Local optima of the Sherrington-Kirkpatrick Hamiltonian

Louigi Addario-Berry [*]
Mathematics and Statistics
McGill University,
Montreal, Canada
louigi.addario@mcgill.ca

Luc Devroye [†]
School of Computer Science
McGill University,
Montreal, Canada
lucdevroye@gmail.com

Gábor Lugosi [‡]
ICREA,
Department of Economics,
Pompeu Fabra University,
Barcelona GSE
Barcelona, Spain;
gabor.lugosi@upf.edu

Roberto Imbuzeiro Oliveira [§]
IMPA, Rio de Janeiro, RJ,
Brazil;
rimfo@impa.br

December 17, 2017

Abstract

We sudy local optima of the Hamiltonian of the Sherrington-Kirkpatrick model. We compute the exponent of the expected number of local optima and determine the "typical" value of the Hamiltonian.

1 Local optima of the Hamiltonian

Let $W = (W_{i,j})_{n \times n}$ be a symmetric matrix with zero diagonal such that the $(W_{i,j})_{1 \leq i < j \leq n}$ are independent standard normal random variables. The *Sherrington-Kirpatrick* model of spin glasses is defined by a *random Hamiltonian*, that is, a random function $H : \{-1, +1\}^n \to \mathbb{R}$. For a configuration $\sigma = (\sigma_i)_{i=1}^n \in \{-1, +1\}^n$, $H(\sigma)$ is defined as follows.

$$H(\sigma) := \sum_{1 \leq i < j \leq n} \sigma_i \, \sigma_j \, W_{ij} \, .$$

We follow the usual convention of calling $\sigma \in \{-1, +1\}^n$ a *spin configuration*, the coordinates of σ *spins*, and the value $H(\sigma)$ the *energy* of configuration σ.

Given $i \in [n]$ (where $[n] = \{1, \ldots, n\}$) and σ as above, we let $\sigma^{(i)}$ denote a new configuration obtained from σ by flipping the i-th spin and leaving other coordinates unchanged. That is, the components of $\sigma^{(i)}$ are defined as

$$\sigma_j^{(i)} := \left\{ \begin{array}{ll} -\sigma_i, & j = i \, ; \\ \sigma_j, & j \in [n] \backslash \{i\} \, . \end{array} \right.$$

We say that σ is a *local minimum* or a *local optimum* of H if

$$H(\sigma^{(i)}) \geq H(\sigma) \quad \text{for all } i \in [n] \, .$$

That is, σ is a local minimum if flipping the sign of any individual spin does not decrease the value of the energy.

The global optimum $\min_{\sigma \in \{-1, +1\}^n} H(\sigma)$—called the "ground-state energy"— has been extensively studied. The problem was introduced by Sherrington and Kirkpatrick [10] as a mean-field model for spin glasses. The value of the optimum was determined non-rigorously in the seminal work of Parisi [9], as a consequence of the so-called "Parisi formula". Parisi's formula was proved by Talagrand [11] in a breakthrough paper, see also Panchenko [8] for an overview. It follows from Talagrand's result that

$$n^{-3/2} \min_{\sigma \in \{-1, +1\}^n} H(\sigma) \to -c \quad \text{in probability,}$$

where c is a constant whose value is numerically estimated to be about $0.7632...$ (Crisanti and Rizzo [4]) and known to be bounded by $\sqrt{2/\pi} \approx 0.797885...$ (Guerra [7]).

In this paper we are interested in locally optimal solutions. An important reason of why local optima are worth considering is because local optima may be computed quickly by simple greedy algorithms, see Etscheid and Röglin [5], Angel, Bubeck, Peres, and Wei [2] and subsection 1.2 below. We show that the

expected number of local optima grows exponentially and we establish the rate of growth. Also, we examine the conditional distribution of $H(\sigma)n^{-3/2}$ given that σ is locally optimal. We prove that the distribution is concentrated on an interval of width $O(n^{-1/4})$ and determine its location.

1.1 Results

In order to state the main result of the paper, we need a few definitions.

Let $\Phi(\lambda) = \mathbb{P}\{N \le \lambda\}$ be the distribution function of a standard normal random variable N and introduce $\phi(\lambda) = \log(2\Phi(\lambda))$. For $x \ge \sqrt{2/\pi}$, we let $\mu^*(x)$ denote the Fenchel-Légendre transform

$$\mu^*(x) := \sup_{\lambda \ge 0} \left(\lambda x - \frac{\lambda^2}{2} - \phi(\lambda) \right) .$$

Lemma 2 below shows that $\mu^* : [\sqrt{2/\pi}, +\infty) \to \mathbb{R}$ is well defined. Lemma 4 shows that the mapping

$$R(x) := \frac{x^2}{4} - \mu^*(x)$$

is strictly concave for $x \ge \sqrt{\frac{2}{\pi}}$ and achieves its global maximum at $x = v_* > \sqrt{\frac{2}{\pi}}$. We let $\alpha^* = R(v^*) > 0$ denote the maximum value of R.

Theorem 1. *For any fixed $n \ge 1$, the probability $\mathbb{P}\{\sigma$ is locally optimal$\}$ is the same for all $\sigma = \sigma(n) \in \{-1, +1\}^n$. Moreover,*

$$\lim_{n \to +\infty} \frac{1}{n} \log \mathbb{P}\{\sigma \text{ is locally optimal}\} = \alpha^* - \log 2 .$$

Also, there exist constants $\epsilon_0 > 0$, $L > 0$, and n_0 such that, for $0 < \epsilon < \epsilon_0$ and $n \ge n_0$,

$$\mathbb{P}\left\{ -\frac{v^*}{2} - \epsilon \le n^{-3/2}H(\sigma) \le -\frac{v^*}{2} + \epsilon \;\middle|\; \sigma \text{ is locally optimal} \right\} \ge 1 - \exp\left(L\sqrt{n} - \epsilon^2 n \right).$$

The values of the constants are numerically evaluated to be $\alpha^* \approx 0.199$ and $v^*/2 \approx 0.506$. Since the global minimum of $H(\sigma)$ is about $-0.763 \, n^{3/2}$, the typical value a local optimum $-0.506 \, n^{3/2}$ comes fairly close.

Also note that Proposition 1 below implies that α^* is between $1/(2\pi) \approx 0.1591\ldots$ and $2/(3\pi) \approx 0.2122\ldots$.

1.2 Local minima, greedy algorithms and MaxCut

Our problem is related to finding a local optimum of weighted MaxCut on the complete graph, which was recently studied in Etscheid and Röglin [5] and Angel,

Bubeck, Peres, and Wei [2]. Given $S \subset [n]$, we denote the value of the *cut* $(S, [n] \backslash S)$ as

$$\text{Cut}(S, [n] \backslash S) := \sum_{i \in S} \sum_{j \in [n] \backslash S} W_{i,j}.$$

Note that there is a direct correspondence between cuts $(S, [n] \backslash S)$ and spin configurations σ_S given by

$$\sigma_S := (2\mathbb{1}_{i \in S} - 1)_{i=1}^n \ ,$$

$$\text{Cut}(S, [n] \backslash S) = \frac{-H(\sigma_S) + \sum_{1 \leq i < j \leq n} W_{ij}}{2} \ .$$

In particular, what [2] calls locally optimal cuts correspond exactly to our notion of local minimum and Theorem 1 may be formulated in terms of locally optimal cuts. (Note that $n^{-3/2} \sum_{1 \leq i < j \leq n} W_{ij} = O_p(n^{-1/2})$ and therefore this term does not play a significant role in the typical value of a locally optimal cut.)

The papers [5] and [2] study the typical running time simple greedy algorithms take to find locally optimal cuts. Such algorithms start from a given σ and perform a sequence of local "greedy moves"–, that is, single spin flips that decrease energy–, until no more such moves are available. The main result of [2] is that this process ends at a local minimum after a polynomial number of moves. Unfortunately, it is not clear that the *distribution* of the value of this local minimum is similar to the one we study in Theorem 1. In fact, simulations suggest that greedy algorithms find spin configurations with significantly smaller energy than the value of a "typical" local minimum suggested by Theorem 1. This intriguing phenomenon has different possible explanations that would be interesting to reveal.

2   The probability of local optimality

In this section we take the first and crucial step to prove Theorem 1. For any fixed spin configuration $\sigma \in \{-1, +1\}^n$, we establish an integral formula for the probability that σ is locally optimal.

For $i \in [n]$, define

$$Z_i(\sigma) := \frac{H(\sigma^{(i)}) - H(\sigma)}{2} = - \sum_{j \in [n] \backslash i} \sigma_i \sigma_j W_{i,j} \ .$$

Note that

$$\sigma \text{ is a local minimum if and only if } Z_i(\sigma) \geq 0 \text{ for all } i \in [n] \ .$$

Moreover,

$$-H(\sigma) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j \in [n] \setminus i} -\sigma_i \sigma_j W_{i,j} = \frac{\sum_{i=1}^{n} Z_i(\sigma)}{2} \ .$$

Since σ is fixed, we write Z_i instead of $Z_i(\sigma)$ most of the time. A key point in our calculations is that the random vector

$$Z = (Z_1, Z_2, \ldots, Z_n)^T$$

is a multivariate normal vector with zero mean and covariance matrix $C = (C_{i,j})_{n \times n}$ such that $C_{i,i} = n - 1$ for all $i \in [n]$ and $C_{i,j} = 1$ for all $i \neq j$. In other words,

$$C = (n-2)\mathrm{Id}_n + \mathbb{1}_n \mathbb{1}_n^T \ ,$$

where Id_n is the $n \times n$ identity matrix and $\mathbb{1}_n = (1,1,\ldots,1)^T$ is the column vector with 1 in each component.

Clearly, the eigenvalues of C are $2n - 2$ with multiplicity 1 and $n - 2$ with multiplicity $n - 1$, and therefore $\det(C) = (2n-2)(n-2)^{n-1}$.

One may use the Sherman-Morrison formula to invert C and obtain

$$C^{-1} = \frac{1}{n-2}\left(\mathrm{Id}_n - \frac{1}{2n-2}\mathbb{1}_n \mathbb{1}_n^T\right) \ .$$

Hence,

$\mathbb{P}\{\sigma \text{ is locally optimal}\}$

$$= \frac{1}{(2\pi)^{n/2} \det(C)^{1/2}} \int_{[0,\infty)^n} \exp\left(\frac{-x^T C^{-1} x}{2}\right) dx$$

$$= \frac{1}{(2\pi)^{n/2}(2n-2)^{1/2}(n-2)^{(n-1)/2}} \int_{[0,\infty)^n} \exp\left(\frac{-\|x\|_2^2}{2(n-2)} + \frac{\|x\|_1^2}{2(n-2)(2n-2)}\right) dx$$

$$= 2^{-n} \frac{1}{(2\pi)^{n/2}(2n-2)^{1/2}(n-2)^{(n-1)/2}} \int_{\mathbb{R}^n} \exp\left(\frac{-\|x\|_2^2}{2(n-2)} + \frac{\|x\|_1^2}{2(n-2)(2n-2)}\right) dx \ .$$

We may rewrite this as:

$$\mathbb{P}\{\sigma \text{ is locally optimal}\} = 2^{-n} \sqrt{\frac{n-2}{2n-2}} \ \mathbb{E} \exp\left(\frac{\|N\|_1^2}{4(n-1)}\right)$$

where N is a vector of n independent standard normal random variables.

In what follows, we derive some simple upper and lower bounds for the integral above.

Lemma 1. *If N is a vector of independent standard normal random variables, then for all $\lambda > 0$,*

$$\lambda \mathbb{E}\|N\|_1^2 \le \log \mathbb{E}\exp\left(\lambda\|N\|_1^2\right) \le \lambda \mathbb{E}\|N\|_1^2\left(1 + \frac{n\lambda}{(1 - n\lambda)}\right).$$

Proof. The inequality on the left-hand side is obvious from Jensen's inequality. To prove the right-hand side, we use the Gaussian logarithmic Sobolev inequality. In particular, writing $f(x) = \|x\|_1^2$ and $F(\lambda) = \mathbb{E}\exp(\lambda f(N))$, the inequality on page 126 of Boucheron, Lugosi, and Massart [3] asserts that

$$\lambda F'(\lambda) - F(\lambda)\log F(\lambda) \le \frac{\lambda^2}{2}\mathbb{E}\left[e^{\lambda f(N)}\|\nabla f(N)\|^2\right].$$

Since $\|\nabla f(N)\|^2 = 4n\|N\|_1^2$, we obtain the differential inequality

$$\lambda F'(\lambda) - F(\lambda)\log F(\lambda) \le 2n\lambda^2 F'(\lambda).$$

This inequality has the same form as the one at the top of page 191 of [3] with $a = 2n$ and $b = 0$ and Theorem 6.19 implies the result above. ∎

Since

$$\mathbb{E}\|N\|_1^2 = n + n(n-1)\frac{2}{\pi},$$

we get

$$\mathbb{P}\{\sigma \text{ is locally optimal}\} \ge 2^{-n}\sqrt{\frac{n-2}{2n-2}}\exp\left(n/(4(n-1)) + \frac{n}{2\pi}\right)$$

and

$$\mathbb{P}\{\sigma \text{ is locally optimal}\} \le 2^{-n}\sqrt{\frac{n-2}{2n-2}}\exp\left(\left(n/(4(n-1)) + \frac{n}{2\pi}\right)\frac{4n-1}{3n-1}\right).$$

Summarizing, we obtain the following bounds

Proposition 1. *For all spin configurations $\sigma \in \{-1, 1\}^n$,*

$$\frac{1}{2\pi} - \log 2 - O\left(\frac{1}{n}\right) \le \frac{1}{n}\log \mathbb{P}\{\sigma \text{ is locally optimal}\} \le \frac{2}{3\pi} - \log 2 + O\left(\frac{1}{n}\right).$$

In the next section we take a closer look at the integral expression of the probability of local optimality. In fact, we prove that $(1/n)\log\mathbb{P}\{\sigma \text{ is locally optimal}\}$ converges to $\alpha^* - \log 2$ defined in the introduction.

3 The value of local optima

In this section we study, for any fixed $\sigma \in \{-1, +1\}^n$ and $\Delta > 0$, the joint probability

$$\mathbb{P}\Big\{\sigma \text{ is locally optimal, } n^{-3/2}H(\sigma) \leq -\Delta\Big\} \; .$$

We let $Z = (Z_i)_{i=1}^n$ with $Z_i = Z_i(\sigma)$ as in the previous section. Recall from the previous section that

$$\sigma \text{ is a local minimum if and only if } Z_i \geq 0 \text{ for all } i \in [n]$$

and

$$-\frac{H(\sigma)}{n^{3/2}} = \frac{1}{2n^{3/2}} \sum_{i=1}^n Z_i \; .$$

Therefore, we may follow the calculations in the previous section and obtain:

$$\mathbb{P}\Big\{\sigma \text{ is locally optimal}, n^{-3/2}H(\sigma) \leq -\Delta\Big\}$$

$$= \; \mathbb{P}\left\{(\cap_{i=1}^n \{Z_i \geq 0\}) \bigcap \left\{\sum_{i=1}^n Z_i \geq 2\Delta n^{3/2}\right\}\right\}$$

$$= \; \frac{1}{(2\pi)^{n/2}\det(C)^{1/2}} \int\limits_{[0,\infty)^n \cap \{x : \sum_i x_i \geq 2\Delta n^{3/2}\}} \exp\left(\frac{-x^T C^{-1} x}{2}\right) dx$$

$$= \; \frac{1}{(2\pi)^{n/2}(2n-2)^{1/2}(n-2)^{(n-1)/2}} \int\limits_{[0,\infty)^n \cap \{x : \sum_i x_i \geq 2\Delta n^{3/2}\}} \exp\left(\frac{-\|x\|_2^2}{2(n-2)} + \frac{\|x\|_1^2}{2(n-2)(2n-2)}\right) dx$$

$$= \; 2^{-n} \frac{1}{(2\pi)^{n/2}(2n-2)^{1/2}(n-2)^{(n-1)/2}} \int\limits_{\{x : \|x\|_1 \geq 2\Delta n^{3/2}\}} \exp\left(\frac{-\|x\|_2^2}{2(n-2)} + \frac{\|x\|_1^2}{2(n-2)(2n-2)}\right) dx \; .$$

Thus, by a change of variables, we get

$$\mathbb{P}\Big\{\sigma \text{ is locally optimal}, n^{-3/2}H(\sigma) \leq -\Delta\Big\}$$

$$= \; 2^{-n} \sqrt{\frac{n-2}{2n-2}} \, \mathbb{E}\left[\mathbb{1}_{\{\|N\|_1 \geq 2\Delta n^{3/2}/\sqrt{n-2}\}} \exp\left(\frac{\|N\|_1^2}{4(n-1)}\right)\right] \; ,$$

where N is a vector of independent standard normal random variables.

We deduce the following identity for the distribution function of the energy $H(\sigma)$ conditional on σ being locally optimal.

Proposition 2. *For all $\sigma \in \{-1, +1\}^n$,*

$$\mathbb{P}\left\{-\frac{H(\sigma)}{n^{3/2}} \leq \Delta \,\middle|\, \sigma \text{ is locally optimal}\right\} = \frac{\mathbb{E}\left[\mathbb{1}_{\{\|N\|_1 \geq 2\Delta n^{3/2}/\sqrt{n-2}\}} \exp\left(\frac{\|N\|_1^2}{4(n-1)}\right)\right]}{\mathbb{E}\exp\left(\frac{\|N\|_1^2}{4(n-1)}\right)} \; . \qquad (3.1)$$

4 Approximating the integral

In order to establish convergence of the exponent $(1/n)\log\mathbb{P}\{\sigma \text{ is locally optimal}\}$ and also the "typical" value of the energy, we need to understand the behavior of the numerator and the denominator of the equation (3.1).

The main idea is to obtain a Laplace-type approximation to the integral. To get an idea of what to expect, make the approximation

$$\mathbb{E}\left[\exp\left(\frac{\|N\|_1^2}{4(n-1)}\right)\right] \approx \mathbb{E}\left[\exp\left(\frac{\|N\|_1^2}{4n}\right)\right].$$

Observe that $\|N\|_1/n = (1/n)\sum_{i=1}^n |N_i|$ is an average of i.i.d. random variables with expectation $\sqrt{2/\pi}$ and light tails. Therefore, it satisfies a Large Deviations Principle with a rate function $\mu^*(x)$

$$\mathbb{P}\{\|N\|_1 \geq nx\} \approx e^{-\mu^*(x)n}.$$

Readers familiar with Varadhan's lemma (see e.g. [6, page 32]) should expect that, as $n \to \infty$,

$$\frac{1}{n}\log\mathbb{E}\left[\exp\left(\frac{\|N\|_1^2}{4n}\right)\right] = \frac{1}{n}\log\mathbb{E}\left[\exp\left(n\frac{(\|N\|_1/n)^2}{4}\right)\right] \to \sup_v\left(\frac{v^2}{4} - \mu_*(v)\right).$$

In fact, the intuition behind the Varadhan's lemma is that most of the "mass" of the expectation concentrates around $\|N\|_1 \sim v_* n$, where v_* achieves the above supremum. This means that the conditional measure described in Proposition 2 should concentrate around $v^*/2$.

Our calculations confirm this reasoning. The usual statement of Varadhan's lemma does not apply directly because $\|N\|_1^2/4$ is an unbounded function of $\|N\|_1$. Another minor technicality is that the function $\|N\|_1$ is divided by $(n-1)$ instead of n. In what follows we have opted for a self-contained approach to our estimates, which gives quantitative bounds. This section collects the corresponding technical estimates. We finish the proof of Theorem 1 in the next section.

The next lemma is a quantitative version of the large deviations principle (or Cramér's theorem) for $\|N\|_1$.

Lemma 2. *For $x \geq \sqrt{2/\pi}$, define $\mu^*(x)$ as in the introduction. Let $N = (N_1,\ldots,N_n)$ be a vector of independent standard normal coordinates. Then*

$$\mathbb{P}\{\|N\|_1 \geq nx\} = e^{-(\mu^*(x)+r_n(x))n}$$

with

$$0 \leq r_n(x) \leq \kappa\left(\frac{x - \sqrt{2/\pi}}{\sqrt{n}} + \frac{1}{n}\right)$$

for some $\kappa > 0$ independent of x and n. Moreover, μ^* is smooth and $\mu^*(\sqrt{2/\pi}) = 0$.

Proof. This follows directly from Lemmas 5, 6 and 7 in Section 6.1. ■

We use this lemma to estimate expectations of the form

$$\mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1\leq an\}}\right] \quad\text{and}\quad \mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1\geq bn\}}\right].$$

The function R_c defined as

$$R_c(x) := \frac{cx^2}{2} - \mu^*(x) \quad\text{for } x \geq \sqrt{2/\pi}$$

appears naturally in our estimates.

Lemma 3. *For $a \geq \sqrt{2/\pi}$, $c \geq 0$*

$$\mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1\leq an\}}\right] = (I) + (II),$$

where

$$1 \leq (I) \leq \exp\left(nR_c(\sqrt{2/\pi})\right)$$

and

$$(II) = cn\int_{\sqrt{2/\pi}}^{a} x\exp(n\,(R_c(x) - r_n(x)))\,dx$$

with $r_n(x)$ is as in Lemma 2. For $b \geq \sqrt{2/\pi}$,

$$\mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1\geq bn\}}\right] = \exp\{n\,(R_c(b) - r_n(b))\}$$

$$+cn\int_{b}^{+\infty} x\exp\{n\,(R_c(x) - r_n(x))\}\,dx.$$

Proof. Let $\phi_{c,n}(x) = e^{cnx^2/2}$. Note that

$$\mathbb{1}_{\{\|N\|_1\leq an\}}\exp\left(\frac{cn\|N\|_1^2}{2n}\right) = \phi_{c,n}\left(\frac{\|N\|_1}{n}\right)\mathbb{1}_{\{\frac{\|N\|_1}{n}\leq a\}}.$$

We may compute the expectation of this expression as follows.

$$\mathbb{E}\left[\mathbb{1}_{\{\|N\|_1\leq an\}}\exp\left(\frac{cn\|N\|_1^2}{2n}\right)\right] = 1 + \int_{0}^{a} \phi'_{c,n}(x)\mathbb{P}\left\{\frac{\|N\|_1}{n}\geq x\right\}dx$$

$$= 1 + cn\int_{0}^{a}\exp\left(\frac{cnx^2}{2}\right)\mathbb{P}\left\{\frac{\|N\|_1}{n}\geq x\right\}dx.$$

We split the above integral in two parts.

$$(I) \quad = \quad 1 + cn \int_0^{\sqrt{2/\pi}} x \exp\left(\frac{cn\,x^2}{2}\right) \mathbb{P}\left\{\frac{\|N\|_1}{n} \geq x\right\} dx$$

$$(II) \quad = \quad cn \int_{\sqrt{2/\pi}}^{a} x \exp\left(\frac{cn\,x^2}{2}\right) \mathbb{P}\left\{\frac{\|N\|_1}{n} \geq x\right\} dx \,.$$

For part (I), we bound the probability in the integral by 1, and obtain:

$$1 \leq (I) \leq 1 + cn \int_0^{\sqrt{2/\pi}} x \exp\left(\frac{cn\,x^2}{2}\right) dx = \exp\left(\frac{cn\,x^2}{2}\right)\Bigg|_{x=\sqrt{\frac{2}{\pi}}} = \exp\left(n\,R_c(\sqrt{2/\pi})\right)$$

because $\mu^*(\sqrt{2/\pi}) = 0$. Term (II) may be evaluated using the estimate from Lemma 2.

$$(II) = cn \int_{\sqrt{2/\pi}}^{a} x \exp\left(\frac{cn\,x^2}{2} - n\,\mu^*(x) - n\,r_n(x)\right) dx \,,$$

which has the desired form because

$$\frac{cn\,x^2}{2} - n\,\mu^*(x) = n\,R_c(x) \,.$$

Similarly,

$$\mathbb{1}_{\{\|N\|_1 \geq b\,n\}} \exp\left(\frac{cn\,\|N\|_1^2}{2n}\right) = \phi_{c,n}\left(\frac{\|N\|_1}{n}\right) \mathbb{1}_{\{\frac{\|N\|_1}{n} \geq b\}} \,,$$

and we finish the proof via the identity

$$\mathbb{E}\left[\mathbb{1}_{\{\|N\|_1 \geq b\,n\}} \exp\left(\frac{cn\,\|N\|_1^2}{2n}\right)\right] \quad = \quad \phi_{c,n}(b)\,\mathbb{P}\left\{\frac{\|N\|_1}{n} \geq b\right\}$$

$$+ \int_b^{+\infty} \phi_{c,n}'(x)\,\mathbb{P}\left\{\frac{\|N\|_1}{n} \geq x\right\} dx$$

and using the bounds in Lemma 2 (which are valid for all $x \geq b \geq \sqrt{2/\pi}$). ∎

5   Proof of Theorem 1

The previous section shows that, in order to estimate the expectations in Lemma 3, we need to understand the function R_c. The case of interest for us is when $c = n/(2(n-1))$, which is when we recover the expectations in (3.1). Since c varies with n, we consider instead

$$R(x) = R_{\frac{1}{2}}(x) := \frac{x^2}{4} - \mu^*(x) \quad \text{for } x \geq \sqrt{2/\pi} \,. \tag{5.1}$$

and note that

$$R(x) \leq R_c(x) \leq R(x) + (2c - 1)\frac{x^2}{4} \ . \tag{5.2}$$

The next lemma contains some information on $R(x)$.

Lemma 4. *Let $x \geq \sqrt{2/\pi}$. Define R as in equation (5.1) and μ^* as in Lemma 2. Then there exists a unique $x = v^* > \sqrt{2/\pi}$ that maximizes $R(x)$ over $x \geq \sqrt{2/\pi}$. Letting $\alpha^* := R(v^*)$ denote the value of the maximum, for any $x \geq \sqrt{2/\pi}$, there exists $\theta(x) \in [1/4, 10]$ with*

$$R(x) - \alpha^* = -\theta(x)(x - v^*)^2 \ .$$

Proof. See subsection 6.2. ∎

We can now obtain good upper and lower estimates on the integral expressions in Lemma 3 and finish the proof of the main theorem.

Proof. [of Theorem 1] In this proof we assume $n \geq 100$ for simplicity. We use the notation L to denote the value of a positive constant independent of n whose value may change from line to line. Finally, we set

$$c = c_n := \frac{n}{2(n-1)} = 1 + \frac{1}{2(n-1)} \ .$$

Lemma 4 and (5.2) imply that for all $x \geq \sqrt{\frac{2}{\pi}}$,

$$R_c(x) - \alpha^* \in \left[-10(x - v^*)^2, -\frac{1}{6}(x - v^*)^2 + \frac{x^2}{(n-1)} \right] \ . \tag{5.3}$$

We now apply this to estimate expectations to the left of v^*. That is, we consider, for $a \in \left[\sqrt{\frac{2}{\pi}}, v^* \right]$,

$$\mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n} \right) \mathbb{1}_{\{\|N\|_1 \leq an\}} \right] \ .$$

In this range $|a - v^*|$ is uniformly bounded, so $x^2 \leq L$ and

$$0 \leq r_n(x) \leq \frac{L}{\sqrt{n}} \quad \text{for all } \sqrt{\frac{2}{\pi}} \leq x \leq v^* \ .$$

10

Combining Lemma 3 with $c \le 1$ and (5.3), we obtain

$$\frac{\mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1 \le an\}}\right]}{\exp(n\alpha^*)} \le \exp(n(R_c(\sqrt{2/\pi}) - \alpha^*))$$

$$+cn\int_{\sqrt{2/\pi}}^{a} x \exp(n(R_c(x) - \alpha^*))\,dx$$

$$\le \exp\left(L - \frac{(v^* - \sqrt{2/\pi})^2}{4}n\right)$$

$$+n\int_{\sqrt{2/\pi}}^{a} x \exp\left(L + n\frac{(v^* - x)^2}{6}\right)dx$$

$$\le L(1 + cn)\exp\left(L - \frac{(a - v^*)^2 n}{4}\right)$$

$$\le \exp\left(L\log n - \frac{(a - v^*)^2 n}{4}\right).$$

At the same time,

$$\frac{\mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1 \le v^*n\}}\right]}{\exp(n\alpha^*)} \ge \exp\left(-L\sqrt{n}\right)\int_{v^*-\frac{1}{n}}^{v^*} x\exp(n(R_c(x) - \alpha^*))\,dx$$

$$\ge \frac{1}{n}\left(v^* - \frac{1}{n}\right)\frac{\exp\left(-L\sqrt{n} - 10(1/n)^2\right)}{n}$$

$$\ge \exp(-L\sqrt{n}).$$

For bounding the expectation for $b \ge v^*$, we cannot simply use $x^2 \le L$ and $r_n(x) \le L/\sqrt{n}$. However, note that

$$-\frac{1}{6}(x - v^*)^2 + \frac{x^2}{4(n-1)} \le \begin{cases} -\frac{1}{5}(x - v^*)^2 + \frac{L}{\sqrt{n}} & \text{for } x \le (n-1)^{1/4}; \\ -\frac{1}{6}(x - v^*)^2 + \frac{2(x-v^*)^2 + 2(v^*)^2}{(n-1)} \le -\frac{1}{5}(x - v^*)^2 + \frac{L}{n} & \text{for larger } x. \end{cases}$$

Also, recalling the expression for r_n in Lemma 2,

$$0 \le r_n(x) \le \kappa\left(\frac{x - \sqrt{2/\pi}}{\sqrt{n}} + \frac{1}{n}\right) \le \frac{L}{\sqrt{n}} + \frac{L(x - v^*)}{\sqrt{n}}.$$

This allows us to obtain, for $b \le v^* + \epsilon_0$,

$$\frac{\mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1 \ge bn\}}\right]}{\exp(n\alpha^*)} \le \left(L\sqrt{n} - \frac{(b - v^*)^2 n}{4}\right);$$

$$\frac{\mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1 \ge v^*n\}}\right]}{\exp(n\alpha^*)} \ge \exp(-L\sqrt{n}).$$

This leads to our main results. Indeed, if we apply the above bounds with $a = b = v^*$, we obtain that, as $n \to \infty$,

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\right] &= \mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1 \leq v^* n\}}\right] \\
&\quad + \mathbb{E}\left[\exp\left(\frac{c\|N\|_1^2}{2n}\right)\mathbb{1}_{\{\|N\|_1 \geq v^* n\}}\right] \\
&= \exp\left(n\alpha^* + \delta_n\right),
\end{aligned}
$$

where $\delta_n \in \left[-L\sqrt{n}, L\sqrt{n}\right]$ for a positive constant L. This implies the first statement in the theorem via Proposition 1.

Secondly, we apply Proposition 2 and obtain

$$
\begin{aligned}
\mathbb{P}\left\{-H(\sigma) \leq -\frac{v^*}{2} - \epsilon \mid \sigma \text{ local optimum}\right\} &\leq \frac{\mathbb{E}\left[\mathbb{1}_{\{\|N\|_1 \geq bn\}}\exp\left(\frac{\|N\|_1^2}{4(n-1)}\right)\right]}{\mathbb{E}\left[\exp\left(\frac{\|N\|_1^2}{4(n-1)}\right)\right]} \\
&\qquad \left(\text{with } b = (v^* + 2\epsilon)\sqrt{\frac{n}{n-1}}\right) \\
&= \exp\left(-L\sqrt{n} - \epsilon^2 n\right),
\end{aligned}
$$

and, for ϵ small enough, so that the value of a below is at most $\sqrt{2/\pi}$,

$$
\begin{aligned}
\mathbb{P}\left\{-H(\sigma) \geq -\frac{v^*}{2} + \epsilon \mid \sigma \text{ local optimum}\right\} &\leq \frac{\mathbb{E}\left[\mathbb{1}_{\{\|N\|_1 \leq an\}}\exp\left(\frac{\|N\|_1^2}{4(n-1)}\right)\right]}{\mathbb{E}\left[\exp\left(\frac{\|N\|_1^2}{4(n-1)}\right)\right]} \\
&\qquad \left(\text{with } a = (v^* - 2\epsilon)\sqrt{\frac{n}{n-1}}\right) \\
&= \exp\left(-L\sqrt{n} - \epsilon^2 n\right).
\end{aligned}
$$

$\blacksquare$

6 Auxiliary results

6.1 Lemmas on large deviations of $\|N\|_1$

The goal of this section is to prove a series of lemmas that together imply Lemma 2. We first find an expression for the Laplace transform of the absolute value of a standard normal random variable:

Lemma 5. *Let N be a standard normal random variable. For all $\lambda > 0$,*

$$\mathbb{E}e^{\lambda|N|} = e^{\lambda^2/2 + \phi(\lambda)},$$

where $\phi(\lambda) = \log(2\Phi(\lambda))$, with $\Phi(\lambda) = \mathbb{P}\{N \le \lambda\}$.

Proof.

$$
\begin{aligned}
\mathbb{E}e^{\lambda|N|} &= \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{\lambda x - x^2/2} dx \\
&= 2e^{\lambda^2/2} \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{(x-\lambda)^2/2} dx \\
&= 2e^{\lambda^2/2} \mathbb{P}\{N > -\lambda\}.
\end{aligned}
$$

■

We need to compute the large deviations rate function for $\sum_{i=1}^n |N_i|$, with N_i independent standard normal random variables. As usual, this is given by the Fenchel-Légendre transform of $\log \mathbb{E}e^{\lambda|N|}$:

$$\mu^*(x) := \sup_{\lambda \ge 0} \left(\lambda x - \log \mathbb{E}e^{\lambda|N|} \right).$$

The next lemma collects technical facts on μ^* and the value $\lambda = \lambda_*$ that achieves the minimum.

Lemma 6. *For each $x \ge \sqrt{2/\pi}$, there exists a unique $\lambda = \lambda_*(x) \ge 0$ such that*

$$\lambda + \phi'(\lambda) = x.$$

Defining

$$\mu^*(x) := \lambda_*(x)x - \frac{\lambda_*(x)^2}{2} - \phi(\lambda_*(x)),$$

for $x \ge \sqrt{\frac{2}{\pi}}$, we have that, for each x in this range, $\mu^*(x)$ is the global maximum of

$$\lambda x - \frac{\lambda^2}{2} - \phi(\lambda)$$

over $\lambda \in [0, \infty)$ which is uniquely achieved at $\lambda = \lambda_*(x)$. We also have the following inequalities.

1. Strict concavity. *For each $\lambda \ge 0$, $x \ge \sqrt{2/\pi}$,*

$$\mu^*(x) - (\lambda x - \log \mathbb{E}e^{\lambda|N|}) \in \left[\frac{(\lambda - \lambda_*(x))^2}{40}, \frac{(\lambda - \lambda_*(x))^2}{2} \right]. \tag{6.1}$$

2. Derivative bounds for λ_*:

$$1 \leq \lambda_*'(x) \leq 20 \,. \tag{6.2}$$

Proof. By the previous Lemma,

$$\lambda + \phi'(\lambda) = \frac{d}{d\lambda} \log \mathbb{E} e^{\lambda|N|} \,,$$

which is a smooth function because $|N|$ has a Gaussian-type tail. Using this "lightness of the tail", one can differentiate under the expectation and obtain

$$\phi'(0) = \frac{d}{d\lambda} \log \mathbb{E} \, e^{\lambda|N|} \Big|_{\lambda=0} = \mathbb{E}|N| = \sqrt{\frac{2}{\pi}} \,.$$

Lemma 9 below implies that

$$-0.95 \leq \phi''(\lambda) \leq 0 \,.$$

Therefore, for all $\lambda \geq 0$,

$$\frac{d}{d\lambda} (\lambda + \phi'(\lambda)) \in [0.05, 1] \,. \tag{6.3}$$

In particular, $\lambda + \phi'(\lambda)$ is an increasing function that is equal to $\sqrt{2/\pi}$ at $\lambda = 0$ and diverges when $\lambda \nearrow +\infty$. It follows that for all $x \geq \sqrt{2/\pi}$ there exists a unique $\lambda = \lambda_*(x)$ with $\lambda_*(x) + \phi'(\lambda_*(x)) = x$, and $\lambda_*(\sqrt{2/\pi}) = 0$. The implicit function theorem guarantees that λ_* is smooth over $[\sqrt{2/\pi}, +\infty)$ and

$$\lambda_*'(x) = \frac{1}{\frac{d}{d\lambda} (\lambda + \phi'(\lambda)) \big|_{\lambda=\lambda_*(x)}} \in [1, 20] \,. \tag{6.4}$$

Equation (6.3) above shows that

$$\lambda x - \frac{\lambda^2}{2} - \phi(\lambda) = \lambda x - \log \mathbb{E} e^{\lambda|N|}$$

is a strictly concave function of λ with second derivative

$$-1 \leq -\frac{d^2}{(d\lambda)^2} \left(\frac{\lambda^2}{2} + \phi(\lambda) \right) \leq -\frac{1}{20} \,.$$

Thus $\lambda_*(x)$, which is a critical point for this function, is the unique global maximum of $\lambda x - \log \mathbb{E} e^{\lambda|N|}$. The value of the function at that point is precisely $\mu^*(x)$.

Let us now prove the estimates in the lemma. The strict concavity property in (6.1) follows from expanding

$$\lambda x - \log \mathbb{E} e^{\lambda|N|}$$

around the critical point $\lambda = \lambda_*(x)$ and applying a second-order Taylor expansion:

$$
\begin{aligned}
\lambda x - \log \mathbb{E}\, e^{\lambda |N|} - \mu^*(x) \;=\;& \frac{d}{d\lambda}\left(\lambda x - \log \mathbb{E}\, e^{\lambda |N|}\right)\Big|_{\lambda = \lambda_*(x)} \\
&+ \frac{1}{2}\frac{d^2}{(d\lambda)^2}\left(\lambda x - \log \mathbb{E}\, e^{\lambda |N|}\right)\Big|_{\lambda = \tilde{\lambda}} (\lambda - \lambda_*(x))^2 \\
&\text{with } \tilde{\lambda} = (1-\alpha)\lambda_*(x) + \alpha\lambda, \text{ for some } \alpha \in [0,1] \,,
\end{aligned}
$$

noting that the first derivative is 0 and the second one is between -1 and $-1/20$. Finally, the derivative bound in item 2 is proven in (6.4). ∎

Lemma 7. *Let $N = (N_1, \ldots, N_n)$ be a vector of n independent standard normal random variables. Let $x \geq \sqrt{2/\pi}$. Define $\mu^*(x)$ as in Lemma 6. Then for all $n \geq 1$,*

$$
\frac{1}{n}\log \mathbb{P}\{\|N\|_1 \geq nx\} = -\mu^*(x) - r_n(x) \,,
$$

where

$$
0 \leq r_n(x) \leq \kappa \left(\frac{x - \sqrt{2/\pi}}{\sqrt{n}} + \frac{1}{n} \right)
$$

for some universal $\kappa > 0$ that is independent of $x \geq \sqrt{2/\pi}$ and $n \geq 1$.

Proof. For any $\lambda > 0$, the usual Cramér-Chernoff trick may be combined with Lemma 6 to obtain

$$
\frac{1}{n}\log \mathbb{P}\{\|N\|_1 \geq nx\} \leq \inf_{\lambda \geq 0}\left(\log \mathbb{E}\, e^{\lambda |N|} - \lambda x\right) = -\mu^*(x) \,.
$$

It remains to give a lower bound for this probability. In order to get a non-asymptotic bound, we use the following lemma that appears in the fourth edition of the book of Alon and Spencer [1, Theorem A.2.1].

Lemma 8. *Let $u, \lambda, \epsilon > 0$ such that $\lambda > \epsilon$. Let X be a random variable such that the moment generating function $\mathbb{E}\, e^{cX}$ exists for $c \leq \lambda + \epsilon$. For any $a \in \mathbb{R}$, define $g_a(c) = e^{-ac}\mathbb{E}\, e^{cX}$. Then*

$$
\mathbb{P}\{X \geq a - u\} \geq e^{-\lambda u}\left(g_a(\lambda) - e^{-\epsilon u}\left(g_a(\lambda + \epsilon) + g_a(\lambda - \epsilon)\right)\right) \,.
$$

We apply Lemma 8 to the random variable $X = \|N\|_1$ with $\lambda = \lambda_*(a/n)$ and a, u, ε to be chosen below. In the notation of Lemma 8, for each $\lambda \geq 0$,

$$
g_a(\lambda) = \exp(-n\mu_a(\lambda)) \quad \text{where} \quad \mu_a(\lambda) = \left(\lambda(a/n) - \log \mathbb{E}\, e^{\lambda |N|}\right) \,.
$$

Using Lemma 6 to bound this expression, we obtain from (6.1) that

$$\frac{g_a(\lambda_*(a/n)+\epsilon)}{g_a(\lambda_*(a/n))} \leq e^{n\epsilon^2/2} \quad \text{and} \quad \frac{g_a(\lambda_*(a/n)-\epsilon)}{g_a(\lambda_*(a/n))} \leq e^{n\epsilon^2/2} .$$

Moreover, $g_a(\lambda_*(a/n)) = e^{-n\mu^*(a/n)}$. So

$$\mathbb{P}\{\|N\|_1 \geq a - u\} \geq e^{-\lambda_*(a/n)u} e^{-n\mu^*(a/n)} \left(1 - 2e^{-\epsilon u + \frac{\epsilon^2 n}{2}} \right) .$$

We now choose $\epsilon = \sqrt{2/n}$ and $u = \epsilon n/2 + 1/\epsilon = \sqrt{2n}$ to obtain

$$\mathbb{P}\{\|N\|_1 \geq a - u\} \geq e^{-\lambda_*(a/n)\sqrt{2n}} e^{-n\mu^*(a/n)} \left(1 - \frac{2}{e} \right) .$$

Letting $a = nx + \sqrt{2n} = n(x + \epsilon)$, we have that

$$\mathbb{P}\{\|N\|_1 \geq nx\} = e^{-\lambda_*(x+\epsilon)\sqrt{2n}} e^{-n\mu^*(x+\epsilon)} \left(1 - \frac{2}{e} \right) .$$

Recall from Lemma 6 that $\lambda_*'(y) \leq 20(y - \sqrt{2/\pi})$ and $y - \sqrt{2/\pi} \leq (\mu^*)'(y) \leq 20(y - \sqrt{2/\pi})$ for all $y \geq \sqrt{2/\pi}$. Thus,

$$\lambda_*(x+\epsilon)\sqrt{2n} \leq 20(x+\epsilon - \sqrt{2/\pi})\sqrt{2n}$$

and

$$\mu^*(x+\epsilon) \leq \mu^*(x) + 20\,\epsilon\,(x+\epsilon - \sqrt{2/\pi}) .$$

Recalling $\epsilon = \sqrt{2/n}$, we may plug theses estimates back in the lower bound for our probability and obtain the theorem. ∎

6.2   Estimates on the optimization problem

In this section we prove Lemma 4.

Proof. [of Lemma 4] We use Lemma 6 several times in the proof. In particular, the properties of μ^* and $\lambda_* = (\mu^*)'$ are used several times.

We first argue that $x \mapsto R(x)$ is a strictly concave function of $x \geq \sqrt{2/\pi}$. To see this, we use Lemma 6 to obtain

$$R''(x) = \frac{1}{2} - \lambda_*'(x) .$$

By the same lemma, we know $1 \leq \lambda_*'(x) \leq 20$. So

$$-20 \leq R_c''(x) \leq -\frac{1}{2} \quad \text{for all } x \geq \sqrt{\frac{2}{\pi}} .$$

We now argue that $R(x)$ is maximized at some $x = v^* > \sqrt{2/\pi}$. To see this, notice that $\lambda_*(\sqrt{2/\pi}) = 0$, so the derivative of R at $x = \sqrt{2/\pi}$ satisfies

$$R'(\sqrt{2/\pi}) = \frac{1}{2}\sqrt{2/\pi} - \lambda_*(\sqrt{2/\pi}) > 0 .$$

We conclude that R is increasing in an interval to the right of $\sqrt{2/\pi}$. At the same time, the second derivative of R in x is at most -1, so there exists a $s_+ \gg \sqrt{2/\pi}$ such that $R'(x) \le 0$ for $x \ge s_+$. So the maximum of R in x must be achieved at a point $v^* \in (\sqrt{2/\pi}, s_+]$. In particular, v^* is a critical point of R, that is, $R'(v^*) = 0$.

Now consider

$$\alpha^* := R(v^*) = \min_{x \ge \sqrt{2/\pi}} R(x) .$$

By Taylor expansion, if $x \ge \sqrt{2/\pi}$,

$$R(x) = \alpha^* + R'(v^*)(x - v^*) + \frac{R''(v^* + \alpha\,(x - v^*))}{2}(x - v^*)^2$$

for some $0 \le \alpha \le 1$. The theorem follows because $R'(v^*) = 0$ and $R'' \in [-20, -1/2]$. ∎

A One more technical estimate

Lemma 9. *Let $f(\lambda) = (2\pi)^{-1/2}e^{-\lambda^2/2}$ be the standard normal density let $\Phi(\lambda) = \int_{-\infty}^{\lambda} f(x)dx$ be the corresponding cumulative distribution function. Then for all $\lambda \ge 0$,*

$$\frac{f'(\lambda)}{\Phi(\lambda)} - \frac{f(\lambda)^2}{\Phi(\lambda)^2} > -0.95 .$$

Proof. Note that $f'(\lambda) = -\lambda f(\lambda)$, so we need only prove

$$\sup_{\lambda \ge 0} \frac{f(\lambda)}{\Phi(\lambda)}\left(\lambda + \frac{f(\lambda)}{\Phi(\lambda)}\right) < 0.95 .$$

We combine three inequalities, considering three ranges of the value of λ, given by $[0, \lambda_1)$, $[\lambda_1, \lambda_2]$, and (λ_2, ∞), where

$$\lambda_1 = \frac{0.95 - \frac{2}{\pi}}{\sqrt{2/\pi}} \approx 0.3927\ldots \quad \text{and} \quad \lambda_2 = \sqrt{\log\frac{2/\pi}{0.95 - \sqrt{2/(\pi e)}}} \approx 0.5584\ldots\, .$$

First, note that $f(\lambda)/\Phi(\lambda) \le \sqrt{2/\pi}$ since f/Φ is a decreasing function. Thus,

$$\frac{f(\lambda)}{\Phi(\lambda)}\left(\lambda + \frac{f(\lambda)}{\Phi(\lambda)}\right) \le \frac{2}{\pi} + \lambda\sqrt{\frac{2}{\pi}} < 0.95 \quad \text{for } \lambda \in [0, \lambda_1).$$

Second, $\lambda e^{-\lambda^2/2} \leq 1/\sqrt{e}$, so

$$\frac{f(\lambda)}{\Phi(\lambda)}\left(\lambda + \frac{f(\lambda)}{\Phi(\lambda)}\right) \leq 2\lambda f(\lambda) + 2f(\lambda)^2 \leq \sqrt{\frac{2}{\pi e}} + \frac{2}{\pi}e^{-\lambda^2} < 0.95 \quad \text{for } \lambda \in (\lambda_2, \infty).$$

Finally, since $\lambda e^{-\lambda^2/2}$ is increasing and $e^{-\lambda^2/2}$ is decreasing on $[\lambda_1, \lambda_2]$, on this interval we have

$$\frac{f(\lambda)}{\Phi(\lambda)}\left(\lambda + \frac{f(\lambda)}{\Phi(\lambda)}\right) \leq 2\lambda_2 f(\lambda_2) + 4f^2(\lambda_1) \approx 0.92685\ldots < 0.95 \,.$$

∎

References

[1] N. Alon and J.H. Spencer. *The Probabilistic Method*. Wiley, New York, 2016. 4th edition.

[2] O. Angel, S. Bubeck, Y. Peres, and F. Wei. Local max-cut in smoothed polynomial time. *arXiv preprint arXiv:1610.04807*, 2016.

[3] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford University Press, 2013.

[4] A. Crisanti and T. Rizzo. Analysis of the-replica symmetry breaking solution of the sherrington-kirkpatrick model. *Physical Review E*, 65(4):046137, 2002.

[5] M. Etscheid and H. Röglin. Smoothed analysis of local search for the maximum-cut problem. *ACM Transactions on Algorithms (TALG)*, 13(2):25, 2017.

[6] Frank Den Hollander. *Large Deviations*. Fields Institute Monographs. American Mathematical Society, 2000.

[7] F. Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Communications in Mathematical Physics*, 233(1):1–12, 2003.

[8] D. Panchenko. *The Sherrington-Kirkpatrick model*. Springer Science & Business Media, 2013.

[9] G. Parisi. A sequence of approximated solutions to the sk model for spin glasses. *Journal of Physics A: Mathematical and General*, 13(4):L115, 1980.

[10] D. Sherrington and S. Kirkpatrick. Solvable model of a spin-glass. *Physical Review Letters*, 35(26):1792, 1975.

[11] M. Talagrand. The Parisi formula. *Annals of Mathematics*, 163:221–263, 2006.