A mean field model for spin glasses based on a 2-level perceptron.

Michel Talagrand [1]
Université Paris VI and the Ohio-State University

Abstract. We introduce a mean field model for spin glasses that is obtained by "iterating" the definition of the perceptron model of [SG] and we prove the validity of the replica-symmetric solution under a condition of the type "high temperature". The replica-symmetric equations involve 6 different parameters. This model illustrates on the one hand that we now have efficient tools to deal with such systems, and on the other hand that it is probably unreasonable to hope for abstract theorems independent of the specific type of the system.

1. Introduction

One of the main motivations for the mathematical study of mean field models for spin glasses is that apparently some new laws of nature (or, at least, of probability) are at work there. A general fact seems to be that "at high temperatures the replica-symmetric solution is correct". This is proved in [2] for the main models considered by the physicists, and it would be very nice to have a general statement to that effect. One of the models considered in [2] is (a suitable abstract version) of the Perceptron model, which is defined as follows.

For $\boldsymbol{\sigma} = (\sigma_i)_{i\le N} \in \{-1,1\}^N$, consider

$$S_k = S_k(\boldsymbol{\sigma}) = \frac{1}{\sqrt{N}} \sum_{i\le N} g_{i,k}\sigma_i, \tag{1.1}$$

where $g_{i,k}$, $i,k \ge 1$, are independent standard normal r.v. (The normalization is chosen so that S_k is typically of order 1.) Then the Hamiltonian is given by

$$-H = \sum_{k\le M} u(S_k), \tag{1.2}$$

when $u : \mathbb{R} \to \mathbb{R}$ is a function. It helps to think of u as being of order 1 and of M as being a proportion of N. In [2] a lot of technical work is devoted to the study of this Hamiltonian under weak regularity conditions on u. However, when one assumes stronger conditions, such as

$$\forall \ell, 0 \le \ell \le 10, \quad |u^{(\ell)}| \le D, \tag{1.3}$$

where D is a parameter, the Hamiltonian (1.2) is already of interest, and its study is technically much easier. (It will be sketched in this case in Section 2.) One of the striking features of the Hamiltonian (1.2) is that

[1] Work partially supported by an NSF grant

(under a suitable condition of the nature of a high-temperature condition), not only is the replica symmetric solution correct, by the quantities S_k (and $u(S_k)$) themselves display certain properties of spins (they "decorrelate"). This raises the possibility that it might be of interest to iterate the process by which one goes from the spins (σ_i) to H. To do this, we consider the quantities

$$T_m = \frac{1}{\sqrt{M}} \sum_{k \le M} h_{k,m} u(S_k) \tag{1.4}$$

where $h_{k,m}$ $k, m \ge 1$ are independent standard normal r.v., that are independent of the $g_{i,k}$. Considering a function $v : \mathbb{R} \to \mathbb{R}$ we will then consider the Hamiltonian

$$-H = \sum_{m \le Q} v(T_m) \tag{1.5}$$

and it is the purpose of this paper to investigate it. In principle "level 2 perceptrons"are important [1], but the present paper represents only a first step in the direction of their rigorous study. Really relevant results remain far away. As the reader will soon realize, the Hamiltonian (1.5) is quite more complicated to study than the Hamiltonian (1.2). Thereby it is legitimate in a first stage to study it under the strongest regularity conditions where it is of interest. We will assume that for a certain number D, we have

$$\forall \ell, 0 \le \ell \le 10, \quad |u^{(\ell)}| \le D, \quad |v^{(\ell)}| \le D. \tag{1.6}$$

The following theorem involves numbers $\overline{r} \ge r \ge 0$, $0 \le q \le 1$, $\tau, \rho \ge 0$ and $\overline{\tau}$. We consider independent standard normal r.v. z, η, ξ and $\xi_\ell, \ell \ge 1$. We consider the r.v.

$$\theta = z\sqrt{q} + \xi\sqrt{1-q}; \ \theta_\ell = z\sqrt{q} + \xi_\ell\sqrt{1-q}; \ \gamma = z\sqrt{r} + \xi\sqrt{\overline{r} - r}. \tag{1.7}$$

We denote by E_ξ expectation in ξ, ξ_ℓ only.

Throughout the paper we denote by K a number depending on D only, that need not be the same at each occurrence.

Theorem 1.1 *There exists a number K (depending on D only) such that if*

$$K\frac{Q}{M} \le 1, \quad K\frac{M}{N} \le 1 \tag{1.8}$$

there is a unique solution to the following system of equations:

$$r = E\left(\frac{E_\xi u(\theta)\exp(u(\theta)\eta\sqrt{\tau} + (\overline{\tau}-\tau)u(\theta)^2/2)}{E_\xi \exp(u(\theta)\eta\sqrt{\tau} + (\overline{\tau}-\tau')u(\theta)^2/2)}\right)^2 \tag{1.9}$$

$$\overline{r} = \ E\ \frac{u^2(\theta)\exp(u(\theta)\eta\sqrt{\tau} + (\overline{\tau}-\tau)u(\theta)^2/2)}{E_\xi \exp(u(\theta)\eta\sqrt{\tau} + (\overline{\tau}-\tau')u(\theta)^2/2)} \tag{1.10}$$

$$\tau = \frac{Q}{M}\ E\left(\frac{E_\xi v'(\gamma)\exp v(\gamma)}{E_\xi \exp v(\gamma)}\right)^2 \tag{1.11}$$

$$\overline{\tau} = \frac{Q}{M}\ E\ \frac{(v''(\gamma)+v'(\gamma)^2)\exp v(\gamma)}{E_\xi \exp v(\gamma)} \tag{1.12}$$

$$\rho = \ \frac{M}{N}\ \Big(\tau\ E\left(\frac{E_\xi u'(\theta)\exp u(\theta)}{E_\xi \exp u(\theta)}\right)^2 + \tau^2 W(1,2) + 2\tau\overline{\tau}W(1,1) \tag{1.13}$$
$$-4\tau(\tau+\overline{\tau})W(1,3) - 2\tau^2 W(3,3) + 6\tau^2 W(3,4)\Big),$$

where, for ℓ, $\ell' \leq 4$

$$W(\ell,\ell') = E\frac{u'(\theta_1)u'(\theta_2)u(\theta_\ell)u(\theta_{\ell'})\mathcal{E}}{E_\xi\mathcal{E}} \tag{1.14}$$

for

$$\mathcal{E} = \exp\sum_{\ell\leq 4}\Big(u(\theta_\ell)\eta\sqrt{\tau} + \frac{\overline{\tau}-\tau}{2}u(\theta_\ell)^2\Big)$$

and

$$q = E\mathrm{th}^2(z\sqrt{\rho})\ . \tag{1.15}$$

These formulas, and in particular (1.13), should make self-apparent that the model is really non-trivial. In fact, we see no apparent limits to the complexity of what could be done using further iterations of the method we used to construct our Hamiltonian.

We denote by $\langle\cdot\rangle$ an average for the Gibbs measure G with Hamiltonian (1.5). We will very often consider replicas, that is powers of the configuration space $\{-1,1\}^N$ provided with the corresponding power of G. When we consider configurations $\boldsymbol{\sigma}^1$, $\boldsymbol{\sigma}^2$, ... and write $\langle f(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2,\dots)\rangle$ we always assume that each configuration $\boldsymbol{\sigma}^\ell$ is averaged independently for G. To shorten notation we write

$$S_k^\ell = S_k(\boldsymbol{\sigma}^\ell);\quad T_m^\ell = T_m(\boldsymbol{\sigma}^\ell), \tag{1.16}$$

and also

$$\nu(f) = E\langle f\rangle, \tag{1.17}$$

where E denotes expectation in the r.v. $g_{i,k}$ and $h_{k,m}$.

Theorem 1.2 *There exists a number K (depending on D only) such that if*

$$K\ \frac{Q}{M} \leq 1 \quad K\ \frac{M}{N} \leq 1 \tag{1.18}$$

the equations (1.9)-(1.15) have a unique solution and that

(1.19)
$$\nu\left(\left(\frac{1}{N}\sum_{i\le N}\sigma_i^1\sigma_i^2-q\right)^2\right)\le\frac{K}{M}$$

(1.20)
$$\nu\left(\left(\frac{1}{M}\sum_{k\le M}u(S_k^1)u(S_k^2)-r\right)^2\right)\le\frac{K}{M}$$

(1.21)
$$\nu\left(\left(\frac{1}{M}\sum_{k\le M}u(S_k)^2-\overline{r}\right)^2\right)\le\frac{K}{M}$$

(1.22)
$$\nu\left(\left(\frac{1}{M}\sum_{m\le Q}v'(T_m^1)v'(T_m^2)-\tau\right)^2\right)\le\frac{K}{M}$$

(1.23)
$$\nu\left(\left(\frac{1}{M}\sum_{m\le Q}(v'(T_m)^2+v''(T_m))-\overline{\tau}\right)^2\right)\le\frac{K}{M}.$$

Of course, these statements explain the meaning of the various parameters. Our last result concerns the computation of the "pressure".

Theorem 1.3. *Under the conditions of Theorem 3 we have*

$$\left|\frac{1}{N}E\ \log\sum_{\boldsymbol{\sigma}}\exp(-H(\boldsymbol{\sigma}))-F\right|\le\frac{K}{M}$$

where

(1.24)
$$\begin{aligned}F=&\log 2+\frac{Q}{N}E\log E_\xi\exp v(z\sqrt{r}+\xi\sqrt{\overline{r}-r})\\&+\frac{M}{N}E\log E_\xi\exp\Big(u(z\sqrt{q}+\xi\sqrt{1-q})\eta\sqrt{z}+\frac{\overline{\tau}-\tau}{2}u^2(z\sqrt{q}+\xi\sqrt{1-q})\Big)\\&+E\,\mathrm{ch}\,z\sqrt{\rho}-\frac{(1-q)\rho}{2}+\frac{\tau r-\overline{\tau}\overline{r}}{2}.\end{aligned}$$

Since the proof of Theorem 1.1 is very simple, we give it right away.

Proof of Theorem 1.1. Let $\rho'=\frac{N}{M}\rho$. Then equations (1.9) to (1.15) are of the type

(1.25)
$$r=F_1(\tau,\overline{\tau},q);\ \overline{r}=F_2(\tau,\overline{\tau},q);\ \rho'=F_3(\tau,\overline{\tau},q)$$

(1.26)
$$\overline{\tau}=\frac{Q}{M}F_4(r,\overline{r});\ \tau=\frac{Q}{M}F_5(r,\overline{r});\ q=F_6\left(\frac{M}{N}\rho'\right)$$

for certain functions $F_1,\ldots,F_6$.

Substitutions of the relations (1.26) into the equations (1.25) yields equations of the type

(1.27)
$$r=G_1(r,\overline{r},\rho');\ \overline{r}=G_2(r,\overline{r},\rho');\ \rho'=G_3(r,\overline{r},\rho'),$$

for certain functions G_1, G_2, G_3. It is straightforward (but tedious) to see that all the first-order partial derivatives of G_1, G_2, G_3 are bounded by

$K(Q/M+M/N)$, where K depends on D only. Therefore when $K(Q/M+M/N)\le 1$, these equations have a unique solution, and this is also the case of the equations (1.25) and (1.26). $\square$

2. The case of the perceptron.

Not surprisingly, the methods required to handle the Hamiltonian (1.5) are a kind of "iteration" of the methods appropriate to study the Hamiltonian (1.2). It therefore seems useful to start the paper by a study of the Hamiltonian (1.2) under condition (1.3), a task that is technically very much easier than the material of [2] (although based on the same principles).

We start by explaining the basic principle on which the entire paper relies. We consider a non-random probability measure G on $\{-1,1\}^N$, and we denote by $\langle\cdot\rangle$ an average for G or for its products $G^{\otimes n}$ on $\{-1,1\}^{Nn}$, where n is a given integer. We consider functions $(a_k)_{k\ge 1}$ on $\{-1,1\}^N$, and independent standard normal r.v. $(g_k)_{k\ge1}$. We consider two functions $w:\mathbb{R}\to\mathbb{R}$ and $W:\mathbb{R}^n\to\mathbb{R}$. We assume that

$$\begin{gathered}\text{The partial derivatives of } w \text{ and } W \text{ of order } \le 2 \text{ are}\\ \text{bounded by } D, \text{where } D \text{ is a parameter.}\end{gathered}\tag{2.1}$$

We consider the random quantity $V=V(\boldsymbol\sigma)=\sum_{k\ge1}g_k a_k(\boldsymbol\sigma)$, and for $(\boldsymbol\sigma^1,\dots,\boldsymbol\sigma^n)\in\{-1,1\}^{Nn}$ we write $V^\ell=V(\boldsymbol\sigma^\ell)$. We consider independent standard normal r.v. z,ξ and $(\xi^\ell)_{\ell\ge1}$ and, for parameters $0\le s\le\overline{s}$ we write $\zeta^\ell=z\sqrt{s}+\xi^\ell\sqrt{\overline{s}-s}$ and $\zeta=z\sqrt{s}+\xi\sqrt{\overline{s}-s}$. (It is a general rule through the paper that "variables labeled ξ should be replaced by independent copies in different replicas".)

We denote by E_ξ expectation in ξ only, and we consider a function $f=f(\boldsymbol\sigma^1,\dots,\boldsymbol\sigma^n)$.

Lemma 2.1 *We have*

$$\begin{aligned}&\left|E\frac{\langle W(V^1,\dots,V^n)f\rangle}{\langle\exp w(V)\rangle^n}-E\ \frac{W(\zeta^1,\dots,\zeta^n)}{(E_\xi\exp w(\zeta))^n}\langle f\rangle\right|\le K(n,D)\langle f^2\rangle^{1/2}\\ &\quad\times\left[\left\langle\Big(\sum_k a_k^2(\boldsymbol\sigma)-\overline{s}\Big)^2\right\rangle^{1/2}+\left\langle\Big(\sum_k a_k(\boldsymbol\sigma^1)a_k(\boldsymbol\sigma^2)-s\Big)^2\right\rangle^{1/2}\right],\end{aligned}\tag{2.2}$$

where $K(n,D)$ depends only on n and D.

Proof. We consider the interpolation parameter $0<c<1$ and we define $V_c^\ell=\sqrt{c}\,V^\ell+\sqrt{1-c}\,\zeta^\ell$, $V_c=\sqrt{c}\,V+\sqrt{1-c}\,\zeta$. We define

$$\varphi(c)=E\frac{\langle W(V_c^1,\dots,V_c^n)f\rangle}{(E_\xi\langle\exp w(V_c)\rangle)^n}\tag{2.3}$$

and we use the bound

$$|\varphi(1)-\varphi(0)| \le \sup_c |\varphi'(c)|. \tag{2.4}$$

We note that the left-hand side of (2.2) is $|\varphi(1)-\varphi(0)|$, so all we have to show is that the right-hand side of (2.2) is a bound for $\varphi'(c)$. To do this we compute $\varphi'(c)$ using integration by parts and the formula

$$EV^\ell V^{\ell'} = \sum_k a_k(\boldsymbol{\sigma}^\ell)a_k(\boldsymbol{\sigma}^{\ell'}).$$

The computation and the estimates are straightforward. □

When now G is random, but that the randomness of the g_k is independent of the randomness of G, we can assume without loss of generality that this is also the case of the randomness of z and ξ. Using (2.2) given the randomness of G and using the Cauchy-Schwarz inequality we then get that

(2.5)

$$\left| E\frac{\langle W(V^1,\dots,V^n)f\rangle}{\langle \exp w(V)\rangle^n} - E\,\frac{W(\zeta^1,\dots,\zeta^n)}{(E_\xi \exp w(\zeta))^n}\langle f\rangle \right| \le K(n,D)(E\langle f^2\rangle)^{1/2}$$

$$\times\left[\left(E\left\langle\left(\sum_k a_k^2(\boldsymbol{\sigma})-\overline{s}\right)^2\right\rangle\right)^{1/2} + \left(E\left\langle\left(\sum_k a_k(\boldsymbol{\sigma}^1)a_k(\boldsymbol{\sigma}^2)-s\right)^2\right\rangle\right)^{1/2}\right].$$

We consider an interpolation parameter $0<t<1$ and throughout the paper we write

$$S_{k,t} = S_{k,t}(\boldsymbol{\sigma}) = \frac{1}{\sqrt{N}}\sum_{i\le N-1} g_{i,k}\sigma_i + \sqrt{\frac{t}{N}}\; g_{N,k}\sigma_N. \tag{2.6}$$

We consider the interpolating Hamiltonian

$$-H_t = \sum_{k\le M} u(S_{k,t}) + \sigma_N\sqrt{1-t}\;Y, \tag{2.7}$$

where $Y = z\sqrt{r}$, and the parameter r will be chosen later. For a function f on $\{-1,1\}^{Nn}$, we denote by $\langle f\rangle_t$ its average for the Gibbs measure with Hamiltonian (2.7), and we write $\nu_t(f) = E\langle f\rangle_t$. For a replica $\boldsymbol{\sigma}^\ell$ we write $S^\ell_{k,t} = S_{k,t}(\boldsymbol{\sigma}^\ell)$. We write $\epsilon_\ell = \sigma^\ell_N$. The following is obtained through a straightforward computation, that is detailed in [2], Proposition 3.2.3.

Lemma 2.2 *We have*

$$2\frac{d}{dt}\nu_t(f) = \mathrm{I}+\mathrm{II}+\mathrm{III} \tag{2.8}$$

where

$$\mathrm{I} = \frac{M}{N}\Big(\sum_{\ell\le n} A_\ell - n\;A_{n+1}\Big),$$

for $A_\ell = \nu_t((u''(S^\ell_{M,t}) + u'^2(S^\ell_{M,t}))f)$*;*

$$\mathrm{II} = \frac{M}{N}\Bigg(\sum_{\ell \neq \ell'} A_{\ell,\ell'} - 2n\sum_{\ell \le n} A_{\ell,n+1} + n(n+1)A_{n+1,n+2}\Bigg),$$

for $A_{\ell,\ell'} = \nu_t(u'(S^\ell_{M,t})u'(S^{\ell'}_{M,t})\epsilon_\ell\epsilon_{\ell'}f)$*;*

$$\mathrm{III} = -r\Bigg(\sum_{\ell \neq \ell'} B_{\ell,\ell'} - 2n\sum_{\ell \le n} B_{\ell,n+1} + n(n+1)B_{n+1,n+2}\Bigg),$$

for $B_{\ell,\ell'} = \nu_t(\epsilon_\ell\epsilon_{\ell'}f)$*.*

A very important difference between the terms I and II is that A_ℓ depends on the function f (which does not depend on ℓ) while $A_{\ell,\ell'}$ depends on the function $\epsilon_\ell\epsilon_{\ell'}f$.

A first observation is that, under (1.3) (and provided $|r| \le K(D)$) for $n = 2$ we have

$$\frac{d}{dt}|\nu_t(f)| \le K\nu_t(|f|)$$

so that, by integration,

$$f \ge 0 \Rightarrow \nu_t(f) \le K\nu(f), \tag{2.9}$$

where for simplicity $\nu(f) = \nu_1(f) = E\langle f\rangle$. Let us write

$$R_{1,2} = \frac{1}{N}\sum_{i \le N}\sigma^1_i\sigma^2_i; \quad R^-_{1,2} = \frac{1}{N}\sum_{i \le N-1}\sigma^1_i\sigma^2_i.$$

Our goal is now to prove that

$$\frac{M}{N}K \le 1 \Rightarrow \quad \nu((R_{1,2} - q)^2) \le \frac{K}{N}, \tag{2.10}$$

for a suitable value of q. By symmetry between the coordinates, we have

$$\nu((R_{1,2} - q)^2) = \nu(f), \tag{2.11}$$

where

$$f = (\epsilon_1\epsilon_2 - q)(R_{1,2} - q). \tag{2.12}$$

Now

$$\nu_0(f) = \nu_0\left(\frac{\epsilon_1\epsilon_2 - q}{N}\epsilon_1\epsilon_2\right) + \nu_0((\epsilon_1\epsilon_2 - q)(R^-_{1,2} - q)). \tag{2.13}$$

It is straightforward that if we assume

$$q = E\mathrm{th}^2(Y) = E\mathrm{th}^2(z\sqrt{q}) \tag{2.14}$$

then the last term of (2.13) is zero, so that

$$|\nu_0(f)| \le \frac{2}{N}. \tag{2.15}$$

Thus

$$\nu((R_{1,2} - q)^2) = \nu(f) \le \frac{2}{N} + \sup_t\left|\frac{d}{dt}\nu_t(f)\right|. \tag{2.16}$$

To bound the last term we will apply Lemma 2.1 to each term provided by Lemma 2.2 (with $n=2$ and f given by (2.12). Consider the Hamiltonian

$$-H_\sim = \sum_{k\le M-1} u(S_{k,t}) + \sigma_N\sqrt{1-t}\,Y, \tag{2.17}$$

and denote by $\langle\cdot\rangle_\sim$ an average for the Gibbs measure with this Hamiltonian. Thus, for $\ell\le 3$,

$$\begin{aligned} A_\ell &= E\langle (u''(S^\ell_{M,t}) + u'^2(S^\ell_{M,t}))f\rangle_t \\ &= E\frac{\langle (u''(S^\ell_{M,t}) + u'^2(S^\ell_{M,t}))\exp\big(\sum_{\ell'\le 3} u(S^{\ell'}_{M,t})\big)f\rangle_\sim}{\langle \exp u(S_{M,t})\rangle_\sim^3}. \end{aligned}$$

We use Lemma 2.1 with $a_i(\boldsymbol{\sigma}) = \sigma_i/\sqrt{N}$ for $i\le N-1$ and $a_N(\boldsymbol{\sigma}) = \sigma_N\sqrt{t/N}$, $\bar{s}=1$, $s=q$, and for G the Gibbs measure with Hamiltonian (2.17). Thus from (2.5) we get, using the value of f in the last line,

$$\begin{aligned} &|A_\ell - C_\ell E\langle f\rangle_\sim| \\ &\le K(E\langle f^2\rangle_\sim)^{1/2}\left(\frac{1}{N} + \left(E\left\langle\left(R_{1,2} - q - \frac{\epsilon_1\epsilon_2(1-t)}{N}\right)^2\right\rangle_\sim\right)^{1/2}\right) \\ &\le K\ (E\langle f^2\rangle_\sim)^{1/2}\left(\frac{1}{N} + (E\langle (R_{1,2}-q)^2\rangle_\sim)^{1/2}\right) \\ &\le K\left(\frac{1}{N} + E\langle (R_{1,2}-q)^2\rangle_\sim\right), \end{aligned} \tag{2.18}$$

where

$$C_\ell = E\frac{(u''(\theta^\ell) + u'^2(\theta^\ell))\exp(\sum_{\ell'\le 3} u(\theta^{\ell'}))}{(E_\xi \exp u(\theta))^3},$$

for $\theta^\ell = z\sqrt{q} + \xi^\ell\sqrt{1-q}$, $\theta = z\sqrt{q} + \xi\sqrt{1-q}$. The value of C_ℓ does not depend on ℓ. In fact, using independence we have

$$\begin{aligned} &E_\xi(u''(\theta^\ell) + u'^2(\theta^\ell))\exp(\sum_{\ell'\le 3} u(\theta^{\ell'})) \\ &= E_\xi\big((u''(\theta) + u'^2(\theta))\exp u(\theta)\big)(E_\xi \exp u(\theta))^2, \end{aligned}$$

and thus

$$C_\ell = C := E\frac{(u''(\theta) + u'^2(\theta))\exp u(\theta)}{E_\xi \exp u(\theta)}.$$

It should be obvious that

$$E\langle (R_{1,2}-q)^2\rangle_\sim \le K\nu_t((R_{1,2}-q)^2).$$

Moreover, from (2.9) we have $\nu_t((R_{1,2}-q)^2) \le K\nu((R_{1,2}-q)^2)$. Thus (2.18) yields that for $\ell\le 3$ we have

$$|A_\ell - CE\langle f\rangle_\sim| \le K\left(\frac{1}{N} + \nu((R_{1,2}-q)^2)\right)$$

and hence

$$|\mathrm{I}| \le K\frac{M}{N}\left(\frac{1}{N} + \nu((R_{1,2}-q)^2)\right). \tag{2.19}$$

Proceeding as in (2.18) we have

$$|A_{\ell,\ell'} - C'E\langle f\epsilon_\ell\epsilon_{\ell'}\rangle_\sim| \le K\left(\frac{1}{N} + \nu((R_{1,2}-q)^2)\right) \tag{2.20}$$

where

$$C' = E\frac{u'(\theta^1)u'(\theta^2)\exp(u(\theta^1)+u(\theta^2))}{(E_\xi \exp u(\theta))^2} = E\left(\frac{E_\xi u'(\theta)\exp u(\theta)}{E_\xi \exp u(\theta)}\right)^2.$$

Also, in the same manner,

$$|\nu_t(f\epsilon_\ell\epsilon_{\ell'}) - E\langle f\epsilon_\ell\epsilon_{\ell'}\rangle_\sim| \le K\left(\frac{1}{N} + \nu((R_{1,2}-q)^2)\right).$$

We then see that it is a very good idea to ensure that

$$r = \frac{M}{N}C' \tag{2.21}$$

because then

$$|\mathrm{II}+\mathrm{III}| \le \frac{M}{N}\left(\frac{1}{N} + \nu((R_{1,2}-q)^2)\right).$$

Thus, if q and r are chosen according to (2.14) and (2.21), we see from (2.16) that

$$\nu((R_{1,2}-q)^2) \le \frac{K}{M} + K\frac{M}{N}\nu((R_{1,2}-q)^2),$$

which proves (2.10).

We now turn to the proof of an auxiliary result that will help us to get the correct rate when we compute the "pressure". We prove that

$$K\frac{M}{N} \le 1 \Rightarrow \nu\left(\left(\frac{1}{N}\sum_{k\le M} u'(S_k^1)u'(S_k^2) - r\right)^2\right) \le \frac{K}{N}. \tag{2.22}$$

To do this, we observe that, by symmetry,

$$\begin{aligned}
&\nu\left(\left(\frac{1}{N}\sum_{k\le M} u'(S_k^1)u'(S_k^2) - r\right)^2\right) \\
&= \nu\left(\left(\frac{M}{N}u'(S_M^1)u'(S_M^2) - r\right)\left(\frac{1}{N}\sum_{k\le M} u'(S_k^1)u'(S_k^2) - r\right)\right) \\
&\le \frac{K}{N} + \nu\left(\left(\frac{M}{N}u'(S_M^1)u'(S_M^2) - r\right)f\right)
\end{aligned} \tag{2.23}$$

where $f = N^{-1}\sum_{k\le M-1} u'(S_k^1)u'(S_k^2) - r$. We denote by $\langle\cdot\rangle_\sim$ an average for the Gibbs measure with Hamiltonian $\sum_{k\le M-1} u(S_k)$, so that

$$\begin{aligned}&\nu\Big(\Big(\frac{M}{N}u'(S_M^1)u'(S_M^2) - r\Big)f\Big)\\ &= E\frac{\langle\big(\frac{M}{N}u'(S_M^1)u'(S_M^2) - r\big)\, f\exp(u(S_M^1) + u(S_M^2))\rangle_\sim}{\langle\exp u(S_M)\rangle_\sim^2},\end{aligned}$$

and by using Lemma 2.1 and (2.21) we get that

$$\begin{aligned}\nu\Big(\Big(\frac{M}{N}u'(S_M^1)u'(S_M^2) - r\Big)f\Big) &\le K(E\langle f^2\rangle_\sim^{1/2})(E\langle(R_{1,2}-q)^2\rangle_\sim)^{1/2}\\ &\le \frac{K}{\sqrt{N}}\nu(f^2)^{1/2},\end{aligned} \tag{2.24}$$

using (2.21) and (2.10). Now by the triangle inequality we have

$$\nu(f^2)^{1/2} \le \frac{K}{N} + \nu\bigg(\bigg(\frac{1}{N}\sum_{k\le M} u'(S_k^1)u'(S_k^2) - r\bigg)^2\bigg)^{1/2},$$

and combining with (2.23) and (2.24) we get, using $xy \le cx^2 + y^2/c$, and if $KM \le N$,

$$\nu\bigg(\bigg(\frac{1}{N}\sum_{k\le M} u'(S_k^1)u'(S_k^2) - r\bigg)^2\bigg) \le \frac{K}{N} + \frac{1}{2}\nu\bigg(\bigg(\frac{1}{N}\sum_{k\le M} u'(S_k^1)u'(S_k^2) - r\bigg)^2\bigg)$$

and this proves (2.22).

Let us now turn to the computation of the "pressure". Rather than reproducing the method of Theorem 3.4.2 of [2] we will follow the ideas of F. Guerra and proceed by interpolation with a simpler system. We denote by q and r the solutions of the equations (2.14) and (2.21) (that are unique provided $KM/N \le 1$). We write $\theta_k = z_k\sqrt{q} + \xi_k\sqrt{1-q}$ where z_k, ξ_k are independent standard normal r.v. (and of course independent of the variables $g_{i,k}$). We consider independent standard normal r.v. $(h_i)_{i\le N}$. For an interpolation parameter $0 \le c \le 1$ we write

$$-H_c = \sum_{k\le M} u(\sqrt{c}\; S_k + \sqrt{1-c}\;\theta_k) + \sqrt{r}\;\sqrt{1-c}\;\sum_{i\le N} h_i\sigma_i, \tag{2.25}$$

and we consider

$$\varphi(c) = \frac{1}{N}\; E\log E_\xi \sum_{\boldsymbol{\sigma}} \exp(-H_c(\boldsymbol{\sigma})),$$

where E_ξ denotes expectation in the r.v. ξ_k only. Thus

$$\begin{aligned}\varphi(0) &= \frac{1}{N}\; E\log\Big(E_\xi\Big(\exp\sum_{k\le M} u(\theta_k)\Big)2^N\prod_{i\le N}\mathrm{ch}\sqrt{r}\; h_i\Big)\\ &= \frac{M}{N}\; E\log E_\xi \exp u(\theta) + E\log \mathrm{ch} z\sqrt{r} + \log 2.\end{aligned}$$

We write

$$\varphi(1) = \varphi(0) + \int_0^1 \varphi'(c)dc, \tag{2.26}$$

and

$$\varphi'(c) = \frac{1}{N}\, E\Big\langle \frac{d}{dc}\, H_c(\boldsymbol{\sigma})\Big\rangle_c, \tag{2.27}$$

where

$$\langle f\rangle_c = \frac{\sum_{\boldsymbol{\sigma}} f(\boldsymbol{\sigma})\exp(-H_c(\boldsymbol{\sigma}))}{E_\xi \sum_{\boldsymbol{\sigma}} \exp(-H_c(\boldsymbol{\sigma}))},$$

and, more generally, for a function $f(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^k)$,

$$\langle f\rangle_c = \frac{\sum_{\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^k} f(\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^k)\exp(-\sum_{\ell\le k} H_c(\boldsymbol{\sigma}^\ell))}{(E_\xi \sum_{\boldsymbol{\sigma}} \exp(-H_c(\boldsymbol{\sigma})))^k}.$$

Now

(2.28)

$$2\frac{d}{dc}H_c(\boldsymbol{\sigma}) = \sum_{k\le M}\left(\frac{S_k}{\sqrt{c}} - \frac{\theta_k}{\sqrt{1-c}}\right)u'(\sqrt{c}\ S_k + \sqrt{1-c}\ \theta_k) - \frac{\sqrt{r}}{\sqrt{1-c}}\sum_{i\le N} h_i\sigma_i.$$

We define $S_{k,c} = \sqrt{c}\ S_k + \sqrt{1-c}\ \theta_k$. We denote by $\xi_k^\ell(\ell = 1,2)$ independent copies of ξ_k, and we define $\theta_k^\ell = z_k\sqrt{q} = \xi_k^\ell\sqrt{1-q}$ and $S_{k,c}^\ell = \sqrt{c}\ S_k^\ell + \sqrt{1-c}\ \theta_k^\ell$, where the superscript ℓ in S_k^ℓ means that we replace $\boldsymbol{\sigma}$ by $\boldsymbol{\sigma}^\ell$. (This is yet another occurrence of the rule that different replicas need independent copies of ξ_k.) Using (2.27), (2.28) and integrating by parts, we get

(2.29)

$$\begin{aligned} 2\varphi'(0) &= -E\left\langle (R_{1,2}-q)\Big(\frac{1}{N}\sum_{k\le M} u'(S_{k,c}^1)u'(S_{k,c}^2)\Big)\right\rangle_c - r(1-E\langle R_{1,2}\rangle_c) \\ &= -E\left\langle (R_{1,2}-q)\Big(\frac{1}{N}\sum_{k\le M} u'(S_{k,c}^1)u'(S_{k,c}^2) - r\Big)\right\rangle_c - r(1-q). \end{aligned}$$

By repeating the proofs of (2.10) and (2.22) for the more general Hamiltonian (2.25) (very little needs to be changed) we get that

$$E\langle (R_{1,2}-q)^2\rangle_c \le \frac{K}{N};\ E\left\langle\left(\frac{1}{N}\sum_{k\le M} u'(S_{k,c}^1)u'(S_{k,c}^2) - r\right)^2\right\rangle_c \le \frac{K}{N}$$

and thus (2.29) yields

$$|2\varphi'(c) - (-r(1-q))| \le \frac{K}{N}$$

and (2.26) yields

(2.30)
$$\frac{1}{N}\ E\log\sum_{\boldsymbol{\sigma}}\exp(-H(\boldsymbol{\sigma}))=\frac{M}{N}\ E\log E_\xi \exp u(\theta)+E\log\operatorname{ch}\ z\sqrt{r}+\log 2 -\frac{1}{2}r(1-q)+\mathcal{R}$$

where $|\mathcal{R}|=K/N$.

3. A first level of control

The basic idea to study the Hamiltonian (1.5) is to interpolate in the last spin as in (2.5). This interpolation uses the Hamiltonian

$$-H_t=\sum_{m\le q} v\left(\frac{1}{\sqrt{M}}\sum_{k\le M} h_{k,m}u(S_{k,t})\right)+\sqrt{1-t}\sigma_N Z, \tag{3.1}$$

where $Z=\eta\sqrt{\rho}$ and ρ is as in (1.13). The formula corresponding to (2.8) one gets in this case is quite complicated, and we will be able to use it only after we gather some preliminary information, which we do in this section. If we think of the quantities $u(S_k)$ as being themselves "spins" (that replace the original spins σ_i), the work of the present section resembles the proof of (2.10). A technical difference is that, while the true spins σ_i satisfy $\sigma_i^2=1$, it need not be true that $u(S_k)^2$ is constant, and a separate effort will be required to control this quantity.

Through the section, we think of $0<t<1$ as being fixed, so the dependence in t will often be kept implicit. Considering an interpolating parameter $0\le c\le 1$ we set

$$T_{m,c}=\frac{1}{\sqrt{M}}\sum_{k\le M-1} h_{k,m}u(S_{k,t})+\sqrt{\frac{c}{M}}\ h_{M,m}u(S_{M,t}), \tag{3.2}$$

and we consider the interpolating Hamiltonian

(3.3)
$$-H_{t,c}=\sum_{m\le Q} v(T_{m,c})+\sqrt{1-c}\ u(S_{M,t})Y+\frac{\tau'}{2}(1-c)u(S_{M,t})^2+\sigma_N\sqrt{1-t}Z,$$

where $Y=\eta\sqrt{\tau}$, $\tau'=\bar{\tau}-\tau$ for the values of τ and $\bar{\tau}$ provided by Theorem 1.1. We write $\langle\cdot\rangle_{t,c}$ an average for the Gibbs measure with Hamiltonian $H_{t,c}$. When $c=1$, we write $\langle\cdot\rangle_t$ such an average. We consider the functions

$$U=\frac{1}{M}\sum_{k\le M} u(S^1_{k,t})u(S^2_{k,t})-r \tag{3.4}$$

$$\overline{U}=\frac{1}{M}\sum_{k\le M} u^2(S_{k,t})-\bar{r}. \tag{3.5}$$

The goal of this section is to prove the following.

Proposition 3.1 *If $KQ/M \le 1$, we have*

$$E\langle U^2\rangle_t \le K\left(\frac{1}{M}+\frac{1}{N}\right) + KE\langle (R_{1,2}-q)^2\rangle_t \tag{3.6}$$

$$E\langle \overline{U}^2\rangle_t \le K\left(\frac{1}{M}+\frac{1}{N}\right) + KE\langle (R_{1,2}-q)^2\rangle_t . \tag{3.7}$$

To prove Theorem 1.2 we assume $KM \le N$, so there is no loss of generality to assume $M \le N$, and the terms K/N are then not needed in (3.6) and (3.7).

For a (possibly random) function f on n replicas, we write $\nu_c(f) = E\langle f\rangle_{t,c}$. The following holds when the quantity f does not depend on the r.v. $h_{k,m}$. For simplicity we use the notation $u_\ell = u\big(S^\ell_{M,t}\big)$, and $T^\ell_{m,c}$ means simply that we have replaced $\boldsymbol{\sigma}$ by $\boldsymbol{\sigma}^\ell$.

Lemma 3.2. *We have*

$$2\frac{d}{dc}\nu_c(f) = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} \tag{3.8}$$

where

$$\begin{aligned} \mathrm{I} = \frac{Q}{M}\Bigg(\sum_{\ell\le n} &\nu_c(u_\ell^2(v''(T^\ell_{Q,c}) + v'^2(T^\ell_{Q,c}))f) \\ &-n\nu_c(u_{n+1}^2(v''(T^\ell_{Q,c}) + v'^2(T^\ell_{Q,c}))f)\Bigg) \end{aligned} \tag{3.9}$$

$$\mathrm{II} = -\overline{\tau}\Bigg(\sum_{\ell\le n}\nu_c(u_\ell^2 f) - n\nu_c(u_{n+1}^2 f)\Bigg) \tag{3.10}$$

$$\mathrm{III} = \frac{Q}{M}\Bigg(\sum_{\ell\ne\ell'} A_{\ell,\ell'} - 2n\sum_{\ell\le n} A_{\ell,n+1} + n(n+1)A_{n+1,n+2}\Bigg) \tag{3.11}$$

for $A_{\ell,\ell'} = \nu_c(u_\ell u_{\ell'} v'(T^\ell_{Q,c})v'(T^{\ell'}_{Q,c})f)$

$$\mathrm{IV} = -\tau\Bigg(\sum_{\ell\ne\ell'} B_{\ell,\ell'} - 2n\sum_{\ell\le n} B_{\ell,n+1} + n(n+1)B_{n+1,n+2}\Bigg) \tag{3.12}$$

for $B_{\ell,\ell'} = \nu_c(u_\ell u_{\ell'} f)$.

The idea is that the choice of $\overline{\tau}$ according to (1.12) ensures that the terms of I nearly cancel out with the corresponding terms of II; while the choice of τ according to (1.11) ensures that the terms of III nearly cancel out with the corresponding terms of IV. This will be shown using Lemma 2.1.

Proof. By straightforward differentiation we have

$$2\frac{d}{dc}\nu_c(f) = \mathrm{V} + \mathrm{VI} + \mathrm{VII}$$

where

$$\mathrm{VII} = -\tau'\Big(\sum_{\ell\le n}\nu_c(u_\ell^2 f) - n\nu_c(u_{n+1}^2 f)\Big)$$

$$\mathrm{V} = \sum_{m\le Q}\Bigg(\sum_{\ell\le n}\nu_c\Big(\frac{1}{\sqrt{cM}}h_{M,m}u_\ell v'(T_{m,c}^\ell)f\Big) - n\nu_c\Big(\frac{1}{\sqrt{cM}}h_{M,m}u_{n+1}v'(T_{m,c}^{n+1})f\Big)\Bigg)$$

$$\mathrm{VI} = -\frac{1}{\sqrt{1-c}}\Big(\sum_{\ell\le n}\nu_c(Yu_\ell f) - n\nu_c(Yu_{n+1}f)\Big).$$

Integration by parts in Y yields that VI=IV+VIII where

$$\mathrm{VIII} = -\tau\Big(\sum_{\ell\le n}\nu_c(u_\ell^2 f) - n\nu_c(u_{n+1}^2 f)\Big).$$

Integration by parts in $h_{M,m}$ and symmetry between the values of $m \le M$ yield that V=I+III, and finally VII+VIII = II since $\tau' + \tau = \bar{\tau}$. □

To start the proof of (3.5) we note that by symmetry between the different values of k we have

$$E\langle U^2\rangle_t = E\langle f\rangle_t \tag{3.13}$$

where

$$f = (u_1u_2 - r)\Big(\frac{1}{M}\sum_{k\le M}u(S_{k,t}^1)u(S_{k,t}^2) - r\Big). \tag{3.14}$$

We write $\varphi(c) = \nu_c(f)$, and we use that

$$|\varphi(1)| \le |\varphi(0)| + \sup_c |\varphi'(c)|. \tag{3.15}$$

The easier part here is the study of $\varphi'(c)$ because this is done exactly as in Section 2. One brings out the dependence of $\langle\cdot\rangle_{t,c}$ in $T_{Q,c}$ by introducing the Hamiltonian

(3.16)

$$-H_\sim = \sum_{m\le Q-1}v(T_{m,c}) + \sqrt{1-c}\,u(S_{M,t})Y + \frac{\tau'}{2}(1-c)u(S_{M,t})^2 + \sigma_N\sqrt{1-t}Z,$$

so that we have formulas such as

$$\langle u_1u_2v'(T_{Q,c}^1)v'(T_{Q,c}^2)f\rangle_{t,c} = \frac{\langle u_1u_2v'(T_{Q,c}^1)v'(T_{Q,c}^2)f\exp(v(T_{Q,c}^1)+v(T_{Q,c}^2))\rangle_\sim}{\langle\exp v(T_{Q,c})\rangle_\sim^2},$$

where of course $\langle\cdot\rangle_\sim$ denotes an average for the Gibbs measure with Hamiltonian (3.16). To estimate such a term, we will use Lemma 2.1 with

$$a_k(\boldsymbol{\sigma}) = \frac{1}{\sqrt{M}}\, u(S_{k,t}) \text{ if } k \le M-1; \quad a_k(\boldsymbol{\sigma}) = 0 \text{ if } k > M;$$

$$a_M(\boldsymbol{\sigma}) = \sqrt{\frac{c}{M}} u(S_{M,t}),$$

and $s = r, \overline{s} = \overline{r}$. We will use this lemma at a given value of all the r.v. other than the $h_{k,Q}$ (and with $g_k = h_{k,Q}$). We observe that

$$\Big(\sum_{k\le M} a_k^2(\boldsymbol{\sigma}) - \overline{r}\Big)^2 \le 2\overline{U}^2 + \frac{K}{M^2};$$

$$\Big(\sum_{k\le M} a_k(\boldsymbol{\sigma}^1)a_k(\boldsymbol{\sigma}^2) - r\Big)^2 \le 2U^2 + \frac{K}{M^2}.$$

We then obtain from (2.5) that

$$|\varphi'(c)| \le \frac{K}{M} + K\frac{Q}{M}(E\langle f^2\rangle_\sim)^{1/2}\big((E\langle U^2\rangle_\sim)^{1/2} + (E\langle \overline{U}^2\rangle_\sim)^{1/2}\big).$$

Since v is bounded, for any function $f \ge 0$ on 2 replicas it is obvious that we have $\langle f\rangle_\sim \le K\nu_c(f)$, we get that

$$|\varphi'(c)| \le \frac{K}{M} + K\frac{Q}{M}\nu_c(f^2)^{1/2}(\nu_c(U^2)^{1/2} + \nu_c(\overline{U}^2)^{1/2}). \tag{3.17}$$

Now, it follows from (3.8) that for a function $f \ge 0$ on 2 replicas, we have

$$|\frac{\mathrm{d}}{\mathrm{d}c}\nu_c(f)| \le K\nu_c(f),$$

so that by integration we have

$$\nu_c(f) \le K\nu_1(f) = E\langle f\rangle_t, \tag{3.18}$$

and combining with (3.17) we get

$$|\varphi'(c)| \le \frac{K}{M} + K\frac{Q}{M}(E\langle f^2\rangle_t)^{1/2}\big((E\langle U^2\rangle_t)^{1/2} + (E\langle \overline{U}^2\rangle)^{1/2}\big). \tag{3.19}$$

We turn to the study of $\varphi(0)$. If

$$\overline{f} = \frac{1}{M}\sum_{k\le M-1} u(S^1_{k,t})u(S^2_{k,t}) - r,$$

we have

$$|f - (u_1u_2 - r)\overline{f}| \le \frac{K}{M},$$

so that

$$|\varphi(0) - A| \le \frac{K}{M} \tag{3.20}$$

where

$$A = \nu_0((u_1u_2 - r)\overline{f}). \tag{3.21}$$

Consider the Hamiltonian

$$-H^* = \sum_{m\le Q} v(T_{m,0}) + \sigma_N\sqrt{1-t}z \tag{3.22}$$

and $\langle\cdot\rangle_*$ an average for the corresponding Gibbs measure, so that

$$A = E\frac{\langle (u_1u_2 - r)\overline{f}\exp((u_1+u_2)Y + \frac{\tau'}{2}(u_1^2+u_2^2))\rangle_*}{\langle \exp(u_1Y + \frac{\tau'}{2}u_1^2)\rangle_*^2}. \tag{3.23}$$

To estimate this quantity, we use the method of Lemma 2.1. We denote by z, ξ_ℓ independent standard normal r.v., and we set

$$u_{\ell,d} = u(\sqrt{d}\ S^\ell_{M,t} + \sqrt{1-d}(z\sqrt{q} + \xi_\ell\sqrt{1-q})),$$

where $0 < d < 1$ is an interpolation parameter. We consider

$$\psi(d) = E\frac{\langle (u_{1,d}u_{2,d} - r)\overline{f}\exp((u_{1,d}+u_{2,d})Y + \frac{\tau'}{2}(u_{1,d}^2+u_{2,d}^2))\rangle_*}{(E_\xi\langle \exp(u_{1,d}Y + \frac{\tau'}{2}u_{1,d}^2)\rangle_*)^2} \tag{3.24}$$

so that $A = \psi(1)$. The choice of r according to (1.9) ensures that $\psi(0) = 0$.

To bound $\psi'(d)$, we cannot proceed as simply as in Lemma 2.1 because the denominator (3.24) is not bounded below, a difficulty that is passed through a differential inequality. We simply notice that if $B \ge 0$ is a function on (say) 10 replicas that is independent of the r.v. $g_{i,M}$ and of Y, the function

$$\psi_1(d) = E\frac{\langle B\exp\sum_{\ell\le 10}(u_{\ell,d}Y + \frac{\tau'}{2}u_{\ell,d}^2)\rangle_*}{(E_\xi\langle \exp(u_{1,d}Y + \frac{\tau'}{2}u_{1,d}^2)\rangle_*)^{10}}$$

satisfies $|\psi_1'(d)| \le K\psi_1(d)$ (as is seen by differentiation and integration by parts, including in Y) so that by integration $\psi_1(d) \le K\psi_1(1)$. We then compute $\psi'(d)$, and we integrate by parts in the r.v. $g_{i,M}$ and Y, using the fact that H^* does not depend on these r.v. and use trivial bounds to reach the estimate

$$\psi'(d) \le K\nu_t\Big(\Big(R_{1,2} - \frac{1-t}{N}\epsilon_1\epsilon_2 - q\Big)^2\Big)^{1/2}\nu_t\big(\overline{f}^2\big)^{1/2},$$

where now (hopefully without creating confusion) we use the notation $\nu_t(f) = E\langle f\rangle_t$. Therefore we have

$$A \le \sup\psi'(d) \le \frac{K}{N} + K\nu_t((R_{1,2}-q)^2)^{1/2}\nu_t\big(\overline{f}^2\big)^{1/2}.$$

Using (3.20), we then see that

$$|\varphi(0)| \le K\Big(\frac{1}{M} + \frac{1}{N}\Big) + K\nu_t((R_{1,2}-q)^2)^{1/2}\nu_t\big(\overline{f}^2\big)^{1/2}. \tag{3.25}$$

Since $\nu_t(\overline{f}^2)^{1/2} \le K/M + K/N + \nu_t(U^2)^{1/2}$, combining with (3.15) and (3.19), we get that

$$\begin{aligned}
\nu_t(U^2) &\le K\left(\frac{1}{M}+\frac{1}{N}\right) + K\nu_t((R_{1,2}-q)^2)^{1/2}\nu_t(U^2)^{1/2} \\
&\quad + K\frac{Q}{M}(\nu_t(U^2)^{1/2} + \nu_t(\overline{U}^2)^{1/2}) \\
&\le K\left(\frac{1}{M}+\frac{1}{N}\right) + \frac{1}{4}\nu_t(U^2)^{1/2} \\
&\quad + K\nu_t((R_{1,2}-q)^2)^{1/2} + K\frac{Q}{M}(\nu_t(U^2)^{1/2} + \nu_t(\overline{U}^2)^{1/2}).
\end{aligned}$$

The same proof shows that the same bound holds for $\nu_t(\overline{U}^2)$; and this finishes the proof of Proposition 3.1. □

4. The cavity method

We will now be able, using Proposition 3.1, to exploit the information we get when differentiating with respect to t. We write

$$T_{m,t} = \frac{1}{\sqrt{M}} \sum_{k \le M} h_{k,m} u(S_{k,t}), \tag{4.1}$$

so that (3.1) becomes

$$-H_t = \sum_{m \le Q} v(T_{m,t}) + \sqrt{1-t}\;\sigma_N Z, \tag{4.2}$$

where we recall that $Z = \eta\sqrt{\rho}$. We note the formulas

$$\frac{dS_{k,t}}{dt} = \frac{\sigma_N}{2\sqrt{tN}}\; g_{N,k}$$

$$\frac{dT_{m,t}}{dt} = \frac{\sigma_N}{2\sqrt{tNM}} \sum_{k \le M} h_{k,m} g_{N,k} u'(S_{k,t})$$

(4.3)

$$-\frac{dH_t}{dt} = \frac{\sigma_N}{2} \sum_{k \le M, m \le Q} \frac{1}{\sqrt{tNM}} h_{k,m} g_{N,k} u'(S_{k,t}) v'(T_{m,t}) - \frac{1}{2\sqrt{1-t}} \sigma_N Z.$$

Also, with some abuse of notation,

$$\frac{\partial T_{m,t}}{\partial g_{N,k}} = \sigma_N \sqrt{\frac{t}{MN}} h_{k,m} u'(S_{k,t}). \tag{4.4}$$

$$\frac{\partial H_t}{\partial g_{N,k}} = \sigma_N \sqrt{\frac{t}{MN}} \sum_{m \le Q} h_{k,m} u'(S_{k,t}) v'(T_{m,t}). \tag{4.5}$$

For simplicity, we write $u_\ell = u(S^\ell_{M,t})$, $u'_\ell = u'(S^\ell_{M,t})$, etc., and $v_\ell = v(T^\ell_{Q,t})$, $v'_\ell = v'(T^\ell_{Q,t})$, etc. We recall the notation $\nu_t(f) = E\langle f \rangle_t$.

Proposition 4.1. *Consider a function f on n replicas, that does not depend on the r.v. $h_{k,m}$ or $g_{i,k}$. Then*

$$\frac{d}{dt}\nu_t(f) = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V} \tag{4.6}$$

where

$$\mathrm{I} = \frac{Q}{N}\sqrt{M}\Big(\sum_{\ell\le n}\nu_t(h_{M,Q}u''_\ell v'_\ell f) - n\nu_t(h_{M,Q}u''_{n+1}v'_{n+1}f)\Big) \tag{4.7}$$

$$\mathrm{II} = \frac{Q}{N}\Big(\sum_{\ell\le n}\nu_t(h^2_{M,Q}u'^2_\ell v''_\ell f) - n\nu_t(h^2_{M,Q}u'^2_{n+1}v''_{n+1}f)\Big) \tag{4.8}$$

$$\mathrm{III} = \frac{1}{N}\Big(\sum_{\ell\le n}C_\ell - nC_{n+1}\Big) \tag{4.9}$$

for

$$C_\ell = \sum_{m,m'\le Q}\nu_t\big(h_{M,m}h_{M,m'}u'^2_\ell v'(T^\ell_{m,t})v'(T^\ell_{m',t})f\big) \tag{4.10}$$

$$\mathrm{IV} = \frac{1}{N}\Big(\sum_{\ell\ne\ell'\le n}A_{\ell,\ell'} - 2n\sum_{\ell\le n}A_{\ell,n+1} + n(n+1)A_{n+1,n+2}\Big) \tag{4.11}$$

for

$$A_{\ell,\ell'} = \sum_{m,m'\le Q}\nu_t\big(h_{M,m}h_{M,m'}u'_\ell u'_{\ell'}v'(T^\ell_{m,t})v'(T^{\ell'}_{m',t})\epsilon_\ell\epsilon_{\ell'}f\big)$$

$$\mathrm{V} = -\rho\Big(\sum_{\ell\ne\ell'\le n}B_{\ell,\ell'} - 2n\sum_{\ell\le n}B_{\ell,n+1} + n(n+1)B_{n+1,n+2}\Big) \tag{4.12}$$

for $B_{\ell,\ell'} = \nu_t(\epsilon_\ell\epsilon_{\ell'}f)$.

Proof. We write

$$\frac{d}{dt}\nu_t(f) = \sum_{\ell\le n}\nu_t\Big(-\frac{\partial H_t}{\partial t}(\boldsymbol{\sigma}^\ell)f\Big) - n\nu_t\Big(-\frac{\partial H_t}{\partial t}(\boldsymbol{\sigma}^{n+1})f\Big),$$

we substitute the value (4.3), we integrate by parts in the r.v. $g_{N,k}$ and Z, using (4.4) and (4.5) and we use symmetry between the values of $k\le Q$. $\square$

We have stated Proposition 4.1 in a manner that makes apparent it resembles Lemma 2.2. We must however face the fact that to make the formulas usable, we will need to integrate by parts in the r.v. $h_{M,m}$. It is this integration by parts that generates complicated formulas.

Using (2.10), we will apply Proposition 4.1 for $f = (\epsilon_1\epsilon_2 - q)(R_{1,2} - q)$, with $n = 2$. Using that $T_{m,0}$ does not depend on σ_N, we see from (1.15) that $|\nu_0(f)| \le 2/N$ (as in (2.15)). Keeping (2.16) in mind, we turn to the control of the quantity (4.7). Without loss of generality we assume $M \le N$. To explain the principle of the approach, we consider the case of the term

$$\begin{aligned}\sqrt{M}\nu_t(h_{M,Q}u''_\ell v'_\ell f) &= \nu_t(u''_\ell u_\ell v''_\ell f) + \sum_{\ell' \le n} \nu_t(u''_\ell v'_\ell u_{\ell'} v'_{\ell'} f) \\ &- n\nu_t(u''_\ell v'_\ell u_{n+1} v'_{n+1} f),\end{aligned} \tag{4.13}$$

by integration by parts, and assuming $\ell \le n$. We explain how to evaluate a typical term, say $\nu_t(u''_1 v'_1 u_2 v'_2 f)$. Consider the Hamiltonian

$$-H_\sim = \sum_{m \le Q-1} v(T_{m,t}) + \sqrt{1-t}\,\sigma_N Z, \tag{4.14}$$

so that

$$\nu_t(u''_1 v'_1 u_2 v'_2 f) = E\frac{\langle u''_1 u_2 v'(T^1_{M,t}) v'(T^2_{M,t}) \exp(v(T^1_{M,t}) + v(T^2_{M,t})) f\rangle_\sim}{\langle \exp v(T_{M,t})\rangle_\sim^2}, \tag{4.15}$$

where of course $\langle\cdot\rangle_\sim$ denotes an average for the Gibbs measure with Hamiltonian (4.14). We use Lemma 2.1 (with $a_k = u(S_{k,t})/\sqrt{M}$, $s = r, \overline{s} = \overline{r}$), Proposition 3.1, and the fact that $|f| \le 2|R_{1,2} - q|$ to get

$$|\nu_t(u''_1 v'_1 u_2 v'_2 f) - CE\langle u''_1 u_2 f\rangle_\sim| \le K\nu_t((R_{1,2}-q)^2) + \frac{K}{M} \tag{4.16}$$

where

$$C = E\left(\frac{E_\xi v'(\gamma)\exp v(\gamma)}{E_\xi \exp v(\gamma)}\right)^2, \tag{4.17}$$

γ being as in Theorem 1.1. We also have by the same method that

$$|\nu_t(u''_1 u_2 f) - E\langle u''_1 u_2 f\rangle_\sim| \le K\nu_t((R_{1,2}-q)^2) + \frac{K}{M}$$

so that, from (4.16)

$$|\nu_t(u''_1 v'_1 u_2 v'_2 f) - C\nu_t(u''_1 u_2 f)| \le K\nu_t((R_{1,2}-q)^2) + \frac{K}{M}. \tag{4.18}$$

We would like to use Lemma 2.1 again to evaluate the contribution of the terms u''_1 and u_2 to the quantity $\nu_t(u''_1 u_2 f)$. A secondary obstacle is that unfortunately each term $T_{m,t}$ in the Hamiltonian H_t contains a term $M^{-1/2}h_{m,M}u(S_{M,t})$, so that the corresponding Gibbs average $\langle\cdot\rangle$ is *not* independent of the r.v. $g_{i,M}$. To clear out this dependence, we recall the interpolating Hamiltonian H_c of (3.2), and we set

$$\varphi(c) = \nu_c(u''_1 u_2 f).$$

Using (3.6), (3.7), (3.19) we see that (assuming $Q \le M \le N$ without loss of generality)

$$|\varphi(1) - \varphi(0)| \le \frac{K}{M} + K\frac{Q}{M}\nu_t((R_{1,2}-q)^2)^{1/2} \le \frac{K}{M} + K\nu_t((R_{1,2}-q)^2).$$

To compute $\varphi(0)$, we use the method of Lemma 2.1, as in the computation of the term (3.23) recalling the Gibbs' average $\langle\cdot\rangle_*$ for the Hamiltonian (3.22), and we find

$$|\varphi(0) - DE\langle f\rangle_*| \le K\nu_t((R_{1,2}-q)^2)^{1/2},$$

where

$$D = E\frac{E_\xi u''(\theta)\exp(u(\theta)\eta\sqrt{\tau}+\frac{\tau'}{2}u(\theta)^2)E_\xi u(\theta)\exp(u(\theta)\eta\sqrt{\tau}+\frac{\tau'}{2}u(\theta)^2)}{(E_\xi\exp(u(\theta)\eta\sqrt{\tau}+\frac{\tau'}{2}u(\theta)^2))^2}.$$

By the same method, on finds

$$|E\langle f\rangle_* - \nu_t(f)| \le \frac{K}{M} + K\nu_t((R_{1,2}-q)^2),$$

so that , we have proved that

$$|\nu_t(u_1''u_2) - D| \le \frac{K}{M} + K\nu_t((R_{1,2}-q)^2).$$

(Thus, in the end, it did not matter that T_m depends on $u(S_{M,t})$.) Combining with (4.18) we get

$$|\nu_t(u_1''v_1'u_2v_2'f) - CD\nu_t(f)| \le \frac{K}{M} + K\nu_t((R_{1,2}-q)^2). \tag{4.19}$$

We would find the same relation (with the same values of C and D) when computing $\nu_t(u_\ell''v_\ell'u_{\ell'}v_{\ell'}'f)$ for $\ell\neq\ell'$, whatever the values of ℓ, $\ell'\le 4$. Proceeding in this manner for all the terms we see that

$$|\sqrt{M}\nu_t(h_{M,Q}u_\ell''v_\ell'f) - A\nu_t(f)| \le \frac{K}{M} + K\nu_t((R_{1,2}-q)^2),$$

where the value of A is independent of ℓ. This establishes the relation

$$|\mathrm{I}| \le \frac{K}{M} + K\frac{Q}{M}\nu_t((R_{1,2}-q)^2). \tag{4.20}$$

and similarly for II and III. (In the case of II there is however no need to integrate by parts in $h_{M,Q}$). The next step is to check that we have the same bound for $|\mathrm{IV}+\mathrm{V}|$, by showing that the terms of IV nearly cancel out with the corresponding terms of V. Without loss of generality, we consider the case of

$$\frac{1}{N}A_{1,2} = \frac{1}{N}\sum_{m,m'\le Q} A(m,m'), \tag{4.21}$$

where

$$A(m,m') = \nu_t(h_{M,m}h_{M,m'}v'(T_{m,t}^1)v'(T_{m',t}^2)\overline{f}), \tag{4.22}$$

for $\overline{f} = \epsilon_1\epsilon_2u_1'u_2'f$. We will proceed to the integration by parts in $h_{M,m}$ and $h_{M,m'}$ using that

$$\frac{\partial T_{m,t}^\ell}{\partial h_{M,m}} = \frac{1}{\sqrt{M}}u_\ell \tag{4.23}$$

and that $T^\ell_{m,t}$ does not depend on $h_{M,m'}$ if $m \neq m'$. We see from (4.23) that integration by parts "brings out a factor $1/\sqrt{M}$", so that it should be clear that

$$|A(m,m) - \nu_t(v'(T^1_{m,t})v'(T^2_{m,t})\overline{f})| \leq \frac{K(n)}{M}, \tag{4.24}$$

and we consider now the case where $m \neq m'$. Integrating by parts first in $h_{M,m'}$ we get

$$\begin{aligned} \sqrt{M}A(m,m') &= \nu_t(h_{M,m}v'(T^1_{m,t})v''(T^2_{m',t})u_2\overline{f}) \\ &+ \sum_{\ell' \leq n} \nu_t(h_{M,m}v'(T^1_{m,t})v'(T^2_{m',t})v'(T^{\ell'}_{m't})u_{\ell'}\overline{f}) \\ &- n\nu_t(h_{M,m}v'(T^1_{m,t})v'(T^2_{m',t})v'(T^{n+1}_{m',t})u_{n+1}\overline{f}). \end{aligned} \tag{4.25}$$

We integrate by parts in $h_{M,m}$ each of these terms, to get

$$MA(m,m') = \mathrm{VI} + \mathrm{VII} + \mathrm{VIII} + \mathrm{IX} + \mathrm{X}, \tag{4.26}$$

where

(4.27) $\mathrm{VI} = \nu_t(v''(T^1_{m,t})v''(T^2_{m',t})u_1u_2\overline{f})$

(4.28) $\mathrm{VII} = \sum_{\ell \leq n} \nu_t(v'(T^1_{m,t})v'(T^\ell_{m,t})v''(T^2_{m',t})u_\ell u_2\overline{f})$

(4.29) $\mathrm{VIII} = -n\nu_t(v'(T^1_{m,t})v'(T^{n+1}_{m,t})v''(T^2_{m',t})u_2u_{n+1}\overline{f})$

(4.30) $\mathrm{IX} = \sum_{\ell' \leq n} \nu_t(v''(T^1_{m,t})v'(T^2_{m',t})v'(T^{\ell'}_{m',t})u_1u_{\ell'}\overline{f})$

(4.31) $\mathrm{X} = -n\nu_t(v''(T^1_{m,t})v'(T^2_{m',t})v'(T^{n+1}_{m',t})u_1u_{n+1}\overline{f})$

(4.32) $\mathrm{XI} = \sum_{\ell,\ell' \leq n} G_{\ell,\ell'} - n\sum_{\ell' \leq n} G_{\ell',n+1} - n\sum_{\ell \leq n+1} G_{\ell,n+1} + n(n+1)G_{n,n+1},$

where

$$G_{\ell,\ell'} = \nu_t(v'(T^1_{m,t})v'(T^\ell_{m,t})v'(T^2_{m',t})v'(T^{\ell'}_{m',t})u_\ell u_{\ell'}\overline{f}).$$

We show how to evaluate a typical term, say

$$\nu_t(v'(T^1_{m,t})v'(T^\ell_{m,t})v''(T^2_{m',t})u_\ell u_2\overline{f}) = \nu_t(v'(T^1_{M,t})v'(T^\ell_{M,t})v''(T^2_{M-1,t})u_\ell u_2\overline{f}).$$

For specificity we assume $\ell \neq 1$. We use the method of (4.18) to see that

$$\begin{aligned} &|\nu_t(v'(T^1_{M,t})v'(T^\ell_{M,t})v''(T^2_{M-1,t})u_\ell u_2\overline{f}) - C\nu_t(v''(T^2_{M-1,t})u_\ell u_2\overline{f})| \\ &\leq \frac{K}{M} + K\nu_t((R_{1,2}-q)^2), \end{aligned}$$

where C is given by (4.17). Now

$$\nu_t(v''(T^2_{M-1,t})u_\ell u_2\overline{f}) = \nu_t(v''(T^2_{M,t})u_\ell u_2\overline{f})$$

and the method of (4.18) shows that

$$|\nu_t(v''(T^2_{M,t})u_2u_\ell\overline{f}) - C'\nu_t(u_2u_\ell\overline{f})| \leq \frac{K}{M} + K\nu((R_{1,2}-q)^2),$$

where, using the notation of (4.17),

$$C' = E\frac{v''(\gamma)\exp v(\gamma)}{E_\xi \exp v(\gamma)}. \tag{4.33}$$

Recalling that $\overline{f} = \epsilon_1\epsilon_2 u_1' u_2' f$, we use the method of proof of (4.19) to see that

$$|\nu_t(u_2 u_\ell \overline{f}) - W(2,\ell)\nu_t(\epsilon_1\epsilon_2 f)| \le \frac{K}{M} + K\nu_t((R_{1,2}-q)^2),$$

where $W(\ell,\ell')$ is given by (1.14). In this manner we see that

$$\nu_t(v'(T^1_{M,t})v'(T^\ell_{m,t})v''(T^2_{m',t})u_\ell u_2\overline{f}) \simeq CC'W(2,\ell)U,$$

where here and below $U = \nu_t(\epsilon_1\epsilon_2 f)$ and where $\simeq$ means that the error is at most $K/M + K\nu_t((R_{1,2}-q)^2)$.

It remains to calculate in this manner the contribution of all the terms in (4.27) to (4.32) and carefully collect them. Consider the quantity

$$C_1 = E\frac{v'(\gamma)^2\exp v(\gamma)}{E_\xi \exp v(\gamma)}. \tag{4.34}$$

By the above method we find (recalling (1.14))

$$\begin{aligned}
&\mathrm{VI} \simeq C'^2 W(1,2)U,\\
&\mathrm{VII} \simeq \Big(C'C_1W(1,2) + C'CW(2,2) + (n-2)C'CW(2,3)\Big)U.
\end{aligned}$$

To obtain this formula we distinguish the cases $\ell = 1$, $\ell = 2$, $\ell \ge 3$, and we use that $W(2,\ell) = W(2,3)$ if $\ell \ge 3$. In a similar manner using the relations $W(\ell,\ell) = W(3,3)$ and $W(1,\ell) = W(1,3)$ for $\ell \ge 3$ we get the relations

$$\begin{aligned}
&\mathrm{VIII} \simeq -nC'CW(2,3)U\\
&\mathrm{IX} \simeq \Big(C'CW(1,1) + C'C_1W(1,2) + (n-2)C'CW(1,3)\Big)U\\
&\mathrm{X} \simeq -nC'CW(1,3)U\\
&n(n+1)G_{n,n+1} \simeq n(n+1)C^2W(3,4)U\\
&-n\sum_{\ell\le n+1} G_{\ell,n+1} \simeq \Big(-nC^2W(3,3) - n(n-2)C^2W(3,4)\\
&\qquad\qquad -nC^2W(2,3) - nCC_1W(1,3)\Big)U\\
&-n\sum_{\ell'\le n} G_{\ell',n+1} \simeq \Big(-n(n-2)C^2W(3,4) - nCC_1W(2,3)\\
&\qquad\qquad -nC^2W(1,3)\Big)U,
\end{aligned}$$

and finally

$$\begin{aligned}\sum_{\ell,\ell'\leq n} G_{\ell,\ell'} &\simeq \Big(CC_1W(1,1)+C^2W(1,2)+(n-2)C^2W(1,3)\\ &+ C_1^2W(1,2)+CC_1W(2,2)+(n-2)CC_1W(2,3)\\ &+ (n-2)(CC_1W(1,3)+CC_1W(2,3)+C^2W(3,3)\\ &+ (n-3)C^2W(3,4))\Big)U,\end{aligned}$$

where the first 2 lines correspond respectively to the contributions of the cases $\ell'=1$, $\ell'=2$, and the last 2 lines to the contributions of $\ell'\geq 3$.

Using (4.26) and collecting the terms, we get

(4.35)

$$\begin{aligned}MA(m,m') \simeq \Big(&W(1,2)((C'+C_1)^2+C^2)+(W(1,1)\\ &+W(2,2))C(C'+C_1)-2(W(1,3)+W(2,3))C(C+C'+C_1)\\ &+6W(3,4)C^2-2W(3,3)C^2\Big)U.\end{aligned}$$

Since by (1.11) and (1.12) we have $\tau=QC/M$ and $\overline{\tau}=Q(C'+C_1)/M$, we see that, since $W(1,1)=W(2,2)$ and $W(1,3)=W(2,3)$,

$$\begin{aligned}\frac{Q^2}{M}A(m,m') &= W(1,2)(\tau^2+\overline{\tau}^2)+2W(1,1)\tau\overline{\tau}-4W(1,3)\tau(\tau+\overline{\tau})\\ &\quad +6W(3,4)\tau^2-2W(3,3)\tau^2+\mathcal{R}\end{aligned} \tag{4.36}$$

where $|\mathcal{R}|\leq KQ^2/M^2+KQ^2\nu((R_{1,2}-q)^2)/M$.

Taking in account that, by (4.24), we have

$$\frac{Q}{M}A(m,m)=\tau E\left(\frac{E_\xi u'(\theta)\exp u(\theta)}{E_\xi \exp u(\theta)}\right)^2+\mathcal{R}, \tag{4.37}$$

where $\mathcal{R}$ is as above, and recalling (4.22) and (1.13), we see that indeed

$$\left|\frac{1}{N}A_{1,2}-\rho\nu_t(\epsilon_1\epsilon_2 f)\right|\leq K\frac{Q^2}{MN}+K\frac{Q^2}{NM}\nu_t((R_{1,2}-q)^2).$$

Thus, for $f=(\epsilon_1\epsilon_2-q)(R_{1,2}-q)$, we have shown that

$$\left|\frac{d}{dt}\nu_t(f)\right|\leq\frac{KQ^2}{MN}+K\frac{Q^2}{NM}\nu_t((R_{1,2}-q)^2).$$

It is straightforward, using (4.6) to obtain through a differential inequality that $\nu_t(f)\leq K\nu(f)$ whenever $f\geq 0$ is a function on 2 replicas, and we finish the proof of (1.19) as in Section 2. The proof of (1.20) to (1.23) is very similar to the proof of (2.22) and is left to the reader.

5. Computing the pressure.

As in the case of the Perceptron, we proceed by interpolation. Not surprisingly, there will be two stages of interpolation.

We consider independent standard Gaussian r.v. z_m, ξ_m and $\gamma_m = z_m\sqrt{r} + \xi_m\sqrt{\overline{r} - r}$. For $0 \le c \le 1$, we set

$$T_{m,c} = \sqrt{c}T_m + \sqrt{1-c}\;\gamma_m. \tag{5.1}$$

(This should not be confused with (3.1); we use another interpolation here.) We consider the Hamiltonian

$$-H_c = \sum_{m\le Q} v(T_{m,c}) + \sum_{k\le M} u(S_k)\eta_k\sqrt{\tau(1-c)} + \frac{\tau'(1-c)}{2}\sum_{k\le M} u^2(S_k), \tag{5.2}$$

where $\tau' = \overline{\tau} - \tau$ and where the r.v. η_k are independent standard normal. Consider

$$\varphi(c) = \frac{1}{N} E \log E_\xi \sum_{\boldsymbol{\sigma}} \exp(-H_c(\boldsymbol{\sigma})),$$

where here and below E_ξ denotes expectation in $\xi_1, \xi_2, \cdots$ only. For a function $f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n)$ we define

$$\nu_c(f) = E\frac{\sum_{\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^n} f(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n)\exp(-\sum_{\ell\le n} H_c(\boldsymbol{\sigma}^\ell))}{(E_\xi \sum_{\boldsymbol{\sigma}} \exp(-H_c(\boldsymbol{\sigma})))^n},$$

where $H_c(\boldsymbol{\sigma}^\ell)$ is obtained form H_c by replacing $\boldsymbol{\sigma}$ by $\boldsymbol{\sigma}^\ell$ and every ξ_m by and independent copy ξ_M^ℓ. Thus

$$\varphi'(c) = \frac{1}{N} E\frac{\sum_{\boldsymbol{\sigma}} -\frac{dH_c}{dc}(\boldsymbol{\sigma})\exp(-H_c(\boldsymbol{\sigma}))}{E_\xi \sum_{\boldsymbol{\sigma}} \exp(-H_c(\boldsymbol{\sigma}))} := \nu_c\Big(-\frac{dH_c}{dc}(\boldsymbol{\sigma})\Big)$$

and

$$\begin{aligned} -2\frac{dH_c}{dc} &= \sum_{m\le Q}\Big(\frac{T_m}{\sqrt{c}} - \frac{\gamma_m}{\sqrt{1-c}}\Big)v'(T_{m,c}) \\ &- \sqrt{\frac{\tau}{1-c}}\sum_{k\le M}\eta_k u(S_k) - \tau'\sum_{k\le M} u^2(S_k). \end{aligned}$$

By integration by parts in the variables $h_{m,k}$,η_k and γ_m we get

$$\begin{aligned} 2\varphi'(c) &= \frac{1}{N}\nu_c\Big(\sum_{m\le Q}(v''(T_{m,c}) + v'^2(T_{m,c}))\Big(\frac{1}{M}\sum_{k\le M} u^2(S_k) - \overline{r}\Big)\Big) \\ &- \frac{1}{N}\nu_c\Big(\sum_{m\le Q} v'(T^1_{m,c})v'(T^2_{m,c})\Big(\frac{1}{M}\sum_{k\le M} u(S^1_k)u(S^2_k) - r\Big)\Big) \\ &- \frac{\tau}{N}\nu_c\Big(\sum_{k\le M} u^2(S_k)\Big) + \frac{\tau}{N}\nu_c\Big(\sum_{k\le M} u(S^1_k)u(S^2_k)\Big) \\ &- \frac{\tau'}{N}\nu_c\Big(\sum_{k\le M} u^2(S_k)\Big). \end{aligned}$$

Using that $\tau' + \tau = \overline{\tau}$, we then get

$$\begin{aligned}
& 2\varphi'(c) \\
= \; & \frac{M}{N}\nu_c\Big(\Big(\frac{1}{M}\sum_{m\le Q}(v''(T_{m,c}) + v'^2(T_{m,c})) - \overline{\tau}\Big)\Big(\frac{1}{M}\sum_{k\le M}u^2(S_k) - \overline{r}\Big)\Big) \\
- \; & \frac{M}{N}\nu_c\Big(\Big(\frac{1}{M}\sum_{m\le Q}v'(T^1_{m,c})v'(T^2_{m,c}) - \tau\Big)\Big(\frac{1}{M}\sum_{k\le M}u(S^1_k)u(S^2_k) - r\Big)\Big) \\
+ \; & \tau r - \overline{\tau}\,\overline{r}.
\end{aligned}$$

To handle this term, one has to generalize Theorem 1.2 to the case of the Hamiltonian (5.2). This is done following the same method, but of course the last two terms of the Hamiltonian (5.2) create more terms when using the cavity method, and one has to check, through integration by parts, that these terms give the correct contribution. Use of the Cauchy-Schwarz inequality and of (1.20) to (1.23) shows that

$$\Big|\varphi'(c) - \frac{1}{2}(\tau r - \overline{\tau}\,\overline{r})\Big| \le \frac{K}{N},$$

so that

$$\Big|\varphi(1) - \varphi(0) - \frac{1}{2}(\tau r - \overline{\tau}\,\overline{r})\Big| \le \frac{K}{N}. \tag{5.3}$$

A new interpolation is required to compute $\varphi(0)$. Consider the r.v. $\theta_k = z_k\sqrt{q} + \xi_k\sqrt{1-q}$, and, for an interpolation parameter $0 < c < 1$, let us write $S_{k,c} = \sqrt{c}\,S_k + \sqrt{1-c}\,\theta_k$. Let

$$-H^*_c = \sum_{k\le M}u(S_{k,c})\eta_k\sqrt{\tau} + \frac{\tau'}{2}\sum_{k\le M}u^2(S_{k,c}) + \sqrt{1-c}\sum_{i\le N}h_i\sigma_i\sqrt{\rho}, \tag{5.4}$$

where the r.v. h_i are independent standard Gaussian. Let

$$\psi(c) = \frac{1}{N}E\log E_\xi\sum_{\boldsymbol{\sigma}}\exp(-H^*_c(\boldsymbol{\sigma})),$$

so that

$$\varphi(0) = \psi(1) + \frac{Q}{N}E\log E_\xi\exp v(\gamma). \tag{5.5}$$

To compute $\psi'(c)$, we proceed as usual,

(5.6) $$2\psi'(c)$$
$$= \frac{1}{N}\nu_c\Big(\sum_{k\le M}\Big(\frac{1}{\sqrt{c}}S_k - \frac{1}{\sqrt{1-c}}\ \theta_k\Big)$$
$$\times\Big(u'(S_{k,c})\eta_k\sqrt{\tau} + \tau'\sum_{k\le M}u'(S_{k,c})u(S_{k,c})\Big)\Big) - \frac{1}{N}\nu_c\Big(\sqrt{\frac{\rho}{1-c}}\sum_{i\le N}h_i\sigma_i\Big)$$
$$= -\frac{1}{N}\nu_c\Big(\tau\sum_{k\le M}\eta_k^2 u'(S^1_{k,c})u'(S^2_{k,c})(R_{1,2}-q)\Big)$$
$$-\frac{1}{N}\nu_c\Big(\tau'\sqrt{\tau}\sum_{k\le M}\eta_k u'(S^1_{k,c})u'(S^2_{k,c})u(S^1_{k,c})u(S^2_{k,c})(R_{1,2}-q)\Big)$$
$$-\frac{1}{N}\nu_c\Big(\tau'^2\sum_{k\le M}u'(S^1_{k,c})u'(S^2_{k,c})u(S^1_{k,c})u(S^2_{k,c})(R_{1,2}-q)\Big)$$
$$-\rho\nu_c(1-R_{1,2}).$$

To handle this quantity we have to extend the results of Section 2 to the case of a Hamiltonian of the type (5.4), showing first that $\nu_c((R_{1,2}-q)^2)\le K/N$, and then that

$$\nu_c\big(B^2_{\ell,\ell'}\big)\le\frac{K}{N}, \tag{5.7}$$

where

$$B_{\ell,\ell'} = \frac{1}{N}\sum_{k\le M}u'(S^1_{k,c})u'(S^2_{k,c})u(S^\ell_{k,c})u(S^{\ell'}_{k,c}) - \frac{M}{N}W(\ell,\ell'), \tag{5.8}$$

a result similar in spirit and in proof to (1.20). Performing the integration by parts in η_k in the first two terms on the right-hand side of (5.6), and using the value of ρ given by (1.13), one then finds after a long computation that

$$\psi'(c) = -\frac{1}{2N}\sum n_{\ell,\ell'}\nu_c\Big(B_{\ell,\ell'}(R_{1,2}-q)\Big) - \frac{(1-q)\rho}{2}, \tag{5.9}$$

where $B_{\ell,\ell'}$ is given by (5.8), where the sum is over $1\le\ell,\ \ell'\le 4$ and where $n_{\ell,\ell'}$ are integers.

It of course makes sense that the value (1.13), found in a somewhat different manner is exactly what is required to obtain (5.9), but the author must admit that he does not fully understand why this is the case.

Using (5.7) one then gets

$$\Big|\psi'(c)+\frac{(1-q)\rho}{2}\Big|\le\frac{K}{N}$$

so that

$$\Big|\psi(1)-\psi(0)+\frac{(1-q)\rho}{2}\Big|\le\frac{K}{N}. \tag{5.10}$$

Now,

$$\psi(0) = \log 2 + \frac{M}{N} E \log E_\xi \exp(u(\theta)\eta\sqrt{\tau} + \frac{\tau'}{2} u^2(\theta)) + E \log \mathrm{ch} z\sqrt{\rho}$$

and combining with (5.9), (5.5), (5.3), this proves (1.24) $\square$

References

[1] J. Hertz; A. Krogh; R. G. Palmer, *Introduction to the theory of neural computation*, Santa Fe Institute studies in the sciences of complexity, 1991.

[2] Talagrand, M., *Spin Glasses, a Challenge to Mathematicians*, Springer Verlag, 2003.

Equipe d'Analyse de l'Institut Mathématique, 4 place Jussieu, 75230 Paris Cedex 05.