Probab. Theory Relat. Fields 119, 187–212 (2001)
Digital Object Identifier (DOI) 10.1007/s004400000099

Michel Talagrand

# The high temperature case for the random $K$-sat problem

Received: 21 April 1998 / Revised version: 24 April 2000 /
Published online: 21 December 2000 – 

**Abstract.** We give a completely rigorous proof that the replica-symmetric solution holds at high enough temperature for the random $K$-sat problem. The most notable feature of this problem is that the order parameter of the system is a function and not a number.

## 1. Introduction

This paper is a step in the author's program to obtain rigorous results about disordered systems related to the theory of spin glasses. This aspect of the paper, and how it relates to the author's previous work will be discussed briefly later. (See [T4] for a more detailed survey.) The specific problem we will study is related to the stochastic version of the famous $K$-sat problem of computer science, on which our results arguably shed some light. We will not formulate the random $K$-sat problem in its usual setting, but rather an equivalent version more suitable to our needs. The reader familiar with the $K$-sat problem will immediately recognize the problem; the reader who is not will only gain by considering directly the aspect of "random geometry" that is relevant here. As we use the notation $K$ for other purposes, we will in fact consider the $p$-sat problem, where $p \geq 2$ is an integer fixed once and for all. We consider the set $\Sigma_N = \{-1, 1\}^N$, and right away point out that we are interested in the case $N$ large.

Let us first fix some notation, that will remain in force throughout the paper. We denote by $[N]^p$ the collection of subsets of $N$ of cardinal $p$. We write, for $J$ in $[N]^p$,

$$J = \{i(J, 1), \cdots, i(J, p)\} \tag{1.1}$$

where $i(J, 1) < \cdots < i(J, p)$.

Consider, for $J \in [N]^p$ and $q \leq p$, random variables $\xi_{J,q}$, $q \leq p$ all independent, with

$$P(\xi_{J,q} = 1) = P(\xi_{J,q} = -1) = \frac{1}{2}.$$

M. Talagrand*: Equipe d'Analyse, Tour 46, ESA au CNRS no. 7064, Université Paris VI, 4, place Jussieu, 75230 Paris Cedex 05, France, (e-mail: mit@ccr.jussieu.fr) and Department of Mathematics, The Ohio State University, 231 West 18th Avenue, Columbus, OH 43210-1174, USA. e-mail: talagran@math.ohio-state.edu

* Work partially supported by an NSF grant.

*Mathematics Subject Classification (2000):* Primary 82D30; Secondary 60G99

Thus given $J$, we can consider the following random subset $A_J$ of $\Sigma_N$:

$$A_J = \{\boldsymbol{\sigma} \in \Sigma_N;\ \exists q \le p,\ \sigma_{i(J,q)} \ne \xi_{J,q}\}.$$

Suppose now that we choose $M$ sets $J_1, \cdots, J_M$ at random, independently and uniformly on $[N]^p$. The object of interest is the random set

$$A = \bigcap_{k \le M} A_{J_k}.$$

We would like to know whether $A$ is typically empty or not; and more generally, what is the typical size of $A$, which can conveniently be measured by the median of

$$\frac{1}{N} \log(2^{-N} \mathrm{card} A). \tag{1.2}$$

We should observe the trivial relation

$$E\mathrm{card} A = 2^N (1 - 2^{-p})^M$$

so that if $M \log(1 - 2^{-p}) > N \log 2$, then $E\mathrm{card} A < 1$ and $A$ is typically empty. Thus, the range of interest is when $M$ is of order $N$; the parameter $\alpha = M/N$ is essential. Leaving $M/N$ fixed, we will study the problem in the limit as $N \to \infty$.

The study of (1.2) is very difficult. Statistical mechanics offers a standard method (used in the present case in [M-Z]) to introduce an easier problem. Rather than studying directly the points that belong to all sets $A_{J_k}$ we count to how many sets $A_{J_k}$ a given $\boldsymbol{\sigma}$ belongs, by setting

$$H_N(\boldsymbol{\sigma}) = -\sum_{k \le M} 1_{A_{J_k}}(\boldsymbol{\sigma}). \tag{1.3}$$

(The purpose of the minus sign is simply to follow the conventions of physics). We then introduce a number $\beta$ (that physically represents an inverse temperature) and which purpose is to weigh how much we will favor the configurations $\boldsymbol{\sigma}$ for which $-H_N(\boldsymbol{\sigma})$ is large. That is, we introduce Gibbs' measure

$$G_N(\{\boldsymbol{\sigma}\}) = Z_N^{-1} \exp -\beta H_N(\boldsymbol{\sigma}) \tag{1.4}$$

where $Z_N$ is the normalization factor

$$Z_N = \sum_{\boldsymbol{\sigma}} \exp -\beta H_N(\boldsymbol{\sigma}). \tag{1.5}$$

We then study the random probability $G_N$, and the corresponding expected "free energy" (per site)

$$F_N(\alpha, \beta) = \frac{1}{N} E \log(2^{-N} Z_N). \tag{1.6}$$

from which the quantity (1.2) can be recovered as

$$\lim_{\beta \to \infty} F_N(\alpha, \beta) + \beta\alpha. \tag{1.7}$$

(There, as usual, $\alpha = M/N$).

The physical approach is to find a formula for $\lim_{N\to\infty} F_N(\alpha,\beta)$, and then hope, as $\beta \to \infty$, to deduce information about (1.1) by taking $\lim_{\beta\to\infty} \beta^{-1} \lim_{N\to\infty} F_N(\alpha,\beta)$. The unjustified interversion of the limit this represents compared to (1.7) is no special reason to worry, since the limit of $F_N(\alpha,\beta)$ is in any case found using the so called “replica method” that involves a number of mathematically rather unjustified assumptions. This is all the more so in the present case, where the “order parameter” of the system (that is, the quantity that specifies it) is a function (or rather, a probability distribution) instead of a number. In the present paper we give a complete and rigorous proof that given $p, \alpha$, if $\beta$ is small enough, the predictions of the replica method are correct. Since our real purpose, is, beyond any specific case, to understand the powerful forces at work here, we will study a more general model. The main estimates for this more general model are harder than in the case of the Hamiltonian (1.3), and being able to perform them represents at least a technical progress. Our basic object is a bounded function

$$f = [0,1] \times \{-1,1\}^p \to \mathbb{R}.$$

The purpose of the first variable is to introduce randomness. If $X$ is a random variable uniformly distributed over $[0,1]$, $f(X,\cdot)$ is a random function on $\{-1,1\}^p$. We consider i.i.d. r.v. $(X_J)$ and i.i.d. r.v. $(\eta_J)$ for $J$ in $[N]^p$. We assume $X_J$ uniform on $[0,1]$; we assume $\eta_J \in \{0,1\}$,

$$P(\eta_J = 1) = \gamma N^{1-p} \tag{1.8}$$

and we consider the random Hamiltonian

$$H_N(\boldsymbol{\sigma}) = \sum_J \eta_J f(X_J, \sigma_{i(J,1)}, \cdots, \sigma_{i(J,p)}). \tag{1.9}$$

The number of terms occurring in (1.9) is $\sum_J \eta_J$, a r.v. sharply concentrated around its mean

$$\gamma N^{1-p} \binom{N}{p} \simeq N \frac{\gamma}{p!}$$

so there are (about) $\alpha N$ terms, for $\alpha = \gamma/p!$. If we wanted exactly $M = \lfloor \alpha N \rfloor$ terms, we could, rather than (1.9), consider the Hamiltonian

$$H_N(\boldsymbol{\sigma}) = \sum_{k\le M} f(X_{J_k}, \sigma_{i(J_k,1)}, \cdots, \sigma_{i(J_k,p)}) \tag{1.10}$$

where $\{J_1, \cdots, J_M\}$ is chosen uniformly among all subsets of $[N]^p$ of cardinal $M$. To simplify some technical detail, we will study the case of (1.9) rather than (1.10); the small extra work to handle (1.10) is left to the reader.

An important feature of mean field models is the symmetry between sites. To ensure this symmetry we will require that

$$f(x, \sigma_1, \cdots, \sigma_p) \text{ is symmetric in the variables } \sigma_1, \cdots, \sigma_p. \tag{1.11}$$

Although we have not checked the details, it seems almost certain that one could with our method treat the more general case where (1.11) does not hold by introducing a random permutation among $i(J, 1), \cdots, i(J, p)$ in (1.9).

We will study $G_N$ and $Z_N$ given by (1.4) and (1.5) respectively. We will always assume $\beta = 1$, because we can think to this parameter as built into $f$. A high temperature hypothesis simply means that the parameter

$$\|f\|_\infty = \sup |f| \tag{1.12}$$

is small. The case of (1.3) is simply the case when

$$f(x, \sigma_1, \cdots, \sigma_p) = -\beta 1_{\{\sigma_1 = a_1(x), \cdots, \sigma_p = a_p(x)\}}$$

where the function $x \to (a_1(x), \cdots, a_p(x))$ is any function that sends Lebesgue measure on $[0, 1]$ onto the uniform measure on $\{-1, 1\}^p$.

Before we state our main result, let us point out that the projection of Gibbs' measure on $\{-1, 1\}^r$ is a random element of the compact set $M(r)$ of probability measures on $\{-1, 1\}^r$. It thereby make sense to say that this distribution converges in law to the law in $M(r)$ of a random probability on $\{-1, 1\}^r$.

**Theorem 1.1.** *Given $\gamma$, $p$ there is a number $a(\gamma, p) > 0$ such that the following occurs. If*

$$\|f\|_\infty \le a(\gamma, p) \tag{1.13}$$

*there is a probability distribution $Q(\gamma, f)$ on $[-1, 1]$ such that, given any integer $r$, as $N \to \infty$ the distribution of $(\sigma_1, \cdots, \sigma_r)$ under Gibbs' measure converges in law to the law of the random product measure $\nu$ on $\{-1, 1\}^r$ such that*

$$\forall i \le r, \quad \int \sigma_i d\nu(\sigma_1, \cdots, \sigma_r) = Y_i$$

*where $(Y_i)_{i \le r}$ are i.i.d. of law $Q(\gamma, f)$.*

In words, this means that two remarkable things happen. First, as $N \to \infty$, the distribution of $(\sigma_1, \cdots, \sigma_r)$ under Gibbs measure resembles a product measure $\nu$ on $\{-1, 1\}^N$. It is then obvious that it must resemble the product measure that gives the same average $\langle \sigma_i \rangle$ to $\sigma_i$ as Gibbs' measure. Moreover, asymptotically, $\langle \sigma_1 \rangle, \cdots, \langle \sigma_r \rangle$ are i.i.d. of law $Q(\gamma, f)$.

**Theorem 1.2.** *The limit*

$$F(\gamma, f) = \lim_{N \to \infty} \frac{1}{N} E \log Z_N$$

*exists (and can in principle be computed as a function of the probability distribution $Q(\gamma, f)$ of Theorem 1.1).*

Let us now comment upon the relationship of this paper with the previous work [T1,2,3]. In three different cases the author has succeeded to prove the validity of the "replica-symmetric solution". Even though there is in the present case an essentially new feature (that the "order parameter" is the distribution $Q(\gamma, f)$) the first and crucial step is to prove that "there is a pure state". This is the aim of Section 2. (The author feels that the new technique presented here is getting close to be able to handle the "most general" case.)

The second main step is to prove that asymptotically, the empirical distribution $N^{-1} \sum_{i \le N} \delta_{\langle \sigma_i \rangle}$ resembles the law of $\langle \sigma_1 \rangle$. This is the aim of Section 3. The difficulty there is specific to the case where the "order parameter is a function" and the work done to overcome this difficulty has no counter part in the previous papers [T1,2,3].

Once the two main obstacles are passed, the rest is easy. The distribution $Q(\gamma, f)$ arises as a fixed point of a certain transformation. It is constructed in Section 4, where the proofs are completed.

## 2. Uniqueness of state

Throughout the paper, we will denote by $\langle \cdot \rangle$ thermal averages, that is, averages with respect to Gibbs' measure. We will use 4-replicas, that is we will consider elements $(\boldsymbol{\sigma}^{(1)}, \cdots, \boldsymbol{\sigma}^{(4)})$ of $\Sigma_N^4$ provided with $G_N^{\otimes 4}$. Averages with respect to this measure are also denoted by $\langle \cdot \rangle$.

The main result of this section is that if $\|f\|_\infty$ is small enough, we have

$$\forall k \ge 1, \qquad \lim_{N \to \infty} C_{N,k} = 0 \tag{2.1}$$

where

$$C_{N,k} = C_{N,k}(\gamma) = E\left\langle \left( \frac{(\boldsymbol{\sigma}^{(1)} - \boldsymbol{\sigma}^{(2)}) \cdot (\boldsymbol{\sigma}^{(3)} - \boldsymbol{\sigma}^{(4)})}{N} \right)^{2k} \right\rangle. \tag{2.2}$$

There, as well as in the sequel,

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i \le N} x_i y_i .$$

It is obvious that $C_{N,k+1} \le 16 C_{N,k}$, so (2.1) is the same as $\lim_{N\to\infty} C_{N,1} = 0$; but the need to consider values of $k \ge 1$ arises for technical reasons. It will require hard work to prove (2.1), so, in order to provide motivation, we show first why (2.1) implies that, for any $r$, the distribution of $(\sigma_1, \cdots, \sigma_r)$ under Gibbs measure is (asymptotically and in average) close to a product measure. This proves "the first half" of Theorem 1.1.

The basic observation is that if $\rho_i \in \{-1, 1\}$ for $i \le r$, then

$$1_{\{\sigma_1 = \rho_1, \cdots, \sigma_r = \rho_r\}} = \prod_{i \le r} \left( \frac{1 + \rho_i \sigma_i}{2} \right).$$

Consider the distribution $\mu$ of $\sigma_1, \cdots, \sigma_r$ under Gibbs measure, and consider the product probability distribution $\nu$ on $\{-1, 1\}^r$ such that $\int \sigma_i d\nu = \langle \sigma_i \rangle$. Then

$$\mu(\{\sigma_1 = \rho_1, \cdots, \sigma_r = \rho_r\}) = \left\langle \prod_{i \le r} \left( \frac{1 + \rho_i \sigma_i}{2} \right) \right\rangle$$
$$\nu(\{\sigma_1 = \rho_1, \cdots, \sigma_r = \rho_r\}) = \prod_{i \le r} \left( \frac{1 + \rho_i \langle \sigma_i \rangle}{2} \right).$$

Thus to show that the expected value of the total variation distance between $\mu$ and $\nu$ goes to zero, it suffices to prove that for each finite set $A$,

$$E \left| \prod_{i \in A} \langle \sigma_i \rangle - \left\langle \prod_{i \in A} \sigma_i \right\rangle \right| \to 0 \ ,$$

and, by symmetry among the variables, that

$$E \left| \prod_{i \le q} \langle \sigma_i \rangle - \left\langle \prod_{i \le q} \sigma_i \right\rangle \right| \to 0. \tag{2.3}$$

Now, by (2.1), for each $q$, we have

$$E \left\langle \left( \frac{(\boldsymbol{\sigma}^{(1)} - \boldsymbol{\sigma}^{(2)}) \cdot (\boldsymbol{\sigma}^{(3)} - \boldsymbol{\sigma}^{(4)})}{N} \right)^q \right\rangle \to 0.$$

After reading the proof of Lemma 2.1 below, it will be obvious that this implies

$$E \left\langle \prod_{i \le q} (\sigma_i^{(1)} - \sigma_i^{(2)}) \right\rangle^2 \to 0$$

and thus

$$E \left| \left\langle \prod_{i \le q} \left( \sigma_i^{(1)} - \sigma_i^{(2)} \right) \right\rangle \right| \to 0$$

from which (2.3) readily follows by induction over $q$.

We now turn towards the proof of (2.1).

**Lemma 2.1.** *We have*

$$C_{N,k} \le E \left\langle \prod_{0 \le m \le 2k-1} (\sigma_{N-m}^{(1)} - \sigma_{N-m}^{(2)})(\sigma_{N-m}^{(3)} - \sigma_{N-m}^{(4)}) \right\rangle + \frac{K(k)}{N}.$$

Throughout the paper, $K$ denotes a universal constant, not necessarily the same at each occurrence. Similarly, $K(k)$ denotes a number depending upon $k$ only, etc.

Even though Lemma 2.1 is proved in [T3], we reproduce the simple proof for the convenience of the reader.

*Proof of Lemma 2.1.* Writing

$$a_i = (\sigma_i^{(1)} - \sigma_i^{(2)})(\sigma_i^{(3)} - \sigma_i^{(4)}) \ ,$$

we have

$$\begin{aligned} C_{N,k} &= E\left\langle \left(\frac{1}{N}\sum_{i\le N} a_i\right)^{2k}\right\rangle \\ &= \frac{1}{N^{2k}} E\left\langle \sum_{i_1,\cdots,i_k} a_{i_1}\cdots a_{i_k}\right\rangle \end{aligned}$$

where the summation is over all choices of $i_1, \cdots, i_k$. The contribution of the terms for which the indexes are not distinct is at most $K(k)/N$; all the other terms are equal by symmetry among the variables. □

In our next step, we learn how to relate the Gibbs measure on $N$ sites with the Gibbs measure on $N-2k$ sites. This is the heart of the cavity method.

Consider the expression $H(\boldsymbol{\sigma}) = H_N(\boldsymbol{\sigma})$ given by (1.9), and the similar expression $H_0(\boldsymbol{\sigma})$ where the summation is restricted to $J \subset \{1, \cdots, N-2k\}$. We note that $H_0$ is the Hamiltonian of an $N-2k$ spin system, except that the parameter $\gamma$ has been replaced by a parameter $\gamma'$ such that

$$\frac{\gamma'}{(N-2k)^{p-1}} = \frac{\gamma}{N^{p-1}}. \tag{2.4}$$

In the sequel, Gibbs' measure relative to the Hamiltonian $H_0$ will simply be called "Gibbs measure on $N-2k$ sites". We will also use "Gibbs measure on $N-1$ sites" or "on $N-r$ sites" with the obvious meaning.

Let us now fix some notation. For $I \in [N-2k]^{p-1}$ and $0 \le m \le 2k-1$, $q \le p$ we set

$$\begin{aligned} \eta_{I,m} &= \eta_{I\cup\{N-m\}} \\ X_{I,m} &= X_{I\cup\{N-m\}} \end{aligned}$$

We write $I = \{i(I,1), \cdots, i(I,p-1)\}$ where $i(I,1) < \cdots < i(I,p-1)$.

For $0 \le m \le 2k-1$, we set

$$\begin{aligned} \mathscr{E}_m &= \mathscr{E}_m(\boldsymbol{\sigma}) \\ &= \exp \sum_{I\in[N-2k]^{p-1}} \eta_{I,m} f(X_{I,m}, \sigma_{i(I,1)}, \cdots, \sigma_{i(I,p-1)}, \sigma_{N-m}) \end{aligned} \tag{2.5}$$

and

$$\mathscr{E} = \prod_{0\le m\le 2k-1} \mathscr{E}_m. \tag{2.6}$$

Let us observe that if the following holds

$$\forall J \in [N]^p, \eta_J = 1 \Rightarrow \operatorname{card}(J \cap [N-2k+1, \cdots, N]) \le 1 \tag{2.7}$$

then

$$\exp -H(\boldsymbol{\sigma}) = \mathscr{E} \exp -H_0(\boldsymbol{\sigma}). \tag{2.8}$$

This is because if $\operatorname{card}(J \cap [N-2k+1, \cdots, N]) \le 1$, the term of $H(\boldsymbol{\sigma})$ corresponding to $J$ occurs in $H_0$ if $\max J \le N-2k$, and occurs in $\mathscr{E}_m$ if $\max J = N-m$. (On the other hand, terms for which $\eta_J = 1$, $\operatorname{card} J \cap [N-2k+1, \cdots, N] \ge 2$ do not occur in the right hand side of (2.8), so that (2.7) is essentially necessary for (2.8) to hold).

Consider now the event $\Omega_1$ defined by (2.7). It should be obvious that

$$P(\Omega_1) \ge 1 - \frac{K(k)}{N} \tag{2.9}$$

and that $\Omega_1$ is probabilistically independent of the r.v. $\mathscr{E}$ and $H_0(\boldsymbol{\sigma})$.

The following fact that is now obvious is fundamental for the sequel.

**Lemma 2.2.** *For any function $h$ on $\{-1, 1\}^{2k}$, on the event $\Omega_1$ we have that*

$$\langle h(\sigma_{N-2k+1}, \cdots, \sigma_N)\rangle = \frac{Av\langle h(\sigma_{N-2k+1}, \cdots, \sigma_N)\mathscr{E}\rangle_0}{Av\langle \mathscr{E}\rangle_0}. \tag{2.10}$$

*There, $Av$ means average with respect to all the possible values of $\sigma_{N-2k+1}, \cdots,$ $\sigma_N = \pm 1$; the bracket $\langle \cdot \rangle_0$ means thermal average in the variables $\sigma_1, \cdots, \sigma_{N-2k}$, with respect to the Gibbs measure on $N-2k$ sites, that is Gibbs measure of Hamiltonian $H_0$.*

We will need the (immediate) extension of (2.10) to 4-replicas. In this extension, we simply replace $\mathscr{E}_m$ by

$$\mathscr{E}'_m = \prod_{\ell \le 4} \mathscr{E}_m(\boldsymbol{\sigma}^{(\ell)}) \tag{2.11}$$

where $\mathscr{E}_m(\boldsymbol{\sigma}^{(\ell)})$ is given by (2.5), and we replace $\mathscr{E}$ by $\mathscr{E}' = \prod_{0 \le m \le 2k-1} \mathscr{E}'_m$.

**Lemma 2.3.** *We have*

$$C_{N,k} \le \frac{K(k)}{N} + E\Bigg(\exp(4\|f\|_\infty \sum \eta_{I,m'})Av \times \Bigg\langle \prod_{0 \le m \le 2k-1} (\sigma^{(1)}_{N-m} - \sigma^{(2)}_{N-m})(\sigma^{(3)}_{N-m} - \sigma^{(4)}_{N-m})\mathscr{E}' \Bigg\rangle_0 \Bigg) \tag{2.12}$$

*where the summation is over $0 \le m' \le 2k-1$, $I \in [N-2k]^{p-1}$.*

*Proof.* We use the version of (2.10) for 4-replicas, with

$$h = \prod_{0 \le m \le 2k-1} (\sigma^{(1)}_{N-m} - \sigma^{(2)}_{N-m})(\sigma^{(3)}_{N-m} - \sigma^{(4)}_{N-m}).$$

We observe that

$$\langle h\mathcal{E}'\rangle_0 = \left\langle \prod_{0 \le m \le 2k-1} (\sigma^{(1)}_{N-m} - \sigma^{(2)}_{N-m})\mathcal{E}(\boldsymbol{\sigma}^{(1)})\mathcal{E}(\boldsymbol{\sigma}^{(2)}) \right\rangle_0^2 \ge 0$$

Moreover,

$$\mathcal{E}(\boldsymbol{\sigma}) \ge \exp(-\|f\|_\infty \sum_{I,m} \eta_{I,m}) \ \ ,$$

so that

$$\langle \mathcal{E}'\rangle_0^{-1} \le \exp(4\|f\|_\infty \sum_{I,m} \eta_{I,m}).$$

The result follows. □

Given $\boldsymbol{\sigma}^{(1)}, \cdots, \boldsymbol{\sigma}^{(4)}$ in $\Sigma_N$, $\ I \in [N-2k]^{p-1}$, we consider the quantity

$$\begin{aligned}\varphi_{I,m} &= \varphi_{I,m}(\boldsymbol{\sigma}^{(1)}, \cdots, \boldsymbol{\sigma}^{(4)}) \\ &= E \exp \sum_{\ell \le 4} f(X_{I,m}, \sigma^{(\ell)}_{i(I,1)}, \cdots, \sigma^{(\ell)}_{i(I,p-1)}, \sigma^{(\ell)}_{N-m})\end{aligned} \tag{2.13}$$

The expectation is of course over the random variables $X_{I,m}$. Let us observe that $\varphi_{I,m}$ depends upon $\boldsymbol{\sigma}^{(\ell)}(\ell \le 4)$ only through $\sigma^{(\ell)}_{i(I,q)}, q \le p-1$ and through $\sigma^{(\ell)}_{N-m}$.

We consider the quantity

$$\begin{aligned}&Av(\sigma^{(1)}_{N-m} - \sigma^{(2)}_{N-m})(\sigma^{(3)}_{N-m} - \sigma^{(4)}_{N-m}) \\ &\exp\left[\frac{\gamma}{N^{p-1}} \sum_I (e^{4\|f\|_\infty}\varphi_{I,m}(\boldsymbol{\sigma}^{(1)}, \cdots, \boldsymbol{\sigma}^{(4)}) - 1)\right]\end{aligned} \tag{2.14}$$

where the summation is over $I \in [N-2k]^{-1}$, and the average over $\sigma^{(\ell)}_{N-m} = \pm 1$. It does not depend upon $m$, but only upon $\boldsymbol{\rho}^{(\ell)} = (\sigma^{(\ell)}_i)_{i \le N-2k}$. We denote by $\Phi(\boldsymbol{\rho}^{(1)}, \cdots, \boldsymbol{\rho}^{(4)})$ this quantity (2.14).

**Lemma 2.4.** *We have*

$$C_{N,k} \le \frac{K(k,f)}{N} + E\langle \Phi(\boldsymbol{\rho}^{(1)}, \cdots, \boldsymbol{\rho}^{(4)})^{2k}\rangle_0. \tag{2.15}$$

*Proof.* The idea is simply to perform integration $E_X$ in the variables $X_{I,m}$ in (2.12), and then integration $E_\eta$ in the variables $\eta_{I,m}$.

To integrate in the variables $X_{I,m}$, we observe that the dependence upon these variables of the right-hand side of (2.12) is only through $\mathcal{E}'$; we also observe that for $\eta \in \{0, 1\}$, we have $EY^\eta = (EY)^\eta$, so that

$$E_X \mathcal{E}'_m = \prod_I \varphi_{I,m}^{\eta_{I,m}} \tag{2.16}$$

where of course the product is over $I \in [N-2k]^p$. Thus (2.12) becomes

$$C_{N,k} \le E\left\langle \prod_{0\le m\le 2k-1} \left( Av(\sigma^{(1)}_{N-m} - \sigma^{(2)}_{N-m})(\sigma^{(3)}_{N-m} - \sigma^{(4)}_{N-m}) \prod_I (\varphi_{I,m} \exp 4\|f\|_\infty)^{\eta_{I,m}} \right)\right\rangle_0 + K(k)/N. \tag{2.17}$$

We will now take expectation in the variables $\eta_{I,m}$. We observe the independence of the terms in the product $\prod_m$. We also observe that

$$Ea^\eta = 1 + P(\eta = 1)(a-1)$$

so that

$$E\prod_I (\varphi_{I,m} \exp 4\|f_\infty\|)^{\eta_{I,m}} = \prod_I \left(1 + \frac{\gamma}{N^{p-1}}(e^{4\|f\|_\infty}\varphi_{I,m} - 1)\right) \tag{2.18}$$

Now, writing

$$1 + x = e^{x+R(x)}$$

where $|R(x)| \le Kx^2$ for $x \ge 0$, we obtain the result (there we use that $p \ge 2$). □

In order to use (2.15), we need to understand $\Phi$. The only terms that contribute to $Av$ in (2.14) are those for which $\sigma^{(2)}_{N-m} = -\sigma^{(1)}_{N-m}$ and $\sigma^{(4)}_{N-m} = -\sigma^{(3)}_{N-m}$. Thus we can make the change of variables

$$\sigma^{(1)}_{N-m} = \epsilon;\ \sigma^{(2)}_{N-m} = -\epsilon;\ \sigma^{(3)}_{N-m} = \epsilon';\ \sigma^{(4)}_{N-m} = -\epsilon'\ , \tag{2.19}$$

and write

$$\Phi(\boldsymbol{\rho}^{(1)}, \cdots, \boldsymbol{\rho}^{(4)}) = \mathop{Av}_{\epsilon,\epsilon'} \epsilon\epsilon' \exp \frac{\gamma}{N^{p-1}} \sum_I (e^{4\|f\|_\infty}\varphi_{I,m} - 1) \tag{2.20}$$

where $\varphi_{I,m}$ is as before, replacing in (2.13) the quantities $\sigma^{(\ell)}_{N-m}$ by their values (2.19). We must now understand $\varphi_{I,m}$ better. To reflect the fact that the dependence of $\varphi_{I,m}$ in $\sigma^{(\ell)}_{N-m}$ is only through $\epsilon$ and $\epsilon'$, we will make a change of notation, and we write

$$\varphi_{I,m} = \varphi_{I,m}(\boldsymbol{\rho}^{(1)}, \cdots, \boldsymbol{\rho}^{(4)}, \epsilon, \epsilon'). \tag{2.21}$$

(We keep the index $m$ even though this quantity does not depend upon $m$). We observe the fundamental properties that

$$\varphi_{I,m}(\boldsymbol{\rho}^{(1)}, \cdots, \boldsymbol{\rho}^{(4)}, \epsilon, \epsilon') = \varphi_{I,m}(\boldsymbol{\rho}^{(2)}, \boldsymbol{\rho}^{(1)}, \boldsymbol{\rho}^{(3)}, \boldsymbol{\rho}^{(4)}, -\epsilon, \epsilon') \tag{2.22}$$

$$\varphi_{I,m}(\boldsymbol{\rho}^{(1)}, \cdots, \boldsymbol{\rho}^{(4)}, \epsilon, \epsilon') = \varphi_{I,m}(\boldsymbol{\rho}^{(1)}, \boldsymbol{\rho}^{(2)}, \boldsymbol{\rho}^{(4)}, \boldsymbol{\rho}^{(3)}, \epsilon, -\epsilon') \tag{2.23}$$

For $s = 0, 1, 2, 3$, we define 4 functions $\psi^s_I$ of $\boldsymbol{\rho}^{(1)}, \cdots, \boldsymbol{\rho}^{(4)}$ by

$$\psi^0_I = \mathop{Av}_{\epsilon,\epsilon'} \varphi_I (= \mathop{Av}_{\epsilon,\epsilon'} \varphi_{I,m}(\boldsymbol{\rho}^{(1)}, \boldsymbol{\rho}^{(2)}, \boldsymbol{\rho}^{(3)}, \boldsymbol{\rho}^{(4)}, \epsilon, \epsilon'))$$

$$\psi^1_I = \mathop{Av}_{\epsilon,\epsilon'} \epsilon\varphi_{I,m}$$

$$\psi^2_I = \mathop{Av}_{\epsilon,\epsilon'} \epsilon'\varphi_{I,m}$$

$$\psi^3_I = \mathop{Av}_{\epsilon,\epsilon'} \epsilon\epsilon'\varphi_{I,m}$$

Thus (2.22) and (2.23) respectively imply the fundamental facts that

$$\text{Exchanging } \boldsymbol{\rho}^{(1)} \text{ and } \boldsymbol{\rho}^{(2)} \text{ changes the sign of } \psi_I^1, \psi_I^3. \tag{2.24}$$

$$\text{Exchanging } \boldsymbol{\rho}^{(3)} \text{ and } \boldsymbol{\rho}^{(4)} \text{ changes the sign of } \psi_I^2, \psi_I^3. \tag{2.25}$$

Moreover, we have the identity

$$\varphi_{I,m} = \psi_I^0 + \epsilon\psi_I^1 + \epsilon'\psi_I^2 + \epsilon\epsilon'\psi_I^3$$

and thus

$$\frac{\gamma}{N^{p-1}} \sum_I (e^{4\|f\|_\infty}\varphi_{I,m} - 1) = W_0 + \epsilon W_1 + \epsilon' W_2 + \epsilon\epsilon' W_3 \tag{2.26}$$

where, for $s = 1, 2, 3$

$$W_s = \frac{\gamma}{N^{p-1}} e^{4\|f\|_\infty} \sum_I \psi_I^s. \tag{2.27}$$

and

$$W_0 = \frac{\gamma}{N^{p-1}} \sum_I (e^{4\|f\|_\infty}\psi_I^0 - 1) \tag{2.28}$$

Now, going back to (2.20) and using (2.27), we get

$$\begin{aligned}\Phi(\boldsymbol{\sigma}^{(1)}, \cdots, \boldsymbol{\sigma}^{(4)}) &= \underset{\epsilon,\epsilon'}{Av} \exp(W_0 + \epsilon W_1 + \epsilon' W_2 + \epsilon\epsilon' W_3) \\ &= \exp W_0(\text{ch } W_1 \text{ ch } W_2 \text{ sh } W_3 + \text{sh } W_1 \text{ sh } W_2 \text{ ch } W_3).\end{aligned} \tag{2.29}$$

Since we want to prove that something happens for $\|f\|_\infty$ small, there is no loss of generality to assume $\|f\|_\infty \le 1$. Then $W_s(s \le 3)$ are bounded by a number depending upon $\gamma$ only, so that (2.29) implies

$$\Phi(\boldsymbol{\sigma}^{(1)}, \cdots, \boldsymbol{\sigma}^{(4)})^{2k} \le K(\gamma, p)^k (W_3^{2k} + (W_1 W_2)^{2k})$$

Combining with (2.15), we have proved the following

**Lemma 2.5.** *We have*

$$C_{N,k} \le \frac{K(f, p, k)}{N} + K(\gamma, p)^k (E\langle W_3^{2k}\rangle_0 + E\langle (W_1 W_2)^{2k}\rangle_0) \tag{2.30}$$

Now we have to relate these terms to $C_{N-2k,2k}$; We will prove the following.

**Lemma 2.6.** *We have*

$$E\langle W_3^{2k}\rangle_0 \le \frac{K(\gamma, p, k)}{\sqrt{N}} + K(\gamma, p)^k \|f\|_\infty^{2k} C_{N-2k,2k}^{1/2}(\gamma') \tag{2.31}$$

$$E\langle (W_1 W_2)^{2k}\rangle_0 \le \frac{K(\gamma, p, k)}{\sqrt{N}} + K(\gamma, p)^k \|f\|_\infty^{2k} C_{N-2k,2k}^{1/2}(\gamma') \tag{2.32}$$

A crucial idea there is that the potentially disastrous power $1/2$ on the right is offset by the index $2k$ rather than $k$. To explain this, we state and prove our main result.

**Theorem 2.7.** *There is a constant $K(\gamma, p)$ such that*

$$\|f\|_\infty K(\gamma, p) \le 1 \Rightarrow \lim_{N\to\infty} C_{N,1}(\gamma) = 0.$$

*Proof.* Given $\gamma_0$ arbitrary, we set

$$c_k = \limsup_{N\to\infty} \sup_{\gamma\le\gamma_0} C_{N,k}(\gamma).$$

Since $\gamma' \le \gamma$, Lemmas 2.5, 2.6 show that

$$c_k \le (K(\gamma_0, p)\|f\|_\infty^2)^k c_{2k}^{1/2}.$$

Thus

$$c_k^{1/k} \le (K(\gamma_0, p)\|f\|_\infty^2) c_{2k}^{1/2k}.$$

Since $c_k \le 16^k$, if $K(\gamma_0, p)\|f\|_\infty < 1$ this implies $c_k = 0$ for each $k$. □

*Proof of Lemma 2.6* We need a better understanding of the quantities $W_1, W_2, W_3$. The functions $\psi_I^s$ depend upon $\boldsymbol{\rho}^{(1)}, \cdots, \boldsymbol{\rho}^{(4)}$ only through $\sigma_{i(I,q)}^{(\ell)}$, $\ell \le 4, q \le p-1$. We use elementary Fourier analysis in $\{-1,1\}^{4(p-1)}$ to write

$$\psi_I^s(\boldsymbol{\rho}^{(1)}, \cdots, \boldsymbol{\rho}^{(4)}) = \sum a^s(B_1, \cdots, B_{p-1}) \prod_{q\le p-1} \prod_{\ell\in B_q} \sigma_{i(I,q)}^{(\ell)} \tag{2.33}$$

where the summation is over all choices of $B_1, \cdots, B_{p-1} \subset \{1, \cdots, 4\}$. To control the Fourier coefficients $a^s(B_1, \cdots, B_{p-1})$ we observe first that $a^s(B_1, \cdots, B_{p-1}) = 0$ if all sets $B_1, \cdots, B_{p-1}$ are empty (and if $s \ne 0$!) It should also be obvious that

$$|a^s(B_1, \cdots, B_{p-1})| \le K\|f\|_\infty \tag{2.34}$$

Indeed (for specificity)

$$a^3(B_1, \cdots, B_{p-1}) = E\,Av\,\epsilon\epsilon' \prod_{q\le p-1} \prod_{\ell\in B_q} \sigma_{i(I,q)}^{(\ell)} (e^F - 1)$$

where the average is over $\epsilon, \epsilon', \sigma_{i(I,q)}^{(\ell)} = \pm 1$ and $F$ is a random function of these numbers that satisfies $|F| \le 4\|f\|_\infty$. The smallness of $\|f\|_\infty$ will of course be used through (2.34). We observe that

$$N^{-(p-1)} \sum_I \prod_{q\le p-1} \prod_{\ell\in B_q} \sigma_{i(I,q)}^{(\ell)} = R \sum_{i_1,\cdots,i_{p-1}} \prod_{q\le p-1} \prod_{\ell\in B_q} \sigma_{i_q}^{(\ell)}$$

where the summation in the right hand side is over $i_1, \cdots, i_{p-1}$ all different, and where

$$R = \frac{1}{N^{p-1}} \binom{N-2k}{p-1} \le 1.$$

It thus follows that

$$\Bigg| \frac{1}{N^{p-1}} \sum_I \prod_{q \le p-1} \prod_{\ell \in B_q} \sigma_{i(I,q)}^{(\ell)} - R \prod_{q \le p-1} \left( \frac{1}{N} \sum_{i \le N-2k} \prod_{\ell \in B_q} \sigma_i^{(\ell)} \right) \Bigg| \le \frac{K(p)}{N}.$$

Let us denote by $T$ (resp. $T'$) the transformation that consists of exchanging $\boldsymbol{\rho}^{(1)}$ and $\boldsymbol{\rho}^{(2)}$ (resp. $\boldsymbol{\rho}^{(3)}$ and $\boldsymbol{\rho}^{(4)}$) so that by (2.24), (2.25)

$$T(\psi_I^3) = T'(\psi_I^3) = -\psi_I^3$$

and thus

$$\psi_I^3 = \frac{1}{4}(\psi_I^3 - T(\psi_I^3) - T'(\psi_I^3) + TT'(\psi_I^3)) \tag{2.35}$$

Combining with (2.27), (2.33) we get

$$\begin{aligned} &|W_3 - R \sum b^3(B_1, \cdots, B_{p-1})(U - T(U) - T'(U) + TT'(U))| \\ &\le \frac{K(p,\gamma)}{N} \end{aligned} \tag{2.36}$$

There,

$$U = U(B_1, \cdots, B_{p-1}) = \prod_{q \le p-1} \left( \frac{1}{N} \sum_{i \le N-2k} \prod_{\ell \in B_q} \sigma_i^{(\ell)} \right) , \tag{2.37}$$

the summation is over all choices of $B_1, \cdots, B_{p-1}$ and

$$|b^3(B_1, \cdots, B_{p-1})| \le K(\gamma) \| f \|_\infty .$$

Thus, to prove (2.31), it suffices to show that given $B_1, \cdots, B_{p-1}$, we have

$$E\langle (U - T(U) - T'(U) + TT'(U))^{2k} \rangle_0 \le \frac{K(p,k)}{N} + K(p)^k C_{N-2k,2k}^{1/2}(\gamma'). \tag{2.38}$$

We write $\dot{\sigma}_i = \sigma_i - b_i$, where $b_i = \langle \sigma_i \rangle_0$. Writing $\sigma_i = \dot{\sigma}_i + b_i$, and expanding the products $\prod_{\ell \in B_q}$ we see now that it suffices to prove (2.58) where rather than (2.37), we have

$$U = \prod_{q \le p-1} \left( \frac{1}{N} \sum_{i \le N-2k} \prod_{\ell \in B_q} \dot{\sigma}_i^{(\ell)} b_i^{n(q)} \right) =: \prod_{q \le p-1} V_q \tag{2.39}$$

(where $0 \le n(q) \le 4$). The essential observation is that

$$\begin{aligned} &U - T(U) - T'(U) + TT'(U) \ne 0 \Rightarrow \\ &\exists q_1; B_{q_1} \cap \{1,2\} \ne \emptyset; \exists q_2; B_{q_2} \cap \{3,4\} \ne \emptyset. \end{aligned} \tag{2.40}$$

Thus, to prove (2.38) it suffices to prove that under (2.40)

$$E\langle U^{2k}\rangle_0 \le \frac{K(p,k)}{N} + K(p)^k C_{N-2k,2k}^{1/2}(\gamma') \tag{2.41}$$

where $U$ is as in (2.39). Since $|\dot{\sigma}_i^{(\ell)}| \le 2$, $|b_i| \le 1$, we have $|V_q| \le K^k$ for each $q$. We fix $q_1, q_2$ as in (2.40).

*case 1*: $q_1 \ne q_2$.

We have

$$\begin{aligned} U^{2k} &\le K(p)^k V_{q_1}^{2k} V_{q_2}^{2k} \le K(p)^k (V_{q_1}^{4k} + V_{q_2}^{4k}) \\ &\le K(p)^k \left[ \sum_{j=1,2} \left( \frac{1}{N} \sum_{i \le N-2k} \prod_{\ell \in B_{q_j}} \dot{\sigma}_i^{(\ell)} b_i^{n(q_j)} \right)^{4k} \right] \end{aligned}$$

and all we have to show is that

$$\begin{aligned} B \ne \emptyset \Rightarrow E\left\langle \left( \frac{1}{N} \sum_{i \le N-2k} \prod_{\ell \in B} \dot{\sigma}_i^{(\ell)} b_i^n \right)^{4k} \right\rangle_0 \\ \le K(p)^k C_{N-2k,2k}^{1/2}(\gamma') + \frac{K(p,k)}{N}. \end{aligned} \tag{2.42}$$

Expanding the power and using trivial bounds, the left hand side is at most

$$K(p)^k \left( E \left| \left\langle \prod_{i \le 4k} \dot{\sigma}_i \right\rangle_0 \right| + \frac{1}{N} \right)$$

Now, by Cauchy-Schwarz,

$$\begin{aligned} E \left| \left\langle \prod_{i \le 4k} \dot{\sigma}_i \right\rangle_0 \right| &\le \left( E \left\langle \prod_{i \le 4k} \dot{\sigma}_i \right\rangle_0^2 \right)^{1/2} \\ &= \left( E \left\langle \prod_{i \le 4k} \dot{\sigma}_i^{(1)} \dot{\sigma}_i^{(2)} \right\rangle_0 \right)^{1/2}, \end{aligned}$$

and

$$\begin{aligned} E \left\langle \prod_{i \le 4k} \dot{\sigma}_i^{(1)} \dot{\sigma}_i^{(2)} \right\rangle_0 &\le E \left\langle \left( \frac{1}{N-2k} \sum_{i \le N-2k} \dot{\sigma}_i^{(1)} \dot{\sigma}_i^{(2)} \right)^{4k} \right\rangle_0 \\ &\le C_{N-2k,2k}(\gamma'). \end{aligned}$$

The last inequality is seen by integrating $\boldsymbol{\rho}^{(2)}$, $\boldsymbol{\rho}^{(4)}$ inside the power $4k$ rather than outside in the definition of $C_{N-2k,2k}$.

*case 2*: $q_1 = q_2$. We will show that

$$\text{card}B \geq 2 \Rightarrow E\left\langle \left( \frac{1}{N} \sum_{i \leq N-2k} \prod_{\ell \in B} \dot{\sigma}_i^{(\ell)} b_i^n \right)^{2k} \right\rangle_0 \leq K(p)^k C_{N-2k,2k}^{1/2}(\gamma') + \frac{K(p,k)}{N}. \tag{2.43}$$

By Cauchy-Schwarz, the left hand side of (2.43) is at most

$$\left( E\left\langle \left( \frac{1}{N} \prod_{i \leq N-2k} \prod_{\ell \in B} \dot{\sigma}_i^{(\ell)} b_i^n \right)^{4k} \right\rangle_0 \right)^{1/2}.$$

Expanding the power, using trivial bounds, and the fact that card$B \geq 2$, this is at most

$$(K(p)^k \left( E\left\langle \prod_{i \leq 4k} \dot{\sigma}_i \right\rangle_0^2 + \frac{1}{N} \right)^{1/2},$$

and we finish as before.

We have proved (2.31). The proof of (2.32) is very similar and is left to the reader. □

## 3. Non-correlation of spin averages

Throughout this section, we denote by $M_1$ the set of probability measures on $[-1, 1]$, provided with the distance

$$d(\nu, \mu) = \sup \left| \int \theta(x) d\nu(x) - \int \theta(x) d\mu(x) \right| \tag{3.1}$$

where the supremum is over the functions $\theta : [0, 1] \to \mathbb{R}$ that have a Lipschitz constant 1, that is satisfy

$$|\theta(x) - \theta(y)| \leq |x - y|$$

for all $x, y$ in $[-1, 1]$. We can assume $\theta(0) = 0$, so that $\theta$ is valued in $[-1, 1]$.

The distance (3.1) is known as the “transportation cost” between $\nu, \mu$. Its use is motivated by the fact that we can find a pair $U, V$ of random variables such that

$$\mathscr{L}(U) = \mu, \mathscr{L}(V) = \nu, E|U - V| = d(\nu, \mu). \tag{3.2}$$

This statement (known as the Monge-Kantorovich theorem) is a concrete, efficient way to use the information provided by $d$. The aim of this section is to prove the following:

**Theorem 3.1.** *For $\gamma_0 > 0$, there exists a constant $K(\gamma_0, p)$ such that if $\|f\|_\infty K(\gamma_0, p) < 1$, then*

$$\lim_{N\to\infty} Ed\left(\frac{1}{N}\sum_{i\le N}\delta_{\langle\sigma_i\rangle}, \mathscr{L}(\langle\sigma_1\rangle)\right) = 0 \tag{3.3}$$

*uniformly in $N$ for $\gamma \le \gamma_0$.*

There $\mathscr{L}(\langle\sigma_1\rangle)$ denotes the law of the r.v. $\langle\sigma_1\rangle$.

Given two r.v. $X, Y$, we write $C(X, Y) = EXY - EXEY$, their correlation.

**Lemma 3.2.** *To prove (3.3) it suffices to prove that for each function $\theta$ with Lipschitz constant $\le 1$, we have*

$$\lim_{N\to\infty} EC(\theta(\langle\sigma_1\rangle), \theta(\langle\sigma_2\rangle)) = 0 \tag{3.4}$$

*uniformly over $\gamma \le \gamma_0$.*

*Proof.* Under (3.4), we have

$$E\left(\frac{1}{N}\sum_{i\le N}(\theta(\langle\sigma_i\rangle) - E\theta(\langle\sigma_i\rangle))\right)^2 \to 0.$$

The result follows, since in (3.1) the sup can be arbitrarily (and uniformly over $\mu, \nu$) approximated by a finite maximum. □

Given a function $\theta$, from $[-1, 1]^k$ to $[-1, 1]$, we define $L(\theta)$ as the smallest number such that we can find numbers $(b(j))_{j\le k}$ for which

$$\forall x, y \in [-1, 1]^k, \quad |\theta(x) - \theta(y)| \le \sum_{j\le k} b(j)|x_j - y_j|$$

and

$$\sum_{j\le k} b(j) \le L(\theta).$$

Thus, for $k = 1$, $L(\theta)$ is simply the Lipschitz constant of $\theta$. The proof of (3.4) relies upon the following statement.

**Proposition 3.3.** *There is a number $K(\gamma_0, p)$ such that if $\|f\|_\infty K(\gamma_0, p) \le 1$ and $\gamma \le \gamma_0$, given $k \ge 1$, given any function $\theta : [-1, 1]^k \to [-1, 1]$, given $\epsilon > 0$, we can find $k' > k$ arbitrarily large and $\theta' : [-1, 1]^{k'} \to [-1, 1]$ with the following properties*

$$L(\theta') \le L(\theta)/2 \tag{3.5}$$

$$\begin{aligned} &C(\theta(\langle\sigma_1\rangle, \cdots, \langle\sigma_k\rangle), \theta(\langle\sigma_{k+1}\rangle \cdots, \langle\sigma_{2k}\rangle)) \\ &\le C(\theta'(\langle\sigma_1\rangle_0, \cdots, \langle\sigma_{k'}\rangle_0), \theta'(\langle\sigma_{k'+1}\rangle_0, \cdots, \langle\sigma_{2k'}\rangle_0)) \\ &+\epsilon + o(1). \end{aligned} \tag{3.6}$$

*There, $\langle\ \cdot\ \rangle_0$ denotes the Gibbs measure for $N - 2k$ sites, and parameter $\gamma' = \gamma((N-2k)/N)^{p-1}$, and $o(1)$ a quantity that goes to zero as $N \to \infty$.*

To provide motivation for this technical statement, we first show why this implies Theorem 3.1. Given a 1-Lipschitz function $\theta : [-1,1] \to [-1,1]$, we iterate the result of Proposition 3.3 to find, given $r$, integers $k', k''$ and a map $\theta' : [-1,1]^{k'} \to [-1,1]$ such that

$$L(\theta') \leq L(\theta)/2^{-r} \leq 2^{-r} \tag{3.7}$$

$$\begin{aligned} & C(\theta(\langle\sigma_1\rangle), \theta(\langle\sigma_2\rangle)) \\ & \leq C(\theta'(\langle\sigma_1\rangle_0, \cdots, \langle\sigma_{k'}\rangle_0), \theta'(\langle\sigma_{k'+1}\rangle_0, \cdots, \langle\sigma_{2k'}\rangle_0)) \\ & + \frac{1}{r} + o(1). \end{aligned} \tag{3.8}$$

where $\langle \ \cdot \ \rangle_0$ is the Gibbs measure for $N - 2k''$ sites. By definition of $L(\theta')$, we can find numbers $(b(j))_{j \leq k'}$ such that $\sum_{j \leq k'} b(j) \leq L(\theta') \leq 2^{-r}$ and that

$$\begin{aligned} \forall x, y \in [-1,1]^{k'}, \quad |\theta'(x_1, \cdots, x_{k'}) - \theta'(y_1, \cdots, y_{k'})| & \leq \sum_{j \leq k'} b(j)|x_j - y_j| \\ & \leq 2 \sum_{j \leq k'} b(j) \leq 2^{1-r} \end{aligned}$$

This implies that $\theta'$ takes its values in an interval of length $\leq 2^{2-r}$. It follows easily that the first term to the right of (3.7) is $\leq 2^{3-2r}$, so that

$$C(\theta(\langle\sigma_1\rangle), \theta(\langle\sigma_2\rangle)) \leq 2^{3-2r} + \frac{1}{r} + o(1).$$

Letting $N \to \infty$ and then $r \to \infty$ completes the proof.

To prove Proposition 3.3, for consistency with the notation of Section 2, we will replace $\langle\sigma_1\rangle, \cdots, \langle\sigma_{2k}\rangle$ by $\langle\sigma_N\rangle, \cdots, \langle\sigma_{N-2k+1}\rangle$.

We will use the cavity method as in Section 2. To evaluate

$$C(\theta(\langle\sigma_N\rangle, \cdots, \langle\sigma_{N-k+1}\rangle), \theta(\langle\sigma_{N-k}\rangle, \cdots, \langle\sigma_{N-2k+1}\rangle)) \ , \tag{3.9}$$

it follows from Lemma 2.2 that we make an error $o(1)$ if replace $\langle\sigma_{N-m}\rangle$ by

$$\frac{Av\langle\sigma_{N-m}\mathcal{E}\rangle_0}{Av\langle\mathcal{E}\rangle_0} \tag{3.10}$$

where $\mathcal{E}$ is given by (2.6). One unpleasant feature of $\mathcal{E}$ is that $\sum_I \eta_{I,m}$ is not bounded. Given an integer $u$, we define a small perturbation $\overline{\eta}_{I,m}$ of the variables $\eta_{I,m}$ for which

$$\forall m \leq 2k-1, \sum_I \overline{\eta}_{I,m} \leq u. \tag{3.11}$$

They are many ways to do this. The reader will choose his own. We simply need that

$$\forall m \leq 2k-1, \sum_I \eta_{I,m} \leq u \Rightarrow \forall I, \eta_{I,m} = \overline{\eta}_{I,m} \tag{3.12}$$

and that the variables $\overline{\eta}_{I,m}$ are independent of the variables $\overline{\eta}_{I,m'}$ if $m \neq m'$.

Let us define $\overline{\mathcal{E}}_m, \overline{\mathcal{E}}$ as in (2.6) when $\eta_{I,m}$ is replaced by $\overline{\eta}_{I,m}$. Then it should be obvious that given $\epsilon$, we can fix $u$ so that (3.9) is at most

$$C(\theta(a_0, \cdots, a_{k-1}), \theta(a_k, \cdots, a_{2k-1})) + \epsilon + o(1) \tag{3.13}$$

where

$$a_m = \frac{Av\langle \sigma_{N-m}\overline{\mathcal{E}}\rangle_0}{Av\langle \overline{\mathcal{E}}\rangle_0}. \tag{3.14}$$

We set $k' = k(p-1)u$. By (3.11), $\overline{\mathcal{E}}$ involves only a bounded number of variables $\sigma_i$. It follows from the work of Section 2 that we make an error at most $o(1)$ if in (3.14) we replace average $\langle\ \cdot\ \rangle_0$ by average for the product measure $\mu_0$ such that

$$\forall i \leq N - 2k, \int \sigma_i d\mu_0(\boldsymbol{\rho}) = \langle \sigma_i \rangle_0.$$

Writing $\mu(f)$ for $\int f d\mu$, we see that in (3.14) we can replace (3.13) by

$$a_m = \frac{\mu_0(Av\sigma_{N-m}\overline{\mathcal{E}})}{\mu_0(Av\overline{\mathcal{E}})}. \tag{3.15}$$

As $N \to \infty$, there is only a vanishing probability such that the sets $\cup\{I; \overline{\eta}_{I,m} = 1\}$ are not disjoint as $m$ ranges from 0 to $2k-1$. Since $\mu_0$ is a product measure, in (3.13) we can replace (3.15) by

$$a_m = \frac{\mu_0(Av\sigma_{N-m}\overline{\mathcal{E}}_m)}{\mu_0(Av\overline{\mathcal{E}}_m)}. \tag{3.16}$$

$\theta(a_0, \cdots, a_{k-1})$ is a function of the numbers $\langle\sigma_i\rangle_0$, but unfortunately this function is random, in that it depends in particular upon the values of the numbers $\overline{\eta}_{I,m}$; moreover we would really like that this function depends only upon $\langle\sigma_i\rangle_0$ for $i \leq k'$; which is certainly not the case a priori. To go over that difficulty, we will show that the symmetry among the sites for $\langle\ \cdot\ \rangle_0$ lets us rearrange coordinates. More precisely, conditionally upon the numbers $\overline{\eta}_{I,m}(I \in [N-2k]^{p-1}, m \leq 2k-1)$, the distribution of the pair

$$\theta(a_0, \cdots, a_{k-1}), \theta(a_k, \cdots, a_{2k-1}) \tag{3.17}$$

depends only upon the numbers $b_m = \sum_I \overline{\eta}_{I,m}, m = 0, \cdots, 2k-1$. Note that $b_m \leq u$. Thus, conditionally upon the numbers $\overline{\eta}_{I,m}$, we do not change the distribution of the pair (3.16) if we now define

$$a_m = \frac{\mu_0(Av\sigma_{N-m}\mathcal{E}_m(b_m))}{\mu_0(Av\mathcal{E}_m(b_m))} \tag{3.18}$$

where

$$\mathcal{E}_m(b_m) = \exp \sum_{1 \leq n \leq b_m} f(X_{m,n}, \sigma_{(p-1)(mu+n-1)+1}, \cdots, \sigma_{(p-1)(mu+n)}, \sigma_{N-m}) \ , \tag{3.19}$$

and where $(X_{m,n})$ are i.i.d uniform over $[-1,1]$. Simply stated, the $b_m(p-1)$ coordinates of $\cup\{I; \overline{\eta}_{I,m}=1\}$ are rearranged as the first $b_m(p-1)$ coordinates of the $m$-th block of length $(p-1)u$. These blocks are long enough to accommodate $b_m(p-1)$ coordinates.

From now on, $a_m$ is given by (3.18).

Given a sequence $\mathbf{y} = (y_i)_{1\le i\le N-2k}$, consider the product measure $\mu_{\mathbf{y}}$ on $\{-1,1\}^{N-2k}$ given by $\mu_{\mathbf{y}}(\sigma_i) = y_i$. Thus $\mu_0 = \mu_{\mathbf{y}}$ for $\mathbf{y} = \langle \boldsymbol{\rho}\rangle_0 = (\langle\sigma_i\rangle_0)_{1\le i\le N-2k}$. We define now

$$a_m(\mathbf{y},\mathbf{X},\eta) = \frac{\mu_{\mathbf{y}}(Av\sigma_{N-m}\mathscr{E}_m(b_m))}{\mu_{\mathbf{y}}(Av\mathscr{E}_m(b_m))} \tag{3.20}$$

where $b_m = \sum_I \eta_{I,m}$, $\mathbf{X} = (X_{m,n})$, $\eta = (\overline{\eta}_{I,m})$. With this notation, we have $a_m = a_m(\mathbf{y},\mathbf{X},\eta)$ for $\mathbf{y} = \langle\boldsymbol{\rho}\rangle_0$.

We now introduce two random functions $\theta_1, \theta_2$ on $[-1,1]^{N-2k}$, given by

$$\theta_1(\mathbf{y}) = \theta(a_0(\mathbf{y},\mathbf{X},\eta),\cdots,a_{k-1}(\mathbf{y},\mathbf{X},\eta)) \tag{3.21}$$

$$\theta_2(\mathbf{y}) = \theta(a_k(\mathbf{y},\mathbf{X},\eta),\cdots,a_{2k-1}(\mathbf{y},\mathbf{X},\eta)). \tag{3.22}$$

Thus $\theta(a_0,\cdots,a_{k-1})$ is $\theta_1(\mathbf{y})$ calculated for $\mathbf{y} = \langle\boldsymbol{\rho}\rangle_0$.

The dependence of $\theta_1$ from $\mathbf{y}$ is only through $y_1,\cdots,y_{k'}$ so that we can define a function

$$\theta'(y_1,\cdots,y_{k'}) = E'\theta_1(\mathbf{y}) \ \ ,$$

the expectation $E'$ being taken in the random variables $X_{n,m}, \overline{\eta}_{I,m}$. It should be quite obvious that

$$\theta'(y_{k'+1},\cdots,y_{2k'}) = E'\theta_2(\mathbf{y})$$

because $\theta_2$ is constructed like $\theta_1$, shifting the dependence through $\mathbf{y}$ by $k'$ places to the right. Let us observe that

$$\theta'(\langle\sigma_1\rangle_0,\cdots,\langle\sigma_{k'}\rangle_0) = E'\theta(a_0,\cdots,a_{k-1}) \tag{3.23}$$

$$\theta'(\langle\sigma_{k'+1}\rangle_0,\cdots,\langle\sigma_{2k'}\rangle_0) = E'\theta(a_k,\cdots,a_{2k-1}). \tag{3.24}$$

The expectation in (3.23) is a smoothing operation, that will be responsible for the fact that $\theta'$ is a smooth function of the variables $y_i$.

Another crucial observation is that $\theta_1(\mathbf{y})$ and $\theta_2(\mathbf{y})$ are probabilistically independent functions of the randomness $(\mathbf{X},\eta)$; this is because $\theta_1$ depends only upon those variables $X_{m,n}, \overline{\eta}_{I,m}$ where $m \le k-1$, while $\theta_2$ depends only upon those with $k \le m \le 2k-1$. Thus, we have

$$\begin{aligned} &C(\theta(a_0,\cdots,a_{k-1}),\theta(a_k,\cdots,a_{2k-1}) \\ &\quad = C(\theta'(\langle\sigma_1\rangle_0,\cdots,\langle\sigma_{k'}\rangle_0),\theta'(\langle\sigma_{k'+1}\rangle_0,\cdots,\langle\sigma_{2k'}\rangle_0)) \end{aligned} \tag{3.25}$$

as we see by first integrating in $\mathbf{X},\eta$ conditionally upon $\langle\ \cdot\ \rangle_0$, and using (3.23), (3.24).

In view of (3.20), (3.25), to prove Proposition 3.3, we only have to prove the required bound on $L(\theta')$.

First, given $\mathbf{X}, \eta$, we will study the dependence of $a_m(\mathbf{y}, \mathbf{X}, \eta)$ upon $\mathbf{y}$. For a function $v$ on $[-1, 1]^N$, it is obvious that

$$\left| \frac{\partial}{\partial y_i} \mu_{\mathbf{y}}(v) \right| \leq \sup v - \inf v. \tag{3.26}$$

Together with the fact that

$$\exp -b_m \| f \|_\infty \leq \mathscr{E}_m(b_m) \leq \exp b_m \| f \|_\infty \ ,$$

crude elementary estimates then imply that

$$\frac{\partial}{\partial y_i} a_m(\mathbf{y}, \mathbf{X}, \eta) \leq 4 b_m \| f \|_\infty \exp 3 \| f \|_\infty b_m .$$

Since $a_m(\mathbf{y}, \mathbf{X}, \eta)$ depends only upon those $y_i$ for which

$$i \in J(m, \eta) := \{(p-1)um + 1, \cdots, (p-1)um + b_m(p-1)\}$$

we have

$$\begin{aligned} &|a_m(\mathbf{y}, \mathbf{X}, \eta) - a_m(\mathbf{y}', \mathbf{X}, \eta)| \\ &\quad \leq 4 b_m \| f \|_\infty (\exp 3 \| f \|_\infty b_m) \sum_{i \in J(m,\eta)} |y_i - y_i'|. \end{aligned} \tag{3.27}$$

Next, we combine this information with the smoothness of $\theta$. We know by definition of $L(\theta)$ that we can find numbers $d_m, m \leq k-1$ for which $\sum_{0 \leq m \leq k-1} d_m \leq L(\theta)$, such that, for all values of $x_0, \cdots, x_{k-1}, x_0', \cdots, x_{k-1}'$ in $[-1, 1]$, we have

$$|\theta(x_0, \cdots, x_{k-1}) - \theta(x_0', \cdots, x_{k-1}')| \leq \sum_{m \leq k-1} d_m |x_m - x_m'|.$$

Combining with (3.27), we get

$$\begin{aligned} &|\theta_1(\mathbf{y}) - \theta_1(\mathbf{y}')| \\ &\quad \leq 4 \sum_{m \leq k-1} d_m b_m \| f \|_\infty (\exp 3 \| f \|_\infty b_m) \sum_{i \in J(m,\eta)} |y_i - y_i'|. \end{aligned} \tag{3.28}$$

We take expectation to obtain

$$\begin{aligned} &|\theta'(y_1, \cdots, y_{k'}) - \theta'(y_{k'+1}, \cdots, y_{k'})| \\ &\quad \leq 4 \sum_{m \leq k-1} d_m \sum_{(p-1)um \leq i < (p-1)u(m+1)} c_i |y_i - y_i'| \end{aligned} \tag{3.29}$$

where

$$c_i = E 4 b_m \| f \|_\infty \exp 3 \| f \|_\infty b_m 1_{(p-1)b_m \geq i-(p-1)um} .$$

Thus

$$\sum_{(p-1)um < i \leq (p-1)u(m+1)} c_i \leq 4(p-1) \| f \|_\infty E(b_m^2 \exp 3 \| f \|_\infty b_m)$$

and

$$L(\theta') \le \sum_{m\le k-1} d_m \sum_{(p-1)um<i\le(p-1)u(m+1)} c_i$$
$$\le \Big(\sum_{m\le k-1} d_m\Big) 4(p-1)\|f\|_\infty E(b_m^2 \exp 3\|f\|_\infty b_m).$$

Since $\sum_{m\le k-1} d_m = L(\theta)$, we will be done if we show that

$$\|f\|_\infty K(\gamma, p) < 1 \Rightarrow \|f\|_\infty E(b_m^2 \exp 3\|f\|_\infty b_m) \le \frac{1}{8(p-1)}. \tag{3.30}$$

Now,

$$E \exp b_m = \Big(1 + \frac{\gamma}{N^{p-1}}(e-1)\Big)^{\binom{N-2k}{p-1}}$$
$$\le \exp \gamma(e-1)$$

so that (3.30) is obvious, and Proposition 3.3 is proved. □

## 4. Construction of the limiting probability

We consider the function $f$ as fixed once and for all, so that dependence in $f$ will not be indicated. Given $\gamma > 0$, we construct a map $T_\gamma$ from $M_1$ to $M_1$ as follows. Given $\nu$ in $M_1$ we consider an i.i.d. sequence $\mathbf{Z} = (Z_i)_{i\ge1}$ distributed like $\nu$. We consider the product measure $\mu_Z$ on $\{-1, 1\}^{\mathbb{N}}$ such that $\int \sigma_i d\mu_Z(\sigma_i) = Z_i$. We consider a Poisson r.v. $b$ of expectation $\gamma/(p-1)!$, and we assume that $b$ and $\mathbf{Z}$ are independent. We consider the random variable

$$Y = \frac{\mu_{\mathbf{Z}}(\underset{\epsilon=\pm1}{\mathrm{Av}}\ \epsilon \exp h)}{\mu_{\mathbf{Z}}(\underset{\epsilon=\pm1}{\mathrm{Av}} \exp h)} \tag{4.1}$$

where

$$h = \sum_{q\le b} f(X_q, \sigma_{(q-1)(p-1)+1}, \cdots, \sigma_{q(p-1)}, \epsilon) \tag{4.2}$$

and where the sequence $(X_q)$ is i.i.d. uniform, independent of $\mathbf{Z}$, $b$. We then define $T_\gamma(\nu)$ as the law of $Y$.

**Lemma 4.1.** *If $\gamma \le \gamma_0$ and*

$$\|f\|_\infty K(\gamma_0, p) \le 1 \ , \tag{4.3}$$

*then for all $\nu, \nu'$ in $M_1$*

$$d(T_\gamma(\nu), T_\gamma(\nu')) \le \frac{1}{2} d(\nu, \nu') \tag{4.4}$$

**Corollary 4.2.** *Under the condition of Lemma 4.1, there exists a unique probability* $Q_\gamma$ *such that*

$$T_\gamma(Q_\gamma) = Q_\gamma .$$

*Proof.* $(M_1, d)$ is a complete metric space. □

*Proof of Lemma 4.1.* Consider another i.i.d. sequence $\mathbf{Z}' = (Z'_i)$, and define $Y'$ as in (4.1), replacing $\mathbf{Z}$ by $\mathbf{Z}'$.

Using (3.36), we see as in (3.27) that

$$|Y - Y'| \le 4b\|f\|_\infty(\exp(3\|f\|_\infty b)) \sum_{i\le b(p-1)} |Z_i - Z'_i|.$$

Since $b$ is independent of $Z_i$, taking expectation, we see under (4.3) (mimicking the estimates of the previous section following (3.29)) that

$$E|Y - Y'| \le \sum c_i E|Z_i - Z'_i| \tag{4.5}$$

where $\sum_{i\ge 1} c_i \le 1/2$. We then use the fundamental property of the distance $d$: we can assume that $E|Z_i - Z'_i| = d(\nu, \nu')$, while $\mathscr{L}(Z_i) = \nu$, $\mathscr{L}(Z'_i) = \nu'$; Thus (4.5) gives $E|Y - Y'| \le \frac{1}{2} d(\nu, \nu')$. Since $\mathscr{L}(Y) = T_\gamma(\nu)$, $\mathscr{L}(Y') = T_{\gamma'}(\nu')$, this proves (4.4). □

**Lemma 4.3.** *We have*

$$\sup_\nu d(T_\gamma(\nu), T_{\gamma'}(\nu)) = o(\gamma - \gamma')$$

*where* $o$ *denotes a function such that* $\lim_{t\to 0_+} o(t) = 0.$

*Proof.* This is a consequence of the obvious fact that given $\gamma, \gamma'$, we can find two Poisson r.v. $b, b'$, with

$$Eb = \gamma/p!, Eb' = \gamma'/p!, \quad P(b \ne b') = o(\gamma - \gamma').$$ □

**Corollary 4.4.** *We have*

$$d(Q_\gamma, Q_{\gamma'}) = o(\gamma - \gamma').$$

*Proof.* We have

$$\begin{aligned} d(Q_\gamma, Q_{\gamma'}) &= d(T_\gamma(Q_\gamma), T_{\gamma'}(Q_{\gamma'})) \\ &\le d(T_\gamma(Q_\gamma), T_\gamma(Q_{\gamma'})) + d(T_\gamma(Q_{\gamma'}), T_{\gamma'}(Q_{\gamma'})) \\ &\le \frac{1}{2} d(Q_\gamma, Q_{\gamma'}) + d(T_\gamma(Q_{\gamma'}), T_{\gamma'}(Q_{\gamma'})) \end{aligned}$$

using Lemma 4.1, so that

$$d(Q_\gamma, Q_{\gamma'}) \le 2d(T_\gamma(Q_{\gamma'}), T_{\gamma'}(Q_{\gamma'})) \ ,$$

and the conclusion by Lemma 4.3. □

We denote by $\langle \ \cdot \ \rangle_0$ Gibbs' measure on $\{-1, 1\}^{N-1}$ with parameter $\gamma' = \gamma(1 - N^{-1})$. We denote by $\mathscr{L}_0(\langle\sigma_N\rangle)$ the conditional law of $\langle\sigma_N\rangle$ given $\langle \ \cdot \ \rangle_0$

**Proposition 4.5.** *If we assume $\|f\|_\infty K(\gamma_0, p) \le 1$, then, uniformly in $\gamma \le \gamma_0$ we have*

$$\lim_{N\to\infty} Ed(\mathscr{L}_0(\langle\sigma_N\rangle), T_\gamma(\mathscr{L}(\langle\sigma_N\rangle_0))) = 0 \tag{4.5}$$

*and in particular*

$$\lim_{N\to\infty} d(\mathscr{L}(\langle\sigma_N\rangle), T_\gamma(\mathscr{L}(\langle\sigma_N\rangle_0)) = 0. \tag{6.7}$$

*Proof.* It seems better here to be slightly informal rather than being unreadable. We start with relation

$$\langle\sigma_N\rangle = \frac{\langle \underset{\epsilon=\pm1}{Av}\ \epsilon \exp h'\rangle_0}{\langle \underset{\epsilon=\pm1}{Av}\ \exp h'\rangle_0} \tag{4.7}$$

where

$$h' = \sum \eta_{I,0} f(X_{I,0}, \sigma_{i(I,s)}, \cdots, \sigma_{i(I,p-1)}, \epsilon)\ ,$$

where the summation is of course over $I \in [N-1]^{p-1}$. According to the results of Section 2, we make only a vanishing error if we replace in (4.7) the average with respect to Gibbs' measure $\langle\ \cdot\ \rangle_0$ by the average for the product measure $\mu_{\mathbf{y}}$ on $\{-1,1\}^{N-1}$ where $\mathbf{y} = (\langle\sigma_i\rangle_0)_{i\le N-1}$. Thus, the law of $\langle\sigma_N\rangle$ asymptotically resembles the law of

$$\frac{\mu_{\mathbf{y}}(\underset{\epsilon=\pm1}{Av}\ \epsilon \exp h')}{\mu_{\mathbf{y}}(\underset{\epsilon=\pm1}{Av}\ \exp h')}. \tag{4.8}$$

Let us define

$$b_N = \sum_I \eta_{I,0} \tag{4.9}$$

and

$$h = \sum_{q\le b_N} f(X_q, \sigma_{(q-1)(p-1)+1}, \cdots, \sigma_{q(p-1)}, \epsilon). \tag{4.10}$$

Consider a uniform random permutation $\tau$ of $\{1, \cdots, N-1\}$, independent of all the other r.v. considered. Let us write

$$\tau h = \sum_{q\le b_N} f(X_q, \sigma_{\tau((q-1)(p-1)+1)}, \cdots, \sigma_{\tau(q(p-1))}\epsilon).$$

Consider the quantity

$$\frac{\mu_{\mathbf{y}}(\underset{\epsilon=\pm1}{Av}\ \epsilon \exp \tau h)}{\mu_{\mathbf{y}}(\underset{\epsilon=\pm1}{Av}\ \exp \tau h)}. \tag{4.11}$$

If we fix $\langle\cdot\rangle_0$, $b_N$ and the $X_J$, that is, if the only randomness is that of $\tau$ and the $\eta_{I,0}$, we see that the random quantities (4.8) and (4.11) have asymptotically the same law. This is because the $b_N$ sets $I$ for which $\eta_{I,0} \ne 0$ are disjoint with a probability $\to 1$ as $N \to \infty$, and, conditionally upon the fact that they are disjoint, they are distributed like the family of the $b_N$ sets

$$\{\tau((q-1)(p-1)+1), \cdots, \tau(q(p-1))\}$$

for $q \leq b_N$. By permuting the indices in (4.11), we can see the random permutation as shuffling the indices of $\mathbf{y}$ rather than those of $h$. Thus, if we write $\tau(\mathbf{y}) = (\langle\sigma_{\tau(i)}\rangle_0)_{i\leq N-1}$ the law of (4.11), given $\langle\ \cdot\ \rangle_0$, is asymptotically the law of

$$\frac{\mu_{\tau(\mathbf{y})}(\underset{\epsilon=\pm 1}{Av}\ \epsilon \exp h)}{\mu_{\tau(\mathbf{y})}(\underset{\epsilon=\pm 1}{Av}\ \exp h)}. \tag{4.12}$$

The crucial point is that, conditionally upon $\langle\ \cdot\ \rangle_0$, as $N \to \infty$, any given number $n$ of components $\langle\sigma_{\tau(1)}\rangle_0, \cdots, \langle\sigma_{\tau(n)}\rangle_0$ of $\tau(\mathbf{y})$ asymptotically resembles a family of $n$ i.i.i. r.v. of law $\xi = (N-1)^{-1} \sum_{i\leq N-1} \delta_{\langle\sigma_i\rangle_0}$. Thus, asymptotically we can replace (4.12) by

$$\frac{\mu_{\mathbf{z}}(\underset{\epsilon=\pm 1}{Av}\ \epsilon \exp h)}{\mu_{\mathbf{z}}(\underset{\epsilon=\pm 1}{Av}\ \exp h)} \tag{4.13}$$

where $\mathbf{z} = (z_i)_{i\leq N-1}$, and $(z_i)_{i\leq N-1}$ are i.i.d of law $\xi$. We proved in Section 3 that (in probability) $\xi$ is close to $\mathscr{L}(\langle\sigma_1\rangle_0)$ so that in (4.13) we can assume that $(\mathbf{z}_i)_{i\leq N-1}$ are i.i.d of law $\mathscr{L}(\langle\sigma_1\rangle_0)$. Moreover, as $N \to \infty$ goes to $\infty$, $b_N$ converges in law to a Poisson r.v. $b$ with $Eb = \gamma/(p-1)!$. This completes the argument. □

**Theorem 4.6.** *Under (4.2), we have*

$$\lim_{N\to\infty} \mathscr{L}(\langle\sigma_1\rangle) = Q_\gamma.$$

*Proof.* We have

$$\begin{aligned} d(\mathscr{L}(\langle\sigma_1\rangle), Q_\gamma) &= d(\mathscr{L}(\langle\sigma_N\rangle), Q_\gamma) \\ &= d(\mathscr{L}(\langle\sigma_N\rangle), T_\gamma(Q_\gamma)) \\ &\leq d(\mathscr{L}(\langle\sigma_N\rangle), T_\gamma(\mathscr{L}(\langle\sigma_1\rangle_0))) \\ &\quad + d(T_\gamma(\mathscr{L}(\langle\sigma_1\rangle_0)), T_\gamma(Q_\gamma)). \end{aligned}$$

Using Lemma 4.1, we have

$$\begin{aligned} &d(T_\gamma(\mathscr{L}(\langle\sigma_1\rangle_0)), T_\gamma(Q_\gamma)) \\ &\leq \frac{1}{2} d(\mathscr{L}(\langle\sigma_1\rangle_0), Q_\gamma) \\ &\leq \frac{1}{2} d(\mathscr{L}(\langle\sigma_1\rangle_0), Q_{\gamma'}) + \frac{1}{2} d(Q_\gamma, Q_{\gamma'}). \end{aligned}$$

Thus

$$d(\mathscr{L}(\langle\sigma_1\rangle), Q_\gamma) \leq \frac{1}{2} d(\mathscr{L}(\langle\sigma_1\rangle_0), Q_{\gamma'}) + R$$

where $R \to 0$ as $N \to \infty$ (by Proposition 4.5 and Lemma 4.1).

Thus, if we fix $\gamma_0$ such that (4.2) holds, and set

$$a_N = \sup_{\gamma\leq\gamma_0} d(\mathscr{L}(\langle\sigma_N\rangle), Q_\gamma)\ ,$$

we deduce from (4.8) that

$$\limsup_N a_N \leq \frac{1}{2} \limsup a_{N-1}$$

so that $a_N \to 0$. □

In order to prove Theorem 1.1, it remains now to show that, given any number $k$, the r.v. $(\langle\sigma_i\rangle)_{i\leq k}$ are asymptotically independent.

If we combine Proposition 4.5, Theorem 4.6, and Corollary 4.4, we see that $Ed(\mathscr{L}_0(\langle\sigma_N\rangle), Q_\gamma) \to 0$ i.e. the conditional law of $\langle\sigma_N\rangle$ given $\langle\ \cdot\ \rangle_0$ is essentially $Q_\gamma$. We observe that there is very little to change to the proof of Proposition 4.5 to prove that the conditional law of $\langle\sigma_N\rangle$ given the Gibbs' measure $G_{N-r}$ on $N-r$ sites also converges in probability to $Q_\gamma$. Thus, it suffices to prove that $\langle\sigma_{N-r+1}\rangle, \cdots, \langle\sigma_N\rangle$ are asymptotically independent given $G_{N-r}$. This is however obvious from (3.11), (3.14). □

To conclude, let us now see how to compute the limiting free energy per site. First, (considering as usual $f$ as fixed) we show that, setting $F_N(\gamma) = \log Z_N(\gamma)$, we have

$$\frac{1}{N}\frac{d}{d\gamma} EF_N(\gamma) = \frac{1}{p!} E \log\langle \exp f(X, \sigma_1, \cdots, \sigma_p)\rangle + o(1). \tag{4.14}$$

To see this, we consider a new independent sequence $\eta'_J$ of r.v., $\eta'_J \in \{0,1\}$, $P(\eta'_J = 1) = \gamma' N^{1-p}$, and we assume that this sequence is independent of all other sequences. Thus $\eta''_J = \max(\eta_J, \eta'_J)$ satisfies

$$P(\eta''_J = 1) = \frac{\gamma}{N^{p-1}} + \frac{\gamma'}{N^{p-1}} - \frac{\gamma\gamma'}{N^{2(p-1)}}.$$

Thus

$$\log \sum_{\boldsymbol{\sigma}} \exp \sum_J \eta''_J f(X_J, \sigma_{i(J,1)}, \cdots, \sigma_{i(J,p)}) \tag{4.15}$$

is distributed like $Z_N(\gamma + \gamma'(1 - \frac{\gamma}{N^{p-1}}))$.
Let $\eta'''_J = \eta''_J - \eta_J$, so that we have the identity

$$\sum_{\boldsymbol{\sigma}} \exp \sum_J \eta''_J f(X_J, \sigma_{i(J,1)}, \cdots, \sigma_{i(J,p)})$$
$$= Z_N(\gamma)\langle \exp \sum_J \eta'''_J f(X_J, \sigma_{i(J,1)}, \cdots, \sigma_{i(J,p)})\rangle$$

and thus

$$EF_N\left(\gamma + \gamma'\left(1 - \frac{\gamma}{N^{p-1}}\right)\right) \tag{4.11}$$
$$= EF_N(\gamma) + E \log\left\langle \exp \sum_J \eta'''_J f(X_J, \sigma_{i(J,1)}, \cdots, \sigma_{i(J,p)})\right\rangle.$$

The last term is zero if no $\eta_J'''$ is equal to 1; if exactly one $\eta_J'''$ is equal to 1, it is

$$E \log\langle \exp f(X, \sigma_1, \cdots, \sigma_p)\rangle$$

by symmetry upon the sites. For $\gamma' << 1/N$, the probability that one exactly $\eta_J'''$ is 1 is (at the first order in $\gamma'$)

$$\frac{\gamma'}{N^{p-1}}\left(\binom{N}{p} - \sum \eta_J\right)$$

which, as $N$ increases, behaves like $N\gamma'/p!$ Together with (4.16) this proves (4.14). Now we deduce from Theorem 1.1 that

$$\lim_{N\to\infty} \frac{\partial}{\partial\gamma}\frac{1}{N} EF_N(\gamma) = \frac{1}{p!} E \log \int f(X, \sigma_1, \cdots, \sigma_p) d\nu_Y$$

where $\nu_Y$ is the product measure on $\{-1, 1\}^p$ for which $\nu(\sigma_i) = Y_i$, and where $(Y_i)_{i\le p}$ is i.i.d. distributed like $Q_\gamma$. This proves Theorem 1.2, since $F_N(0) = N \log 2$.

## References


[M-P-V] Mézard, M., Parisi, G., Virasiro, M.: Spin glass theory and beyond, World Scientific, Singapore, (1987)

[M-Z] Monasson, R., Zecchina, R.: Statistical Mechanics of the Random $K$-sat model, Phys. Rev. B, 1352–1370 (1997)

[T1] Talagrand, M.: The Sherrington Kirkpatrick model: a challenge for mathematicians Probab. Theory Rel. Fields., **110**, 109–176 (1998)

[T2] Talagrand, M.: Rigorous results for the Hopfield Model, Probab. Theory Rel. Fields., **110**, 177–276 (1998)

[T3] Talagrand, M.: Intersecting random half-spaces: towards the Gardner-Nerrida formula, Ann. Probab. (to be published)

[T4] Talagrand, M.: Huge Random Structures and mean field models for spin glasses, Proceedings of the Berlin International Congress of Mathematics, Documenta Math., Extra Vol. I. (1998) (electronic)