Eur. Phys. J. B **20**, 217–233 (2001)

THE EUROPEAN PHYSICAL JOURNAL B

EDP Sciences

The Bethe lattice spin glass revisited

M. Mézard[1,a] and G. Parisi[2]

[1] Laboratoire de Physique Théorique et Modèles Statistiques, Université Paris Sud, Bâtiment 100, 91405 Orsay Cedex, France
[2] Dipartimento di Fisica, Sezione INFN and Unità INFM, Università di Roma "La Sapienza", Piazzale Aldo Moro 2, 00185 Rome, Italy

Received 27 September 2000

Abstract. So far the problem of a spin glass on a Bethe lattice has been solved only at the replica symmetric level, which is wrong in the spin glass phase. Because of some technical difficulties, attempts at deriving a replica symmetry breaking solution have been confined to some perturbative regimes, high connectivity lattices or temperature close to the critical temperature. Using the cavity method, we propose a general non perturbative solution of the Bethe lattice spin glass problem at a level of approximation which is equivalent to a one step replica symmetry breaking solution. The results compare well with numerical simulations. The method can be used for many finite connectivity problems appearing in combinatorial optimization.

PACS. 75.10.Nr Spin-glass and other random models

1 Introduction

The spin glass problem has been around for twenty five years, but its understanding has turned out to be remarkably complicated. It is generally considered as solved only in its fully connected version introduced by Sherrington and Kirkpatrick [1]. The first consistent solution was derived with the replica method [2,3] and it was then confirmed using a probabilistic approach, the cavity method, which avoids the strange (and powerful) mathematical subtleties of the replica approach [3,4]. A rigorous proof of the validity of the solution is still lacking, in spite of recent progress [5–7].

A slightly more realistic theory of spin glasses, still of the mean field type, deals with the situation in which each spin interacts only with a finite number of neighbours. Models of this type include the spin glass on a Cayley tree, a Bethe lattice and a disordered random lattice with fixed or with fluctuating connectivity. There are many motivations for studying such problems. On one hand one may hope to get a better knowledge on the finite dimensional problem, since these models include a notion of neighborhood which is absent in the infinite range case. But another motivation is the possibility to solve these problems using different methods, like iterative methods which are typical of statistical mechanics on tree-like structures. In fact the cavity method is a generalization of the Bethe Peierls iterative method to the case in which there may exist several pure states, and it is therefore very natural to work out the details of this generalisation, and to test its validity. Another important aspect comes from the connection between the statistical mechanics of disordered systems and the optimization problems: many of the interesting random optimization problems turn out to have a finite connectivity structure. This is the case for instance of the travelling salesman problem [8], the matching [9] the graph partitioning [10,11] or the K-satisfiability problem [12].

While the problems of a spin-glass on tree-like lattices were naturally studied very soon after the discovery of spin glasses, the present status of the knowledge on these systems is still rather poor compared to that on the SK model. A lot of efforts have been devoted to the simple Bethe Peierls method which builds up a solution in terms of the distribution of local magnetic fields [13–18]. However this simple iterative solution, which may be relevant for a Cayley tree with a certain type of boundary conditions [18] is wrong for the Bethe lattice spin-glass. When the replica formalism is used, this simple iterative solution turns out to be equivalent to the replica symmetric (RS) solution. However one knows that there exists a replica symmetry breaking (RSB) instability similar to the one found in the SK model [17,19–22]. Unfortunately, in the replica formalism, the RSB solution could be found only in some rather limited regimes: expansion around the high connectivity (SK-like) limit [23], or close to the critical temperature [20]. The main problem encountered in all these attempts is a very general one, common to all disordered problems with a finite connectivity. Roughly speaking it can be summarized as follows: the distribution of local fields, even within one pure state, is not a simple Gaussian as in the infinite range problems,

[a] e-mail: mezard@ipno.in2p3.fr

but a more complicated function [9,24,25]. When one takes into account the existence of several pure states, the natural order parameter, even within the simplest "one step" replica symmetry breaking solution, becomes the probability distribution of these local field probabilities [27]. This is a functional order parameter which is difficult to handle. Several interesting attempts at solving the one step RSB equations have been done in the past [26,32,33], but they all restricted this functional order parameter to some particular subspace, within which a variational approach was used.

In this paper we present an improved solution of the Bethe lattice spin glass. This solution is nothing but the application to this problem of the cavity method, treated at a level which is equivalent to the one step RSB solution. It is valid for any connectivity and any temperature. In the next section we discuss the various tree-like lattices which are usually studied and precise our definition of the Bethe lattice problem. In Section 3 we recall the basic steps of the simple Bethe-Peierls approach, and we discuss its instability in Section 4. In Section 5 we discuss the formalism of the cavity approach at the one step RSB level. Section 6 describes the algorithm used to determine the distribution of local fields within this approach. The implementation of the algorithm is discussed in Section 7, where we derive explicit results for a lattice with six neighbours per point and compare the analytic prediction to those of numerical simulations. Finally, Section 8 contains a brief discussion and mentions the perspectives.

2 The Bethe lattice

We consider a system of N Ising spins, $\sigma_i = \pm 1$, $i \in \{1, ..., N\}$, interacting with random couplings, the energy being:

$$E = -\sum_{\langle ij \rangle} J_{ij}\sigma_i\sigma_j \ . \tag{1}$$

The sum is over all links of a lattice. For each link $\langle ij \rangle$ the coupling J_{ij} is an independent random variable chosen with the same probability distribution $P(J)$. The various types of tree-like lattices which have been considered are:

- A) The Cayley tree: starting from a central site $i = 0$, one builds a first shell of $k + 1$ neighbours. Then each of the first shell spins is connected to k new neighbours in the second shell etc. until one reaches the L'th shell which is the boundary. There is no overlap among the new neighbours, so that the graph is a tree.
- B) The random graph with fluctuating connectivity: for each pair of indices (ij), a link is present with probability c/N and absent with probability $1 - c/N$. The number of links connected to a point is a random variable with a Poisson distribution, its mean being equal to c.
- C) The random graph with fixed connectivity, equal to $k + 1$. The space of allowed graphs are all graphs such that the number of links connected to each point is equal to $k + 1$. The simplest choice, which we adopt here, is the case where every such graph has the same probability.

On a Cayley tree a finite fraction of the total number of spins lie on the boundary. The Cayley tree is thus a strongly inhomogeneous system, the properties of which are often remote from those of a usual finite dimensional problem. For this reason people generally consider instead a *Bethe lattice*, which consists of a subset of the Cayley tree containing the first L' shells. Taking the limits $L \to \infty$, $L' \to \infty$ with $L/L' \to \infty$ allows to isolate the central part of the tree, away from the boundary. This procedure is OK when one considers a ferromagnetic problem. In the case of a spin glass this definition of the Bethe lattice is not free from ambiguities: one can not totally forget the boundary conditions which are imposed on the boundary of the Cayley tree, since they are fixing the degree of frustration [22]. For this reason we prefer to define the Bethe lattice as the random lattice with fixed connectivity (lattice C defined above). Clearly on such a graph the local structure is that of a tree with a fixed branching ratio. Small loops are rare, the typical size of a loop is of order $\log N$. Therefore in the large N limit the random graph with fixed connectivity provides a well defined realisation of a Bethe lattice, *i.e.* a statistically homogeneous, locally tree-like structure. This is the lattice which we study in this paper (the case of fluctuating connectivities will be studied in a forthcoming work). Numerical simulations of this system can be found in [11,22,28,29].

Historically, spin glasses on the Bethe lattice and on diluted lattices with a fixed finite connectivity (type C) were often discussed as separate issues. The reason for these separate discussions of the same problem is the type of techniques which are used. Generally speaking the Bethe lattice papers rely on the Bethe Peierls method while the random lattice papers use the replica method. One exception is the use of the cavity method for the random lattice case [25,26]. Hereafter we shall basically develop the iterative/cavity approach, but we shall also mention at each step its connections to the replica approach.

3 The simple Bethe-Peierls 'solution'

This section will give a brief review to the standard approach to the spin glass on the Bethe lattice, defined as lattice C in the above classification. This solution is wrong, because, as we shall see later, it does not consider the phenomenon of replica symmetry breaking, however it sets the stage for the correct solution that will be presented in the next section.

3.1 The iterative approach

As is well known, on tree-like structures the problem can be solved by iteration. Let us consider in general the merging of k branches of a tree onto one site σ_0 as in Figure 1. The partition function can be computed exactly if one introduces, for each of the outside spins $\sigma_i, i \in \{1, ..., k\}$,

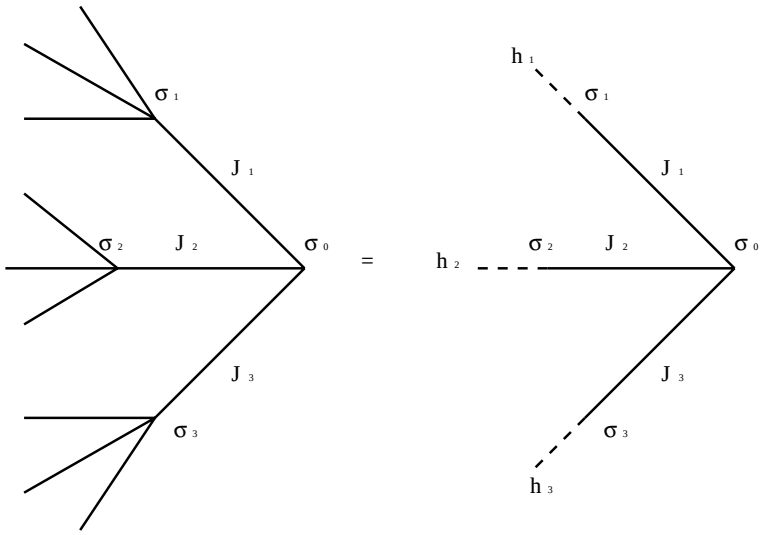

Fig. 1. The merging of k branches off the tree (here $k = 3$) onto the spin σ_0. The cavity field h_i is the total field acting on spin σ_i in the absence of the central spin σ_0.

the effective "cavity" field h_i representing the action onto the spin σ_i of all the other spins, in the absence of the central spin σ_0. In other words the magnetization of the ith spin in absence of the central spin is given by $m_i = \tanh(\beta h_i)$. The variables h_i (and consequently the variables m_i) are uncorrelated in the limit $N \to \infty$. It is crucial to consider the magnetization before the introduction of the spin σ_0, because after its introduction all the spins that are coupled to the spin σ_0 become correlated.

Calling J_i the coupling between spins σ_0 and σ_i, the partition function of the spin σ_0 is expressed as

$$\sum_{\sigma_0,\sigma_1,\ldots\sigma_k} \exp\left(\beta\sigma_0 \sum_{i=1}^{k} J_i\sigma_i + \beta \sum_{i=1}^{k} h_i\sigma_i\right). \qquad (2)$$

Let us recall here the basic identity which allows to forward the effect of the fields h_i onto spin σ_0, and which is used repeatedly in this work. For an Ising spin $\sigma_0 = \pm 1$, one has:

$$\sum_{\sigma=\pm 1} \exp\left(\beta\sigma_0 J\sigma + \beta h\sigma\right) = c(J,h)\exp(\beta u(J,h)\sigma_0) \qquad (3)$$

where we define the two functions u and c as:

$$u(J,h) = \frac{1}{\beta}\mathrm{atanh}\left[\tanh(\beta J)\tanh(\beta h)\right];$$
$$c(J,h) = 2\frac{\cosh(\beta J)\cosh(\beta h)}{\cosh(\beta u(J,h))}. \qquad (4)$$

The magnetization on site 0 is thus $m_0 = \langle\sigma_0\rangle = \tanh(\beta h_0)$, where

$$h_0 = \sum_{i=1}^{k} u(J_i, h_i); \qquad (5)$$

From this equation one gets the basic recursion relation for the probability density $Q(h)$ of local fields:

$$Q(h) = E_J \int \prod_{i=1}^{k} \left[\mathrm{d}h_i Q(h_i)\right]\ \delta\left(h - \sum_{i=1}^{k} u(J_i, h_i)\right). \qquad (6)$$

Here and throughout the paper, we denote by E_J the expectation value with respect to all the exchange coupling constants J_i: $E_J = \int \prod_i [\mathrm{d}J_i P(J_i)]$.

It will be useful for future use to introduce the probability distribution $R(u)$ of the propagated field variable $u(J,h)$:

$$R(u) = E_J \int Q(h)\mathrm{d}h\ \delta(u - u(J,h)). \qquad (7)$$

The field distribution is nothing but the convolution: $Q(h) = \int \mathrm{d}u_1 \ldots \mathrm{d}u_k\ R(u_1)\ldots R(u_k)\delta(u_1 + \ldots + u_k - h)$.

In order to relate the true distribution of local fields, $Q_t(H)$, to the distribution $Q(h)$ of local fields on one branch[1], one needs to consider the merging of $k+1$ branches onto one site. The true local field H_0 on a given site 0 is simply given by a sum of contributions from each of its $k+1$ neighbours,

$$H_0 = \sum_{j=1}^{k+1} u(J_j, h_j), \qquad (8)$$

where as before h_j is the local field on j in the absence of the spin s_0. This gives the distribution of true local fields $Q_t(H)$ as the convolution:

$$Q_t(H) = \int \prod_{i=1}^{k+1} \left[\mathrm{d}u_i R(u_i)\right] \delta\left(H - \left[\sum_{i=1}^{k+1} u_i\right]\right). \qquad (9)$$

[1] We denote by upper case letters the true local fields and by small case letters the local fields on one branch.

Let us now compute the internal energy with this method. We add a new link [30] with a coupling constant J_{ij} between two spins σ_i and σ_j, where the local fields in the absence of the new link are respectively $h_i^{(j)}$ and $h_j^{(i)}$. Then the energy of this link is:

$$E_{ij} = -J_{ij}\langle\sigma_i\sigma_j\rangle, \tag{10}$$

where the expectation value is computed using the Hamiltonian $H_{ij}(\sigma_i,\sigma_j)$, which is given by

$$H_{ij}(\sigma_i,\sigma_j) = -\left(J_{ij}\sigma_i\sigma_j + h_i^{(j)}\sigma_i + h_j^{(i)}\sigma_j\right). \tag{11}$$

A simple computation shows that

$$E_{ij} = -J_{ij}\frac{\tanh(\beta J_{ij}) + \tanh(\beta h_i^{(j)})\tanh(\beta h_j^{(i)})}{1 + \tanh(\beta J_{ij})\tanh(\beta h_i^{(j)})\tanh(\beta h_j^{(i)})}\,. \tag{12}$$

Computing the total free energy of the system is slightly more involved. Using the fact that the Bethe-Peierls approximation is exact on the tree-like lattices one can write the free energy as the sum of site and bond contributions [14–16]:

$$F = -k\sum_i F_i^{(1)} + \sum_{\langle ij\rangle} F_{\langle ij\rangle}^{(2)}, \tag{13}$$

where the contribution from the bond ij is

$$-\beta F_{\langle ij\rangle}^{(2)} = \ln\sum_{\sigma_i,\sigma_j}\exp\left(-\beta H_{ij}(\sigma_i,\sigma_j)\right) \tag{14}$$

and that from the site i is:

$$-\beta F_i^{(1)} = \ln\sum_{\sigma_i}\exp\left(\beta H_i\sigma_i\right), \tag{15}$$

where H_i is the total spin acting on spin σ_i. One can prove the validity of the expression (13) by the following two steps: 1) it clearly gives the correct free energy at high temperature; 2) using the fact that $\sum_{j(i)} h_i^{(j)} = kH_i$, where the sum is over all the neighbours j of site i, one finds that $\partial(\beta F)/\partial\beta$ gives back the correct expression for the internal energy obtained in (12). We notice *en passant* that this free energy is nothing but the generalization to a finite coordination number of the TAP free energy (and reduces to the usual TAP free energy in the limit of infinite coordination number)[15,16].

The Edwards-Anderson order parameter [34], $q = (1/N)\sum_i\langle\sigma_i\rangle^2$ can be written as the magnetization squared of a spin coupled to $k+1$ neighbours and is given by

$$q = \int \mathrm{d}H Q_t(H)\ [\tanh^2(\beta H)]. \tag{16}$$

We shall also compute the link overlap, $q^{(l)} = (2/(N(k+1))\sum_{\langle ij\rangle}\langle\sigma_i\sigma_j\rangle^2$, which is deduced from the $Q(h)$ distribution as:

$$q^{(l)} = E_J\int \mathrm{d}h\mathrm{d}h' Q(h)Q(h') \times\left(\frac{\tanh(\beta J)+\tanh(\beta h)\tanh(\beta h')}{1+\tanh(\beta J)\tanh(\beta h)\tanh(\beta h')}\right)^2. \tag{17}$$

3.2 A variational formulation

We have just seen in the previous section that, if one neglects the possibility of RSB, all the thermodynamic quantities of the Bethe lattice spin glass can be computed in terms of the probability distribution $Q(h)$ of the effective field h. This probability distribution is obtained by solving the self-consistency equation (6).

It is interesting for many reasons, some of which will become clear later, to reformulate this problem in a variational way. One can write a free energy $F[Q]$, which is a functional of the probability distribution $Q(h)$, such that:

1. The equation $\delta F/\delta Q(h) = 0$ is equivalent to the self-consistency equation (6) for $Q(h)$.
2. Calling Q^* the solution of the previous equation, the equilibrium free energy (13) is equal to $F[Q^*]$.

This free energy functional is given by:

$$\frac{F[Q]}{N} = \frac{k+1}{2}\int\prod_{i=1}^{k}\left[\mathrm{d}h_i\mathrm{d}g_i Q(h_i)Q(g_i)\right]\ F^{(2)}(h_1\dots h_k, g_1\dots g_k) - k\int\prod_{i=1}^{k+1}\left[\mathrm{d}h_i Q(h_i)\right]\ F^{(1)}(h_1\dots h_{k+1}) \tag{18}$$

where

$$-\beta F^{(1)}(h_1\dots h_{k+1}) = E_J\ln\left(\left[\prod_{i=1}^{k+1}\frac{1}{d(J_i,h_i)}\right] \times\sum_{\sigma_0,\sigma_1,\dots\sigma_{k+1}}\exp\left[\beta\sigma_0\sum_{i=1}^{k+1}J_i\sigma_i + \beta\sum_{i=1}^{k+1}h_i\sigma_i\right]\right),$$

$$-\beta F^{(2)}(h_1\dots h_k, g_1\dots g_k) = E_J E_K\ln\left(\left[\prod_{i=1}^{k}\frac{1}{d(J_i,h_i)d(K_i,g_i)}\right] \times\sum_{\sigma_0,\sigma_1,\dots\sigma_k}\ \sum_{\tau_0,\tau_1,\dots\tau_k}\exp\left[\beta J_0\sigma_0\tau_0 + \beta\sigma_0\sum_i J_i\sigma_i + \beta\sum_i h_i\sigma_i + \beta\tau_0\sum_i K_i\tau_i + \beta\sum_i g_i\tau_i\right]\right). \tag{19}$$

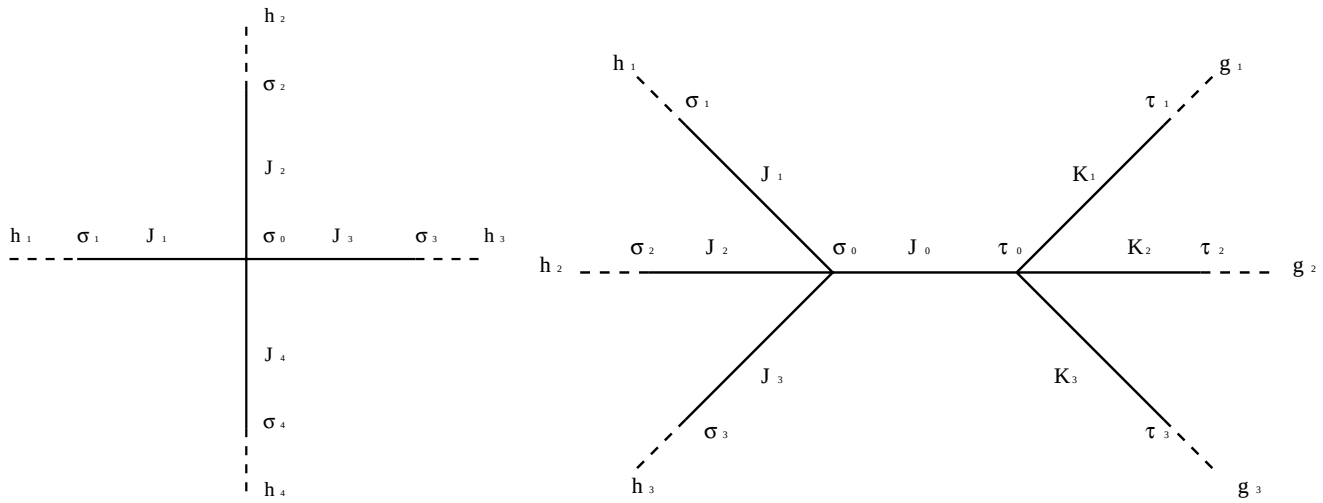

Fig. 2. A pictorial representation of the two contributions (19) to the free energy. The 'site' contribution on the top is obtained by merging $k+1$ lines onto one site (here $k=3$), and the 'bond' contribution pictured on the bottom figure is obtained by adding one new link J_0, and two new spins σ_0 and τ_0, to the lattice, together with the other k branches arriving onto each of these spins.

These two expressions are represented pictorially in Figure 2. In this formula the function $d(J,h)$ is an arbitrary positive function, since its contributions to the two pieces $F^{(1)}$ and $F^{(2)}$ cancel. We shall mainly use it with $d(J,h)=c(J,h)$, which allows an easy connection with the expression (13), but some other choice will also be useful in order to make contact with the result of the replica method, as we shall see below.

Let us now show that this free energy has the desired properties. In order to check that $Q^*(h)$ is a stationarity point of the free energy in the space of normalized probability distributions $Q(h)$(such that $\int \mathrm{d}hQ(h)=1$), we need to show that $\delta F/\delta Q(h)=$ const. when $Q=Q^*$. This functional derivative is equal to

$$\frac{1}{N}\frac{\delta F[Q]}{\delta Q(h)} = k(k+1)\int \mathrm{d}h_2\ldots\mathrm{d}h_k Q(h_2)\ldots Q(h_k) \times\left(\int \mathrm{d}g_1\ldots\mathrm{d}g_k Q(g_1)\ldots Q(g_k)\ F^{(2)}(h,h_2\ldots h_k,g_1\ldots g_k) - \int \mathrm{d}h_{k+1}Q(h_{k+1})F^{(1)}(h,h_2\ldots h_{k+1})\right). \quad (20)$$

Using (3), one easily sees that, if $Q(h)=Q^*(h)$ satisfies the self consistency equation (6), one has, for any $h_1\ldots h_k$:

$$\int \mathrm{d}g_1\ldots dg_k Q^*(g_1)\ldots Q^*(g_k)F^{(2)}(h_1\ldots h_k,g_1\ldots g_k) = A+\int \mathrm{d}h_{k+1}Q^*(h_{k+1})F^{(1)}(h_1\ldots h_{k+1}), \quad (21)$$

where A is a constant (independent of $h_1\ldots h_k$), given by

$$A=(-1/\beta)\int \mathrm{d}g_0 Q^*(g_0)E_J\left(\ln\left[d(J_0,g_0)\right] + k\ln\left[c(J_0,g_0)/d(J_0,g_0)\right]\right). \quad (22)$$

This shows that the functional derivative (20) is a constant. The repeated use of (3) allows to show similarly that the saddle point free energy $F[Q^*(h)]$ is indeed equal to the free energy (13) (the factor $(k+1)/2$ in (18) is nothing but the number of links per site).

As an extra check, one can see that the derivative of $\beta F[Q]$ we respect to β gives the internal energy of the previous section (the derivative is quite simple if we absorb most of the β's redefining the h and notice that the only explicit dependence on β comes from the term βJ).

It is interesting to note that, using the basic recursion relation (3), and the special choice $d(J,h)=2\cosh(\beta h)$, we can write the free energy under the simple form:

$$\frac{F[Q]}{N} = \frac{k+1}{2}\int\prod_{i=1}^{k}[\mathrm{d}h_i Q(h_i)]\ \ F^{(1')}(h_1\ldots h_k) - \frac{k-1}{2}\int\prod_{i=1}^{k+1}[\mathrm{d}h_i Q(h_i)]\ \ F^{(1)}(h_1\ldots h_{k+1}) \quad (23)$$

where

$$-\beta F^{(1')}(h_1\ldots h_k) = E_J\ln\left(\left[\prod_{i=1}^{k}\frac{1}{2\cosh(\beta h_i)}\right] \times\sum_{\sigma_0,\sigma_1,\ldots\sigma_k}\exp\left[\beta\sigma_0\sum_{i=1}^{k}J_i\sigma_i+\beta\sum_{i=1}^{k}h_i\sigma_i\right]\right). \quad (24)$$

One must keep in mind that this expression for the free energy is correct only if Q satisfies equation (6) and should *not* be used as a variational free energy (see also next section).

3.3 Equivalence with the replica formalism

We shall not present here the details of the replica approach to this problem, for which we refer the reader

to [19,20,23,27,32]. Let us recall the main results of [23]. In the replica approach one introduces a probability distribution $\rho(\sigma)$, where the variables σ are n Ising variables (n eventually goes to zero). One can introduce a free energy functional $F_{\rm rep}[\rho(\sigma)]$. The equilibrium free energy is given by $F(\beta) = F_{\rm rep}[\rho^*]$, where ρ^* is the solution of the stationarity equation $\delta F_{\rm rep}/\delta\rho = 0$.

The expression of the free energy functional in replica space can be derived following exactly the same steps as in the previous section. The final result is [23]

$$-\beta n \frac{F_{\rm rep}[\rho]}{N} = \frac{k+1}{2} \times \ln\left(\mathrm{Tr}_{\sigma,\tau}\left[\rho(\sigma)^k\rho(\tau)^k \exp\left(\sum_{a=1}^{n}\beta J\sigma_a\tau_a\right)\right]\right) - k\ln\left(\mathrm{Tr}_\sigma\left[\rho(\sigma)^{k+1}\right]\right) \quad (25)$$

where Tr_σ denotes the average over the 2^n configurations of the variables σ or τ, and the correct result is obtained in the $n \to 0$ limit. The same value for the free energy is obtained if we multiply the function ρ by a constant, so that the ρ does not need to be normalized, although it is more convenient to work with a normalized ρ. The advantage of the replica approach is that the system is homogeneous and the distribution $\rho(\sigma)$ is the same in all the points. This advantage is partially compensated by the fact the the number of variables n is going to zero.

As can be readily checked, ρ satisfies a very simple equation:

$$\rho(\sigma) = \frac{E_J \mathrm{Tr}_\tau\left[\rho(\tau)^k \exp\left(\sum_{a=1}^{n}\beta J\sigma_a\tau_a\right)\right]}{\mathrm{Tr}_\tau\left[\rho(\tau)^k\right]}\,. \quad (26)$$

Using this equation the free energy (25) can be simplified to [23]:

$$-\beta n\frac{F_{\rm rep}[\rho]}{N} = \frac{k+1}{2}\ln\left(\mathrm{Tr}_\sigma\left[\rho(\sigma)^k\right]\right) - \frac{k-1}{2}\ln\left(\mathrm{Tr}_\sigma\left[\rho(\sigma)^{k+1}\right]\right) \quad (27)$$

where as before this new form of the free energy cannot be used in a variational formulation. The result (25) for the replicated free energy is correct in general, whether the replica symmetry is broken or not. The problem is to find the solution $\rho^*(\sigma)$. In the replica symmetric situation this task is easy: the $\rho(\sigma)$ is a function of only $\Sigma = \sum_a \sigma_a$. We can thus write in general:

$$\rho(\sigma) = \int \mathrm{d}u\; R(u)\exp(\beta u\Sigma), \quad (28)$$

where the normalization condition of $\rho(\sigma)$ imposes:

$$\int \mathrm{d}u\; R(u)\left(2\cosh(\beta u)\right)^n = 1. \quad (29)$$

Using this expression for ρ, we obtain in the small n limit:

$$\ln\left(\mathrm{Tr}_\sigma\left[\rho(\sigma)^{k+1}\right]\right) = n\int\prod_{i=1}^{k+1}\left[\mathrm{d}u_i R(u_i)\right] \times \ln\left(\left[\prod_{i=1}^{k+1}\frac{1}{2\cosh(\beta u_i)}\right]\sum_{\sigma_0}\exp\left(\beta\sigma_0\sum_i u_i\right)\right) \quad (30)$$

and:

$$\ln\left(\mathrm{Tr}_{\sigma,\tau}\left[\rho(\sigma)^k\rho(\tau)^k\exp(\sum_{a=1,n}\beta J\sigma_a\tau_b)\right]\right) = n\int\prod_{i=1}^{k}\left[\mathrm{d}u_i\mathrm{d}v_i R(u_i)R(v_i)\right]\ln\left(\left[\prod_{i=1}^{k}\frac{1}{4\cosh(\beta u_i)\cosh(\beta v_i)}\right] \times\sum_{\sigma_0,\tau_0}\exp\left(\beta\sigma_0\sum_{i=1}^{k}u_i+\beta\tau_0\sum_{i=1}^{k}v_i+\beta J_0\sigma_0\tau_0\right)\right)\,. \quad (31)$$

Putting these expressions back into the replica free energy (25), one gets exactly the functional $F[Q]$ which we had written previously in (18), provided we identify the $n \to 0$ limit of $R(u)$ with the probability distribution (7) of the variable $u(J,h)$, and we use in (19) a function $d(J,h) = c(J,h)[2\cosh(\beta u(J,h))]$.

In other words we have seen three equivalent ways to solve the Bethe lattice spin glass in the replica symmetric approximation:

- One can derive the recursion equations (6) for the probability distribution $Q^*(h)$ of the local 'cavity' field h, and evaluate the free energy and the internal energy using this distribution.
- Alternatively one can introduce the free energy functional(13), which depends on the probability distribution $Q(h)$ and satisfies a variational principle: the distribution $Q^*(h)$ is obtained as the one which makes the functional stationary.
- One can obtain the same functional starting from the replica approach (25), making explicitly an assumption of replica symmetry, and doing some simple algebra. It is typical of the replica approach that a probability distribution is traded with a function of n variables, in the $n \to 0$ limit.

3.4 Free energy shifts

It may be instructive to compare this approach with the more usual cavity method and to check that we obtain the same results. In the cavity method, one computes the free energy by averaging the various free energy shifts obtained when adding a new site or a new bond to the lattice.

The first quantity which we compute is the free energy shift $\Delta F_{\rm iter}$ obtained my adding a new spin σ_0 connected to k branches, as we do in the iterative procedure. Using

the same notations as in (5), this free energy shift is:

$$-\beta\Delta F_{\rm iter}(J_1\dots J_k,h_1\dots h_k)= \ln\left[2\cosh\left(\beta\sum_{i=1}^{k}u(J_i,h_i)\right)\right]+\sum_{i=1}^{k}\ln\left[\frac{\cosh(\beta J_i)}{\cosh(\beta u(J_i,h_i))}\right]. \quad (32)$$

In order to compute the total free energy, we also need the free energy shift when adding the new spin σ_0, onto which merge $k+1$ branches (see Fig. 2). This free energy shift is equal to the same quantity with k changed into $k+1$:

$$\begin{aligned}-\beta\Delta F^{(1)}(J_1\dots J_{k+1},h_1\dots h_{k+1}) &= \\ -\beta F^{(1)}(h_1\dots h_{k+1})+\beta\sum_{i=1}^{k+1}\ln[2\cosh(\beta h_i)] &\\ =\ln\left[2\cosh\left(\beta\sum_{i=1}^{k+1}u(J_i,h_i)\right)\right]+\sum_{i=1}^{k+1}\ln\left[\frac{\cosh(\beta J_i)}{\cosh(\beta u(J_i,h_i))}\right]. & \quad (33)\end{aligned}$$

The free energy shift when adding the two new spins σ_0,τ_0 (see Fig. 2) is equal to:

$$\Delta F^{(2)}(J_1\dots J_k,K_1\dots K_k,h_1\dots h_k,g_1\dots g_k)= F^{(2)}(h_1\dots h_k,g_1\dots g_k)-\sum_{i=1}^{k}\ln[4\cosh(\beta h_i)\cosh(\beta g_i)] \quad (34)$$

and is given by:

$$\begin{aligned}-\beta\Delta F^{(2)}(J_0,J_1\dots J_k,K_1\dots K_k,h_1\dots h_k,g_1\dots g_k) &= \\ \sum_{i=1}^{k}\ln\left[\frac{\cosh(\beta J_i)}{\cosh(\beta u(J_i,h_i))}\frac{\cosh(\beta K_i)}{\cosh(\beta u(K_i,g_i))}\right] &\\ +\ln\left[\sum_{\sigma_0,\tau_0}\exp\left(\beta J_0\sigma_0\tau_0+\beta\sigma_0\sum_{i=1}^{k}u(J_i,h_i)\right.\right. &\\ \left.\left.+\beta\tau_0\sum_{i=1}^{k}u(K_i,g_i)\right)\right]. & \quad (35)\end{aligned}$$

In the process of adding new sites or new bonds randomly, one can thus compute the total free energy as the average over the distribution of fields and couplings of $[(k+1)/2]\ \Delta F^{(2)}-k\Delta F^{(1)}$. It is a simple exercise to check that this indeed gives back the free energy (18).

4 The RSB instability

The recursion relation of the local fields (5) has been the subject of a lot of studies in the past twenty years [13–18,25,31]. The distribution $Q(h)$ is a simple δ function at the origin, indicating a paramagnetic phase, at high temperatures $\beta<\beta_{\rm c}$, where the critical inverse temperature $\beta_{\rm c}$ is the solution of [17]:

$$E_J\tanh^2(\beta_{\rm c}J)=1/k\ . \quad (36)$$

In the low temperature phase the specific heat becomes negative at low enough temperatures at least for some distributions of couplings [16], and the solution for $Q(h)$ becomes identical to the replica symmetric field distribution of the SK model in the large k limit [14,16,18], which is known to be wrong. Another indication that the above procedure gives a wrong result for the Bethe lattice (while it might be correct for the Cayley tree with some sets of boundary conditions [18,21]) is the fact that it fails to identify a transition in a magnetic field H. This transition exists though, on a line in the H-T plane similar to the A-T line, and can be identified by considering the onset of correlations between two replicas of the system [17].

One can investigate the instability of the previous solution using the replica method. Writing the recursion relations for the replicated system, Mottishaw has shown that the replica symmetric solution, which coincides with the simple Bethe-Peierls iteration described above, is unstable at $\beta>\beta_{\rm c}$ (or, in a field, beyond the A-T line) [19]. Unfortunately, getting the replica symmetry broken solution in the low temperature phase is difficult. In general the problem involves an infinity of order parameters which are multi-spin overlaps [9,19,20,23,24]. As we saw, the replica symmetric solution already involves an order parameter which is a whole function (the distribution of local fields); going to a 'one step RSB' solution [3], the replica order parameter becomes now a functional, the probability distribution over the space of local field distributions [27] (the reason will be discussed in details in the next section). While one can write formally some integral equations satisfied by this order parameter, solving them is in general a formidable task. The solution is known only in the neighborhood of the critical temperature [20], or in the limit of large connectivities [23], where the overlaps involving three spins or more are small and can be treated perturbatively, allowing for some expansion around the SK solution. In the general case, the only tractable method so far has been an approximation which parametrizes the functional by a small enough number of parameters and optimizes the one step RSB free energy inside this subspace [32,33]. We shall develop in the next two sections a solution to this problem.

5 The formulation of the 'one step RSB' solution

5.1 The iterative approach

In this section we shall explain the physical nature of the RSB instability and work out the equivalent of the 'one step RSB solution' using the cavity method [4,3], *i.e.* the same type of iterative approach which was used in Section 3.1.

The reason for the failure of the simple iterative solution is that it neglects the possibility of the existence of several pure states [3]. We shall proceed by first assuming some properties of the states on one branch, and then imposing the self-consistency of these hypotheses when one joins k branches to a new site. Let us assume that there exist many pure states, labelled by an index α going from 1 to ∞, with the following properties: Looking at one branch of the tree as in fig. 1, the total local field seen by the site $i = 0$ at the extremity of this branch depends on the state α and is denoted by h_0^α; the free energies of the states on one branch, F^α, are independent identically distributed (iid) random variables, with an exponential density behaving as

$$\rho(F) = \exp(\beta x(F - F^{\mathrm{R}})), \tag{37}$$

where F^{R} is a reference free energy. This assumption of iid exponentially distributed free energies is one of the basic ingredients of the one step rsb solutions to spin glasses [3]. The differences between the free energies remain finite when the volume goes to infinity, which means that the various states have non-zero statistical weights:

$$W^\alpha = \frac{\exp(-\beta F^\alpha)}{\sum_\gamma \exp(-\beta F^\gamma)}, \tag{38}$$

The fact that the W can be normalized in this way is possible only if $x < 1$.

Let us consider as before a point i on one branch of the Bethe lattice, *i.e.* a point connected to k other points. In each phase α of the system the magnetization m_i^α will be different and therefore the effective fields h_i^α depend on α. The description of the properties of this point will include the list of fields h_i^α and the free energies of the branch F^α, for all states α. Here we shall assume a relatively simple situation namely that the free energies and the magnetic fields are not correlated, and the distribution of free energies is the one described above, leading to the density (37). It is convenient (to avoid possible difficulties in dealing with measures in infinite dimensional spaces) to order the states in an increasing order of free energy, and to consider only the set of the first $\mathcal{M}$ states with lowest free energies (in the end $\mathcal{M}$ will be sent to infinity). Let us introduce the set of the local fields in all the $\mathcal{M}$ states, $\mathbf{h} = \{h_i^\alpha\}$. When changing the sample (or equivalently changing the site i), these fields fluctuate and our task is to compute the corresponding probability distribution $Q(\mathbf{h})$, which is a function of the $\mathcal{M}$ effective magnetic fields which is left invariant by the permutations of these fields. This task is simplified if we assume that the $\mathcal{M}$ fields h_i^α on one point can be characterized as independent random variables. We thus assume that there exists a probability function $Q_i(\mathbf{h})$ which can be written in a factorized form:

$$Q_i(\mathbf{h}) = \prod_{\alpha=1}^{\mathcal{M}} Q_i(h_\alpha). \tag{39}$$

The total probability function $Q(\mathbf{h})$ is thus given by

$$Q(\mathbf{h}) = N^{-1} \sum_{i=1}^{N} \left[\prod_{\alpha=1}^{\mathcal{M}} Q_i(h_\alpha) \right]. \tag{40}$$

In other words on any given point the fields are independent variables, which become correlated on the global level after we average over the samples. More generally one can assume that the total probability function $Q(\mathbf{h})$ is of the form:

$$Q(\mathbf{h}) = \int \mathrm{d}\lambda \; m(\lambda) \prod_{\alpha=1}^{M} q(h^\alpha|\lambda), \tag{41}$$

where λ is an appropriate set, $m(\lambda)$ is a probability distribution, and $q(h|\lambda)$ is a probability distribution on h, conditioned to a given value of λ. A possible representation of the distribution (41) is given by (40), where each point i is characterized by a parameter λ (extracted with the measure $m(\lambda)$).

We shall now check that this hypothesis is self consistent, *i.e.* that it is reproduced when one iterates the construction of the tree by merging k lines to a new spin σ_0 [35]. For each state α, the local field on this site h_0^α is expressed in terms of those on the branches, h_i^α by (5), giving:

$$h_0^\alpha = \sum_{i=1}^{k} u(J_i, h_i^\alpha). \tag{42}$$

The free energy shift ΔF^α for the state α during this process is given by the function ΔF_{iter} defined in (32):

$$\Delta F^\alpha = \Delta F_{\mathrm{iter}}(J_1 \dots J_k, h_1^\alpha, \dots, h_k^\alpha). \tag{43}$$

We must be careful at this stage because the free energy shifts and the local fields on the new spin σ_0 are correlated. More precisely, for a given state α, h_0^α and ΔF^α are two correlated variables, but they are not correlated with the local fields or free-energy shifts in the other states. Because of our ordering process of the free energies, we need to compute and order the new free energies $G^\alpha = F^\alpha + \Delta F^\alpha$. G^α and G^γ in two different states are obviously independent random variables. Furthermore, a standard argument of the cavity method [4,3], relying on the exponential distribution of the free energies, allows to show that the new free energy G^α is in fact *uncorrelated* with the local field h_0^α. To show this, let us introduce the joint distribution $P_0(h_0, \Delta F)$ of h_0^α and ΔF^α. The joint distribution $R_0(h_0, G)$ of the local field and the new free energy is given by:

$$R_0(h_0, G) \propto \int \mathrm{d}F \mathrm{d}(\Delta F) \exp(\beta x(F - F^{\mathrm{R}})) P_0(h_0, \Delta F) \times \delta(G - F - \Delta F) \propto \exp(\beta x(G - F^{\mathrm{R}})) Q_0(h_0), \tag{44}$$

where

$$Q_0(h_0) = C \int \mathrm{d}(\Delta F) \; P_0(h_0, \Delta F) \exp(-\beta x \Delta F) \tag{45}$$

the constant C being fixed in such a way that $Q_0(h_0)$ is a normalized probability distribution.

In our ordering process of the new free energies G_α we pick up the $\mathcal{M}$ lowest ones, sending in the end $\mathcal{M}$ to infinity. This ordering process thus gives rise to a probability distribution for the h_0^α in which the fields for different α are not correlated and have the distribution $Q_0(h_0)$. The reader should notice that this distribution is in general different from the naive result $\int \mathrm{d}\Delta F_0\ P_0(h, \Delta F_0)$. In this way we have constructed, for one given new spin σ_0 with a fixed environment of coupling constants, the new distribution of all local fields: $\prod_\alpha Q_0(h_0^\alpha)$. By averaging over the coupling constants, one thus generates the probability distribution $Q_0(\mathbf{h})$ which is a functional of the probability distribution $Q(\mathbf{h})$ of the other k sites. Imposing that

$$Q_0(\mathbf{h}) = Q(\mathbf{h}) \tag{46}$$

gives a self-consistency equation for the probability distribution $Q(\mathbf{h})$. We shall see in Section 6 how one can actually find a solution to this self-consistency equation with a good accuracy.

Let us suppose for the time being that we know the self-consistent distribution $Q(\mathbf{h})$ and let us evaluate the free energy and the internal energy. The computation is quite similar to the one in the replica symmetric case. There are two contributions to the free energy: the site contribution and the bond contribution.

The local site contribution to the free energy is evaluated as a weighted average of the free energy shift when adding one new spin:

$$F^{(1)} = -\frac{1}{\beta} E_J \left\langle \ln \left(\sum_{\alpha=1}^{\mathcal{M}} W^\alpha \exp[-\beta \Delta F_\alpha^{(1)}] \right) \right\rangle, \tag{47}$$

where $\Delta F_\alpha^{(1)} = \Delta F^{(1)}(J_1 \dots J_{k+1}, h_1^\alpha \dots h_{k+1}^\alpha)$ is the site free-energy-shift in state α computed from (33), and the bracket denotes an average over the distribution of weights W^α derived from (37, 38), and over the fields (with distribution $\prod_{i=1}^{k+1} Q(\mathbf{h}_i)$).

The local bond contribution to the free energy is evaluated as a weighted average of the free energy shift when adding a new bond and the corresponding two spins:

$$F^{(2)} = -\frac{1}{\beta} E_J E_K \left\langle \ln \left(\sum_{\alpha=1}^{\mathcal{M}} W^\alpha \exp[-\beta \Delta F_\alpha^{(2)}] \right) \right\rangle, \tag{48}$$

where

$$\Delta F_\alpha^{(2)} = \Delta F^{(2)}(J_1 \dots J_k, K_1 \dots K_k, h_1^\alpha \dots h_k^\alpha, g_1^\alpha \dots g_k^\alpha)$$

is the bond free-energy-shift in state α computed from (35), and the bracket denotes an average over the weights, and over the fields (with distribution $\prod_{i=1}^{k} [Q(\mathbf{h}_i) Q(\mathbf{g}_i)]$).

The total free energy is given, according to (13–15), by:

$$F = \frac{k+1}{2} F^{(2)} - k F^{(1)}. \tag{49}$$

Similarly to what happened in the RS case (23), one can also write here a simplified form of the free energy, valid only on the saddle point:

$$F = \frac{k+1}{2} F^{(1')} - \frac{k-1}{2} F^{(1)} \tag{50}$$

where

$$F^{(1')} = -\frac{1}{\beta} E_J \left\langle \ln \left(\sum_{\alpha=1}^{\mathcal{M}} W^\alpha \exp[-\beta \Delta F^{(1)}(J_1 \dots J_k, h_1^\alpha \dots h_k^\alpha)] \right) \right\rangle. \tag{51}$$

For reasons which are beyond our control this second form of the free energy has empirically smaller errors (about by a factor 3) and less systematic errors (at finite $\mathcal{M}$) than the first one.

The computation of the internal energy is done by considering what happens when we couple two sites which where previously connected each to k branches of the tree. We obtain, with the same notations as in formula (18) and in Figure 2:

$$U = -\left\langle \sum_{\alpha=1}^{\mathcal{M}} \tilde{W}_2^\alpha J_0 \frac{\tanh(\beta J_0) + \tanh(\beta h_0^\alpha) \tanh(\beta g_0^\alpha)}{1 + \tanh(\beta J_0) \tanh(\beta h_0^\alpha) \tanh(\beta g_0^\alpha)} \right\rangle \tag{52}$$

where $h_0^\alpha = \sum_{i=1}^k u(J_i, h_i^\alpha)$, $g_0^\alpha = \sum_{i=1}^k u(K_i, g_i^\alpha)$ and we have introduced the shorthand notation:

$$\tilde{W}_2^\alpha \equiv \frac{W^\alpha \exp[-\beta \Delta F_\alpha^{(2)}]}{\sum_{\gamma=1}^{\mathcal{M}} W^\gamma \exp[-\beta \Delta F_\gamma^{(2)}]}. \tag{53}$$

The various overlaps can be obtained in the same way. There are now two site overlaps, the self-overlap q_1 and the inter-state-overlap q_0, which are given by:

$$q_1 = E_J \left\langle \sum_{\alpha=1}^{\mathcal{M}} \tilde{W}_1^\alpha \tanh^2[\beta \sum_{i=1}^{k+1} u(J_i, h_i^\alpha)] \right\rangle$$

$$q_0 = E_J \left\langle \sum_{\alpha \neq \beta} \tilde{W}_1^\alpha \tilde{W}_1^\beta \tanh[\beta \sum_{i=1}^{k+1} u(J_i, h_i^\alpha)] \tanh[\beta \sum_{i=1}^{k+1} u(J_i, h_i^\beta)] \right\rangle \tag{54}$$

where we have kept the same notations as in (47) but we have introduced $\tilde{W}_1^\alpha$ which is a notation for:

$$\tilde{W}_1^\alpha \equiv \frac{\exp[-\beta F^\alpha - \beta \Delta F^{(1)}(J_1 \dots J_{k+1}, h_1^\alpha \dots h_{k+1}^\alpha)]}{\sum_\gamma \exp[-\beta F^\gamma - \beta \Delta F^{(1)}(J_1 \dots J_{k+1}, h_1^\gamma \dots h_{k+1}^\gamma)]}. \tag{55}$$

Similarly, we have two link overlaps $q_1^{(l)}$ and $q_0^{(l)}$ which are given by:

$$q_1^{(l)} = E_J \left\langle \sum_{\alpha=1}^{\mathcal{M}} \tilde{W}_2^\alpha \left(\frac{\tanh(\beta J_0) + \tanh(\beta h_0^\alpha) \tanh(\beta g_0^\alpha)}{1 + \tanh(\beta J_0) \tanh(\beta h_0^\alpha) \tanh(\beta g_0^\alpha)} \right)^2 \right\rangle \tag{56}$$

and

$$q_0^{(l)} = E_J \left\langle \sum_{\alpha\neq\beta} \tilde{W}_2^\alpha \tilde{W}_2^\beta \left(\frac{\tanh(\beta J_0) + \tanh(\beta h_0^\alpha)\tanh(\beta g_0^\alpha)}{1+\tanh(\beta J_0)\tanh(\beta h_0^\alpha)\tanh(\beta g_0^\alpha)} \right) \times \left(\frac{\tanh(\beta J_0) + \tanh(\beta h_0^\beta)\tanh(\beta g_0^\beta)}{1+\tanh(\beta J_0)\tanh(\beta h_0^\beta)\tanh(\beta g_0^\beta)} \right) \right\rangle \quad (57)$$

where we keep the same notations as in (52, 53).

At this stage we have written the hole formalism of the cavity method at the level of one step RSB. The self-consistency equation (46) fixes the distribution $Q(\mathbf{h})$, from which one can deduce the free energy and internal energy through (49) and (52). One can actually find a self-consistent solution for any value of the parameter $x \in [0,1]$, and the free energy depends on x through the distribution of free energies. It is well known that, in order to describe the thermal equilibrium, one must fix x by maximising the free energy with respect to x [3,36,38]. (Actually the whole dependence on x carries some information [36–38], particularly interesting for optimization problems, which we shall not try to study here since this paper is restricted to the study of equilibrium thermodynamics.) As we shall see, the variation of free energy with respect to x is small and one needs a better computation of the derivative than just a naive difference.

We have improved the precision on the computation of the free energy and its x derivative by using the theorem of Appendix A, which allows to compute explicitly the derivative of the free energy with respect to x. From the structure of equation (49) one finds that the total derivative $d(x) = \mathrm{d}F/\mathrm{d}x$ takes the form:

$$d(x) = -\frac{1}{x}F - kd^{(1)} + \frac{k+1}{2}d^{(2)}, \quad (58)$$

where

$$d^{(1)} = \frac{1}{x} E_J \left\langle \frac{\sum_{\alpha=1}^{\mathcal{M}} \exp[-\beta \Delta F_\alpha^{(1)}] \Delta F_\alpha^{(1)}}{\sum_{\alpha=1}^{\mathcal{M}} \exp[-\beta \Delta F_\alpha^{(1)}]} \right\rangle, \quad (59)$$

using the notations of (47), and

$$d^{(2)} = \frac{1}{x} E_J E_K \left\langle \frac{\sum_{\alpha=1}^{\mathcal{M}} \exp[-\beta \Delta F_\alpha^{(2)}] \Delta F_\alpha^{(2)}}{\sum_{\alpha=1}^{\mathcal{M}} \exp[-\beta \Delta F_\alpha^{(2)}]} \right\rangle, \quad (60)$$

using the notations of (48).

5.2 A variational formulation

As in the case where no replica symmetry breaking is present we can write a free energy functional of the field distribution $Q(\mathbf{h})$ such that the self-consistency equations for Q are equivalent to the stationarity condition of this functional.

This free energy functional is a simple generalisation of the replica symmetric one, given by:

$$\frac{F[Q]}{N} = \frac{k+1}{2} \int \prod_{i=1}^{k} [\mathrm{d}\mathbf{h}_i d\mathbf{g}_i Q(\mathbf{h}_i) Q(\mathbf{g}_i)] \; F^{(2)}(\mathbf{h}_1 \dots \mathbf{h}_k, \mathbf{g}_1 \dots \mathbf{g}_k) - k \int \prod_{i=1}^{k+1} [\mathrm{d}\mathbf{h}_i Q(\mathbf{h}_i)] \; F^{(1)}(\mathbf{h}_1 \dots \mathbf{h}_{k+1}) \quad (61)$$

where

$$-\beta F^{(1)}(\mathbf{h}_1 \dots \mathbf{h}_{k+1}) = E_J \left\langle \ln \left(\sum_\alpha W^\alpha \left[\prod_{i=1}^{k+1} \frac{1}{2\cosh(\beta h_i^\alpha)} \right] \times \sum_{\sigma_0, \sigma_1, \dots \sigma_{k+1}} \exp\left[\beta \sigma_0 \sum_{i=1}^{k+1} J_i \sigma_i + \beta \sum_{i=1}^{k+1} h_i^\alpha \sigma_i \right] \right) \right\rangle, \quad (62)$$

$$-\beta F^{(2)}(\mathbf{h}_1 \dots \mathbf{h}_k, \mathbf{g}_1 \dots \mathbf{g}_k) = E_J E_K \left\langle \ln \left(\sum_\alpha W^\alpha \times \left[\prod_{i=1}^{k} \frac{1}{4\cosh(\beta h_i^\alpha)\cosh(\beta g_i^\alpha)} \right] \sum_{\sigma_0, \sigma_1, \dots \sigma_k} \sum_{\tau_0, \tau_1, \dots \tau_k} \times \exp\left[\beta J_0 \sigma_0 \tau_0 + \beta \sigma_0 \sum_i J_i \sigma_i + \beta \sum_i h_i^\alpha \sigma_i + \beta \tau_0 \sum_i K_i \tau_i + \beta \sum_i g_i^\alpha \tau_i \right] \right) \right\rangle. \quad (63)$$

The weights W_α are given by (38), and the brackets stand for an average over the distribution of free energies (37).

The proof of the equivalence of the stationarity equation of this functional with the self-consistency condition of the iterative procedure can be done exactly as in the replica symmetric case. The advantage of this variational formulation is that we can compute directly the various derivatives of the free energy (*e.g.* with respect to x and with respect to β) by taking into account only the explicit dependence.

5.3 Equivalence with the replica formalism

We can compare what we have obtained in the previous sections with the results from the replica formalism. In the one step RSB formalism the n replicas are divided into n/x groups (labeled by $\mathcal{C}$) of x replicas each, and the function $\rho(\sigma)$ depends on the n/x 'block' variables

$$\Sigma_{\mathcal{C}} = \sum_{a \in \mathcal{C}} \sigma_a, \quad (64)$$

each sum containing x terms.

In this case we face the problem that $\rho(\sigma)$ may depend on the n/x variables $\Sigma_{\mathcal{C}}$ in a rather complex way. Similarly to what we have done in the iterative approach, we shall not try to describe the most general dependence, but we restrict to the class of probability distributions $\rho(\sigma)$ which can be written as:

$$\rho(\sigma) = \int \mathrm{d}\lambda \mu(\lambda) \int \prod_{\mathcal{C}=1}^{n/x} [\mathrm{d}u_{\mathcal{C}} \phi(u_{\mathcal{C}}|\lambda)] \exp\left(\beta \sum_{\mathcal{C}=1}^{n/x} u_{\mathcal{C}} \Sigma_{\mathcal{C}}\right). \tag{65}$$

with a positive probability distribution $\mu(\lambda)$ [32] and a positive function $\phi(u|\lambda)$. The normalisation condition on $\rho(\sigma)$ is implemented by imposing that:

$$\forall \lambda : \quad \int \mathrm{d}u \phi(u|\lambda)(2\cosh(\beta u))^x = 1, \tag{66}$$

so that the function

$$\Phi(u|\lambda) = \phi(u|\lambda)(2\cosh(\beta u))^x \tag{67}$$

is a probability distribution on the u variable, for any value of λ.

It is easy to check that this form for the function is consistent with the stationarity equations for the free energy (26) by proving that, if $\rho(\sigma)$ has the form (65), so does

$$E_J \mathrm{Tr}_\tau \left(\rho(\tau)^k \exp[\sum_{a=1}^{n} \beta J \sigma_a \tau_a]\right). \tag{68}$$

Using the Ansatz (65) for ρ, one can write the 'site' term in the replica free energy (25) as:

$$\mathrm{Tr}_\sigma \left(\rho(\sigma)^{k+1}\right) = \int \left[\prod_{i=1}^{k+1} \mathrm{d}\lambda_i \mu(\lambda_i)\right] A\left(\lambda_1, \ldots, \lambda_{k+1}\right)^{n/x}, \tag{69}$$

where:

$$A\left(\lambda_1, \ldots, \lambda_{k+1}\right) = \int \prod_{i=1}^{k+1} [\mathrm{d}u_i \Phi(u_i|\lambda_i)] \left[\frac{2\cosh(\beta \sum_{i=1}^{k+1} u_i)}{\prod_i [2\cosh \beta u_i]}\right]^x. \tag{70}$$

Using the small theorem proven in Appendix A, the previous expression can be written as

$$\ln A\left(\lambda_1, \ldots, \lambda_{k+1}\right) = x \int \prod_{i=1}^{k+1} \prod_{\alpha} [\mathrm{d}u_i^\alpha \Phi(u_i^\alpha|\lambda_i)] \times \left\langle \left[\ln\left(\sum_\alpha W_\alpha \frac{2\cosh(\beta \sum_{i=1}^{k+1} u_i^\alpha)}{\prod_i [2\cosh \beta u_i^\alpha]}\right)\right]\right\rangle, \tag{71}$$

where the bracket means an average over the distribution of the weights W_α with the usual distribution parametrized by the parameter x (see Appendix A). One gets in the end:

$$\ln \mathrm{Tr}_\sigma \left(\rho(\sigma)^{k+1}\right) = n \int \prod_{i=1}^{k+1} [\mathrm{d}\lambda_i \mu(\lambda_i)] \times \int \prod_{i=1}^{k+1} \prod_{\alpha} [\mathrm{d}u_i^\alpha \Phi(u_i^\alpha|\lambda_i)] \left\langle \ln\left(\sum_\alpha W_\alpha \frac{2\cosh(\beta \sum_{i=1}^{k+1} u_i^\alpha)}{\prod_i [2\cosh \beta u_i^\alpha]}\right)\right\rangle. \tag{72}$$

The previous quantity is exactly the expression of the 'site contribution' to the variational free energy found in (62), provided we identify the probability distributions defined in the iterative approach (41) and in the replica approach (65):

$$\mu(\lambda) = m(\lambda); \quad \Phi(u|\lambda) = E_J \int \mathrm{d}h \; q(h|\lambda) \delta(u - u(J,h)). \tag{73}$$

A similar computation shows that the 'bond' term in the free energy, calculated either through the iterative procedure or through the replica method also coincide.

We have thus derived the variational equations for the probability distribution of local fields in the one step RSB case using two different methods, the cavity iterative approach on the one hand, and the algebraic replica formalism on the other hand.

6 Solving the one step RSB equations

Our method consists in following the population of local fields h_i^α when one iterates the merging process of k branches onto one site. In some sense it thus amounts to solving the complicated equation for the functional order parameter by a method of population dynamics. In other words we parametrize the probability distribution by presenting a large number of instances of variables distributed according to this probability distribution. A similar method has been first used in the context of mean field equations in [39].

To explain it in more details, let us first state how the procedure works in the case of the 'replica symmetric' approximation of Section 3. There, one just chooses a population of $\mathcal{N}$ local fields h_i. At each iteration, one picks up k such fields at random among the $\mathcal{N}$, and computes the new field h_0 according to (5). Then one field is removed at random from the population and is substituted by h_0.

In this way one defines a Markov chain on the space of the $\mathcal{N}$ magnetic fields. This chain has a stationary distribution which is reached after some transient time. In the $\mathcal{N} \to \infty$ limit, the stationary distribution satisfies the self-consistency equation (6). It is possible to argue that the corrections to this limit are proportional to $\mathcal{N}^{-1}$ and could also be computed analytically. Our procedure consists in fixing the value of $\mathcal{N}$, iterating the merging transformation many times in such a way as to obtain the average over the asymptotic distribution at fixed $\mathcal{N}$ with

a high precision, and finally extrapolating the results to $\mathcal{N} \to \infty$.

If we consider the case where there exist many states, we have the same problem as before, with the only difference that at each point we have a probability distribution $Q_i(\mathbf{h})$. We must therefore consider a population dynamics in which the elements of the population are probability distributions. We use the population method to represent the probability distribution in each point i by a populations of fields. In this way we have a population of $\mathcal{N}$ populations of $\mathcal{M}$ fields (a total of $\mathcal{N}\mathcal{M}$ fields h_i^α, $i \in \{1, ..., \mathcal{N}\}$, $\alpha \in \{1, ..., \mathcal{M}\}$), where both $\mathcal{N}$ and $\mathcal{M}$ have to go to infinity.

The basic step of the algorithm is the merging of k lines. One chooses k sites $i_1, ..., i_k$ in $\{1, ..., \mathcal{N}\}$, and one generates, for each of the $\mathcal{M}$ states, a new local field h_0^α obtained by merging k branches, using (5), as well as the corresponding free-energy shift ΔF^α calculated using (32). But the field h_0^α is not the one which will enter the population of fields. The reason is that one needs to reweigh the various states by a factor which depends on their free energy shifts: as seen in (45), the field distribution, at a fixed new free energy, is modified by a factor $\exp(-\beta x \Delta F)$. From the knowledge of the h_0^α one can infer an approximate form of the distribution $P_0(h_0)$ from which they are extracted. For instance a simple form for P_0 is a smoothly interpolated version of the identity

$$\int_{-\infty}^{h} P_0(h_0) \mathrm{d}h_0 = \frac{1}{\mathcal{M}} \sum_\alpha \theta(h_0 - h_0^\alpha) \tag{74}$$

where $\theta(x)$ is Heaviside's function (in practice we smooth this staircase function by a linear interpolation procedure). According to (45), the real field distribution $Q_0(h)$ is well approximated by a smoothly interpolated version of the identity

$$\int_{-\infty}^{h} Q_0(h) \mathrm{d}h = \frac{1}{\mathcal{M}} \sum_\alpha \exp(-\beta x \Delta F^\alpha) \theta(h - h_0^\alpha). \tag{75}$$

We can use two different methods in order to achieve the reweighing, which will lead to two different algorithms.

- Method A: The idea is to generate, from the set of $\mathcal{M}$ fields $h_{i_l}^\alpha$, on each of the sites $i_1, ..., i_k$ which is used in the merging, a *larger* population, of $r\mathcal{M}$ local fields $h_{i_l}^{\alpha'}$, $\alpha' = 1, ..., r\mathcal{M}$, taken from the same distribution. This will be realised by having

$$\frac{1}{\mathcal{M}} \sum_{\alpha=1}^{\mathcal{M}} \theta(h - h_{i_l}^\alpha) \simeq \frac{1}{r\mathcal{M}} \sum_{\alpha'=1}^{r\mathcal{M}} \theta(h - h_{i_l}^{\alpha'}), \tag{76}$$

  at the level of linearly interpolated functions. Simultaneously one generates $r\mathcal{M}$ independent random free energies $F^{\alpha'}$, $\alpha' = 1, ..., r\mathcal{M}$, with the exponential density (37). For each of the $r\mathcal{M}$ states, one then compute the local field $h_0^{\alpha'}$ and the free energy shift $\Delta F^{\alpha'}$. The correct reweighing is obtained by the selection of low lying states: one computes the new $r\mathcal{M}$ free energies $F^{\alpha'} + \Delta F^{\alpha'}$, orders them, and keeps only the $\mathcal{M}$ states with the lowest new free energies. Their local fields, called h^α, with $\alpha \in \{1, \cdots, \mathcal{M}\}$, have the correctly reweighed distribution provided that r is large enough so that there is a zero probability for the states α' with the highest free energy $F^{\alpha'}$ to enter the list of the $\mathcal{M}$ selected states after reweighing.
- Method B: The idea is to generate directly the fields with the reweighed distribution (75). Knowing the $\mathcal{M}$ local fields h_0^α and their free-energy shifts ΔF^α, we generate $\mathcal{M}$ new local fields h^α in such a way that the following identity holds at the level of linear interpolation:

$$\frac{1}{\sum_\alpha \exp(-\beta x \Delta F^\alpha)} \sum_\alpha \exp(-\beta x \Delta F^\alpha) \theta(h - h_0^\alpha) \simeq \frac{1}{\mathcal{M}} \sum_\alpha \theta(h - h^\alpha). \tag{77}$$

Having generated the new fields h^α which are typical of the properly reweighed distribution, we then substitute in the population the set of $\mathcal{M}$ local fields h_i^α, $\alpha = 1, .., \mathcal{M}$ by the set of new fields h^α. The site index i on which this substitution is performed is chosen sequentially.

While the merging of k lines is enough to build up the Markov chain which generates the population of local fields, one also needs to consider some different merging processes in order to compute the various observables, free energy, energy, local overlap and link overlap.

By merging $k+1$ lines instead of k, one generates with exactly the same procedure as above the three sets of $\mathcal{M}$ local fields h_0^α, free-energy shifts ΔF^α and new fields h^α (we call new fields the fields which are typical of the reweighed distribution, obtained either through procedure A or B). Using the little theorem of Appendix A, the site contribution (47) to the free energy is computed as

$$F^{(1)} = -\frac{1}{\beta x} \ln \left[\frac{1}{\mathcal{M}} \sum_\alpha \exp(-\beta x \Delta F^\alpha) \right], \tag{78}$$

and the site overlaps receive a contribution

$$q_1 = \frac{1}{\mathcal{M}} \sum_\alpha \tanh^2(\beta h^\alpha);$$
$$q_0 = \frac{1}{\mathcal{M}(\mathcal{M}-1)} \sum_{\alpha \neq \gamma} \tanh(\beta h^\alpha) \tanh(\beta h^\gamma). \tag{79}$$

The x-derivative of the free energy (59) receives a site contribution equal to:

$$d^{(1)} = \frac{1}{x} \frac{\sum_\alpha \exp(-\beta x \Delta F^\alpha) \Delta F^\alpha}{\sum_\alpha \exp(-\beta x \Delta F^\alpha)}. \tag{80}$$

These contributions are then averaged over many iterations.

One can also add a new link to the system. This is done by merging k lines on the left side of the new link, generating the local fields h_0^α, and similarly merging k lines on the right side of the new link, generating the local fields g_0^α h_0^α. The corresponding free energy shift ΔF^α is now computed using (35). The bond contribution (48) to the

free energy is computed by the same expression as (78), but using this free energy shift ΔF^α corresponding to a bond-addition.

The x-derivative of the free energy receives a bond contribution (60) equal to:

$$d^{(2)} = \frac{1}{x}\frac{\sum_\alpha \exp(-\beta x \Delta F^\alpha)\Delta F^\alpha}{\sum_\alpha \exp(-\beta x \Delta F^\alpha)} \cdot \tag{81}$$

In order to compute the internal energy and the link overlaps, it is useful to introduce the local fields v_0^α defined by

$$v_0^\alpha = \frac{1}{\beta}\text{atanh}\frac{\tanh(\beta J_0) + \tanh(\beta h_0^\alpha)\tanh(\beta g_0^\alpha)}{1 + \tanh(\beta J_0)\tanh(\beta h_0^\alpha)\tanh(\beta g_0^\alpha)} \cdot \tag{82}$$

From this population one generates a population of new fields v^α which takes into account the appropriate weights of the fields with one of the two procedures A or B, similarly to what was done in (76) or in (77) for going from the fields h_0^α to h^α. From (52, 56, 57), the contributions to the internal energy and link overlaps are given by:

$$U = -J_0 \sum_\alpha \tanh(\beta v^\alpha), \tag{83}$$

$$q_1^{(l)} = \frac{1}{\mathcal{M}}\sum_\alpha \tanh^2(\beta v^\alpha),$$

$$q_0^{(l)} = \frac{1}{\mathcal{M}(\mathcal{M}-1)}\sum_{\alpha\neq\gamma} \tanh(\beta v^\alpha)\tanh(\beta v^\gamma). \tag{84}$$

In order to compute the free energy with the simplified formula (50), one also needs the contribution $F^{(1')}$ to the free energy, which is obtained by the same expression as (78), but using the free energy shift ΔF^α obtained by merging k lines.

Let us therefore summarize the main lines of our population dynamics algorithms for solving the Bethe lattice spin-glass problem at the level of one-step RSB. We have used two algorithms, A and B, which differ in the reweighing procedure used, but have otherwise the same skeleton:

1. Start from the population of $\mathcal{N}\times\mathcal{M}$ local fields h_i^α.
2. Merge $k+1$ lines and compute the site observables:
   a) Choose at random $k+1$ sites $i_1, ..., i_{k+1}$ in $\{1, ..., \mathcal{N}\}$.
   b) For each of these $k+1$ sites, say on site $j \in \{i_1, ..., i_{k+1}\}$, one has a population of $\mathcal{M}$ local fields h_j^α, $\alpha = 1, ..., \mathcal{M}$. For each of the $\mathcal{M}$ states, compute the new local field H_0^α obtained by merging $k+1$ branches, using (8). Compute the corresponding free energy shift ΔF^α using (33).
   c) Knowing the sets of fields H_0^α and free energy shift ΔF^α, generate a new set of fields H^α according to (76) in algorithm A (resp. (77) if one uses algorithm B).
   d) Compute the site contribution to the free energy using (78), its x-derivative using (80) and the contribution to the site overlaps (79).
3. Merge $2k$ lines onto a new bond and compute the bond overlaps:
   a) Choose at random $2k$ sites $i_1, ..., i_k, j_1, ..., j_k$ in $\{1, ..., \mathcal{N}\}$.
   b) From the sites $i_1, ..., i_k$, compute the $\mathcal{M}$ local fields h_0^α obtained by merging the k branches, using (5). From the sites $j_1, ..., j_k$, compute the $\mathcal{M}$ local fields g_0^α obtained by merging the k branches, using (5). Deduce the $\mathcal{M}$ local fields v_0^α using (82).
   c) Compute the free energy shifts ΔF^α using (35).
   d) Knowing the sets of fields v_0^α and free energy shifts ΔF^α, generate a new set of fields v^α according to (76) in algorithm A (resp. (77) if one uses algorithm B).
   e) Compute the bond contribution to the free energy using (78), its x-derivative using (81) and the contribution to the internal energy (83) and link overlaps (84).
4. Merge k lines and update the population of fields:
   a) Choose at random k sites $i_1, ..., i_k$ in $\{1, ..., \mathcal{N}\}$.
   b) For each of these k sites, say on site $j \in \{i_1, ..., i_k\}$, one has a population of $\mathcal{M}$ local fields h_j^α, $\alpha = 1, ..., \mathcal{M}$. For each of the $\mathcal{M}$ states, compute the new local field h_0^α obtained by merging k branches, using (5). Compute the corresponding free energy shift ΔF^α using (32).
   c) Knowing the sets of fields h_0^α and free energy shift ΔF^α, generate a new set of fields h^α according to (76) in algorithm A (resp. (77) if one uses algorithm B).
   d) Compute the contribution $F^{(1')}$ to the simple form (50) of the free energy using (78).
   e) Pick up the site $i \in \{1, ..., \mathcal{N}\}$ sequentially, and substitute the local fields $h_i^1, ..., h_i^{\mathcal{M}}$ by the new local fields h^α.
5. Start again the iteration from 2.

Obviously one needs not really perform the above three merging procedures sequentially. In our actual algorithm, we select $2k+2$ random points, merge two groups of $k+1$ to compute site observable, two groups of k to compute bond observables and to update two new sets of $\mathcal{M}$ fields.

A word about the difference between the two algorithms. In algorithm A, the value of r must be chosen large enough so that the probability of a state with the highest old free energy $F^{\alpha'}$ to enter the set of selected $\mathcal{M}$ states be negligible. Both $\mathcal{M}$ and r must go to infinity and in the numerical computations we have taken $r = \mathcal{M}$. Algorithm B is faster than algorithm (A) by a factor that is asymptotically proportional to r for large r. In algorithm B there is no need of introducing $r\mathcal{M}$ fields at an intermediate stage: it corresponds to the discretization of equation (45) and we care take of the effects of the reshuffling by explicitly reweighing the fields. Unfortunately, as we shall see in the next section, the finite $\mathcal{M}$ corrections are empirically larger in algorithm B that in algorithm A. In our case algorithm B turned out to be faster by a factor about 10, but we mentioned both algorithms because algorithm A is somewhat simpler conceptually (and closer to the original discussion of the cavity method), and also because in different situations (*e.g.* depending on the value of x) the relative advantages may be reversed.

7 Numerical results

Here we present the numerical results for one case in order to study the dependence of the algorithm on the various parameters involved in the numerical computation.

We consider the Ising spin glasses with binary couplings ($J = \pm 1$) on a random lattice of fixed coordination 6 ($k = 5$). High precisions measurements [29] of the internal energy are available at temperatures greater or equal to 0.8 for different values of the number of spins N (up to $N = 4096$). The data of the energy at $T = 0.8$ *versus* size can be very well fitted by a power law correction:

$$U(N) = U + AN^{-\omega} \tag{85}$$

with a quite reasonable value of $\omega \sim 0.767 \pm 0.008$ and $A \sim 2.59 \pm 0.02$. The value of the internal energy at infinite size is estimated to be

$$U = -1.799 \pm 0.001. \tag{86}$$

We have done a replica symmetric computation for different values of the population size $\mathcal{N}$. For large $\mathcal{N}$ there are corrections proportional to $1/\mathcal{N}$ (as expected) and for $\mathcal{N} > 10^3$ the $1/\mathcal{N}$ corrections are negligible within our accuracy. With $I = 100$, $S = 1000$, $\mathcal{N} = 4000$, iterating $1000 \times \mathcal{N}$ times (and dropping the first $100 \times \mathcal{N}$ results for allowing transients to decay), we obtain the following replica symmetric results for the free energy, internal energy, entropy, site and link Edwards-Anderson parameters:

$$\begin{aligned} F = -1.863 \pm 0.002,\ U = -1.8160 \pm 0.001 \\ S = 0.058 \pm 0.004,\ q = 0.6863 \pm 0.0002 \\ q_{\rm link} = 0.6385 \pm 0.0003. \end{aligned} \tag{87}$$

Notice that the value of the internal energy totally disagrees with the one found in the simulations (86).

We have done a computation at the one step RSB level using the two algorithms described in the previous section, always at temperature $T = 0.8$.

The crucial point is to find the value of the parameter x such that the derivative of the free energy with respect to x vanishes. In Figure 3 we show our results for the derivative $d(x)$ at $x = 0.21$, plotted versus different values of $\mathcal{M}$. We have used both algorithms A and B. With algorithm B we have used $\mathcal{M} = 2^l$, $l = 3 \ldots 12$ and we plot the results obtained for $l \leq 10$ for a clearer figure. With algorithm A we have used $r = \mathcal{M} \leq 400$. Unfortunately the finite $\mathcal{M}$ corrections are empirically much larger in the a-priori-faster algorithm B. Moreover, in this case although the finite $\mathcal{M}$ corrections seem to be asymptotically proportional to $1/\mathcal{M}$, high order corrections cannot be neglected unless $\mathcal{M}$ is very large. In the end both algorithms give compatible asymptotic results at large $\mathcal{M}$ as seen on Figure 3, with similar computer efforts (for this temperature and the values of x which are relevant).

In Figure 4 we plot the extrapolated result at $\mathcal{M} = \infty$ for the free energy derivative $d(x)$, obtained using algorithm B. The data has been extrapolated with a second order polynomial of $\mathcal{M}^{-1}$ in the interval $\mathcal{M} = [256–4096]$.

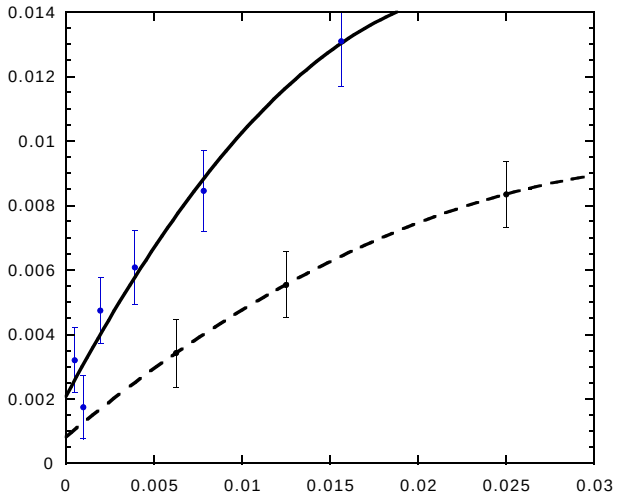

Fig. 3. The x-derivative of the free energy, $d(x)$, evaluated at $x = 0.21$, is plotted as function of $\mathcal{M}^{-1}$ for the algorithm A (upper curve) and the algorithm B (lower curve).

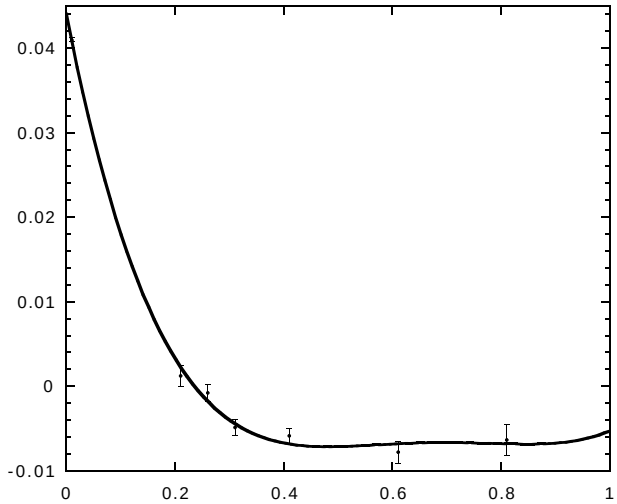

Fig. 4. The x-derivative of the free energy, $d(x)$, computed using the algorithm (B) and extrapolated at large values of $\mathcal{M}$, plotted *versus* x.

Replica symmetry breaking is clearly present. The value of x where the free energy is maximum, which can be obtained by estimating the zero of the function $d(x)$, is given by $x^* = 0.24 \pm 0.02$. We use a similar procedure for all other observables: we extrapolate the x dependent results at $\mathcal{M} = \infty$ and evaluate the errors due to the imprecise location of x^* (which is by far the largest source of error). This gives the following values of the free energy, internal energy, entropy, site and link overlaps:

$$\begin{aligned} F = -1.858 \pm 0.002,\ U = -1.799 \pm 0.001 \\ S = 0.074 \pm 0.004,\ q_1 = 0.779 \pm 0.006,\ q_0 = 0.30 \pm 0.02 \end{aligned} \tag{88}$$

$$q_1^{(l)} = 0.706 \pm 0.007, q_0^{(l)} = 0.408 \pm 0.01 \tag{89}$$

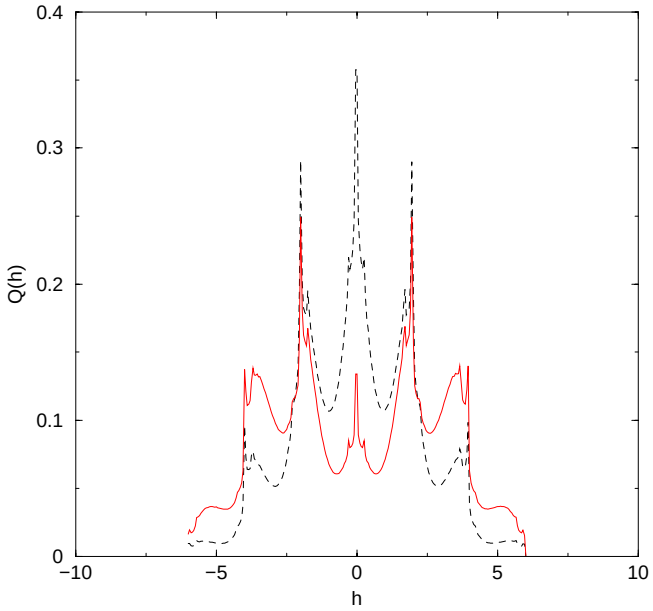

Fig. 5. The probability distribution of the field H_0 at $x = 1$ before (dashed black curve) and after the reweighing (full red curve).

The energy is in very good agreement with the results from simulations. In order to compare the values of the overlaps, one can study the quantity $\langle q^2 \rangle = \int \mathrm{d}q P(q) q^2$. In our RSB theory we find $\langle q^2 \rangle = (1 - x^*) q_1^2 + x^* q_0^2 = 0.485 \pm 0.015$ which agrees well with the numerical value $\langle q^2 \rangle = 0.49 \pm 0.02$. In order to perform a finer comparison it is useful to consider a quantity which is sensitive to replica symmetry breaking. A natural choice is

$$R = \int \mathrm{d}q P(q) q^4 - \left(\int \mathrm{d}q P(q) q^2 \right)^2 . \qquad (90)$$

We find in our RSB theory $R = 0.046 \pm 0.002$ which is again in good agreement with the result of the simulations extrapolated at infinite volume: $R = 0.051 \pm 0.002$. The agreement is remarkable if we recall that in the replica symmetric case $R = 0$. The possible small difference between our value and the simulation data is likely due to the the approximation of one step replica symmetry breaking (it is quite likely that replica symmetry should be broken in a continuous way, as happens in the limit of infinite coordination number). One should notice that the order parameter q_0 is non zero, which explains why some previous attempts at solving the one step RSB problem within a restricted subspace with $q_0 = 0$ did not improve much on the RS solution [32] (the necessity of having a non vanishing q_0 was already noticed in [26]).

Finally let us point out the crucial effect of the reweighing of the states which modifies the local fields as in (45). In Figure 5 we plot the probability distribution of the field H_0 before and after the reweighing, at $x = 1$.

8 Conclusion

We have presented a general solution of finite connectivity spin glass problems at the level of one step RSB. This solution uses the cavity method, together with a kind of population dynamic algorithm to solve the complicated functional equations. As exemplified by a detailed numerical study of the spin glass on a random lattice with fixed connectivity, this method allows to obtain good agreement with numerical simulations. It should be rather easily extendable to many other disordered systems with a finite connectivity, including the fluctuating valence spin glass and various optimization problems such as the K-satisfiability problem.

The possibility to go to higher order of RSB should be explored. In principle the method we have presented can be extended to higher order. At second order one will need a population of $\mathcal{M}_s$ states within a given cluster, and a population of $\mathcal{M}_c$ clusters of states. Therefore the algorithm must follow a total population of $\mathcal{N}\mathcal{M}_c\mathcal{M}_s$ states. The problem will be to see if the resulting algorithm reaches accurate results within the numerically accessible values of $\mathcal{N}$, $\mathcal{M}_c$, $\mathcal{M}_s$.

Some variants of the method should also be explored. In particular we have not exploited the variational formulation in the computation of the probability $Q(\mathbf{h})$. One could also study in details the shape of the probability distributions of local fields in order to understand how they could parametrized in a simple but efficient way so that the free energy to be minimized does not involve a too large number of parameters.

We thank Giulio Biroli, Cristiana Carrus, Enzo Marinari, Rémi Monasson, Riccardo Zecchina and Francesco Zuliani for useful discussions and comments. This research has been supported in part by the European Science Foundation through the SPHINX programme.

Appendix A: Weighted sums of uncorrelated variables

In this appendix we want to prove a useful little theorem, which applies to the computation of $F^{(1)}$ given in (47) and $F^{(2)}$ in (48).

Theorem:

Consider a set of $\mathcal{M}(\gg 1)$ iid random free energies f_α, ($\alpha \in \{1 \dots M\}$) distributed with the exponential density equation (37), and a set of $\mathcal{M}$ positive numbers a_α. Then,

neglecting terms which go to zero when M goes to infinity, the following relation holds:

$$\left\langle \ln\left(\frac{\sum_\alpha a_\alpha \exp(-\beta f_\alpha)}{\sum_\alpha \exp(-\beta f_\alpha)}\right)\right\rangle_f \equiv \left\langle \ln\left(\sum_\alpha w_\alpha a_\alpha\right)\right\rangle_f = \frac{1}{x}\ln\left(\frac{1}{M}\sum_\alpha a_\alpha^x\right) \quad (91)$$

where $\langle . \rangle_f$ denotes an average over the distribution of f.

Corollary 1:

In the same conditions as the theorem, for any set of M real numbers b_α, one has:

$$\left\langle \frac{\sum_\alpha a_\alpha b_\alpha \exp(-\beta f_\alpha)}{\sum_\alpha a_\alpha \exp(-\beta f_\alpha)}\right\rangle_f = \frac{\sum_\alpha a_\alpha^x b_\alpha}{\sum_\alpha a_\alpha^x}\,. \quad (92)$$

Corollary 2:

In the same conditions as the theorem, for any set of M real numbers b_α, one has:

$$\left\langle \frac{\sum_\alpha a_\alpha b_\alpha^2 \exp(-\beta f_\alpha)}{\sum_\alpha a_\alpha \exp(-\beta f_\alpha)}\right\rangle_f - \left\langle \left(\frac{\sum_\alpha a_\alpha b_\alpha \exp(-\beta f_\alpha)}{\sum_\alpha a_\alpha \exp(-\beta f_\alpha)}\right)^2\right\rangle_f = x\left[\frac{\sum_\alpha a_\alpha^x b_\alpha^2}{\sum_\alpha a_\alpha^x} - \left(\frac{\sum_\alpha a_\alpha^x b_\alpha}{\sum_\alpha a_\alpha^x}\right)^2\right]. \quad (93)$$

Corollary 3:

If the numbers a_α are M iid positive random variables, such that the average of a^x exists, which are uncorrelated with the f_α, then one has:

$$\left\langle \ln\left(\frac{\sum_\alpha a_\alpha \exp(-\beta f_\alpha)}{\sum_\alpha \exp(-\beta f_\alpha)}\right)\right\rangle_f \equiv \left\langle \ln\left(\sum_\alpha w_\alpha a_\alpha\right)\right\rangle_f = \frac{1}{x}\ln\left(\langle a_\alpha^x\rangle_a\right) \quad (94)$$

where $\langle . \rangle_a$ denotes an average over the distribution of a.

Proof:

we follow some of the techniques exposed in [3]. We start from the identity

$$\ln\left(\sum_\alpha \exp(-\beta f_\alpha)a_\alpha\right) = \int_0^\infty \frac{\mathrm{d}t}{t}\left[\exp(-t) - \exp\left(-t\sum_\alpha \exp(-\beta f_\alpha)a_\alpha\right)\right]. \quad (95)$$

We choose to work with a regularised distribution of the M iid random variables f_α:

$$P(f_\alpha) = \beta x \exp(\beta x(f_\alpha - f_c))\ \theta(f_c - f), \quad (96)$$

where in the end we shall send $M \to \infty$, $f_c \to \infty$, with $r = M\exp(-\beta f_c)$ fixed (the value of r is irrelevant). In this limit one has:

$$\langle \exp\left(-t\exp(-\beta f_\alpha)a_\alpha\right)\rangle_f \simeq 1 - (ta_\alpha)^x \exp(-\beta x f_c)\Gamma(1-x), \quad (97)$$

from which one deduces:

$$\left\langle \ln\left(\sum_\alpha \exp(-\beta f_\alpha)a_\alpha\right)\right\rangle_f = \int_0^\infty \frac{\mathrm{d}t}{t}\left[\exp(-t) - \exp(-\Gamma(1-x)t^x \exp(-\beta x f_c)\sum_\alpha a_\alpha^x\right] = \frac{1}{x}\ln\left(\Gamma(1-x)t^x M\exp(-\beta x f_c)\sum_\alpha a_\alpha^x\right) + \frac{1-x}{x}C, \quad (98)$$

where C is Euler's constant. The quantity we need to compute involves subtracting the same expression with a_α substituted by one, which gives the desired result:

$$\left\langle \ln\left(\sum_\alpha \exp(-\beta f_\alpha)a_\alpha\right) - \ln\left(\sum_\alpha \exp(-\beta f_\alpha)\right)\right\rangle_f = \frac{1}{x}\ln\left(\frac{1}{M}\sum_\alpha a_\alpha^x\right). \quad (99)$$

The proof of Corollary 1 is easily obtained by applying the theorem to the set of numbers $a_\alpha \exp(\lambda b_\alpha)$, and taking the derivative with respect to λ at $\lambda = 0$. Similar generalised formulas can be obtained by taking higher order derivatives. The second derivative gives Corollary 2. Corollary 3 is trivial.

Remark:

We notice that in the two limits $x \to 0$ and $x \to 1$ the equations can be simply understood:

- In the limit $x = 0$, in a typical realization of the random free energies, only one weight w is equal to one and all the others are zero. Averaging over the realizations of free energies amounts to spanning uniformly the set of indices of this special non zero weight.
- In the limit $x = 1$ the number of relevant w goes to infinity and each individual contribution goes to zero. An infinite number of term is present in the l.h.s. of equation (91) and the r.h.s. of the equation (91) becomes $\ln\left[(1/M)\sum_\alpha a_\alpha\right]$, as it should.

References

1. D. Sherrington, S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).
2. G. Parisi, J. Phys. A **13**, L115, 1101, 1887 (1980).
3. M. Mézard, G. Parisi, M.A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
4. M. Mézard, G. Parisi, M.A. Virasoro, Europhys. Lett. **1**, 77 (1986). See also [3].
5. F. Guerra, Int. J. Mod. Phys. B **10**, 1675 (1997).
6. M. Aizenman, P. Contucci, J. Stat. Phys. **92**, 765 (1998).
7. M. Talagrand, Probab. Theory Related Fields **110**, 109 (1998).
8. M. Mézard, G. Parisi, J. Phys. France **47**, 1285 (1986); W. Krauth, M. Mézard, Europhys. Lett. **8**, 213 (1989).
9. M. Mézard, G. Parisi, J. Phys. Lett. **46**, L771 (1985).
10. Y. Fu, P.W. Anderson, J. Phys. A **19**, 1605 (1986).
11. J.R. Banavar, D. Sherrington, N. Sourlas, J. Phys. A **20**, L1 (1987).
12. S. Kirkpatrick, B. Selman, Science **264**, 1297 (1994); R. Monasson, R. Zecchina, Phys. Rev. E **56**, 1357 (1997); R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, L. Troyansky, Nature **400**, 133 (1999).
13. M.W. Klein, L.J. Schowalter, P. Shukla, Phys. Rev. B **19**, 1492 (1979).
14. S. Katsura, S. Inawashiro, S. Fujiki, Physica A **99**, 193 (1979).
15. K. Nakanishi, Phys. Rev. B **23**, 3514 (1981).
16. D.R. Bowman, K. Levin, Phys. Rev. B **25**, 3438 (1982).
17. D.J. Thouless, Phys. Rev. Lett. **56**, 1082 (1986).
18. J.T. Chayes, L. Chayes, J.P. Sethna, D.J. Thouless, Commun. Math. Phys. **106**, 41 (1986); J.M. Carlson, J.T. Chayes, L. Chayes, J.P. Sethna, D.J. Thouless, Europhys. Lett. **5**, 355 (1988).
19. C. De Dominicis, P. Mottishaw, J. Phys. A **20**, L375 (1987).
20. P. Mottishaw, Europhys. Lett. **4**, 333 (1987).
21. R.C. Dewar, P. Mottishaw, J. Phys. A **21**, L1135 (1988).
22. P.Y. Lai, Y.Y. Goldschmidt, J. Phys. A **22**, 399 (1989).
23. C. De Dominicis, Y.Y. Goldschmidt, J. Phys. A **22**, L775 (1989), Phys. Rev. B **41**, 2184 (1990).
24. H. Orland, J. Physique Lett. **46**, L768 (1985).
25. M. Mézard, G. Parisi, Europhys. Lett. **3**, 1067 (1987).
26. Y.Y. Goldschmidt, P.Y. Lai J. Phys. A **23**, L775 (1990).
27. R. Monasson, J. Phys. A **31**, 513 (1998); Phil. Mag. B **77**, 1515 (1998).
28. G. Parisi, F. Ritort J. Phys. I France **3**, 969 (1993).
29. C. Carrus, E. Marinari, F. Zuliani (in preparation).
30. One could develop a similar discussion by removing a link at random.
31. I. Kanter, H. Sompolinsky, Phys. Rev. Lett. **58**, 164 (1987).
32. K.Y. Wong, D. Sherrington, J. Phys. A **21**, L459 (1988).
33. A detailed such investigation was carried out in the K-sat problem by G. Biroli, R. Monasson, M. Weigt, Eur. Phys. J. B **14**, 551 (2000).
34. S.F. Edwards, P.W. Anderson, J. Phys. F **5**, 965 (1975).
35. We believe that assumption equation (41) is not only consistent in the present model, but that its consistency is a general feature of one step replica symmetry breaking.
36. R. Monasson, Phys. Rev. Lett. **75** 2847 (1995).
37. S. Franz, G. Parisi, J. Phys. I France **5**, 1401 (1995).
38. R. Baviera, M.A. Virasoro, Physica D **107**, 151 (1997).
39. R. Abou-Chacra, P.W. Anderson, D.J. Thouless, J. Phys C **6**, 1734 (1973), R. Abou-Chacra, D.J. Thouless, J. Phys C **7**, 65 (1974).