Large deviations of the free energy in diluted mean-field spin-glass

2010 J. Phys. A: Math. Theor. 43 045001

(http://iopscience.iop.org/1751-8121/43/4/045001)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **43** (2010) 045001 (18pp) doi:10.1088/1751-8113/43/4/045001

Large deviations of the free energy in diluted mean-field spin-glass

Giorgio Parisi[1,2] **and Tommaso Rizzo**[1]

[1] Dipartimento di Fisica, Università di Roma 'La Sapienza', P.le Aldo Moro 2, 00185 Roma, Italy
[2] Statistical Mechanics and Complexity Center (SMC)-INFM-CNR, Italy

E-mail: giorgio.parisi@roma1.infn.it

Received 23 October 2009
Published 23 December 2009
Online at stacks.iop.org/JPhysA/43/045001

Abstract

Sample-to-sample free-energy fluctuations in spin-glasses display a markedly different behaviour in finite-dimensional and fully connected models, namely Gaussian versus non-Gaussian. Spin-glass models defined on various types of random graphs are in an intermediate situation between these two classes of models and we investigate whether the nature of their free-energy fluctuations is Gaussian or not. It has been argued that Gaussian behaviour is present whenever the interactions are locally non-homogeneous, i.e. in most cases with the notable exception of models with fixed connectivity and random couplings $J_{ij} = \pm\tilde{J}$. We confirm these expectations by means of various analytical results concerning the large deviations of the free energy. In particular we unveil the connection between the spatial fluctuations of the populations of fields defined at different sites of the lattice and the Gaussian nature of the free-energy fluctuations. In contrast, on locally homogeneous lattices the populations do not fluctuate over the sites and as a consequence the small deviations of the free energy are non-Gaussian and scale as in the Sherrington–Kirkpatrick model.

PACS number: 75.50.Lk

1. Introduction

The problem of sample-to-sample free-energy fluctuations in spin-glasses has attracted large interest in recent years both from the theoretical and numerical point of view [1–15].

In general at fixed system size N the free energy *per spin* of a given sample is a random variable with mean f_N and variance σ_N^2. In the thermodynamic limit, f_N approaches a definite value f_{typ} while σ_N approaches zero: the free energy is self-averaging and does not depend on the sample. It is expected that in the thermodynamic limit, the rescaled variable $x = (f - f_N)/\sigma_N$ has a limiting probability distribution that describes the *small deviations* of the free energy, i.e. those that occur with finite probability and whose scale decreases with N. On the other hand large deviations of the free energy, i.e. those that remain finite in the large N limit, have a probability that vanishes exponentially with the size of the system and are described by a large deviation function [13–16].

In finite-dimensional models one expects the small deviations of the free energy to have a Gaussian distribution with variance proportional to the volume of the system, i.e. $\sigma_N \propto N^{-1/2}$ [7, 12, 17]. In mean-field models, however, the standard arguments leading to this expectation no longer hold and more exotic situations can be observed. Indeed the scaling of σ_N with N in the Sherrington–Kirkpatrick (SK) model [18] has been largely studied numerically in recent years [1–6, 10] and there is growing consensus that $\sigma_N \propto N^{-5/6}$. This scaling had been conjectured early in [16] using a large deviation result by Kondor [19] under the assumption that there is a matching between large and small deviations (for a recent discussion see [14]). The large deviation result of Kondor had been questioned in [7, 8] but was later proved to be correct [20]. The distribution of *negative* large deviations was computed down to zero temperature and excellent agreement with numerical results was found [13, 14]. Currently there is no analytical tool to compute directly σ_N; nevertheless, the recent computation of *positive* large deviations [15] adds further support to the 5/6 scaling. Furthermore, the evaluation of the small-deviation distribution of the SK model is an open problem. We just know that it is not a Gaussian function, but it has been argued that the behaviour of its tails can be deduced from the large deviations behaviour [15].

The mean-field spin-glass models defined on random graphs are considered as something intermediate between fully connected models (notably the SK model) and finite-dimensional models. Indeed they have a mean-field nature but are more realistic because every spin interacts with a finite number of neighbours. Therefore, it is natural to ask what is the behaviour of the free-energy fluctuations in these models. In particular we can ask if the small deviation distribution is Gaussian and if the large deviation function is similar to the one found in the SK model. The ground-state properties of spin-glass models defined on the Bethe Lattice have been studied intensively in recent times [2, 11, 21–23]. Our investigation was motivated by the findings of [2]. The authors of [2] observed that both the scaling of σ_N and the shape of probability distribution of the ground-state energy of a spin-glass defined on random graphs with fixed connectivity depend on the distribution of the couplings. In particular on a random graph with fixed connectivity and Gaussian distributed couplings J_{ij} it turned out that $\sigma_N \propto N^{-1/2}$ (as in finite dimensional models) and that the skewness of the small-deviation distribution tends to zero at large N, consistently with the assumption that it is a Gaussian. However, in the case of fixed connectivity and bimodal distribution of the couplings ($J_{ij} = \pm\tilde{J}$) they observed a scaling of σ_N definitively different from $N^{-1/2}$ (possible values being $N^{-3/4}$ or $N^{-4/5}$). Furthermore, it turned out that the skewness of the small deviation distribution instead of vanishing at large N tends to go to a finite value consistent with that of the SK model.

In this paper we compute the large deviations of the models considered by [2], i.e. spin-glass models defined on graphs with fixed connectivity. We consider the following functional [13, 14, 16]:

$$\Phi_N(n, \beta) = -\frac{1}{\beta n N} \ln \overline{Z_J(\beta)^n}, \tag{1}$$

where different systems (or samples) are labelled by J, $Z_J(\beta)$ is the partition function of a given sample and the bar denotes the average over different disordered samples. The above functional is the generating function of the cumulants of the sample-to-sample fluctuations of the free energy. In order to determine the moments of the small deviation distribution we should first compute the derivatives with respect to n of $\Phi_N(n)$ at $n = 0$, and then take the limit $N \to \infty$. This apparently simple step is a complex and open problem also in the SK model [13–15]; instead the opposite case in which one takes first the $N \to \infty$ limit and then the $n \to 0$ limit is tractable, see [14] for a complete discussion of this issue. This amounts to compute

$$\Phi(n, \beta) = \lim_{N\to\infty} \Phi_N(n, \beta). \tag{2}$$

It is well known that the probability of large deviations is related to the function $\Phi(n, \beta)$. Indeed

$$\exp(-\beta n N \Phi(n, \beta)) = \overline{Z_J(\beta)^n} = \overline{\exp(-nN\beta f_J(\beta))}, \tag{3}$$

where f_J is the system-dependent free energy *per spin*. The region of positive n corresponds to fluctuations where the free energy is smaller than the typical one and the region on negative n corresponds to fluctuations where the free energy is larger than the typical one.

We define the large deviation function for the free energy, $L(f)$, (which we will call in the following the sample complexity because it is related to the number of samples with free energy equal to f), as the logarithm divided by N of the probability density of samples with free energy per spin f in the thermodynamic limit:

$$L(f) \equiv \lim_{N\to\infty} \frac{\log(P_N(f))}{N}. \tag{4}$$

For large N the majority of the samples have free energy per spin equal to f_{typ}, and all other values have exponentially small probability. Consistently $L(f)$ is less than or equal to zero, the equality holding $f = f_{\text{typ}}$, i.e. $L(f_{\text{typ}}) = 0$. For some values of f it is possible that $L(f) = -\infty$, meaning that the probability of large deviations goes to zero faster than exponentially with N. In the thermodynamic limit the function $\Phi(n, \beta)$ defined in equation (1) yields the Legendre transform of $L(f)$ [16]; indeed we have

$$-\beta n \Phi(n) = -\beta n f + L(f) \tag{5}$$

where f is determined by the condition

$$\beta n = \frac{\partial L}{\partial f} \tag{6}$$

and equivalently we have

$$L(f) = \beta n f - \beta n \Phi(n) \tag{7}$$

where βn is determined by the condition

$$f = \frac{\partial n\Phi}{\partial n}. \tag{8}$$

Note that while numerical methods are best suited to study small deviations (precisely because they are typical) present day theoretical methods deal mainly with large deviations. This is because at the theoretical level large deviations require essentially treating a replica field theory at the mean-field level, while small deviations require the computation of the loop corrections. Furthermore, the problem is made even more complicated by the fact that these corrections are singular [15]. Nevertheless, it is usually assumed that information on small deviations can be extracted from large deviations [14, 16]. In particular, if the expansion of $n\Phi(n)$ around $n = 0$ reads $n\Phi(n) = nf_{\text{typ}} + c_2 n^2$ one expects that $\sigma_N \approx (-2c_2/\beta)^{1/2} N^{-1/2}$ and that the small deviation function is a Gaussian. In the SK model instead $n\Phi(n) = nf_{\text{typ}} + c_6 n^6$ for positive n and $n\Phi(n) = nf_{\text{typ}}$ for negative n [24] and this has led to the 5/6 scaling prediction for σ_N [16]. Correspondingly the small deviation distribution is not expected to be a Gaussian, as confirmed by the numerics [2]. In this case however the constant c_6 is not related to the sixth moment of the small deviation distribution but rather to its right tail [14]. Note that the fact that the function $\Phi(n)$ is constant for negative n leads to $L(f) = -\infty$ for $f > f_{\text{typ}}$; indeed in this region the probability vanishes as $\exp[-O(N^2)]$ [15]. This interpretation scheme allows us to make contact between our results and those of [2]. Indeed we find that in random graphs with fixed connectivity the presence of an n^2 term in the expansion of $n\Phi(n)$ depends on the nature of the coupling distribution. For a generic distribution (e.g. Gaussian) of the coupling, the term is present and the small deviations are expected to be Gaussian. Nevertheless, if the couplings have a bimodal distribution $J_{ij} = \pm\tilde{J}$, the n^2 term is absent and $\Phi(n)$ has the same qualitative properties of the SK model, namely $n\Phi(n) = nf_{\text{typ}} + c_6 n^6$ (with a different c_6) for positive n and $n\Phi(n) = nf_{\text{typ}}$ for negative n. Thus, in the bimodal case we expect that $\sigma_N \propto N^{-5/6}$ and that the small deviation distribution is not a Gaussian, possibly the same of the SK model [15]. With respect to the data from [2], we think that the expected 5/6 scaling is seen only at a high enough system size as in the case of the SK model [10].

Our results on $\Phi(n)$ are thus in agreement with the findings of [2]. Interestingly, the authors of [2] suggested that the peculiar behaviour of free-energy fluctuations for bimodal distribution of the couplings is caused by the fact that the Hamiltonian is locally homogeneous. Indeed around a given site the negative couplings can be transformed into positive coupling through a Gauge transformation. The process must stop when one of the sites already encountered in the process is reached again due to a loop of the graph; however, since loops are typically large, the disordered nature of the model appears 'at infinity'. As we will see in the following our results put this intuition on a firm ground. Indeed we will show that they are precisely the spatial inhomogeneities of the interactions that generate an n^2 term in $n\Phi(n)$. In contrast in the bimodal case, the local homogeneity allows us to obtain a solution that does not fluctuate over the sites (in a sense to be specified below) and this guarantees that $\Phi(n) = nf_{\text{eq}}$ for positive and negative n both in the replica-symmetric (RS) case and for $n < 0$ in the replica-symmetry-breaking (RSB) case similarly, as in the SK model.

The plan of the paper is the following. In section 2 we will write down the variational expression of $\Phi(n)$ for the spin-glass on a random lattice with fixed connectivity (the Bethe lattice). In sections 3 and 4 we will consider the RS and RSB solution, respectively. In both cases we will show that local inhomogeneities lead to the presence of an $O(n^2)$ term whose coefficient can be expressed in terms of the spatial fluctuations of the local fields. In contrast when the interactions have a bimodal distribution, the resulting local homogeneity allows us to obtain a solution that does not fluctuate over the sites and leads to the vanishing of the n^2 term. This also implies that $\Phi(n) = nf_{\text{eq}}$ for positive and negative n in the RS case and for $n < 0$ in the full-RSB case much as in the SK model. In the latter case in order to compute the first non-trivial term for positive n we resort to an expansion of $\Phi(n)$ in powers of the order parameter. This is presented in section 5. We will confirm that in the bimodal case the $O(n^2)$ term vanishes and that the first non-trivial term in the expansion of $\Phi(n)$ is $O(n^5)$ much as in the SK model. At the end we will give our conclusions and discuss some interesting consequences of our results.

2. The functional $\Phi(n)$ of the Bethe lattice spin-glass

In this section we discuss the potential $\Phi(n)$ of the spin-glass defined on the Bethe lattice with fixed connectivity $M + 1$. Following [25] we express $\Phi(n)$ as a variational functional of the order parameter $\rho(\sigma)$ that is a function defined on n Ising spins σ. The variational expression of the free energy reads

$$n\beta\Phi(n) = M \ln \text{Tr}_{\{\sigma\}}\rho^{M+1}(\sigma) - \frac{M+1}{2} \ln \int \text{Tr}_{\{\sigma\}}\,\text{Tr}_{\{\tau\}}\rho^M(\sigma)\rho^M(\tau)\left\langle \exp \beta J \sum_\alpha \sigma_\alpha \tau_\alpha \right\rangle \tag{9}$$

where the square brackets mean the average with respect to the distribution of J. The above expression has to be extremized with respect to $\rho(\sigma)$. We note that it is invariant under a rescaling of $\rho(\sigma)$ so that we can choose any normalization for it. If we normalize $\rho(\sigma)$ to the corresponding variational equation in terms of $\rho(\sigma)$, then

$$\rho(\sigma) = \frac{\langle \mathrm{Tr}_\tau \rho^M(\tau) \exp J\sigma\tau \rangle}{\langle \mathrm{Tr}_{\tau,\sigma} \rho^M(\tau) \exp \beta J\sigma\tau \rangle} \tag{10}$$

where $\sigma\tau \equiv \sum_\alpha \sigma_\alpha \tau_\alpha$.

In the next two sections we discuss the RS and RSB ansatz of $\rho(\sigma)$ that are characterized respectively by fields and distributions of fields. In the RS case we will find that a crucial condition, in order *not* to have $O(N^{-1/2})$ Gaussian fluctuations, is that the fields do not fluctuate over the sites (which is possible in the low-temperature phase only if the interactions are locally homogeneous). In the RSB case this condition becomes a condition of homogeneity of the populations of fields, meaning that the populations do not fluctuate. Note that at one-step RSB level this corresponds to the fact that we have to consider the so-called factorized solution and the fact that fluctuating solutions are actually obtained at this level [26] confirms that the true solution is full-RSB.

In the RS case the homogeneity condition guarantees that the same solution valid at $n = 0$ can be used at n different from zero yielding $\Phi(n) = n f_{\mathrm{eq}}$ exactly. Similarly, as in the SK model we expect that this statement holds in the RSB case only for negative n, because the full-RSB solution at positive n cannot be the same at $n = 0$ if we require that $x_{\min} \geqslant n$ where $x_{\min}$ is the first breaking point of $q(x)$. Indeed the expansion in the order parameter of section 5 shows that the model is mapped into the SK model with different coefficients and an explicit computation shows that $\Phi(n)$ for positive n has on $O(n^5)$ behaviour as SK.

3. The replica-symmetric solution

In this section we study the replica-symmetric ansatz on $\rho(\sigma)$. We normalize $\rho(\sigma)$ to 1, following [26]. In the RS case $\rho(\sigma)$ is a function of $\sum_a \sigma_a$ and is parameterized by a function $R(u)$ as

$$\rho(\sigma) = \int \mathrm{d}u\, R(u) \frac{\exp \beta u \sum_a \sigma_a}{(2\cosh\beta u)^n} \tag{11}$$

where $R(u)$ must satisfy $\int \mathrm{d}u\, R(u) = 1$ because of the normalization of $\rho(\sigma)$. Accordingly we have

$$\ln \mathrm{Tr}_{\{\sigma\}} \rho^{M+1}(\sigma) = \ln \int \left(\frac{2\cosh\beta \sum_i^{M+1} u_i}{\prod_i^{M+1} 2\cosh\beta u_i} \right)^n \prod_{i=1}^{M+1} R(u_i)\, \mathrm{d}u_i. \tag{12}$$

We are interested in evaluating what is the dependence on n of the previous quantity for fixed $R(u)$. Expanding in powers of n we get

$$\ln \mathrm{Tr}_{\{\sigma\}} \rho^{M+1}(\sigma) = n[\langle\langle A \rangle\rangle] + \frac{n^2}{2}[(\langle\langle A^2 \rangle\rangle - \langle\langle A \rangle\rangle^2)] + O(n^3) \tag{13}$$

where we have used

$$\langle\langle A^p \rangle\rangle \equiv \int \prod_{i=1}^{M+1} R(u_i)\, \mathrm{d}u_i \left(\ln \frac{2\cosh\beta \sum_i^{M+1} u_i}{\prod_i^{M+1} 2\cosh\beta u_i} \right)^p; \tag{14}$$

thus, we see that iff $R(u) = \delta(u - u_0)$ (i.e. $R(u)$ is concentrated on some value u_0) there is no $O(n^2)$ term. On the other hand it is easily seen that in this case there are no higher terms as well, and the following relationship is valid at all orders in n:

$$\ln \mathrm{Tr}_{\{\sigma\}} \rho^{M+1}(\sigma) = n(\ln 2\cosh\beta(M+1)u_0 - (M+1)\ln 2\cosh\beta u_0). \tag{15}$$

So the crucial condition in order not to have an $O(n^2)$ term is that $R(u)$ is a delta function, and a sufficient condition for this is that the interactions are locally homogeneous. The other term entering the free energy can be expressed as

$$\ln \mathrm{Tr}_{\{\sigma\}} \mathrm{Tr}_{\{\tau\}} \rho^M(\sigma)\rho^M(\tau) \left\langle \exp\left(\beta \sum_a \sigma_a \tau_a\right)\right\rangle$$
$$= \ln \int \prod_{i=1}^M R(u_i)\,\mathrm{d}u_i \prod_{i=1}^M R(v_i)\,\mathrm{d}v_i \left\langle\left(\sum_{\sigma,\tau} \frac{\exp\left[\beta\sum_i^M u_i\sigma + \beta\sum_i^M v_i\tau + \beta J\sigma\tau\right]}{\prod_i^M 4\cosh\beta u_i \cosh\beta v_i}\right)^n\right\rangle \tag{16}$$

where the square brackets mean the average with respect to the distribution of J. As above we can expand in powers of n and obtain

$$\ln \mathrm{Tr}_{\{\sigma\}} \mathrm{Tr}_{\{\tau\}} \rho^M(\sigma)\rho^M(\tau) \left\langle \exp\left(\beta \sum_a \sigma_a \tau_a\right)\right\rangle = n[\langle\langle B\rangle\rangle] + \frac{n^2}{2}[(\langle\langle B^2\rangle\rangle - \langle\langle B\rangle\rangle^2)] + O(n^3) \tag{17}$$

where we have used

$$\langle\langle B^p\rangle\rangle \equiv \int \prod_{i=1}^M [R(u_i)\,\mathrm{d}u_i R(v_i)\,\mathrm{d}v_i] \left\langle\left(\ln\sum_{\sigma,\tau} \frac{\exp\left[\beta\sum_i^M u_i\sigma + \beta\sum_i^M v_i\tau + \beta J\sigma\tau\right]}{\prod_i^M 4\cosh\beta u_i \cosh\beta v_i}\right)^p\right\rangle; \tag{18}$$

thus we see that the $O(n^2)$ is absent if $R(u_i) = \delta(u_i)$ *and* $\rho(J) = \pm\tilde{J}$, according to the criterion of local homogeneity of the interactions. On the other hand, the $O(n^2)$ term is present also if $R(u_i) = \delta(u_i)$ but $\rho(J)$ is not bimodal, e.g. in the high-temperature phase of the corresponding model.

3.1. Fluctuations of the free energy in the replica-symmetric solution

The presence of an $O(n^2)$ term in the large deviation function leads naturally to assume that the small deviations of the free energy are Gaussian. A straightforward computation shows that the variance of the small deviation is proportional to the coefficient c_2 of the $O(n^2)$ in $n\Phi(n)$ according to

$$\langle\Delta F^2\rangle - \langle\Delta F\rangle^2 = -\frac{2c_2}{\beta}N. \tag{19}$$

Around $n = 0$ the coefficient c_2 can be computed noting that since $\Phi(n)$ is stationary with respect to $R(u)$, the derivative with respect to n of $\Phi(n)$ is given by its partial derivative with respect to n.

Using the definition of $\Phi(n)$, equation (9), and the expansions equations (13) and (17) we get

$$\beta n\Phi(n) = n\left[M\langle\langle A\rangle\rangle - \frac{M+1}{2}\langle\langle B\rangle\rangle\right]$$
$$+ \frac{n^2}{2}\left[M(\langle\langle A^2\rangle\rangle - \langle\langle A\rangle\rangle^2) - \frac{M+1}{2}(\langle\langle B^2\rangle\rangle - \langle\langle B\rangle\rangle^2)\right] + O(n^3) \tag{20}$$

where we have used definitions (14) and (18). The previous expression has to be evaluated using the variational $R(u)$ obtained at $n = 0$.

The RS solution with $R(u) = \delta(u)$ is correct above the critical temperature specified by the condition $\langle \tanh^2 \beta_c J \rangle = 1/M$. Therefore, we conclude that above the critical temperature $\Phi(n)$ in general has a term $O(n^2)$ different from zero and its coefficient c_2 is given by

$$c_2 = -\frac{M+1}{4}(\langle(\ln\cosh\beta J)^2\rangle - \langle \ln\cosh\beta J\rangle^2). \tag{21}$$

Clearly this coefficient vanishes in the case of a bimodal distribution. Above the critical temperature the solution is $R(u) = \delta(u)$ also for $n \neq 0$ and $\Phi(n)$ reads

$$\Phi(n) = -\frac{\ln 2}{\beta} - \frac{M+1}{2\beta n}\ln\langle\cosh^n \beta J\rangle; \tag{22}$$

again we see that in the case of a bimodal distribution $\Phi(n)$ does not depend on n. In the low-temperature phase we know that $\rho(\sigma)$ is no longer a constant. The correct parameterization in the low-temperature spin-glass phase is full-RSB. In the following section we will describe the RSB ansatz and show that in general the expansion of $n\Phi(n)$ has an $O(n^2)$ term. Nevertheless, we will see that in the case of a bimodal distribution the $O(n^2)$ term vanishes and that $\Phi(n)$ is constant for $n < 0$ much as in the SK model. In section (5) we will compute the first non-trivial term in $\Phi(n)$ for $n > 0$ and show that it is $O(n^5)$ much as in the SK model.

4. The replica-symmetry-breaking solution

The RSB parameterization of $\rho(\sigma)$ in terms of distributions of fields was presented in [26] and we refer to that paper for an explanation of the main ideas underlying it. In particular we will work in the replica framework rather than using the cavity method. The resulting equations are the same but while the former allows a quicker derivation the latter unveils the physical meaning of the populations and the appearance of free-energy shifts.

We introduce the field u that parameterizes a distribution over the values of an Ising spin σ according to the formula $P(\sigma) = \exp(\beta u\sigma)/2\cosh\beta u$. We define a probability distribution (population) $P^{(0)}(u)$ of such fields a 0-distribution, correspondingly a 1-distribution is a probability distribution on probability distributions (population of populations) and so on. In the following a k-distribution will be written as $P^{(k)}$, and it defines a measure $P^{(k)}\,\mathrm{d}P^{(k-1)}$ over the space of $k-1$-distributions.

In order to parameterize the function $\rho(\sigma)$ with K steps of RSB we need

- a K-distribution $P^{(K)}$;
- K integers $1 \leqslant x_1, \ldots, x_K \leqslant n$ (as usual for $n < 1$ they become real and the inequalities change sign).

In the following we will consider the parameters $1 \leqslant x_1, \ldots, x_K \leqslant n$ fixed and consider just the dependence on the distributions. The construction is iterative and requires a set of functions $\rho_{P^{(k)}}(\sigma)$ of x_{k+1} spins with $k = 1, \ldots, K+1$ (we define $x_{k+1} \equiv n$ and $x_0 \equiv 1$). The normalization of $\rho_{P^{(k)}}$ is crucial; we choose to normalize all of them to 1. We define $\rho_{P^{(k)}}(\sigma)$ starting from $\rho_{P^{(k-1)}}(\sigma)$, we first divide the x_{k+1} spins in x_{k+1}/x_k groups $\{\sigma_{\mathcal{C}}\}$ of x_k spins labelled by an index $\mathcal{C} = 1, \ldots, x_{k+1}/x_k$. Then we have

$$\rho_{P^{(k)}}(\sigma) = \int P^{(k)}\,\mathrm{d}P^{(k-1)} \prod_{\mathcal{C}=1}^{x_{k+1}/x_k} \rho_{P^{(k-1)}}(\{\sigma\}_{\mathcal{C}}). \tag{23}$$

Thus, $\rho(\sigma) \equiv \rho_{P^{(K)}}(\sigma)$ is defined iteratively starting from the replica-symmetric case corresponding to $k = 0$:

$$\rho_{P^{(0)}}(\sigma) = \int P^{(0)}(u)\,\mathrm{d}u \prod_{i=1}^{x_1} \frac{\exp\beta u\sigma_i}{2\cosh\beta u}. \tag{24}$$

With the above definitions it is possible to express the variational free energy (9) in terms of the populations of populations in the same way as we derived equations (12) and (16). In order to do that we need to introduce two functions: $\Delta F_1^{(k)}\big[P_1^{(k)},\ldots,P_{M+1}^{(k)}\big]$ is a function of $M+1$ k-populations and $\Delta F_{12}^{(k)}\big[P_1^{(k)},\ldots,P_{2M}^{(k)},J\big]$ is a function of $2M$ k-populations and a coupling constant J. Their definition is iterative, i.e. the function at level k is defined in terms of the function at level $k-1$:

$$\Delta F_1^{(k)}\big(P_1^{(k)},\ldots,P_{M+1}^{(k)}\big)\equiv-\frac{1}{\beta x_{k+1}}\ln\int\left[\prod_{i=1}^{M+1}P_i^{(k)}\,\mathrm{d}P_i^{(k-1)}\right]\mathrm{e}^{-\beta x_{k+1}\Delta F_1^{(k-1)}(P_1^{(k-1)},\ldots,P_{M+1}^{(k-1)})} \tag{25}$$

$$\Delta F_{12}^{(k)}\big(P_1^{(k)},\ldots,P_{2M}^{(k)},J\big)\equiv-\frac{1}{\beta x_{k+1}}\ln\int\left[\prod_{i=1}^{2M}P_i^{(k)}\,\mathrm{d}P_i^{(k-1)}\right]\mathrm{e}^{-\beta x_{k+1}\Delta F_{12}^{(k-1)}(P_1^{(k-1)},\ldots,P_{2M}^{(k-1)},J)}. \tag{26}$$

The above definitions have to be supplemented with the definitions for $k=0$ that read

$$\Delta F_1^{(0)}\big(P_1^{(0)},\ldots,P_{M+1}^{(0)}\big)\equiv-\frac{1}{\beta x_1}\ln\int\left[\prod_{i=1}^{M+1}P_i^{(0)}\,\mathrm{d}u_i\right]\left(\frac{2\cosh\beta\sum_{i=1}^{M+1}u_i}{\prod_{i=1}^{M+1}2\cosh\beta u_i}\right)^{x_1}. \tag{27}$$

$$\Delta F_{12}^{(0)}\big(P_1^{(0)},\ldots,P_{2M}^{(0)},J\big)\equiv-\frac{1}{\beta x_1}\ln\int\left[\prod_{i=1}^{2M}P_i^{(0)}\,\mathrm{d}u_i\right]\times\left(\frac{\sum_{\sigma,\tau}\exp\big[\beta\sum_{i=1}^{M}u_i\sigma+\beta\sum_{i=M+1}^{2M}u_i\tau+\beta J\sigma\tau\big]}{\prod_{i=1}^{2M}2\cosh\beta u_i}\right)^{x_1}. \tag{28}$$

4.1. *The functional $\Phi(n)$ in the replica-symmetry-breaking solution*

The variational free energy expressed in terms of the K-population $P^{(K)}$ that parameterizes $\rho(\sigma)$ reads

$$n\beta\Phi(n)=nM\beta\Delta F_1^{(K)}(P^{(K)},\ldots,P^{(K)})+\frac{M+1}{2}\ln\big\langle\mathrm{e}^{-\beta n\Delta F_{12}^{(K)}(P^{(K)},\ldots,P^{(K)},J)}\big\rangle \tag{29}$$

where the square brackets mean average over the coupling constant J and we have used the functions defined above. The proof that the above expression is equivalent to (9) is not very complicated and we will not report it. We just mention that it can be obtained in an iterative way much as we will do in the appendix for the variational equation. In order to determine the K-population that extremizes (29) we need to solve the corresponding variational equations obtained differentiating it with respect to $P^{(K)}$. An equivalent way to obtain $P^{(K)}$ is to consider the variational equation (10) and rewrite it in terms of $P^{(K)}$; we will show how to do this in the appendix. For practical purposes this second method is to be preferred because the corresponding equations can be solved by means of a population dynamics algorithm [26]; however, in order to study the small-n behaviour of $\Phi(n)$ it is more useful to consider that variational expression (29).

In general, the K-population that extremizes (29) depends on the value of n. In order to determine the expression of $\Phi(n)$ at small values of n we can expand expression (29) at the second order in n around $n=0$. This expression is variational in $P^{(K)}$ and therefore the total second derivative of $n\phi(n)$ with respect to n is equal to the partial second derivative. However, we must keep in mind that $\Delta F_1^{(K)}$ and $\Delta F_{12}^{(K)}$ both have an implicit dependence from n;

therefore, in order to derive with respect to n we make this dependence explicit by rewriting expression (29) as

$$n\beta\Phi(n) = -M \ln \int \left[\prod_{i=1}^{M+1} P^{(K)}\, \mathrm{d}P_i^{(K-1)}\right] \mathrm{e}^{-\beta n \Delta F_1^{(K-1)}(P_1^{(K-1)},\ldots,P_{M+1}^{(K-1)})} + \frac{M+1}{2} \ln \int \left[\prod_{i=1}^{2M} P^{(K)}\, \mathrm{d}P_i^{(K-1)}\right] \left\langle \mathrm{e}^{-\beta n \Delta F_{12}^{(K-1)}(P_1^{(K-1)},\ldots,P_{2M}^{(K-1)},J)}\right\rangle \tag{30}$$

Expanding in powers of n we get

$$\beta n\Phi(n) = n\left[M\langle\langle A_1\rangle\rangle - \frac{M+1}{2}\langle\langle A_{12}\rangle\rangle\right] - \frac{n^2}{2}\left[M\left(\langle\langle A_1^2\rangle\rangle - \langle\langle A_1\rangle\rangle^2\right) - \frac{M+1}{2}\left(\langle\langle A_{12}^2\rangle\rangle - \langle\langle A_{12}\rangle\rangle^2\right)\right] + O(n^3)$$

where we have defined

$$\langle\langle A_1^p\rangle\rangle \equiv \int \prod_{i=1}^{M+1} P^{(K)}\, \mathrm{d}P_i^{(K-1)} \left[\Delta F_1^{(K-1)}\left(P_1^{(K-1)},\ldots,P_{M+1}^{(K-1)}\right)\right]^p$$

$$\langle\langle A_{12}^p\rangle\rangle \equiv \int \prod_{i=1}^{2M} \left[P^{(K)}\, \mathrm{d}P_i^{(K-1)}\right] \left\langle\left[\Delta F_{12}^{(K-1)}\left(P_1^{(K-1)},\ldots,P_{2M}^{(K-1)},J\right)\right]^p\right\rangle$$

Thus, we conclude that the coefficient c_2 of the $O(n^2)$ term in the small-n power series of $n\Phi(n)$ is given by the coefficient of the above $O(n^2)$ term evaluated with the population $P^{(K)}$ corresponding to $n = 0$. Therefore, in general c_2 will be non-zero and we will expect Gaussian fluctuations with variance given by equation (19).

We note that the physical interpretation of the K-RSB ansatz is that on a given sample (i.e. a graph with a given disorder realization) the local fields on a given sites are described by a $K-1$-population and the population $P^{(K)}$ represents the distribution over the sites of these $K-1$-populations [26]. The functions $\Delta F_1^{(K-1)}$ and $\Delta F_{12}^{(K-1)}$ are interpreted as the free-energy variations (shifts) that are observed in the process of adding respectively a spin and a bond to a given graph [26]. Thus, the interpretation of the above equation is that *the Gaussian fluctuations of the free energy are determined by the local fluctuations of the free-energy shifts*.

4.2. The factorized solution

The results of the preceding subsection tell us that the $O(n^2)$ term will be absent only if there are no spatial fluctuations in the distribution of the fields. This corresponds to the fact that $P^{(K)}$ is given by the so-called *factorized solution* $P^{(K)} = \delta\left(P^{(K-1)} - P_0^{(K-1)}\right)$. Indeed it is easy to check that if $P^{(K)}$ is factorized the term $\left(\langle\langle A_1^2\rangle\rangle - \langle\langle A_1\rangle\rangle^2\right)$ vanishes. In order to obtain this as well the term $\left(\langle\langle A_{12}^2\rangle\rangle - \langle\langle A_{12}\rangle\rangle^2\right)$ vanishes and we also need the condition $J = \pm\tilde{J}$ (in the high-temperature phase the $O(n^2)$ term is determined solely by the fluctuations of J).

The variational equation for $P^{(K)}$ reported in the appendix shows that $P^{(K)}$ can be factorized only if the couplings have a bimodal distribution $J = \pm\tilde{J}$. We argue that in this case the correct distribution is indeed factorized. Note that for $K = 1$ a non-factorized solution was found in [26] for the bimodal case. We believe that this is an artefact because the correct solution has an infinite number of RSB steps $K = \infty$. This is similar to what happens for the function $q(x)$ of the SK model: in the 1RSB ansatz we find $q(0) \neq 0$ while $q(0) = 0$ in the full-RSB solution [18].

In the bimodal case one can see that the factorized solution $P^{(K)} = \delta\big(P^{(K-1)} - P_0^{(K-1)}\big)$ is such that $P_0^{(K-1)}$ is independent of n. As a consequence $\Phi(n)$ is constant in n. Likewise in the SK model we argue that $\Phi(n)$ is constant only for $n < 0$, while for $n > 0$ we have to abandon the factorized solution corresponding to $n = 0$ because of the condition that the smallest RSB parameter $x_{\min}$ must be larger than n. In order to study this effect and to determine the first non-trivial term of $\Phi(n)$ for positive n in the following section we will study an expansion of $\Phi(n)$ in the order parameter.

5. Expansion of $\Phi(n)$ in powers of the order parameter

In this section we report the expansion of the potential $\Phi(n)$ of the spin-glass defined on the Bethe lattice with fixed connectivity $M + 1$ in powers of the order parameter. Note that following [25] we will use a different normalization of $\rho(\sigma)$ with respect to the previous subsection; we write it as

$$\rho(\{\sigma\}) = \sum_{k=0}^{n} b_k \sum_{(\alpha_1\ldots\alpha_k)} q_{\alpha_1\ldots\alpha_k}\sigma_{\alpha_1}\ldots\sigma_{\alpha_k}, \tag{31}$$

with

$$b_k \equiv \langle\cosh^n \beta J \tanh^k \beta J\rangle \qquad \text{and} \qquad \tilde{b}_k \equiv b_k/b_0. \tag{32}$$

The variational equation reads

$$\rho(\{\sigma\}) = \frac{\mathrm{Tr}_{\{\tau\}}\rho^M(\tau)\big\langle \exp\beta J \sum_\alpha \sigma_\alpha\tau_\alpha\big\rangle}{\mathrm{Tr}_{\{\tau\}}\rho^M(\tau)} \tag{33}$$

and when expressed in terms of the $q_{\alpha_1\ldots\alpha_k}$ it reads

$$q_{\alpha_1\ldots\alpha_k} = \frac{\mathrm{Tr}_{\{\sigma\}}\sigma_{\alpha_1}\ldots\sigma_{\alpha_k}\rho^M(\sigma)}{\mathrm{Tr}_{\{\sigma\}}\rho^M(\sigma)}. \tag{34}$$

We have expanded expression (9) in powers of the order parameter q_{ab} at fourth order. The four-index order parameter q_{abcd} has been expressed in terms of q_{ab} by means of its variational equation. In the appendix we give some details while here we report the results:

$$n\beta\Phi = -\frac{M+1}{2}\ln b_0 - n\ln 2 + F_{\mathrm{var}}. \tag{35}$$

where

$$F_{\mathrm{var}} \equiv -\frac{\tau}{2}\,\mathrm{Tr}\,q^2 - \frac{\omega}{6}\,\mathrm{Tr}\,q^3 - \frac{v}{8}\,\mathrm{Tr}\,q^4 + \frac{y}{4}\sum_{abc} q_{ab}^2 q_{ac}^2 - \frac{u}{12}\sum_{ab} q_{ab}^4 + \frac{\tilde{y}}{4}(\mathrm{Tr}\,q^2)^2 + O(q^5) \tag{36}$$

where the various coefficients depend on b_k with $k = 2, 4$, see their explicit expressions below.

In the high-temperature region we have $q = 0$ and therefore $F_{\mathrm{var}} = 0$; thus, the only relevant term is the first term in equation (35). If J can take just two values $J = \pm\tilde{J}$ we have $b_0 = \cosh^n \beta\tilde{J}$; therefore, for $T > T_c$ $\Phi(n)$ is just a constant in n. From a high-temperature expansion we can verify that in this case the fluctuations of the extensive free-energy scale as $\Delta F = O(1)$. In contrast, if the distribution of J is not just peaked at $\pm\tilde{J}$, $\Phi(n)$ is linear in n and the fluctuations of the extensive free energy are Gaussian with $\Delta F = O(\sqrt{N})$. Therefore, if the graph is locally homogeneous the first term immediately yields small $O(1)$ fluctuations above the critical temperature.

Below the critical temperature we have to check whether $n\Phi(n)$ is quadratic in n around $n = 0$. We note that the various coefficients $\{\tau, \omega, u, v, y\}$ in equation (36) depend on n and on the temperature; however, we first study the model with fixed coefficients and variable number of replicas n; we also reabsorb the coefficient $\tilde{y}$ in y, which is possible if we restrict the choice of the matrix q_{ab} to those that verify the condition that $(q^2)_{aa}$ does not depend on a. This produces an additional dependence of y from n because $y + n\,\tilde{y} \to y$ but, like the other coefficients, we first consider it as independent of n.

We have computed the free energy of this model, basically generalizing Kondor's original computation to general values of the coefficients and including all the fourth-order terms. We have found that similarly as in the SK model: (i) the first nonlinear term in the expansion of $n\,\Phi(n)$ for $n \geqslant 0$ is $O(n^6)$, due to non-trivial cancellations at order $O(n^2)$ and $O(n^4)$ and (ii) $n\,\Phi(n)$ is just linear in n for $n < 0$. More explicitly we have for $n > 0$

$$F_{\text{var}} = n\left(\frac{\tau^3}{6\omega^2} + \frac{2u + 9v + 6y}{24\omega^4}\tau^4 + O(\tau^5)\right) \pm \frac{9}{5120}n^6\left(40\frac{\omega^7}{u^6} - 75\frac{\omega^8}{u^7} + 36\frac{\omega^9}{u^8} + O(\tau)\right) + O(n^7) \qquad \text{for} \quad n > 0. \tag{37}$$

In the SK model we have $\omega = u = v = y = 1$ and we recover Kondor's value $-9/5120$ for the $O(n^6)$ coefficient. Now to study the actual n-dependence of the model we have to take into account the fact that the coefficients $\{\tau, \omega, u, v, y\}$ depend on n; indeed the various coefficients read

$$\begin{aligned}
\tau &= \frac{1}{2}M(1+M)\tilde{b}_2^2(-1+M\tilde{b}_2).\\
\omega &= M(-1+M^2)\tilde{b}_2^3(-2+3M\tilde{b}_2).\\
v &= \frac{M(-1+M^2)\tilde{b}_2^4(3(-2+M)+M(-1+4M)\tilde{b}_4+2M\tilde{b}_2(5-3M+(-3+M)M\tilde{b}_4))}{-1+M\tilde{b}_4}.\\
y &= -\frac{M(-1+M^2)\tilde{b}_2^4(6-3M+(8-11M)M\tilde{b}_4+M\tilde{b}_2(-9+5M+M(1+3M)\tilde{b}_4))}{-1+M\tilde{b}_4}.\\
u &= -\frac{M(-1+M^2)\tilde{b}_2^4(-6(-2+M)+3(3-5M)M\tilde{b}_4+4M\tilde{b}_2(-2+M+M(-1+2M)\tilde{b}_4))}{(-1+M\tilde{b}_4)}.\\
\tilde{y} &= \frac{M(1+M)\tilde{b}_2^4(-3+4M+(2-3M)M^2\tilde{b}_4+M^3\tilde{b}_2^2(M\tilde{b}_4-1)+4(M-1)M\tilde{b}_2(M^2\tilde{b}_4-1))}{4(M\tilde{b}_4-1)}.
\end{aligned} \tag{38}$$

In order to recover the SK limit of the above coefficients, we have to rescale the couplings as $J = \tilde{J}/M^{1/2}$ where $\tilde{J}$ is a random variable with unit variance and take the limit $M \to \infty$. We obtain $\tau = 1 - T + O(1-T)^2$ and $\omega = u = v = y = 1$, while $\tilde{y} = 0$, i.e. as it should.

Each coefficient depends on β and n through $\tilde{b}_k$ and we have

$$\tilde{b}_k = \langle \tanh^k \beta J\rangle + n(\langle \ln\cosh\beta J \tanh^k \beta J\rangle - \langle \ln\cosh\beta J\rangle\langle \tanh^k \beta J\rangle) + O(n^2). \tag{39}$$

Therefore, the coefficients for a generic distribution of the J's have a linear dependence on n that, included in expression (37), leads to an $O(n^2)$ dependence of $n\Phi(n)$. Note that the coefficients have a dependence from n that is regular around $n = 0$ and therefore the first term in equation (37) is regular in n around $n = 0$, but there still is the $O(n^6)$ term which produces a non-regular dependence from n around $n = 0$. In other words in the general diluted model the function $n\phi(n)$ develops a regular $O(n^2)$ dependence but there still is a singularity at $n = 0$ in the sixth derivative. Near the critical temperature the leading $O(n^2)$ term is given by the n dependence of τ in equation (37) and therefore is $O(\tau^2)$; expanding equation (38) in powers of n we get for the $O(n^2)$ term of F_{var}:

$$n^2\left(\frac{M^4}{4(M-1)^2(M+1)}\tau_0^2(\langle\ln\cosh\beta_c J\tanh^2\beta_c J\rangle-\langle\ln\cosh\beta_c J\rangle\langle\tanh^2\beta_c J\rangle)+O\left(\tau_0^3\right)\right) \tag{40}$$

where τ_0 is given by equation (38) computed at $n = 0$.

We consider now the locally homogeneous case in which $J = \pm\tilde{J}$ with equal probability. We note first that in this case the coefficients $\tilde{b}_k$ no more depend on n (see equation (32)) and as a consequence τ, ω, u and v also do not depend on n. Thus, the only dangerous coefficient is y in which we have reabsorbed the coefficient $\tilde{y}$ through $y + n\,\tilde{y} \to y$. The $\tilde{y}$ coefficient turns out to be zero and therefore the behaviour of the model is the same as that of the SK model, the first nonlinear term in $n\phi(n)$ being $O(n^6)$. To be more precise we have checked that there are no $O(n^2)$ terms in $n\phi(n)$ at the first non-trivial order in τ, because $\tilde{y}$ defined according to equation (38) actually is zero only at the critical temperature where $\tilde{b}_2 = 1/M$ and $\tilde{b}_4 = 1/M^2$ (because $J = \pm\tilde{J}$) but has small non-zero corrections $O(\tau)$ at higher orders. These $O(n^2)$ corrections are likely cancelled by the $O(n^2)$ higher order term $\operatorname{Tr} Q^2 \operatorname{Tr} Q^3$ not included in the computation.

In the appendices we report a similar expansion for the variational equation and various quantities relevant for the computation.

6. Conclusion

We have investigated the large deviations free-energy functional $\Phi(n)$ in the case of the Bethe lattice spin-glass and we have confirmed that Gaussian behaviour of the free-energy fluctuations has to be expected whenever there is no local homogeneity of the interactions. In particular, the only case in which we found a non-Gaussian SK-like behaviour when the random couplings can take only two possible opposite values $J_{ij} = \pm\tilde{J}$ with equal probability.

In general the quantity $\Phi(n)$ can be expressed in terms of a distribution of fields. In the RS case we have a single distribution corresponding to the possible values of the cavity fields at different sites of the lattice for a given disorder realization. In the RSB phase we have a population of populations, i.e. on each site we have a population of fields corresponding to the presence of many states. We have found that if the system is locally homogeneous we can find a locally homogeneous distribution of the fields and this leads to the vanishing of the $O(n^2)$ term in $\Phi(n)$. Thus, we argue that the correct RSB solution in the bimodal case is the so-called factorized solution. Instead if the system is not locally homogeneous the $O(n^2)$ terms in $\Phi(n)$ can be evaluated using the $n = 0$ solution because of stationarity.

We have also verified that the expansion in power of the order parameter near the critical temperature in the locally homogeneous case is formally equivalent to that of the SK model and found that $\Phi(n)$ has the same $O(n^5)$ behaviour of SK for small positive n.

We note that the fact that in the bimodal case $n\Phi(n) = nf_{\text{typ}}$ for $n < 0$ has some interesting consequences. Indeed since $\Phi(n)$ is the Legendre transform of the large deviations function $L(f)$ (see equations (5)–(8)) it follows that $L(f) = -\infty$ for free energies *per spin* larger than the typical one f_{typ}. This means that the probability of finding a sample with $f > f_{\text{typ}}$ is smaller than $\exp[O(N)]$. Indeed for the SK model a recent computation [15] has shown that $P(f) \propto \exp[O(N^2)]$. This scaling cannot hold for the Bethe lattice because while the total number of samples is actually $\exp[O(N^2)]$ in the SK model, the total number of samples on the Bethe lattice is $\exp[(M+1)N\ln N]$ at leading order. Thus, we argue that in the Bethe lattice with bimodal distribution of the couplings $P(f) \propto \exp[O(N\ln N)]$ for $f > f_{\text{typ}}$ although the actual computation is beyond the scope of this work. For $M = 1$ detailed computations are easy. Nevertheless, we note that free energies larger than the typical one can only be observed on graphs with topologies different from the typical one (e.g. a regular lattice). In other words, *the probability of observing a free energy (and in particular a ground-state energy) larger than the typical one on a graph with typical topology is strictly zero*. Indeed suppose that by just changing the signs of the interactions of a typical graph (i.e. without modifying the incidence matrix) we could raise the free energy *per spin*. Since the number of links on a graph is precisely $\frac{M+1}{2}N$ the probability of such a sample will be $\exp[O(N)]$ and this would lead to a non-constant $\Phi(n)$ for $n < 0$ contrarily to what we have computed.

Appendix A. The variational equations in terms of populations

In this appendix we write the variational equations in terms of populations. These equations have been obtained at the level of one-step RSB in [26] using the cavity method. In the following we write them down for a generic number of RSB steps using the replica method. The variational equation that extremizes the free energy (9) reads

$$\rho(\sigma) = \frac{\langle \mathrm{Tr}_\tau \rho^M(\tau) \exp J\sigma\tau \rangle}{\langle \mathrm{Tr}_{\tau,\sigma} \rho^M(\tau) \exp \beta J\sigma\tau \rangle}. \tag{A.1}$$

In terms of the K-population the above equation reads

$$P^{(K)} \equiv \frac{1}{\langle \mathrm{e}^{-\beta n \Delta F^{(K)}(P^{(K)},\dots,P^{(K)},J)} \rangle} \int \left[\prod_{i=1}^M P^{(K)} \, \mathrm{d}P_i^{(K-1)} \right] \times \left\langle \delta(P^{(K-1)} - \tilde{P}^{(K-1)}) \, \mathrm{e}^{-\beta n \Delta F^{(K-1)}(P_1^{(K-1)},\dots,P_M^{(K-1)},J)} \right\rangle \tag{A.2}$$

where the square brackets mean average over the disorder. In the above equation we have used the following functions of populations: (i) a function $\tilde{P}^{(k)}\left[P_1^{(k)},\dots,P_M^{(k)},J\right]$ that yields a k-population from M other k-populations and (ii) a function $\Delta F^{(k)}\left[P_1^{(k)},\dots,P_M^{(k)},J\right]$ (also called the free-energy shift [26]) that yields a real number from M k-populations. The definition is iterative: the functions $\tilde{P}^{(k)}$ and $\Delta F^{(k)}$ at level k of RSB are defined starting from the functions $\tilde{P}^{(k-1)}$ and $\Delta F^{(k-1)}$:

$$\tilde{P}^{(k)}\left(P_1^{(k)},\dots,P_M^{(k)},J\right) \equiv \frac{1}{\mathrm{e}^{-\beta x_{k+1}\Delta F^{(k)}\left(P_1^{(k)},\dots,P_M^{(k)},J\right)}} \int \left[\prod_{i=1}^M P_i^{(k)} \, \mathrm{d}P_i^{(k-1)}\right] \delta(P^{(k-1)} - \tilde{P}^{(k-1)}) \times \mathrm{e}^{-\beta x_{k+1}\Delta F^{(k-1)}(P_1^{(k-1)},\dots,P_M^{(k-1)},J)} \tag{A.3}$$

and

$$\Delta F^{(k)}\left(P_1^{(k)},\dots,P_M^{(k)},J\right) = -\frac{1}{\beta x_{k+1}} \ln \int \left[\prod_{i=1}^M P_i^{(k)} \, \mathrm{d}P_i^{(k-1)}\right] \mathrm{e}^{-\beta x_{k+1}\Delta F^{(k-1)}(P_1^{(k-1)},\dots,P_M^{(k-1)},J)}. \tag{A.4}$$

The iterative definition has to be supplemented with the $k = 0$ case that reads

$$\tilde{P}^{(0)}\left(P_1^{(0)},\dots,P_M^{(0)},J\right) \equiv \frac{1}{\mathrm{e}^{-\beta x_1 \Delta F^{(0)}(P_1^{(0)},\dots,P_M^{(0)},J)}} \int \left[\prod_{i=1}^M P_i^{(0)} \, \mathrm{d}u_i\right] \times \left(\frac{4\cosh\beta J \cosh\beta \sum_i u_i}{\prod_{i=1}^M 2\cosh\beta u_i}\right)^{x_1} \delta\left(u - \tilde{u}\left(\sum_i u_i, J\right)\right) \tag{A.5}$$

and

$$\Delta F^{(0)}\big(P_1^{(0)},\ldots,P_M^{(0)},J\big)\equiv-\frac{1}{\beta x_1}\ln\int\left[\prod_{i=1}^{M}P_i^{(0)}\,\mathrm{d}u_i\right]\left(\frac{4\cosh\beta J\cosh\beta\sum_i u_i}{\prod_{i=1}^{M}2\cosh\beta u_i}\right)^{x_1} \qquad (A.6)$$

where we used the definition [26]

$$\tilde{u}(h,J)=\frac{1}{\beta}\,\mathrm{arctanh}[\tanh\beta J\tanh\beta h]. \qquad (A.7)$$

We recall also the relationship between the populations and $\rho(\sigma)$:

$$\rho_{P^{(k)}}(\sigma)=\int P^{(k)}\,\mathrm{d}P^{(k-1)}\prod_{\mathcal{C}=1}^{x_{k+1}/x_k}\rho_{P^{(k-1)}}(\{\sigma\}_{\mathcal{C}}). \qquad (A.8)$$

In the following we will prove the equivalence between equation (A.1) and equation (A.2). Our basic step is to prove that the following fundamental equation holds at any level k:

$$\left[\mathrm{Tr}_{\tau_{\mathcal{C}}}\left(\prod_{i=1}^{M}\rho_{P_i^{(k)}}(\tau_{\mathcal{C}})\right)\exp\beta J\sigma_{\mathcal{C}}\tau_{\mathcal{C}}\right]=\mathrm{e}^{-\beta x_{k+1}\Delta F^{(k)}\left(P_1^{(k)},\ldots,P_M^{(k)},J\right)}\rho_{\tilde{P}^{(k)}}(\sigma_{\mathcal{C}}). \qquad (A.9)$$

In the above equations $\sigma_{\mathcal{C}}$ and $\tau_{\mathcal{C}}$ are two sets of x_{k+1} spins and $\tau_{\mathcal{C}}\sigma_{\mathcal{C}}=\sum_{a=1}^{x_{k+1}}\sigma_a\tau_a$. The proof is iterative: assuming that the equation is satisfied at level $k-1$ we will show that it is also satisfied at level k. In order to do that we divide the x_{k+1} spins $\sigma_{\mathcal{C}}$ in x_{k+1}/x_k groups $\sigma_{\mathcal{C}'}$ of x_k spins and we use definition (A.8):

$$\begin{aligned}\mathrm{Tr}_{\tau_{\mathcal{C}}}\left(\prod_{i=1}^{M}\rho_{P_i^{(k)}}(\tau_{\mathcal{C}})\right)\exp\beta J\sigma_{\mathcal{C}}\tau_{\mathcal{C}}&=\mathrm{Tr}_{\tau_{\mathcal{C}}}\left(\prod_{i=1}^{M}\int P_i^{(k)}\,\mathrm{d}P_i^{(k-1)}\prod_{\mathcal{C}'=1}^{x_{k+1}/x_k}\rho_{P_i^{(k-1)}}(\tau_{\mathcal{C}'})\right)\exp\beta J\sigma_{\mathcal{C}}\tau_{\mathcal{C}}\\&=\int\left[\prod_{i=1}^{M}P_i^{(k)}\,\mathrm{d}P_i^{(k-1)}\right]\prod_{\mathcal{C}'=1}^{x_{k+1}/x_k}\mathrm{Tr}_{\tau_{\mathcal{C}'}}\left[\left(\prod_{i=1}^{M}\rho_{P_i^{(k-1)}}(\tau_{\mathcal{C}'})\right)\exp\beta J\sigma_{\mathcal{C}'}\tau_{\mathcal{C}'}\right].\end{aligned} \qquad (A.10)$$

Now assuming that equation (A.9) holds true at level $k-1$ and integrating over a delta function $\delta(P^{(k-1)}-\tilde{P}^{(k-1)})$ we get

$$\mathrm{Tr}_{\tau_{\mathcal{C}}}\left(\prod_{i=1}^{M}\rho_{P_i^{(k)}}(\tau_{\mathcal{C}})\right)\exp\beta J\sigma_{\mathcal{C}}\tau_{\mathcal{C}} \qquad (A.11)$$

$$=\int\mathrm{d}P^{(k-1)}\left\{\left[\prod_{i=1}^{M}P_i^{(k)}\,\mathrm{d}P_i^{(k-1)}\right]\delta(P^{(k-1)}-\tilde{P}^{(k-1)})\,\mathrm{e}^{-\beta x_{k+1}\Delta F^{(k-1)}(P_1^{(k-1)},\ldots,P_M^{(k-1)},J)}\right\}$$
$$\times\prod_{\mathcal{C}'=1}^{x_{k+1}/x_k}\rho_{P^{(k-1)}}(\sigma_{\mathcal{C}'}) \qquad (A.12)$$

we see that the term in curly brackets corresponds to the one in the definition (A.3), and using definition (A.8) we conclude that equation (A.9) holds true at level k.

In order to complete the proof we need to show that equation (A.9) holds for $k=0$. In this case $\sigma_{\mathcal{C}}$ is a group of x_1 spins, using equation (A.10) we have

$$\mathrm{Tr}_{\tau_{\mathcal{C}}}\left(\prod_{i=1}^{M}\rho_{P_i^{(0)}}(\tau_{\mathcal{C}})\right)\exp\beta J\sigma_{\mathcal{C}}\tau_{\mathcal{C}}=\int\left[\prod_{i=1}^{M}P_i^{(0)}\,\mathrm{d}u_i\right]\prod_{a=1}^{x_1}\sum_{\tau_a}\left[\left(\prod_{i=1}^{M}\frac{\exp\beta u_i\tau_a}{2\cosh\beta u_i}\right)\exp\beta J\sigma_a\tau_a\right]$$

now summing over each τ_a and introducing a delta function $\delta\big(u-\tilde{u}\big(\sum_{i=1}^{M}u_i,J\big)\big)$ and using the definitions (A.5) and (A.6) we can see that equation (A.9) holds true also at level $k=0$. The equation (A.9) can now be used to prove the equivalence between (A.1) and (A.2).

Appendix B. The order-parameter equation

In this appendix we report an order parameter expansion of the variational equation (34). Expanding equation (34) for q_{abcd} in powers of the order parameters we get (see appendices C and D):

$$q_{abcd} = \frac{M(M-1)}{1-M\tilde{b}_4}\tilde{b}_2^2(q_{ab}q_{cd}+q_{ac}q_{db}+q_{ad}q_{cb})+O(q^3) \tag{B.1}$$

Substituting this expression in equation (34) for q_{ab} we get at the third order in the order parameter q_{ab}:

$$0 = c_1 q_{ab} + c_2(q^2)_{ab} + c_{3,1}(q^3)_{ab} + c_{3,2}q_{ab}((q^2)_{bb}+(q^2)_{aa}) + c_{3,3}q_{ab}^3 + c_{3,4}q_{ab}\,\mathrm{Tr}\,Q^2 \tag{B.2}$$

$$\begin{aligned}
c_1 &= M\tilde{b}_2 - 1\\
c_2 &= \tilde{b}_2^2(M^2-M)\\
c_{3,1} &= -\tilde{b}_2^3\frac{(M-1)M(M\tilde{b}_4+M-2)}{M\tilde{b}_4-1}\\
c_{3,2} &= \tilde{b}_2^3\frac{(M-1)M(M^2\tilde{b}_4+M-2)}{M\tilde{b}_4-1}\\
c_{3,3} &= -\frac{2\tilde{b}_2^3}{3}\frac{(M-1)M(M(2M-1)\tilde{b}_4+M-2)}{M\tilde{b}_4-1}\\
c_{3,4} &= -\frac{\tilde{b}_2^3}{2}\frac{(M-1)M(M^2\tilde{b}_4-1)}{M\tilde{b}_4-1}.
\end{aligned}$$

The coefficients of the previous expansion are different from what could be obtained by differentiating the variational expression equation (35) derived above. This can be understood noting that the equation for the order parameter corresponds to the following expression:

$$0 = \mathrm{Tr}\left[\sigma_a\sigma_b\left(\rho(\sigma) - \frac{\mathrm{Tr}_\tau\,\rho^M(\tau)\langle\exp J\sum_c\sigma_c\tau_c\rangle}{\mathrm{Tr}\,\rho^M(\tau)}\right)\right] \tag{B.3}$$

while the equation one obtains by differentiating equation (9) corresponds to

$$0 = \mathrm{Tr}\left[\rho^{M-1}(\{\sigma\})\sigma_a\sigma_b\left(\rho(\sigma) - \frac{\mathrm{Tr}_\tau\,\rho^M(\tau)\langle\exp J\sum_c\sigma_c\tau_c\rangle}{\mathrm{Tr}\,\rho^M(\tau)}\right)\right]. \tag{B.4}$$

Thus, the two expressions are equivalent in the sense that they have the same solution at the order at which they are valid. It can be checked explicitly that the coefficient $c_{3,4}$ (as much as $a_{4,4}$) vanishes at zeroth order in the expansion in τ, noting that at $T = T_c$ we have $\tilde{b}_2 = 1/M$ and $\tilde{b}_4 = 1/M^2$ (because $J = \pm\tilde{J}$).

In the Sherrington–Kirkpatrick limit $M \to \infty$ and $J = (\tilde{J})/\sqrt{M}$ with $\overline{J^2} = 1$ the coefficients of the order parameter equation go to the corresponding SK limit as can also be seen noting that in this limit equation (B.3) reduces to the corresponding SK equation:

$$q_{ab} = \frac{\mathrm{Tr}\,\sigma_a\sigma_b\exp\left[\beta^2\sum_{a<b}Q_{ab}\sigma_a\sigma_b\right]}{\mathrm{Tr}\exp\left[\beta^2\sum_{a<b}Q_{ab}\sigma_a\sigma_b\right]} \tag{B.5}$$

Appendix C. Traces of $\rho(\sigma)$

In this section we report various quantities that are relevant to compute the expansions in the order parameter. We define

$$\rho(\sigma) = b_0(1 + \tilde{g}(\sigma)) \tag{C.1}$$

$$\tilde{g}(\sigma) = \tilde{b}_2 \sum_{a<b} q_{ab}\sigma_a\sigma_b + \tilde{b}_4 \sum_{a<b<c<d} q_{abcd}\sigma_a\sigma_b\sigma_c\sigma_d + \cdots. \tag{C.2}$$

Then the following traces are needed to compute the expansion of the variational expression of $\Phi(n)$. In order to compute them we need also the traces over spin reported in the next appendix.

$$\frac{1}{2^n}\operatorname{Tr}\tilde{g} = 0 \tag{C.3}$$

$$\frac{1}{2^n}\operatorname{Tr}\tilde{g}^2 = \frac{\tilde{b}_2^2}{2}\operatorname{Tr}q^2 + \tilde{b}_4^2\left(\frac{M(M-1)}{1-M\tilde{b}_4}\tilde{b}_2^2\right)^2 \times \left(\frac{1}{8}(\operatorname{Tr}q^2)^2 + \frac{1}{4}\operatorname{Tr}q^4 - \sum_{abc} q_{ab}^2 q_{ac}^2 + \frac{1}{2}\sum_{ab} q_{ab}^4\right) + O(q^5) \tag{C.4}$$

$$\frac{1}{2^n}\operatorname{Tr}\tilde{g}^3 = \tilde{b}_2^3\operatorname{Tr}q^3 + 3\tilde{b}_4\tilde{b}_2^2\left(\frac{M(M-1)}{1-M\tilde{b}_4}\tilde{b}_2^2\right) \times \left(\frac{1}{4}(\operatorname{Tr}q^2)^2 + \operatorname{Tr}q^4 - 4\sum_{abc} q_{ab}^2 q_{ac}^2 + 2\sum_{ab} q_{ab}^4\right) + O(q^5) \tag{C.5}$$

$$\frac{1}{2^n}\operatorname{Tr}\tilde{g}^4 = \tilde{b}_2^4\left(\frac{3}{4}(\operatorname{Tr}q^2)^2 + 3\operatorname{Tr}q^4 - 6\sum_{abc} q_{ab}^2 q_{ac}^2 + 4\sum_{ab} q_{ab}^4\right) + O(q^5). \tag{C.6}$$

In order to sum over q_{abcd} in $\operatorname{Tr}\tilde{g}^2$ we used the following identity valid for a general A_{abcd} symmetric with respect to permutations of its indices

$$\sum_{a<b<c<d} A_{abcd} = \frac{1}{24}\left(\sum_{abcd} A_{abcd} - 6\sum_{abc} A_{aabc} + 3\sum_{ab} A_{aabb} + 8\sum_{ab} A_{aaab} - 6\sum_{a} A_{aaaa}\right). \tag{C.7}$$

The following traces are needed to compute the expansion of the equation for the order parameter:

$$\frac{1}{2^n}\operatorname{Tr}\sigma_a\sigma_b\tilde{g} = \tilde{b}_2 q_{ab} \tag{C.8}$$

$$\frac{1}{2^n}\operatorname{Tr}\sigma_a\sigma_b\tilde{g}^2 = 2\tilde{b}_2^2(q^2)_{ab} + 2\tilde{b}_2\tilde{b}_4\left(\frac{M(M-1)}{1-M\tilde{b}_4}\tilde{b}_2^2\right) \times \left((q^3)_{ab} - 2q_{ab}((q^2)_{aa} + (q^2)_{bb}) + 2q_{ab}^3 + \frac{1}{2}q_{ab}\operatorname{Tr}q^2\right) + O(q^4) \tag{C.9}$$

$$\frac{1}{2^n}\operatorname{Tr}\sigma_a\sigma_b\tilde{g}^3 = \tilde{b}_2^3\left(6(q^3)_{ab} - 6q_{ab}((q^2)_{aa} + (q^2)_{bb}) + 4q_{ab}^3 + \frac{3}{2}q_{ab}\operatorname{Tr}q^2\right) + O(q^4) \tag{C.10}$$

$$b_0^M / \mathrm{Tr}\, \rho^M = 1 - \frac{1}{4} M(M-1)\tilde{b}_2^2 \, \mathrm{Tr}\, q^2 + O(q^3). \tag{C.11}$$

The next traces are needed to compute the equation for q_{abcd}:

$$\frac{1}{2^n} \mathrm{Tr}\, \sigma_a \sigma_b \sigma_c \sigma_d \tilde{g} = \tilde{b}_4 q_{abcd} \tag{C.12}$$

$$\frac{1}{2^n} \mathrm{Tr}\, \sigma_a \sigma_b \sigma_c \sigma_d \tilde{g}^2 = 2\tilde{b}_2^2 (q_{ab} q_{cd} + q_{ac} q_{bd} + q_{ad} q_{cb}) + O(q^3) \tag{C.13}$$

Appendix D. Spin traces

In the following we report the values of traces over the spins. They have been computed using the following general formula

$$\frac{1}{2^n} \mathrm{Tr}\, \sigma_a \sigma_b \sigma_c \sigma_d \ldots \sigma_e \sigma_f \sigma_g \sigma_h = \sum_\pi \delta_{ab} \delta_{cd} \ldots \delta_{ef} \delta_{gh} - 2 \sum_\pi \delta_{abcd} \ldots \delta_{ef} \delta_{gh}$$
$$+ 16 \sum_\pi \delta_{abcdef} \ldots \delta_{gh} + 4 \sum_\pi \delta_{abcd} \delta_{efgh} \ldots + \cdots \tag{D.1}$$

The above expression represents the fact that each of the spins $\sigma_a, \sigma_b, \ldots$ must appear an even number of times in order for the trace to be non-zero. The first term describes the case in which each spin appears just two times in the sum and the index π runs over all different permutations of the indices that change $\delta_{ab}\delta_{cd} \ldots \delta_{ef}\delta_{gh}$. The second term describes the case in which one spin appears four times and all the other appear two times. However, if this is the case the first term also gives a non-zero contribution; for this reason the second term has the factor -2 in front of it, because the l.h.s. of (D.1) is either 0 or 1. Again the index π runs over all permutations of the indices that change the summand. The third term corresponds to the case in which one spin appears six times in the sum, while the fourth corresponds to the case in which two different spins appears four times each in the sum. To give an example, in the case of four spins expression (D.1) specializes to

$$\frac{1}{2^n} \mathrm{Tr}\, \sigma_a \sigma_b \sigma_c \sigma_d = \delta_{ab}\delta_{cd} + \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{cb} - 2\delta_{abcd}. \tag{D.2}$$

Using these expression to couple the replica indices we get

$$\frac{1}{2^n} \mathrm{Tr} \left(\sum_{ab} q_{ab} \sigma_a \sigma_b \right) = 0 \tag{D.3}$$

$$\frac{1}{2^n} \mathrm{Tr} \left(\sum_{ab} q_{ab} \sigma_a \sigma_b \right)^2 = 2\, \mathrm{Tr}\, q^2 \tag{D.4}$$

$$\frac{1}{2^n} \mathrm{Tr} \left(\sum_{ab} q_{ab} \sigma_a \sigma_b \right)^3 = 8\, \mathrm{Tr}\, q^3 \tag{D.5}$$

$$\frac{1}{2^n} \mathrm{Tr} \left(\sum_{ab} q_{ab} \sigma_a \sigma_b \right)^4 = 48\, \mathrm{Tr}\, q^4 - 96 \sum_{abc} q_{ab}^2 q_{ac}^2 + 64 \sum_{ab} q_{ab}^4 + 12 (\mathrm{Tr}\, q^2)^2. \tag{D.6}$$

Other traces necessary to the expansions are

$$\frac{1}{2^n}\,\mathrm{Tr}\left(\sum_{ab}(q^2)_{ab}\sigma_a\sigma_b\right)\left(\sum_{ab}q_{ab}\sigma_a\sigma_b\right)^2 = 4\mathrm{Tr}\,q^4 - 8\sum_{abc}q_{ab}^2q_{ac}^2 + 2(\mathrm{Tr}\,q^2)^2 \tag{D.7}$$

$$\frac{1}{2^n}\,\mathrm{Tr}\left(\sum_{mn}q_{mn}\sigma_m\sigma_n\right)\sigma_a\sigma_b = 2q_{ab} \tag{D.8}$$

$$\frac{1}{2^n}\,\mathrm{Tr}\left(\sum_{mn}q_{mn}\sigma_m\sigma_n\right)^3\sigma_a\sigma_b = 48(q^3)_{ab} - 48q_{ab}((q^2)_{aa}+(q^2)_{bb}) + 32q_{ab}^3 + 12q_{ab}\,\mathrm{Tr}\,q^2 \tag{D.9}$$

$$\frac{1}{2^n}\,\mathrm{Tr}\left(\sum_{mn}(q^2)_{mn}\sigma_m\sigma_n\right)\left(\sum_{mn}q_{mn}\sigma_m\sigma_n\right)\sigma_a\sigma_b = 8(q^3)_{ab} + 2q_{ab}\mathrm{Tr}\,q^2 - 4q_{ab}((q^2)_{aa}+(q^2)_{bb}). \tag{D.10}$$

References

[1] Boettcher S 2004 *Europhys. Lett.* **67** 453
[2] Bouchaud J-P, Krzakala F and Martin O C 2003 *Phys. Rev.* B **68** 224404
[3] Palassini M 2003 arXiv:cond-mat/0307713
[4] Boettcher S 2005 *Eur. Phys. J.* B **46** 501
[5] Katzgraber H G, Korner M, Liers F, Junger M and Hartmann A K 2005 *Phys. Rev.* B **72** 094421
[6] Pal K F 2006 *Physica* A **367** 261
[7] Aspelmeier T and Moore M A 2003 *Phys. Rev. Lett.* **90** 177201
[8] De Dominicis C and Di Francesco P 2003 arXiv:cond-mat/0301066
[9] Aspelmeier T, Moore M A and Young A P 2003 *Phys. Rev. Lett.* **90** 127202
[10] Aspelmeier T, Billoire A, Marinari E and Moore M A Finite size corrections in the Sherrington–Kirkpatrick model *J. Phys. A. Math. Theor.* **41** 324008
[11] Andreanov A, Barbieri F and Martin O C 2004 *Eur. Phys. J.* B **41** 365
[12] Temesvari T 2009 The Ising spin glass in finite dimensions: a perturbative study of the free energy arXiv:0911.0362v1
[13] Parisi G and Rizzo T 2008 *Phys. Rev. Lett.* **101** 117205
[14] Parisi G and Rizzo T 2009 *Phys. Rev.* B **79** 134205
[15] Parisi G and Rizzo T 2009 arXiv:0901.1100
[16] Crisanti A, Paladin G, Sommers H-J and Vulpiani A 1992 *J. Phys. I France* **2** 1325
[17] Wehr J and Aizenman M 1990 *J. Stat. Phys.* **60** 287
[18] Mezard M,, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[19] Kondor I 1983 *J. Phys. A: Math. Gen.* **16** L127
[20] Talagrand M 2007 *J. Stat. Phys.* **126** 837
[21] Boettcher S 2003 *Eur. Phys. J.* B **31** 29–39
[22] Boettcher S 2003 *Phys. Rev.* B **67** 060403 (Rapid Communications)
[23] Liers F, Palassini M, Hartmann A K and Juenger M 2003 *Phys. Rev.* B **68** 094406
[24] Dotsenko Vik, Franz S and Mezard M 1994 *J. Phys. A. Math. Gen.* **27** 2351–65
[25] Goldschmidt Y and De Dominicis C 1989 *Phys. Rev.* B **41** 2186
[26] Mezard M and Parisi G 2001 *Eur. Phys. J.* B **20** 217