Journal of Statistical Physics, Vol. 100, Nos. 5/6, 2000

On the Distribution of Overlaps in the Sherrington–Kirkpatrick Spin Glass Model

M. Ledoux[1]

Received December 20, 1999

This paper describes some of the analytic tools developed recently by Ghirlanda and Guerra in the investigation of the distribution of overlaps in the Sherrington–Kirkpatrick spin glass model and of Parisi's ultrametricity. In particular, we introduce to this task a simplified (but also generalized) model on which the Gaussian analysis is made easier. Moments of the Hamiltonian and derivatives of the free energy are expressed as polynomials of the overlaps. Under the essential tool of self-averaging, we describe with full rigour, various overlap identities and replica independence that actually hold in a rather large generality. The results are presented in a language accessible to probabilists and analysts.

KEY WORDS: Sherrington–Kirkpatrick spin glass model; overlap; free energy; Gaussian analysis; replica equivalence; Parisi's ultrametricity.

1. INTRODUCTION

Let $N \geqslant 2$. The Hamiltonian of the Sherrington–Kirkpatrick (SK) model (without external field) is defined as

$$H = H_N(\varepsilon, x) = \frac{1}{\sqrt{N}} \sum_{1 \leqslant i < j \leqslant N} \varepsilon_i \varepsilon_j x_{ij}$$

where $\varepsilon = (\varepsilon_i) \in \{-1, +1\}^N$ and $x = (x_{ij}) \in \mathbb{R}^{N(N-1)/2}$. We consider here the Hamiltonian as a function of both the spins ε_i and the "random" interactions x_{ij} between two spins ε_i and ε_j with $i < j$. The randomness of the x_{ij}'s will be represented by the canonical Gaussian measure γ on $\mathbb{R}^{N(N-1)/2}$. Integration with respect to γ is denoted by $\int$.

[1] Département de Mathématiques, Laboratoire de Statistique et Probabilités associé au C.N.R.S., Université Paul-Sabatier, 31062 Toulouse, France.

0022-4715/00/0900-0871$18.00/0

Let

$$Z = Z_N(x;\beta) = \int e^{\beta H_N(\varepsilon, x)}\, d\varepsilon$$

be the SK partition function where $d\varepsilon$ denotes uniform measure on the discrete cube $\{-1, +1\}^N$ and where $\beta > 0$ represents the inverse of the temperature. We will usually suppress the dependence in N and β in both H and Z. We emphasize when necessary dependence on x or ε. Integration with respect to the Gibbs measure with density

$$\frac{1}{Z}\, e^{\beta H}$$

with respect to $d\varepsilon$ (that depends on N, β and x) will be denoted by $\langle \cdot \rangle_J$. The notation J is the commonly used one to describe the dependence of the Gibbs measure with respect to the Gaussian interaction. The SK free energy is defined as $F = \log Z$. It is plain that, for fixed N, Z and F are C^∞ functions in x and β. Integration with respect to

$$\frac{1}{Z}\, e^{\beta H}\, d\varepsilon\, d\gamma$$

will be denoted by $\langle \cdot \rangle$ (in other words, $\int \langle \cdot \rangle_J = \langle \cdot \rangle$).

Consider now two independent copies (replicas) ε^1 and ε^2 of ε and define the so-called overlap of ε_1 and ε_2 as

$$q_{1,2} = \frac{1}{N}\, \varepsilon^1 \cdot \varepsilon^2 = \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i^1 \varepsilon_i^2.$$

Note that $|q_{1,2}| \leqslant 1$. When we speak of the distribution of such an overlap, it has to be understood with respect to the (annealed) Gibbs measure

$$\int \frac{1}{Z^2(x)}\, e^{\beta H(\varepsilon^1, x)} e^{\beta H(\varepsilon^2, x)}\, d\varepsilon^1\, d\varepsilon^2\, d\gamma(x)$$

averaging on the Gaussian realizations, represented similarly by $\langle \cdot \rangle$, and similarly if overlaps between a higher number of replicas has to be considered (such as in products $q_{1,2}q_{2,3}$, $q_{1,2}q_{3,4}$, $q_{1,2}q_{2,3}q_{3,1}$). The $\langle \cdot \rangle$ averages are obviously invariant under permutations and relabeling of the replicas.

Overlaps arise naturally as moments of the Hamiltonian or derivatives of the free energy (with respect to β). For example,

$$\frac{\partial F}{\partial \beta} = \frac{1}{Z} \frac{\partial Z}{\partial \beta} = \int H \frac{1}{Z} \mathrm{e}^{\beta H} \, d\varepsilon = \langle H \rangle_J.$$

Now, by definition of H,

$$\int \frac{\partial F}{\partial \beta} = \langle H \rangle = \frac{1}{\sqrt{N}} \sum_{i<j} \iint \varepsilon_i \varepsilon_j x_{ij} \frac{1}{Z} \mathrm{e}^{\beta H} \, d\varepsilon \, d\gamma.$$

Integration by parts along each coordinate x_{ij} shows that for every smooth function φ on $\mathbb{R}^{N(N-1)/2}$,

$$\int x_{ij} \varphi \, d\gamma = \int \partial_{x_{ij}} \varphi \, d\gamma.$$

Therefore,

$$\begin{aligned}\int \frac{\partial F}{\partial \beta} = \langle H \rangle &= \frac{1}{\sqrt{N}} \sum_{i<j} \iint \varepsilon_i \varepsilon_j \partial_{x_{ij}} \left(\frac{1}{Z} \mathrm{e}^{\beta H} \right) d\gamma \, d\varepsilon \\ &= \frac{1}{\sqrt{N}} \sum_{i<j} \iint \varepsilon_i \varepsilon_j \left(\frac{\beta}{\sqrt{N}} \varepsilon_i \varepsilon_j - \frac{\partial_{x_{ij}} Z}{Z} \right) \frac{1}{Z} \mathrm{e}^{\beta H} \, d\gamma \, d\varepsilon.\end{aligned}$$

Now,

$$\partial_{x_{ij}} F = \frac{\partial_{x_{ij}} Z}{Z} = \frac{\beta}{\sqrt{N}} \int \varepsilon'_i \varepsilon'_j \frac{1}{Z} \mathrm{e}^{\beta H(\varepsilon')} \, d\varepsilon' = \frac{\beta}{\sqrt{N}} \langle \varepsilon'_i \varepsilon'_j \rangle_J \tag{1.1}$$

so that, by Fubini's theorem,

$$\int \frac{\partial F}{\partial \beta} = \langle H \rangle = \frac{\beta}{N} \sum_{i<j} \langle (\varepsilon_i \varepsilon_j)^2 - \varepsilon_i \varepsilon_j \varepsilon'_i \varepsilon'_j \rangle = \frac{\beta N}{2} [1 - \langle q^2 \rangle] \tag{1.2}$$

where we set, for simplicity, $q = q_{1,2}$ to describe the basic overlap between two replicas.

The preceding analytic procedure has been performed similarly at the level of second moment and derivative by Guerra in ref. 1. In particular, it is shown there that

$$\langle H^2 \rangle = \frac{N-1}{2} + \frac{\beta^2 N^2}{4} [1 - 2\langle q^2 \rangle - \langle q^4 \rangle + 2\langle q_{1,2}^2 q_{2,3}^2 \rangle]. \tag{1.3}$$

Similarly,

$$\int \frac{\partial^2 F}{\partial \beta^2} = \int \langle H^2 \rangle_J - \int \langle H \rangle_J^2$$

$$= \frac{N}{2} [1 - \langle q^2 \rangle] - \frac{\beta^2 N^2}{2} [\langle q^4 \rangle + 3 \langle q_{1,2}^2 q_{3,4}^2 \rangle - 4 \langle q_{1,2}^2 q_{2,3}^2 \rangle]. \quad (1.4)$$

Note that $\int \langle q^2 \rangle_J^2 = \langle q_{1,2}^2 q_{3,4}^2 \rangle$. Together with self-averaging in quadratic mean, this led Guerra to some remarkable overlap identities (in the thermodynamical limit as $N \to \infty$). Namely, it was shown by Pastur and Sherbina[2] (see also ref. 3) and Guerra[1] that, at least along a subsequence N and for almost every $\beta > 0$,

$$\lim_{N \to \infty} \frac{1}{N^2} [\langle H^2 \rangle - \langle H \rangle^2] = 0. \quad (1.5)$$

The conjunction of (1.2)–(1.5) then leads to the overlap identities (in the thermodynamical limit, along a subsequence and for almost every β),

$$\langle q_{1,2}^2 q_{2,3}^2 \rangle = \tfrac{1}{2} \langle q^2 \rangle^2 + \tfrac{1}{2} \langle q^4 \rangle \quad (1.6)$$

and

$$\langle q_{1,2}^2 q_{3,4}^2 \rangle = \tfrac{2}{3} \langle q^2 \rangle^2 + \tfrac{1}{3} \langle q^4 \rangle. \quad (1.7)$$

Equations (1.6) and (1.7) are due to Guerra.[1] They go in the direction of the Parisi predictions on ultrametricity of overlaps (see below). Note in particular the somewhat surprising feature of (1.7) since the overlaps $q_{1,2}$ and $q_{3,4}$ are independent for each fixed Gaussian realization.

The preceding identities have been extended recently by Ghirlanda and Guerra[4] to show, again under self-averaging, that the overlap $q_{1,\ell+1}$ between one amongst ℓ replicas and the added one $\ell+1$ is, conditionally to the first ℓ replicas, either independent of the former ones, or identical to one of the overlap $q_{1,m}$, $1 < m \leqslant \ell$, each of these cases having equal probability $1/\ell$. In other words, for any bounded function Φ of the first ℓ replicas (and not depending upon the Gaussian interaction),

$$\langle \Phi q_{1,\ell+1}^2 \rangle = \frac{1}{\ell} \langle \Phi \rangle \langle q^2 \rangle + \frac{1}{\ell} \sum_{m=2}^{\ell} \langle \Phi q_{1,m}^2 \rangle \quad (1.8)$$

and similarly for any power of the overlaps. This property in particular allows us to reduce overlaps to combination of overlaps involving a smaller number of replicas.

The purpose of this work is a thorough and hopefully rigorous exposition of the preceding results by means of a simplified model on which the underlying analytical procedure (mainly Gaussian integration by parts) takes an easy form. The model is not only simplified but also generalized, and covers in a convenient language the classical SK model (with or without external field), the p-spin model, $p > 2$, the Derrida random energy model and so on. While we outrageously loose with this model the spin structure, it is however good enough, to some extent, to retain the analytical properties on the overlap distributions as investigated in refs. 1, 4–6. In particular, the preceding overlap identities hold in a surprising generality. The model is presented in Section 2 that contains the basic relations between the Hamiltonian and the overlaps. In particular, we describe there moments of the Hamiltonian in terms of polynomials of overlaps. By related arguments, we describe similarly derivatives in terms of β of the free energy. These relations lead to various overlap identities that we further investigate in the next sections. In Section 3, we present through a simplified approach, the self-averaging properties of ref. 2, which are shown to hold at least for almost every temperature along a subsequence, We then present the result (1.8) of Ghirlanda and Guerra[(4)] in our generalized setting. Moreover, we describe, under some further self-averaging property, how to reduce sets of overlaps to complete overlaps involving a minimal number of replicas (replica equivalence): the joint distribution of ℓ overlaps can be obtained by considering only ℓ replicas. These aspects are related to the recent contribution by Aizenman and Contucci[(5)] where continuity in temperature is used to produce similar conclusions. In the final section, we discuss the complete overlap $q_{1,2}q_{2,3}q_{3,1}$ (the simplest one after $q_{1,2}$) and the Parisi ultrametricity as developed in ref. 7.

We do not consider here the difficult question of the existence of the various limits as $N \to \infty$ of F, $\langle H \rangle$, q^2, etc. As is classical, these limits exist at high temperature ($0 < \beta \leqslant 1$), in which case all the overlaps are essentially 0. The delicate low temperature regime is investigated in refs. 3 and 8.

2. THE GENERALIZED MODEL

Let $n \geqslant 1$, and let $\mu = \mu_n$ be a probability measure on the standard sphere $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$. For every $\xi, x \in \mathbb{R}^n$, let

$$H = H_n(\xi, x) = a(\xi \cdot x)$$

where $a = a_n > 0$ and where $\xi \cdot x$ denotes the scalar product in $\mathbb{R}^n$ of ξ and x. For $\beta > 0$, set

$$Z = Z_n(x; \beta) = \int e^{\beta H}\, d\mu = \int e^{\beta a(\xi \cdot x)}\, d\mu(\xi).$$

Set $F = \log Z$. Integration with respect to the Gibbs measure with density

$$\rho = \frac{1}{Z}\, \mathrm{e}^{\beta H}$$

with respect to μ (that depends on n, β and x) is denoted by $\langle \cdot \rangle_J$. Besides, we equipped $\mathbb{R}^n$ with the canonical Gaussian measure γ, integration with respect to which is denoted by $\int$. By the integration by parts formula with respect to γ,

$$\int x \cdot u = \int \mathrm{div}(u) \tag{2.1}$$

for every smooth function $u: \mathbb{R}^n \to \mathbb{R}^n$. In particular,

$$\int (\xi \cdot x)\, v = \int \xi \cdot \nabla v \tag{2.2}$$

for every $\xi \in \mathbb{R}^n$ and smooth function $v: \mathbb{R}^n \to \mathbb{R}$. Integration with respect to $\int \langle \cdot \rangle_J$ is denoted by $\langle \cdot \rangle$. As in the introduction, we usually suppress dependence on n and β, and emphasize if necessary dependence on ξ or x.

Overlaps in this framework are simply scalar product $\xi^1 \cdot \xi^2$ between independent elements (ξ^1, ξ^2) on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$ (with respect to the product measure $\mu \otimes \mu$). We adopt again the notation $\langle \cdot \rangle_J$ and $\langle \cdot \rangle$ when a higher number of replicas, and thus of Gibbs measures, is involved.

To emphasize the technical simplifications with respect to the usual SK model, it might be worthwhile to reproduce for this generalized model the expression of $\langle H \rangle$. Namely,

$$\begin{aligned}
\langle H \rangle = \int H \frac{1}{Z}\, \mathrm{e}^{\beta H}\, d\mu\, d\gamma &= \int \left(\int a(\xi \cdot x) \frac{1}{Z}\, \mathrm{e}^{\beta H}\, d\gamma \right) d\mu \\
&= a \iint \left[\beta a\, |\xi|^2 - \xi \cdot \frac{\nabla Z}{Z} \right] \frac{1}{Z}\, \mathrm{e}^{\beta H}\, d\gamma\, d\mu \\
&= \beta a^2 [1 - \langle \xi \cdot \xi' \rangle]
\end{aligned}$$

where we used (2.2) and that

$$\nabla Z = \beta a \int \xi' \mathrm{e}^{\beta H(\xi')}\, d\mu(\xi'). \tag{2.3}$$

Writing more simply $p = \xi \cdot \xi'$ for the basic overlap between ξ and ξ', we thus have

$$\langle H \rangle = \beta a^2 [1 - \langle p \rangle]. \tag{2.4}$$

To transpose this simplified model to the usual SK model, let $n = N(N-1)/2$, $a = \sqrt{(N-1)/2}$, and let μ be the image measure of the uniform measure on the cube $\{-1, +1\}^N$ in $\mathbb{R}^N$ by the map

$$\varepsilon \in \{-1, +1\}^N \mapsto \xi = \frac{1}{\sqrt{n}} (\varepsilon_i \varepsilon_j)_{1 \leqslant i < j \leqslant N} \in \mathbb{S}^{n-1}$$

Then,

$$H(\varepsilon) = \frac{1}{\sqrt{N}} \sum_{1 \leqslant i < j \leqslant N} \varepsilon_i \varepsilon_j x_{ij} = a(\xi \cdot x)$$

Note furthermore that

$$N(\varepsilon \cdot \varepsilon')^2 - 1 = (N-1)(\xi \cdot \xi') \tag{2.5}$$

that allows us to transfer scalar products to overlaps of the SK model. In particular, it is then immediate to recover (1.2) from (2.4).

The generalized model includes a number of further examples of interest. The same construction may indeed be applied to the p-spin SK model, $p > 2$. In this case,

$$H = \left(\frac{p!}{2N^{p-1}} \right) \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant N} \varepsilon_{i_1} \cdots \varepsilon_{i_p} x_{i_1 \cdots i_p}$$

and we transfer the uniform measure on $\{-1, +1\}^N$ by the map

$$\varepsilon \in \{-1, +1\}^N \mapsto \xi = \frac{1}{\sqrt{n}} (\varepsilon_{i_1} \cdots \varepsilon_{i_p})_{1 \leqslant i_1 < \cdots < i_p \leqslant N} \in \mathbb{S}^{n-1}$$

where now $n = \binom{N}{p}$ and

$$a^2 = \frac{N(N-1) \cdots (N-p+1)}{2N^{p-1}} \sim \frac{N}{2}.$$

In the random energy model of Derrida,[9] that corresponds to the limit value $p = \infty$ in the p-spin model, we let $x = (x_\varepsilon)_{\varepsilon \in \{-1, +1\}^N} \in \mathbb{R}^n$, $n = 2^N$, be independent standard Gaussian under γ on $\mathbb{R}^n$, and $H = (1/\sqrt{N})\, x_\varepsilon$ This model is handled similarly by mapping $\varepsilon \in \{-1, +1\}^N$ to the εth vector e_ε

of the canonical basis of $\mathbb{R}^n$. Here $a=\sqrt{N}$. The overlap structure is however trivial in this example since $\varepsilon \cdot \varepsilon' = e_\varepsilon \cdot e_{\varepsilon'} = 0$ whenever $\varepsilon \neq \varepsilon'$.

Another case of interest is the presence of an external field. For a probability measure μ now on $\mathbb{S}^{n-1} \times \mathbb{R}$ and $\beta > 0$, $h \in \mathbb{R}$, consider the Gibbs measure with density

$$\frac{1}{Z_n(x; \beta, h)} \, e^{\beta H_n(\xi, x) + h\zeta}$$

with respect to $d\mu(\xi, \zeta)$. When pushing uniform measure on the cube $\{-1, +1\}^N$ by the map

$$\varepsilon \in \{-1, +1\}^N \mapsto (\xi, \zeta) = \left(\frac{1}{\sqrt{n}} (\varepsilon_i \varepsilon_j)_{1 \leqslant i < j \leqslant N}, \sum_{i=1}^N \varepsilon_i \right) \in \mathbb{S}^{n-1} \times \mathbb{R}$$

we recover the SK model with external field described by the Hamiltonian

$$H = \frac{1}{\sqrt{N}} \sum_{1 \leqslant i < j \leqslant N} \varepsilon_i \varepsilon_j x_{ij} + h \sum_{i=1}^N \varepsilon_i$$

(changing h into h/β). In this way, all the results presented below for $h=0$ readily extend to any value of h simply replacing μ on $\mathbb{S}^{n-1}$ by μ on $\mathbb{S}^{n-1} \times \mathbb{R}$. For simplicity in the notation, we however only deal below with the case $h=0$.

Our first task will be to describe moments of the Hamiltonian in terms of overlaps. One basic lemma in the analysis of overlaps is the following result, consequence of the Gaussian integration by parts formula (2.1).

Let I be a finite subset of integers $\geqslant 1$, $\Phi = \Phi(\xi_i, i \in I; x)$ bounded (say) on $\mathbb{S}^{n-1} \times \cdots \times \mathbb{S}^{n-1} \times \mathbb{R}^n$, $j \geqslant 1$, $\lambda \in \mathbb{R}$. Set

$$K(\Phi; \lambda) = \langle \Phi e^{\lambda H(\xi^j)} \rangle$$

Here $\langle \cdot \rangle$ is thus understood with respect to $\rho^J \, d\mu^J \, d\gamma$ where we set $J = I \cup \{j\}$ and $\rho^J = \prod_{m \in J} \rho(\xi^m)$, $d\mu^J = \bigotimes_{m \in J} d\mu(\xi^m)$.

Lemma 2.1. Under the preceding notation,

$$\frac{\partial}{\partial \lambda} K(\Phi; \lambda) = a K(\xi^j \cdot \nabla \Phi; \lambda) + a^2 \lambda K(\Phi; \lambda)$$
$$+ \beta a^2 \left[\sum_{m \in J} K(\Phi(\xi^j \cdot \xi^m); \lambda) - |J| \, K(\Phi(\xi^j \cdot \xi^k); \lambda) \right]$$

where $|J|$ is the cardinal of J and $k \notin J$.

Proof. By (2.2) with $\xi = \xi^j$ and $v = \Phi e^{\lambda H(\xi^j)}$,

$$\frac{\partial}{\partial \lambda} K(\Phi; \lambda) = a\langle (\xi^j \cdot x)\, \Phi e^{\lambda H(\xi^j)} \rangle$$

$$= a \iint (\xi^j \cdot x)\, \Phi e^{\lambda H(\xi^j)} \rho^J \, d\gamma \, d\mu^J$$

$$= a \iint (\xi^j \cdot \nabla \Phi e^{\lambda H(\xi^j)})\, \rho^J \, d\gamma \, d\mu^J + a^2 \lambda \iint \Phi e^{\lambda H(\xi^j)} \rho^J \, d\gamma \, d\mu^J$$

$$+ a \iint (\xi^j \cdot \nabla(\log \rho^J))\, \Phi e^{\lambda H(\xi^j)} \rho^J \, d\gamma \, d\mu^J$$

where we used that $\xi^j \in \mathbb{S}^{n-1}$. Since

$$\log \rho^J = \beta a \sum_{m \in J} \xi^m \cdot x - |J| \, \nabla \log Z$$

and since

$$\nabla \log Z = \frac{\nabla Z}{Z} = \beta a \int \xi^k \rho(\xi^k) \, d\mu(\xi^k)$$

the lemma easily follows. ∎

The preceding lemma may be used to provide a simple induction formula for the moments of $H = a(\xi \cdot x)$ with respect to $\langle \cdot \rangle$. Namely, for any integer $r \geqslant 0$,

$$\langle H^r \rangle = \frac{\partial^r}{\partial \lambda^r} K(1; \lambda)|_{\lambda = 0}$$

For example (with $\Phi \equiv 1$, $j = 1$, $k = 2$),

$$\langle H \rangle = \frac{\partial}{\partial \lambda} K(1; \lambda)|_{\lambda = 0} = \beta a^2 [K(1; 0) - K(\xi^1 \cdot \xi^2; 0)] = \beta a^2 [1 - \langle p \rangle]$$

and (with $\Phi = \xi^1 \cdot \xi^2$, $j = 1$, $k = 3$),

$$\begin{aligned}\langle H^2 \rangle &= a^2 K(1; 0) + \beta a^2 \left[\frac{\partial}{\partial \lambda} K(1; \lambda)|_{\lambda = 0} - \frac{\partial}{\partial \lambda} K(\xi^1 \cdot \xi^2; \lambda)|_{\lambda = 0} \right] \\ &= a^2 K(1; 0) + \beta a^2 [\beta a^2 (K(1; 0) - K(\xi^1 \cdot \xi^2; 0)) \\ &\quad - \beta a^2 (K(\xi^1 \cdot \xi^2; 0) + K((\xi^1 \cdot \xi^2)^2; 0) - 2K((\xi^1 \cdot \xi^2)(\xi^1 \cdot \xi^3); 0))] \\ &= a^2 + \beta^2 a^4 [1 - 2\langle p \rangle - \langle p^2 \rangle + 2\langle (\xi^1 \cdot \xi^2)(\xi^1 \cdot \xi^3) \rangle] \qquad (2.6)\end{aligned}$$

in accordance with (1.3) (via (2.5)). More generally, Lemma 2.1 may be used to develop in terms of overlaps mixed moments such that

$$\int \langle H^r \rangle_J \langle H \rangle_J^s, \qquad r, s \in \mathbb{N}.$$

It follows from Lemma 2.1 that moments of H may be developed as polynomials in a^2 with overlaps as coefficients, with degree $2r$ for the rth moment. In particular, since overlaps are bounded (by 1), for any $r \geqslant 0$ and $a \geqslant 1$,

$$\langle |H|^r \rangle \leqslant C a^{2r} \tag{2.7}$$

where $C > 0$ is polynomial in β. This property is actually well-known on the classical SK model. Indeed

$$\begin{aligned} \langle |H|^r \rangle &= N^{-r/2} \iint \left| \sum_{i<j} \varepsilon_i \varepsilon_j x_{ij} \right|^r \frac{1}{Z} \mathrm{e}^{\beta H(\varepsilon, x)} \, d\varepsilon \, d\gamma(x) \\ &\leqslant N^{-r/2} \int \max_{\varepsilon \in \{-1, +1\}^N} \left| \sum_{i<j} \varepsilon_i \varepsilon_j x_{ij} \right|^r d\gamma(x) \end{aligned}$$

Under γ, $X(\varepsilon) = \sum_{i<j} \varepsilon_i \varepsilon_j x_{ij}$ is a centered Gaussian random variable with variance $N(N-1)/2$. Therefore, by classical Gaussian comparison theorems,[10]

$$\int \max_{\varepsilon \in \{-1, +1\}^N} |X(\varepsilon)|^r \, d\gamma \leqslant C \left(\frac{N(N-1)}{2} \right)^{r/2} (\log \operatorname{Card}(\{-1, +1\}^N))^{r/2}$$

where $C > 0$ only depends on r. Hence (2.7) holds in this case (recall that $a = \sqrt{(N-1)/2}$).

The following corollary is an immediate consequence of Lemma 2.1.

Corollary 2.2. Let $\Phi = \Phi(\xi^1, ..., \xi^\ell)$, say bounded, on $\mathbb{S}^{n-1} \times \cdots \times \mathbb{S}^{n-1}$, $\ell \geqslant 1$. Then

$$\langle \Phi H(\xi^1) \rangle = \beta a^2 \left[\sum_{m=1}^{\ell} \langle \Phi(\xi^1 \cdot \xi^m) \rangle - \ell \langle \Phi(\xi^1 \cdot \xi^{\ell+1}) \rangle \right]$$

Following a somewhat more analytical procedure, one may describe similarly the derivatives in β of the free energy F as cumulants. By definition, the rth cumulant κ_r of

$$K(t) = K_H(t) = \log \langle e^{tH} \rangle_J, \qquad t \geqslant 0,$$

is the coefficient of $t^r/r!$ in the Taylor expansion of K as a function of t. If we center H at its mean $\langle H \rangle_J$,

$$K_{H - \langle H \rangle_J}(t) = K_H(t) - t\langle H \rangle_J$$

for every t so that κ_r may be expressed, for any $r \geqslant 0$, as an algebraic expression of the moments of $H - \langle H \rangle_J$. Now, for each $r \geqslant 0$,

$$\kappa_r = \frac{\partial^r K}{\partial t^r}(0) = \frac{\partial^r F}{\partial \beta^r}.$$

In particular,

$$\frac{\partial F}{\partial \beta} = \langle H \rangle_J,$$

$$\frac{\partial^2 F}{\partial \beta^2} = \langle (H - \langle H \rangle_J)^2 \rangle_J,$$

$$\frac{\partial^3 F}{\partial \beta^3} = \langle (H - \langle H \rangle_J)^3 \rangle_J,$$

$$\frac{\partial^4 F}{\partial \beta^4} = \langle (H - \langle H \rangle_J)^4 \rangle_J - 3\langle (H - \langle H \rangle_J)^2 \rangle_J^2,$$

$$\vdots$$

We may express the derivatives of F as gradients along x. By definition of $Z = \int e^{\beta H}\, d\mu$ and since $H = a(\xi \cdot x)$, it is easily seen by induction that, for every $r \geqslant 1$,

$$\beta^r \frac{\partial^r Z}{\partial \beta^r} = (x \otimes \cdots \otimes x) \cdot \nabla^r Z = x^{\otimes r} \cdot \nabla^r Z.$$

Since $F = \log Z$, we have similarly

$$\beta^r \frac{\partial^r F}{\partial \beta^r} = x^{\otimes r} \cdot \nabla^r F.$$

Integrate now the right-hand side with respect to the canonical Gaussian measure γ on $\mathbb{R}^n$ and make use of (2.1). It follows that

$$\beta^r \int \frac{\partial^r F}{\partial \beta^r} = \int x^{\otimes r} \cdot \nabla^r F$$

$$= (r-1) \int x^{\otimes (r-2)} \cdot \nabla^{r-2}(\Delta F) + \int x^{\otimes (r-1)} \cdot \nabla^{r-1}(\Delta F).$$

In particular, $\beta \int (\partial F / \partial \beta) = \int \Delta F$. In other words

Proposition 2.3. Set $A_r(\phi) = \beta^r \int (\partial^r \phi / \partial \beta^r)$ with $A_0(\phi) = \int \phi$. Then, for every $r \geqslant 1$,

$$A_r(F) = (r-1)\, A_{r-2}(\Delta F) + A_{r-1}(\Delta F).$$

For example,

$$\beta \int \frac{\partial F}{\partial \beta} = \int \Delta F,$$

$$\beta^2 \int \frac{\partial^2 F}{\partial \beta^2} = \int [\Delta^2 F + \Delta F],$$

$$\beta^3 \int \frac{\partial^3 F}{\partial \beta^3} = \int [\Delta^3 F + 3\Delta^2 F],$$

$$\beta^4 \int \frac{\partial^4 F}{\partial \beta^4} = \int [\Delta^4 F + 6\Delta^3 F + 3\Delta^2 F],$$

$$\vdots$$

As a consequence of the preceding, identifying the derivatives of the free energy amounts to identify $\Delta^r F$, $r \geqslant 1$. To this task, note that, as a function of $x \in \mathbb{R}^n$,

$$\partial^2_{ii} Z = \beta^2 a^2 \int \xi_i^2 \mathrm{e}^H \, d\mu(\xi)$$

where ∂_i is partial differentiation along the ith coordinate. Since μ is concentrated on S^{n-1},

$$\Delta Z = \sum_{i=1}^n \partial^2_{ii} Z = \beta^2 a^2 Z. \tag{2.8}$$

It follows, to start with, that

$$\Delta F = \Delta(\log Z) = \frac{\Delta Z}{Z} - \frac{|\nabla Z|^2}{Z^2} = \beta^2 a^2 - \frac{|\nabla Z|^2}{Z^2} = \beta^2 a^2 - |\nabla F|^2. \tag{2.9}$$

It is important to observe that such a differential formula also leads to overlaps since, by Fubini's theorem and (2.3),

$$|\nabla F|^2 = \frac{|\nabla Z|^2}{Z^2} = \beta^2 a^2 \iint \xi \cdot \xi' \, \frac{1}{Z^2} \, \mathrm{e}^{\beta H(\xi)} \mathrm{e}^{\beta H(\xi')} \, d\mu(\xi) \, d\mu(\xi') = \beta^2 a^2 \langle p \rangle_J. \tag{2.10}$$

In particular, we recover in this way (2.4). In order to develop $\Delta^r F$ for $r \geqslant 2$, we borrow from ref. 11 a convenient notation for the iterated gradients. Namely, set, for smooth functions φ, ψ on $\mathbb{R}^n$, $\Gamma_0(\varphi, \psi) = \varphi\psi$, and, for every $k \geqslant 1$,

$$\Gamma_k(\varphi, \psi) = \tfrac{1}{2}[\Delta(\Gamma_{k-1}(\varphi, \psi)) - \Gamma_{k-1}(\varphi, \Delta\psi) - \Gamma_{k-1}(\Delta\varphi, \psi)] = \nabla^k \varphi \cdot \nabla^k \psi.$$

Write furthermore $\Gamma_k(\varphi) = \Gamma_k(\varphi, \varphi)$. Let us test this notation for $\Delta^2 F$. We have, by (2.9),

$$\Delta^2 F = -\Delta(|\nabla F|^2) = -2\Gamma_2(F) - 2\Gamma_2(F, \Delta F) = -2\Gamma_2(F) + 2\Gamma_1(F, \Gamma_1(F)).$$

Thus by the chain rule formula ($F = \log Z$) and (2.8),

$$\Delta^2 F = -\frac{2}{Z^2} \Gamma_2(Z) - \frac{6}{Z^4} \Gamma_1(Z)^2 + \frac{4}{Z^3} \Gamma_1(Z, \Gamma_1(Z)). \tag{2.11}$$

Again, the definition of the Γ_k's allows us to express (2.11) as overlaps. For example,

$$\Gamma_2(Z) = |\nabla^2 Z|^2 = \beta^4 a^4 \iint (\xi \otimes \xi) \cdot (\xi' \otimes \xi') \frac{1}{Z^2} \mathrm{e}^{\beta H(\xi)} \mathrm{e}^{\beta H(\xi')} \, d\mu(\xi) \, d\mu(\xi')$$
$$= \beta^4 a^4 \langle p^2 \rangle_J.$$

We get similarly

$$\Delta^2 F = -2\beta^4 a^4 [2\langle p^2 \rangle_J + 3\langle p \rangle_J^2 - 4\langle (\xi^1 \cdot \xi^2)(\xi^2 \cdot \xi^3) \rangle_J].$$

Since $\beta^2 \int (\partial^2 F / \partial \beta^2) = \int [\Delta^2 F + \Delta F]$, we recover in this way (1.4). Iteration of the preceding shows that each $\Delta^r F$ may be developed in terms of iterated gradients of F. Since $F = \log Z$, these can be turned into iterated gradients of Z, and thus into overlaps. The important aspect drawn from this development is that it is actually "homogeneous" in $\beta^2 a^2$, that is, $\Delta^r F$ is equal to $\beta^{2r} a^{2r}$ times a linear combination of overlaps. To identify precisely this linear combination would require a more careful analysis of the induction procedure. We do not pursue in this direction.

3. SELF-AVERAGING CONDITIONS

In what follows, we let $a = a_n \to \infty$. Following the self-averaging property (1.5), we consider the condition

$$\langle H^2 \rangle - \langle H \rangle^2 = o(a^4) \tag{3.1}$$

We first show, following refs. 1 and 2 (see also ref. 3), how positivity and convexity may be used to check such conditions, at least along a subsequence and for almost every $\beta > 0$.

To this task, write

$$\begin{aligned} &\int_0^{\beta_0} [\langle H^2 \rangle - \langle H \rangle^2] \, d\beta \\ &\quad = \int_0^{\beta_0} \left[\int (\langle H^2 \rangle_J - \langle H \rangle_J^2) \right] d\beta + \int_0^{\beta_0} \left[\int \langle H^2 \rangle_J - \left(\int \langle H \rangle_J \right)^2 \right] d\beta \\ &\quad = \int_0^{\beta_0} \int \frac{\partial^2 F}{\partial \beta^2} \, d\beta + \int_0^{\beta_0} \left[\int \left(\frac{\partial F}{\partial \beta} \right)^2 - \left(\int \frac{\partial F}{\partial \beta} \right)^2 \right] d\beta. \end{aligned} \tag{3.2}$$

Now, by (2.4), for any $\beta_0 > 0$,

$$\int_0^{\beta_0} \int \frac{\partial^2 F}{\partial \beta^2} \, d\beta = \left(\int \frac{\partial F}{\partial \beta} \right)_{|\beta_0} = \langle H \rangle_{|\beta_0} = \beta_0 a^2 [1 - \langle p \rangle] \tag{3.3}$$

so that

$$\frac{1}{a^4} \int_0^{\beta_0} \int \frac{\partial^2 F}{\partial \beta^2} \, d\beta \to 0 \tag{3.4}$$

as $n \to \infty$. Turning to the second term on the right-hand side of (3.2), the classical Poincaré inequality for Gaussian measures (see, e.g., ref. 11) first indicates that

$$\int F^2 - \left(\int F \right)^2 = \tfrac{1}{2} \iint |F(x) - F(y)|^2 \, d\gamma(x) \, d\gamma(y) \leqslant \int |\nabla F|^2. \tag{3.5}$$

As we have seen in (2.10),

$$|\nabla F|^2 = \beta^2 a^2 \langle p \rangle_J \leqslant \beta^2 a^2. \tag{3.6}$$

By Taylor's formula, for any $\beta > 0$, $\delta > 0$, and every $x \in \mathbb{R}^n$,

$$F(x; \beta + \delta) = F(x; \beta) + \delta F'(x; \beta) + \delta^2 \int_0^1 (1-t)\, F''(x; \beta + \delta t)\, dt$$

where we denote for simplicity by F' and F'' the first and second derivatives of $F(x; \beta)$ in β. It follows that

$$|F'(x; \beta) - F'(y; \beta)| \leqslant \frac{1}{\delta}\, |F(x; \beta + \delta) - F(y; \beta + \delta)| + \frac{1}{\delta}\, |F(x; \beta) - F(y; \beta)|$$

$$+ \delta \int (1-t)\, F''(x; \beta + \delta t)\, dt$$

$$+ \delta \int_0^1 (1-t)\, F''(y; \beta + \delta t)\, dt$$

where we used that $F'' \geqslant 0$. By (3.5) and (3.6),

$$\iint |F(x; \beta) - F(y; \beta)|\, d\gamma(x)\, d\gamma(y) \leqslant \sqrt{2}\, \beta a$$

and similarly with $\beta + \delta$ instead of β. Furthermore, for any $\beta_0 > 0$,

$$\int_0^{\beta_0} \int_0^1 (1-t)\, F''(x; \beta + \delta t)\, dt\, d\beta = \int_0^1 (1-t) [F'(x; \beta_0 + \delta t) - F'(x; \delta t)]\, dt$$

so that, by (3.3),

$$\int_0^{\beta_0} \int_0^1 \int (1-t)\, F''(x; \beta + \delta t)\, d\gamma(x)\, dt\, d\beta \leqslant (\beta_0 + \delta)^2\, a^2.$$

Hence, summarizing the preceding estimates,

$$\int_0^{\beta_0} \iint |F'(x; \beta) - F'(y; \beta)|\, d\gamma(x)\, d\gamma(y)\, d\beta$$

$$\leqslant \frac{\sqrt{2}}{\delta}\, \beta_0 a_0 (\beta_0 + \delta) + 2\delta (\beta_0 + \delta)^2\, a^2.$$

Letting n go to infinity, and then δ go to 0, shows that

$$\frac{1}{a^2}\int_0^{\beta_0}\iint |F'(x;\beta)-F'(y;\beta)|\,d\gamma(x)\,d\gamma(y)\,d\beta\to 0.$$

But $F'=\langle H\rangle_J$ and we have seen in (2.7) that, for every $r\geqslant 0$, $\langle |H|^r\rangle = O(a^{2r})$ where O may be bounded uniformly on every interval $[0,\beta_0]$. Therefore, $\int |\langle H\rangle_J|^r = O(a^{2r})$ and it follows that we also have that

$$\frac{1}{a^4}\int_0^{\beta_0}\iint |F'(x;\beta)-F'(y;\beta)|^2\,d\gamma(x)\,d\gamma(y)\,d\beta\to 0$$

and thus

$$\frac{1}{a^4}\int_0^{\beta_0}\left[\int\left(\frac{\partial F}{\partial\beta}\right)^2-\left(\int\frac{\partial F}{\partial\beta}\right)^2\right]d\beta\to 0. \tag{3.7}$$

Therefore, from (3.2), (3.4) and (3.7),

$$\frac{1}{a^4}\int_0^{\beta_0}[\langle H^2\rangle-\langle H^2\rangle]\,d\beta\to 0$$

for every $\beta_0>0$. Hence, at least along a subsequence and for almost every $\beta>0$,

$$\langle H^2\rangle-\langle H\rangle^2=o(a^4).$$

We now investigate, following refs. 1 and 4, the consequences of self-averaging to overlap identities. In what follows, we assume that (3.1) holds, possibly only for almost every β and along a subsequence, as it was shown above. By (2.7), we thus also have that for every integer $r\geqslant 1$,

$$\langle H^r\rangle-\langle H\rangle^r=o(a^{2r}). \tag{3.8}$$

In particular, since by (2.4), $\langle H\rangle=\beta a^2[1-\langle p\rangle]$, it follows that

$$\langle H^r\rangle-\beta^r[1-\langle p\rangle]^r=o(a^{2r}). \tag{3.9}$$

For the derivatives of the free energy, we have similarly that for every $r\geqslant 2$,

$$\int\frac{\partial^r F}{\partial\beta^r}=o(a^{2r}),\qquad \int\Delta^r F=o(a^{2r}). \tag{3.10}$$

Since as we have seen in the preceding section, both $\langle H^r \rangle$ and $\int (\partial^r F/\partial \beta^r)$ may be expressed in terms of overlaps, (3.9) and (3.10) describe various overlap identities extending (1.6) and (1.7).

Now, we turn to the results of ref. 4. As a consequence of the self-averaging condition (3.1) and the Cauchy–Schwarz inequality, for every (say bounded) $\Phi = \Phi(\xi^1,..., \xi^\ell)$ depending on ℓ replicas $\xi^1,..., \xi^\ell$ (although possibly not the first one ξ^1),

$$\langle \Phi H(\xi^1) \rangle = \langle \Phi \rangle \langle H \rangle + o(a^2).$$

Hence, together with Corollary 2.2, we see that

$$\beta a^2 \left[\sum_{m=1}^{\ell} \langle \Phi(\xi^1 \cdot \xi^m) \rangle - \ell \langle \Phi(\xi^1 \cdot \xi^{\ell+1}) \rangle \right] = \langle \Phi \rangle \, \beta a^2 [1 - \langle p \rangle] + o(a^2).$$

We may therefore conclude to the main observation of the work by Ghirlanda and Guerra.[4] In the next statement and below, we use for simplicity the notation (i, j) for $\xi^i \cdot \xi^j$ to represent the overlap between the replicas ξ^i and ξ^j. Somewhat surprisingly, this result holds in the generality of our simplified model, for arbitrary probability measures μ on $\mathbb{S}^{n-1}$.

Proposition 3.1. For every bounded $\Phi = \Phi(\xi^1,..., \xi^\ell)$,

$$\langle \Phi(1, \ell+1) \rangle = \frac{1}{\ell} \langle \Phi \rangle \langle p \rangle + \frac{1}{\ell} \sum_{m=2}^{\ell} \langle \Phi(1, m) \rangle + o(1).$$

Proposition 3.1 is a fundamental tool to reduce sets of overlaps to more simple ones. A first example consists of course of the overlap identities (1.6) and (1.7) that read here

$$\langle (1,2)(2,3) \rangle = \tfrac{1}{2} \langle p \rangle^2 + \tfrac{1}{2} \langle p^2 \rangle \tag{3.11}$$

and

$$\langle (1,2)(3,4) \rangle = \tfrac{2}{3} \langle p \rangle^2 + \tfrac{1}{3} \langle p^2 \rangle. \tag{3.12}$$

For simplicity, we omit here and below, the $o(1)$ term. The identity (3.11) is immediate from Proposition 3.1. For (3.12), apply Proposition 3.1 with $\ell = 3$ to get, by relabeling the replicas,

$$\langle (1,2)(3,4) \rangle = \langle (2,3)(1,4) \rangle = \tfrac{1}{3} \langle p \rangle^2 + \tfrac{2}{3} \langle (1,2)(2,3) \rangle$$

and use (3.11). It might be worthwhile noting that by (2.4) and (2.6), the self-averaging condition (3.1) is actually equivalent to (3.11).

The identities of Proposition 3.1 extend to arbitrary powers of the overlaps, however under stronger self-averaging properties, in the form of, for any $s \geqslant 1$,

$$\langle \Phi(1, \ell+1)^s \rangle = \frac{1}{\ell} \langle \Phi \rangle \langle p^s \rangle + \frac{1}{\ell} \sum_{m=2}^{\ell} \langle \Phi(1, m)^s \rangle + o(1). \tag{3.14}$$

To this task and according to ref. 4, we may add to H auxiliary Hamiltonians $H^{(s)}$ and consider, for every integer $s \geqslant 2$ and $\lambda > 0$,

$$\begin{aligned} H + \lambda H^{(s)} &= H_n(\xi, x) + \lambda H_n^{(s)}(\xi, y) \\ &= a(\xi \cdot x) + \lambda a(\xi^{\otimes s} \cdot y), \qquad \xi \in \mathbb{S}^{n-1}, \quad x \in \mathbb{R}^n, \quad y \in \mathbb{R}^{ns}, \end{aligned}$$

where $\xi^{\otimes s} = \xi \otimes \cdots \otimes \xi$ s-times and the scalar product $\xi^{\otimes s} \cdot y$ takes place in $\mathbb{R}^{ns}$. Performing integration by parts along the y variable only, for example shows as before that

$$\langle H^{(s)} \rangle = \beta a^2 [1 - \langle p^s \rangle]$$

where the averages are now taken with respect to the Hamiltonian $H + \lambda H^{(s)}$ annnealed in the x, y Gaussian variables.

It may be proved similarly that, along a subsequence, almost everywhere in $\lambda > 0$,

$$\langle (H^{(s)})^2 \rangle - \langle H^{(s)} \rangle^2 = o(a^4) \tag{3.15}$$

for every $s \geqslant 2$. The line of reasoning leading to Proposition 3.1 then allows us to raise overlaps to the s power so to get (3.14). However, in order to recover the initial model as $\lambda \to 0$, one has to strengthen the self-averaging properties (3.15) uniformly in $\lambda \to 0$, for example such as

$$\liminf_{\lambda \to 0} \limsup_{n \to \infty} \frac{1}{\lambda^2 a^4} [\langle (H^{(s)})^2 \rangle - \langle H^{(s)} \rangle^2] = 0. \tag{3.16}$$

Such type of conditions are close to continuity in temperature as discussed in ref. 5. It is not clear however when and how these stronger self-averaging conditions can be satisfied.

As we have seen, relations (3.14) are true under the addition of auxiliary Hamiltonians. These small perturbations can deeply change the system (to free SK to p-spin models, $p > 2$), however possibly without changing the free energy: the intuition would be that these perturbations are generic. Relations (3.14) are satisfied under the replica symmetry breaking solution and Parisi's ultrametricity,[7] and are strongly related to their validity. We are grateful to Talagrand for helpful comments on this point.

Taken the relations (3.14) for granted, it gives, by the moment theorem, some ledgitimity to assert that the overlap $(1, \ell+1)$ between one amongst ℓ replicas and the added one $\ell+1$ is, conditionally to the first ℓ replicas, either independent of the former one, or identical to one of the overlap $(1, m)$, $1 < m \leqslant \ell$, each of these cases having equal probability $1/\ell$. In particular, for any integers s, r,

$$\langle (1,2)^s \, (2,3)^r \rangle = \tfrac{1}{2} \langle p^s \rangle \langle p^r \rangle + \tfrac{1}{2} \langle p^{s+r} \rangle.$$

Therefore, the overlaps (1, 2) and (2, 3) are, with equal probability, either independent or equal. In other words, letting Q the distribution (with respect to the annealed measure $\int \rho(\xi^1)\, \rho(\xi^2)\, d\gamma\, \rho(\xi^3)\, d\mu(\xi^1)\, d\mu(\xi^2)\, d\mu(\xi^3)$) of the couple ((1, 2), (2, 3)) on $[-1, +1]^2$, for any bounded measurable φ on $[-1, +1]^2$,

$$\iint \varphi(x, y)\, dQ(x, y) = \tfrac{1}{2} \iint \varphi(x, y)\, dP(x)\, dP(y) + \tfrac{1}{2} \int \varphi(x, x)\, dP(x) \tag{3.17}$$

where we denote by P the distribution of the basic overlap $p = (1, 2)$ on $[-1, +1]$. A similar result holds for the couple ((1, 2), (3, 4)).

4. THE COMPLETE OVERLAP (1, 2)(2, 3)(3, 1) AND PARISI'S ULTRAMETRICITY

Proposition 3.1 allows us to reduce general overlaps to more simple ones. Let us first illustrate this observation on sets of 3 overlaps that, by invariance by permutation and labeling, reduce to the five overlaps (1, 2)(1, 3)(1, 4), (1, 2)(2, 3)(3, 4), (1, 2)(3, 4)(4, 5), (1, 2)(3, 4)(5, 6) and the complete overlap (1, 2)(2, 3)(3, 1). By means of Proposition 3.1, it is easily checked that

$$\langle (1,2)(1,3)(1,4) \rangle = \tfrac{1}{6} \langle p \rangle^3 + \tfrac{1}{2} \langle p \rangle \langle p^2 \rangle + \tfrac{1}{3} \langle p^3 \rangle$$

$$\langle (1,2)(2,3)(3,4) \rangle = \tfrac{1}{6} \langle p \rangle^3 + \tfrac{1}{3} \langle p \rangle \langle p^2 \rangle + \tfrac{1}{6} \langle p^3 \rangle + \tfrac{1}{6} \langle (1,2)(2,3)(3,1) \rangle$$

$$\langle (1,2)(3,4)(4,5) \rangle = \tfrac{1}{4} \langle p \rangle^3 + \tfrac{5}{12} \langle p \rangle \langle p^2 \rangle + \tfrac{1}{6} \langle p^3 \rangle + \tfrac{1}{6} \langle (1,2)(2,3)(3,1) \rangle$$

$$\langle (1,2)(3,4)(5,6) \rangle = \tfrac{1}{3} \langle p \rangle^3 + \tfrac{6}{15} \langle p \rangle \langle p^2 \rangle + \tfrac{2}{15} \langle p^3 \rangle + \tfrac{2}{15} \langle (1,2)(2,3)(3,1) \rangle.$$

Let us check for example the second identity. Setting $\Phi=(1,2)(2,3)$, $\ell=3$ and relabeling, by Proposition 3.1,

$$\begin{aligned}\langle(1,2)(2,3)(3,4)\rangle &= \langle(1,2)(2,3)(1,4)\rangle \\ &= \tfrac{1}{3}\langle(1,2)(2,3)\rangle\langle p\rangle+\tfrac{1}{3}\langle(1,2)^2\,(2,3)\rangle \\ &\quad+\tfrac{1}{3}\langle(1,2)(2,3)(1,3)\rangle\end{aligned}$$

Now, as for (3.11),

$$\langle(1,2)^2\,(2,3)\rangle=\tfrac{1}{2}\langle p\rangle\langle p^2\rangle+\tfrac{1}{2}\langle p^3\rangle$$

from which the result follows.

We observe that the distribution of 3 overlaps reduces to the complete overlaps $p=(1,2)$ and $((1,2),(2,3),(3,1))$. Under (3.14), one may state a general result in this regard.

Proposition 4.1. Under the relations (3.14), the joint distribution of ℓ overlaps is determined by the distributions of j overlaps, $j\leqslant\ell$, involving at most ℓ replicas.

Proof. The statement is somewhat abusive since, according to (3.14), we only express, in the limit, a product $(i_1,i_2)^{\alpha_2}\cdots(i_{2\ell-1},i_{2\ell})^{\alpha_{2\ell}}$ of powers of ℓ disctinct overlaps by similar expressions involving at most ℓ replicas.

The proof goes by induction on ℓ and the number b of replicas. Consider

$$(i_1,i_2)^{\alpha_2}\cdots(i_{2\ell-1},i_{2\ell})^{\alpha_{2\ell}}$$

and assume that it involves $b>\ell$ replicas. It is impossible that all the replicas repeat twice so that there is at least of the i_m's that only occurs once. Assume it is i_1. By (3.14) applied to $\Phi=(i_3,i_4)^{\alpha_3}\cdots(i_{2\ell-1},i_{2\ell})^{\alpha_{2\ell}}$ that involves a set I involving $a<b$ replicas, we get that

$$\begin{aligned}&\langle(i_1,i_2)^{\alpha_2}\cdots(i_{2\ell-1},i_{2\ell})^{\alpha_{2\ell}}\rangle\\ &\quad=\frac{1}{a}\langle\Phi\rangle\langle p\rangle+\frac{1}{a}\sum_{m\in I\setminus\{i_2\}}\langle(i_2,m)(i_3,i_4)^{\alpha_3}\cdots(i_{2\ell-1},i_{2\ell})^{\alpha_{2\ell}}\rangle.\end{aligned}$$

Repeating the procedure concludes the proof. ▮

We conclude this work by comments around Parisi's ultrametricity of the overlap distributions. Predictions based on the replica trick reveal ultrametric structures of the overlap distributions. Generally speaking, ultrametricity implies that the probability distribution of the overlaps over ℓ replicas, which is a priori a function of $\ell(\ell-1)/2$ variables, depends only on $\ell-1$ variables (cf. ref. 7). At the simple level of 3 replicas, ultrametricity indicates that if $(1,2)>(2,3)$ with strictly positive probability, then $(3,1)=(2,3)$. Recalling P the distribution of $p=(1,2)$ and denoting by R the distribution of $((1,2),(2,3),(3,1))$ on $[-1,+1]^3$, this may be expressed by saying that, for any, say bounded, measurable function φ on $[-1,+1]^3$,

$$\int \varphi(x,y,z)\,dR(x,y,z)$$

$$=\iint_{\{x<y\}} \varphi(x,y,x)\,dA(x,y)+\iint_{\{y<z\}} \varphi(y,y,z)\,dA(y,z)$$

$$+\iint_{\{z<x\}} \varphi(x,z,z)\,dA(x,z)+\int \varphi(x,x,x)\,dB(x)$$

where A is a symmetric distribution on $[-1,+1]^2$. (We assume implicitely that overlaps have continuous distributions.) Since $\int dR(x,y,\cdot) = dQ(x,y)$ and $\iint dR(x,\cdot,\cdot)=dP(x)$, integrating successively with respect to one or two variables and making use of Guerra's relations (3.17) allows us to easily identify A and B. Namely, $dA(x,y)=\frac{1}{2}dP(x)\,dP(y)$ and $dB(x) =\frac{1}{2}(\int_{\{y<x\}} dP(y))\,dP(x)$. Therefore,

$$\int \varphi(x,y,z)\,dR(x,y,z)$$

$$=\frac{1}{2}\iint_{\{x<y\}} [\varphi(x,y,x)+\varphi(x,x,y)+\varphi(y,x,x)+\varphi(y,y,y)]\,dP(x)\,dP(y).$$

In particular,

$$\langle(1,2)(2,3)(3,1)\rangle=\frac{1}{2}\iint_{\{x<y\}} (3x^2y+y^3\Big)\,dP(x)\,dP(y).$$

By the reduction of Proposition 4.1, we thus deduce the distribution of any 3 overlap. However, to interpret analytically the ultrametric structure, even in this simple case, seems a challenging question.

ACKNOWLEDGMENTS

I am grateful to M. Talagrand and the referee for several useful comments.

REFERENCES

1. F. Guerra, About the overlap distribution in mean field spin glass model, *Int. J. Mod. Phys. B* **10**:1675–1684 (1996).
2. L. Pastur and M. Sherbina, Absence of self-averaging of the order parameter in the Sherrington–Kirkpatrick model, *J. Stat. Phys.* **62**:1–19 (1991).
3. M. Sherbina, On the replica symmetric solution for the Sherrington-Kirkpatrick model, *Helv. Phys. Acta* **70**:838–853 (1997).
4. S. Ghirlanda and F. Guerra, General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity, preprint (1998).
5. M. Aizenman and P. Contucci, On the stability of the quenched state in mean field spin glass models, *J. Stat. Phys.* **92**:765–783 (1998).
6. G. Parisi, On the probabilitic formulation of the replica approach to spin glasses, preprint (1998).
7. G. Parisi, *Field Theory, Disorder and Simulations* (World Scientific, 1992).
8. M. Talagrand, The Sherrington–Kirkpatrick model: A challenge for mathematicians, *Prob. Th. Rel. Fields* **110**:109–176 (1998).
9. B. Derrida, Random energy model: An exactly solvable model of disordered systems, *Phys. Rev. B* **24**:2613–2626 (1981).
10. X. Fernique, *Régularité des trajectoires des fonctions aléatoires gaussiennes*, École d'Été de Probabilités de St-Flour 1974, Lecture Notes in Math., Vol. 480, pp. 1–96 (Springer, 1975).
11. D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes*, École d'Été de Probabilités de St-Flour 1992, Lecture Notes in Math., Vol. 1581, pp. 1–114 (Springer, 1994).