ON THE ORIGIN OF ULTRAMETRICITY

Giorgio Parisi

and

Federico Ricci-Tersenghi

3 1 / 2 0

IC/2000/34

ON THE ORIGIN OF ULTRAMETRICITY

Giorgio Parisi
*Dipartimento di Fisica, Università "La Sapienza" and INFN Sezione di Roma I,
Piazzale Aldo Moro 2, 00185 Roma, Italy,*

and

Federico Ricci-Tersenghi
The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy.

Abstract

In this paper we show that in systems where the probability distribution of the the overlap is non trivial in the infinity 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*).

MIRAMARE – TRIESTE
April 2000

1 Introduction

Since its introduction in 1975 [1] the Sherrington-Kirkpatrick (SK) model for spin glasses has been one of the major challenges for the 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 non trivial. 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 happens in usual model (e.g. ferromagnetic model). The existence of fluctuating intensive quantities (like 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 that one equilibrium state is present.

The presence of many equilibrium states implies that any analytic solution of the model should tell us something on 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 probabilistic nature given the presence of the quenched disorder in the system. Using the replica formalism an Ansatz was proposed almost twenty years ago [5] which makes some hypothesis on the nature of the states, the most notable being ultrametricity. Roughly speaking ultrametricity implies that the distance among the different states is such that they can be put in a taxonomic or genealogical tree such that the distance among two states is consistent with their position on the tree.

This hierarchical Ansatz seems every day more reliable and, although the physical origin of ultrametricity was not fully evident, it is widely believed that it provides the correct solution of the SK model. The ultrametric solution has passed many numerical tests and it is in agreement with all the 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, 9, 10, 11].

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

A considerable progress has been done in recent years when it was realized that one of the main hypothesis at 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 ultrametricity. Indeed it can be shown that if we assume stochastic stability and ultrametricity, the whole hierarchical Ansatz

can be reconstructed [9].

The aim of this note is to show that there is a simpler property which is equivalent to ultrametricity. 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 may think 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 is interesting because they show the root of ultrametricity: ultrametricity 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 fluctuates also when the volume is very large and consequently replica symmetry is broken. For systems in which the overlap does not fluctuates and replica symmetry is exact we have nothing to say (ultrametricity 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 happens or not in a particular system. We do not discuss here the criticism that have been done to the replica approach by [12, 13] (who cast some doubts on the viability of replica symmetry breaking in finite dimensional systems): the reader may find a quite long discussion in [4].

In the second section we will recall the replica formalism, in the third section we will present our assumptions. Next in the forth section we will present our main results on the relation among overlap equivalence and ultrametricity. Finally we will present our conclusions. Some of the arguments needed to show that overlap equivalence implies separability are presented in Appendix I and a part of the tedious algebra we have to do to reach the results is confined in appendix II.

2   The replica formalism

In this paper we make use of the replica formalism (we address the reader to Refs. [2, 3, 14, 15] for an introduction on the issue). Let for simplicity restrict the discussion to systems with quenched random disorder in the Hamiltonian. A similar discussion can be done for system without quenched disorder (like 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, the

overlap q being defined as

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

Let us assume that, in the large N limit, the function $P_J(q)$ does not becomes a delta function (otherwise we have nothing interesting to say) and therefore the function $P_J(q)$ has a non trivial shape also for large N. When this happens, we are interested to find 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 the J have a given probability distribution). Eventually we would like to know

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

We are also interested to control the behavior 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 is due to the fact that they control many others physical properties of the 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 behavior of the 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. $P_J(q)$, $P_J^{12,23,31}(q_{12}, q_{23}, q_{31})$, ...).

In the hierarchical Ansatz the n replicas are divided into many groups of equal sizes, such that, if the replica indices a and b belong to the same group, then Q_{ab} has a higher value than whether a and b are in different groups. The groups are then divided in 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 in the matrix Q an element of value q. To every ultrametric matrix Q corresponds one and only one probability distribution function $P_R(q)$.

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 the $P_R(q)$ acquires a finite width.

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

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

where the bar 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 probability distribution functions of the overlaps and the matrix Q.

More complicated probability distribution functions (pdf) can be defined, considering the joint probability of more than one overlap. For example a crucial role is played by the joint pdf of 3 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

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

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

If ultrametricity holds, this probability distribution has the following property

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

as soon as the ultrametricity relations

$$q_{12} \geq \min(q_{23}, q_{31}) \quad ,$$
$$q_{23} \geq \min(q_{31}, q_{12}) \quad , \tag{6}$$
$$q_{31} \geq \min(q_{12}, q_{23}) \quad ,$$

are not satisfied. The UM property can be easier understood by a geometrical picture. Given three configurations, that is three points in the configurational space, UM implies that they can be the vertices of two kind of triangles only: equilateral or isosceles, with the two equal edges larger than the third one. The other kind of isosceles triangles, together with the scalene triangles, can not be obtained with any tern of equilibrium configurations, if these configurations are organized in an UM fashion.

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

Firstly, as far as probabilities cannot be negative, the previous relations implies that also for any J (with probability one) we have that $P_J^{12,23,31}(q_{12}, q_{23}, q_{31}) = 0$ as soon as the relations in Eq.(6) are not satisfied. This result has the consequence that any equilibrium configuration can be assigned (for fixed J) to a leave of a tree constructed in such a way that the overlap among two equilibrium configuration is related to the distance of the two configurations on the tree. Ultrametricity implies for example that, if two equilibrium configurations 1 and 2 are at overlap $q_{12} > q$, any equilibrium configuration 3 such that $q_{13} > q$ satisfies also the relation $q_{23} > q$. Ultrametricity is very interesting because it implies that many pdf 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 ultrametricity completely determines the $P^{12,23,31}$ given the $P(q)$. One can show that [9]

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

where

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

5

and

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

In other words stochastic stability and ultrametricity allow us to obtain all the pdf 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 to some general conclusions on these probability distributions without doing extra assumptions. We now show that two rather simple assumptions: replica equivalence (or equivalently stochastic stability) and separability are giving very strong constraints.

Let us firstly consider replica equivalence as formulated in 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 observable which involve only one replica are replica symmetric, i.e. the 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 arbitrary small random long range Hamiltonian (see [6] for a more careful discussion).

Eq.(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

$$P^{12,13}(q_{12}, q_{13}) = \frac{1}{2} P(q_{12}) P(q_{13}) + \frac{1}{2} P(q_{12}) \delta(q_{12} - q_{13}) \quad ,$$

$$P^{12,34}(q_{12}, q_{34}) = \frac{2}{3} P(q_{12}) P(q_{34}) + \frac{1}{3} P(q_{12}) \delta(q_{12} - q_{34}) \quad . \tag{11}$$

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

$$P^{12,13}(q_{12}, q_{13}) = \lim_{n \to 0} \frac{\sum'_{a,b,c=1,n} Q_{a,b} Q_{a,c}}{n(n-1)(n-2)} \quad ,$$

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

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

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

6

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

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

If we now look to the consequences of the previous equation and we use the relations in Eq.(12) both for $a = b$ and $a \neq b$, we obtain the two relations in Eq.(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. Eq.(11) determine all the joint pdf of 2 real replicas in terms of the $P(q)$. The consequences of stochastic stability have been lengthy discussed in [15, 17, 18]. In the 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 know as non-degeneracy) of the matrix Q [15], which correspond to the following statement. Let us consider all the matrices which can be generated from the matrix Q in a permutational covariant fashion. Some example are

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

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

$$Q_{ab} = Q_{cd} \implies M_{ab} = M_{cd} \qquad (16)$$

where M is a generic matrix of the set generated by the rules shown in Eq.(15). In other words pairs of indices which have different properties have different values of the overlap. It 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

$$O_i = \sum_k A_{i-k} \sigma_k,$$
$$O_i = \sum_{k,l} B_{i-k,i-l} \sigma_k \sigma_l, \qquad (17)$$
$$O_i = \sum_k \sigma_i \sigma_k J_{i,k}$$

where A and B are appropriate functions (e.g. they decrease sufficiently fast 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]:

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

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 fluctuates also 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 by 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 fluctuates also 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 happens only 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 a spontaneous magnetization.

A direct check of overlap equivalence can be done in 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 among 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 among 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 among two *objects* is fixed, all the mutual relations among the two objects are also fixed. The difference among 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 I we present some more detailed considerations which points 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 ultrametricity, 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. On the other end, 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 parameterized by λ (where $0 < \lambda \le 1$) and defined as

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

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

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

Therefore in this simply 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 Eq.(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.

Maybe the simplest non-trivial example of a non-ultrametric system is given by the union of separately ultrametric system with a non-trivial distribution of the overlaps [22]. It is easy to check that $q = \lambda q_1 + (1 - \lambda)q_2$ does not satisfy the ultrametric condition also if q_1 and q_2 do satisfy it. However it is clear that in this example q_1 and q_2 are generalized overlaps which are not function of q. Both ultrametricity and overlaps equivalence disappear at the same time [1].

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 let us consider two matrices M and R constructed with the rules in Eq. (15). It is evident that we can write

$$\sum_b Q_{ab}^k M_{ab} = \sum_b \int dq \delta(q - Q_{ab}) Q_{ab}^k M_{ab} = \int dq P(q) M(q) q^k \quad ,$$

$$\sum_b Q_{ab}^k R_{ab} = \sum_b \int dq \delta(q - Q_{ab}) Q_{ab}^k R_{ab} \int dq P(q) R(q) q^k \quad , \tag{21}$$

where we have defined

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

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

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

Therefore separability implies that quantities like those in Eq.(23) can be computed from the knowledge of those in Eq.(21).

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

$$\sum_b Q_{ab}^k M_{ab} = \int dq P_M(q) q^k \quad ,$$

---

[1] In the last example both ultrametricity 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 ultrametricity.

$$\sum_b Q^k_{ab} R_{ab} = \int dq P_R(q) q^k \quad , \tag{24}$$

$$\sum_b Q^k_{ab} M_{ab} R_{ab} = \int dq P_{MR}(q) q^k \quad , $$

the previous equations implies that

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

The last equation can also be written as

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

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

$$M_{ab} = \sum_c Q^{k_1}_{ac} Q^{k_2}_{cb} \quad , \tag{27}$$

$$R_{ab} = \sum_c Q^{k_3}_{ac} Q^{k_4}_{cb} \quad , \tag{28}$$

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

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

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

Eq.(29) is particular interesting because integrating over q it implies that

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

The previous equation is remarkable not only because it give the full expression of the probability with four overlap in terms of the probability with three overlaps but also because it puts hard constraints on the possible values of the function $P^{12,23,31}(q, q_1, q_2)$. Indeed the l.h.s. of Eq.(30) is by definition invariant under cyclic permutation of the q's, while the r.h.s. 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 satisfy Eq.(30)? We will argue in the next section 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, Eq.(11)] and with separability [and then in particular with Eq.(30)]. We will show that the most general matrix is the ultrametric one.

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

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

where the weights p_i are, by definition, positive and such that $\sum_i p_i = 1$, and the values q_i are different. Also the joint pdf of 3 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 tern (q_i, q_j, q_l) must be the same of any permutation of it, i.e. $p_{122} = p_{212} = p_{221}$. Then the number of really independent parameters in $P^{(3)}$ is $k(k+1)(k+2)/6$. More relations between the p_{ijl} can be obtained exploiting the following equation, which is based essentially on the first Guerra's relation

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

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 Eq.(30).

4.1 Three overlaps ($k = 3$)

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

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

while the Guerra's relations imply some equalities like

$$\begin{cases} \sum_j p_{11j} = \frac{1}{2} p_1^2 + \frac{1}{2} p_1 \quad , \\ \ldots \\ \sum_j p_{12j} = \frac{1}{2} p_1 p_2 \quad , \\ \ldots \end{cases} \tag{34}$$

11

We end with only 4 free parameters $(s, a_{32}, a_{31}, a_{21})$:

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

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, whether we exchange two of the overlap values, the forbidden isosceles triangle changes. For example, if $q_1 < q_2 < q_3$ then ultrametricity 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

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

while the reversed implication is not true, because UM also holds for different parameters values, e.g. $s = a_{31} = a_{21} = 0$ and $a_{32} = p_3 p_2 / 2$ (that 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 be all identically zero in order that UM hold ($\{s_i = 0\} \Longleftrightarrow \text{UM}$), 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 Eq.(30) to determine the values of all these parameters. The l.h.s. of Eq.(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 the right hand sides. The number of non-trivial equations we can obtain in this way is large enough to fix all the parameters.

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

$$-2a_{32}a_{31}p_2 p_1 + 2a_{32}a_{21}p_3 p_1 - 2a_{31}a_{21}p_3 p_2 + a_{31}p_3 p_2^2 p_1 + 2a_{32}p_3 p_1 s +$$
$$-2a_{32}p_2 p_1 s - 2a_{31}p_2 p_1 s + p_3 p_2 p_1 s - p_3^2 p_2 p_1 s + 2p_3 p_2 s^2 + 2p_3 p_1 s^2 = 0 \tag{37}$$

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

$$a_{31} = \frac{1}{2} p_3 p_1 - s - p_{311} \quad, \tag{38}$$

we can write Eq.(37) as

$$E_0 + 2(a_{32}p_3 p_1 + a_{21}p_3 p_2 + p_{311}p_2 p_1)s + 2(p_3 p_2 + p_3 p_1 + p_2 p_1)s^2 = 0 \quad, \tag{39}$$
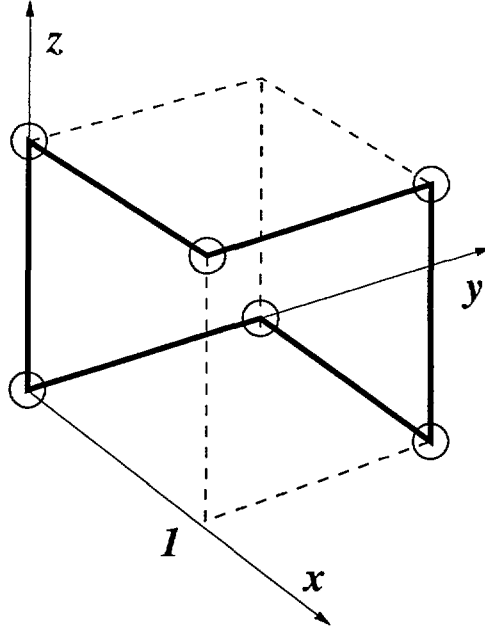
12

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

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

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

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

Let us now introduce the following symbol

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

For $x, y, z \in [0, 1]$ we have that $((x; y, z)) \geq 0$ and the equality $((x; y, z)) = 0$ only holds on 6 of the 12 edges of the cube (those in bold face in Fig. 1).

If we introduce the new parameters

$$a'_{ij} \equiv \frac{2a_{ij}}{p_i p_j} \quad , \tag{42}$$

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

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

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

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

13

For example, on the segment $a'_{32} = a'_{31} = 0$ and $0 < a'_{21} < 1$ they seem to co-exist non-zero probabilities p_{221} and p_{211} and it would be a small violation of UM. However we know that $a'_{32} = a'_{31} = a'_{21} = 0$ correspond to the ordering $q_1 < q_2 < q_3$, while $a'_{32} = a'_{31} = 0$ and $a'_{21} = 1$ correspond to $q_2 < q_1 < q_3$. Then we believe that the points on the segment between this two UM points corresponds to the case $q_1 = q_2 < q_3$, when there is no difference between p_{221} and p_{211} (but we can not still 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 ($k \geq 4$)

In this section we would like to sketch how the information we need about $P^{(3)}$ can be systematically derived from Eq.(30). To make this section more readable, the formulæ relatives to the cases $k = 4, 5$ will be presented in Appendix II. 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 Eq.(30) can be divided into three classes: those with 2, 3 and 4 different values of the overlaps. These equation are not independent: those with 2 (respectively 4) different overlaps can be expressed as the sum (resp. difference) of those with 3 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 2 and 3 overlaps, i.e. those which respectively correspond to the equalities

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

and

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

Each one of these equations can be identified giving a pair or a tern of numbers: (i, j) or (i, j, l). The l.h.s. 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 kind of Eq.(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 Eq.(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 II) 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 it is non-negative. Moreover in some of the $E^{(i,j)}$ expressions all the scalene parameters have positive defined coefficients, and then we should set them to zero in order to solve the equation, $E^{(i,j)} = 0$. In Appendix II we present a possible way of 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 Eq.(45), because each equation identified by (i, j, l) takes a very simple form:

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

where we choose the indexes such that $q_i < q_j < q_l$. In general, given 3 different overlap values, we can easily write down the corresponding equation [of the kind of Eq.(45)], which gives, once we set all the scalene parameters to zero, the corresponding double-parenthesis symbol [of the kind of Eq.(46)].

What about the equation with 2 and 4 different overlaps? When we set all the scalene parameters to zero, we have that the expressions $E^{(i,j)}$ are the sum of $k-2$ of these double-parenthesis symbols, those derived from the overlap terns (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 4 different overlaps are identically satisfied and they are useless.

Then we conclude that, in the more general solution, Eq.(46) must hold for every overlaps tern $q_i < q_j < q_l$. 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, 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 neighbors overlaps 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 the $P(q)$ is dense on a single compact domain, the distance between any pair of nearest neighbors overlaps tends to zero for $k \to \infty$ and then strict UM holds for any finite overlap difference $|q_i - q_j|$.

Lastly it should be noted that we have not exploited all the available information and maybe 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.

Maybe it should be also 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, using maybe 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 non trivial and the overlap is a fluctuating quantity, stochastic stability and separability imply ultrametricity. The reader should notice that the proofs presented here are likely to be too involved and it is quite possible that there is a direct proof of the fact that replica equivalence implies ultrametricity. At this end we

recall that in a dynamical approach it was show that one can identify a dynamical equivalent of separability: i.e. we can assume that in the aging regime all the possible overlaps among two configurations at two quite different times (t_1 and t_2) are functions of the usual overlap among 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 ultrametricity.

This result implies that, if we do not give up stochastic stability (which is a general property of generic equilibrium systems) violations of ultrametricity 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 of 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 done at high precision using present numerical technology. This could be done just doing numerical simulations in the ensemble with fixed overlap and looking to the fluctuations of the other overlaps.

Acknowledgments

Appendix I

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

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

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

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

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

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

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

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

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

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

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

with

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

where by the overbar we denote 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 end it is evident that

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

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

At this level the relation among 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 the same quantities does not fluctuate when the distance among i and k is fixed.

The two properties seems to be slightly different if the interaction is short range. On the contrary 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 and 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 too much the properties according to the principle of stochastic stability which tells us that the system should be stable with respect to a small random perturbation.

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

Appendix II

In this appendix we show some details of the computation we have done in the cases $k = 4$ and $k = 5$. 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 case $k = 4$ we have 4 scalene parameters $(p_{432}, p_{431}, p_{421}, p_{321})$ and 6 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 these 10 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's consider the equation of the kind of Eq.(44) with the two greatest overlaps (q_4 and q_3 in the $k = 4$ case),

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

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

$$
cE^{(4,3)} = cE_0^{(4,3)} +
$$

$$
+[2p_1 p_3 p_4 a_{43} + p_1 p_2 p_3(p_4 p_3 - 2a_{43} + p_4 p_2 - 2a_{42} + p_4 p_1 - 2a_{41})](p_{432} + p_{431}) +
$$

$$
+[4p_1 p_3 p_4]p_{432}p_{431} + [2p_1 p_4(p_2 + p_3)]p_{432}^2 + [2p_2 p_4(p_1 + p_3)]p_{431}^2 = 0 \quad . \tag{56}
$$

In $E_0^{(4,3)}$ there are 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-negative defined (see below).

The coefficients of p_{432} and p_{431} in Eq.(56) are positive defined, thanks to the inequalities $p_4 p_3 \geq 2a_{43}, p_4 p_2 \geq 2a_{42}, p_4 p_1 \geq 2a_{41}$, that are direct consequence of Eqs.(54) and of probabilities positiveness. Then we have that Eq.(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 enough to consider the equation analogous to Eq.(55) with q_1 and q_2 instead of q_3 and q_4. Once we set $p_{432} = p_{431} = 0$ we obtain

$$
cE^{(2,1)} = cE_0^{(2,1)} + [2p_1 p_3 p_4(a_{42} + a_{32}) + p_1 p_2 p_3(p_4 p_2 - 2a_{42} + p_4 p_1 - 2a_{41})]p_{421} +
$$

$$
+[2p_1 p_3 p_4(a_{42} + a_{32}) + p_1 p_2 p_4(p_3 p_2 - 2a_{32} + p_3 p_1 - 2a_{31})]p_{321} +
$$

$$
+[2p_3 p_4(p_1 + p_2)](p_{421} + p_{321})^2 = 0 \quad . \tag{57}
$$

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

Then we conclude that the more general solution to Eq.(30) in the case of $k = 4$ different overlaps forbids 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

$$
E_0^{(4,3)} = \frac{p_2 p_3 p_4}{4}((a'_{42}; a'_{43}, a'_{32})) + \frac{p_1 p_3 p_4}{4}((a'_{41}; a'_{43}, a'_{31})) \quad , \tag{58}
$$

$$
E_0^{(2,1)} = \frac{p_1 p_2 p_4}{4}((a'_{41}; a'_{42}, a'_{21})) + \frac{p_1 p_2 p_3}{4}((a'_{31}; a'_{32}, a'_{21})) \quad . \tag{59}
$$

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 easier to work with equations of the kind of Eq.(45). For example, considering the 3 overlaps $q_1 < q_2 < q_3$, we have that

$$0 = P^{12,23,34,41}(q_1, q_1, q_2, q_3) - P^{12,23,34,41}(q_1, q_2, q_3, q_1) = \frac{p_1 p_2 p_3}{4}((a'_{31}; a'_{32}, a'_{21})) \quad . \tag{60}$$

In general for every 3 given overlaps we end with an equation like Eq.(46).

The $k = 5$ case

The way to force the scalene parameters to zero should be now clear: it exploit the coefficient positiveness in the equations with 2 different overlaps. Maybe it is still not so clear if there is a systematic way to set all those 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's always consider first the equation with the two greatest overlaps (q_5 and q_4 in this particular case). It implies

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

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

Then let's substitute the just found solution ($p_{543} = p_{542} = p_{541} = 0$) into all the other equations and let's go forward in the same way

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

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

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

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

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

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

References

[1] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).

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

[3] K. H. Fischer and J. A. Hertz, *Spin Glasses*, Cambridge University Press (Cambridge, 1991).

[4] J.J. Ruiz-Lorenzo, E. Marinari and G. Parisi, F. Ricci-Tersenghi and F. Zuliani, preprint cond-mat/9906076, submitted to J. Stat. Phys.

[5] G. Parisi, Phys. Lett. A **73**, 203 (1979); Phys. Rev. Lett. **43**, 1754 (1979); J. Phys. A **13**, L115 (1980); J. Phys. A **13**, 1101 (1980); J. Phys. A **13**, 1887 (1980); Phys. Rev. Lett. **50**, 1946 (1983).

[6] F. Guerra, Int. J. Mod. Phys. B **10**, 1675 (1997) and references therein.

[7] M. Aizenman and P. Contucci, preprint cond-mat/9712129.

[8] A. Cacciuto, E. Marinari and G. Parisi, J. Phys. A **30**, L263 (1997). A. Cacciuto, preprint cond-mat/9704053.

[9] D. Iñiguez, G. Parisi and J.J. Ruiz-Lorenzo, J.Phys. A **29**, 4337 (1996).

[10] A. K. Hartmann, Europhys. Lett. **44**, 249 (1998).

[11] S. Franz and F. Ricci-Tersenghi, preprint cond-mat/9903130, to appear in Phys. Rev. E.

[12] D. S. Fisher and D. A. Huse, Phys. Rev. Lett. **56**, 1601 (1986); Phys. Rev. B **38**, 386 (1988).

[13] C. M. Newman and D. L. Stein, Phys. Rev. B **46** 973 (1992).

[14] G. Parisi, J. Stat. Phys. **51**, 51 (1993).

[15] G. Parisi, preprint cond-mat/9801081.

[16] G. Parisi, F. Ricci-Tersenghi and J. J. Ruiz-Lorenzo, Phys. Rev. B **57**, 13617 (1998).

[17] S. Franz, M. Mézard, G. Parisi and L. Peliti, Phys. Rev. Lett. **81**, 1758 (1998).

[18] S. Franz, M. Mézard, G. Parisi and L. Peliti, preprint cond-mat/9903370.

[19] G. Parisi, Physica Scripta **35**, 123 (1987).

[20] G. G. Athanasiu, C. Bachas and W. F. Wolff, Phys. Rev. B **35**, 1965 (1987).

[21] R. Rammal, G. Toulouse and M.A. Virasoro, Rev. Mod. Phys. **58**, 765 (1986).

[22] S. Franz, G. Parisi and M.A. Virasoro, Europhys. Lett. **17**, 5 (1992).

[23] D. Ruelle, Comm. Math. Phys. **108**, 225 (1987).