008** P10005.
- [49] PANCHENKO, D. (2013). *The Sherrington–Kirkpatrick Model*. Springer, New York. [MR3052333](#)
- [50] PASTUR, L. and SHCHERBINA, M. (2011). *Eigenvalue Distribution of Large Random Matrices. Mathematical Surveys and Monographs* **171**. Amer. Math. Soc., Providence, RI. [MR2808038](#)
- [51] PELED, R. and SPINKA, Y. Lectures on the Spin and Loop  $O(n)$  Models. Preprint. Available at <https://arxiv.org/abs/1708.00058>.
- [52] PEMANTLE, R. and PERES, Y. (1994). Planar first-passage percolation times are not tight. In *Probability and Phase Transition (Cambridge, 1993)*. NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **420** 261–264. Kluwer Academic, Dordrecht. [MR1283187](#)
- [53] PETROV, V. V. (1975). *Sums of Independent Random Variables. Ergebnisse der Mathematik und Ihrer Grenzgebiete* **82**. Springer, New York-Heidelberg. Translated from the Russian by A. A. Brown. [MR0388499](#)
- [54] PFISTER, C. E. (1981). On the symmetry of the Gibbs states in two-dimensional lattice systems. *Comm. Math. Phys.* **79** 181–188. [MR0612247](#)
- [55] RHEE, W. T. (1991). On the fluctuations of the stochastic traveling salesperson problem. *Math. Oper. Res.* **16** 482–489. [MR1120465](#)
- [56] SHERRINGTON, D. and KIRKPATRICK, S. (1975). Solvable model of a spin glass. *Phys. Rev. Lett.* **35** 1792–1796.
- [57] STEELE, J. M. (1997). *Probability Theory and Combinatorial Optimization. CBMS-NSF Regional Conference Series in Applied Mathematics* **69**. SIAM, Philadelphia, PA. [MR1422018](#)
- [58] TALAGRAND, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Publ. Math. Inst. Hautes Études Sci.* **81** 73–205. [MR1361756](#)
- [59] TALAGRAND, M. (2003). *Spin Glasses: A Challenge for Mathematicians: Cavity and Mean Field Models. Ergebnisse der Mathematik und Ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]* **46**. Springer, Berlin. [MR1993891](#)
- [60] TALAGRAND, M. (2011a). *Mean Field Models for Spin Glasses. Vol. I. Basic Examples. Ergebnisse der Mathematik und Ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]* **54**. Springer, Berlin. [MR2731561](#)

- [61] TALAGRAND, M. (2011b). *Mean Field Models for Spin Glasses. Vol. II. Advanced Replica-Symmetry and Low Temperature. Ergebnisse der Mathematik und Ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]* **55**. Springer, Heidelberg. [MR3024566](#)
- [62] TAO, T. and VU, V. (2012). A central limit theorem for the determinant of a Wigner matrix. *Adv. Math.* **231** 74–101. [MR2935384](#)
- [63] WÄSTLUND, J. (2005). The variance and higher moments in the random assignment problem. *Linköping Studies in Mathematics*, no. 8.
- [64] WÄSTLUND, J. (2010). The mean field traveling salesman and related problems. *Acta Math.* **204** 91–150. [MR2600434](#)
- [65] WÄSTLUND, J. (2012). Replica symmetry of the minimum matching. *Ann. of Math. (2)* **175** 1061–1091. [MR2912702](#)
- [66] WEHR, J. and AIZENMAN, M. (1990). Fluctuations of extensive functions of quenched random couplings. *J. Stat. Phys.* **60** 287–306. [MR1069633](#)
- [67] ZHANG, Y. (2006). The divergence of fluctuations for shape in first passage percolation. *Probab. Theory Related Fields* **136** 298–320. [MR2240790](#)

DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
SEQUOIA HALL, 390 SERRA MALL  
STANFORD, CALIFORNIA 94305  
USA  
E-MAIL: [souravc@stanford.edu](mailto:souravc@stanford.edu)