Stationary points of the Thouless-Anderson-Palmer free energy

Andrea Cavagna,* Irene Giardina,† and Giorgio Parisi‡
Dipartimento di Fisica, Università di Roma I ''La Sapienza,'' P.le A. Moro 5, 00185 Roma, Italy
and INFN Sezione di Roma I, Roma, Italy
(Received 4 November 1997)

In the context of the p-spin spherical model, we introduce a method for the computation of the number of stationary points of any nature (minima, saddles, etc.) of the Thouless-Anderson-Palmer free energy. In doing this we clarify the ambiguities related to the approximations usually adopted in the standard calculations of the number of states in mean-field spin-glass models. [S0163-1829(98)00118-0]

I. INTRODUCTION

Mean-field spin-glass models are characterized in their low-temperature phase by the great number of metastable as well as equilibrium states. A question which naturally arises in this context is the computation of the number $\mathcal{N}$ of these states or, more precisely, the analysis of how this number increases with the size N of the system.

In models with a continuous transition, such as the Sherrington-Kirkpatrick (SK) model,[1] the equilibrium thermodynamics is dominated by a number of states that remains finite when $N\to\infty$, while there is an exponentially high number of metastable states,[2] which do not contribute to the thermodynamics of the system. On the other hand, models with a discontinuous transition, such as the p-spin spherical model,[3–5] exhibit a temperature range where the number of metastable *and* equilibrium states with a given energy density E grows exponentially, i.e., $\mathcal{N}(E)\sim\exp[N\Sigma(E)]$.[6,7] In this last case knowledge of the *complexity* $\Sigma(E)$ is crucial, since it gives a finite entropic contribution to the global free energy.[8] It is therefore particularly important in this case to have a well-defined method to compute the number of states of the system.

The standard strategy to perform this calculation is grounded on the formulation of mean-field equations for the local magnetizations, the Thouless-Anderson-Palmer (TAP) equations.[9] The solutions of these equations are identified with equilibrium or metastable states of the system, and therefore one simply resorts to counting the number of these solutions.

This standard approach contains, however, some ambiguities. The TAP solutions can be viewed as the stationary points of a TAP free energy f_{TAP}, a function of the magnetizations.[2,6–8] Therefore only the minima of this free energy can actually be identified with metastable or equilibrium states of the system. Yet there are surely many other kinds of stationary points different from minima. When in the standard approach one counts the number of TAP solutions, it is not clear whether only the genuine states of the system are taken into consideration.

Moreover, a typical approximation of the standard method is related to the modulus of the determinant of the free-energy Hessian (i.e., the Jacobian of the equations), which appears in the integral over all the solutions.[10] The presence of this modulus is fundamental to avoid a trivial result: if one tries to count the number of stationary points of a function *without* this modulus, each stationary point is weighted with the sign of the Hessian and one obtains a simple topological constant, by virtue of the Morse theorem.[11] Nonetheless, in the standard approach this modulus is always disregarded to simplify the computation.

From what is said above we are led to say that the standard procedure is not really under control. Nonetheless, at least in the case of the p-spin spherical model, this standard calculation gives a result[7] that has been exactly confirmed by a completely different approach.[12] This result is therefore correct, although all the approximations involved are not well justified. On the other hand, for the case of the SK model there is no confirmation of the standard result of Ref. 2.

The aim of this paper is to clarify this subject, at least in the case of the p-spin spherical model. In the context of the replica approach, we show that different solutions of the saddle-point equations for the overlap matrix are related to different kinds of stationary points (minima, saddles, etc.). Grouping them into classes characterized by the number k of their instable directions, we find that each class has a different complexity $\Sigma_k(E)$. By virtue of this result, we are able to extract separately from the total number of solutions the contribution of minima and of saddles of various indices k, discovering that there is an ordering of the complexities $\Sigma_k(E)$: at a generic energy E, only *one* kind of stationary point (minimum or saddles, depending on the energy) is exponentially dominant over all the others, so that in the thermodynamic limit the weight of the sign of the determinant has no influence. Only at exceptional energies do we find that all the stationary points give an equal contribution to the complexity. Therefore, as long as the energy is kept fixed, the modulus can be disregarded and the standard approach gives the correct result. The results of the present work give also some insights into the nature of the glassy transition for this kind of system.

II. COMPLEXITY

The p-spin spherical model is defined by the Hamiltonian

$$H(s)=-\sum_{i_1<\cdots<i_p} J_{i_1\cdots i_p}s_{i_1}\cdots s_{i_p}. \tag{2.1}$$

The spins s are real variables satisfying the spherical constraint $\Sigma_i s_i^2 = N$, where N is the size of the system. The couplings J are Gaussian variables with zero mean and variance $p!/2N^{p-1}$. In the context of the TAP approach,[9] one formulates a set of mean-field equations for the local magnetizations $m_i = \langle s_i \rangle$. In Ref. 8 a free energy density $f_{\rm TAP}$ has been introduced, a function of the magnetizations m_i. The minimization of $f_{\rm TAP}$ with respect to m_i gives the TAP equations of the system. We can express the magnetization vector m in terms of its angular part σ and of its self-overlap $q = 1/N\Sigma_i m_i^2$:

$$m_i = \sqrt{q}\,\sigma_i, \quad \sigma \cdot \sigma = \sum_i \sigma_i^2 = N. \tag{2.2}$$

The TAP equation for σ reads[8]

$$0 = -p \sum_{i_2 < \cdots < i_p} J_{l,i_2 \cdots i_p} \sigma_{i_2} \cdots \sigma_{i_p} - pE\sigma_l \overset{\rm def}{=} \mathcal{T}_l(\sigma;E),$$

$$l = 1,\ldots,N, \tag{2.3}$$

where E is the zero-temperature energy density,

$$E = -\frac{1}{N} \sum_{i_1 < \cdots < i_p} J_{i_1 \cdots i_p} \sigma_{i_1} \cdots \sigma_{i_p}. \tag{2.4}$$

In the following we shall always refer to the zero-temperature energy density. The equations for σ do not depend on the temperature, while the equation for q does.[8] Moreover, the q equation has a solution as long as the energy density is lower than a maximum value of the energy, called the *threshold* energy $E_{\rm th}$. The dependence on temperature of the set of TAP solutions $\{m(T)\}_{\alpha=1,\ldots,\mathcal{N}}$ comes entirely from q, while their multiplicity $\mathcal{N}$ is encoded in Eq. (2.3) and thus does not depend on the temperature. It turns out that there is an exponentially high number of solutions of Eq. (2.3) for each given value of the energy density E, $\mathcal{N}(E) \sim \exp[N\Sigma(E)]$, where $\Sigma(E)$ is the complexity, computed for this model in Ref. 7. $\Sigma(E)$ is an increasing function of E, which reaches a finite value for $E = E_{\rm th}$. To avoid any confusion, we note that the TAP free-energy density of a solution at temperature T is unambiguously determined by its zero-temperature energy density E. Therefore in the following we shall always use E to label TAP solutions.

We start our analysis with the computation of $\Sigma(E)$, paying special attention to the nature of the stationary points actually considered. By definition we write

$$\Sigma(E) \overset{\rm def}{=} \lim_{N\to\infty} \frac{1}{N} \overline{\ln \mathcal{N}(E)}. \tag{2.5}$$

We average the logarithm of $\mathcal{N}$ since this is the extensive quantity. To perform this average it is necessary to introduce replicas already at this level of the calculation. However, it can be shown that the correct ansatz for the overlap matrix is symmetric and diagonal, and this is equivalent to average directly the number $\mathcal{N}$ of the solutions. Therefore we will perform the annealed computation

$$\Sigma(E) = \lim_{N\to\infty} \frac{1}{N} \overline{\ln \mathcal{N}(E)}. \tag{2.6}$$

In terms of the angular parts (2.2), we have

$$\Sigma(E) = \lim_{N\to\infty} \frac{1}{N} \ln \overline{\int \mathcal{D}\sigma\, \delta(\sigma\cdot\sigma - N) \prod_{l=1}^{N} \delta(\mathcal{T}_l(\sigma;E)) |\det \mathcal{H}(\sigma;E)|}, \tag{2.7}$$

where $\mathcal{H}(\sigma;E)$ is the Hessian of the TAP equations evaluated in the solution σ of energy density E. It is given by

$$\mathcal{H}_{r,l}(\sigma;E) = \frac{\partial \mathcal{T}_r(\sigma;E)}{\partial \sigma_l} = -p(p-1) \sum_{i_3 < \cdots < i_p} J_{r,l,i_3 \cdots i_p} \sigma_{i_3} \cdots \sigma_{i_p} - pE\,\delta_{r,l}. \tag{2.8}$$

We stress that by means of formula (2.7) we are counting only the solutions with a *given* energy density E. This is a crucial point: the principal effort of our discussion will be to show that, as long as $E < E_{\rm th}$, if we keep the energy fixed, the modulus in Eq. (2.7) can be dropped without affecting the result in the limit $N \to \infty$. We shall return to this point with greater detail at the end of our discussion. We therefore perform the calculation without the modulus, showing *a posteriori* which are the justifications of this procedure. Let us introduce a bosonic representation both for the determinant and the δ functions that implement the TAP equations:

$$\det \mathcal{H} = \lim_{n\to -2} \{\det \mathcal{H}\}^{-n/2} = \lim_{n\to -2} \int \mathcal{D}\phi^a \exp\left(-\frac{1}{2} \sum_{a=1}^{n} (\phi^a \mathcal{H} \phi^a) \right)$$

$$\prod_{l=1}^{N} \delta(\mathcal{T}_l(\sigma;E)) = \int \mathcal{D}\mu \exp(i\mu\cdot\mathcal{T}), \tag{2.9}$$

where the sums over repeated site indices are understood. The average over the disorder generates couplings between the fields ϕ, σ, and μ. A crucial approximation is to set equal to zero the couplings $\phi^a \cdot \sigma$ and $\phi^a \cdot \mu$ which depend on one replica index and which break the rotational invariance in the space of the replicas. We will see that this approximation is consistent with all the solutions we shall consider for the saddle-point equations. Thus we retain only the terms $\phi_a \cdot \phi_b$, $\mu \cdot \mu$, and $\mu \cdot \sigma$. It is easy to see that this approximation is equivalent to writing

$$\Sigma(E)=\lim_{N\to\infty}\frac{1}{N}\ln\int \mathcal{D}\sigma\,\delta(\sigma\cdot\sigma-N) \times\overline{\prod_{l=1}^{N}\delta(\mathscr{J}_l(\sigma;E))\det\mathcal{H}(\sigma;E)}. \tag{2.10}$$

Once averaged over the disorder, because of the spherical constraint, the part of the determinant does not depend on σ any more. Therefore we have

$$\Sigma(E)=A(E)+B(E), \tag{2.11}$$

$$A(E)=\lim_{N\to\infty}\frac{1}{N}\ln\int \mathcal{D}\sigma\,\delta(\sigma\cdot\sigma-N)\overline{\int \mathcal{D}\mu\,\exp(i\mu\mathscr{J})},$$

$$B(E)=\lim_{N\to\infty}\frac{1}{N}\ln\int \mathcal{D}\sigma\,\delta(\sigma\cdot\sigma-N) \times\overline{\lim_{n\to-2}\int \mathcal{D}\phi^a\exp\left(-\frac{1}{2}\sum_{a=1}^{n}(\phi^a\mathcal{H}\phi^a)\right)}. \tag{2.12}$$

The first integral does not involve replicas and gives the contribution

$$A(E)=\frac{1}{2}-\frac{1}{2}\ln\frac{p}{2}-E^2. \tag{2.13}$$

The second integral is more subtle to solve because it contains replicas and an appropriate ansatz has to be chosen to solve the saddle-point equations. Moreover, this integral is the one related to the Hessian of the TAP solutions, and thus it contains information on the nature of the solutions (minima, saddles, or maxima) that we are counting. Once averaged over the disorder and introducing the overlap matrix $Q_{ab}=-p(p-1)(\phi^a\cdot\phi^b)/2N$, we obtain

$$B(E)=\lim_{N\to\infty}\frac{1}{N}\ln\lim_{n\to-2}\int \mathcal{D}Q_{ab}\exp\left\{-N\left(\frac{\mathrm{Tr}Q^2}{2p(p-1)}+\frac{1}{2}\ln\det(-pE+Q)\right)\right\}. \tag{2.14}$$

As an ansatz for the matrix Q, we take $Q_{ab}=q_a\delta_{ab}$. In this way the exponent of Eq. (2.14) splits into n independent parts, each one giving the same saddle-point equation for q_a, whose possible solutions are

$$q_a=q_\pm=\frac{p}{2}(E\pm\sqrt{E^2-E_{\mathrm{th}}^2}),\quad E_{\mathrm{th}}=-\sqrt{\frac{2(p-1)}{p}}. \tag{2.15}$$

We restrict our discussion to $E\leqslant E_{\mathrm{th}}$, so that $q_\pm$ are real. Note that $q_+=q_-$ at the threshold energy $E=E_{\mathrm{th}}$. Since each q_a can assume one of these two values, we have a multiplicity of different solutions. The analysis of the fluctuations shows that there is a stable solution $\mathcal{S}_0$, given by

$$\mathcal{S}_0:\quad q_a=q_+,\quad a=1,...,n. \tag{2.16}$$

This solution is invariant under rotations in the replica space, and thus the approximation we made setting to zero the terms depending on one replica index turns out to be consistent. The solution $\mathcal{S}_0$ gives the complexity

$$\Sigma_0(E)=\frac{q_+^2}{p(p-1)}+\ln(-pE+q_+)+A(E), \tag{2.17}$$

with A given in Eq. (2.13). This is the known result of Ref. 7. It is important to note that this result has been confirmed in the analysis of Refs. 12 and 13 where, by means of a completely different method, it has been shown that Σ_0 is equal to the logarithm of the number of genuine states of the system and, thus, that Σ_0 is the complexity of the *minima* of the TAP free energy.

Nonetheless, we note the presence of many other solutions of the saddle-point equations, involving both the values $q_\pm$. In particular, we are interested in the solution $\mathcal{S}_1$ with the lowest degree of instability, that is,

$$\mathcal{S}_1:\quad q_1=q_-,\quad q_a=q_+,\quad a=2,...,n \tag{2.18}$$

(and permutations). This solution presents a one-step breaking of the rotational invariance in the replica space. Therefore one can be concerned about the fact that we have disregarded terms breaking this invariance. To check this point, we have performed the whole computation retaining the terms $\phi^a\cdot\sigma$ and $\phi^a\cdot\mu$ and we have looked for a solution breaking the rotational invariance in the replica space. We found analytically that the saddle-point equations give as a unique solution $\phi^a\cdot\sigma=0$ and $\phi^a\cdot\mu=0$ and thus that solution $\mathcal{S}_1$ is recovered. The complexity Σ_1 arising from $\mathcal{S}_1$ is

$$\Sigma_1(E)=\frac{3}{2}\frac{q_+^2}{p(p-1)}-\frac{1}{2}\frac{q_-^2}{p(p-1)}+\frac{3}{2}\ln(-pE+q_+) -\frac{1}{2}\ln(-pE+q_-)+A(E), \tag{2.19}$$

which is *lower* than Σ_0, since $|q_-|\geqslant|q_+|$,

$$\Sigma_1(E)<\Sigma_0(E)\quad\text{for } E<E_{\mathrm{th}}, \tag{2.20}$$

while

$$\Sigma_1(E_{\mathrm{th}})=\Sigma_0(E_{\mathrm{th}}). \tag{2.21}$$

In this context it is not clear which is the physical meaning of the complexity Σ_1 or if there is one. Moreover, apart from the fact that the complexity Σ_0 is confirmed by a different method to be related to the number of minima, we have given no justification for dropping the modulus in the original formula. We shall see in the next sections that the analysis of the average spectrum of the TAP Hessian gives an answer to both these questions.

III. HESSIAN SPECTRUM

In the previous section we made the approximation of setting to zero the couplings $\phi^a\cdot\sigma$ and $\phi^a\cdot\mu$. We stress that this approximation is consistent when considering the solutions $\mathcal{S}_0$ and $\mathcal{S}_1$. As a consequence, what appears in expression (2.10) is the Hessian function evaluated in a *generic* vector σ, and not in a TAP solution. This means that, in the

context of this approximation, the properties of the TAP Hessian that are relevant in determining the behavior of Σ are well encoded in the matrix $\mathcal{H}(\sigma;E)$, which has the same functional form of the TAP Hessian, but requires σ only to satisfy the spherical constraint. The average spectrum is then defined in the following way:

$$\rho(\lambda;E)=\lim_{N\to\infty}\overline{\int \mathcal{D}\sigma\ \delta(\sigma\cdot\sigma-N)\rho_J(\lambda;\sigma)}, \tag{3.1}$$

where $\rho_J(\lambda;\sigma)$ is the spectrum for a given realization of the disorder whose expression is

$$\rho_J(\lambda;\sigma)=-\frac{1}{N\pi}\,\mathrm{Im}\,\mathrm{Tr}(\mathcal{H}-\lambda+i\epsilon)^{-1}. \tag{3.2}$$

We can write the trace in the following way:

$$\begin{aligned}\mathrm{Tr}(\mathcal{H}-\lambda+i\epsilon)^{-1}&=\sum_{l=1}^{N}[(\mathcal{H}-\lambda+i\epsilon)^{-1}]_{ll}\\&=\lim_{n\to 0}\int \mathcal{D}\phi^a\phi^1\cdot\phi^1\\&\quad\times\exp\left\{-\frac{1}{2}\sum_{a=1}^{n}\phi^a(\mathcal{H}-\lambda+i\epsilon)\phi^a\right\}.\end{aligned} \tag{3.3}$$

Once averaged over the disorder J and the spherical constraint on σ is exploited, this computation becomes analogous to the one of the average spectrum of a Gaussian ensemble of symmetric random matrices.[14] If we introduce the overlap matrix $Q_{ab}=-p(p-1)(\phi_a\cdot\phi_b)/2N$, we finally get

$$\begin{aligned}\rho(\lambda;E)=\lim_{N\to\infty}-\frac{1}{N\pi}\,\mathrm{Im}\lim_{n\to 0}\int \mathcal{D}Q_{ab}(-pE-\lambda+Q)^{-1}_{11}\\\times\exp\left\{-N\left(\frac{\mathrm{Tr}Q^2}{2p(p-1)}+\frac{1}{2}\ln\det(-pE-\lambda+Q)\right)\right\}.\end{aligned} \tag{3.4}$$

It is important to note the great similarity between Eqs. (3.4) and (2.14). If we choose once again a diagonal ansatz $Q_{ab}=w_a\delta_{ab}$, we get the following solutions of the saddle-point equations:

$$w_a=w_\pm(\lambda)=\frac{p}{2}\left(\frac{\lambda}{p}+E\pm\sqrt{\left(\frac{\lambda}{p}+E\right)^2-E_{\mathrm{th}}^2}\right), \tag{3.5}$$

where E_{th} is the same as in Eq. (2.15). For $\lambda=0$ the integrand in Eq. (3.4) is identical to the one of Eq. (2.14) and $w_\pm(0)=q_\pm$. As in the case of the complexity, we have a multiplicity of different solutions. To get a finite contribution to ρ, it is necessary that the argument of the exponential in Eq. (3.4) be zero. Since $n\to 0$, this can be achieved taking the same value for each w_a. Moreover, the condition $\rho\geqslant 0$ shows that we must take the solution

$$\mathcal{S}_0:\quad w_a=w_+(\lambda),\quad a=1,...,n, \tag{3.6}$$

which is exactly the same kind of solution that led to Σ_0. If we look at Eq. (3.5), we can see that the nonzero contribution to ρ comes from the region $-pE+pE_{\mathrm{th}}<\lambda<-pE-pE_{\mathrm{th}}$, where w_+ develops an imaginary part. Thus

$$\rho_0(\lambda;E)=\frac{1}{\pi p(p-1)}\sqrt{p^2E_{\mathrm{th}}^2-(\lambda+pE)^2}. \tag{3.7}$$

We stress that the solution $\mathcal{S}_0$ is the only one that gives a finite contribution ρ_0 to ρ. Formula (3.7) is the well-known Wigner semicircle law,[15] which can be obtained for symmetric Gaussian random matrices also without using replicas.[14] This result tells us that for $E<E_{\mathrm{th}}$ the averaged spectrum has a strictly positive support, and thus the typical determinant of the Hessian is positive, i.e., that the dominant part of TAP solutions with energy density $E<E_{\mathrm{th}}$ are *minima*. On the other hand, when E approaches E_{th} the lowest eigenvalue $\lambda=p(E_{\mathrm{th}}-E)$ goes to zero. Therefore the typical solutions with $E=E_{\mathrm{th}}$ have some flat directions.[16]

We understand now the reason why the complexity of Eq. (2.17) is related to the number of minima: the solution $\mathcal{S}_0$ of the saddle-point equations leading to Σ_0 is exactly the same as the one leading to the eigenvalue distribution ρ_0, which has positive support.

The important thing is that in this context it is possible to give a precise physical interpretation of the solution $\mathcal{S}_1$ of Eq. (2.18): as we are going to show in the next section, $\mathcal{S}_1$ is related to the exponentially small corrections to the distribution ρ_0 and therefore gives information on those TAP solutions which are not minima.

IV. EXPONENTIAL TAILS AND COMPLEXITY OF THE SADDLES

For an ensemble of symmetric random matrices with a Gaussian distribution, it is possible to compute corrections to the semicircle law, when N is large but finite. In particular, it is possible to compute the correction to the averaged spectrum related to the probability of having a single eigenvalue outside the semicircle support.

In the context of our calculation, this can be achieved by considering solutions of the saddle-point equations for ρ different from $\mathcal{S}_0$. In particular, we are interested in corrections to ρ_0 in the eigenvalue region on the left of the semicircle region, i.e., for $\lambda<-pE+pE_{\mathrm{th}}$, since this tail contains the contribution of the negative eigenvalues. In this region we consider the solution $\mathcal{S}_1$,

$$\mathcal{S}_1:\quad w_1=w_-(\lambda),\quad w_a=w_+(\lambda),\quad a=2,...,n \tag{4.1}$$

(and permutations); from Eq. (3.4), we get

$$\begin{aligned}&\rho_1(\lambda,E)=r(\lambda,E)e^{-N\Delta(\lambda,E)},\\&\Delta(\lambda,E)>0\quad\text{for }\lambda<-pE+pE_{\mathrm{th}},\end{aligned} \tag{4.2}$$

which goes exponentially to zero as $N\to\infty$. In the computation of ρ_1, a crucial role is played by the fluctuations around the saddle-point solution $\mathcal{S}_1$, since the fluctuations matrix has an instable direction which provides the imaginary part necessary for ρ_1 to be nonzero outside the semicircle. One

can easily check that both $r(\lambda,E)$ and $\Delta(\lambda,E)$ coincide with the expressions obtained for the Gaussian random matrices with other methods.[14] This is therefore a correct result. The important quantity for our analysis is $\Delta(\lambda,E)$,

$$\Delta(\lambda,E)=\frac{w_-^2}{2p(p-1)}-\frac{w_+^2}{2p(p-1)}+\frac{1}{2}\ln\left(\frac{-\lambda-pE+w_-}{-\lambda-pE+w_+}\right). \tag{4.3}$$

Solution $\mathcal{S}_1$ then gives the exponentially vanishing left tail, due to the probability of having one eigenvalue outside the semicircle. Since this tail is different from zero also in the negative semiaxis, we can calculate the probability of having a negative eigenvalue, i.e., the exponentially small probability of finding a TAP solution which is a saddle with one negative eigenvalue and has energy density E. This probability is

$$P_{(-)}=\int_{-\infty}^{0} d\lambda\ \rho_1(\lambda,E)\sim e^{-N\Delta(0,E)},\quad N\to\infty. \tag{4.4}$$

In this context solution $\mathcal{S}_1$ has a clear physical interpretation: it is related to the contribution of TAP saddles with one negative eigenvalue, in the energy range $E<E_{\rm th}$. Given this, we can try to push further this interpretation. As we have seen in Sec. II, the same solution $\mathcal{S}_1$ gives rise to a complexity Σ_1 smaller than Σ_0, whose meaning was not clear. Now we can make the hypothesis that Σ_1 is the complexity of the saddles with one negative eigenvalue. To prove this statement we note that once we have the number $\mathcal{N}_1(E)\sim\exp[N\Sigma_1(E)]$ of saddles with one negative eigenvalue and energy density E, we can easily compute the probability $P_{(-)}$ of having one of these saddles,

$$P_{(-)}=\frac{\mathcal{N}_1(E)}{\mathcal{N}_{\rm total}(E)}=\frac{e^{N\Sigma_1(E)}}{e^{N\Sigma_0(E)}+e^{N\Sigma_1(E)}}\sim e^{-N[\Sigma_0(E)-\Sigma_1(E)]}, \tag{4.5}$$

where we used the relation $\Sigma_0(E)>\Sigma_1(E)$. From a comparison between Eqs. (4.5) and (4.4), we see that

$$\Delta(0,E)=\Sigma_0(E)-\Sigma_1(E) \tag{4.6}$$

must hold. It is not difficult to see from Eqs. (2.17), (2.19), and (4.3) that this equation is fulfilled. Our hypothesis is therefore correct, and we can then write

$$\Sigma_1(E)=\lim_{N\to\infty}\frac{1}{N}\ln\overline{\mathcal{N}_1(E)}, \tag{4.7}$$

where, as already said, $\mathcal{N}_1(E)$ is the number of TAP solutions of energy density E, which are saddles with one negative eigenvalue. This result can be generalized. If we consider the following solution $\mathcal{S}_k$ of the saddle-point equations for Σ,

$$\mathcal{S}_k:\quad q_a=q_-,\quad a=1,\ldots,k,\quad q_a=q_+,\quad a=k+1,\ldots,n \tag{4.8}$$

(and permutations), we obtain from Eq. (2.14) the complexity

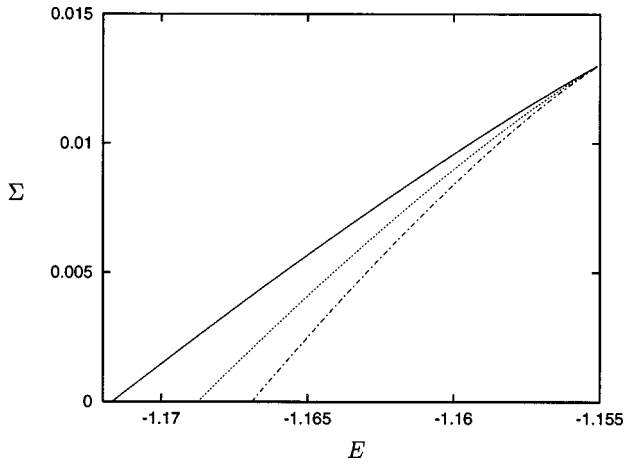

FIG. 1. Complexity Σ_0 of the TAP minima (solid line) and the complexities Σ_1 and Σ_2 of the TAP saddles with one and two negative eigenvalues (respectively, dotted and dot-dashed lines), as a function of the zero-temperature energy density E. The three curves reach the same value at the threshold energy, which is $E_{\rm th}=-1.1547$, for $p=3$. The minimum saddle energy, where $\Sigma_1=0$, is $E_0=-1.1688$.

$$\Sigma_k(E)=\frac{k+2}{2}\frac{q_+^2}{p(p-1)}-\frac{k}{2}\frac{q_-^2}{p(p-1)}+\frac{k+2}{2}\ln(-pE+q_+)-\frac{k}{2}\ln(-pE+q_-)+A(E). \tag{4.9}$$

It is not a surprise the fact that Σ_k is related to the number of TAP solutions which are saddles with k negative eigenvalues. Indeed, the probability of finding such a solution is

$$P_{(k,-)}=[P_{(-)}]^k\sim e^{-Nk\Delta(0,E)}, \tag{4.10}$$

so that to prove our assertion it is sufficient to verify that the relation

$$k\Delta(0,E)=\Sigma_0(E)-\Sigma_k(E) \tag{4.11}$$

holds, as it does. In writing Eq. (4.10) we can disregard the correlations between different negative eigenvalues, as long as k is much smaller than N. We conclude that, as a general result, $\Sigma_k(E)$ is the complexity of TAP saddles with k negative eigenvalues and energy density E.

Since for $E<E_{\rm th}$ $|q_-|>|q_+|$ holds [see Eq. (2.15)], we have from Eq. (4.9) that $\Sigma_0(E)>\Sigma_1(E)>\cdots>\Sigma_k(E)>\Sigma_{k+1}(E)\cdots$. Thus all the TAP solutions, also those with some negative eigenvalues, are exponentially numerous in N. Nevertheless, the number of minima is exponentially higher than the number of saddles with one negative eigenvalue, which is exponentially higher than the number of saddles with two negative eigenvalues, and so on. This is the very reason why, as long as $E<E_{\rm th}$, the approximation of dropping the modulus in Eq. (2.7) is justified. In Fig. 1 we have plotted Σ_0, Σ_1, and Σ_2 as a function of E.

From Eq. (4.9) we note that $\Sigma_k(E_{\rm th})=\Sigma_0(E_{\rm th})$, for each k, since $q_+=q_-$ at the threshold energy [see Eq. (2.15) and Fig. 1]. This equality is very important. If we try to count the *total* number of solutions neglecting the modulus in Eq.

(2.7), a trivial result is obtained,[10] since we are weighting each stationary point with the sign of the determinant (this is the Morse theorem). This is the reason why we considered solutions with a *given* fixed energy E. Yet for what is said above, if we integrate our result over all the energies E, we must recover the result predicted by the Morse theorem. Remembering that the q part of the TAP equations admits solutions only for $E<E_{\rm th}$, we have, from our calculations,

$$\int^{E_{\rm th}} dE \overline{\int \mathcal{D}\sigma\ \delta(\sigma\cdot\sigma-N)\prod_{l=1}^{N}\delta(\mathcal{J}_l(\sigma;E))\det\mathcal{H}(\sigma;E)} = a_0 e^{N\Sigma_0(E_{\rm th})}+a_1 e^{N\Sigma_1(E_{\rm th})}+a_2 e^{N\Sigma_2(E_{\rm th})}+\cdots . \qquad (4.12)$$

In this formula we must introduce all the Σ_k's coming from all the solutions of the saddle-point equations for Σ, which refer to stationary points of any nature. One can easily see from Eq. (4.9) that all the Σ_k's are monotonously increasing functions of E which reach their maximum value at $E_{\rm th}$, so that we can substitute the integral in Eq. (4.12) with the maximum of the integrand. The prefactors $a_0, a_1, \ldots$ come from the fluctuations around each saddle-point solution and contain the sign of the determinant. It is exactly the combination of these signs that gives rise to the Morse theorem. From Eq. (4.12) it is then clear that a necessary condition to get a trivial topological constant is that $\Sigma_0(E_{\rm th})=\Sigma_1(E_{\rm th})=\Sigma_2(E_{\rm th})=\cdots$, so that we can sum *all* the terms on the same footing. As said above, this necessary condition is fulfilled by our calculation.

Besides, from Eq. (4.12) it is finally clear what the role is of the modulus in the calculation: taking $|\det\mathcal{H}|$ is equivalent to taking the absolute value of the prefactors a_k, thus preventing us from obtaining a trivial result. Yet at fixed energy $E<E_{\rm th}$, one of the terms $\exp[N\Sigma_k(E)]$ is always strictly greater than all the others, and therefore in the limit $N\to\infty$ the signs of the prefactors a_k have no influence on the final result. As we have said, this dominant term turns out to be the one with $k=0$, which gives exactly the contribution of the minima.

We note that it should be possible to show that Σ_k is related to the number of saddles with a k negative eigenvalue directly from Eq. (2.14). If we keep n finite, this integral is equivalent to $(\det\mathcal{H})^n$. Taking the saddle-point solution $\mathcal{S}_k$ and appropriately computing the Gaussian fluctuations around it, it should be possible to single out a factor $(-1)^{kn}$ related to the sign of the determinant. Unfortunately, we did not succeed in performing this quite complex computation.

From Fig. 1 we see that there is a minimum energy density E_0 below which no saddles with finite complexity are found. Therefore, when considering a state with energy density $E<E_0$, the value $\Delta E=E_0-E$ is a lower bound for the energy density barrier between this state and any other state of the system. In Ref. 13 a potential function has been introduced, whose minima are by construction equivalent to metastable or equilibrium states of the system. With this method it has therefore been possible to give an estimate for the barriers separating two states.[17] It turns out that this estimate is fully consistent with the result of the present work.

V. CONCLUSIONS

The main result of this paper concerns the organization of the stationary points of the TAP free energy in the p-spin spherical model. If we classify these points according to the number k of negative eigenvalues of their Hessian, we find that each class is characterized by a complexity $\Sigma_k(E)$, which gives the exponentially high number of TAP solutions of energy E in that class, $\mathcal{N}_k(E)\sim\exp[N\Sigma_k(E)]$. In the energy range $E<E_{\rm th}$, we find that $\Sigma_k(E)>\Sigma_{k+1}(E)$ for each value of k. This means that in this energy range minima are exponentially dominant in number over all the other stationary points.

From what is said above, we conclude two things: First, if we compute, even in the most rigorous way, the complexity $\Sigma(E)$ at a given fixed energy, according to formula (2.5), we automatically recover $\Sigma_0(E)$, i.e., the complexity of the minima. Second, the modulus of the determinant simply contributes to the sign of the prefactor of the dominant contribution, since at fixed energy all the other terms are vanishing in the thermodynamic limit. Therefore, when such a structure of the stationary points is present, it is clear that the naive calculations which do not discriminate among minima, saddles, etc., and which disregard the modulus are, notwithstanding this, consistent.[6,7,18]

We stress that it is crucial to keep the energy fixed in the calculation, but more important is the fact that all the complexities are different, so that only one of them survives in the limit $N\to\infty$. This becomes clear when E is equal to the threshold energy $E_{\rm th}$: here all the Σ_k's are equal and a trivial result is recovered.

From a technical point of view, we note that the use of a bosonic representation for the determinant and the consequent replica approach introduces a degree of arbitrariness in the choice of the saddle-point solutions which makes it possible to extract the contributions of different classes of stationary points.

The results stated above give us a rigorous way to analyze the behavior of the Hessian spectrum. This turns out to be particularly interesting for the description of the dynamical transition which occurs in this kind of model. At high temperature equilibrium is given by the ergodic paramagnetic state. This phase corresponds to a Hessian spectrum which has a nonvanishing contribution in the negative semiaxis, so that many escape directions (negative eigenvalues) exist. At the dynamical transition temperature T_d, equilibrium is given by threshold TAP solutions.[8,16,19] As we have shown, at the threshold energy all the complexities Σ_k coincide (see Fig. 1), so that the free-energy landscape at the transition is strictly related to the point where the complexities of all the different TAP stationary points bifurcate. Here the Hessian spectrum exits the negative semiaxis and has a completely

positive support, so that escape directions become exponentially hard to find, while minima dominate the landscape: the system is trapped and a dramatic slowing down of the dynamics occurs.

To conclude, it would be intriguing to describe the glassy transition also in structural glasses simply in terms of the evolution of the Hessian spectrum, by means of an analysis similar to the one we have performed here.

*Electronic address: andrea.cavagna@roma1.infn.it
†Electronic address: irene.giardina@roma1.infn.it
‡Electronic address: giorgio.parisi@roma1.infn.it

[1] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).
[2] A. J. Bray and M. A. Moore, J. Phys. C **13**, L469 (1980).
[3] D. J. Gross and M. Mézard, Nucl. Phys. B **240**, 431 (1984).
[4] T. R. Kirkpatrick and D. Thirumalai, Phys. Rev. B **36**, 5388 (1987).
[5] A. Crisanti and H.-J. Sommers, Z. Phys. B **87**, 341 (1992).
[6] H. Rieger, Phys. Rev. B **46**, 14 655 (1992).
[7] A. Crisanti and H.-J. Sommers, J. Phys. I **5**, 805 (1995).
[8] J. Kurchan, G. Parisi, and M. A. Virasoro, J. Phys. I **3**, 1819 (1993).
[9] D. J. Thouless, P. W. Anderson, and R. G. Palmer, Philos. Mag. **35**, 593 (1977).
[10] J. Kurchan, J. Phys. A **24**, 4969 (1991).
[11] B. Doubrovine, S. Novikov, and A. Fomenko, *Geometrie Contemporaine* (Mir, Moscow, 1982), Vol. 2, Chap. 14.
[12] S. Franz and G. Parisi, J. Phys. I **5**, 1401 (1995).
[13] A. Cavagna, I. Giardina, and G. Parisi, J. Phys. A **30**, 4449 (1997).
[14] M. L. Metha, *Random Matrices* (Academic Press, New York, 1967).
[15] E. P. Wigner, *Proceedings of the Canadian Mathematical Congress* (University of Toronto Press, Toronto, 1957), reprinted in C. E. Porter, *Statistical Theories of Spectra: Fluctuations: a collection of reprints and original papers* (Academic Press, New York, 1965).
[16] L. F. Cugliandolo and J. Kurchan, Phys. Rev. Lett. **71**, 173 (1993).
[17] A. Cavagna, I. Giardina, and G. Parisi, cond-mat/9702069 (unpublished).
[18] A. Cavagna, I. Giardina, and G. Parisi, J. Phys. A **30**, 7021 (1997).
[19] A. Crisanti, H. Horner, and H.-J. Sommers, Z. Phys. B **92**, 257 (1993).