A diluted version of the perceptron model

David Márquez-Carreras[1], **Carles Rovira**[1] and **Samy Tindel**[2]

[1] *Facultat de Matemàtiques, Universitat de Barcelona,*
Gran Via 585, 08007-Barcelona, Spain
e-mail: davidmarquez@ub.edu, carles.rovira@ub.edu

[2] *Institut Elie Cartan, Université de Nancy 1,*
BP 239, 54506-Vandoeuvre-lès-Nancy, France
e-mail: tindel@iecn.u-nancy.fr

Abstract

This note is concerned with a diluted version of the perceptron model. We establish a replica symmetric formula at high temperature, which is achieved by studying the asymptotic behavior of a given spin magnetization. Our main task will be to identify the order parameter of the system.

Keywords: spin glasses, perceptron model, magnetization, order parameter.

MSC: 60G15, 82D30

[1]Partially supported by MEC grant MTM2006-01351.

1 Introduction

A wide number of spectacular advances have occurred in the spin glasses theory during the last past years, and it could easily be argued that this topic, at least as far as the Sherrington-Kirkpatrick model is concerned, has reached a certain level of maturity from the mathematical point of view: the cavity method has been set in a clear and effective way in [10], some monotonicity properties along a smart path have been discovered in [5], and these elements have been combined in [11] in order to obtain a completely rigorous proof of the Parisi solution [8].

However, there are some canonical models of mean field spin glasses for which the basic theory is far from being complete, and this paper proposes to study the high temperature behavior of one of them, namely the diluted perceptron model, which can be described as follows: for $N \geq 1$, consider the configuration space $\Sigma_N = \{-1, 1\}^N$, and for $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \Sigma_N$, define a Hamiltonian $-H_{N,M}(\boldsymbol{\sigma})$ by

$$-H_{N,M}(\boldsymbol{\sigma}) = \sum_{k \leq M} \eta_k \; u\left(\sum_{i \leq N} g_{i,k} \; \gamma_{i,k} \; \sigma_i\right). \tag{1}$$

In this Hamiltonian, M stands for a positive integer such that $M = \alpha N$ for a given $\alpha \in (0, 1)$; u is a bounded continuous function defined on $\mathbb{R}$; $\{g_{i,k}, \; i \geq 1, k \geq 1\}$ and $\{\gamma_{i,k}, \; i \geq 1, k \geq 1\}$ are two independent families of independent random variables, $g_{i,k}$ following a standard Gaussian law and $\gamma_{i,k}$ being a Bernoulli random variable with parameter $\frac{\gamma}{N}$, which we denote by $B(\frac{\gamma}{N})$. Finally, $\{\eta_k, \; k \geq 1\}$ stands for an arbitrary family of numbers, with $\eta_k \in \{0, 1\}$, even if the case of interest for us will be $\eta_k = 1$ for all $k \leq M$. Associated to this Hamiltonian, define a random Gibbs measure G_N on Σ_N, whose density with respect to the uniform measure μ_N is given by $Z_{N,M}^{-1} \exp(-H_{N,M}(\boldsymbol{\sigma}))$, where the partition function $Z_{N,M}$ is defined by

$$Z_{N,M} = \sum_{\boldsymbol{\sigma} \in \Sigma_N} \exp\left(-H_{N,M}(\boldsymbol{\sigma})\right).$$

In the sequel, we will denote by $\langle f \rangle$ the average of a function $f : \Sigma_N^n \longrightarrow \mathbb{R}$ with respect to $dG_N^{\otimes n}$, i.e.

$$\langle f \rangle = Z_{N,M}^{-n} \sum_{(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \in \Sigma_N^n} f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \exp\left(-\sum_{l \leq n} H_{N,M}(\boldsymbol{\sigma}^l)\right).$$

The measure described above is of course a generalization of the usual perceptron model, which has been introduced for neural computation purposes (see [6]), and whose high temperature behavior has been described in [10, Chapter 3], or [9] for an approach based on convexity properties of the Hamiltonian. Indeed the usual perceptron model is induced by a Hamiltonian $\hat{H}_{N,M}$ on Σ_N given by

$$-\hat{H}_{N,M}(\boldsymbol{\sigma}) = \sum_{k\leq M} u\left(\frac{1}{N^{1/2}}\sum_{i\leq N} g_{i,k}\ \sigma_i\right), \tag{2}$$

where we have kept the notations introduced for equation (1). Thus, our model can be seen as a real diluted version of (2), in the sense that in our model, each condition $\sum_{i\leq N} g_{i,k}\gamma_{i,k}\sigma_i \geq 0$ only involves, in average, a finite number of spins, uniformly in N. It is worth noticing at that point that this last requirement fits better to the initial neural computation motivation, since in a one-layer perceptron, an output is generally obtained by a threshold function applied to a certain number of spins, that does not grow linearly with the size of the system. Furthermore, our coefficient γ is arbitrarily large, which means that the global interaction between spins is not trivial. Another motivation for the study of the system induced by (1) can be found in [3]. Indeed, in this latter article, a social interaction model is proposed, based on a Hopfield-like (or perceptron-like) diluted Hamiltonian with parameter N and M, where N represents the number of social agents, and M the diversity of these agents, the number of interactions of each agent varying with the dilution parameter. However, in [3], the equilibrium of the system is studied only when M is a fixed number. The result we will explain later on can thus be read as follows: as soon as the diversity M does not grow faster than a small proportion of N, the capacity of the social interaction system is not attained

Let us turn now to a brief description of the results contained in this paper: in fact, we will try to get a replica symmetric formula for the system when M is a small proportion of N, which amounts to identify the limit of $\frac{1}{N}\log(Z_{N,M})$ when $N \to \infty$, $M = \alpha N$. This will be achieved, as in the diluted SK model studied through the cavity method (see [4] for a study based on monotonicity methods), once the limiting law for the magnetization $\langle\sigma_i\rangle$ is obtained. This will thus be our first aim, and in order to obtain that result, we will try to adapt the method designed in [10, Chapter 7].

However, in our case, the identification of the limiting law for $\langle\sigma_i\rangle$ will be done through an intricate fixed point argument, involving a map $T : \mathbf{P} \to \mathbf{P}$ (where $\mathbf{P}$ stands for the set of probability measures on $[-1, 1]$), which in turn involves a kind of $\mathcal{P}(\lambda)^{\otimes\mathcal{P}(\mu)}$ measure, for two independent Poisson measures $\mathcal{P}(\lambda)$ and $\mathcal{P}(\mu)$. Notice that this kind of complexity, inherent to diluted inhomogeneous systems, is also illustrated e.g. in the context of random assignments in [1]. For sake of readability, we will give the details of (almost) all the computations we will need in order to establish our replica symmetric formula, but it should be mentioned at that point that our main contribution, with respect to [10, Chapter 7], is that construction of the invariant measure.

More specifically, our paper is divided as follows:

- At Section 2, we will establish a decorrelation result for two arbitrary spins. Namely, setting $U_\infty = \|u\|_\infty$, for αU_∞ small enough, we will show that
$$\mathbf{E}\left[|\langle\sigma_1\sigma_2\rangle - \langle\sigma_1\rangle\langle\sigma_2\rangle|\right] \leq \frac{K}{N},$$
for a constant $K > 0$.

- At Section 3, we will study the asymptotic behavior of the magnetization of m spins, where m is an arbitrary integer. Here again, if αU_∞ is small enough, and in the particular case of interest where all $\eta_k = 1$, we will see that
$$\mathbf{E}\left[\sum_{i\leq m} |\langle\sigma_i\rangle - z_i|\right] \leq \frac{Km^3}{N},$$
where $z_1, \ldots, z_m$ is a family of i.i.d random variable, with law $\mu_{\alpha,\gamma}$, and $\mu_{\alpha,\gamma}$ is the fixed point of the map T alluded to above, whose precise description will be given at the beginning of Section 3.

- Finally, at Section 4, we obtain the replica symmetric formula for our model (where all $\eta_k = 1$): set
$$\begin{aligned}\bar{V}_p &= \int \Big\langle \exp(u(\sum_{i\leq p} g_{i,M}\sigma_i)) \Big\rangle_{(x_1,\ldots,x_p)} d\mu_{\alpha,\gamma}(x_1)\cdots d\mu_{\alpha,\gamma}(x_p) \\ G(\gamma) &= \alpha \log\left(\sum_{p=0}^{\infty} \exp(-\gamma)\frac{\gamma^p}{p!}\mathbf{E}\left[\frac{\bar{V}_{p+1}}{\bar{V}_p}\right]\right),\end{aligned}$$

where $\langle\cdot\rangle_x$ means integration with respect to the product measure ν on $\{-1,1\}^p$ such that $\int \sigma_i d\nu = x_i$. Let $F:[0,1]\to\mathbb{R}^+$ be defined by $F(0)=\log 2-\alpha u(0)$ and $F'(\gamma)=G(\gamma)$. Then, if αU_∞ is small enough, we will get that

$$\left|\frac{1}{N}\mathbf{E}\left[\log(Z_{N,M})\right]-F(\gamma)\right|\leq\frac{K}{N},$$

for a finite constant K.

All these results will be described in greater detail in the corresponding sections.

2 Spin correlations

As in [10, Chapter 7], the first step towards a replica symmetric formula will be to establish a decorrelation result for two arbitrary spins in the system. However, a much more general property holds true, and we will turn now to its description: for $j\leq N$, let T_j be the transformation of Σ_N^n that, for a configuration $(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)$ in Σ_N^n, exchanges the j-th coordinates of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$. More specifically, let $f:\Sigma_N^n\to\mathbb{R}$, with $n\geq 2$, and let us write, for $j\leq N$,

$$f=f\left(\boldsymbol{\sigma}_{j^c}^1,\sigma_j^1;\boldsymbol{\sigma}_{j^c}^2,\sigma_j^2;\ldots;\boldsymbol{\sigma}_{j^c}^n,\sigma_j^n\right),$$

where, for $l=1,\ldots,n$, $\boldsymbol{\sigma}_{j^c}^l=(\sigma_1^l\ldots,\sigma_{j-1}^l,\sigma_{j+1}^l,\ldots,\sigma_N^l)$. Then define $f\circ T_j$ by

$$f\circ T_j(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n)=f\left(\boldsymbol{\sigma}_{j^c}^1,\sigma_j^2;\boldsymbol{\sigma}_{j^c}^2,\sigma_j^1;\ldots;\boldsymbol{\sigma}_{j^c}^n,\sigma_j^n\right). \quad (3)$$

For $j\leq N-1$, we will call U_j the equivalent transformation on Σ_{N-1}^n.

Definition 2.1 *We say that Property* $\boldsymbol{P}(N,\gamma_0,B)$ *is satisfied if the following requirement is true: let* f *and* f' *be two functions on* Σ_N^n *depending on* m *coordinates, such that* $f\geq 0$, $f'\circ T_N=-f'$, *and there exists* $Q\geq 0$ *such that* $|f'|\leq Qf$*; then if* $\gamma\leq\gamma_0$ *we have*

$$\mathbf{E}\left|\frac{\langle f'\rangle}{\langle f\rangle}\right|\leq\frac{mQB}{N},$$

for any Hamiltonian of the form (1), uniformly in η*.*

Set now $U_\infty = \|u\|_\infty$. With Definition 2.1 in hand, one of the purposes of this section is to prove the following Theorem.

Theorem 2.2 *Let γ_0 be a positive number, and U_∞ be small enough, so that*

$$4U_\infty \ \alpha\gamma_0^2 \ e^{4U_\infty} \ e^{\alpha\gamma_0(e^{4U_\infty}-1)} \ \left(3 + 2\gamma_0 + \alpha(\gamma_0^2 + \gamma_0^3)e^{4U_\infty}\right) < 1. \qquad (4)$$

Then there exists a number $B_0(\gamma_0, U_\infty)$ such that if $\gamma \leq \gamma_0$, the property $\boldsymbol{P}(N, \gamma_0, B_0)$ holds true for each $N \geq 1$.

In the previous theorem, notice that the value of γ_0 has been picked arbitrarily. Then we have to choose U_∞, which also contains implicitly the temperature parameter, accordingly. Let us also mention that the spin decorrelation follows easily from the last result:

Corollary 2.3 *Assuming (4) there exists $K > 0$ such that, for all $\gamma < \gamma_0$,*

$$\mathbf{E}\,|\langle\sigma_1\sigma_2\rangle - \langle\sigma_1\rangle\langle\sigma_2\rangle| \leq \frac{K}{N}.$$

Proof: It is an easy consequence of property $\mathbf{P}(N, \gamma_0, B_0)$ applied to $n = 2$, $f = 1$ and $f'(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) = \sigma_1^1(\sigma_2^1 - \sigma_2^2)$.

□

We will prepare now the ground for the proof of Theorem 2.2, which will be based on an induction argument over N. A first step in this direction will be to state the cavity formula for our model: for $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \Sigma_N$, we write

$$\boldsymbol{\rho} \equiv (\rho_1, \ldots, \rho_{N-1}) = (\sigma_1, \ldots, \sigma_{N-1}) \in \Sigma_{N-1}.$$

Then the Hamiltonian (1) can be decomposed into

$$-H_{N,M}(\boldsymbol{\sigma}) = \sum_{k\leq M} \eta_k \gamma_{N,k} u \left(\sum_{i\leq N-1} g_{i,k}\gamma_{i,k}\sigma_i + g_{N,k}\sigma_N \right) - H^-_{N-1,M}(\boldsymbol{\rho}),$$

with

$$-H^-_{N-1,M}(\boldsymbol{\rho}) = \sum_{k\leq M} \eta_k^- u \left(\sum_{i\leq N-1} g_{i,k}\gamma_{i,k}\sigma_i \right), \quad \text{and} \quad \eta_k^- = \eta_k(1-\gamma_{N,k}). \quad (5)$$

Note that in $H^-_{N-1,M}$, the coefficients $\eta^-_k = \eta_k(1-\gamma_{N,k})$ are not deterministic, and hence $H^-_{N-1,M}$ is not really of the same kind as $H_{N,M}$. However, this problem can be solved by conditioning on $\{\gamma_{N,k}, k \le M\}$. Then, given the randomness contained in the $\gamma_{N,k}$, the expression $H^-_{N-1,M}(\boldsymbol{\rho})$ is a Hamiltonian of a $(N-1)$-spin system with $\gamma_{i,k} \sim B(\frac{\gamma^-}{N-1})$, where $\gamma^- = \gamma\frac{N-1}{N}$ and so $\gamma^- \le \gamma \le \gamma_0$.

Thus, given a function $f : \Sigma_N^n \longrightarrow \mathbb{R}$, we easily get the following decomposition of the mean value of f with respect to $G_N^{\otimes n}$:

$$\langle f \rangle = \frac{\langle \mathbf{Av} f \xi \rangle_-}{\langle \mathbf{Av} \xi \rangle_-}, \tag{6}$$

with

$$\xi = \exp\left(\sum_{l \le n} \sum_{k \le M} \eta_k \gamma_{N,k} u \left(\sum_{i \le N-1} g_{i,k}\gamma_{i,k}\sigma_i^l + g_{N,k}\sigma_N^l \right)\right), \tag{7}$$

and with $\langle \bar{f} \rangle_-$ defined, for a given $\bar{f} : \Sigma_{N-1}^n \to \mathbb{R}$, by

$$\langle \bar{f} \rangle_- = \frac{\sum_{(\boldsymbol{\rho}^1,\ldots,\boldsymbol{\rho}^n)\in\Sigma_{N-1}^n} \bar{f}(\boldsymbol{\rho}^1,\ldots,\boldsymbol{\rho}^n) \exp\left(-\sum_{l\le n} H^-_{N-1,M}(\boldsymbol{\rho}^l)\right)}{\sum_{(\boldsymbol{\rho}^1,\ldots,\boldsymbol{\rho}^n)\in\Sigma_{N-1}^n} \exp\left(-\sum_{l\le n} H^-_{N-1,M}(\boldsymbol{\rho}^l)\right)}.$$

Notice also that in expression (6), $\mathbf{Av}$ stands for the average with respect to the last component of the system, namely if $f = f(\boldsymbol{\rho}^1, \sigma_N^1, \ldots, \boldsymbol{\rho}^n, \sigma_N^n)$, then

$$\mathbf{Av} f(\boldsymbol{\rho}^1 \ldots, \boldsymbol{\rho}^n) = \frac{1}{2^n} \sum_{\sigma_N^j = \pm 1, j \le n} f(\boldsymbol{\rho}^1, \sigma_N^1, \ldots, \boldsymbol{\rho}^n, \sigma_N^n).$$

Let us introduce now a little more notation: in the sequel we will have to take expectations for a fixed value of ξ given at (7). Let us denote thus by $\mathbf{E}_{\gamma_N}$ the expectation given $\gamma_{N,k},\ k \le M$, and define

$$\mathbf{E}_{-,\gamma_N}[\,\cdot\,] = \mathbf{E}_{\gamma_N}\left[\,\cdot \mid g_{N,k},\ g_{i,k},\ \gamma_{i,k},\ i \le N-1,\ k \in D^M_{N,1}\,\right], \tag{8}$$

where $D^M_{N,1}$ is given by

$$D^M_{N,1} = \{k \le M :\ \gamma_{N,k} = 1\}.$$

One has to be careful about the way all these conditioning are performed, but it is worth observing that the set $D_{N,1}^M$ is not too large: indeed, it is obvious that, setting $|A|$ for the size of a set A, we have

$$|D_{N,1}^M| = \sum_{k\le M} \gamma_{N,k}, \tag{9}$$

and thus

$$\mathbf{E}|D_{N,1}^M| = M\frac{\gamma}{N} = \alpha\gamma.$$

Let us go on now with the first step of the induction procedure for the proof of Theorem 2.2: in $\mathbf{P}(N,\gamma_0,B)$ we can assume without loss of generality that f and f' depend on the coordinates $1,\dots,m-1,N$. Moreover, since $|f'\xi| \le Qf\xi$, we have

$$|\langle \mathbf{Av} f'\xi\rangle_-| \le \langle \mathbf{Av}|f'\xi|\rangle_- \le \langle Q\mathbf{Av} f\xi\rangle_-,$$

and hence

$$\left|\frac{\langle \mathbf{Av} f'\xi\rangle_-}{\langle \mathbf{Av} f\xi\rangle_-}\right| \le Q. \tag{10}$$

We now define the following two events:

$$\begin{aligned}
\Omega_1 &= \{\exists p\le m-1, k\in D_{N,1}^M : \gamma_{p,k}=1\} \\
&= \{\exists p\le m-1, k\le M : \gamma_{p,k}=\gamma_{N,k}=1\}, \\
\Omega_2 &= \{\exists j\le N-1, k_1,k_2\in D_{N,1}^M : \gamma_{j,k_1}=\gamma_{j,k_2}=1\} \\
&= \{\exists j\le N-1, k_1,k_2\le M : \gamma_{j,k_1}=\gamma_{j,k_2}=\gamma_{N,k_1}=\gamma_{N,k_2}=1\}.
\end{aligned}$$

These two events can be considered as exceptional. Indeed, it is readily checked that

$$P(\Omega_1) \le \alpha\frac{\gamma^2}{N}(m-1), \qquad P(\Omega_2) \le \alpha^2\gamma^4\frac{N-1}{N^2}.$$

Thus, if $\Omega = \Omega_1\cup\Omega_2$, we get

$$P(\Omega) \le \frac{\alpha\gamma^2(m-1)+\alpha^2\gamma^4}{N},$$

and using this fact together with (10), we have

$$\begin{aligned}\mathbf{E}\left|\frac{\langle f'\rangle}{\langle f\rangle}\right| &= \mathbf{E}\left|\frac{\langle \mathbf{Av} f'\xi\rangle_-}{\langle \mathbf{Av} f\xi\rangle_-}\right| \\ &= \mathbf{E}\left(\mathbf{1}_{\Omega}\left|\frac{\langle \mathbf{Av} f'\xi\rangle_-}{\langle \mathbf{Av} f\xi\rangle_-}\right|\right) + \mathbf{E}\left(\mathbf{1}_{\Omega^c}\left|\frac{\langle \mathbf{Av} f'\xi\rangle_-}{\langle \mathbf{Av} f\xi\rangle_-}\right|\right) \\ &\leq Q\frac{\alpha\gamma^2(m-1)+\alpha^2\gamma^4}{N} + \mathbf{E}\left(\mathbf{1}_{\Omega^c}\left|\frac{\langle \mathbf{Av} f'\xi\rangle_-}{\langle \mathbf{Av} f\xi\rangle_-}\right|\right). \qquad (11)\end{aligned}$$

Consequently, in order to prove Theorem 2.2 we only need to bound accurately the expectation of the right-hand side of (11) by means of the induction hypothesis. To this purpose, we will introduce some new notations and go through a series of lemmas: set

$$J_1 = \{j \leq N-1 :\ \gamma_{j,k} = 1 \,\text{for some}\, k \in D_{N,1}^M\},$$

and observe that, when Ω_1 does not occur,

$$J_1 \cap \{1, \ldots, m-1\} = \emptyset.$$

Denote $|J_1| = \text{card}(J_1)$ and write an enumeration of J_1 as follows: $J_1 = \{j_1, \ldots, j_{|J_1|}\}$.

Lemma 2.4 *Let U_j be the transformation defined at (3), and $f' : \Sigma_N^n \to \mathbb{R}$ such that $f' \circ T_N = -f'$. When Ω_1 does not occur, we have*

$$(\mathbf{Av} f'\xi) \circ \prod_{j\in J_1} U_j = -\mathbf{Av} f'\xi.$$

Proof: The proof of this lemma can be done following the steps of [10, Lemma 7.2.4], and we include it here for sake of readability. Set $T = \prod_{j\in J_1} T_j$. Since f' depends only on the coordinates $\{1, \ldots, m-1, N\}$ and this set is disjoint from J_1, we have $f' \circ T = f'$. Moreover,

$$f' \circ T \circ T_N = f' \circ T_N = -f'.$$

On the other hand, ξ only depends on $J_1 \cup \{N\}$ and using

$$\xi(\sigma^1, \sigma^2, \ldots, \sigma^n) = \xi(\sigma^2, \sigma^1, \ldots, \sigma^n),$$

we obtain

$$\xi \circ T \circ T_N = \xi.$$

Hence

$$(f'\xi) \circ T \circ T_N = -f'\xi,$$

and, since $T_N^2 = \mathrm{Id}$, we get

$$(f'\xi) \circ T = -(f'\xi) \circ T_N. \tag{12}$$

Finally,

$$\begin{aligned} \mathbf{Av}[(f'\xi) \circ T_N] &= \mathbf{Av} f'\xi, &(13)\\ \mathbf{Av}[(f'\xi) \circ T] &= (\mathbf{Av} f'\xi) \circ \prod_{j \in J_1} U_j. &(14) \end{aligned}$$

The proof is now easily concluded by plugging (13) and (14) into (12).

□

Let us now go on with the proof of Theorem 2.2: thanks to Lemma 2.4, when Ω_1 does not occur, we can write

$$\mathbf{Av} f'\xi = \frac{1}{2}\left[\mathbf{Av} f'\xi - (\mathbf{Av} f'\xi) \circ \prod_{s \le |J_1|} U_{j_s}\right] = \frac{1}{2} \sum_{1 \le s \le |J_1|} f_s, \tag{15}$$

with

$$f_s = (\mathbf{Av} f'\xi) \circ \prod_{l \le s-1} U_{j_l} - (\mathbf{Av} f'\xi) \circ \prod_{l \le s} U_{j_l}.$$

Notice that $U_j^2 = \mathrm{Id}$, and that f_s enjoys the same kind of antisymmetric property as f', since $f_s \circ U_{j_s} = -f_s$.

Define $R_1 = |D_{N,1}^M|$. Then, recalling relation (9), we have

$$R_1 = |D_{N,1}^M| = \sum_{k \le M} \gamma_{N,k},$$

and let us enumerate as $k_1, \ldots, k_{R_1}$ the values $k \le M$ such that $\gamma_{N,k} = 1$. We also define $I_1^1, \ldots, I_{R_1}^1$ as follows:

$$I_v^1 = \{j \le N-1 : \gamma_{j,k_v} = 1\}, \quad \text{for} \quad v \le R_1,$$

and observe that we trivially have

$$J_1 = \bigcup_{v \le R_1} I_v^1. \tag{16}$$

Moreover, when Ω_2 does not occur, we have

$$I_{v_1}^1 \cap I_{v_2}^1 = \emptyset, \quad \text{if } v_1 \neq v_2.$$

Then, on Ω^c, we get

$$|J_1| = \mathrm{Card}(J_1) = \sum_{v \le R_1} |I_v^1|. \tag{17}$$

Furthermore, it is easily checked that, for each v, and conditionally on the $\gamma_{N,k}$, the quantity $|I_v^1|$ is a binomial random variable with parameters $N-1$ and $\frac{\gamma}{N}$, which we denote by $\mathrm{Bin}(N-1, \frac{\gamma}{N})$.

With all these notations in mind, our next step will be to bound f_s in function of f, in order to get a similar condition to that of Definition 2.1:

Lemma 2.5 *Recall that $U_\infty = \|u\|_\infty$. Then, on Ω^c, for $j_s \in I_v^1$, we have*

$$|f_s| \le \hat{Q}\mathbf{Av} f \xi,$$

where

$$\hat{Q} \equiv 4QU_\infty \exp\left(4U_\infty R_1\right).$$

Proof: Let us decompose ξ as $\xi = \xi' \, \xi''$, with

$$\begin{aligned}
\xi' &= \exp\left(\sum_{3 \le m \le n} \sum_{k \le M} \eta_k \gamma_{N,k} u \left(\sum_{i \le N-1} g_{i,k}\gamma_{i,k}\sigma_i^m + g_{N,k}\sigma_N^m \right) \right), \\
\xi'' &= \exp\left(\sum_{m \le 2} \sum_{k \le M} \eta_k \gamma_{N,k} u \left(\sum_{i \le N-1} g_{i,k}\gamma_{i,k}\sigma_i^m + g_{N,k}\sigma_N^m \right) \right). \qquad (18)
\end{aligned}$$

Thus

$$\begin{aligned}
\xi &\ge \xi' \exp\left(-\sum_{m \le 2} \sum_{\bar{v} \le R_1} \left| u \left(\sum_{i \le N-1} g_{i,k_{\bar{v}}}\gamma_{i,k_{\bar{v}}}\sigma_i^m + g_{N,k_{\bar{v}}}\sigma_N^m \right) \right| \right) \\
&\ge \xi' \exp\left(-2U_\infty R_1\right),
\end{aligned}$$

and hence

$$\mathbf{Av} f\xi \geq (\mathbf{Av} f\xi') \exp(-2U_\infty R_1)\,. \tag{19}$$

On the other hand, since f' only depends on $\{1, \ldots, m-1, N\}$, we have $f' \circ T_{j_l} = f'$ for any $l \leq |J_1|$ on Ω^c, which yields

$$\begin{aligned} f_s &= (\mathbf{Av} f'\xi) \circ \prod_{l \leq s-1} U_{j_l} - (\mathbf{Av} f'\xi) \circ \prod_{l \leq s} U_{j_l} \\ &= \mathbf{Av}\left((f'\xi) \circ \prod_{l \leq s-1} T_{j_l} - (f'\xi) \circ \prod_{l \leq s} T_{j_l} \right) \\ &= \mathbf{Av}\left(f' \left(\xi \circ \prod_{l \leq s-1} T_{j_l} - \xi \circ \prod_{l \leq s} T_{j_l} \right) \right), \end{aligned} \tag{20}$$

where we have used the fact that J_1 can be written as $J_1 = \{j_1, \ldots, j_{|J_1|}\}$. Moreover, for any l, by construction of ξ', we have $\xi' \circ T_{j_l} = \xi'$. Thus,

$$\xi \circ \prod_{l \leq s-1} T_{j_l} - \xi \circ \prod_{l \leq s} T_{j_l} = \xi' \left[\xi'' \circ \prod_{l \leq s-1} T_{j_l} - \xi'' \circ \prod_{l \leq s} T_{j_l} \right]. \tag{21}$$

Set now

$$\Gamma = \sup_\sigma \left| \xi'' \circ \prod_{l \leq s-1} T_{j_l} - \xi'' \circ \prod_{l \leq s} T_{j_l} \right| = \sup_\sigma |\xi'' - \xi'' \circ T_{j_s}|.$$

Then, from (20) and (21), and invoking the fact that $|f'| \leq Qf$, we get

$$|f_s| \leq \Gamma \mathbf{Av}(|f'|\xi') \leq Q\Gamma \mathbf{Av} f\xi'. \tag{22}$$

We now bound Γ: recall that ξ'' is defined by (18), and thus

$$\xi'' = \prod_{\bar{v} \leq R_1} \xi_{\bar{v}},$$

with

$$\xi_{\bar{v}} = \exp\left(\sum_{m \leq 2} \eta_{k_{\bar{v}}} u \left(\sum_{i \leq N-1} g_{i,k_{\bar{v}}} \gamma_{i,k_{\bar{v}}} \sigma_i^m + g_{N,k_{\bar{v}}} \sigma_N^m \right) \right).$$

Recall now that we have assumed that $j_s \in I_v^1$. Therefore, we have $j_s \notin I_{\bar{v}}^1$ if $\bar{v} \neq v$, according to the fact that $I_v^1 \cap I_{\bar{v}}^1 = \emptyset$ on Ω^c. Hence

$$\xi_{\bar{v}} \circ T_{j_s} = \xi_{\bar{v}},$$

and

$$\xi'' - \xi'' \circ T_{j_s} = (\xi_v - \xi_v \circ T_{j_s}) \prod_{\bar{v} \neq v} \xi_{\bar{v}}. \tag{23}$$

On the other hand, since $|e^x - e^y| \leq |x-y|e^a$ for $|x|, |y| \leq a$, we obtain

$$|\xi_v - \xi_v \circ T_{j_s}| \leq 4U_\infty e^{2U_\infty}, \tag{24}$$

and we also have the trivial bound

$$\xi_{\bar{v}} \leq e^{2U_\infty}. \tag{25}$$

Thus, plugging (24) and (25) into (23), we get

$$\Gamma \leq 4U_\infty e^{2U_\infty R_1}.$$

Combining this bound with (19) and (22), the proof is now easily completed. □

We are now ready to start the induction procedure on $\mathbf{P}(N, \gamma_0, B)$, which will use the following elementary lemma (whose proof is left to the reader).

Lemma 2.6 *Let R be a random variable following the* $\mathrm{Bin}(M, \frac{\gamma}{N})$ *distribution, and λ be a positive number. Then*

$$\mathbf{E}\left[Re^{\lambda R}\right] \leq \alpha\gamma e^{\lambda} e^{\alpha\gamma(e^\lambda - 1)}, \tag{26}$$

$$\mathbf{E}\left[R^2 e^{\lambda R}\right] \leq \alpha^2\gamma^2 e^{2\lambda} e^{\alpha\gamma(e^\lambda - 1)} + \alpha\gamma e^{\lambda} e^{\alpha\gamma(e^\lambda - 1)}. \tag{27}$$

Let us proceed now with the main step of the induction:

Proposition 2.7 *Assume that $\boldsymbol{P}(N-1, \gamma_0, B)$ holds for $N \geq 2$ and $\gamma \leq \gamma_0$. Consider f and f' as in Definition 2.1. Then*

$$\mathbf{E}\left[\left|\frac{\langle f' \rangle}{\langle f \rangle}\right|\right] \leq \frac{mQ}{N}\left(\alpha\gamma^2 + \alpha^2\gamma^4 + 4B\Upsilon(\alpha, \gamma, U_\infty)\right), \tag{28}$$

where

$$\Upsilon(\alpha, \gamma, U_\infty) = U_\infty \alpha\gamma^2 \; e^{4U_\infty} \; e^{\alpha\gamma(e^{4U_\infty} - 1)} \; \left(3 + 2\gamma + \alpha(\gamma^2 + \gamma^3)e^{4U_\infty}\right).$$

Proof: Using (11) and (15), we have

$$\mathbf{E}\left[\left|\frac{\langle f'\rangle}{\langle f\rangle}\right|\right] \le Q\frac{\alpha\gamma^2(m-1)+\alpha^2\gamma^4}{N} + \frac{1}{2}\mathbf{E}\left(\mathbf{1}_{\Omega^c}\sum_{s\le|J_1|}\frac{|\langle f_s\rangle_-|}{\langle \mathbf{Av} f\xi\rangle_-}\right).$$

However, on Ω^c, the functions f_s and $\mathbf{Av} f\xi$ depend on $m-1+|J_1|$ coordinates. Since $\gamma^- \le \gamma$ and $m-1+|J_1| \le m(1+|J_1|)$, the definition of the expectation $\mathbf{E}_{-,\gamma_N}$, the property $\mathbf{P}(N-1,\gamma_0,B)$, (17) and Lemma 2.5 imply

$$\begin{aligned}
&\mathbf{E}\left[\mathbf{1}_{\Omega^c}\sum_{s\le|J_1|}\frac{|\langle f_s\rangle_-|}{\langle \mathbf{Av} f\xi\rangle_-}\right] = \mathbf{E}\left[\mathbf{1}_{\Omega^c}\sum_{s\le|J_1|}\mathbf{E}_{-,\gamma_N}\left[\frac{|\langle f_s\rangle_-|}{\langle \mathbf{Av} f\xi\rangle_-}\right]\right]\\
&\le \mathbf{E}\left[\mathbf{1}_{\Omega^c}\sum_{s\le|J_1|}\frac{(m-1+|J_1|)B\hat{Q}}{N-1}\right]\\
&\le 4\frac{m}{N-1}BQU_\infty\ \mathbf{E}\left[\mathbf{1}_{\Omega^c}|J_1|(1+|J_1|)\ e^{4U_\infty R_1}\right]\\
&\le 8\frac{m}{N}BQU_\infty\ \mathbf{E}\left[\mathbf{1}_{\Omega^c}|J_1|(1+|J_1|)\ e^{4U_\infty R_1}\right].
\end{aligned}$$

Recall that, according to (16), we have

$$|J_1| \le \sum_{v\le R_1}|I_v^1|,$$

and that the quantity R_1 is a $\mathrm{Bin}(M,\frac{\gamma}{N})$ random variable. Thus

$$\begin{aligned}
\mathbf{E}\left[\mathbf{1}_{\Omega^c}|J_1|\ e^{\lambda R_1}\right] &= \mathbf{E}\left\{\mathbf{E}\left[\mathbf{1}_{\Omega^c}|J_1|\ e^{\lambda R_1}\middle|R_1\right]\right\} = \mathbf{E}\left\{e^{\lambda R_1}\ \mathbf{E}\left[\mathbf{1}_{\Omega^c}|J_1|\middle|R_1\right]\right\}\\
&\le \gamma\mathbf{E}\left[R_1 e^{\lambda R_1}\right],\\
\mathbf{E}\left[\mathbf{1}_{\Omega^c}|J_1|^2\ e^{\lambda R_1}\right] &= \mathbf{E}\left\{\mathbf{E}\left[\mathbf{1}_{\Omega^c}|J_1|^2\ e^{\lambda R_1}\middle|R_1\right]\right\}\\
&= \mathbf{E}\left\{e^{\lambda R_1}\ \mathbf{E}\left[\mathbf{1}_{\Omega^c}|J_1|^2\middle|R_1\right]\right\}\\
&\le (\gamma+\gamma^2)\mathbf{E}\left[(R_1+R_1^2)\ e^{\lambda R_1}\right]. \qquad (29)
\end{aligned}$$

The proof of this proposition is now easily concluded by applying the previous bounds, together with Lemma 2.6, to the quantity

$$\mathbf{E}\left[\mathbf{1}_{\Omega^c}|J_1|(1+|J_1|)\ e^{4U_\infty R_1}\right].$$

□

We can turn now to the main aim of this Section:

Proof of Theorem 2.2: The result is now an immediate consequence of (4) and Proposition 2.7, applied to

$$B = B_0 = \frac{\alpha\gamma^2 + \alpha^2\gamma^4}{1-\varepsilon},$$

where ε satisfies

$$4U_\infty\ \alpha\gamma_0^2\ e^{4U_\infty}\ e^{\alpha\gamma_0(e^{4U_\infty}-1)}\ \left(3 + 2\gamma_0 + \alpha(\gamma_0^2+\gamma_0^3)e^{4U_\infty}\right) < \varepsilon < 1.$$

□

Before closing this Section, we will give an easy consequence of Theorem 2.2: we will see that, as N grows to ∞, the Gibbs measure G_N taken on a finite number of spins looks like a product measure. To this purpose, let us denote by $\langle\cdot\rangle_\bullet$ the average with respect to the product measure ν on Σ_{N-1} such that

$$\forall i \leq N-1, \qquad \int \sigma_i\ d\nu(\boldsymbol{\rho}) = \langle \sigma_i \rangle_- .$$

Equivalently, for a function $\bar{f}$ on Σ_{N-1}, we can write

$$\langle \bar{f} \rangle_\bullet = \langle \bar{f}(\sigma_1^1, \ldots, \sigma_{N-1}^{N-1})\rangle_-,$$

where σ_i^i is the i-th coordinate of the i-th replica $\boldsymbol{\rho}^i$. Recall also that, for $v \leq R_1$, I_v^1 has been defined as

$$I_v^1 = \{i \leq N-1 :\ \gamma_{i,k_v} = 1\}.$$

We now introduce the enumeration $\{i_1^v, \ldots, i_{|I_v^1|}^v\}$ of this set. Furthermore, given the randomness contained in the $\gamma_{N,k}$, the law of $|I_v^1|$ is a $\mathrm{Bin}(N-1, \frac{\gamma}{N})$.

Proposition 2.8 *Assume (4) and $\gamma \leq \gamma_0$, and consider*

$$\Theta = \exp\left(\sum_{v\leq R_1} \eta_{k_v} u \left(\sum_{p \leq |I_v^1|} g_{i_p^v, k_v}\sigma_{i_p^v} + g_{N,k_v}\sigma_N\right)\right).$$

Then, when Ω does not occur, we have

$$\mathbf{E}_{-,\gamma_N}\left|\frac{\langle \mathbf{Av}\sigma_N\Theta\rangle_-}{\langle \mathbf{Av}\Theta\rangle_-}-\frac{\langle \mathbf{Av}\sigma_N\Theta\rangle_\bullet}{\langle \mathbf{Av}\Theta\rangle_\bullet}\right|\leq 2B_0\ (|J_1|-1)\ \frac{|J_1|+1}{N-1}(e^{2U_\infty}-1),$$

where the conditional expectation $\mathbf{E}_{-,\gamma_N}$ has been defined at (8).

Remark 2.9 *The quantity Θ appears naturally in the decomposition of the Hamiltonian $H_{N,M}$. Indeed, on Ω_2^c, we have*

$$\begin{aligned}
-H_{N,M}(\boldsymbol{\sigma}) &= \sum_{k\leq M}\eta_k u\left(\sum_{i\leq N} g_{i,k}\gamma_{i,k}\sigma_i\right)\\
&= \sum_{k\notin D_{N,1}^M}\eta_k u\left(\sum_{i\leq N-1} g_{i,k}\gamma_{i,k}\sigma_i\right)+\sum_{k\in D_{N,1}^M}\eta_k u\left(\sum_{i\leq N-1} g_{i,k}\gamma_{i,k}\sigma_i+g_{N,k}\sigma_N\right)\\
&= \sum_{k\notin D_{N,1}^M}\eta_k u\left(\sum_{i\leq N-1} g_{i,k}\gamma_{i,k}\sigma_i\right)+\sum_{v\leq R_1}\eta_{k_v} u\left(\sum_{p\leq |I_v^1|} g_{i_p^v,k_v}\sigma_{i_p^v}+g_{N,k_v}\sigma_N\right).
\end{aligned}$$

Observe also that ξ defined by (7) evaluated for $n=1$ gives $\xi=\Theta$.

Proof of Proposition 2.8: The proof is similar to Proposition 7.2.7 in [10], and we include it here for sake of completeness: On Ω^c, since the sets I_v^1 are disjoint, the values i_p^v, for any v and p, are different and we can write

$$\bigcup_{v\leq R_1} I_v^1 = J_1 \equiv \left\{j_1,\dots j_{|J_1|}\right\}.$$

Set

$$f'=f'\big(\sigma_{j_1}^1,\dots,\sigma_{j_{|J_1|}}^1\big)\equiv \mathbf{Av}\sigma_N\Theta,\qquad f=f\big(\sigma_{j_1}^1,\dots,\sigma_{j_{|J_1|}}^1\big)\equiv \mathbf{Av}\Theta.$$

Let us also define, for $2\leq l\leq |J_1|$,

$$f'_{j_l}=f'\big(\sigma_1^1,\dots,\sigma_{j_l}^{j_l},\sigma_{j_{l+1}}^1,\dots\sigma_{j_{|J_1|}}^1\big)$$

and f_{j_l} in a similar way. Then

$$\begin{aligned}
&\mathbf{E}_{-,\gamma_N}\left|\frac{\langle \mathbf{Av}\sigma_N\Theta\rangle_-}{\langle \mathbf{Av}\Theta\rangle_-}-\frac{\langle \mathbf{Av}\sigma_N\Theta\rangle_\bullet}{\langle \mathbf{Av}\Theta\rangle_\bullet}\right| \\
&\quad= \mathbf{E}_{-,\gamma_N}\left|\frac{\langle f'_{j_1}\rangle_-}{\langle f_{j_1}\rangle_-}-\frac{\langle f'_{j_{|J_1|}}\rangle_-}{\langle f_{j_{|J_1|}}\rangle_-}\right| \le \sum_{2\le l\le |J_1|}\mathbf{E}_{-,\gamma_N}\left[\frac{\langle f'_{j_l-1}\rangle_-}{\langle f_{j_l-1}\rangle_-}-\frac{\langle f'_{j_l}\rangle_-}{\langle f_{j_l}\rangle_-}\right] \\
&\quad\le \sum_{2\le l\le |J_1|}\left[\mathbf{E}_{-,\gamma_N}\left|\frac{\langle f'_{j_l-1}-f'_{j_l}\rangle_-}{\langle f_{j_l-1}\rangle_-}\right| + \mathbf{E}_{-,\gamma_N}\left|\frac{\langle f'_{j_l}\rangle_-\langle f_{j_l-1}-f_{j_l}\rangle_-}{\langle f_{j_l-1}\rangle_- -\langle f_{j_l}\rangle_-}\right|\right]. \qquad (30)
\end{aligned}$$

Let us concentrate now on the first term of the right-hand side of (30), since the other term can be bounded similarly: observe that, for $2 \le l \le |J_1|$, we have

$$f'_{j_l} = f'_{j_{l-1}}\Delta, \quad \text{with} \quad e^{-2U_\infty} \le \Delta \le e^{2U_\infty}.$$

Furthermore, it is easily seen that $f'_{j_l-1} - f'_{j_l}$ enjoys the antisymmetric property assumed in Definition 2.1. Thus, applying $\mathbf{P}(N-1,\gamma_0,B_0)$, we get

$$\mathbf{E}_{-,\gamma_N}\left|\frac{\langle f'_{j_l-1}-f'_{j_l}\rangle_-}{\langle f_{j_l-1}\rangle_-}\right| \le \frac{B_0(|J_1|+1)}{N-1}(e^{2U_\infty}-1),$$

which ends the proof.

□

3 Study of the magnetization

For the non-diluted perceptron model, in the high temperature regime, the asymptotic behavior of the magnetization can be summarized easily: indeed, it has been shown in [7] that $\langle\sigma_1\rangle$ converges in L^2 to a random variable of the form $\tanh^2(z\sqrt{r})$, where r is a solution to a deterministic equation, and $z \sim N(0,1)$. Our goal in this section is to analyze the same problem for the diluted perceptron model. However, in the current situation, the limiting law is a more complicated object, and in order to present our asymptotic result, we will go through a series of notations and preliminary lemmas.

Let $\mathbf{P}$ be the set of probability measures on $[-1,1]$. We start by constructing a map $T : \mathbf{P} \to \mathbf{P}$ in the following way: for any integer $\theta \ge 1$, let $(\tau_1,\ldots,\tau_\theta)$ be θ arbitrary integers. Then, for $k = 1,\ldots,\theta$, let t_k be the cumulative sum of the τ_k; that is, $t_0 = 0$ and $t_k = \sum_{\hat{k} \le k} \tau_{\hat{k}}$ for $k \ge 1$. Let also $\{\bar{g}_{i,k},\ i,k \ge 1\}$ and $\{\bar{g}_k,\ k \ge 1\}$ be two independent families of independent standard Gaussian random variables. Define then a random variable $\xi_{\theta,\tau}$ by

$$\begin{aligned} \xi_{\theta,\tau} &= \xi_{\theta,\tau}(\sigma_1, \ldots, \sigma_{t_\theta}, \varepsilon) \\ &= \exp \sum_{k=1}^{\theta} u \left(\sum_{i=1}^{\tau_k} \bar{g}_{i,k}\ \sigma_{t_{k-1}+i}\ + \bar{g}_k\ \varepsilon \right). \end{aligned} \tag{31}$$

Whenever $\theta = 0$, set also $\xi_\theta = 1$, which is equivalent to the convention $\sum_{k=1}^{0} w_k = 0$ for any real sequence $\{w_k; k \ge 0\}$.

Consider now $\boldsymbol{x} = (x_1, \ldots, x_{\sum_{k=1}^{\theta} \tau_k})$ with $|x_i| \le 1$ and a function

$$f : \{-1, 1\}^{\sum_{k=1}^{\theta} \tau_k} \to \mathbb{R}.$$

We denote by $\langle f \rangle_{\boldsymbol{x}}$ the average of f with respect to the product measure ν on $\{-1,1\}^{\sum_{k=1}^{\theta} \tau_k}$ such that $\int \sigma_i d\nu(\boldsymbol{\delta}) = x_i$, where $\boldsymbol{\delta} = (\sigma_1, \ldots, \sigma_{\sum_{k=1}^{\theta} \tau_k})$. Using this notation, when $\theta \ge 1$, we define $T_{\theta,\tau} : \mathbf{P} \to \mathbf{P}$ such that, for $\mu \in \mathbf{P}$, $T_{\theta,\tau}(\mu)$ is the law of the random variable

$$\frac{\langle \mathbf{Av} \varepsilon \xi_{\theta,\tau} \rangle_{\boldsymbol{X}}}{\langle \mathbf{Av} \xi_{\theta,\tau} \rangle_{\boldsymbol{X}}}, \tag{32}$$

where $\boldsymbol{X} = (X_1, \ldots, X_{\sum_{k=1}^{\theta} \tau_k})$ is a sequence of i.i.d. random variables of law μ independent of the randomness in $\xi_{\theta,\tau}$ and $\mathbf{Av}$ denotes the average over $\varepsilon = \pm 1$. Notice that when $\theta = 0$, $T_{\theta,\tau}(\mu)$ is the Dirac measure at point 0.

Finally, we can define the map $T : \mathbf{P} \to \mathbf{P}$ by

$$T(\mu) = \sum_{\theta \ge 0} \sum_{\tau_1, \ldots, \tau_\theta \ge 0} \kappa(\theta, \tau_1, \ldots, \tau_\theta)\ T_{\theta,\tau}(\mu), \tag{33}$$

with

$$\kappa(\theta, \tau_1, \ldots, \tau_\theta) = e^{-\alpha\gamma}\ \frac{(\alpha\gamma)^\theta}{\theta!}\ e^{-\theta\gamma}\ \frac{\gamma^{\sum_{l \le \theta} \tau_l}}{\tau_1! \cdots \tau_\theta!}, \tag{34}$$

and where the coefficients α, γ are the parameters of our perceptron model. We will see that the asymptotic law μ of the magnetization $\langle \sigma_1 \rangle$ will satisfy the relation $\mu = T(\mu)$. Hence, a first natural aim of this section is to prove that the equation $\mu = T(\mu)$ admits a unique solution:

Theorem 3.1 *Assume*

$$2U_\infty \; e^{2U_\infty} \; \alpha\gamma^2 < \frac{1}{2}. \tag{35}$$

Then there exists a unique probability distribution μ on $[-1,1]$ such that $\mu = T(\mu)$.

Remark 3.2 *Notice that (4) implies (35).*

In order to establish the fixed point argument for the proof of Theorem 3.1, we will need a metric on $\mathbf{P}$, and in fact it will be suitable for computational purposes to choose the Monge-Kantorovich transportation-cost distance (or equivalently the total variation distance) for the compact metric space $([-1,1], |\cdot|)$: for two probabilities μ_1 and μ_2 on $[-1,1]$, the distance between μ_1 and μ_2 will be defined as

$$d(\mu_1, \mu_2) = \inf \mathbf{E}|X_1 - X_2|, \tag{36}$$

where this infimum is taken over all the pairs (X_1, X_2) of random variables such that the law of X_j is μ_j, $j = 1, 2$. This definition is equivalent to say that

$$d(\mu_1, \mu_2) = \inf \int d(x_1, x_2) d\zeta(x_1, x_2), \qquad \text{with} \qquad d(x_1, x_2) = |x_2 - x_1|,$$

where this infimum is now taken over all probabilities ζ on $[-1,1]^2$ with marginals μ_1 and μ_2 (see Section 7.3 in [10] for more information about transportation-cost distances). Finally, throughout this section, we also use a local definition of distance between two probabilities, with respect to an event Ω:

$$d_\Omega(\mu_1, \mu_2) = \inf \mathbf{E} \left| (X_1 - X_2) \mathbf{1}_\Omega \right|, \tag{37}$$

where this infimum is as in (36).

Proof of Theorem 3.1: Assume that $\theta \geq 1$ and $\tau_k \geq 1$ for some $k = 1, \cdots, \theta$. Then, using similar arguments to Lemma 7.3.5 in [10] we can prove, for $1 \leq i \leq \sum_{k=1}^{\theta} \tau_k$, that

$$\left| \frac{\partial}{\partial x_i} \frac{\langle \mathbf{Av} \varepsilon \xi_{\theta,\tau} \rangle_{\boldsymbol{x}}}{\langle \mathbf{Av} \xi_{\theta,\tau} \rangle_{\boldsymbol{x}}} \right| \leq 2U_\infty \; e^{2U_\infty}, \tag{38}$$

with $\boldsymbol{x} = (x_1, \ldots, x_{\sum_{k=1}^{\theta} \tau_k})$. Then if $\boldsymbol{y} = (y_1, \ldots, y_{\sum_{k=1}^{\theta} \tau_k})$, the bound (38) implies that

$$\left| \frac{\langle \mathbf{Av} \varepsilon \xi_{\theta,\tau} \rangle_{\boldsymbol{x}}}{\langle \mathbf{Av} \xi_{\theta,\tau} \rangle_{\boldsymbol{x}}} - \frac{\langle \mathbf{Av} \varepsilon \xi_{\theta,\tau} \rangle_{\boldsymbol{y}}}{\langle \mathbf{Av} \xi_{\theta,\tau} \rangle_{\boldsymbol{y}}} \right| \leq 2U_\infty \ e^{2U_\infty} \ \sum_{k=1}^{\theta} \sum_{i=1}^{\tau_k} |x_{t_k+i} - y_{t_k+i}|. \tag{39}$$

Remark that if $\theta = 0$ or $\theta \neq 0$ but $\tau_k = 0$ for any $k = 1, \ldots, \theta$, then the left-hand side of (39) is zero.

Let now (X, Y) be a pair of random variables such that the laws of X and Y are μ_1 and μ_2, respectively (μ_1 and μ_2 are independent of the randomness in $\xi_{\theta,\tau}$). Consider independent copies $(X_i, Y_i)_{i \leq \sum_{k=1}^{\theta} \tau_k}$ of this couple of random variables. Then, if $\boldsymbol{X} = (X_i)_{i \leq \sum_{k=1}^{\theta} \tau_k}$ and $\boldsymbol{Y} = (Y_i)_{i \leq \sum_{k=1}^{\theta} \tau_k}$, we have that

$$\frac{\langle \mathbf{Av} \varepsilon \xi_{\theta,\tau} \rangle_{\boldsymbol{X}}}{\langle \mathbf{Av} \xi_{\theta,\tau} \rangle_{\boldsymbol{X}}} \stackrel{(d)}{=} T_{\theta,\tau}(\mu_1) \quad \text{and} \quad \frac{\langle \mathbf{Av} \varepsilon \xi_{\theta,\tau} \rangle_{\boldsymbol{Y}}}{\langle \mathbf{Av} \xi_{\theta,\tau} \rangle_{\boldsymbol{Y}}} \stackrel{(d)}{=} T_{\theta,\tau}(\mu_2).$$

Hence, applying (39) for $\boldsymbol{x} = \boldsymbol{X}$ and $\boldsymbol{y} = \boldsymbol{Y}$ and taking first expectation and then infimum over the choice of (X, Y), we obtain

$$d(T_{\theta,\tau}(\mu_1), T_{\theta,\tau}(\mu_2)) \leq 2U_\infty \ e^{2U_\infty} \ d(\mu_1, \mu_2) \sum_{k=1}^{\theta} \tau_k. \tag{40}$$

Finally, recall (see [10, Lemma 7.3.2]) that for a given sequence $\{c_n; n \geq 1\}$ of positive numbers such that $\sum_{n \geq 1} c_n = 1$, and two sequences $\{\mu_n, \nu_n; n \geq 1\}$ of elements of $\mathbf{P}$, we have

$$d\left(\sum_{n \geq 1} c_n \mu_n, \sum_{n \geq 1} c_n \nu_n \right) \leq \sum_{n \geq 1} c_n \, d\left(\mu_n, \nu_n\right). \tag{41}$$

Applying this elementary result to $c_{\theta,\tau} = \kappa(\theta, \tau_1, \ldots, \tau_\theta)$, $\mu_{\theta,\tau} = T_{\theta,\tau}(\mu_1)$ and $\nu_{\theta,\tau} = T_{\theta,\tau}(\mu_2)$, we get

$$\begin{aligned}
d(T(\mu_1), T(\mu_2)) &\leq \sum_{\theta \geq 0} \sum_{\tau_1, \ldots, \tau_\theta \geq 0} \kappa(\theta, \tau_1, \ldots, \tau_\theta) \ d(T_{\theta,\tau}(\mu_1), T_{\theta,\tau}(\mu_2)) \\
&\leq 2U_\infty \ e^{2U_\infty} \ \sum_{\theta \geq 0} \sum_{\tau_1, \ldots, \tau_\theta \geq 0} \sum_{k=1}^{\theta} \kappa(\theta, \tau_1, \ldots, \tau_\theta) \, \tau_k \, d(\mu_1, \mu_2) \\
&= 2U_\infty \ e^{2U_\infty} \ \left(\sum_{\theta \geq 0} e^{-\alpha\gamma} \ \frac{(\alpha\gamma)^\theta}{\theta!} \ \theta\gamma \right) \ d(\mu_1, \mu_2) \\
&= 2U_\infty \ e^{2U_\infty} \ \alpha\gamma^2 \ d(\mu_1, \mu_2),
\end{aligned}$$

where we have used the fact that the mean of a Poisson random variable with parameter ρ is ρ. Then, under assumption (35), T is a contraction and there exists a unique probability distribution such that $\mu = T(\mu)$.

$\square$

Notice that the solution to the equation $\mu = T(\mu)$ depends on the parameters α and γ. Furthermore, in the sequel, we will need some continuity properties for the application $(\alpha, \gamma) \mapsto \mu_{\alpha,\gamma}$. Thus, we will set $\mu = \mu_{\alpha,\gamma}$ when we want to stress the dependence on the parameters α and γ, and the following holds true:

Lemma 3.3 *If (α, γ) and (α', γ') satisfy (35), then*

$$d(\mu_{\alpha,\gamma}, \mu_{\alpha',\gamma'}) \le 4\left[|\gamma - \gamma'|\alpha'\gamma' e^{|\gamma-\gamma'|} + |\alpha\gamma - \alpha'\gamma'|e^{|\alpha\gamma-\alpha'\gamma'|}\right].$$

Proof: Since $\mu_{\alpha,\gamma} = T_{\alpha,\gamma}(\mu_{\alpha,\gamma})$ and $\mu_{\alpha',\gamma'} = T_{\alpha',\gamma'}(\mu_{\alpha',\gamma'})$, using the triangular inequality and Theorem 3.1 we have

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

So

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

and we only need to deal with $d(T_{\alpha,\gamma}(\mu_{\alpha',\gamma'}), T_{\alpha',\gamma'}(\mu_{\alpha',\gamma'}))$. However, Lemma 7.3.3 in [10], which is a direct consequence of (41), implies that

$$\begin{aligned} & d(T_{\alpha,\gamma}(\mu_{\alpha',\gamma'}), T_{\alpha',\gamma'}(\mu_{\alpha',\gamma'})) \\ &\qquad \le 2\sum_{\theta\ge 0}\sum_{\tau_1,\dots,\tau_\theta\ge 0} |\kappa_{\alpha,\gamma}(\theta, \tau_1, \dots, \tau_\theta) - \kappa_{\alpha',\gamma'}(\theta, \tau_1, \dots, \tau_\theta)| \\ &\qquad \le 2(V_1 + V_2), \end{aligned}$$

with κ defined in (34) and

$$\begin{aligned} V_1 &= \sum_{\theta\ge 0}\sum_{\tau_1,\dots,\tau_\theta\ge 0} e^{-\theta\gamma}\frac{\gamma^{\sum_{l\le\theta}\tau_l}}{\theta!\ \tau_1!\cdots\tau_\theta!}\left|e^{-\alpha\gamma}(\alpha\gamma)^\theta - e^{-\alpha'\gamma'}(\alpha'\gamma')^\theta\right|, \\ V_2 &= \sum_{\theta\ge 0}\sum_{\tau_1,\dots,\tau_\theta\ge 0} e^{-\alpha'\gamma'}\frac{(\alpha'\gamma')^\theta}{\theta!\ \tau_1!\cdots\tau_\theta!}\left|e^{-\theta\gamma}\gamma^{\sum_{l\le\theta}\tau_l} - e^{-\theta\gamma'}\gamma'^{\sum_{l\le\theta}\tau_l}\right|. \end{aligned}$$

Now, following the arguments of (7.53) in [10], we get

$$V_1 = \sum_{\theta \geq 0} \frac{1}{\theta!} \left| e^{-\alpha\gamma}(\alpha\gamma)^\theta - e^{-\alpha'\gamma'}(\alpha'\gamma')^\theta \right| \leq |\alpha\gamma - \alpha'\gamma'| e^{|\alpha\gamma - \alpha'\gamma'|},$$

$$V_2 \leq |\gamma - \gamma'| e^{|\gamma-\gamma'|} \sum_{\theta \geq 0} \theta e^{-\alpha'\gamma'} \frac{(\alpha'\gamma')^\theta}{\theta!} = \alpha'\gamma' |\gamma - \gamma'| e^{|\gamma-\gamma'|},$$

which ends the proof of this lemma.

□

From now on, we will specialize our Hamiltonian to the case of interest for us:

Hypothesis 3.4 *The parameters η_k, $k = 1, \ldots, M$ in the Hamiltonian (1) are all equal to one.*

This assumption being made, we can now turn to the main result of the section:

Theorem 3.5 *Let γ_0 be a positive number such that*

$$4U_\infty \, \alpha\gamma_0^2 \, e^{4U_\infty} \, e^{\alpha\gamma_0(e^{4U_\infty}-1)} \left(3 + 2\gamma_0 + \alpha(\gamma_0^2 + \gamma_0^3)e^{4U_\infty}\right) < 1, \tag{42}$$

and assume that there exists a positive number C_0 satisfying

$$C_0 \alpha \, \gamma_0^6 \, U_\infty \, e^{2U_\infty} \leq 1. \tag{43}$$

Then for any $\gamma \leq \gamma_0$, given any integer m, we can find i.i.d. random variables $z_1, \ldots, z_m$ with law $\mu_{\alpha,\gamma}$ such that

$$\mathbf{E}\left[\sum_{i \leq m} |\langle \sigma_i \rangle - z_i|\right] \leq \frac{Km^3}{N}, \tag{44}$$

for a constant $K > 0$ independent of m.

Remark 3.6 *The two conditions in the above theorem are met when the following hypothesis is satisfied: there exists $L > 0$ such that*

$$L \, U_\infty \, \alpha \, \gamma_0^6 \, \exp\left\{8U_\infty + \alpha\gamma_0\left(e^{4U_\infty} - 1\right)\right\} < 1.$$

As in the case of Theorem 2.2, the proof of Theorem 3.5 will require the introduction of some notations and preliminary Lemmas. Let us first recast relation (44) in a suitable way for an induction procedure: consider the metric space $[-1,1]^m$, equipped with the distance given by

$$d((x_i)_{i\leq m},(y_i)_{i\leq m})=\sum_{i\leq m}|x_i-y_i|.$$

We also denote by d the transportation-cost distance on the space of probability measures on $[-1,1]^m$, defined as in (36). Define now

$$D(N,M,m,\gamma_0)=\sup_{\gamma\leq\gamma_0} d\left(\mathcal{L}(\langle\sigma_1\rangle,\ldots,\langle\sigma_m\rangle),\mu_{\alpha,\gamma}^{\otimes m}\right), \tag{45}$$

where $\mathcal{L}(X)$ stands for the law of the random variable X. Then the statement of Theorem 3.5 is equivalent to say that, under Hypothesis (43), we have

$$D(N,M,m,\gamma_0)\leq\frac{Km^3}{N},$$

for any fixed integer $m\geq 1$.

It will also be useful to introduce a cavity formula for m spins, which we proceed to do now: generalizing some aspects of the previous section, we consider, for $p\in\{1,\ldots,m\}$, the random sets

$$D_{N,p}^M=\{k\leq M:\ \gamma_{N-p+1,k}=1\},$$

and

$$F_{N,M}^m=\bigcup_{p=1}^m D_{N,p}^M\ .$$

We also define the following two rare events:

$$\begin{aligned}\tilde\Omega_1 &= \{\exists k\leq M,\ p_1,p_2\leq m:\ \gamma_{N-p_1+1,k}=\gamma_{N-p_2+1,k}=1\},\\ \tilde\Omega_2 &= \{\exists i\leq N-m,\ k_1,k_2\in F_{N,M}^m:\ \gamma_{i,k_1}=\gamma_{i,k_2}=1\},\end{aligned}$$

satisfying

$$P(\tilde\Omega_1)\leq\frac{\alpha\gamma^2m^2}{N}\qquad\text{and}\qquad P(\tilde\Omega_2)\leq\frac{\alpha^2\gamma^4m^2}{N}. \tag{46}$$

Then, the following properties hold true: first, for a fixed k, if $\tilde\Omega_1^c$ is realized, we have

$$\mathrm{Card}\{p\leq m,\ \gamma_{N-p+1,k}=1\}\leq 1.$$

Moreover, still on $\tilde{\Omega}_1^c$, for $p_1 \neq p_2$,

$$D_{N,p_1}^M \cap D_{N,p_2}^M = \emptyset;$$

and hence,

$$R_m \equiv |F_{N,M}^m| = \sum_{p=1}^m |D_{N,p}^M| = \sum_{k \le M} \sum_{p \le m} \gamma_{N-p+1,k}.$$

Actually, notice that we always have

$$R_m \le \sum_{p=1}^m |D_{N,p}^M|.$$

Let us introduce now an enumeration of $F_{N,M}^m$:

$$F_{N,M}^m = \{k_1, \dots, k_{R_m}\},$$

and for any $v \le R_m$ set

$$I_v^m = \{j \le N-m : \gamma_{j,k_v} = 1\}.$$

Then, on $\tilde{\Omega}_2^c$, we get

$$I_{v_1}^m \cap I_{v_2}^m = \emptyset, \quad \text{if } v_1 \neq v_2, \tag{47}$$

and we can also write

$$J_m = \bigcup_{v \le R_m} I_v^m = \bigcup_{v \le R_m} \{j \le N-m : \ \gamma_{j,k_v} = 1\}.$$

Let us separate now the m last spins in the Hamiltonian $H_{N,M}$: if $\tilde{\Omega}_1^c$ is realized, for $\boldsymbol{\rho} = (\sigma_1, \dots, \sigma_{N-m})$, we have the following decomposition:

$$-H_{N,M}(\boldsymbol{\sigma}) = -H_{N-m,M}^-(\boldsymbol{\rho}) + \log \xi,$$

with

$$-H_{N-m,M}^-(\boldsymbol{\rho}) = \sum_{k \in (F_{N,M}^m)^c} u\left(\sum_{i \le N-m} g_{i,k}\gamma_{i,k}\sigma_i\right),$$

$$\xi = \exp\left(\sum_{p=1}^m \sum_{k \in D_{N,p}^M} u\left(\sum_{i \le N-m} g_{i,k}\gamma_{i,k}\sigma_i + g_{N-p+1,k}\sigma_{N-p+1}\right)\right). \tag{48}$$

Observe that, in the last formula, $H^-_{N-m,M}(\boldsymbol{\rho})$ is not exactly the Hamiltonian of a $(N-m)$-spin system changing γ into γ^-, because the set $F^n_{N,M}$ is not deterministic. But this problem will be solved again by conditioning upon the random variables $\{\gamma_{N-p+1,k},\ p=1,\ldots,m,\ k\leq M\}$. For the moment, let us just mention that the m cavity formula will be the following: given f on Σ_N, we have

$$\mathbf{1}_{\tilde{\Omega}_1^c}\langle f\rangle = \mathbf{1}_{\tilde{\Omega}_1^c}\frac{\langle \mathbf{Av} f\xi\rangle_-}{\langle \mathbf{Av}\xi\rangle_-}, \tag{49}$$

where $\langle\cdot\rangle_-$ is the average with respect to $H^-_{N-m,M}$ and $\mathbf{Av}$ is the average with respect to last m spins. Moreover, in the last formula, we have kept the notation ξ from Section 2, which hopefully will not lead to any confusion. Finally, we denote by $\mathcal{L}_0$ the law of a random variable conditioned by $\{\gamma_{N-p+1,k},\ p=1,\ldots,m,\ k\leq M\}$, and by $\mathbf{E}_{\gamma_{N,m}}$ the associated conditional expectation.

We can start now stating and proving the lemmas and propositions that will lead to the proof of Theorem 3.5. Recall that given $\boldsymbol{x}=(x_1,\ldots,x_{N-m})$, $|x_i|\leq 1$, and a function f on Σ_{N-m}, $\langle f\rangle_{\boldsymbol{x}}$ means the average of f with respect to the product measure ν on Σ_{N-m} such that $\int \sigma_i d\nu(\boldsymbol{\rho}) = x_i$, for $0\leq i\leq N-m$. Recall also that $\gamma^-=\gamma\frac{N-m}{N}$. Then, as a direct consequence of the definition of the operator $T_{\theta,\tau}$, we have the following result:

Lemma 3.7 *Let $\boldsymbol{X}=(X_1,\ldots,X_{N-m})$ be an independent sequence of random variables, where the law of each X_l is μ_{α,γ^-}. Set*

$$w_p = \frac{\langle \mathbf{Av}\sigma_{N-p+1}\xi\rangle_{\boldsymbol{X}}}{\langle \mathbf{Av}\xi\rangle_{\boldsymbol{X}}}, \qquad p=1,\ldots,m.$$

Then, on $\tilde{\Omega}^c=(\tilde{\Omega}_1\cup\tilde{\Omega}_2)^c$, we have

$$\mathcal{L}_0(w_1\ldots,w_m) = T_{|D^M_{N,1}|,(|I^m_k|,k\in D^M_{N,1})}(\mu_{\alpha,\gamma^-})\otimes\cdots\otimes T_{|D^M_{N,m}|,(|I^m_k|,k\in D^M_{N,m})}(\mu_{\alpha,\gamma^-}).$$

We will now try to relate the random variables w_p with the magnetization of the m last spins. A first step in that direction is the following lemma where we use the random value of the parameter α^- associated to the Hamiltonian of a $(N-m)$-spin system.

Lemma 3.8 *On $\tilde{\Omega}^c$, set*

$$\Gamma_m = d\left(\mathcal{L}_0(w_1\ldots,w_m),\mathcal{L}_0(\bar{w}_1\ldots,\bar{w}_m)\right),$$

where, for $p = 1, \ldots, m$,

$$\bar{w}_p = T_{|D^M_{N,p}|,(|I^m_k|, k \in D^M_{N,p})}(\mu_{\alpha^-,\gamma^-}), \quad \text{with } \alpha^- = \frac{M - R_m}{N - m}.$$

Then, on $\tilde{\Omega}^c$, *we have*

$$\Gamma_m \leq 2U_\infty \; e^{2U_\infty} \; \gamma_0 \; \frac{|R_m - m\alpha|}{N} \exp\left\{\gamma_0 \; \frac{|R_m - m\alpha|}{N}\right\} \; \sum_{k=1}^{R_m} |I^m_k|.$$

Proof: Using (40) we obtain

$$\Gamma_m \leq 2U_\infty \; e^{2U_\infty} \; \sum_{p=1}^{m} \sum_{k \in D^M_{N,p}} |I^m_k| \; d(\mu_{\alpha,\gamma^-}, \mu_{\alpha^-,\gamma^-}).$$

The proof of this lemma is then easily finished thanks to Lemma 3.3, and taking the following equality into account:

$$\gamma^- \; |\alpha - \alpha^-| = \gamma \frac{|R_m - m\alpha|}{N}.$$

□

Notice that we have introduced the random variables $\bar{w}_p$ for the following reason: given the randomness contained in the $\{\gamma_{N-p+1,k}, \; p = 1, \ldots, m, k \leq M\}$, $\bar{w}_p$ can be interpreted as

$$\bar{w}_p = \frac{\langle \mathbf{Av} \sigma_{N-p+1} \xi \rangle_{\bar{\boldsymbol{X}}}}{\langle \mathbf{Av} \xi \rangle_{\bar{\boldsymbol{X}}}}, \qquad p = 1, \ldots, m,$$

where $\bar{\boldsymbol{X}} = (\bar{X}_1, \ldots, \bar{X}_{N-m})$ is an independent sequence of random variables with law μ_{α^-,γ^-}.

Lemma 3.9 *Consider* $\boldsymbol{Z} = (\langle \sigma_1 \rangle_-, \ldots, \langle \sigma_{N-m} \rangle_-)$, *and denote*

$$u_p = \frac{\langle \mathbf{Av} \sigma_{N-p+1} \xi \rangle_{\boldsymbol{Z}}}{\langle \mathbf{Av} \xi \rangle_{\boldsymbol{Z}}}, \qquad p = 1, \ldots, m.$$

Then, on $\tilde{\Omega}^c$,

$$d\left(\mathcal{L}_0(u_1, \ldots, u_m), \mathcal{L}_0(\bar{w}_1, \ldots, \bar{w}_m)\right) \\ \leq 4D(N - m, |(F^m_{N,M})^c|, |J_m|, \gamma_0) \; U_\infty \; e^{2U_\infty},$$

where the quantity D *has been defined at relation (45).*

Proof: As in (38) we can obtain, for any $i \leq N-m$,

$$\left| \frac{\partial}{\partial x_i} \frac{\langle \mathbf{Av} \sigma_{N-p+1} \xi \rangle_{\boldsymbol{x}}}{\langle \mathbf{Av} \xi \rangle_{\boldsymbol{x}}} \right| \leq 2U_\infty \, e^{2U_\infty}. \tag{50}$$

But in fact, these derivatives are vanishing, unless

$$i \in I_{N,p}^m \equiv \bigcup_{v;\ k_v \in D_{N,p}^M} I_v^m,$$

for some $p = 1, \ldots, m$. Indeed, on $\tilde{\Omega}^c$, from (47), we have

$$I_{N,p_1}^m \cap I_{N,p_2}^m = \emptyset, \qquad \text{if } p_1 \neq p_2.$$

Then, for a given $p \in \{1, \ldots, m\}$, we can decompose ξ into $\xi = \xi_{N,p} \, \bar{\xi}_{N,p}$, with

$$\begin{aligned} \xi_{N,p} &= \exp \sum_{k \in D_{N,p}^M} u \left(\sum_{i \leq N-m} g_{i,k} \gamma_{i,k} \sigma_i + g_{N-p+1,k} \sigma_{N-p+1} \right) \\ &= \xi_{N,p} \left(\{\sigma_i, i \in I_{N,p}^m\}, \sigma_{N-p+1} \right), \\ \bar{\xi}_{N,p} &= \bar{\xi}_{N,p} \left(\{\sigma_i, i \in J_m \backslash I_{N,p}^m\}, \sigma_{N-\bar{p}+1}, \bar{p} \leq m, \bar{p} \neq p\} \right). \end{aligned}$$

Then

$$\frac{\langle \mathbf{Av} \sigma_{N-p+1} \xi \rangle_{\boldsymbol{x}}}{\langle \mathbf{Av} \xi \rangle_{\boldsymbol{x}}} = \frac{\langle \mathbf{Av} \sigma_{N-p+1} \xi_{N,p} \rangle_{\boldsymbol{x}} \ \langle \mathbf{Av} \bar{\xi}_{N,p} \rangle_{\boldsymbol{x}}}{\langle \mathbf{Av} \xi_{N,p} \rangle_{\boldsymbol{x}} \ \langle \mathbf{Av} \bar{\xi}_{N,p} \rangle_{\boldsymbol{x}}} = \frac{\langle \mathbf{Av} \sigma_{N-p+1} \xi_{N,p} \rangle_{\boldsymbol{x}}}{\langle \mathbf{Av} \xi_{N,p} \rangle_{\boldsymbol{x}}},$$

and clearly the derivative $\frac{\partial}{\partial x_i}$ is zero when i does not belong to $I_{N,p}^m$, for any $p \in \{1, \ldots, m\}$.

Now, invoking inequality (50), we get

$$\sum_{p=1}^m \left| \frac{\langle \mathbf{Av} \sigma_{N-p+1} \xi \rangle_{\bar{\boldsymbol{X}}}}{\langle \mathbf{Av} \xi \rangle_{\bar{\boldsymbol{X}}}} - u_p \right| \leq \left(\sum_{p=1}^m \sum_{i \in I_{N,p}^m} \left| \bar{X}_i - \langle \sigma_i \rangle_- \right| \right) \ 2U_\infty \, e^{2U_\infty}.$$

Then, the definition of $\mathbf{E}_{\gamma_{N,m}}$ and (45) easily yield

$$\mathbf{E}_{\gamma_{N,m}} \left(\sum_{p=1}^m \sum_{i \in I_{N,p}^m} \left| \bar{X}_i - \langle \sigma_i \rangle_- \right| \right) \leq 2D(N-m, |(F_{N,M}^m)^c|, |J_m|, \gamma_0),$$

which ends the proof.

□

Set now, for $1 \leq p \leq m$,

$$\bar{u}_p = \frac{\langle \mathbf{Av} \sigma_{N-p+1} \xi \rangle_-}{\langle \mathbf{Av} \xi \rangle_-}. \tag{51}$$

Then $\bar{u}_p$ is closer to the real magnetization in the sense that $\bar{u}_p = \langle \sigma_{N-p+1} \rangle$ on $\tilde{\Omega}^c$, and the following Lemma claims that the distance between $\bar{u}_p$ and u_p vanishes as $N \to \infty$.

Lemma 3.10 *For $1 \leq p \leq m$, let $\bar{u}_p$ be defined by (51). Then, on $\tilde{\Omega}^c$, we have*

$$d\left(\mathcal{L}_0(\bar{u}_1, \ldots, \bar{u}_m), \mathcal{L}_0(u_1, \ldots, u_m)\right) \leq 2B_0 \, \frac{|J_m|^2 - 1}{N - m + 1} \, (e^{2U_\infty} - 1),$$

where the constant B_0 has been defined in the previous section.

Proof: The computations can be leaded here almost like in the proof of Proposition 2.8, and the details are left to the reader.

□

We will now identify the law of the $\bar{u}_p$ in terms of laws of the type $T(\mu_{\alpha,\gamma^-})$:

Lemma 3.11 *Recall that $d_{\tilde{\Omega}^c}$ has been defined by relation (37). Then, for $m \geq 1$, set*

$$\delta_m = d_{\tilde{\Omega}^c}\Big(\mathcal{L}(\bar{u}_1, \ldots, \bar{u}_m), \sum_{(b)} \sum_{(\boldsymbol{v})} a\left((b_1, \boldsymbol{v_1}), \ldots, (b_m, \boldsymbol{v_m})\right)$$
$$T_{b_1, \boldsymbol{v_1}}(\mu_{\alpha,\gamma^-}) \otimes \cdots \otimes T_{b_m, \boldsymbol{v_m}}(\mu_{\alpha,\gamma^-})\Big),$$

where we have used the following conventions: for $j \leq m$, $\boldsymbol{v_j}$ is a multi-index of the form $\boldsymbol{v_j} = (v_1^j, \ldots, v_{b_j}^j)$; the first summation $\sum_{(b)}$ is over $b_j \geq 0$, for $j = 1, \ldots, m$; the second one $\sum_{(\boldsymbol{v})}$ is over $v_1^j, \ldots, v_{b_j}^j \geq 0$, for $j = 1, \ldots, m$; and $a((b_1, \boldsymbol{v_1}), \ldots, (b_m, \boldsymbol{v_m}))$ is defined by

$$a\left((b_1, \boldsymbol{v_1}), \ldots, (b_m, \boldsymbol{v_m})\right) = P\left(|D_{N,j}^M| = b_j, (|I_k^m|, k \in D_{N,j}^M) = \boldsymbol{v_j}, \forall j \leq m\right).$$

Then, under the conditions of Lemma 3.10, we have

$$\delta_m \leq c_1(N, m),$$

with

$$\begin{aligned} c_1(N,m) &= 4\, U_\infty\, e^{2U_\infty}\, \mathbf{E}\left(D(N-m, |(F^m_{N,M})^c|, |J_m|, \gamma_0)\right) \\ &+2B_0\, \frac{\mathbf{E}|J_m|^2 - 1}{N-m+1}\, (e^{2U_\infty} - 1) \\ &+2U_\infty\, e^{2U_\infty}\, \frac{\gamma_0}{N}\mathbf{E}\left(|R_m - m\alpha|\; |J_m| \exp\left\{\frac{\gamma_0}{N}|R_m - m\alpha|\right\}\right). \end{aligned}$$

Proof: This result is easily obtained by combining Lemmas 3.7, 3.8, 3.9, 3.10 and taking expectations.

□

With Lemma 3.11 in hand, we can see that the remaining task left to us is mainly to compare the coefficients $a((b_1, \boldsymbol{v_1}), \ldots, (b_m, \boldsymbol{v_m}))$ with the coefficients $\kappa_{\alpha,\gamma^-}(b_j, \boldsymbol{v_j})$. This is done in the following lemma.

Lemma 3.12 *With the conventions of Lemma 3.11, we have*

$$\sum_{(b)} \sum_{(\boldsymbol{v})} \left| a\left((b_1, \boldsymbol{v_1}), \ldots, (b_m, \boldsymbol{v_m})\right) - \prod_{j=1}^{m} \kappa_{\alpha,\gamma^-}(b_j, \boldsymbol{v_j}) \right| \leq \frac{m L_0(\gamma)}{N}. \qquad (52)$$

Proof: In fact, it is easily seen that we only need to prove that

$$\sum_{t,\boldsymbol{v}\geq 0} |a(b, \boldsymbol{v}) - \kappa_{\alpha,\gamma^-}(b, \boldsymbol{v})| \leq \frac{L_0(\gamma)}{N},$$

with $\boldsymbol{v} = (v_1, \ldots, v_b)$. However, notice that

$$a(b, \boldsymbol{v}) = \binom{M}{b} \left(\frac{\gamma}{N}\right)^b \left(1 - \frac{\gamma}{N}\right)^{M-b} \prod_{l=1}^{b} \binom{N-m}{v_l} \left(\frac{\gamma}{N}\right)^{v_l} \left(1 - \frac{\gamma}{N}\right)^{N-m-v_l},$$

and recall that

$$\kappa_{\alpha,\gamma^-}(b, \boldsymbol{v}) = e^{-\alpha\gamma^-} \frac{(\alpha\gamma^-)^b}{b!}\; e^{-b\gamma^-} \frac{(\gamma^-)^{\sum_{l\leq b} v_l}}{v_1! \cdots v_b!}.$$

Then

$$\sum_{b,\boldsymbol{v}\geq 0} |a(b,\boldsymbol{v}) - \kappa_{\alpha,\gamma^-}(b,\boldsymbol{v})| \leq A + B,$$

with

$$\begin{aligned}
A &= \sum_{b,\boldsymbol{v}\geq 0} \left| e^{-\alpha\gamma^-} \frac{(\alpha\gamma^-)^b}{b!} \bar{A}_{b,\boldsymbol{v}} \right|, \\
\bar{A}_{b,\boldsymbol{v}} &= \left| e^{-b\gamma^-} \frac{(\gamma^-)^{\sum_{l\leq b} v_l}}{v_1! \cdots v_b!} - \prod_{l=1}^{b} \binom{N-m}{v_l} \left(\frac{\gamma}{N}\right)^{v_l} \left(1-\frac{\gamma}{N}\right)^{N-m-v_l} \right|, \\
B &= \sum_{b,\boldsymbol{v}\geq 0} \left| \bar{B}_b \prod_{l=1}^{b} \binom{N-m}{v_l} \left(\frac{\gamma}{N}\right)^{v_l} \left(1-\frac{\gamma}{N}\right)^{N-m-v_l} \right|, \\
\bar{B}_b &= \left| e^{-\alpha\gamma^-} \frac{(\alpha\gamma^-)^b}{b!} - \binom{M}{b} \left(\frac{\gamma}{N}\right)^{b} \left(1-\frac{\gamma}{N}\right)^{M-b} \right|.
\end{aligned}$$

Now, following the estimates for the approximation of a Poisson distribution by a Binomial given in [10, Lemma 7.4.6], we can bound $\bar{A}_{b,\boldsymbol{v}}$ and $\bar{B}_b$ by a quantity of the form $\frac{c}{N}$. The proof is then easily finished.

□

Let us relate now the law of $(\bar{u}_1, \ldots, \bar{u}_m)$ with $\mu_{\alpha,\gamma^-}^{\otimes m}$.

Lemma 3.13 *We have*

$$d_{\tilde{\Omega}^c}\Big(\mathcal{L}(\bar{u}_1, \ldots, \bar{u}_m), \mu_{\alpha,\gamma^-}^{\otimes m}\Big) \leq c_2(N,m),$$

with

$$\begin{aligned}
c_2(N,m) &= 4\, U_\infty\, e^{2U_\infty}\, \mathbf{E}\left(D(N-m, |(F_{N,M}^m)^c|, |J_m|, \gamma_0)\right) \\
&\quad + 2B_0\, \frac{\mathbf{E}|J_m|^2 - 1}{N-m+1}\, (e^{2U_\infty} - 1) + \frac{2m^2 L_0(\gamma_0)}{N} \\
&\quad + 2U_\infty\, e^{2U_\infty}\, \frac{\gamma_0}{N} \mathbf{E}\left(|R_m - m\alpha|\; |J_m| \exp\left\{\frac{\gamma_0}{N}|R_m - m\alpha|\right\}\right).
\end{aligned}$$

Proof: Notice that, invoking relation (33) and Theorem 3.1, we get

$$\sum_{(b)} \sum_{(\boldsymbol{v})} \Big(\prod_{j=1}^{m} \kappa_{\alpha,\gamma^-}(b_j, \boldsymbol{v_j})\Big) T_{b_1,\boldsymbol{v_1}}(\mu_{\alpha,\gamma^-}) \otimes \cdots \otimes T_{b_m,\boldsymbol{v_m}}(\mu_{\alpha,\gamma^-}) = \mu_{\alpha,\gamma^-}^{\otimes m}.$$

Then, the results follows easily from Lemmas 3.11 and 3.12, Lemma 7.3.3 in [10] and the triangular inequality.

□

We are now ready to end the proof of the main result concerning the magnetization of the system.

Proof of Theorem 3.5: First of all, notice that by symmetry we have

$$\mathcal{L}(\langle\sigma_1\rangle,\ldots,\langle\sigma_m\rangle) = \mathcal{L}(\langle\sigma_{N-m+1}\rangle,\ldots,\langle\sigma_N\rangle).$$

Furthermore, thanks to (49) and (46) and Lemma 3.3, we can write

$$\begin{aligned}
D(N,M,m,\gamma_0) &= \sup_{\gamma\le\gamma_0} d\left(\mathcal{L}(\langle\sigma_1\rangle,\ldots,\langle\sigma_m\rangle),\mu_{\alpha,\gamma}^{\otimes m}\right)\\
&\le \sup_{\gamma\le\gamma_0} d_{\tilde\Omega^c}\left(\mathcal{L}\left(\frac{\langle\mathbf{Av}\sigma_{N-m+1}\xi\rangle_-}{\langle\mathbf{Av}\xi\rangle_-},\ldots,\frac{\langle\mathbf{Av}\sigma_N\xi\rangle_-}{\langle\mathbf{Av}\xi\rangle_-}\right),\mu_{\alpha,\gamma}^{\otimes m}\right)\\
&\quad+\frac{2m^3\alpha\gamma_0^2(1+\alpha\gamma_0^2)}{N}\\
&\le \sup_{\gamma\le\gamma_0} d_{\tilde\Omega^c}\left(\mathcal{L}\left(\frac{\langle\mathbf{Av}\sigma_{N-m+1}\xi\rangle_-}{\langle\mathbf{Av}\xi\rangle_-},\ldots,\frac{\langle\mathbf{Av}\sigma_N\xi\rangle_-}{\langle\mathbf{Av}\xi\rangle_-}\right),\mu_{\alpha,\gamma^-}^{\otimes m}\right)\\
&\quad+\frac{2m^3\alpha\gamma_0^2(1+\alpha\gamma_0^2)}{N}\\
&\quad+\frac{4m\alpha\gamma_0}{N}\left(\gamma_0\exp\left\{\frac{m\gamma_0}{N}\right\}+\exp\left\{\frac{m\alpha\gamma_0}{N}\right\}\right).
\end{aligned}$$

Then, Lemma 3.13 implies

$$\begin{aligned}
D(N,M,m,\gamma_0) &\le 4\;U_\infty\;e^{2U_\infty}\;\mathbf{E}\left(D(N-m,|(F_{N,M}^m)^c|,|J_m|,\gamma_0)\right)\\
&\quad+2B_0\;\frac{\mathbf{E}|J_m|^2-1}{N-m+1}\;(e^{2U_\infty}-1)\\
&\quad+2U_\infty\;e^{2U_\infty}\;\frac{\gamma_0}{N}\mathbf{E}\left(|R_m-m\alpha|\;|J_m|\exp\left\{\frac{\gamma_0}{N}|R_m-m\alpha|\right\}\right)\\
&\quad+\frac{2m^2L_0(\gamma_0)}{N}+\frac{12m^3\alpha\gamma_0^4\exp(\gamma_0)}{N}.
\end{aligned}$$

It is readily checked, as we did in (29), that

$$\begin{aligned}
\mathbf{E}\left(|J_m|\right) &\leq \frac{N-m}{N}\ \alpha m\gamma_0^2,\\
\mathbf{E}\left(|J_m|^2\right) &\leq \frac{N-m}{N}(\gamma_0+\gamma_0^2)\big(\alpha m\gamma_0+(\alpha m\gamma_0)^2\big),\\
\mathbf{E}\left(|J_m|^3\right) &\leq \frac{N-m}{N}(\gamma_0+3\gamma_0^2+\gamma_0^3)\big(\alpha m\gamma_0+3(\alpha m\gamma_0)^2+(\alpha m\gamma_0)^3\big).
\end{aligned}$$

Thus, using the fact that $R_m \leq Y$ where $Y \sim B(mM, \frac{\gamma}{N})$, together with the trivial bound $R_m \leq M$, there exists a constant $K_0 \geq 1$ such that

$$\begin{aligned}
D(N,M,m,\gamma_0) &\leq 4\ U_\infty\ e^{2U_\infty}\mathbf{E}\left(D(N-m,|(F^m_{N,M})^c|,|J_m|,\gamma_0)\right)\\
&\quad+\frac{K_0 m^3\left[\alpha\gamma_0^4\exp(\frac{3}{2}\gamma_0)+L_0(\gamma_0)\right]}{N}.
\end{aligned}\tag{53}$$

Now we are able to prove, by induction over N, that

$$D(N,M,m,\gamma_0) \leq \frac{2K_0 m^3\left[\alpha\gamma_0^4\exp(\frac{3}{2}\gamma_0)+L_0(\gamma_0)\right]}{N}, \qquad \text{for all} \quad m \leq \frac{N}{2}.$$

Indeed, in order to check the induction step from $N-1$ to N, notice that $|(F^m_{N,M})^c| \leq M$ and that

$$\mathbf{E}\left(|J_m|^3\right) \leq 25\frac{N-m}{N}(m^3\alpha\gamma_0^6).$$

So, using also that

$$P\left(|J_m| \geq \frac{N}{2}\right) \leq \frac{4}{N^2}\mathbf{E}\left(|J_m|^2\right) \leq \frac{16m^2\alpha\gamma_0^4}{N^2},$$

and by our induction hypothesis and (53), we have

$$\begin{aligned}
&D(N,M,m,\gamma_0) \leq \frac{K_0 m^3\left[\alpha\gamma_0^4\exp(\frac{3}{2}\gamma_0)+L_0(\gamma_0)\right]}{N}\\
&+4\ U_\infty\ e^{2U_\infty}\left(\frac{2K_0\mathbf{E}\left(|J_m|^3\right)\left[\frac{M}{N-m}\ \gamma_0^4\exp(\frac{3}{2}\gamma_0)+L_0(\gamma_0)\right]}{N-m}+\frac{32m^3\alpha\gamma_0^4}{N^2}\right)\\
&\leq \frac{K_0 m^3\left[\alpha\gamma_0^4\exp(\frac{3}{2}\gamma_0)+L_0(\gamma_0)\right]}{N}\\
&+4\ U_\infty\ e^{2U_\infty}\left(\frac{50K_0 m^3\alpha\gamma_0^6\left[2\alpha\ \gamma_0^4\exp(\frac{3}{2}\gamma_0)+L_0(\gamma_0)\right]}{N}+\frac{32m^3\alpha\gamma_0^4}{N^2}\right).
\end{aligned}$$

Finally, since $M < N-m$, the proof easily follows from hypothesis (43). □

4 Replica symmetric formula

Now that the limiting law of the magnetization has been computed, we can try to evaluate the asymptotic behavior of the free energy of our system, namely

$$p_N(\gamma) = \frac{1}{N}\mathbf{E}\left[\log\left(\sum_{\boldsymbol{\sigma}\in\Sigma_N}\exp\left(-H_{N,M}(\boldsymbol{\sigma})\right)\right)\right].$$

To this purpose, set

$$G(\gamma) = \alpha\log\left(\sum_{p=0}^{\infty}\exp(-\gamma)\frac{\gamma^p}{p!}\mathbf{E}\left[\frac{\bar{V}_{p+1}}{\bar{V}_p}\right]\right),$$

where

$$\bar{V}_p := \int\left\langle\exp\left(u\Big(\sum_{i\leq p} g_{i,M}\sigma_i\Big)\right)\right\rangle_{(x_1,\ldots,x_p)} d\mu_{\alpha,\gamma}(x_1)\times\ldots\times d\mu_{\alpha,\gamma}(x_p)$$

and $\langle\cdot\rangle_x$ means integration with respect to the product measure ν on $\{-1,1\}^p$ such that $\int\sigma_i d\nu = x_i$. Then, the main result of this part states that:

Theorem 4.1 *Set F such that $F'(\gamma) = G(\gamma)$ and $F(0) = \log 2 - \alpha u(0)$. Then, if $\gamma\leq\gamma_0$ and (42) and (43) hold true, we have*

$$|p_N(\gamma) - F(\gamma)| \leq \frac{K}{N},$$

where K does not depend on γ and N.

Since $p_N(0) = \log 2 - \alpha u(0)$, the proof of the theorem is a consequence of the following proposition.

Proposition 4.2 *If $\gamma\leq\gamma_0$ and (42) and (43) hold, we have*

$$|p'_N(\gamma) - G(\gamma)| \leq \frac{K}{N},$$

where $p'_N(\gamma)$ is the right derivative of $p_N(\gamma)$.

Proof: We divide the proof into two steps.

Step 1: We will check that

$$|p_N'(\gamma) - G^1(\gamma)| \leq \frac{K}{N}. \tag{54}$$

where $G^1(\gamma)$ is defined as

$$\alpha \mathbf{E}\left[\log\left\langle \exp\left(u\Big(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i + g_{N,M}\sigma_N\Big) - u\Big(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i\Big)\right)\right\rangle\right].$$

Following the method used in Lemma 7.4.11 in [10], we introduce the Hamiltonians

$$-H^1_{N,M}(\boldsymbol{\sigma}) = \sum_{k\leq M} u\left(\sum_{i\leq N} g_{i,k}\ (\gamma_{i,k} + \delta_{i,k})\ \sigma_i\right),$$

$$-H^2_{N,M}(\boldsymbol{\sigma}) = \sum_{k\leq M} u\left(\sum_{i\leq N} g_{i,k}\ \min(1, (\gamma_{i,k} + \delta_{i,k}))\ \sigma_i\right),$$

where $\{\delta_{i,k}\}_{1\leq i\leq N, 1\leq k\leq M}$ is a family of i.i.d. random variables with $P(\delta_{i,k} = 1) = \frac{\delta}{N}$, $P(\delta_{i,k} = 0) = 1 - \frac{\delta}{N}$. We also assume that this sequence is independent of all the random sequences previously introduced. Observe that the random variables $\min(1, (\gamma_{i,k} + \delta_{i,k}))$ are i.i.d with Bernoulli law of parameter $\frac{\gamma'}{N}$, where $\gamma' \equiv \gamma + \delta - \frac{\gamma\delta}{N}$. Set now, for $j = 1, 2$,

$$p_N^j(\delta) = \frac{1}{N}\mathbf{E}\left[\log\left(\sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp\left(-H^j_{N,M}(\boldsymbol{\sigma})\right)\right)\right].$$

Obviously, $p_N^2(\delta) = p_N(\gamma')$, and our first task will be to show that $p_N^1(\delta) - p_N^2(\delta)$ is of order δ^2: notice that

$$p_N^1(\delta) - p_N^2(\delta) = \frac{1}{N}\mathbf{E}\left[\log\langle\exp\left(-H^1_{N,M}(\boldsymbol{\sigma}) + H^2_{N,M}(\boldsymbol{\sigma})\right)\rangle_2\right],$$

where $\langle\cdot\rangle_2$ denotes the average for the Gibbs' measure defined by the Hamiltonian $H^2_{N,M}$. Consider now $Y^1_{N,M} = \sum_{i,k}\gamma_{i,k}\delta_{i,k}$. Since, $\gamma_{i,k}+\delta_{i,k} = \min(1, \gamma_{i,k}+$

$\delta_{i,k}) + \gamma_{i,k}\delta_{i,k}$, on the set $\{Y^1_{N,M} = 0\}$, we have $H^1_{N,M} = H^2_{N,M}$. So, we can write

$$\begin{aligned} p^1_N(\delta) - p^2_N(\delta) &= \frac{1}{N}\mathbf{E}\left[\mathbf{1}_{\{Y^1_{N,M}=1\}} \log\langle \exp\left(-H^1_{N,M}(\boldsymbol{\sigma}) + H^2_{N,M}(\boldsymbol{\sigma})\right)\rangle_2\right] \\ &+\frac{1}{N}\mathbf{E}\left[\mathbf{1}_{\{Y^1_{N,M}\geq 2\}} \log\langle \exp\left(-H^1_{N,M}(\boldsymbol{\sigma}) + H^2_{N,M}(\boldsymbol{\sigma})\right)\rangle_2\right]. \end{aligned}$$

Using that

$$P(Y_{N,M} \geq 2) = 1 - \left(1 - \frac{\gamma\delta}{N^2}\right)^{NM} - NM\left(1 - \frac{\gamma\delta}{N^2}\right)^{NM-1}\frac{\gamma\delta}{N^2} \leq \alpha^2\delta^2\gamma^2$$

and

$$P(Y^1_{N,M} = 1) = NM\left(1 - \frac{\gamma\delta}{N^2}\right)^{NM-1}\frac{\gamma\delta}{N^2} \leq \alpha\gamma\delta,$$

it is easily checked that

$$\lim_{\delta \longrightarrow 0^+} \frac{p^1_N(\delta) - p^2_N(\delta)}{\delta} \leq \frac{K}{N}, \tag{55}$$

which means that we can evaluate the difference $p^1_N(\delta) - p_N(\gamma)$ instead of $p^2_N(\delta) - p_N(\gamma)$.

However, following the same arguments as above, we can write

$$p^1_N(\delta) - p_N(\gamma) = \frac{1}{N}\mathbf{E}\left[\log\langle \exp\left(-H^1_{N,M}(\boldsymbol{\sigma}) + H_{N,M}(\boldsymbol{\sigma})\right)\rangle\right].$$

We consider now $Y_{N,M} = \sum_{i,k}\delta_{i,k}$. Notice that on the set $\{Y_{N,M} = 0\}$, $H_{N,M} = H^1_{N,M}$. So, we can write

$$\begin{aligned} p^1_N(\delta) - p_N(\gamma) &= \frac{1}{N}\mathbf{E}\left[\mathbf{1}_{\{Y_{N,M}=1\}} \log\langle \exp\left(-H^1_{N,M}(\boldsymbol{\sigma}) + H_{N,M}(\boldsymbol{\sigma})\right)\rangle\right] \\ &+\frac{1}{N}\mathbf{E}\left[\mathbf{1}_{\{Y_{N,M}\geq 2\}} \log\langle \exp\left(-H^1_{N,M}(\boldsymbol{\sigma}) + H_{N,M}(\boldsymbol{\sigma})\right)\rangle\right] \\ &\equiv V_1(\delta) + V_2(\delta). \end{aligned}$$

Let us bound now $V_1(\delta)$ and $V_2(\delta)$: since

$$P(Y_{N,M} \geq 2) = 1 - \left(1 - \frac{\delta}{N}\right)^{NM} - NM\left(1 - \frac{\delta}{N}\right)^{NM-1}\frac{\delta}{N} \leq \alpha(NM-1)\delta^2,$$

we have

$$|V_2(\delta)| \leq 2\alpha^2(NM-1)U_\infty\delta^2.$$

On the other hand, using a symmetry argument, we get

$$V_1(\delta) = NM\left(1-\frac{\delta}{N}\right)^{NM-1}\frac{\delta}{N^2}$$
$$\times\, \mathbf{E}\left[\log\left\langle \exp\left(u\Big(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i + g_{N,M}\sigma_N\Big) - u\Big(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i\Big)\right)\right\rangle\right].$$

Hence, we obtain that

$$\lim_{\delta\longrightarrow 0^+}\frac{p_N^1(\delta)-p_N(\gamma)}{\delta} = \lim_{\delta\longrightarrow 0^+}\frac{V_1(\delta)+V_2(\delta)}{\delta} \tag{56}$$
$$= \alpha\mathbf{E}\left[\log\left\langle \exp\left(u\Big(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i + g_{N,M}\sigma_N\Big) - u\Big(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i\Big)\right)\right\rangle\right].$$

Finally, since

$$p_N'(\gamma) = \lim_{\gamma'\longrightarrow\gamma^+}\frac{p_N(\gamma')-p_N(\gamma)}{\gamma'-\gamma} = \lim_{\delta\longrightarrow 0^+}\frac{p_N^2(\delta)-p_N(\gamma)}{\delta\left(1-\frac{\gamma}{N}\right)},$$

putting together (55) and (56), we obtain (54).

Step 2: Let us check now that

$$|G(\gamma)-G^1(\gamma)| \leq \frac{K}{N}. \tag{57}$$

To this purpose, set

$$\Psi := \left\langle \exp\left(u\Big(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i + g_{N,M}\sigma_N\Big) - u\Big(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i\Big)\right)\right\rangle,$$

and let us try to evaluate first $\mathbf{E}[\Psi]$: notice that

$$\begin{aligned}\Psi &= \frac{\sum_{\boldsymbol{\sigma}\in\Sigma_N}\exp\left(u(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i + g_{N,M}\sigma_N)\right)\exp\left(-H_{N,M-1}(\boldsymbol{\sigma})\right)}{\sum_{\boldsymbol{\sigma}\in\Sigma_N}\exp\left(-H_{N,M}(\boldsymbol{\sigma})\right)}\\ &= \frac{\langle\exp\left(u(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i + g_{N,M}\sigma_N)\right)\rangle_{M-1}}{\langle\exp\left(u(\sum_{i\leq N} g_{i,M}\gamma_{i,M}\sigma_i)\right)\rangle_{M-1}},\end{aligned}$$

where $\langle\cdot\rangle_{M-1}$ denotes the usual average using the Hamiltonian $H_{N,M-1}$. Set $B_p := \{\sum_{i=1}^{N-1}\gamma_{i,M} = p, \gamma_{N,M} = 0\}$ and $B := \{\gamma_{N,M} = 1\}$, and let us denote by $\mathbf{E}_M$ the conditional expectation given $\{\gamma_{i,M}, 1 \leq i \leq N\}$. Then

$$\begin{aligned}\mathbf{E}[\Psi] &= \mathbf{E}\left[\sum_{p=0}^{N-1}\mathbf{1}_{B_p}\mathbf{E}_M[\Psi]\right] + \mathbf{E}[\mathbf{1}_B\mathbf{E}_M[\Psi]] \qquad (58)\\ &= \sum_{p=0}^{N-1}\binom{N-1}{p}\left(\frac{\gamma}{N}\right)^p\left(1-\frac{\gamma}{N}\right)^{N-p+1}\mathbf{E}\left[\frac{\langle\exp(V_{p+1})\rangle_{M-1}}{\langle\exp(V_p)\rangle_{M-1}}\right] + \frac{\gamma}{N}e^{2U_\infty},\end{aligned}$$

where

$$V_p := u\left(\sum_{i\leq p} g_{i,M}\sigma_i\right).$$

Set $\boldsymbol{X}_p = (\langle\sigma_1\rangle, \ldots, \langle\sigma_p\rangle)$. Then, using the triangular inequality and following the same arguments as in Proposition 2.8, we get, for a strictly positive constant K,

$$\begin{aligned}&\left|\mathbf{E}\left[\frac{\langle\exp(V_{p+1})\rangle_{M-1}}{\langle\exp(V_p)\rangle_{M-1}}\right] - \mathbf{E}\left[\frac{\langle\exp(V_{p+1})\rangle_{\boldsymbol{X}_{p+1}}}{\langle\exp(V_p)\rangle_{\boldsymbol{X}_p}}\right]\right|\\ &= \left|\mathbf{E}\left[\frac{\langle\exp(V_{p+1})\rangle_{M-1}\langle\exp(V_p)\rangle_{\boldsymbol{X}_p} - \langle\exp(V_{p+1})\rangle_{\boldsymbol{X}_{p+1}}\langle\exp(V_p)\rangle_{M-1}}{\langle\exp(V_p)\rangle_{M-1}\langle\exp(V_p)\rangle_{\boldsymbol{X}_p}}\right]\right|\\ &\leq e^{3U_\infty}\mathbf{E}\big[\big|\langle\exp(V_{p+1})\rangle_{M-1} - \langle\exp(V_{p+1})\rangle_{\boldsymbol{X}_{p+1}}\big|\\ &\qquad + \big|\langle\exp(V_p)\rangle_{M-1} - \langle\exp(V_p)\rangle_{\boldsymbol{X}_p}\big|\big]\\ &\leq e^{3U_\infty}\frac{p^2K}{N}. \qquad (59)\end{aligned}$$

Consider now some i.i.d. random variables $z_1, \ldots, z_p$ of law $\mu_{\alpha,\gamma}$ such that (44) holds. Set $\boldsymbol{Y}_p = (z_1, \ldots, z_p)$. Then, following the same arguments as above, we get, for a strictly positive constant K,

$$\begin{aligned}&\left|\mathbf{E}\left[\frac{\langle\exp(V_{p+1})\rangle_{\boldsymbol{X}_{p+1}}}{\langle\exp(V_p)\rangle_{\boldsymbol{X}_p}}\right] - \mathbf{E}\left[\frac{\langle\exp(V_{p+1})\rangle_{\boldsymbol{Y}_{p+1}}}{\langle\exp(V_p)\rangle_{\boldsymbol{Y}_p}}\right]\right|\\ &\leq e^{3U_\infty}\mathbf{E}\big[\big|\langle\exp(V_{p+1})\rangle_{\boldsymbol{X}_{p+1}} - \langle\exp(V_{p+1})\rangle_{\boldsymbol{Y}_{p+1}}\big|\\ &\qquad + \big|\langle\exp(V_p)\rangle_{\boldsymbol{X}_p} - \langle\exp(V_p)\rangle_{\boldsymbol{Y}_p}\big|\big]\\ &\leq e^{3U_\infty}\frac{p^3K}{N}, \qquad (60)\end{aligned}$$

where in the last inequality we have used (44) and the fact that

$$\frac{\partial}{\partial x_i}\langle \exp(V_p)\rangle_x \le e^{U_\infty}.$$

Notice that if W is a random variable with law $\mathrm{Bin}(N-1, \frac{\gamma}{N})$, then $\mathbf{E}(W^3) \le K$, where K does not depend on N. So, putting together (58), (59) and (60), we get

$$\mathbf{E}\left[\Psi\right] = \sum_{p=0}^{N-1} \binom{N-1}{p} \left(\frac{\gamma}{N}\right)^p \left(1-\frac{\gamma}{N}\right)^{N-p} \mathbf{E}\left[\frac{\langle \exp(V_{p+1})\rangle_{\boldsymbol{Y}_{p+1}}}{\langle \exp(V_p)\rangle_{\boldsymbol{Y}_p}}\right] + \frac{K}{N}.$$

Using now similar arguments to those ones used in the proof of Lemma 3.11, we get

$$\mathbf{E}\left[\Psi\right] = \sum_{p=0}^{\infty} \exp(-\gamma)\frac{\gamma^p}{p!}\mathbf{E}\left[\frac{\langle \exp(V_{p+1})\rangle_{\boldsymbol{Y}_{p+1}}}{\langle \exp(V_p)\rangle_{\boldsymbol{Y}_p}}\right] + \frac{K}{N}. \tag{61}$$

Finally, once (61) has been obtained, (57) can be established following the method used in Proposition 7.4.10 in [10], the remaining details being left to the reader. □

References

[1] Aldous, D.: The $\zeta(2)$ limit in the random assignment problem. *Random Structures Algorithms* **18** (2001), no. 4, 381–418.

[2] Bardina, X.; Márquez-Carreras, D.; Rovira, C.; Tindel, S.: The p-spin interaction model with external field. *Potential Analysis* **21** no. 4 (2004) 311–362.

[3] Cont, R.; Lowe, M.: Social distance, heterogeneity and social interactions. Preprint.

[4] Franz, S.; Toninelli, F.: The Kac limit for diluted spin glasses. *Internat. J. Modern Phys. B* **18** (2004), no. 4-5, 675–679.

[5] Guerra, F.; Toninelli, F.: The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.* **230** (2002), no. 1, 71–79.

[6] Hertz, J.; Krogh, A.; Palmer, R.: *Introduction to the Theory of Neural Computation*. Addison-Wesley Publishing Company, 1991.

[7] Márquez-Carreras, D.; Rovira, C.; Tindel, S.: Assymptotic behavior of the magnetization for the perceptron model.*Ann. Inst. H. Poincaré Probab. Statist.* **42** (2006), no. 3, 327–342.

[8] Mezard, M.; Parisi, G.; Virasoro, M.A.: *Spin glass theory and beyond.* World Scientific, 1987.

[9] Shcherbina, M.; Tirozzi, B.: Rigorous solution of the Gardner problem. *Comm. Math. Phys.* **234** (2003), no. 3, 383–422.

[10] Talagrand, M.: *Spin Glasses: a Challenge for Mathematicians.* Springer, 2003.

[11] Talagrand, M.: The Parisi formula. *Ann. of Math.* **163** (2006), no. 1, 221–263.