Journal of Statistical Physics, Vol. 116, Nos. 5/6, September 2004 (© 2004)

A Self-Averaging "Order Parameter" for the Sherrington-Kirkpatrick Spin Glass Model

Walter F. Wreszinski[1] **and Oscar Bolina**[2]

Received November 21, 2003; accepted February 12, 2004

Following an idea of van Enter and Griffiths, we define a self-averaging parameter for the Sherrington-Kirkpatrick (SK) spin glass which is a self-averaging version of the order parameter introduced by Aizenman, Lebowitz and Ruelle. It is strictly positive at low temperature and zero at sufficiently high temperature. The proof is based on the recent construction of the thermodynamic limit of the free energy by Guerra and Toninelli. We also discuss how our definition compares with various existing definitions of order-parameter like quantities.

KEY WORDS: Spin glass model; thermodynamic limit; self-averaging parameter.

1. INTRODUCTION

The Sherrington-Kirkpatrick (SK) model is a model of disordered spin system in which a spin variable taking on the values $\sigma_i = \pm 1$ is assigned to each lattice site $i = 1, 2, \ldots, N$. The Hamiltonian of the model in an external field h is given by

$$H_N(\{J\}) = -\frac{1}{\sqrt{N}} \sum_{1 \leqslant i < j \leqslant N} J_{ij}\sigma_i\sigma_j - h\sum_{i=1}^{N}\sigma_i, \tag{1.1}$$

where the J_{ij}, for $1 \leqslant i < j \leqslant N$, are independent identically distributed random variables with mean zero. The distributions dealt with in this

[1]Departamento de Física-Matemática, Universidade de São Paulo, Caixa Postal 66318, São Paulo 05315-970 Brasil; e-mail: wreszins@if.usp.br
[2]Department of Mathematics, University of California, Davis California 95616-8633 USA; e-mail: oscar@math.ucdavis.edu

paper are the discrete one, with

$$p(dJ_{ij}) = \frac{1}{2}\left(\delta\left(J_{ij} - J\right) + \delta\left(J_{ij} + J\right)\right) dJ_{ij}, \tag{1.2}$$

and the Gaussian one, with

$$p(dJ_{ij}) = \frac{1}{J\sqrt{2\pi}} e^{-\frac{J_{ij}^2}{2J^2}} dJ_{ij}. \tag{1.3}$$

For a particular configuration of the random variables $\{J_{ij}\}_{1 \leqslant i,j \leqslant N}$ the corresponding free-energy per site associated with the Hamiltonian (1.1) is defined by

$$f_N(\{J\}) = -\frac{\beta^{-1}}{N} \ln Z_N(\{J\}). \tag{1.4}$$

where $Z_N(\{J\})$ is the partition function of the model, given by

$$Z_N(\{J\}) = \sum_{\sigma_1} \cdots \sum_{\sigma_N} e^{-\beta H_N(\{J\})} = \mathrm{Tr}\ e^{-\beta H_N(\{J\})}, \tag{1.5}$$

and the trace means the sum over all 2^N possible spin configurations.

As usual, we also define the `quenched` average of a thermodynamic quantity to be the average with respect to the random variables $\{J_{ij}\}$, which will always be denoted by $\langle\langle \cdot \rangle\rangle$. The quenched average of the free-energy associated with the Hamiltonian (1.1) is given by

$$f_N^c = -\frac{\beta^{-1}}{N} \langle\langle \ln Z_N \rangle\rangle. \tag{1.6}$$

It has been proved by Guerra and Toninelli[1] that the thermodynamic limit

$$\lim_{N\to\infty} f_N^c(\{J\}) \tag{1.7}$$

of the quenched free energy exists almost everywhere with respect to the random variables $\{J_{ij}\}$, and that

$$\lim_{N\to\infty} -\frac{\beta^{-1}}{N} \ln Z_N(\{J\}) = \lim_{N\to\infty} -\frac{\beta^{-1}}{N} \langle\langle \ln Z_N(\{J\}) \rangle\rangle. \tag{1.8}$$

This equality is the crucial and physically indispensable self-averaging property of the free energy. Aizenman, Lebowitz and Ruelle[6] have also proved that for the Hamiltonian (1.1) in zero external field and for $\beta < 1$ the equality

$$\lim_{N\to\infty} -\frac{\beta^{-1}}{N} \ln\langle\langle Z_N(\{J\})\rangle\rangle = \lim_{N\to\infty} -\frac{\beta^{-1}}{N} \langle\langle \ln Z_N(\{J\})\rangle\rangle, \tag{1.9}$$

is equivalent to the vanishing of the mean value of the order parameter

$$q_N(\{\beta J\}) = \frac{2}{N(N-1)} \sum_{1\leqslant i<j\leqslant N} \langle\sigma_i\sigma_j\rangle^2 \tag{1.10}$$

as $N\to\infty$, where

$$\langle\sigma_i\sigma_j\rangle = \frac{1}{Z_N(\{J\})} \sum_{\{\sigma\}} \sigma_i\sigma_j \exp\left(\frac{\beta}{\sqrt{N}} \sum_{1\leqslant i<j\leqslant N} J_{ij}\sigma_i\sigma_j\right). \tag{1.11}$$

In this paper we introduce a parameter different from the one which is suitable for high temperature considered by Guerra and Toninelli in[(2)] and Toninelli in.[(3)] It is inspired by the work of van Enter and Griffiths[(12)] and the recent theory of the thermodynamic limit due to Guerra and Toninelli.[(1)] One main point is that – if the *equilibrium* theory of the SK model is taken seriously, the order parameter must be self-averaging.[(4)] It is not clear that the Parisi overlap distributions, which have now been proved to describe the properties of the SK model after the work by Guerra and Guerra and Toninelli, are related to our parameter. Thus, we do not call it an order parameter; hence the quotation marks in our title which should emphasize that different sensible definitions are possible. On the other hand, unlike the (short-range) Edwards-Anderson model,[(5)] some important results are known for the SK model at low temperature: the Aizenman-Lebowitz-Ruelle (ALR) inequality (see[(6)] and Appendix B), the very important recent work by Talagrand[(7)] (see also,[(8)] where a proof of the Parisi solution has been announced), and the theory of the thermodynamic limit at the level of states.[(9)] Thus our parameter is a self-averaging version of the one devised by ALR in.[(6)] The self-averaging property is related to the existence of the thermodynamic limit, which was not controlled in.[(6)] This parameter is strictly positive at low temperatures and zero at high temperature under a standard assumption. This is proved in Section III. In Section IV we present another interpretation of the given parameter and conclude with some conjectures. In Appendix A we quote, for the reader's convenience, the result of[(11)] which is used in the text. In Appendix B we present, also for the reader's convenience, a proof of the ALR inequality.

A recent review where the role of self-averaging arguments is discussed is the one by C. M. Newman and D. L. Stein.[10] Note that our parameter is different from the important Pastur-Scherbina parameter,[20] proved to be non self-averaging.

2. THE TWO-REPLICA HAMILTONIAN

Combining an idea of van Enter and Griffiths[12] with the order parameter (1.10), we are led to define the two-replica Hamiltonian

$$\begin{aligned} H_N(\lambda, \{J\}) = {} & H_N^{(1)}(\{J\}) + H^{(2)}(\{J\}) \\ & + \frac{2\lambda}{N-1} \sum_{1 \leqslant i < j \leqslant N} \sigma_i^{(1)} \sigma_j^{(1)} \sigma_i^{(2)} \sigma_j^{(2)}, \end{aligned} \tag{2.12}$$

where

$$H_N^{(k)}(\{J\}) = -\frac{1}{\sqrt{N}} \sum_{1 \leqslant i < j \leqslant N} J_{ij} \cdot \sigma_i^{(k)} \sigma_j^{(k)} \tag{2.13}$$

for $k = 1, 2$. The corresponding partition function and free energy are given by

$$Z_N(\lambda, \{J\}) = \sum_{\{\sigma_i^{(1)}, \sigma_i^{(2)}\}} e^{-\beta H_N(\lambda, \{J\})} = \mathrm{Tr}\ e^{-\beta H_N(\lambda, \{J\})} \tag{2.14}$$

and

$$f_N(\lambda, \{J\}) = -\frac{\beta^{-1}}{2N} \cdot \ln Z_N(\lambda, \{J\}). \tag{2.15}$$

As a function of λ, the free-energy is a concave function (convex upwards in the terminology of reference[13]). This follows from the fact that the second derivative of (2.15) is given by

$$\frac{d^2 f_N(\lambda, \{J\})}{d\lambda^2} = -\frac{\beta}{2N} \langle (\mathcal{A} - \langle \mathcal{A} \rangle)^2 \rangle \tag{2.16}$$

for

$$\mathcal{A} = \frac{2}{N-1} \sum_{1 \leqslant i < j \leqslant N} \sigma_i^{(1)} \sigma_j^{(1)} \sigma_i^{(2)} \sigma_j^{(2)}, \tag{2.17}$$

and (2.16) is always negative.

To prove that the limit (1.7) exists, Guerra and Toninelli employ a very nice integration by parts technique. By a generalization due to Guerra and Toninelli[11] it can be shown that

$$\lim_{N\to\infty} f_N(\lambda, \{J\}), \forall\lambda \in \mathbb{R} \tag{2.18}$$

exists and is *self-averaging*, that is, has almost everywhere the same value.

Remark 2.1. For $\lambda > 0$, the proof of the generalization is quite simple: it suffices to interpolate linearly for the term (2.16), in addition to the square root interpolation of the random term in (1.1) used in reference.[1] For $\lambda < 0$, however, which is crucially needed in the main text, we must rely on the general result of,[11] which requires a somewhat more elaborate method. See Appendix A.

The proof of (2.18) requires that $\{J_{ij}\}$ satisfy some mild conditions. Under these conditions,

$$f(\lambda, \{J\}) = \lim_{N\to\infty} f_N(\lambda, \{J\}) \tag{2.19}$$

and

$$f(\lambda, \{J\}) = \langle\langle f(\lambda, \{J\})\rangle\rangle = \lim_{N\to\infty} \langle\langle f_N(\lambda, \{J\})\rangle\rangle \tag{2.20}$$

almost everywhere. Since f is concave as a function of λ, as a pointwise limit of concave functions, we have that, for $h > 0$,

$$\frac{f_N(\lambda + h, \{J\}) - f_N(\lambda, \{J\})}{h} \leqslant f_N'(\lambda, \{J\}) \tag{2.21}$$

and

$$\frac{f_N(\lambda, \{J\}) - f_N(\lambda - h, \{J\})}{h} \geqslant f_N'(\lambda, \{J\}). \tag{2.22}$$

On taking the thermodynamic limit and then the limit $h \to 0$ from the right on both sides of (2.21) we obtain, by Griffiths's lemma,[13,14],

$$\begin{aligned} f_+'(\lambda, \{J\}) &= \lim_{h\to 0} \frac{f(\lambda + h, \{J\}) - f(\lambda, \{J\})}{h} \\ &\leqslant \liminf_{N\to\infty} f_N'(\lambda, \{J\}). \end{aligned} \tag{2.23}$$

The same reasoning, applied to the inequality (2.22), now taking $h \to 0$ from the left, yields

$$\begin{aligned} f_-'(\lambda, \{J\}) &= \lim_{h\to 0} \frac{f(\lambda, \{J\}) - f(\lambda - h, \{J\})}{h} \\ &\geqslant \limsup_{N\to\infty} f_N'(\lambda, \{J\}). \end{aligned} \tag{2.24}$$

Note that the derivatives of f with respect to λ from the right and left exist everywhere. It then follows from (2.23) and (2.24) that

$$\begin{aligned} f_+'(\lambda, \{J\}) &\leqslant \liminf_{N\to\infty} f_N'(\lambda, \{J\}) \\ &\leqslant \limsup_{N\to\infty} f_N'(\lambda, \{J\}) \leqslant f_-'(\lambda, \{J\}). \end{aligned} \tag{2.25}$$

Note that

$$\begin{aligned} \frac{df_N(\lambda, \{J\})}{d\lambda} &= \frac{1}{N(N-1)} \frac{1}{Z_N(\lambda, \{J\})} \\ &\times \sum_{1\leqslant i<j\leqslant N} \sum_{\{\sigma^{(1)}, \sigma^{(2)}\}} \sigma_i^{(1)}\sigma_j^{(1)}\sigma_i^{(2)}\sigma_j^{(2)} e^{-\beta H_N(\lambda, \{J\})}, \end{aligned} \tag{2.26}$$

and this derivative, when evaluated at $\lambda = 0$, yields

$$\begin{aligned} \frac{df_N(0, \{J\})}{d\lambda} &= \frac{1}{N(N-1)} \sum_{1\leqslant i<j\leqslant N} \\ &\times \left(\frac{\sum_{\{\sigma^{(1)}\}} \sigma_i^{(1)}\sigma_j^{(1)} \cdot e^{\frac{\beta}{\sqrt{N}} \sum_{1\leqslant i<j\leqslant N} J_{ij}\sigma_i^{(1)}\sigma_j^{(1)}}}{\sum_{\{\sigma^{(1)}\}} e^{\frac{\beta}{\sqrt{N}} \sum_{1\leqslant i<j\leqslant N} J_{ij}\sigma_i^{(1)}\sigma_j^{(1)}}} \right)^2, \end{aligned} \tag{2.27}$$

where we have factored the partition function (2.14) in order to write the denominator of this last expression. Therefore (2.27) is, up to a factor, the ALR order parameter (see[6] and Appendix B).

3. A SELF-AVERAGING VERSION OF THE ALR ORDER PARAMETER

If we take the equilibrium theory of the SK spin glass seriously, the order parameter should satisfy the self-averaging property, which means that the experiment performed on any one sample must be typical (see, however,[15] and[16]).

Since $f_-'(\lambda, \{J\})$ has, by (2.20), almost everywhere the same value, we propose the following definition for our parameter q:

Definition 3.1. We define a self-averaging version of the ALR order parameter for the Sherrington-Kirkpatrick Hamiltonian (2.12) by

$$q = f'_{-}(0, \{J\}). \tag{3.28}$$

It turns out that q has two nice properties expected of an order parameter, although its role as such is debatable (see the introduction). Another clue to its meaning is discussed in section IV.

Proposition 3.2. For $\beta > 1$,

$$q > c\left(1 - \mathcal{O}\left(\frac{1}{\beta}\right)\right), \tag{3.29}$$

where c is a constant.

Proof. Since $f_N(\lambda, \{J\})$ is a concave function, it follows from (2.21) that

$$\frac{f_N(0, \{J\}) - f_N(h, \{J\})}{-h} \geqslant f'_N(0, \{J\}) \tag{3.30}$$

for any $h < 0$. Taking now the $\limsup_{N \to \infty}$ of the left-hand side of (3.30), and taking into account that the limit of the left-hand side of (3.30) as $N \to \infty$ exists, we find that

$$\frac{f(0, \{J\}) - f(h, \{J\})}{-h} \geqslant \limsup_{N \to N} f'_N(0, \{J\}), \forall h < 0 \tag{3.31}$$

Since the right-hand side of (3.31) is independent of $h < 0$, we may take the limit as $h \to 0_-$ of the left-hand side and arrive at

$$f'_{-}(0, \{J\}) \geqslant \limsup_{N \to \infty} f'_N(0, \{J\}). \tag{3.32}$$

It follows from[6] that the following *ALR inequality* holds:

$$\mathcal{D} - \liminf_{N \to \infty} f'_N(0, \{J\}) \geqslant c\left(1 - O\left(\frac{1}{\beta}\right)\right), \tag{3.33}$$

where $\mathcal{D}$ denotes the limit in distribution. Thus, given $\epsilon > 0$,

$$\liminf_{N \to \infty,\ N \geqslant N_0(\epsilon)} f'_N(0, \{J\}) \geqslant c\left(1 - O\left(\frac{1}{\beta}\right)\right) \tag{3.34}$$

with probability larger than $1-\epsilon$. A proof of (3.33) is given in Appendix B. It then follows from (3.32) and (3.34) that

$$f_{-}^{'}(0,\{J\}) \geqslant c\left(1-\mathcal{O}\left(\frac{1}{\beta}\right)\right)$$

holds almost everywhere because $f_{-}^{'}(0,\{J\})$ has the same value almost everywhere, thus proving (3.29). ∎

For β, λ sufficiently small, it follows (see the second reference in[17]) that $f(\lambda,\{J\})$ is differentiable in its coupling parameters $(\beta,\lambda,\dots)$, and therefore there exist $\beta_0>0$, $\lambda_0>0$ such that for $\beta\in[-\beta_0,\beta_0]$ and $\lambda\in[-\lambda_0,\lambda_0]$, $f(\lambda,\{J\})$ is differentiable in λ, which is equivalent to

$$f_{+}^{'}(\lambda,\{J\}) = f_{-}^{'}(\lambda,\{J\})\,. \tag{3.35}$$

By (2.25), this equality implies that at $\lambda=0$ we have, for β sufficiently small,

$$q = \lim_{N\to\infty} f_N^{'}(0,\{J\})$$

and thus, again by,[6]

$$q = 0. \tag{3.36}$$

We have thus

Proposition 3.3. Equation (3.36) holds for $\beta\in[-\beta_0,\beta_0]$.

Note that only differentiability was required above: analyticity is more delicate due to the existence of Griffiths singularities.[17]

4. AN ALTERNATIVE APPROACH TO THE PARAMETER q

We now show that q is *almost*

$$q \equiv \lim_{\lambda\to 0_-}\lim_{N\to\infty} f_N^{'}(\lambda,\{J\})\,. \tag{4.37}$$

It is possible to prove (4.37) only with several provisos. Firstly, the first limit ($N\to\infty$) is only along a special subsequence, and the second one ($\lambda\to 0_-$) only along a *sequence* $\{s_n\}_{n=1}^{\infty}$ of values for λ with $s_n\to 0_-$ as $n\to\infty$. We then have

Proposition 4.1. For any sequence $\{s_n\}_{n=1}^{\infty}$ with $s_n \to 0_-$ as $n \to \infty$, there exists a subsequence $\{N_i\}_{i=1}^{\infty}$ of $N = 1, 2, 3, \ldots$ such that

$$\lim_{n\to\infty} \lim_{i\to\infty} f'_{N_i}(s_n, \{J\}) = q. \tag{4.38}$$

Remark 4.2. Of course, if the limit

$$\lim_{N\to\infty} f'_N(\lambda, \{J\})$$

exists for all λ in a neighborhood of zero, we come back to (4.37).

Proof. Since $f'_N(s_n, \{J\})$ is uniformly bounded in N, there exists, by the diagonal method, a subsequence $\{N_i\}_{i=1}^{\infty}$ of $N = 1, 2, 3, \ldots$ such that

$$\lim_{i\to\infty} f'_{N_i}(s_n, \{J\})$$

exists for each $n = 1, 2, 3, \ldots$. By (2.25), we have

$$\lim_{i\to\infty} f'_{N_i}(s_n, \{J\}) \leqslant f'_-(s_n, \{J\})$$

for $n = 1, 2, 3, \ldots$, and therefore

$$\limsup_{n\to\infty} \lim_{i\to\infty} f'_{N_i}(s_n, \{J\}) \leqslant f'_-(0, \{J\}). \tag{4.39}$$

By definition,

$$\begin{aligned} f_{N_i}(\lambda, \{J\}) &= \inf_{\rho} (-TS(\rho) + \rho(H)) \\ &= -TS(\rho^{(\lambda)}) + \rho^{(\lambda)}(H(\lambda, \{J\})), \end{aligned} \tag{4.40}$$

where

$$\begin{aligned} &\rho^{(\lambda)}(H_N(\lambda, \{J\})) \\ &\quad = \frac{1}{Z_N(\lambda, \{J\})} \operatorname{Tr} \left(H_N(\lambda, \{J\}) \cdot e^{-\beta(H_N(\lambda, \{J\}))} \right). \end{aligned} \tag{4.41}$$

By the variational principle,[12] we have

$$\begin{aligned} f_{N_i}(0, \{J\}) &= \inf_{\rho} \left(-TS(\rho) + \rho(H_{N_i}(0, \{J\})) \right) \\ &\leqslant -TS(\rho^{(\lambda)}) + \rho^{(\lambda)} \left(H_{N_i}(0, \{J\}) \right). \end{aligned} \tag{4.42}$$

Since

$$\frac{\rho^{(\lambda)}\left(\mathcal{H}(0,\{J\})\right)-\rho^{(\lambda)}\left(\mathcal{H}(\lambda,\{J\})\right)}{-\lambda}=f_{N_i}^{'}(\lambda,\{J\}) \tag{4.43}$$

for $\lambda<0$, it follows (4.43), (4.40) and (4.42) that

$$\frac{f_{N_i}(0,\{J\})-f_{N_i}(\lambda,\{J\})}{-\lambda}\leqslant f_{N_i}^{'}(\lambda,\{J\}). \tag{4.44}$$

Considering (4.44) with $\lambda=s_n$ and taking the limit $i\to\infty$

$$\lim_{i\to\infty} f_{N_i}^{'}(s_n,\{J\})\geqslant \frac{f(0,\{J\})-f(s_n,\{J\})}{-s_n}. \tag{4.45}$$

Taking now $\liminf_{n\to\infty}$ of (4.45), we obtain

$$\liminf_{n\to\infty}\lim_{i\to\infty} f_{N_i}^{'}(s_n,\{J\})\geqslant f_{-}^{'}(0,\{J\}). \tag{4.46}$$

By (4.39) and (4.46) the proposition follows. ∎

Remark 4.3. By[9] it seems clear that it is impossible to make sense of the notion of infinite-volume Gibbs states for the SK model. The infinite volume states of random-sites mean-field models, as treated in[18,19] can be well-defined objects.

Remark 4.4. Since the term

$$\frac{2\lambda}{N-1}\sum_{1\leqslant i<j\leqslant N}\sigma_i^{(1)}\sigma_j^{(1)}\sigma_i^{(2)}\sigma_j^{(2)}$$

in (2.12) represents a (special) coupling between replicas, it is not clear that the limits in (4.37) commute. It should be interesting to investigate this question, which might be related to the replica symmetry breaking. If the limits in (4.37) commute, the result obtained would be the average of just the overlap q_{12} (in the notation of Appendix A). Since its self-averaging nature says nothing about the average of other overlaps which occur in the Parisi theory[7,8], it is – in this case – impossible that our q is an order parameter, in the sense that the quantity we have introduced does not provide sufficient information to express the free energy in. However, it distinguishes between the paramagnetic high-temperature phase and the ordered, low-temperature spin-glass phase. In this sense, it only plays part of the role order parameters of mean-field models usually play.

APPENDIX A. THE THERMODYNAMIC LIMIT OF THE FREE ENERGY

In,[1] Guerra and Toninelli proved the existence of the thermodynamic limit of the free energy for the SK model and a number of closely related models, among which is the multi-replica SK model. In this appendix, we state a special case of their results, which is applicable to our model (2.12): The multi replica SK model, with coupled replicas. Let

$$H_N\left(\sigma^{(1)}, ..., \sigma^{(n)}, \{J\}\right) = -\frac{1}{\sqrt{N}} \sum_{1\leqslant i<j\leqslant N} J_{ij}\left(\sigma_i^{(1)}\sigma_j^{(1)} + \cdots + \sigma_i^{(n)}\sigma_j^{(n)}\right) + Ng\left(\{q_{ab}\}\right), \tag{A.1}$$

where

$$q_{ab} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i^{(a)}\sigma_i^{(b)} \tag{A.2}$$

are the overlaps, $a, b = 1, 2, \ldots, n$, and g is a smooth function of class C^1. Then (Theorem 1 of[6]):

Theorem A.1. Under the above conditions the thermodynamic limit of the quenched free energy exists:

$$\lim_{N\to\infty} -\frac{1}{N\beta} E\left(\ln Z_N(\beta)\right) = f(\beta). \tag{A.3}$$

Moreover, the free energy converges almost surely with respect to the disorder realization:

$$\lim_{N\to\infty} -\frac{1}{N\beta}\left(\ln Z_N(\beta)\right) = f(\beta), \quad J-\text{almost surely}, \tag{A.4}$$

and there exists $L < \infty$ such that the disorder fluctuations satisfy the bound

$$P\left(\left|-\frac{1}{N\beta}\ln Z_N(\beta) - f_N(\beta)\right| \geqslant u\right) \leqslant 2\exp\left(-\frac{Nu^2}{2L}\right). \tag{A.5}$$

We see that (2.12) is of the form (A.1), (A.2), with $g(x) = x^2$ up to an unimportant constant term.

APPENDIX B. THE ALR INEQUALITY

The inequality of Aizenman-Lebowitz-Ruelle was given in[6] without detailed proof. In this appendix we provide a short proof since it plays a major role in our considerations.

We first prove the following preliminary result based on the idea of setting the interaction J_{xy} equal to zero. Let

$$H(\{J\}) = H_0 - \frac{J_{xy}}{\sqrt{N}} \sigma_x \sigma_y, \tag{B.1}$$

where $H_0(\{J\})$ is a Hamiltonian obtained from original Hamiltonian of the SK model by removing the interaction $J_{xy}\sigma_x\sigma_y$ corresponding to the particular pair of spins at sites x and y.

Lemma B.1. If we define the correlation

$$\langle \sigma_x \sigma_y \rangle_0 = \frac{\sum_{\{\sigma\}} \sigma_x \sigma_y e^{-\beta H_0}}{\sum_{\{\sigma\}} e^{-\beta H_0}}. \tag{B.2}$$

then we have

$$\langle \sigma_x \sigma_y \rangle = \langle \sigma_x \sigma_y \rangle_0 + \frac{\beta J_{xy}}{\sqrt{N}} \left(1 - \langle \sigma_x \sigma_y \rangle_0^2\right) + \mathcal{O}\left(\frac{\beta^2 J_{xy}^2}{N}\right). \tag{B.3}$$

Proof. Note that the Hamiltonian (B.1) can be written as

$$H_0(\{J\}) = h_1 \sigma_x + h_2 \sigma_y + h,$$

where

$$h_1 = \sum_{\substack{1 \leqslant i \leqslant N \\ i \neq y}} J_{ix} \sigma_i, \qquad h_2 = \sum_{\substack{1 \leqslant i \leqslant N \\ i \neq x}} J_{iy} \sigma_i \qquad \text{and} \qquad h = \sum_{\substack{1 \leqslant i, j \leqslant N \\ i \neq x \text{ and } j \neq y}} J_{ij} \sigma_i \sigma_j.$$

It follows from the definition (1.11) that

$$\langle\sigma_x\sigma_y\rangle = \frac{1}{Z_N(\{J\})}\sum_{\{\sigma\}}{}' \left\{ e^{\frac{\beta}{\sqrt{N}}(h+J_{xy})}\left(e^{\frac{\beta}{\sqrt{N}}(h_1+h_2)} + e^{-\frac{\beta}{\sqrt{N}}(h_1+h_2)}\right) - e^{\frac{\beta}{\sqrt{N}}(h-J_{xy})}\left(e^{\frac{\beta}{\sqrt{N}}(h_1-h_2)} + e^{-\frac{\beta}{\sqrt{N}}(h_1-h_2)}\right)\right\},$$

where the prime on the summation sign indicates that we have carried out the summation on the possible values $\sigma=\pm1$ of the spins at x and y. Factoring out the exponential term containing J_{xy}, we obtain

$$\langle\sigma_x\sigma_y\rangle = \frac{e^{\frac{\beta J_{xy}}{\sqrt{N}}}}{Z_N(\{J\})}\sum_{\{\sigma\}}{}' e^{\frac{\beta h}{\sqrt{N}}}\left(e^{\frac{\beta}{\sqrt{N}}(h_1+h_2)} + e^{-\frac{\beta}{\sqrt{N}}(h_1+h_2)}\right) - \frac{e^{\frac{-\beta J_{xy}}{\sqrt{N}}}}{Z_N(\{J\})}\sum_{\{\sigma\}}{}' e^{\frac{\beta h}{\sqrt{N}}}\left(e^{\frac{\beta}{\sqrt{N}}(h_1-h_2)} + e^{-\frac{\beta}{\sqrt{N}}(h_1-h_2)}\right).$$

For fixed J_{xy} and β, and N sufficiently large, first order expansion of the exponential factors yields (from now on and up to (B.7), a reminder term $\mathcal{O}(\beta^2 J_{xy}^2/N)$ is omitted):

$$\langle\sigma_x\sigma_y\rangle = \frac{1+\frac{\beta J_{xy}}{\sqrt{N}}}{Z_N(\{J\})}\sum_{\{\sigma\}}{}' e^{\frac{\beta h}{\sqrt{N}}}\left(e^{\frac{\beta}{\sqrt{N}}(h_1+h_2)} + e^{-\frac{\beta}{\sqrt{N}}(h_1+h_2)}\right) - \frac{1-\frac{\beta J_{xy}}{\sqrt{N}}}{Z_N(\{J\})}\sum_{\{\sigma\}}{}' e^{\frac{\beta h}{\sqrt{N}}}\left(e^{\frac{\beta}{\sqrt{N}}(h_1-h_2)} + e^{-\frac{\beta}{\sqrt{N}}(h_1-h_2)}\right).$$

This can be now be rewritten as

$$\langle\sigma_x\sigma_y\rangle = \frac{1}{Z_N(\{J\})}\left(A + \frac{\beta J_{xy}}{\sqrt{N}} B\right), \tag{B.4}$$

where

$$A = \sum_{\{\sigma\}}{}' e^{\frac{\beta h}{\sqrt{N}}}\left(e^{\frac{\beta}{\sqrt{N}}(h_1+h_2)} + e^{-\frac{\beta}{\sqrt{N}}(h_1+h_2)} - e^{\frac{\beta}{\sqrt{N}}(h_1-h_2)} - e^{-\frac{\beta}{\sqrt{N}}(h_1-h_2)}\right),$$

and

$$B=\sum_{\{\sigma\}}' e^{\frac{\beta h}{\sqrt{N}}}\left(e^{\frac{\beta}{\sqrt{N}}(h_1+h_2)}+e^{-\frac{\beta}{\sqrt{N}}(h_1-h_2)}+e^{\frac{\beta}{\sqrt{N}}(h_1-h_2)}+e^{-\frac{\beta}{\sqrt{N}}(h_1+h_2)}\right).$$

We now rewrite the partition function $Z_N(\{J\})$ itself in a similar fashion. A repetition of the arguments leading to (B.4) yields

$$Z_N(\{J\})=B+\frac{\beta J_{xy}}{\sqrt{N}}A. \tag{B.5}$$

Substituting (B.5) back into the expression (B.4) we obtain

$$\langle\sigma_x\sigma_y\rangle=\frac{A+\frac{\beta J_{xy}}{\sqrt{N}}B}{B+\frac{\beta J_{xy}}{\sqrt{N}}A} \tag{B.6}$$

We now put the correlation $\langle\sigma_x\sigma_y\rangle_0$ as in (B.6). This is done by simply setting $J_{xy}=0$ in the expression (B.6), thus obtaining

$$\langle\sigma_x\sigma_y\rangle_0=\frac{A}{B}. \tag{B.7}$$

This fact allows us to write (B.6) as

$$\begin{aligned}\langle\sigma_x\sigma_y\rangle &= \left(\frac{A}{B}+\frac{\beta J_{xy}}{\sqrt{N}}\right)\left(1-\frac{\beta J_{xy}}{\sqrt{N}}\cdot\frac{A}{B}\right)\\ &= \frac{A}{B}+\frac{\beta J_{xy}}{\sqrt{N}}\left(1-\left(\frac{A}{B}\right)^2\right)+\mathcal{O}\left(\frac{\beta^2 J_{xy}^2}{N}\right),\end{aligned} \tag{B.8}$$

which, by (B.7), is (B.3). This completes the proof of the lemma. ∎

We are now ready to prove the following result.

Theorem B.2. **[ALR equality]** For each N and β,

$$\frac{d}{d\beta}\frac{1}{N}\langle\langle\ln Z_N\rangle\rangle=\frac{1}{2}\beta J^2\left(1-\langle\langle q_N(\beta J)\rangle\rangle\right)+\mathcal{R}_N, \tag{B.9}$$

where $\mathcal{R}_N$ is a remainder term that vanishes as $N\to\infty$, satisfying

$$\mathcal{R}_N\leqslant\text{constant}\frac{\beta^2}{\sqrt{N}}. \tag{B.10}$$

Proof. We have

$$\frac{d}{d\beta}\frac{1}{N}\ln Z_N = \frac{1}{N}\sum_{1\leqslant i<j\leqslant N}\frac{J_{ij}}{\sqrt{N}}\langle\sigma_i\sigma_j\rangle, \tag{B.11}$$

where $\langle\sigma_i\sigma_j\rangle$ is defined in (1.11). By substituting the relation (B.3) into the expression (B.11), we obtain

$$\begin{aligned}\frac{d}{d\beta}\frac{1}{N}\ln Z_N &= \frac{1}{N^{3/2}}\sum_{1\leqslant i<j\leqslant N} J_{ij}\langle\sigma_i\sigma_j\rangle_0\\ &\quad+\frac{1}{N^2}\sum_{1\leqslant i<j\leqslant N}\beta J_{ij}^2\left(1-\langle\sigma_i\sigma_j\rangle_0^2\right)+\tilde{\mathcal{R}}_N\end{aligned} \tag{B.12}$$

with a remainder $\tilde{\mathcal{R}}_N = \mathcal{O}\left(\beta^2/\sqrt{N}\right)$. We now take the average value of the above expression. The first term on the right-hand side is zero. In order to calculate the average value of the second term on the right-hand side we use (B.3) to write

$$\langle\sigma_x\sigma_y\rangle^2 = \langle\sigma_x\sigma_y\rangle_0^2 + \frac{2\beta J_{ij}}{\sqrt{N}}\langle\sigma_x\sigma_y\rangle_0\cdot\left(1-\langle\sigma_x\sigma_y\rangle_0^2\right)+\mathcal{O}\left(\frac{\beta^2 J_{xy}^2}{N}\right).$$

and, on taking the average value with respect to the J_{ij}'s, we obtain simply

$$\langle\langle\langle\sigma_x\sigma_y\rangle^2\rangle\rangle = \langle\langle\langle\sigma_x\sigma_y\rangle_0^2\rangle\rangle + \mathcal{O}\left(\frac{\beta^2 J_{xy}^2}{N}\right). \tag{B.13}$$

By combining (B.12) and (B.13) the theorem follows. ∎

By using Lemma B.1, (3.34), which we called the ALR inequality in the main text, follows in a straightforward fashion, as in.[6]

ACKNOWLEDGMENT

W.F.W. and O.B. thank Pierluigi Contucci for very useful remarks and conversations. W.F.W. is grateful to A. C. D. van Enter for an illuminating correspondence, and the referees for very important remarks. O.B. thanks Fapesp – FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO – for financial support under grant 01/08485-6. O.B. permanent address is: Departamento de Física-Matemática, Universidade de São Paulo, Caixa Postal 66318, São Paulo 05315-970 Brasil.

REFERENCES

1. F. Guerra and F. L. Toninelli, The thermodynamic limit in mean field spin glass models, *Commun. Math. Phys*. **230**:71–79 (2002).
2. F. Guerra and F. L. Toninelli, Central limit theorem for fluctuations in the high temperature region of the Sherrington-Kirkpatrick spin glass model, *J. Math. Phys*. **43**:6224–6237 (2002).
3. F. L. Toninelli, About the Almeida-Thouless transition line in the Sherrington-Kirkpatrick mean field spin glass model, *Europhys. Lett*. **60**:764–767 (2002).
4. A. C. D. van Enter and J. L. van Hemmen, Statistical mechanical formalism for spin glasses, *Phys. Rev. A* **29**:355–365 (1984).
5. D. Petritis, *Equilibrium statistical mechanics of frustrated spin glasses*, Annales de l'Institute Henri Poincaré, **64**:255–288 (1996).
6. M. Aizenman, J. L. Lebowitz, D. Ruelle, Some rigorous results on the Sherrington-Kirkpatrick spin glass model, *Commun. Math. Phys.* **112**:3–20 (1987).
7. M. Talagrand, *Spin glasses, a challenge to mathematicians*, (Springer, 2003).
8. M. Talagrand, The generalized Parisi formula, *C. R. Math. Acad. Sci. Paris, Ser. I*, **337**:111–114 (2003).
9. C. M. Newman, D. L. Stein, Nonrealistic behavior of mean field spin glasses. *Phys. Rev. Lett*. **91**:197205 (2003).
10. C. M. Newman and D. L. Stein, Ordering and broken symmetry in short-ranged spin glasses, *J. Phys.: Condens. Matter* **15**:R1319–R1364 (2003).
11. F. Guerra and F. L. Toninelli, The infinite volume limit in generalized mean field disordered models, *Markov Proc. Rel. Fields* **9**:195–207 (2003).
12. A. C. D. van Enter and Robert B. Griffiths, The order parameter in a spin glass, *Commun. Math. Phys*. **20**:655–659 (1983).
13. Robert B. Griffiths, A proof that the free energy of a spin system is extensive, *J. Math. Phys*. **5**:1215–1222 (1964).
14. Klaus Hepp and Elliott H. Lieb, Equilibrium statistical mechanics of matter interacting with the quantized radiation field, *Phys. Rev. A* **8**:2517–2525 (1973).
15. S. Franz, M. Mézard, G. Parisi and L. Peliti, The response of glassy systems to random perturbations: A bridge between equilibrium and off-equilibrium, *J. Stat. Phys*. **97**:459–488 (1999).
16. M. Aizenman and P. Contucci, On the stability of the quenched state in mean field spin glass models, *J. Stat. Phys*. **95**:765–783 (1998).
17. J. Frohlich and B. Zegarlinski, The high temperature phase of long range spin glasses, *Commun. Math. Phys*. **110**:122–155; ibid **112**:553–566 (1987).
18. A. Bovier, V. Gayrard and P. Picco, Gibbs states of the Hopfield model in the regime of perfect memory, *Probab. Theor. Relat. Fields*, **100**:329–363 (1994).
19. J. M. G. Amaro de Matos, A. E. Patrick and V. A. Zagrebnov, Random infinite volume Gibbs states for the Curie Weiss random field Ising model, *J. Stat. Phys*. **66**:139–164 (1992).
20. L. A. Pastur and M. V. Shcherbina, Absence of self-averaging of the order-parameter in the Sherrington-Kirkpatrick model, *J. Stat. Phys*. **62**:1–19 (1991).