Quadratic replica coupling in the Sherrington-Kirkpatrick mean field spin glass model

Francesco Guerra *

Dipartimento di Fisica, Università di Roma "La Sapienza"

Istituto Nazionale di Fisica Nucleare, Sezione di Roma1

Piazzale Aldo Moro, 2, I-00185 Roma, Italy

Fabio Lucio Toninelli †

Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

Istituto Nazionale di Fisica Nucleare, Sezione di Pisa

October 7, 2004

Abstract

We develop a very simple method to study the high temperature, or equivalently high external field, behavior of the Sherrington-Kirkpatrick mean field spin glass model. The basic idea is to couple two different replicas with a quadratic term, trying to push out the two replica overlap from its replica symmetric value. In the case of zero external field, our results reproduce the well known validity of the annealed approximation, up to the known critical value for the temperature. In the case of nontrivial external field, we can prove the validity of the Sherrington-Kirkpatrick replica symmetric solution up to a line, which falls short of the Almeida-Thouless line, associated to the onset of the spontaneous replica symmetry breaking, in the Parisi Ansatz. The main difference with the method, recently developed by Michel Talagrand, is that we employ a quadratic coupling, and not a linear one. The resulting flow equations, with respect to the parameters of the model, turn out to be much simpler, and more tractable. As a straightforward application of cavity methods, we show also how to determine free energy and overlap fluctuations, in the region where replica symmetry has been shown to hold. It is a major open problem to give

* e-mail: `francesco.guerra@roma1.infn.it`

† e-mail: `toninell@cibs.sns.it`

a rigorous mathematical treatment of the transition to replica symmetry breaking, necessarily present in the model.

1 Introduction

The mean field spin glass model, introduced by Sherrington and Kirkpatrick in [1], is here considered in the high temperature regime, or, equivalently, for a large external field. It is very well known, on physical grounds, that in this region the replica symmetric solution holds, as shown for example in [2], and references quoted there. However, due to the very large fluctuations present in the model, it is not so simple to give a complete, mathematically rigorous, characterization of this region, especially when there are external fields. Rigorous work on this subject include [3], [4], [5], [6], [7], [8]. For other rigorous results on the structure of the model, we refer to [9], [10], [11], [12], [13], [14], [15].

The method developed in [8], by Michel Talagrand, is particularly interesting. The starting point is given by the very sound physical idea that the spontaneous replica symmetry breaking phenomenon can be understood by exploring the properties of the model, under the application of auxiliary interactions, which explicitly break the replica symmetry. In [8], the replica symmetric solution is shown to hold in a region, which (*probably*) coincides with the region found in the theoretical physics literature, as shown for example in [2], *i.e.* up to the Almeida-Thouless line [16].

The main tool in Talagrand's treatment is an additional minimal replica coupling, linear in the overlap between two replicas. Then, a kind of quadratic stability for the so modified free energy leads immediately, through a generalization of the methods developed in [15], to establish the validity of the replica symmetric solution in a suitable parameter region.

Here we propose a very different strategy, by introducing a quadratic replica coupling, attempting to push the overlap away from its replica symmetric value. In a sense, our method is the natural extension, with applications, of the ideas put forward in [15], where sum rules were introduced for the free energy, by expressing its deviation from the replica symmetric solution in terms of appropriate quadratic fluctuations for the overlap. We choose exactly these quadratic fluctuation terms to act as additional interaction between two replicas, thus explicitly breaking replica symmetry. Then, a generalization of the sum rules, given in [15], for this modified model, allows us immediately to prove that the free energy of the original model converges, in the infinite volume limit, to its replica symmetric value, at least in a parameter region, explicitly determined.

The organization of the paper is as follows. In Section 2, we recall the basic definitions of the mean field spin glass model, and introduce the overlap distribution structure. As a first introduction of our method of quadratic coupling, in Section 3, we treat the well known case of zero external field, by showing that the annealed approximation holds, in the infinite volume limit, up to the true critical inverse temperature $\beta_c = 1$. Our proof shows explicitly that there is a strong connection between the critical value of the transition temperature for the zero external field model, and the analogous, and numerically equivalent, temperature for the well known ferromagnetic Curie-Weiss mean field model.

In Section 4, we consider the model with external field, and introduce the associated model with quadratic replica coupling. Then, simple stability estimates give immediately the convergence of the free energy to its replica symmetric value, in a suitable, well defined, region of the parameters.

Section 5 reports about results on the free energy and overlap fluctuations, in our determined replica symmetric region. We also sketch the method of proof, based on cavity considerations, as developed for example in [2] and [5]. A more complete treatment will be found in a forthcoming paper [17].

Finally, Section 6 is dedicated to a short outlook about open problems and further developments.

For the relevance of the mean field spin glass model for the understanding of the physical properties of realistic spin glasses, we refer to [18], but see also [19].

2 The general structure of the mean field spin glass model

The generic configuration of the mean field spin glass model is defined by Ising spin variables $\sigma_i = \pm 1$, attached to each site $i = 1, 2, \dots, N$. The external quenched disorder is given by the $N(N-1)/2$ independent and identical distributed random variables J_{ij}, defined for each couple of sites. For the sake of simplicity, we assume each J_{ij} to be a centered unit Gaussian with averages

$$E(J_{ij}) = 0, \quad E(J_{ij}^2) = 1.$$

The Hamiltonian of the model, in some external field of strength h, is given by

$$H_N(\sigma, J) = -\frac{1}{\sqrt{N}} \sum_{(i,j)} J_{ij}\sigma_i\sigma_j - h\sum_i \sigma_i. \tag{1}$$

The first sum extends to all site couples, an the second to all sites. The normalizing factor $1/\sqrt{N}$ is typical of the mean field character of the model, and guarantees a good thermodynamic limit for the free energy per spin, i.e., the existence of a finite and non trivial limit for the free energy as $N \to \infty$. The first term in (1) is a long range random two body interaction, while the second represents the interaction of the spins with a fixed external magnetic field h.

For a given inverse temperature β, we introduce the disorder dependent partition function $Z_N(\beta, J)$, the (quenched average of the) free energy per site $f_N(\beta)$, the internal energy per site $u_N(\beta)$, the Boltzmann state ω_J, and the auxiliary function $\alpha_N(\beta)$, according to the definitions

$$Z_N(\beta, J) = \sum_{\sigma_1 \dots \sigma_N} \exp(-\beta H_N(\sigma, J)), \tag{2}$$

$$-\beta f_N(\beta) = N^{-1} E \log Z_N(\beta, J) = \alpha_N(\beta), \tag{3}$$

$$\omega_J(A) = Z_N(\beta, J)^{-1} \sum_{\sigma_1 \dots \sigma_N} A \exp(-\beta H_N(\sigma, J)), \tag{4}$$

$$u_N(\beta) = N^{-1} E \omega_J(H_N(\sigma, J)) = \partial_\beta(\beta f_N(\beta)) = -\partial_\beta \alpha_N(\beta), \tag{5}$$

where A is a generic function of the σ's. In the notation ω_J, we have stressed the dependence of the Boltzmann state on the external noise J, but, of course, there is also a dependence on β, h and N.

We are interested in the thermodynamic limit $N \to \infty$.

Let us now introduce the important concept of replicas. Consider a generic number s of independent copies of the system, characterized by the Boltzmann variables $\sigma_i^{(1)}$, $\sigma_i^{(2)}$, …, distributed according to the product state

$$\Omega_J = \omega_J^{(1)} \omega_J^{(2)} \dots \omega_J^{(s)},$$

where all $\omega_J^{(\alpha)}$ act on each one $\sigma_i^{(\alpha)}$'s, and are subject to the *same* sample J of the external noise. Clearly, the *Boltzmannfaktor* for the replicated system is given by

$$\exp\left(-\beta(H_N(\sigma^{(1)}, J) + H_N(\sigma^{(2)}, J) + \dots + H_N(\sigma^{(s)}, J))\right). \tag{6}$$

The overlaps between two replicas a, b are defined according to

$$q_{ab}(\sigma^{(a)}, \sigma^{(b)}) = \frac{1}{N} \sum_i \sigma_i^{(a)} \sigma_i^{(b)},$$

and they satisfy the obvious bounds

$$-1 \leq q_{ab} \leq 1.$$

For a generic smooth function F of the overlaps, we define the $\langle\rangle$ averages

$$\langle F(q_{12}, q_{13}, \dots)\rangle = E\Omega_J(F(q_{12}, q_{13}, \dots)),$$

where the Boltzmann averages Ω_J acts on the replicated σ variables, and E is the average with respect to the external noise J.

We remark here that the noise average E introduces correlations between different groups of replicas, which would be otherwise independent under the Boltzmann averages Ω_J, as for example q_{12} and q_{34}.

The $\langle\rangle$ averages are obviously invariant under permutations of the replicas.

Overlap distributions play a very important role in the theory. For example, a simple direct calculation shows [2], [11] that

$$\partial_\beta \alpha_N(\beta) = \frac{\beta}{2}(1 - \langle q_{12}^2 \rangle). \tag{7}$$

In order to introduce our treatment, based on flow equations with respect to the parameters of the theory, it is convenient to start from a *Boltzmannfaktor* given by

$$\exp\left(\sqrt{\frac{t}{N}} \sum_{(i,j)} J_{ij}\sigma_i\sigma_j + \beta h \sum_i \sigma_i + \sqrt{x} \sum_i J_i \sigma_i\right). \tag{8}$$

Notice that we have introduced an auxiliary additional one body random interaction ruled by the strength $\sqrt{x}$, $x \geq 0$, and N quenched independent and identically distributed centered unit Gaussian random variables J_i, so that

$$E(J_i) = 0, \quad E(J_i^2) = 1.$$

In order to get the original model, we have to put $x = 0$ at the end. Moreover, we have written $\beta = \sqrt{t}$, with $t \geq 0$. The variables t, and x, will play the role of time variable, and space variable, respectively, in our flow equations. Now we define the partition function Z by using the *Boltzmannfaktor* (8), and the auxiliary function $\alpha_N(x,t)$ in the form

$$\alpha_N(x,t) = N^{-1} E \log Z_N.$$

Then, as in the proof of (7), we have

$$\partial_t \alpha_N(x,t) = \frac{1}{4}(1 - \langle q_{12}^2 \rangle) \tag{9}$$

$$\partial_x \alpha_N(x,t) = \frac{1}{2}(1 - \langle q_{12} \rangle). \tag{10}$$

It is very simple to calculate explicitly the average $N^{-1} E \log Z_N$ for $t = 0$, at a generic strength x_0 of the one body random interaction. In fact, at $t = 0$, the interaction factorizes, and the spins at different sites become independent. Therefore we have

$$\alpha(x_0, 0) = \log 2 + \int \log\cosh(\beta h + z\sqrt{x_0})\, d\mu(z), \tag{11}$$

independently of N, where $d\mu$ is the centered unit Gaussian, representing each single J_i,

$$d\mu(z) = \exp(-\frac{z^2}{2}) dz/\sqrt{2\pi}.$$

Starting from (11), (10) at $t = 0$, we can immediately calculate the order parameter $\overline{q}(x_0)$ according to

$$\overline{q}(x_0) = \langle q_{12} \rangle (x_0, 0) = \int \tanh^2(\beta h + z\sqrt{x_0}) d\mu(z).$$

Let us now consider consider linear trajectories at constant velocity given by

$$x(t) = x_0 - \overline{q}(x_0)t, \tag{12}$$

where x_0 is the initial starting point. It is immediate to verify that there is only one trajectory passing for a generic point (x, t), $x \geq 0$, $t \geq 0$, *i.e.* the initial point is uniquely determined by (x, t), and so is for $\overline{q}$. In fact, the following Theorem holds.

Theorem 1. *Consider first the case of non zero external field h. Then, for a generic point (x, t), $x \geq 0$, $t \geq 0$, there exists a unique $x_0(x, t)$ such that*

$$x = x_0(x, t) - \overline{q}(x_0(x, t))t,$$

and a unique $\overline{q}(x, t) = \overline{q}(x_0(x, t))$, such that

$$\overline{q}(x, t) = \int \tanh^2(\beta h + z\sqrt{x + \overline{q}(x, t)t}) d\mu(z).$$

If $h = 0$, then invertibility is assured, with some $x_0(x, t) > 0$, in the region $x \geq 0$, $t \geq 0$, with the exclusion of the segment $x = 0$, $0 \leq t \leq 1$.

The proof is very simple, and can be found in [15].

By following the methods of [15], sum rules connecting α and its SK approximation can be easily found by using transport equations. They involve overlap fluctuations. In fact, let us consider α along the trajectories in (12), *i.e.* $\alpha(x(t), t)$. An easy calculation gives

$$\frac{d}{dt}\alpha_N(x(t), t) = \frac{1}{4}(1 - \overline{q}(x_0))^2 - \frac{1}{4}\langle (q_{12} - \overline{q})^2 \rangle. \tag{13}$$

Now we integrate along t, take into account the initial condition (11), and the invertibility of (12), and find the sum rule

$$\overline{\alpha}(x, t) = \alpha_N(x, t) + \frac{1}{4}\int_0^t \langle (q_{12} - \overline{q})^2 \rangle_{x(t'), t'} dt'. \tag{14}$$

Here, we have defined the replica symmetric Sherrington-Kirkpatrick solution [1], [2], in the form

$$\overline{\alpha}(x, t) = \log 2 + \int \log\cosh(\beta h + z\sqrt{x}_0) d\mu(z) + \frac{t}{4}(1 - \overline{q}(x_0))^2. \tag{15}$$

A very simple, but important, consequence of the sum rule is that α_N is dominated by its replica symmetric solution, uniformly in N,

$$\alpha_N(x,t) = N^{-1} E \log Z_N \leq \overline{\alpha}(x,t). \tag{16}$$

This is a simple consequence of the positivity of the term under integration in (14).

Now, we are ready to explain our method of quadratic coupling, starting with the simple case of zero external field, and then going to the case of nontrivial external fields.

3 Quadratic coupling for zero external field

The high temperature region ($\beta < 1$) of the zero external field SK model is a very particular case where everything can be computed. As it is well known [3], in this case the annealed approximation is exact in the infinite volume limit. In fact, we have

$$\lim_{N\to\infty} \frac{1}{N} E \ln Z_N(t,J) = \overline{\alpha}(t) \equiv \ln 2 + \frac{t}{4} = \frac{1}{N} \ln E Z_N(t,J) + \frac{t}{4N}. \tag{17}$$

In this section we give a new proof of Eq. (17), based on sum rules for the free energy. Our method is very simple and can be easily extended to the case of nontrivial external field, considered in the next Section.

When $h=0$, $x=0$, then also $\bar{q}=0$ and $x_0=0$. Then, the sum rule in (14) reads

$$\alpha_N(t) = \overline{\alpha}(t) - \frac{1}{4}\int_0^t \langle q_{12}^2\rangle_{t'}\, dt'. \tag{18}$$

The presence of $\langle q_{12}^2\rangle$, as order parameter, suggests to couple two replicas with a term proportional to the square of the overlap, the corresponding partition function being

$$\tilde{Z}_N(t,\lambda,J) = \sum_{\{\sigma,\sigma'\}} \exp\left(\sqrt{\frac{t}{N}} \sum_{(i,j)} J_{ij}(\sigma_i\sigma_j + \sigma'_i\sigma'_j) + \frac{\lambda}{2} N q_{12}^2 \right), \tag{19}$$

with $\lambda \geq 0$. The effect of the added term is to give a larger weight to the configurations having $q_{12} \neq 0$, thus favoring nonselfaveraging of the overlap. Of course, the system possesses spin-flip symmetry also for $\lambda \neq 0$, so that $\langle q_{12}\rangle = 0$. Therefore, if $\langle q_{12}^2\rangle \neq 0$ then the overlap is nonselfaveraging. Now replica symmetry is explicitly broken.

The basic idea of our method is to show that, as long as $t < 1$ and λ is small enough, the term $\lambda N q_{12}^2$ does not change the value of the free energy in the thermodynamic limit. Therefore, "most" configuration must have $q_{12} = 0$ and the overlap must be selfaveraging. In order to implement this intuitive idea, one introduces the λ dependent auxiliary function

$$\tilde{\alpha}_N(t,\lambda) = \frac{1}{2N} E \ln \tilde{Z}_N,$$

where the normalization factor 1/2 is chosen so that $\tilde{\alpha}_N(t,0) = \alpha_N(t)$. Through a simple explicit calculation, we can easily calculate the t derivative in the form

$$\partial_t \tilde{\alpha}_N = \frac{1}{4}\left(1 + \langle q_{12}^2\rangle - 2\langle q_{13}^2\rangle\right), \tag{20}$$

where now all averages $\langle\,\rangle$ involve the λ-dependent state with *Boltzmannfaktor* given in agreement with (19). Moreover, it is obvious that

$$\partial_\lambda \tilde{\alpha}_N = \frac{1}{4}\langle q_{12}^2\rangle.$$

Starting from some point $\lambda_0 > 0$, consider the linear trajectory $\lambda(t) = \lambda_0 - t$, with obvious invertibility in the form $\lambda_0 = \lambda + t$. Let us take the t derivative of $\tilde{\alpha}_N$ along this trajectory

$$\frac{d}{dt}\tilde{\alpha}_N(t,\lambda(t)) = (\partial_t - \partial_\lambda)\,\tilde{\alpha}_N = \frac{1}{4} - \frac{1}{2}\langle q_{13}^2\rangle_{t,\lambda(t)}.$$

Notice that the term containing $\langle q_{12}^2\rangle$ disappeared. By integration we get the sum rule and the inequality

$$\tilde{\alpha}_N(t,\lambda) = \frac{t}{4} + \tilde{\alpha}_N(0,\lambda_0) - \frac{1}{2}\int_0^t \langle q_{13}^2\rangle_{t',\lambda(t')}\,dt' \le \frac{t}{4} + \tilde{\alpha}_N(0,\lambda_0), \tag{21}$$

where $\langle q_{13}^2\rangle_{t',\lambda(t')}$ refers to $\lambda(t') = \lambda_0 - t' = \lambda + t - t'$.

Next, we compute $\tilde{\alpha}_N(0,\lambda_0)$. We introduce an auxiliary unit Gaussian z, and perform simple rescaling, in order to obtain

$$\begin{aligned}
\tilde{\alpha}_N(0,\lambda_0) &= \frac{1}{2N}\ln\sum_{\{\sigma,\sigma'\}} e^{\frac{1}{2}\lambda_0 N q_{12}^2} = \frac{1}{2N}\ln\sum_{\{\sigma,\sigma'\}}\int e^{\sqrt{\lambda_0 N}q_{12}z}\,d\mu(z) && (22)\\
&= \ln 2 + \frac{1}{2N}\ln\int\left(\cosh z\sqrt{\frac{\lambda_0}{N}}\right)^N d\mu(z)\\
&= \ln 2 + \frac{1}{2N}\ln\int dy\sqrt{\frac{N\lambda_0}{2\pi}}\exp N\left(-\lambda_0\frac{y^2}{2} + \ln\cosh(y\lambda_0)\right), && (23)
\end{aligned}$$

where we performed the change of variables z to $y\sqrt{N\lambda_0}$ in the last step. It is immediately recognized that the integral in (23) appears in the ordinary treatment of the well known ferromagnetic mean field Curie-Weiss model. The saddle point method gives immediately

$$\lim_{N\to\infty}\tilde{\alpha}_N(0,\lambda_0) = \ln 2 + \frac{1}{2}\max_y\left(-\lambda_0\frac{y^2}{2} + \ln\cosh(y\lambda_0)\right). \tag{24}$$

Therefore, the critical value for λ_0 is $\lambda_c = 1$. For $\lambda_0 > 1$ we have

$$\lim_{N\to\infty} \tilde{\alpha}_N(0,\lambda_0) > \ln 2,$$

while for $\lambda_0 < 1$, one can use the elementary property $2\ln\cosh x \leq x^2$ to find

$$\tilde{\alpha}_N(0,\lambda_0) \leq \ln 2 + \frac{1}{4N} \ln \frac{1}{1-\lambda_0}. \tag{25}$$

Notice that, when λ_0 approaches the value 1^-, the term of order $1/N$ diverges, since Gaussian fluctuations around the saddle point become larger and larger.

Thanks to (25), the inequality in (21) becomes

$$\tilde{\alpha}_N(t,\lambda) \leq \overline{\alpha}(t) + \frac{1}{4N} \ln \frac{1}{1-\lambda_0},$$

which holds for $0 \leq \lambda_0 < 1$, i.e., for $0 \leq t + \lambda < 1$.

Next, we use convexity of $\tilde{\alpha}_N(t,\lambda)$ with respect to λ and the fact that

$$\partial_\lambda \tilde{\alpha}_N(t,\lambda)|_{\lambda=0} = \frac{1}{4}\langle q_{12}^2\rangle_t$$

to write

$$\alpha_N(t) + \frac{\lambda}{4}\langle q_{12}^2\rangle_t \leq \tilde{\alpha}_N(t,\lambda) \leq \overline{\alpha}(t) + \frac{1}{4N} \ln \frac{1}{1-\lambda-t},$$

for $\lambda > 0$. For $0 \leq t \leq \bar{t} < 1$, choose $\lambda = (1-\bar{t})/2$, so that

$$\lambda + t \leq \bar{\lambda}_0 \equiv (1+\bar{t})/2 < 1,$$

and

$$\frac{1}{4}\langle q_{12}^2\rangle_t \leq \frac{1}{\lambda}(\overline{\alpha}(t) - \alpha_N(t)) + \frac{1}{4N\lambda} \ln \frac{1}{1-\bar{\lambda}_0}. \tag{26}$$

Recalling Eq. (18), one has

$$\frac{d}{dt}(\overline{\alpha}(t) - \alpha_N(t)) = \frac{1}{4}\langle q_{12}^2\rangle_t \leq \frac{1}{\lambda}(\overline{\alpha}(t) - \alpha_N(t)) + \frac{1}{4N\lambda} \ln \frac{1}{1-\bar{\lambda}_0}, \tag{27}$$

so that

$$\alpha_N(t) = \overline{\alpha}(t) + O(1/N),$$

uniformly for $0 \leq t \leq \bar{t} < 1$. Of course, from Eq. (21) and convexity of $\tilde{\alpha}_N$ one also has

$$\tilde{\alpha}_N(t,\lambda) = \overline{\alpha}(t) + O(1/N),$$
$$\langle q_{13}^2\rangle_{t,\lambda} = O(1/N),$$

for $0 \leq t + \lambda \leq \bar{\lambda}_0 < 1$.

We have gained a complete control of the system in the triangular region $0 \leq t < 1$, $0 \leq \lambda < 1 - t$. Note that we have not only proved Eq. (17) but we also shown that the leading correction to annealing is of order at most $1/N$.

4 The general case

The method we follow for the general case, where the *Boltzmannfaktor* is given by (8), is a direct generalization of the one explained in the previous section. In fact, by taking into account the t derivative in (13), we are led to introduce the auxiliary function

$$\tilde{\alpha}_N(x, \lambda, t) = \frac{1}{2N} E \ln \tilde{Z}_N(x, \lambda, t; J),$$

where $\tilde{Z}_N$ is the partition function for a system of two replicas coupled by the term

$$\frac{\lambda}{2} N (q_{12} - \bar{q}(x, t))^2,$$

with $\lambda \geq 0$. In order to simplify notation, we omit the argument h.

Now the t derivative is given by

$$\partial_t \tilde{\alpha}_N = \frac{1}{4} \left(1 + \langle q_{12} \rangle - 2 \langle q_{13} \rangle\right) + \frac{\lambda}{2} (\bar{q} - \langle q_{12} \rangle) \frac{\partial \bar{q}}{\partial t},$$

while the x and λ derivatives appear as

$$\partial_x \tilde{\alpha}_N = \frac{1}{2} \left(1 + \langle q_{12} \rangle - 2 \langle q_{13} \rangle\right) + \frac{\lambda}{2} (\bar{q} - \langle q_{12} \rangle) \frac{\partial \bar{q}}{\partial x},$$
$$\partial_\lambda \tilde{\alpha}_N = \frac{1}{4} \left\langle (q_{12} - \bar{q}(x, t))^2 \right\rangle .$$

Starting from points $\lambda_0 > 0$, x_0, consider the linear trajectories $\lambda(t) = \lambda_0 - t$, $x(t) = x_0 - \bar{q}(x_0) t$, as in (12), with obvious invertibility as explained before. Since $\bar{q}$ is constant along the trajectory, one finds for the total time derivative of $\tilde{\alpha}_N$

$$\frac{d}{dt} \tilde{\alpha}_N(x(t), \lambda(t), t) = (\partial_t - \bar{q}\, \partial x - \partial_\lambda)\, \tilde{\alpha}_N = \frac{1}{4}(1 - \bar{q})^2 - \frac{1}{2} \left\langle (q_{13} - \bar{q})^2 \right\rangle .$$

Notice that in this case the term containing $\langle (q_{12} - \bar{q})^2 \rangle$ disappeared.

By integration, we get the sum rule and the inequality

$$\begin{aligned} \tilde{\alpha}_N(x, \lambda, t) &= \frac{t}{4}(1 - \bar{q})^2 + \tilde{\alpha}_N(x_0, \lambda_0, 0) - \frac{1}{2} \int_0^t \left\langle (q_{13} - \bar{q})^2 \right\rangle_{t', x(t'), \lambda(t')} \, dt' \\ &\leq \frac{t}{4}(1 - \bar{q})^2 + \tilde{\alpha}_N(x_0, \lambda_0, 0), \end{aligned}$$

where $\langle (q_{13} - \bar{q})^2 \rangle_{t', x(t'), \lambda(t')}$ refers to

$$\begin{aligned} \lambda(t') &= \lambda_0 - t' = \lambda + t - t', \\ x(t') &= x_0 - \bar{q} t' = x + \bar{q}(t - t'). \end{aligned}$$

If Ω_J is the product state for two replicas with the original *Boltzmannfaktor* given by (8), then we can write

$$\tilde{\alpha}_N(x,\lambda,t) - \alpha_N(x,t) \equiv \frac{1}{2N} E \ln \Omega_J \left(\exp \frac{1}{2} \lambda N (q_{12} - \bar{q})^2 \right).$$

Therefore, by exploiting Jensen inequality, we have, for $\lambda \geq 0$,

$$\frac{\lambda}{4} \left\langle (q_{12} - \bar{q})^2 \right\rangle_{x,t} \leq \tilde{\alpha}_N(x,\lambda,t) - \alpha_N(x,t).$$

Let us also define

$$\Delta_N(x_0,\lambda_0) \equiv \tilde{\alpha}_N(x_0,\lambda_0,0) - \alpha(x_0,0) = \frac{1}{2N} E \ln \Omega_J^0 \left(\exp \frac{1}{2} \lambda_0 N (q_{12} - \bar{q})^2 \right), \tag{28}$$

where we have introduced the state Ω_J^0 for two replicas, corresponding to $t = 0$, and $x = x_0$, in (8). Notice that Ω_J^0 is a factor state over the sites i.

By collecting all our definitions and inequalities, and taking into account the definition (15), we have

$$\frac{\lambda}{4} \left\langle (q_{12} - \bar{q})^2 \right\rangle_{x,t} \leq \Delta_N(x_0,\lambda_0) + \bar{\alpha}(x,t) - \alpha_N(x,t).$$

Let us now introduce $\lambda_c(x_0)$ such that, for any $\lambda_0 \leq \lambda_c(x_0)$, one has

$$\lim_{N\to\infty} \Delta_N(x_0,\lambda_0) = 0.$$

Then, by the same reasoning already exploited starting from (26), and taking into account (13), we obtain the proof of the following

Theorem 2. *For any $t \leq \lambda_c(x_0(x,t))$, where $x_0(x,t)$ is defined as in Theorem 1, we have the convergence*

$$\lim_{N\to\infty} \alpha_N(x,t) = \bar{\alpha}(x,t). \tag{29}$$

For the specification of $\lambda_c(x_0)$, we can easily establish the complete characterization of the Δ_N limit. In fact, the following holds.

Theorem 3. *The infinite volume limit of Δ_N is given by*

$$\lim_{N\to\infty} \Delta_N(x_0,\lambda_0) = \Delta(x_0,\lambda_0).$$

Here, $\Delta(x_0,\lambda_0)$ is defined through the variational expression

$$\Delta(x_0,\lambda_0) \equiv \frac{1}{2} \max_{\mu} \left(\int \ln(\cosh\mu + \tanh^2(\beta h + z\sqrt{x_0}) \sinh\mu) d\rho(z) - \mu\bar{q} - \frac{\mu^2}{2\lambda_0} \right), \tag{30}$$

where $d\rho(z)$ is the centered unit Gaussian measure.

Of course, the expression (30) is in agreement with (24), when there are no external fields.

It is easy to realize that the value λ_c, in the general case, is strictly less than the expected value t_c, following from the Almeida-Thouless argument,

$$t_c \int \cosh^{-4}(\beta h + z\sqrt{x_0}) d\mu(z) = 1.$$

The proof of the Theorem 3 is easy. First of all let us establish the elementary bound, uniform in N,

$$\Delta_N(x_0, \lambda_0) \geq \Delta(x_0, \lambda_0). \tag{31}$$

In fact, starting from the definition of $\Delta_N(x_0, \lambda_0)$ given in (28), we can write, for $\lambda_0 \neq 0$, and any μ,

$$(q_{12} - \bar{q})^2 \geq 2\frac{\mu}{\lambda_0}(q_{12} - \bar{q}) - \left(\frac{\mu}{\lambda_0}\right)^2,$$

and conclude that

$$\Delta_N(x_0, \lambda_0) \geq \alpha_0(\mu) - \frac{\mu^2}{4\lambda_0}, \tag{32}$$

where we have defined

$$\begin{aligned}
\alpha_0(\mu) &\equiv \frac{1}{2N} E \ln \Omega_J^0 \left(\exp \mu N (q_{12} - \bar{q})\right) \\
&= \frac{1}{2} \int \ln \left(\cosh \mu + \tanh^2(\beta h + z\sqrt{x_0}) \sinh \mu\right) \, d\rho(z) - \frac{1}{2}\mu\bar{q}.
\end{aligned}$$

Of course, it is convenient to take the $\max_\mu$ in the r.h.s. of (32), so that the bound in (31) is established. The proof that the bound is in effect the limit, as $N \to \infty$, can be obtained in a very simple way by using a Gaussian transformation on (28), as it was done in (22). In fact, we now have

$$\begin{aligned}
&\frac{1}{2N} E \ln \Omega_J^0 \left(\exp \frac{1}{2}\lambda_0 N (q_{12} - \bar{q})^2\right) \\
&= \frac{1}{2N} E \ln \int \Omega_J^0 \left(\exp \sqrt{\lambda_0 N}(q_{12} - \bar{q}) z\right) \, d\rho(z)
\end{aligned} \tag{33}$$

Therefore, by exploiting the fact that also Ω_J^0 factorizes with respect to the sites i, we can write

$$\begin{aligned}
\Delta_N(x_0, \lambda_0) = &\frac{1}{2N} E \ln \int \prod_i \left(\cosh \sqrt{\frac{\lambda_0}{N}} z + \tanh^2(\beta h + J_i\sqrt{x_0}) \sinh \sqrt{\frac{\lambda_0}{N}} z\right) \\
&\times \exp\left(-\sqrt{\lambda_0 N}\bar{q} z\right) \, d\rho(z).
\end{aligned} \tag{34}$$

Now, we find convenient to introduce a small $\epsilon > 0$, so that

$$\frac{1}{\lambda_0} = \frac{1}{\lambda_0'} + \epsilon. \tag{35}$$

Notice that $\lambda_0 < \lambda_0'$. We also introduce the auxiliary (random) function

$$\phi_N(y, \lambda_0') \equiv \frac{1}{N} \sum_i \ln \left(\cosh y + \tanh^2(\beta h + J_i \sqrt{x_0}) \sinh y \right) - \bar{q} y - \frac{1}{2} \frac{y^2}{\lambda_0'}.$$

By the strong law of large numbers, as $N \to \infty$, for any y, we have the J almost sure convergence of $\phi_N(y, \lambda_0')$ to $\phi(y, \lambda_0')$ defined by

$$\begin{aligned} \phi(y, \lambda_0') &\equiv \int \ln \left(\cosh y + \tanh^2(\beta h + z \sqrt{x_0}) \sinh y \right) d\rho(z) - \bar{q} y - \frac{1}{2} \frac{y^2}{\lambda_0'} \\ &= E \phi_N(y, \lambda_0'). \end{aligned}$$

Here $d\rho(z)$ performs the averages with respect to the J_i variables. Let us also remark that the convergence is J almost surely uniform for any finite number of values of the variable y.

Now we can go back to (34), write explicitly the unit Gaussian measure $d\rho(z)$, perform the change of variables $y = z\sqrt{\lambda_0 N^{-1}}$, make the transformation (35), take the $\sup_y$ for the ϕ_N, and perform the residual Gaussian integration over y. We end up with the estimate

$$\Delta_N(x_0, \lambda_0) \leq \frac{1}{2} E \sup_y \phi_N(y, \lambda_0') + \frac{1}{2N} \ln \frac{1}{\sqrt{\lambda_0 \epsilon}}. \tag{36}$$

Since the J-dependent $\sup_y$ is reached in some finite interval, for any fixed λ_0', and the function ϕ_N is continuous with respect to y, with bounded derivatives, we can perform the $\sup_y$ with y running over a finite discrete mesh of values, by tolerating a small error, which becomes smaller and smaller as the mesh interval is made smaller. But in this case the strong law of large numbers allows us to substitute ϕ_N with ϕ, in the infinite volume limit $N \to \infty$. On the other hand, the second term in the r.h.s. of (36) vanishes in the limit. Therefore, we conclude that

$$\limsup_{N \to \infty} \Delta_N(x_0, \lambda_0) \leq \frac{1}{2} \sup_y \phi(y, \lambda_0').$$

From continuity with respect to λ_0, we can let λ_0' approach λ_0, and the Theorem is proven.

5 Fluctuations of overlaps and free energy

In the previous sections we proved that, in a certain region of the parameters $t, x, \beta h$, the typical values of the free energy $\ln Z_N/N$ and of the overlap q_{ab} are the replica symmetric expressions $\bar{\alpha}$ and $\bar{q}$, respectively. In the same region, one can obtain a more precise characterization of the fluctuations of these quantities, for $N \to \infty$, showing that a central limit-type theorem holds, after suitable rescaling. This will be analyzed in detail in a subsequent paper [17]. Here, we just give the main results and sketch the ideas underlying the proof.

Concerning the fluctuations of the overlap around the Sherrington-Kirkpatrick order parameter $\bar{q}$, we prove the following

Theorem 4. *[17] The rescaled overlap variables*

$$\xi_{ab}^N \equiv \sqrt{N}(q_{ab} - \bar{q})$$

tend in distribution, as $N \to \infty$, to centered, jointly Gaussian variables with covariances

$$\begin{aligned}\langle \xi_{ab}^2 \rangle &= A(t, x, h)\\ \langle \xi_{ab}\xi_{ac} \rangle &= B(t, x, h)\\ \langle \xi_{ab}\xi_{cd} \rangle &= C(t, x, h),\end{aligned}$$

where $b \neq c$, $c \neq a, b$ and $d \neq a, b$. The expressions of A, B, C are explicitly given and coincide with those found in [15].

Recently, an analogous result was proved independently by Talagrand [20], who computed the $N \to \infty$ limit for all moments of the ξ variables.

The scheme of our proof is as follows: The control we obtained on the coupled two replica system and concentration of measure inequalities for the free energy [21] imply that the fluctuations of q_{ab} from $\bar{q}$ are exponentially suppressed for N large. Then, by means of the cavity method [2] one can write a self-consistent closed equation for the characteristic function of the variables ξ_{ab}^N. This equation turns out to be linear, apart from error terms which vanish asymptotically for $N \to \infty$, thanks to the strong suppression of the overlap fluctuations. The solution, which is easily found, coincides with the characteristic function of a Gaussian distribution, with the correct covariance structure.

Concerning the free energy, the result we prove is the following:

Theorem 5. *[17] Define the rescaled free energy fluctuation as*

$$\hat{f}_N(t, x, h; J) \equiv \sqrt{N}\left(\frac{\ln Z_N(t, x, h; J)}{N} - \bar{\alpha}(t, x, h)\right).$$

Then,

$$\hat{f}_N(t, x, h; J) \xrightarrow{d} \mathcal{N}(0, \sigma^2(t, x, h)),$$

where $\mathcal{N}(m, \sigma^2)$ denotes the Gaussian random variable of mean m and variance σ^2, and

$$\sigma^2(t, x, h) = Var\left(\ln\cosh(z\sqrt{\bar{q}t + x} + \beta h)\right) - \frac{\bar{q}^2 t}{2}.$$

Here, $Var(.)$ denotes the variance of a random variable and $z = \mathcal{N}(0, 1)$.

This result is a consequence of Theorem 4 and of concentration of measure inequalities for the free energy.

6 Outlook and conclusions

We obtained control on the thermodynamic limit of the model, in a region above the Almeida-Thouless line, by suitably coupling two replicas of the system and studying stability with respect to the coupling parameter. The question naturally arises, whether and how this method can be extended up to the expected critical line. This problem seems to be common to all approaches proposed so far.

The method can be also further generalized to the case where more and more replicas are mutually coupled. In this case, replica symmetry is explicitly broken at various levels, and it is possible to give a generalization of the Ghirlanda-Guerra relations [13]. We plan to report soon on these generalizations [22].

Acknowledgments

We gratefully acknowledge useful conversations with Pierluigi Contucci, Roberto D'Autilia, Sandro Graffi, Enzo Marinari, Giorgio Parisi, Masha Scherbina, and Michel Talagrand.

This work was supported in part by MIUR (Italian Minister of Instruction, University and Research), and by INFN (Italian National Institute for Nuclear Physics).

References

[1] Sherrington, D. and Kirkpatrick, S. *Solvable model of a spin glass*, Phys. Rev. Lett. **35**, 1792 (1975); Sherrington, D. and Kirkpatrick, S. *Infinite-ranged models of spin-glasses*, Phys. Rev. B **17**, 4384 (1978).

[2] Mézard, M., Parisi, G. and Virasoro, M. A. *Spin Glass Theory and Beyond*, World Scientific, Singapore (1987).

[3] Aizenman, M., Lebowitz, J. and Ruelle, D. *Some rigorous results on the Sherrington-Kirkpatrick model of spin glasses*, Commun. Math. Phys. **112**, 3 (1987).

[4] Fröhlich, J. and Zegarlinski, B. *Some comments on the Sherrington-Kirkpatrick model of spin glasses*, Commun. Math. Phys. **112**, 553 (1987).

[5] Guerra, F. *Fluctuations and thermodynamic variables in mean field spin glass models*, in: *Stochastic Processes, Physics and Geometry*, S. Albeverio *et al.*, eds, World Scientific, Singapore (1995).

[6] Shcherbina, M. *On the replica symmetric solution for the Sherrington-Kirkpatrick model*, Helv. Phys. Acta **70**, 838 (1997).

[7] Comets, F. and Neveu, J. *The Sherrington-Kirkpatrick model of spin glasses and stochastic calculus: the high temperature case*, Commun. Math. Phys. **166**, 549 (1995).

[8] Talagrand, M. *On the high temperature phase of the Sherrington-Kirkpatrick model*, Annals of Probability, to appear.

[9] Pastur, L. A. and Shcherbina, M. V. *The Absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model*, J. Stat. Phys. **62**, 1 (1991).

[10] Talagrand, M. *The Sherrington-Kirkpatrick model: a challenge to mathematicians*, Probability Theory and Related Fields **110**, 109 (1998).

[11] Guerra, F. *About the overlap distribution in mean field spin glass models*, International Journal of Modern Physics B **10**, 1675 (1996).

[12] Aizenman, M. and Contucci, P. *On the stability of the quenched state in mean field spin glass models*, J. Stat. Phys. **92**, 765 (1998).

[13] Ghirlanda, S. and Guerra, F. *General properties of overlap probability distributions in disordered spin systems. Toward Parisi ultrametricity*, J. Phys. A: Math. Gen. **31**, 9149 (1998).

[14] Baffioni, F. and Rosati, F. *On the ultrametric overlap distribution for mean field spin glass models (I)*, The European Physical Journal B **17**, 439 (2000).

[15] Guerra, F. *Sum rules for the free energy in the mean field spin glass model*, Fields Institute Communications **30**, 161 (2001).

[16] de Almeida, J. R. L. and Thouless, D. J. *Stability of the Sherrington-Kirkpatrick solution of a spin glass model*, J. Phys. A **11**, 983 (1978).

[17] Guerra, F. and Toninelli, F. *Central limit theorem for fluctuations in the high temperature region of the Sherrington-Kirkpatrick spin glass model*, to appear.

[18] Marinari, E., Parisi, G., Ricci-Tersenghi, F., Ruiz-Lorenzo, J. J. and Zuliani, F. *Replica symmetry breaking in short range spin glasses: a review of the theoretical foundations and of the numerical evidence*, J. Stat. Phys **98**, 973 (2000).

[19] Newman, C. and Stein, D. *Non-mean-field behavior of realistic spin glasses*, Phys. Rev. Lett. **76**, 515 (1996).

[20] Talagrand, M. private communication.

[21] Talagrand, M. *Mean field models for spin glasses: a first course*, course given in Saint Flour in summer 2000, to appear.

[22] Guerra, F. and Toninelli, F. L. in preparation.