J. Phys. A: Math. Gen. **33** (2000) 113–129. Printed in the UK PII: S0305-4470(00)04437-1

On the origin of ultrametricity

Giorgio Parisi† and Federico Ricci-Tersenghi‡

† Dipartimento di Fisica, Università La Sapienza and INFN Sezione di Roma I, Piazzale Aldo Moro 2, 00185 Roma, Italy

‡ Abdus Salam ICTP, Condensed Matter Group, Strada Costiera 11, PO Box 586, 34100 Trieste, Italy

Received 20 May 1999, in final form 29 October 1999

Abstract. In this paper we show that in systems where the probability distribution of the overlap is nontrivial in the infinite volume limit, the property of ultrametricity can be proved, in general, starting from two very simple and natural assumptions: each replica is equivalent to the others (*replica equivalence* or *stochastic stability*) and all the mutual information about a pair of equilibrium configurations is encoded in their mutual distance or overlap (*separability* or *overlap equivalence*).

1. Introduction

Since its introduction in 1975 [1] the Sherrington–Kirkpatrick (SK) model for spin glasses has been one of the major challenges for physicists interested in complex systems.

Although it is a mean-field model the exact solution is still not completely certain. Nonetheless, it is known [2, 3] that in the low-temperature phase the replica symmetry is spontaneously broken and this makes the solution highly nontrivial. In other words, it can be rigorously proved that it is not possible that the connected correlation function of the spins at different points goes to zero when the total number of spins, N, goes to infinity: consequently, the probability distribution of the overlap q (defined below) cannot be a single delta function as occurs in the usual model (e.g. the ferromagnetic model). The existence of fluctuating intensive quantities (such as q) implies that it is not possible that only one equilibrium state is present in the thermodynamic limit. As usual, extensive quantities do fluctuate when more than one equilibrium state is present and, consequently, we could say, with some abuse of language (see [4] for a more precise discussion), that in the SK model, for large values of N more than one equilibrium state is present.

The presence of many equilibrium states implies that any analytic solution of the model should tell us something about the nature of these states, on their relative relations and on the probability at equilibrium of finding the system in one of these states. Of course, this information should be of a probabilistic nature given the presence of the quenched disorder in the system. Using replica formalism an Ansatz was proposed almost 20 years ago [5] which makes some hypothesis on the nature of the states, the most notable being ultrametricity (UM). Roughly speaking, UM implies that the distance between the different states is such that they can be put in a taxonomic or genealogical tree such that the distance between two states is consistent with their position on the tree.

This hierarchical Ansatz seems more reliable as each day passes and, although the physical origin of UM is not fully evident, it is widely believed that it provides the correct solution of

0305-4470/00/010113+17$30.00

the SK model. The ultrametric solution has passed many numerical tests and it is in agreement with all known analytical results [6,7]. It is quite possible, and in agreement with the numerical simulations, that the ultrametric organization of the equilibrium configurations is also present in finite-dimensional spin glasses [8–11].

It is certainly very interesting, for two reasons, to find which are the physical assumptions forming the basis of the hierarchical Ansatz: it would be easier to understand if the assumptions also make sense in the finite-dimensional case and it would be easier to prove them or to extract their consequences.

Considerable progress has been made in recent years after it was realized that one of the main hypotheses forming the basis of the hierarchical Ansatz was stochastic stability: many compulsory arguments can be given for the validity of stochastic stability, and the correctness of this hypothesis can be directly tested in experiments measuring the fluctuations and the response to a perturbation of the appropriate quantities. In the replica language *stochastic stability* is equivalent to the usual assumption of *replica equivalence* (i.e. each replica is equivalent to the others).

The other 'pillar' of the hierarchical Ansatz turned out to be UM. Indeed, it can be shown that if we assume stochastic stability and UM, the whole hierarchical Ansatz can be reconstructed [9].

The aim of this paper is to show that there is a simpler property which is equivalent to UM. This property can be called *separability*, in the replica language, or *overlap equivalence*: it states that all the mutual information about a pair of equilibrium configurations is encoded in their mutual distance or overlap. In other words, according to the principle of overlap equivalence any possible definition of overlaps should not give information additional to that of the usual overlap. It is not clear if there are strong compulsory reasons for assuming overlap equivalence, however, the results presented here show that the hierarchical Ansatz is certainly the simplest one that we can consider for a stochastically stable system with many equilibrium states: any other possible proposal should include the presence of at least two inequivalent definitions of distance.

These two assumptions, stochastic stability and overlap equivalence, are quite general and can be applied to many other systems beyond the SK model. A direct test or an analytic proof of the validity of both properties would have direct implications on the validity of the ultrametric solution. Moreover, as we have already remarked, the results that we are going to present in this paper are interesting because they show the root of UM: UM is the unique possibility we have if we stay within the simple framework where stochastic stability and separability hold.

It should be clear that the whole discussion applies to systems in which the overlap also fluctuates when the volume is very large and consequently, replica symmetry is broken. For systems in which the overlap does not fluctuate and replica symmetry is exact, we have nothing to say (UM is satisfied, but in a trivial way). It should also be clear that the arguments presented here cannot be used to argue if replica symmetry breaking occurs or not in a particular system. Here, we do not discuss the criticisms that have been applied to the replica approach in [12,13] (which cast some doubts on the viability of replica symmetry breaking in finite-dimensional systems): the reader may find a quite lengthy discussion in [4].

In section 2 we recall the replica formalism and in section 3 we present our assumptions. Next, in section 4 we present our main results on the relation between overlap equivalence and UM. Finally, we present our conclusions. Some of the arguments needed to show that overlap equivalence implies separability are presented in appendix A and a part of the tedious algebra required to reach the results is confined to appendix B.

2. The replica formalism

In this paper we make use of the replica formalism (we address the reader to [2, 3, 14, 15] for an introduction to the issue). For simplicity, we restrict the discussion to systems with quenched random disorder in the Hamiltonian. A similar discussion can be performed for systems without quenched disorder (such as structural glasses) if we substitute the average over the quenched disorder with the average over the size of the system.

We consider a system with N spins characterized by an Hamiltonian $H_J(\sigma)$ (where J represents the quenched disorder). We define $P_J(q)$ to be the probability distribution of finding two equilibrium configurations σ and τ at the same inverse temperature β, with overlap q, where overlap q is defined as

$$q = \frac{1}{N} \sum_{i=1,N} \sigma_i \tau_i . \tag{1}$$

Let us assume that, in the large-N limit, the function $P_J(q)$ does not become a delta function (otherwise we have nothing interesting to say) and, therefore, the function $P_J(q)$ also has a nontrivial shape for large N. When this happens, we are interested in finding out the probability distribution of the function $P_J(q)$ in the limit where N goes to infinity. For a given value of N, different choices of J may produce different functions $P_J(q)$, and we can introduce the functional $\mathcal{P}_N(P)$ as the probability distribution of $P_J(q)$ at fixed N (we assume that J have a given probability distribution). Eventually, we would like to know

$$\mathcal{P}_\infty(P) \equiv \lim_{N\to\infty} \mathcal{P}_N(P). \tag{2}$$

We are also interested in controlling the behaviour of the probability distribution of the mutual overlaps among three or more equilibrium configurations (e.g. the probability $P_J^{12,23,31}(q_{12}, q_{23}, q_{31})$, which will be properly defined later).

The origin of our interest in the probability distribution of the overlap stems from the fact that it controls many others physical properties of systems: e.g. in some models one finds [16] that the magnetic susceptibility is given by $\chi = \beta(1 - \overline{\langle q \rangle})$, where $\overline{\langle q \rangle}$ is the average over J of the equilibrium expectation value of q.

In the replica formalism the behaviour of these probability distributions is encoded by an $n \times n$ symmetric matrix Q in the limit of $n \to 0$ (taken after the analytical continuation of n from integer to real values). The limiting matrix depends on all the matrices with any value of n and so the general solution has an infinite number of parameters and the analytical continuation of the matrix Q is, in general, dependent on an extremely high number of parameters. This is quite natural as far as the matrix Q encodes the properties of the functional which controls the probability distribution of finding, for a random J, a set of probability distributions for the overlaps (i.e. $P_J(q)$, $P_J^{12,23,31}(q_{12}, q_{23}, q_{31}), \ldots$).

In the hierarchical Ansatz the n replicas are divided into many groups of equal size, such that, if the replica indices a and b belong to the same group, then Q_{ab} has a higher value than if a and b are in different groups. The groups are then divided into subgroups, and so on, for an infinite number of times. This kind of solution can be summarized in an infinite set of parameters (the size of the group and the value of the overlap at each level). In the limit $n \to 0$ these parameters can be conveniently represented by a function $P_R(q)$ defined for $q \in [0, 1]$, where $P_R(q)$ is the probability of finding an element of value q in the matrix Q. To every ultrametric matrix Q corresponds one and only one probability distribution function (PDF) $P_R(q)$.

In the paramagnetic phase all the elements of Q are equal and the function $P_R(q)$ is a delta function. In the spin glass phase the elements Q_{ab} take different values and $P_R(q)$ acquires a finite width.

The relation of this function to the probability distribution function of the overlap is

$$P_R(q) = P(q) \equiv \lim_{N\to\infty} \overline{P_J(q)} \tag{3}$$

where the overbar denotes the average over J at fixed N. The equality of the two functions $P_R(q)$and $P(q)$ is one of the many relations among PDFs of the overlaps and the matrix Q.

More complicated PDFs can be defined, considering the joint probability of more than one overlap. For example, a crucial role is played by the joint PDF of three real replicas $P^{12,23,31}(q_{12}, q_{23}, q_{31})$, defined as the J-average of $P_J^{12,23,31}(q_{12}, q_{23}, q_{31})$, where σ^1, σ^2 and σ^3 are three equilibrium configurations and

$$q_{\alpha,\beta} = \frac{1}{N} \sum_{i=1,N} \sigma_i^\alpha \sigma_i^\beta \tag{4}$$

with the indices α and β running from 1 to 3.

If UM holds, this probability distribution has the following property:

$$P^{12,23,31}(q_{12}, q_{23}, q_{31}) = 0 \tag{5}$$

as soon as the UM relations

$$\begin{aligned} q_{12} &\geqslant \min(q_{23}, q_{31}) \\ q_{23} &\geqslant \min(q_{31}, q_{12}) \\ q_{31} &\geqslant \min(q_{12}, q_{23}) \end{aligned} \tag{6}$$

are no longer satisfied. The UM property can be more easily understood from a geometrical viewpoint. Given three configurations, that is three points in the configurational space, UM implies that they can only be the vertices of two kinds of triangle: equilateral or isosceles, with the two equal edges larger than the third one. The other kind of isosceles triangle, together with the scalene triangles, cannot be obtained with any term of the equilibrium configurations; if these configurations are organized in an UM fashion.

This property, which is satisfied in the hierarchical Ansatz, has rather important consequences.

Firstly, as far as probabilities cannot be negative, the previous relations implies that for any J (with probability one) we also have that $P_J^{12,23,31}(q_{12}, q_{23}, q_{31}) = 0$ as soon as the relations in equation (6) are not satisfied. This result has the consequence that any equilibrium configuration can be assigned (for fixed J) to a leaf of a tree constructed in such a way that the overlap between two equilibrium configurations is related to the distance of the two configurations on the tree. UM implies, for example, that if two equilibrium configurations 1 and 2 are at overlap $q_{12} > q$, any equilibrium configuration 3 such that $q_{13} > q$ also satisfies the relation $q_{23} > q$. UM is very interesting because it implies that many PDFs of more than three overlaps are zero in a wide region and reduces the whole problem to the construction of the statistical properties of the aforementioned tree.

Moreover, if stochastic stability is valid, the UM completely determines $P^{12,23,31}$ given the $P(q)$. One can show that [9]

$$\begin{aligned} P^{12,23,31}(q, q', q'') = {} & A(q)\delta(q - q')\delta(q - q'') + B(q, q')\theta(q - q')\delta(q' - q'') \\ & + B(q', q'')\theta(q' - q'')\delta(q'' - q) + B(q'', q)\theta(q'' - q)\delta(q - q') \end{aligned} \tag{7}$$

where

$$A(q) = \tfrac{1}{2} P(q) \int_0^q \mathrm{d}q'\, P(q') \tag{8}$$

and

$$B(q, q') = \tfrac{1}{2} P(q) P(q'). \tag{9}$$

In other words, stochastic stability and UM allow us to obtain all the PDFs of the overlap starting from the knowledge of the function $P(q)$.

3. The assumptions

It is clear that it is extremely difficult to arrive at some general conclusions about these probability distributions without making extra assumptions. We now show that two rather simple assumptions: replica equivalence (or equivalently stochastic stability) and separability provide very strong constraints.

Firstly, let us consider replica equivalence as formulated in the replica formalism.

As we have already stated, the properties of the probability distribution of the overlaps can be obtained in terms of the matrix Q_{ab}. Even in the low-temperature phase, when the matrix elements Q_{ab} are not constant, we may expect no physical difference between the replicas (which have been introduced as a mathematical trick). Replica equivalence states that the observables which involve only one replica are replica symmetric, i.e. they assume the same value. For example, replica equivalence implies that we must have that

$$\sum_b f(Q_{ab}) \tag{10}$$

does not depend on a.

Replica equivalence is equivalent to the stochastic stability property introduced by Guerra [6] and Aizenman and Contucci [7] which is valid under general conditions, i.e. if we introduce an arbitrarily small random long-range Hamiltonian (see [6] for a more careful discussion).

Equation (10) implies that each line (column) of the matrix Q is a permutation of the other lines (columns). Moreover, it has interesting consequences: with some algebra the following equalities can be proven:

$$\begin{aligned} P^{12,13}(q_{12}, q_{13}) &= \tfrac{1}{2} P(q_{12}) P(q_{13}) + \tfrac{1}{2} P(q_{12}) \delta(q_{12} - q_{13}) \\ P^{12,34}(q_{12}, q_{34}) &= \tfrac{2}{3} P(q_{12}) P(q_{34}) + \tfrac{1}{3} P(q_{12}) \delta(q_{12} - q_{34}). \end{aligned} \tag{11}$$

The proof of these equations can be performed by recalling some relations between the matrix Q_{ab} and the probability functions

$$\begin{aligned} P^{12,13}(q_{12}, q_{13}) &= \lim_{n\to 0} \frac{\sum'_{a,b,c=1,n} Q_{a,b} Q_{a,c}}{n(n-1)(n-2)} \\ P^{12,34}(q_{12}, q_{34}) &= \lim_{n\to 0} \frac{\sum'_{a,b,c,d=1,n} Q_{a,b} Q_{d,c}}{n(n-1)(n-2)(n-3)} \end{aligned} \tag{12}$$

where the primed sum is performed over all different replica indices. Let us denote

$$q^{(k)} = -\sum_{b=1,n} q^k_{a,b}. \tag{13}$$

The sum does not depend on a as a consequence of replica equivalence. It is also evident that

$$\sum_{a,b=1,n} q^{k_1}_{a,c} q^{k_2}_{b,d} = q^{(k_1)} q^{(k_2)}. \tag{14}$$

If we now look at the consequences of the previous equation and use the relations in equation (12) for both $a = b$ and $a \neq b$, we obtain the two relations in equation (11).

Identical relations have been proven by Guerra [6], using stochastic stability. We can very safely assume that they must be valid in any scheme of replica symmetry breaking. Equation (11) determines all the joint PDFs of two real replicas in terms of $P(q)$. The consequences of stochastic stability have been discussed at length in [15, 17, 18]. In a nutshell, stochastic stability implies that the system is a generic random system and it does not have any special properties: its properties are smooth functions of any external random perturbation.

The second assumption we made is the separability (also known as nondegeneracy) of the matrix Q [15], which corresponds to the following statement. Let us consider all the matrices which can be generated from the matrix Q in a permutational covariant fashion. Some examples are

$$Q_{ab}^k \qquad \sum_c Q_{ac} Q_{cb} \qquad \sum_{c,d} Q_{ac} Q_{ad} Q_{cd} Q_{cb} Q_{db}. \tag{15}$$

Separability states that, if we take two pairs of indices (ab and cd), we have that

$$Q_{ab} = Q_{cd} \Longrightarrow M_{ab} = M_{cd} \tag{16}$$

where M is a generic matrix of the set generated by the rules shown in equation (15). In other words, pairs of indices which have different properties have different overlap values. This means that we can classify a pair of replicas in terms of their mutual overlap [18] and that no finer classification of their mutual properties is possible.

The physical meaning of separability can be understood if we introduce another concept, the overlap equivalence. Let us consider an arbitrary local observable $O_i(\sigma)$. Simple examples of such an observable are

$$\begin{aligned} O_i &= \sum_k A_{i-k} \sigma_k \\ O_i &= \sum_{k,l} B_{i-k,i-l} \sigma_k \sigma_l \\ O_i &= \sum_k \sigma_i \sigma_k J_{i,k} \end{aligned} \tag{17}$$

where A and B are appropriate functions (e.g. they decrease sufficiently quickly at infinity). Many more complex choices of the local observable O can be constructed, for example, those involving more than two spins.

For any choice of the operator O we could define a generalized overlap [19]:

$$q_O = \frac{1}{N} \sum_{i=1,N} O_i(\sigma) O_i(\tau). \tag{18}$$

In the hierarchical Ansatz it turns out [20] that for any reasonable choice of the observable O, q_O is a function of q. In other words, when we change the two equilibrium configurations and the couplings J, the values of q and q_O also fluctuate for very large N, while the value of q_O restricted on those pairs of configurations with a fixed value of q does not fluctuate when N goes to infinity (that is, a scattered plot of q and q_O should collapse on a curve in the limit of large N). In other words, overlap equivalence implies that in the case where replica symmetry is broken and all overlaps fluctuate in the usual thermodynamical ensemble, these fluctuations disappear in the fixed-q ensemble.

In other words, overlap equivalence states that for a system composed of two replicas the overlap is a good, complete order parameter in the same way as the magnetization is for ferromagnetic systems. If we stay in the usual thermodynamic ensemble, there are many quantities that also fluctuate at large distances, however, if we consider the restricted ensemble where the order parameter takes a given value, all fluctuations at large distance disappear, and the connected correlation functions go to zero at infinity. This only happens if the order parameter has been chosen in such a way to carry enough information: in a ferromagnetic Heisenberg model the order parameter must be the three-component vector of the magnetization, one or two components of the magnetization would be not enough to fully characterize the state of the system in case of spontaneous magnetization.

A direct check of overlap equivalence can be done with the usual numerical simulations and it would be very interesting to see the results.

This property is called overlap equivalence because it states that all possible definitions of the overlap are equivalent, and there is an unique correspondence among the values of the different overlaps.

It is clear that the overlap equivalence is a very strong simplification. In general, we could have that the mutual relations between two equilibrium configurations are characterized by a large, possibly infinite set of independent overlaps and, therefore, their mutual relations are characterized by a large (or infinite) set of parameters. The property of overlap equivalence implies a much simpler situation, where only one parameter (the overlap q) characterizes the mutual relations between two equilibrium configurations.

We can argue that separability is the way to code overlap equivalence in the replica formalism. Both properties state that once the overlap between two *objects* is fixed, all the mutual relations between the two objects are also fixed. The difference between these two statements is that in the case of replica equivalence the two objects are equilibrium configurations, while in the case of separability the two objects are replicas. The identification of separability with overlap equivalence is quite natural because the structure of the matrix Q, in replica space, mirrors the structure of the mutual overlaps of equilibrium configurations. In appendix A we present some more detailed considerations which point toward the correctness of the identification of these two properties, however, a more general and formal proof of this statement would be welcome.

It is interesting to note that in the simplest model leading to UM, i.e. a branching random process in the infinite-dimensional space, the condition of overlap equivalence is satisfied [21]. Indeed, if we consider a random vector x_α in a finite-dimensional space (of dimension N) the quantities x_α^2 convey different information when α changes from 1 to N and can be used as different measures of the distance. At the other end of the spectrum, when N goes to infinity at fixed $x^2 \equiv \sum_{\alpha=1,N} x_\alpha^2$, thanks to the rotational invariance, we have that for each α, $\langle x_\alpha^2 \rangle = \frac{x^2}{N} \to 0$. Then if we introduce generalized distances parametrized by λ (where $0 < \lambda \leqslant 1$) and defined as

$$x_\lambda^2 \equiv \lambda^{-1} \sum_{\alpha=1,\lambda N} x_\alpha^2 \tag{19}$$

it is easy to check that with probability 1 (if the probability distribution is rotationally invariant) in the limit $N \to \infty$

$$x_\lambda^2 = x_1^2 = x^2. \tag{20}$$

Therefore, in this simple model overlap equivalence is automatically satisfied.

Let us consider what happens in the usual hierarchical Ansatz. In this case, when replica symmetry is broken, there is a subgroup of the group of permutations that commutes with the matrix Q. Let us consider the orbits in the space of pairs of indices. It can be checked that the values of the elements of the matrix Q and of any matrix derived using the rules in equation (15) are constant of the orbits and that different values of q do correspond, in general, to different orbits. Moreover, it can be checked that both separability and overlap equivalence hold in this case.

Perhaps the simplest nontrivial example of a nonultrametric system is given by the union of separately ultrametric systems with a nontrivial overlap distribution [22]. It is easy to check that $q = \lambda q_1 + (1 - \lambda) q_2$ does not also satisfy the ultrametric condition if q_1 and q_2 satisfy it. However, it is clear that in this example q_1 and q_2 are generalized overlaps which are not functions of q. Both UM and overlap equivalence disappear at the same time†.

† In the last example both UM and stochastic stability are violated. There are no known examples of stochastically stable states which are not ultrametric. As far as we know, it is still possible that stochastic stability implies UM.

The separability condition is extremely powerful in determining the expectation values of higher-order moments of the probability distribution. Let us study a simple example and consider two matrices M and R constructed following the rules given in equation (15). It is evident that we can write

$$\begin{aligned}\sum_b Q_{ab}^k M_{ab} &= \sum_b \int \mathrm{d}q\, \delta(q - Q_{ab}) Q_{ab}^k M_{ab} = \int \mathrm{d}q\, P(q) M(q) q^k \\ \sum_b Q_{ab}^k R_{ab} &= \sum_b \int \mathrm{d}q\, \delta(q - Q_{ab}) Q_{ab}^k R_{ab} \int \mathrm{d}q\, P(q) R(q) q^k\end{aligned} \tag{21}$$

where we have defined

$$P(q) = \sum_b \delta(q - Q_{ab}). \tag{22}$$

Indeed, separability implies that the matrix elements M_{ab} and R_{ab} are constant in the region where $Q_{ab} = q$; their values are denoted by $M(q)$ and $R(q)$, respectively. In the same way we have that

$$\sum_b Q_{ab}^k M_{ab} R_{ab} = \int \mathrm{d}q\, P(q) M(q) R(q) q^k. \tag{23}$$

Therefore, separability implies that quantities such as those in equation (23) can be computed from the knowledge of those in equation (21).

If we introduce the functions $P_M(q)$, $P_R(q)$ and $P_{MR}(q)$ such that

$$\begin{aligned}\sum_b Q_{ab}^k M_{ab} &= \int \mathrm{d}q\, P_M(q) q^k \\ \sum_b Q_{ab}^k R_{ab} &= \int \mathrm{d}q\, P_R(q) q^k \\ \sum_b Q_{ab}^k M_{ab} R_{ab} &= \int \mathrm{d}q\, P_{MR}(q) q^k\end{aligned} \tag{24}$$

the previous equations imply that

$$P_M(q) = P(q)M(q) \qquad P_R(q) = P(q)R(q) \qquad P_{MR}(q) = P(q)M(q)R(q). \tag{25}$$

The last equation can also be written as

$$P_{MR}(q) = \frac{P_M(q) P_R(q)}{P(q)}. \tag{26}$$

If we apply the previous formula to the case where M and R have the form

$$M_{ab} = \sum_c Q_{ac}^{k_1} Q_{cb}^{k_2} \tag{27}$$

$$R_{ab} = \sum_c Q_{ac}^{k_3} Q_{cb}^{k_4} \tag{28}$$

and consider all the possible k values, we find (after separating the contributions where some of the indices are equal) the rather surprising formula

$$\begin{aligned}3P^{12,13,32,24,41}(q, q_1, q_2, q_3, q_4) &= \delta(q_1 - q_4)\delta(q_2 - q_3) P^{12,23,31}(q, q_1, q_2) \\ &\quad + 2\frac{P^{12,23,31}(q, q_1, q_2) P^{12,23,31}(q, q_3, q_4)}{P(q)}.\end{aligned} \tag{29}$$

Similar results can be obtained for other probability distributions with more overlap.

Equation (29) is particularly interesting because, integrating over q, it implies that

$$3P^{13,32,24,41}(q_1, q_2, q_3, q_4) = \tfrac{1}{2}\delta(q_1 - q_4)\delta(q_2 - q_3)[P(q_1)P(q_2) + \delta(q_1 - q_2)P(q_2)] \\ +2\int \mathrm{d}q\, \frac{P^{12,23,31}(q, q_1, q_2)P^{12,23,31}(q, q_3, q_4)}{P(q)}. \tag{30}$$

The previous equation is remarkable, not only because it gives the full expression of the probability with four overlaps in terms of the probability with three overlaps, but also because it enforces hard constraints on the possible values of the function $P^{12,23,31}(q, q_1, q_2)$. Indeed, the lhs of equation (30) is, by definition, invariant under cyclic permutation of q, while the rhs of the same equation is not invariant for a generic choice of the function $P^{12,23,31}$.

What is the form of the generic function $P^{12,23,31}$ that satisies equation (30)? In the next section we will argue that it must be ultrametric.

4. Results

Our problem is now that of finding the most general matrix Q (or equivalently, the most general probability distribution) compatible with the replica equivalence (and then with Guerra's relations, equation (11)) and with separability (and then, in particular, with equation (30)). We will show that the most general matrix is the ultrametric one.

We consider the case when only a few values ($k = 3$, 4 or 5) are allowed for the matrix elements. The generalization to more than five values is straightforward and we hope that our conclusions will still be valid for a generic $P(q)$ having a continuous distribution of possible values.

When the overlap (or the matrix elements Q_{ab}) can take only k different values the function $P(q)$is the sum of k delta functions

$$P(q) = \sum_{i=1}^{k} p_i\ \delta(q - q_i) \tag{31}$$

where the weights p_i are, by definition, positive and such that $\sum_i p_i = 1$, and the q_i values are different. In addition, the joint PDF of three overlaps $P^{12,23,31}$ (that, hereafter, we will call $P^{(3)}$ for brevity) is the the sum of k^3 delta functions on the points (q_i, q_j, q_l), with $i, j, l = 1, \ldots, k$, and so we should give the k^3 weights p_{ijl} in order to determine $P^{(3)}$.

We can lower the number of free parameters p_{ijl} using some symmetries and the Guerra relations. The weight of the term (q_i, q_j, q_l) must be the same as any permutation of it, i.e. $p_{122} = p_{212} = p_{221}$. Then the number of truly independent parameters in $P^{(3)}$is $k(k + 1)(k + 2)/6$. More relations between p_{ijl} can be obtained exploiting the following equation, which is essentially based on the first Guerra relation:

$$\int \mathrm{d}q\ P^{(3)}(q, q_1, q_2) = P^{12,23}(q_1, q_2) = \tfrac{1}{2}P(q_1)P(q_2) + \tfrac{1}{2}P(q_1)\delta(q_1 - q_2). \tag{32}$$

These are $k(k + 1)/2$ relations that lower the degrees of freedom of $P^{(3)}$to $(k - 1)k(k + 1)/6$.

Then we have to determine the values of these $(k - 1)k(k + 1)/6$ parameters which are compatible with equation (30).

4.1. Three overlaps ($k = 3$)

To fix the ideas, let us write down some formulae for the easier case ($k = 3$) where we have 27 p_{ijl} parameters: $p_{111}, p_{112}, p_{113}, p_{121}, \ldots, p_{333}$. The symmetries imply that

$$\begin{aligned} & p_{112} = p_{121} = p_{211} \\ & \ldots \\ & p_{123} = p_{132} = p_{213} = p_{231} = p_{312} = p_{321} \end{aligned} \tag{33}$$

while Guerra's relations imply some equalities such as

$$\begin{aligned} & \sum_j p_{11j} = \tfrac{1}{2}p_1^2 + \tfrac{1}{2}p_1 \\ & \ldots \\ & \sum_j p_{12j} = \tfrac{1}{2}p_1 p_2 \\ & \ldots . \end{aligned} \tag{34}$$

We end with only four free parameters ($s, a_{32}, a_{31}, a_{21}$):

$$\begin{aligned} p_{321} &= s \\ p_{332} &= a_{32} \\ p_{331} &= a_{31} \\ p_{221} &= a_{21} \\ p_{322} &= p_3 p_2/2 - a_{32} - s \\ p_{311} &= p_3 p_1/2 - a_{31} - s \\ p_{211} &= p_2 p_1/2 - a_{21} - s \\ p_{333} &= p_3(1 + p_3)/2 - a_{31} - a_{32} \\ p_{222} &= p_2(1 + p_2)/2 - p_3 p_2/2 + a_{32} - a_{21} + s \\ p_{111} &= p_1(1 + p_1)/2 - p_3 p_1/2 - p_2 p_1/2 + a_{31} + a_{21} + 2s. \end{aligned} \tag{35}$$

The way we have ordered the probabilities is meaningful: we call s the weight of the scalene triangle (which is forbidden in the ultrametric solution) and we call a_{32}, a_{31}, a_{21} the weights of the isosceles triangles (which are also forbidden in the UM Ansatz if we assume the overlap ordering $q_1 < q_2 < q_3$).

If we do not fix any order in the values of the q_i, we have to keep in mind that, if we exchange two of the overlap values, the forbidden isosceles triangle changes. For example, if $q_1 < q_2 < q_3$, then UM implies $p_{332} = p_{331} = p_{221} = 0$, while when we reverse the second inequality, i.e. $q_1 < q_3 < q_2$, we have that $p_{322} = p_{331} = p_{221} = 0$. Then we note that

$$s = a_{32} = a_{31} = a_{21} = 0 \Longrightarrow \text{UM} \tag{36}$$

while the reversed implication is not true, because UM also holds for different parameter values, e.g. $s = a_{31} = a_{21} = 0$ and $a_{32} = p_3 p_2/2$ (which corresponds to the ordering $q_1 < q_3 < q_2$).

For a generic k we have $k(k-1)(k-2)/6$ scalene parameters s_i, which must all be identically zero in order that UM hold ($\{s_i = 0\} \Longleftrightarrow$ UM), while the $k(k-1)/2$ isosceles parameters a_{ij} must be zero or $a_{ij} = \frac{1}{2}p_i p_j$, depending on the order of q_i and q_j.

We will now use equation (30) to determine the values of all these parameters. The lhs of equation (30) is invariant under cyclic permutations of the four overlaps. This allows us to obtain useful relations simply taking two of these equations (the second one with the overlaps cycled with respect to the first one) and equating their rhs. The number of nontrivial equations we can obtain in this way is large enough to fix all the parameters.

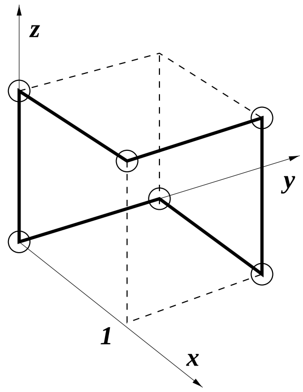

Figure 1. Schematic representation of the properties of the symbol $((x;\, y, z))$, it takes non-negative values in all of the unitary cube and is zero only on the bold edges.

In the particular case of $k = 3$ we have that all the nontrivial equations are equal (this is highly fortuitous) and read

$$\begin{aligned}&-2a_{32}a_{31}p_2p_1 + 2a_{32}a_{21}p_3p_1 - 2a_{31}a_{21}p_3p_2 + a_{31}p_3p_2^2p_1 + 2a_{32}p_3p_1s\\&-2a_{32}p_2p_1s - 2a_{31}p_2p_1s + p_3p_2p_1s - p_3^2p_2p_1s + 2p_3p_2s^2 + 2p_3p_1s^2 = 0.\end{aligned} \tag{37}$$

Using the relations that come from the sixth equation in (35)

$$a_{31} = \tfrac{1}{2}p_3p_1 - s - p_{311} \tag{38}$$

we can write equation (37) as

$$E_0 + 2(a_{32}p_3p_1 + a_{21}p_3p_2 + p_{311}p_2p_1)s + 2(p_3p_2 + p_3p_1 + p_2p_1)s^2 = 0 \tag{39}$$

where in E_0 we put all the terms that survive once we set $s = 0$. The coefficients of s and s^2 are positively defined (thanks to the positiveness of all the probabilities) and E_0 in non-negative (as we will show in a while). Then equation (39) is equivalent to

$$\begin{aligned}&s = 0\\&E_0 = a_{31}p_3p_2^2p_1 - 2a_{32}a_{31}p_2p_1 - 2a_{31}a_{21}p_3p_2 + 2a_{32}a_{21}p_3p_1 = 0.\end{aligned} \tag{40}$$

As a first result we obtain that scalene triangles are completely forbidden.

Let us now introduce the following symbol:

$$((x;\, y, z)) \equiv x - xy - xz + yz = x(1-y)(1-z) + (1-x)yz. \tag{41}$$

For $x, y, z \in [0, 1]$ we have that $((x;\, y, z)) \geqslant 0$, and the equality $((x;\, y, z)) = 0$ only holds on six of the 12 edges of the cube (those in bold face in figure 1).

If we introduce the new parameters

$$a'_{ij} \equiv \frac{2a_{ij}}{p_ip_j} \tag{42}$$

that belong to the range [0, 1] thanks to the positiveness of the probabilities, then the second equality in equation (40) can be rewritten in a very compact form as

$$((a'_{31};\, a'_{32}, a'_{21})) = 0. \tag{43}$$

This form makes clear that E_0 is non-negative, as we claimed above.

Equation (43) is not as stringent as UM would like, but the deviations from UM are small. In fact, in the cube $a'_{32}, a'_{31}, a'_{21} \in [0, 1]$ (see figure 1) strict UM only holds on the vertices marked by a circle, while equation (43) is also satisfied along the bold lines.

For example, on the segment $a'_{32} = a'_{31} = 0$ and $0 < a'_{21} < 1$ they seem to co-exist with nonzero probabilities p_{221} and p_{211} and it would be a small violation of UM. However, we know that $a'_{32} = a'_{31} = a'_{21} = 0$ corresponds to the ordering $q_1 < q_2 < q_3$, while $a'_{32} = a'_{31} = 0$ and $a'_{21} = 1$, correspond to $q_2 < q_1 < q_3$. We believe that the points on the segment between these two UM points correspond to the case $q_1 = q_2 < q_3$, when there is no difference between p_{221} and p_{211} (but we still cannot prove it).

In conclusion, in the case with $k = 3$ overlaps, we have that the scalene triangle and two of the three 'wrong' (in the UM sense) isosceles triangles are forbidden. As we will see below, the UM violations become smaller and smaller as k is increased.

4.2. Four or more overlaps ($k \geqslant 4$)

In this section we would like to sketch how the information we need about $P^{(3)}$can be systematically derived from equation (30). To make this section more readable, the formulae related to the cases $k = 4, 5$ will be presented in appendix B. The method we use to obtain the results does not depend on k and so we will be able to generalize our findings to whichever $P(q)$ that is the sum of a finite number of delta functions.

The many equations derivable from equation (30) can be divided into three classes: those with two, three and four different overlap values. These equations are not independent: those with two (resp. four) different overlaps can be expressed as the sum (resp. difference) of those with three overlaps.

There are many ways of solving the equations. Here we present the simplest one we were able to find: we have to consider only the equations with two and three overlaps, i.e. those which correspond, respectively, to the equalities

$$P^{12,23,34,41}(q_i, q_j, q_j, q_i) - P^{12,23,34,41}(q_i, q_i, q_j, q_j) = 0 \tag{44}$$

and

$$P^{12,23,34,41}(q_i, q_i, q_j, q_l) - P^{12,23,34,41}(q_i, q_j, q_l, q_i) = 0. \tag{45}$$

Each one of these equations can be identified giving a pair or a term of numbers: (i, j) or (i, j, l). The lhs of these equations will be called, respectively, $E^{(i,j)}$ and $E^{(i,j,l)}$, for brevity.

Our demonstration follows two steps: first we show that the equations of the same kind as equation (44) can be solved only if all the scalene parameters p_{ijl} (with i, j, l different) are zero, then we find all the solutions for the simplified set of equations corresponding to equation (45). Our demonstration is essentially based on the non-negativity of the $E^{(i,j)}$ expressions and on the properties of the double-parenthesis symbol, previously introduced.

First of all, we note (see appendix B) that when we set to zero all the scalene parameters p_{ijl} (with i, j, l different), every $E^{(i,j)}$ becomes the sum of some double-parenthesis symbols, and so are non-negative. Moreover, in some of the $E^{(i,j)}$ expressions all the scalene parameters have positively defined coefficients, which we should set to zero in order to solve the equation, $E^{(i,j)} = 0$. In appendix B we present a possible method for choosing the $E^{(i,j)}$ expressions in order to systematically set to zero all the scalene parameters.

Once the scalene parameters have been set to zero, we prefer working with equation (45), because each equation identified by (i, j, l) takes a very simple form:

$$((a'_{il}; a'_{ij}, a'_{jl})) = 0 \tag{46}$$

where we choose the indices such that $q_i < q_j < q_l$. In general, given three different overlap values, we can easily write down the corresponding equation (of the same kind as equation (45)), which gives, once we set all the scalene parameters to zero, the corresponding double-parenthesis symbol (of the same kind as equation (46)).

What about the equation with two and four different overlaps? When we set all the scalene parameters to zero, we have that the $E^{(i,j)}$ expressions are the sum of $k - 2$ of these double-parenthesis symbols, those derived from the overlap terms (q_i, q_j, q_h) with $q_h \neq q_i$ and $q_h \neq q_j$ (that is, those where the parameter a'_{ij} appears). On the other hand, the equations with four different overlaps are identically satisfied and are useless.

Then we conclude that, in the more general solution, equation (46) must hold for every overlap term $q_i < q_j < q_l$. What does it imply in terms of the ultrametric properties of $P^{(3)}$?

For any pair of overlaps, $q_i > q_j$, we have two different isosceles triangles: a 'right' one (i.e. allowed by UM) with probability p_{ijj} and a 'wrong' one (i.e. forbidden by UM) with probability $p_{iij} \propto a'_{ij}$. In the solution we have found that, almost all the wrong isosceles triangles are forbidden. More precisely, for any pair of overlaps $q_i > q_j$, such that there is an overlap q_h in between ($q_i > q_h > q_j$), we have that $a'_{ij} = 0$ and the wrong isosceles triangle is not allowed. That can be easily proved noting that, for any $q_i > q_h > q_j$, the equation $((a'_{ij}; a'_{ih}, a'_{hj})) = 0$ forces $a'_{ij} = 0$.

Small UM violations can appear only when one considers nearest-neighbour overlap pairs. In this case both the right and the wrong isosceles triangles are allowed. However, for any fixed k, the maximum number of wrong isosceles triangles allowed is $[\frac{k}{2}]$, while the total number of isosceles triangles is proportional to k^2. So in the limit $k \to \infty$ the probability of having wrong isosceles triangles tends to zero.

Moreover, if in the continuum limit $P(q)$ is dense on a single compact domain, the distance between any pair of nearest-neighbour overlaps tends to zero for $k \to \infty$ and then strict UM holds for any finite overlap difference $|q_i - q_j|$.

Finally, it should be noted that we have not exploited all the available information and it is possible that even these small UM violations could be ruled out with some more work. In fact, we believe that the solution with both p_{iij} and p_{ijj} different from zero, actually corresponds to the case $q_i = q_j$ and it is not really an UM violation.

It might also be possible to find a direct proof of our result directly in the continuum limit, without considering the intermediate case in which the number of steps are finite, perhaps using the techniques introduced by Ruelle [23], however we have not succeeded in this task.

5. Conclusions

We have seen that in systems where the function $P(q)$ is nontrivial and the overlap is a fluctuating quantity, stochastic stability and separability imply UM. The reader should note that the proofs presented here are likely to be too involved and it is quite possible that there is a direct proof that replica equivalence implies UM. To this end we recall that in a dynamical approach it was shown that one can identify a dynamical equivalent of separability, i.e. we can assume that in the aging regime all the possible overlaps between two configurations at two quite different times (t_1 and t_2) are functions of the usual overlap between the two configurations at the same times (t_1 and t_2). It was possible to prove that this dynamical overlap equivalence implies a dynamical form of UM.

This result implies that, if we do not give up stochastic stability (which is a general property of generic equilibrium systems) violations of UM may be found only in systems for which the separability conditions does not hold and the mutual relations among equilibrium configurations is described by two or more overlaps. The probability distribution of such a system (if it exists in the framework of equilibrium statistical mechanics) would be much more complex than that of the usual ultrametric Ansatz. We can thus conclude that the ultrametric solution is the simplest one.

Our arguments imply that it would be particularly interesting to check the numerical validity of overlap equivalence. This task can be performed at high precision using presently available numerical technology. This could be done by performing numerical simulations in the ensemble with fixed overlap and looking to the fluctuations of the other overlaps.

Acknowledgments

We thank S Franz and M A Virasoro for many interesting discussions relating to the subject matter of this paper.

Appendix A

In this appendix we present some arguments in order to show that *separability* implies *overlap equivalence*.

To this end let us consider a specific spin glass model

$$H_J = \sum_{i,k} J_{ik}\sigma_i\sigma_k \tag{47}$$

where the variables J are Gaussian uncorrelated random variables with zero average and variance

$$\overline{J_{ik}^2} = K_{ik}. \tag{48}$$

In short-range models K_{ik} is a rapidly decreasing function of the distance between the two points i and k, while in the SK model $K_{ik} = N^{-1}$, where N denotes the total number of spins.

In this model we can define not only the usual overlap but also a modified overlap between two configurations σ and τ which we denote by r:

$$r = \frac{1}{N}\sum_{i,k} J_{ik}\sigma_i\tau_k. \tag{49}$$

In the same way we can define an overlap between two replicas, which we denote by r_{ab}

$$r_{ab} = \frac{1}{N}\sum_{i,k} J_{ik}\langle\sigma_i^a\sigma_k^b\rangle. \tag{50}$$

We notice that by simple integration by parts on the Gaussian variables J, one can prove that

$$\overline{\langle r\rangle} = \frac{1}{n(n-1)}\sum_{a,b=1,n} t_{ab} \tag{51}$$

with

$$t_{ab} = \langle r_{ab}\rangle \sum_{i,k} K_{ik} \sum_c \frac{1}{n(n-1)} \langle\sigma_i^a\sigma_k^b\sigma_i^c\sigma_k^c\rangle \tag{52}$$

where the overbar denotes the average over the random quenched variables J.

A similar computation tells us that the fluctuations of the quantity r at fixed q are (neglecting terms which go to zero with the volume) the same as the fluctuations of t_{ab} at fixed q_{ab}. On the other hand, it is evident that

$$\sum_c \sum_{i,k} \frac{1}{N^2}\langle\sigma_i^a\sigma_k^b\sigma_i^c\sigma_k^c\rangle = \sum_c Q_{ac}Q_{bc}. \tag{53}$$

Separability states that the last sum takes a fixed value which does not fluctuate if we stay in the ensemble of fixed Q_{ab}.

At this level the relation between the separability and overlap equivalence is clear: the first is equivalent to the statement that the quantity $N^{-1}\sum_i\sum_c\langle\sigma_i^a\sigma_k^b\sigma_i^c\sigma_k^c\rangle$ does not fluctuate (in the fixed q_{ab} ensemble) when $i - k$ is large, while for overlap equivalence we need that the same quantity does not fluctuate when the distance between i and k is fixed.

The two properties seems to be slightly different if the interaction is short range. In contrast, if the interaction is long range the two formulations are the same. In order to be more precise we can consider a model in which a small long-range interaction has been added—the same argument as before can be used to prove that overlap equivalence implies separability. Moreover, the presence of a long-range term should not affect the properties according to the principle of stochastic stability too much which tells us that the system should be stable with respect to a small random perturbation.

The argument we have presented here tells us that replica equivalence implies separability.

Appendix B

In this appendix we show some details of the computation we performed for the cases $k = 4$ and 5. In particular, we show exactly how to derive the solution for the $k = 4$ case, while we simply sketch it in the $k = 5$ case.

The $k = 4$ case

In the $k = 4$ case we have four scalene parameters ($p_{432}, p_{431}, p_{421}, p_{321}$) and six isosceles parameters ($p_{443} = a_{43}, p_{442} = a_{42}, p_{441} = a_{41}, p_{332} = a_{32}, p_{331} = a_{31}, p_{221} = a_{21}$). The remaining probabilities are functions of the following ten parameters:

$$\begin{aligned}
p_{433} &= p_4 p_3/2 - a_{43} - p_{432} - p_{431} \\
p_{422} &= p_4 p_2/2 - a_{42} - p_{432} - p_{421} \\
p_{411} &= p_4 p_1/2 - a_{41} - p_{431} - p_{421} \\
p_{322} &= p_3 p_2/2 - a_{32} - p_{432} - p_{321} \\
p_{311} &= p_3 p_1/2 - a_{31} - p_{431} - p_{321} \\
p_{211} &= p_2 p_1/2 - a_{21} - p_{421} - p_{321} \\
p_{444} &= p_4(1 + p_4)/2 - a_{43} - a_{42} - a_{41} \\
p_{333} &= p_3(1 + p_3 - p_4)/2 - a_{32} - a_{31} + a_{43} + p_{432} + p_{431} \\
p_{222} &= p_2(1 + p_2 - p_3 - p_4)/2 - a_{21} + a_{42} + a_{32} + 2p_{432} + p_{421} + p_{321} \\
p_{111} &= p_1^2 + a_{41} + a_{31} + a_{21} + 2(p_{431} + p_{421} + p_{321}).
\end{aligned} \tag{54}$$

Let us consider the equation of the same kind as equation (44) with the two greatest overlaps (q_4 and q_3 in the $k = 4$ case),

$$\begin{aligned}
E^{(4,3)} = \frac{1}{2} p_4 p_3 + \frac{p_{431}^2}{p_1} + \frac{p_{432}^2}{p_2} + \frac{p_{433}^2}{p_3} + \frac{a_{43}^2}{p_4} \\
- \left(\frac{a_{41} a_{31}}{p_1} + \frac{a_{42} a_{32}}{p_2} + \frac{a_{43} p_{333}}{p_3} + \frac{p_{444} p_{433}}{p_4} \right) = 0.
\end{aligned} \tag{55}$$

Using some of equations (54) and multiplying the previous equation by $c = 2p_1 p_2 p_3 p_4$, we end with the following equation:

$$\begin{aligned}
cE^{(4,3)} = cE_0^{(4,3)} + [2p_1 p_3 p_4 a_{43} + p_1 p_2 p_3 (p_4 p_3 - 2a_{43} + p_4 p_2 \\
-2a_{42} + p_4 p_1 - 2a_{41})](p_{432} + p_{431}) + [4p_1 p_3 p_4] p_{432} p_{431} \\
+[2p_1 p_4 (p_2 + p_3)] p_{432}^2 + [2p_2 p_4 (p_1 + p_3)] p_{431}^2 = 0.
\end{aligned} \tag{56}$$

$E_0^{(4,3)}$ contains all the terms that survive from the expression $E^{(4,3)}$ when we set all the scalene parameters to zero. In general, all the expressions $E_0^{(i,j)}$ are non-negatively defined (see below).

The coefficients of p_{432} and p_{431} in equation (56) are positively defined, thanks to the inequalities $p_4 p_3 \geqslant 2a_{43}$, $p_4 p_2 \geqslant 2a_{42}$, $p_4 p_1 \geqslant 2a_{41}$, that are direct consequences of equations (54) and the positiveness of the probabilities. Then we have that equation (56) is equivalent to $p_{432} = p_{431} = 0$ and $E_0^{(4,3)} = 0$.

To force the two remaining scalene parameters to zero, it is sufficient to consider the equation analogous to equation (55) with q_1 and q_2 instead of q_3 and q_4. Once we set $p_{432} = p_{431} = 0$, we obtain

$$\begin{aligned} cE^{(2,1)} = cE_0^{(2,1)} &+ [2p_1p_3p_4(a_{42} + a_{32}) + p_1p_2p_3(p_4p_2 - 2a_{42} + p_4p_1 - 2a_{41})]p_{421} \\ &+[2p_1p_3p_4(a_{42} + a_{32}) + p_1p_2p_4(p_3p_2 - 2a_{32} + p_3p_1 - 2a_{31})]p_{321} \\ &+[2p_3p_4(p_1 + p_2)](p_{421} + p_{321})^2 = 0. \end{aligned} \tag{57}$$

Again, we note that all the coefficients are positively defined thanks to equations (54) and to the positiveness of the probabilities. Equation (57) implies $p_{421} = p_{321} = 0$ and $E_0^{(2,1)} = 0$.

Then we conclude that in the more general solution to equation (30) in the case of $k = 4$, different overlaps forbid any scalene triangle.

In order to obtain this result we have made only one assumption, about the non-negativity of $E_0^{(4,3)}$ and of $E_0^{(2,1)}$, which we now show to be correct. Using the rescaled variables, $a'_{ij} = \frac{2a_{ij}}{p_i p_j}$, those expressions read

$$E_0^{(4,3)} = \frac{p_2p_3p_4}{4}((a'_{42}; a'_{43}, a'_{32})) + \frac{p_1p_3p_4}{4}((a'_{41}; a'_{43}, a'_{31})) \tag{58}$$

$$E_0^{(2,1)} = \frac{p_1p_2p_4}{4}((a'_{41}; a'_{42}, a'_{21})) + \frac{p_1p_2p_3}{4}((a'_{31}; a'_{32}, a'_{21})). \tag{59}$$

Note that $E_0^{(4,3)}$ (resp. $E_0^{(2,1)}$) is the sum of the two double-parenthesis symbols containing a'_{43} (resp. a'_{21}).

Once we set to zero all the scalene parameters $p_{432} = p_{431} = p_{421} = p_{321} = 0$, we found it easier to work with equations of the same kind as equation (45). For example, considering the three overlaps $q_1 < q_2 < q_3$, we have that

$$0 = P^{12,23,34,41}(q_1, q_1, q_2, q_3) - P^{12,23,34,41}(q_1, q_2, q_3, q_1) = \frac{p_1p_2p_3}{4}((a'_{31}; a'_{32}, a'_{21})). \tag{60}$$

In general, for every three overlaps given we end with an equation such as (46).

The $k = 5$ case

The way to force the scalene parameters to zero should now be clear: exploit the coefficient positiveness in the equations with two different overlaps. Maybe it is still not so clear if there is a systematic way to set all these parameters to zero, without 'getting lost' in the many $E^{(i,j)}$ expressions.

We found such a systematic way and we will illustrate it in the case with $k = 5$ different overlaps. Let us always consider first the equation with the two greatest overlaps (q_5 and q_4, in this particular case). It implies

$$E^{(5,4)} = 0 \Longrightarrow p_{543} = p_{542} = p_{541} = 0. \tag{61}$$

Note that all the scalene probabilities forced to zero contain both q_5 and q_4.

Then let us substitute the newly found solution ($p_{543} = p_{542} = p_{541} = 0$) into all the other equations and move forward in the same way:

$$E^{(5,3)} = 0 \Longrightarrow p_{532} = p_{531} = 0 \tag{62}$$

$$E^{(5,2)} = 0 \Longrightarrow p_{521} = 0. \tag{63}$$

At this point we end with the same scalene parameters we work with in the $k = 4$ case and then follow the same steps as in the previous section:

$$E^{(4,3)} = 0 \Longrightarrow p_{432} = p_{431} = 0 \tag{64}$$

$$E^{(2,1)} = 0 \Longrightarrow p_{421} = p_{321} = 0. \tag{65}$$

Once all the scalene probabilities have been forced to zero, the demonstration is straightforward and follows the same method as outlined in the previous sections for the $k = 3, 4$ cases.

References

[1] Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **35** 1792
[2] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[3] Fischer K H and Hertz J A 1991 *Spin Glasses* (Cambridge: Cambridge University Press)
[4] Marinari E, Parisi G, Ricci-Tersenghi F, Ruiz-Lorenzo J J and Zuliani F 1999 *Preprint* cond-mat/9906076 submitted to *J. Stat. Phys.*
[5] Parisi G 1979 *Phys. Lett.* A **73** 203
Parisi G 1979 *Phys. Rev. Lett.* **43** 1754
Parisi G 1980 *J. Phys. A: Math. Gen.* **13** L115
Parisi G 1980 *J. Phys. A: Math. Gen.* **13** 1101
Parisi G 1980 *J. Phys. A: Math. Gen.* **13** 1887
Parisi G 1983 *Phys. Rev. Lett.* **50** 1946
[6] Guerra F 1997 *Int. J. Mod. Phys.* B **10** 1675 and references therein
[7] Aizenman M and Contucci P 1997 *Preprint* cond-mat/9712129
[8] Cacciuto A, Marinari E and Parisi G 1997 *J. Phys. A: Math. Gen.* **30** L263
A Cacciuto 1997 *Preprint* cond-mat/9704053
[9] Iñiguez D, Parisi G and Ruiz-Lorenzo J J 1996 *J. Phys. A: Math. Gen.* **29** 4337
[10] Hartmann A K 1998 *Europhys. Lett.* **44** 249
[11] Franz S and Ricci-Tersenghi F 1999 *Preprint* cond-mat/9903130 *Phys. Rev.* E to appear
[12] Fisher D S and Huse D A 1986 *Phys. Rev. Lett.* **56** 1601
Fisher D S and Huse D A 1988 *Phys. Rev.* B **38** 386
[13] Newman C M and Stein D L 1992 *Phys. Rev.* B **46** 973
[14] Parisi G 1993 *J. Stat. Phys.* **51** 51
[15] Parisi G 1998 *Preprint* cond-mat/9801081
[16] Parisi G, Ricci-Tersenghi F and Ruiz-Lorenzo J J 1998 *Phys. Rev.* B **57** 13617
[17] Franz S, Mézard M, Parisi G and Peliti L 1998 *Phys. Rev. Lett.* **81** 1758
[18] Franz S, Mézard M, Parisi G and Peliti L 1999 *Preprint* cond-mat/9903370
[19] Parisi G 1987 *Phys. Scr.* **35** 123
[20] Athanasiu G G, Bachas C and Wolff W F 1987 *Phys. Rev.* B **35** 1965
[21] Rammal R, Toulouse G and Virasoro M A 1986 *Rev. Mod. Phys.* **58** 765
[22] Franz S, Parisi G and Virasoro M A 1992 *Europhys. Lett.* **17** 5
[23] Ruelle D 1987 *Commun. Math. Phys.* **108** 225