Marginal states in mean-field glasses

Markus Müller,[1] Luca Leuzzi,[2,3] and Andrea Crisanti[3]
[1]*Department of Physics and Astronomy, Rutgers University, 136 Frelinghuysen Road, Piscataway, New Jersey 08854, USA*
[2]*INFM-CNR, Via dei Taurini 19, 00185 Rome, Italy*
[3]*Department of Physics, University "La Sapienza," Piazzale Aldo Moro 5, 00185 Rome, Italy*
(Received 24 July 2006; published 31 October 2006)

We study mean-field systems whose free-energy landscape is dominated by marginally stable states. We review and develop various techniques to describe such states, elucidating their physical meaning and the interrelation between them. In particular, we give a physical interpretation of the two-group replica symmetry-breaking scheme and confirm it by establishing the relation to the cavity method and to the counting of solutions of the Thouless-Anderson-Palmer equations. We show how these methods all incorporate the presence of a soft mode in the free-energy landscape and interpret the occurring order-parameter functions in terms of correlations between the soft mode and the local magnetizations. The general formalism is applied to the prototypical case of the Sherrington-Kirkpatrick-model, where we reexamine the physical properties of marginal states under a new perspective.

DOI: 10.1103/PhysRevB.74.134431 PACS number(s): 75.10.Nr, 11.30.Pb, 64.60.Cn

I. INTRODUCTION

Disordered, frustrated systems often possess a multitude of nearly degenerate metastable states that are at the basis of glassy phenomena such as slow relaxation to equilibrium and aging. In order to understand the dynamics of such systems, it is important to characterize their corrugated free-energy landscape, as well as the physical properties and abundance of metastable states, usually described by the complexity function (also called configurational entropy in the context of amorphous solids).

In finite-dimensional, short-range interacting systems, it is not possible to rigorously define metastable states since nucleation phenomena always restore ergodicity on sufficiently long time scales. However, in mean-field models where an analytical description of the free-energy landscape in terms of local order parameters (magnetizations) is available, metastable states can be identified as stable stationary points of the free-energy functional (see, e.g., Ref. 1 for an instructive review of the state of the art). The properties of the local neighborhood of the latter allow for a natural classification: (i) genuine minima (with a positive definite free-energy Hessian),[2] and (ii) *marginal states* (with eigenvalues of the Hessian tending to zero in the thermodynamic limit).[3–5]

Recent studies of the Ising p-spin model at low temperature have led to the following picture for the metastable states as a function of their free-energy density: at low free-energy densities, most metastable states are genuine minima, while above some threshold, free-energy entropy is dominated by marginal states with a single soft mode.[6,7] The same phenomenology is found in mixed spherical $s+p$-spin models.[8] In the special case of the spherical p-spin model,[9,10] the dominant metastable states are genuine minima at all free-energy densities up to the threshold of dynamic arrest, above which there are *no* stable states anymore.[2]

In certain models, however, marginal states are exponentially more numerous than genuine minima *at all free-energy densities*. This situation is expected to occur generically in models whose thermodynamics is described by continuous replica symmetry breaking, the best studied case being the Sherrington-Kirkpatrick (SK) model.[3,5,11]

The studies of the above models have shown that very often marginal states possess only one single soft mode. Such states require a special treatment both in replica and cavity approaches, and we will focus on their description in the present study. Under certain circumstances, however, marginal states with a large number (diverging in the thermodynamic limit) of soft directions in the free-energy landscape may occur, the best known example being the thermodynamically dominant state of the SK model below the critical temperature T_c.[12] This global free-energy minimum is a stationary point with many almost flat directions in the free-energy landscape. Since there is a multitude of marginal directions, none of them can be singled out *a priori*. It is thus not clear whether the methods presented here, tailored to the presence of a single marginal mode, may still apply to that situation.[13]

The dynamical behavior of a given model will strongly depend on the local environment of the metastable states that dominate the landscape in the range of dynamically accessible free energies (usually energies where marginal state dominate). To date, the role of higher-lying marginal states in the slow dynamics of mean-field glasses remains unclear. However, the characterization of their local properties presented here should help to identify their traces in future analytical or numerical studies of glassy dynamics.

In this paper, we focus on the analytical description of marginal states with a single soft mode and show what physical information is contained in three choice methods to describe them: the two-group replica formalism, the generalized cavity method, and the counting of the solutions of the Thouless-Anderson-Palmer (TAP) mean-field equations. In particular, we show that the emerging order parameters contain information on the correlation between the marginal mode and the local magnetizations. Further, we discuss the computation of the distribution of (frozen) local fields and infer further key characteristics of the local environment of

marginal states, such as the spectrum of the free-energy Hessian.

From a detailed analysis of the SK model, it will become clear that the compact two-group replica method[14,15] is an effective means to describe marginal metastable states in mean-field systems. We compare this approach with other methods and exhibit their equivalence at the level of the annealed approximation.

The paper is organized as follows. In Sec. II, we briefly review various approaches to stable and marginal states and summarize the current knowledge. In Sec. III, we introduce an exactly solvable toy model whose physics is dominated by marginal states. By analyzing it with the help of a simplified two-group ansatz, we gain an understanding of the physical meaning of the formalism that appears in the more complicated mean-field models studied in Sec. IV. In Sec. V, we relate the two-group calculation to the direct counting of solutions of the TAP equations. Rederiving the TAP complexity following Bray and Moore, we exhibit the equivalence with the two-group approach. In Sec. VI, we review and extend the cavity method adapted to marginal states. We show its equivalence with the two-group formalism and confirm the interpretation of the order parameters. In Sec. VII, we build the formalism for a *quenched* two-group computation and discuss its physical content, in particular the distribution of local fields and soft-mode components. In Sec. VIII, we recall the criteria for internal thermodynamic consistency and local stability, and discuss possible scenarios for the evolution of the local properties of metastable states as their free energy decreases. In Sec. IX, we analyze the local field distribution in the uncorrelated, high-energy regime at low temperatures, and speculate on its consequences for the dynamics. Finally, in Sec. X, we discuss various open questions and possible future extensions of the presented methods, concluding with a brief summary in Sec. XI.

II. CLASSIFICATION OF METASTABLE STATES

The possible techniques to investigate metastable states in mean-field systems with quenched disorder are (i) the replica method with an ultrametric replica symmetry breaking (RSB) Parisi ansatz,[16] (ii) the direct counting of the solutions of the Thouless-Anderson-Palmer (TAP) equations,[17,18] (iii) and the cavity method.[19] All these approaches can be used to describe stable states and to compute their properties as well as their complexity, i.e., the logarithm of the number of states at given free-energy density.

However, in situations where the most numerous states are marginal, these techniques need to be generalized.[20,21] As explained in the Introduction, this generally occurs at sufficiently high free-energy densities, and in certain models even at all free energies (in the SK model below T_c, for instance). The characteristics of the three equivalent approaches to stable states, and their generalizations adapted to marginal states, are summarized in Table I.

TABLE I. Table summarizing the methods to study properties of metastable states in mean-field glasses. The second column refers to genuine minima, while the third column describes the necessary generalization in the case of *marginal states* with a single soft mode.

	Minima	Marginal
Replica	Parisi RSB ansatz	Two-group+RSB ansatz
Cavity	States robust to addition of a spin Single cavity field	Fragile pairs of saddles and minima Additional field related to soft mode
Counting of TAP solutions	Saddle point of TAP action conserves fermionic symmetry	Saddle point of TAP action breaks fermionic symmetry

A. Prevalence of genuine minima

At free-energy densities where the vast majority of metastable states are minima of the free-energy landscape, any thermodynamic function (including the complexity) is correctly described by Parisi's RSB ansatz, by the standard cavity method, or by counting stable solutions of the TAP equations (imposing a saddle point that preserves a fermionic symmetry of the problem[23,24]).

Mean-field models with an ultrametric organization of states in two levels (a "one-step" structure) usually exhibit this kind of landscape at low enough free energies. Their static properties can be computed either using a one-step RSB ansatz[25] or the cavity method including cluster correlations via the so-called "reweighting."[19] Whenever the one-step RSB solution is stable, the complexity of stable minima can be calculated using Monasson's method as the Legendre transform of the free energy of m coupled clones.[26]

It is a nontrivial fact that the same result can be obtained by counting the number of solutions of the TAP equations,[18,27] which requires the saddle point extremization of a certain action functional. Under the assumption that every solution represents a stable minimum, the corresponding saddle point has to preserve the fermionic Becchi-Rouet-Stora-Tyutin (BRST) symmetry of the action[22] which generically occurs in the description of stochastic equations.[24,28]

B. Prevalence of marginal states

In those cases in which marginal states dominate (being exponentially more numerous than genuine minima), the approaches mentioned above fail because they assume the states to be stable minima. However, all approaches can be suitably generalized to deal with the marginal case, too. The purpose of this paper is to examine the physical meaning of these generalizations, namely (i) the extension of the RSB scheme to a two-group ansatz,[14,15,29] (ii) the inclusion of an extra auxiliary field in the cavity method,[20,21] and (iii) the counting of TAP solutions via a saddle point that breaks the BRST symmetry.[3,18] While some aspects of the generalized cavity method and the breaking of the BRST symmetry have been interpreted previously in the literature, the meaning of the two-group ansatz, as well as the relation between the above methods, have not been established. The present paper

tries to fill this gap. Furthermore, we will show how to extract so far hidden information from the formalism and give the interpretation of the emerging order-parameter functions.

Before applying the two-group approach to the SK model, we shall introduce a simple, exactly solvable toy model that illustrates the basic features of the marginal states and their appearance within a replica language. This will be a conceptual guide to the physics discussed later in the context of the more complicated SK model.

III. A SIMPLE MODEL TO UNDERSTAND THE TWO-GROUP ANSATZ AND MARGINAL STATES

The two-group ansatz was first introduced by Bray and Moore in Ref. 14, in an attempt to resolve the instability of the replica symmetric solution found by Sherrington and Kirkpatrick[11] in the mean-field approximation of the Edwards-Anderson model.[30] Even though it did not turn out to be the correct way to describe the equilibrium (the stable equilibrium solution with ultrametric replica symmetry breaking was found soon after by Parisi[16]), this two-group ansatz endowed with an ultrametric structure can be used to study the free-energy landscape *above* the thermodynamic ground state.

This replica symmetry breaking scheme consists in dividing m replicas in two nonequivalent groups of $m+K$ (group "+") and $-K$ (group "−") elements, respectively, and computing the corresponding "replicated free energy." A subsequent Legendre transform with respect to the total number m of replicas[26] was found to reproduce Bray and Moore's calculation of the TAP complexity.[15,18,31]

A very similar scheme of replica symmetry breaking with groups of $K \to \pm\infty$ replicas occurred in the mean-field description of the random field Ising model,[28,32,33] where "instantons" with this kind of two-group structure were used to identify a certain class of Griffith-like singularities of the free energy,[34] the physics of which will become clear from the following toy model.[35]

A. The meaning of the two-group limit in a toy model

In this section, we examine an exactly solvable zero-dimensional model that contains the basic ingredients for an understanding of the two-group ansatz.

From the replica solution of this toy model, we will infer that the two groups should be interpreted as representing a minimum and a saddle point that merge into a single marginally stable configuration in the limit $K \to \infty$. Furthermore, this picture will allow us to obtain a physical interpretation of the order parameters appearing in the two-group ansatz, as we will later confirm using the equivalence with the physically more intuitive cavity approach.

Let us consider the simple model of a particle in a potential $V(\phi)$ and subject to a random field h with probability distribution $P(h)$,

$$\mathcal{H}(\phi;h) = V(\phi) - h\phi. \tag{1}$$

This can be considered as the zero-dimensional case of the problem of pinned manifolds, such as the random field or random temperature Ising model, where a very similar analysis leads to the description of Griffith singularities.[34]

We consider potentials such that for typical fields h the Hamiltonian $\mathcal{H}$ possesses only one minimum at $\phi_{\mathrm{I}}(h)$. A secondary minimum occurs at $\phi_{\mathrm{II}}(h)$ only for rare fields that are larger than a critical value $h > h_c$. A simple example for such a potential is given by

$$V(\phi) = h_c(\phi - \phi_c) + \frac{V_3}{6}(\phi - \phi_c)^3, \tag{2}$$

subject to the constraint $\phi \geqslant 0$. In this case, we have $\phi_{\mathrm{I}}(h) \equiv 0$ for all random fields of practical interest, and for $h \geqslant h_c$ there is a secondary minimum $\phi_{\mathrm{II}}(h)$. The latter becomes marginal as $h \to h_c$, while ϕ_{II} approaches $\phi_{\mathrm{II}}(h_c) = \phi_c$.

At sufficiently low temperature $[T \ll V(\phi_c) - V(0)]$, the disorder average of the free energy behaves as

$$\bar{F} \equiv -\beta^{-1}\,\overline{\ln Z} = -\beta^{-1}\int dh\, P(h) \ln\left[\int d\phi e^{-\beta\mathcal{H}(\phi,h)}\right]$$
$$\approx V(0) + \int_{h_c}^{\infty} dh P(h) e^{-\beta[\mathcal{H}(\phi_{\mathrm{II}}(h);h) - \mathcal{H}(\phi_{\mathrm{I}}(h);h)]}, \tag{3}$$

where we expanded the logarithm and made a saddle-point approximation to obtain the second term describing the Griffith contribution from the secondary minimum in rare random fields.

For simplicity, we will carry out explicit calculations for the case of Gaussian random fields with distribution

$$P(h) = \frac{1}{\sqrt{2\pi h_0^2}} \exp\left[-\frac{h^2}{2h_0^2}\right]. \tag{4}$$

In this case, the integral in Eq. (3) is dominated by its lower boundary for $T \gtrsim T^* \equiv \phi_c h_0^2 / h_c$,

$$\bar{F} \approx V(0) + \text{const} \times P(h_c) e^{-\beta \Delta E_c}, \tag{5}$$

where ΔE_c is the energy difference between the primary and secondary minimum, evaluated at the critical field strength, h_c,

$$\Delta E_c = V(\phi_c) - V(\phi_{\mathrm{I}}) - h_c(\phi_c - \phi_{\mathrm{I}}). \tag{6}$$

1. Derivation with vectorial replica symmetry breaking

It is instructive to rederive this simple result with a replica calculation. Dotsenko and Mézard[32] found that the exact low-temperature partition function of similar disordered single-particle models could be reproduced by summing a class of saddle points that break replica symmetry in a nonstandard (i.e., nonultrametric), "vectorial" way. The physical content of this recipe will become clearer below, where we will show that a generalization of their scheme applied to the model Eq. (1) indeed gives back the anticipated result of Eq. (5).

For a Gaussian field distribution, Eq. (4), the replicated and averaged partition function reads

$$\overline{Z^n} = \int \prod_{a=1}^{n} d\phi_a \exp[-\beta F(\{\phi_a\})]$$
$$= \int \prod_{a=1}^{n} d\phi_a \exp\left[-\beta \sum_{a=1}^{n} V(\phi_a) + \frac{h_0^2}{2}\left(\beta \sum_{a=1}^{n} \phi_a\right)^2\right]. \tag{7}$$

For arbitrary distributions $P(h)$, characterized by their cumulants $c_r \equiv \langle h^r \rangle_c$, the replica free-energy functional generalizes to

$$F(\{\phi_a\}) = \sum_{a=1}^{n} V(\phi_a) - \sum_{r=1}^{\infty} \frac{1}{\beta}\frac{c_r}{r!}\left(\beta \sum_{a=1}^{n} \phi_a\right)^r. \tag{8}$$

Following the recipe by Dotsenko and Mézard,[32] we approximate the integral in Eq. (7) by determining all stable saddle points of the free-energy functional $F(\{\phi_a\})$, and summing their respective Boltzmann weights,

$$-\beta \bar{F} \equiv \overline{\ln Z} = \lim_{n\to 0} \frac{\overline{Z^n} - 1}{n} \approx \frac{Z_{\mathrm{RS}} - 1}{n} + \frac{Z_{\mathrm{RSB}}}{n}. \tag{9}$$

The partition function $Z_{\mathrm{RS}} = \exp[-\beta F_{\mathrm{RS}}]$ denotes the contribution of the replica symmetric saddle point with $F_{\mathrm{RS}} \equiv F(\{\phi_a = \phi_{\mathrm{RS}}\}) = n f_{\mathrm{RS}}$, and Z_{RSB} is the sum over all saddle points breaking the replica symmetry. As we will see shortly, the latter always comes with a combinatorial factor proportional to n, which cancels the denominator.

Let us show that this recipe allows us to rederive Eq. (5). We make the most general ansatz for the configuration $\{\phi_a\}$ of a saddle point, collecting the ϕ's with identical values into M groups labeled by $i = 1, \ldots, M$, each with k_i replicas, i.e.,

$$\phi_a = \phi_1, \quad a = 1, \ldots, k_1,$$
$$\phi_a = \phi_2, \quad a = k_1 + 1, \ldots, k_1 + k_2,$$
$$\vdots$$
$$\phi_a = \phi_M, \quad a = \left(\sum_{i=1}^{M-1} k_i\right) + 1, \ldots, n.$$

In the Gaussian case, the corresponding replica free energy evaluates to

$$F_{\{k_i\}} = \sum_{i=1}^{M} k_i V(\phi_i) - \frac{\beta h_0^2}{2}\left(\sum_{i=1}^{M} k_i \phi_i\right)^2. \tag{10}$$

The saddle-point equations with respect to variation of ϕ_i read

$$V'(\phi_i) = h_\phi, \quad \forall\, i, \tag{11}$$

$$h_\phi \equiv \beta h_0^2 \sum_{i=1}^{M} k_i \phi_i. \tag{12}$$

For a regular potential $V(\phi)$, the number M of different solutions of Eq. (11) is limited. In particular, for the potential in Eq. (2) we have $M=3$ for $h_\phi > h_c$, $M=2$ for $h_\phi = h_c$, and $M=1$ for $h_\phi < h_c$. The latter corresponds to the replica symmetric saddle point with $h_\phi = 0$, $\phi_a = 0$, $F_{\mathrm{RS}} = V(0)$.

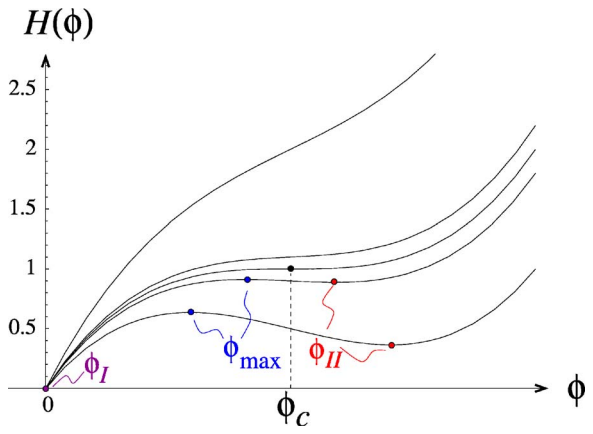

FIG. 1. (Color online) Hamiltonian $H(\phi) = V(\phi) - h\phi$ with $V(\phi)$ from Eq. (2), plotted for $h/h_c = 0$, 0.9, 1, 1.1, and 1.5 (from top to bottom). The secondary minimum ϕ_{II} and the saddle $\phi_{\max}$ merge for $h \to h_c$.

Replica symmetry breaking saddle points exist for $h_\phi \geqslant h_c$, where we label the global minimum, the secondary minimum, and the local maximum of $\mathcal{H}$ by ϕ_{I}, ϕ_{II}, and $\phi_{\max}$, respectively, leaving their dependence on h implicit (cf. Fig. 1). In the limit of $n \to 0$, Eq. (12) yields the self-consistency equation for h_ϕ,

$$k_1(\phi_{\mathrm{I}} - \phi_{\max}) + k_2(\phi_{\mathrm{II}} - \phi_{\max}) = \frac{h_\phi}{\beta h_0^2}. \tag{13}$$

Notice that we choose the number of replicas in the minima, k_1 and k_2, to be positive, leaving a negative number $n - k_1 - k_2 \to -(k_1 + k_2)$ of replicas in $\phi_{\max}$. This is necessary to ensure a positive Hessian of the free-energy functional Eq. (8).

Let us now consider the saddle-point free energy F_{k_1,k_2}, Eq. (10), as a function of k_2 for fixed k_1. For not too low temperatures $T > T^*$, F_{k_1,k_2} decreases as k_2 increases. Hence, the most important contribution to the sum over saddles comes from large $k_2 \gg 1$. As is evident from Eq. (13), this requires ϕ_{II} and $\phi_{\max}$ to approach each other, and hence h_ϕ must be nearly critical. More precisely, one finds

$$\phi_{\mathrm{II}} \approx \phi_c + \frac{h_c}{2\beta h_0^2}\frac{1}{k_2} + \cdots, \tag{14}$$

$$\phi_{\max} \approx \phi_c - \frac{h_c}{2\beta h_0^2}\frac{1}{k_2} + \cdots, \tag{15}$$

$$h_\phi \approx h_c + \frac{V_3}{2}\left(\frac{h_c}{2\beta h_0^2}\right)^2 \frac{1}{k_2^2} + \cdots, \tag{16}$$

$$F_{k_1,k_2} = k_1 \Delta E_c + \frac{h_c^2}{2\beta h_0^2} + O\left(\frac{1}{k_2^2}\right). \tag{17}$$

The total Griffith contribution to the free energy results from the sum over all RSB saddle points with the corresponding multiplicity,

$$
\begin{aligned}
-\beta F_{\mathrm{RSB}} &= \lim_{n\to 0} \frac{Z_{\mathrm{RSB}}}{n} \\
&= \lim_{n\to 0} \frac{1}{n} \sum_{k_1>0,k_2\geqslant 0} \binom{n}{k_1,k_2} e^{-\beta F_{k_1,k_2}} \\
&= \sum_{k_1>0,k_2\geqslant 0} \frac{(-1)^{k_1+k_2-1}(k_1+k_2-1)!}{k_1!k_2!} e^{-\beta F_{k_1,k_2}}.
\end{aligned}
\tag{18}
$$

In order to recover Eq. (5), we need to exclude saddles with $k_1=0$. This can be understood on physical grounds: the saddle-point free energies $F_{k_1=0,k_2}$ are independent of the ground-state level ϕ_{I}. The corresponding terms are not suppressed by the Boltzmann weight associated with the excitation probability to the state ϕ_{II}. In other words, these terms do not know about the ground state and hence must be discarded for the calculation of the Griffith correction.[36]

We note that for any fixed $k_1>0$, the sum over k_2 is weakly divergent and has to be regularized appropriately. Since the sum is dominated by large k_2, we use Eq. (17) to approximate the Griffith contribution

$$
\begin{aligned}
-\beta F_{\mathrm{RSB}} &\approx \left\{ \sum_{K>0} \frac{(-1)^{K-1}}{K} \sum_{k_2=0}^{K} \binom{K}{k_2} e^{-\beta(K-k_2)\Delta E_c} \right. \\
&\quad \left. - \sum_{k_2>0} \frac{(-1)^{k_2-1}}{k_2} \right\} e^{-h_c^2/2h_0^2} \\
&= e^{-h_c^2/2h_0^2} \left\{ \sum_{K>0} \frac{(-1)^{K-1}}{K} (1+e^{-\beta\Delta E_c})^K - \ln 2 \right\} \\
&= e^{-h_c^2/2h_0^2} \ln\left(1+\frac{e^{-\beta\Delta E_c}}{2}\right) \approx \frac{1}{2} e^{-h_c^2/2h_0^2} e^{-\beta\Delta E_c} \\
&\approx P(h_c) e^{-\beta\Delta E_c}.
\end{aligned}
\tag{19}
$$

and we recover indeed the Griffith correction, Eq. (5), due to rare secondary minima.

An analogous analysis can be carried out in the case of non-Gaussian distributions $P(h)$. Instead of Eq. (17), one obtains the saddle-point free energy

$$
F_{k_1,k_2\to\infty} = k_1 \Delta E_c - \beta^{-1} S_c + O(1/k_2^2), \tag{20}
$$

where $S_c \approx \log[P(h_c)]$. More precisely, $\exp(S_c)$ is given by the saddle-point approximation of the integral representation,

$$
\begin{aligned}
P(h_c) &= \int \frac{d\lambda}{2\pi} \exp\left(-i\lambda h_c + \sum_{r=1}^{\infty} \frac{c_r}{r!} (i\lambda)^r \right) \\
&\approx \operatorname*{ext}_{\lambda_*} \left[\exp\left(-\lambda_* h_c + \sum_{r=1}^{\infty} \frac{c_r}{r!} \lambda_*^r \right) \right],
\end{aligned}
\tag{21}
$$

which is a good approximation provided that h_c belongs to the tail of $P(h)$. The Griffith term is again dominated by $k_1=1$ and $k_2\to\infty$, yielding $\Delta F_c \approx P(h_c)\exp[-\beta\Delta E_c]$.

In summary, the above toy model describes a simple Griffith phase that is dominated by marginal configurations. This physics is exactly reproduced by a vectorial replica symmetry breaking that divides the replicas into two (infinite) groups describing a coalescing pair of a minimum and a saddle point (ϕ_{II} and $\phi_{\max}$).

B. Generalization to higher dimensions

To make contact with more complicated models, it is instructive to generalize our simple model to a particle in d dimensions, $\vec{\phi} \in \mathbb{R}^d$, subject to Gaussian random fields

$$
\mathcal{H}(\vec{\phi};\vec{h}) = V(\vec{\phi}) - \vec{h}\cdot\vec{\phi}. \tag{22}
$$

The minima of Eq. (22) satisfy $\vec{\nabla} V(\vec{\phi}) = \vec{h}$. The Griffith contributions to the quenched free energy are dominated by rare fields $\vec{h}_c$ that are just strong enough to admit a marginal state $\vec{\phi}_c = \vec{\phi}(\vec{h}_c)$. The marginality implies the presence of a soft mode in the Hessian of $\mathcal{H}$, i.e.,

$$
\det[\mathrm{Hess}(\vec{\phi}_c)] \equiv \det\left[\frac{\partial^2 V(\vec{\phi}_c)}{\partial\phi_i \partial\phi_j} \right] = 0. \tag{23}
$$

This condition defines a hypersurface $\mathcal{S}$ in the space of fields $\vec{h}$. The dominating Griffith contribution derives from $\vec{h}_c \in \mathcal{S}$, which maximizes $P(\vec{h}_c)\exp\{-\beta[\mathcal{H}(\vec{\phi}_c)-\mathcal{H}(\vec{\phi}_{\mathrm{I}})]\}$. In the replica formalism, this condition is conveniently encoded in the saddle-point equation for $k_1=1$ and $k_2\to\infty$, analogous to Eq. (13),

$$
(\vec{\phi}_{\mathrm{I}} - \vec{\phi}_c) + \vec{\psi} = -\frac{1}{\beta} \vec{\nabla} \log[P(\vec{h}_c)] = \frac{\vec{h}_c}{\beta h_0^2}, \tag{24}
$$

where

$$
\vec{\psi} = \lim_{k_2\to\infty} k_2 [\vec{\phi}_{\mathrm{II}} - \vec{\phi}_{\max}] \tag{25}
$$

is proportional to the soft mode of the Hessian (computed in $\vec{\phi}_c$), and $\vec{\phi}_{\mathrm{I}}$ is the primary minimum in the presence of the field $\vec{h}_c$.

C. Stabilizing marginal states

Marginal states are rather delicate to work with since they are very sensitive to perturbations. A way to circumvent this problem is to introduce a regularizer favoring (slightly) stable states, and let it tend to zero at the end of the calculation. Such an approach has been implemented in Refs. 21 and 37 for the SK model and the Viana-Bray model. Here, we show explicitly the mechanism of this procedure for the toy model studied above. A similar analysis for the SK model will be presented in Sec. VI (see also Appendix C).

Let us reconsider the one-dimensional model, Eq. (1). For slightly supercritical random fields, $h=h_c+\delta h$, there is a nearly marginal secondary minimum $\phi_{\mathrm{II}} \approx \phi_c + (2\delta h/V_3)^{1/2}$. In order to lift the marginality of the dominating states, we impose a stabilizing "constraint" by introducing a weight factor $\exp[\lambda_X X(\phi)]$ in the phase-space average of the free

energy. The function $X(\phi)$ is arbitrary up to the conditions $X(0)=0$ (in order not to couple to the ground state) and $X'(\phi_c)\neq 0$. Any other specifics of this regularizer will disappear in the end. The Griffith part of the quenched free energy Eq. (3) then reads

$$F_G = \int_{h_c}^{\infty} dh P(h) e^{-\beta[\mathcal{H}(\phi_{\mathrm{II}};h)-\mathcal{H}(\phi_{\mathrm{I}};h)]+\lambda_X X(\phi_2)}. \tag{26}$$

The integrand is maximal at the saddle point $h^* = h_c + \delta h$ with

$$\delta h^{1/2} \approx -\frac{\lambda_X X'(\phi_c)}{(2V_3)^{1/2}} \frac{P(h_c)}{P'(h_c)}, \tag{27}$$

corresponding to a state ϕ_{II} that is a genuine local minimum with a minimal eigenvalue of the energy Hessian that is positive,

$$\lambda_{\min} = V''(\phi(h^*)) = \sqrt{\frac{V_3 \delta h}{2}} \sim \lambda_X. \tag{28}$$

The linear response to a Hamiltonian perturbation $\mathcal{H} \to \mathcal{H} - h_X X(\phi)$ diverges as $\partial\phi/\partial h_X \sim \lambda_X^{-1}$ as the constraint is lifted. However, the product

$$z \equiv \lim_{\lambda_X \to 0} \lambda_X \frac{\partial \phi}{\partial h_X} = -\frac{P(h_c)}{P'(h_c)} \tag{29}$$

tends to a finite limit.

The above reasoning holds for any function X provided that $X'(\phi_c)\neq 0$, which is necessary to lift the marginality. In the higher-dimensional case, this generalizes to the requirement

$$\vec{\nabla}_{\phi} X(\vec{\phi}_c) \cdot \vec{\psi} \neq 0, \tag{30}$$

where $\vec{\psi}$ is the soft mode of the Hessian.

As in the one-dimensional case, there is a finite limit for the combination

$$\lim_{\lambda_X \to 0} \lambda_X \frac{\partial \vec{\phi}}{\partial h_X} = \{\vec{\nabla} \log[P(h_c)] \cdot \vec{\psi}\}\vec{\psi}. \tag{31}$$

The avoidance of marginality by means of a control parameter coupled to a generic weight function X will be used again in Sec. VI, where we will apply the same trick to the SK model.

D. Discussion and connection to the two-group ansatz for mean-field models

The vectorial replica analysis presented in this section demonstrates that sending the number of replicas of two separate groups to plus and minus infinity, respectively, encodes the marginality of the metastable states they describe. In particular, we observe that the sum over RSB saddles in Eq. (19) is dominated by terms with $k_1=1$ and $k_2 \to \infty$. The latter reflects the fact that the leading (Griffith-like) contribution to the free energy is due to field realizations in which the secondary minimum is just marginal.

Furthermore, in the d-dimensional case, we have seen that the difference vector $\vec{\psi}$ between the secondary minimum $\vec{\phi}_{\mathrm{II}}$ and the saddle $\vec{\phi}_{\max}$ corresponds to the direction of the soft mode in that marginal state. This is shown to emerge (in Sec. III B) assuming that the integral over the field $\vec{h}$ is dominated by a small vicinity of a single saddle point $\vec{h}_c$, where $P(\vec{h})$ takes its maximum on the marginal surface $\mathcal{S}$ defined by Eq. (23).

The contribution to the partition function is dominated by a single state only as long as the number of minima does not grow exponentially with increasing disorder strength. This assumption breaks down in high-dimensional mean-field models such as the SK model where the dimensionality grows with number of spins $d \sim N$. Nevertheless, in these cases the most numerous states in a given quenched realization of random couplings turn out to be marginal, even though the origin of marginality is different. For any disorder configuration, there are exponentially many states (solutions of the TAP equations) in magnetization space, but stable states are usually less abundant than marginal ones since true stability imposes extra constraints on the TAP solutions. This argument certainly holds if no constraint is imposed on the free-energy density. However, in many models (e.g., Ising p-spin models), there is a range of low free energies where typical states are stable, while at higher energies the majority of states are marginal. The SK model is special in that it does not possess such a low-energy regime, so that the dominant states at all free energies are marginal.

The above mechanism for marginality in mean-field glasses is to be contrasted with the toy model where marginality is a consequence of maximizing Griffith contributions over the rare disorder. Despite this difference, the insight from the toy model will help us in the next section to obtain a deeper understanding of the physical significance of the two-group ansatz for mean-field glasses. In later sections, we will confirm this interpretation by revisiting the counting of TAP states and the cavity approach.

IV. THE TWO GROUP ANSATZ FOR MEAN FIELD SPIN GLASSES

A. Marginal states in the SK model

Bray and Moore[31] discovered that their computation of the TAP complexity[18] could be exactly reproduced by substituting for each entry in a standard Parisi matrix a two-group matrix (with $m+K$ and $-K$ replicas, respectively), and Legendre transforming the result with respect to m, anticipating Monasson's approach.[26] Later, Parisi and Potters[15] showed that this equivalence extends to the low-energy regime of the SK model, where full replica symmetry breaking occurs, and also holds in a model of random orthogonal matrices.[38] However, the meaning of this remarkable result remained unclear.

The preceding analysis of our toy model suggests to interpret the two-group ansatz as a means to force representatives of the two groups into pairs of almost coalescing minima and saddles. This picture is strongly supported by the recent analytical[3] and numerical[4,39] finding that solutions of the TAP equations always appear in pairs, one being a local

minimum and the other a saddle of rank 1. The straight connection of such a pair of stationary points defines a path in the energy landscape that is increasingly flat with increasing system size. In the thermodynamic limit, the Hessian matrix computed at the minimum has a zero eigenvalue with a soft mode pointing in the direction of the adjacent saddle.

The only TAP state without a "partner state" is the paramagnetic TAP solution $\{m_i=0\}$, which has to be discarded because it is unphysical. The absence of a "trivial" ground state constitutes an (inessential) difference between the SK model and the toy model considered in the previous section: in the SK model, there is no physically relevant analog of the global minimum ϕ_1 of Sec. III. The marginal states do not merely occur as high-energy excitations above some trivial ground state, but they are *the* dominant metastable states at a given free energy. Therefore, a vectorial replica symmetry breaking with only two groups of replicas (associated with a minimum and a saddle) captures all the information we need.

B. Order parameters of the two-group ansatz

In the presence of pairs of minima and saddles, the concept of overlaps between different states (similarity of their magnetizations) needs to be generalized to cover three cases: overlaps between two minima, between two saddles, and between a minimum and a saddle. They can be described by matrices Q_{ab}^{++}, Q_{ab}^{--}, and Q_{ab}^{+-}, respectively. The replica indices a, b indicate the distance within an ultrametric Parisi tree of the respective pairs.

Such order parameters indeed appear in the two-group ansatz,[29] where one works with two groups of $m+K$ and $-K$ replicas. K plays now the same role as k_2 for the toy model of Sec. III. We thus expect that a set of m replicas corresponds to a pair of a minimum and a saddle whose magnetizations differ by a quantity of the order of $O(1/K)$,

$$m_i^{\pm} = \bar{m}_i \pm \frac{\delta m_i}{2K}, \tag{32}$$

where $\{\bar{m}_i\}$ is the set of average site magnetizations of the saddle-minimum pair and $\{\delta m_i\}$ is the ith component of the soft mode connecting the minimum to the nearby saddle in configuration space. Consequently, we expect the overlap matrices to be given by

$$Q_{ab}^{+-} \equiv Q_{ab} = \bar{Q}_{ab} - \frac{C_{ab}}{4K^2}, \tag{33}$$

$$\begin{aligned} Q_{ab}^{++} &= \bar{Q}_{ab} + \frac{A_{ab}}{K} + \frac{C_{ab}}{4K^2} \\ &= Q_{ab} + \frac{A_{ab}}{K} + \frac{C_{ab}}{2K^2}, \end{aligned} \tag{34}$$

$$Q_{ab}^{--} = \bar{Q}_{ab} - \frac{A_{ab}}{K} + \frac{C_{ab}}{4K^2} = Q_{ab} - \frac{A_{ab}}{K} + \frac{C_{ab}}{2K^2}, \tag{35}$$

with the following interpretation of the order parameters:

$$\bar{Q}_{ab} = \frac{1}{N}\sum_{i=1}^{N} \bar{m}_i^a \bar{m}_i^b, \tag{36}$$

$$A_{ab} = \frac{1}{N}\sum_{i=1}^{N} \bar{m}_i^a \delta m_i^b, \tag{37}$$

$$C_{ab} = \frac{1}{N}\sum_{i=1}^{N} \delta m_i^a \delta m_i^b, \tag{38}$$

which are assumed to be self-averaging and only dependent on the distance between the minimum-saddle pairs labeled by a and b. Note that in particular, $Q \equiv \bar{Q}_{aa}$, $A \equiv A_{aa}$, and $C \equiv C_{aa}$ describe the internal overlaps of a single marginal minimum-saddle pair.

Here we have introduced $Q_{ab} \equiv Q_{ab}^{+-}$ to match the notation of Ref. 29, but the difference between Q_{ab} and $\bar{Q}_{ab}$ of order $1/K$ is immaterial in the two-group limit where K is sent to infinity.

In a situation in which the dominant states are stable, the matrix $\bar{Q}$ is the only order parameter (formally there is no soft mode, $\delta\vec{m} \equiv 0$, and thus $A_{ab}=C_{ab}=0$). In the case of marginal states, the matrix A_{ab} measures the correlations between the magnetization of a state ($\{\bar{m}_i^a\}$) and the direction of the soft mode of another state ($\{\delta m_i^b\}$) at a phase space distance labeled by a and b (and vice versa). The matrix C_{ab} measures the similarity between the soft modes of such states. Note that the picture of merging minima and saddles gives a clear interpretation only of the *direction* of the soft mode, while it does not determine the magnitude of δm_i. In order to extract physical information, we should thus normalize the soft modes by $\delta\hat{m}_i = \delta m_i / \sqrt{\langle \delta m_i^2 \rangle} = \delta m_i / \sqrt{C_{aa}}$. In particular, we may infer that C_{ab}/C_{aa} describes the decreasing correlations between soft modes with increasing distance of pairs in phase space. Moreover, the angle γ_{ab} between the magnetization vector and the soft mode in states labeled by a, b is given by

$$\cos(\gamma_{ab}) = \frac{\langle \delta m_i^a m_i^b \rangle}{\sqrt{\langle \delta m_i^2 \rangle \langle m_i^2 \rangle}} = \frac{A_{ab}}{\sqrt{Q_{aa} C_{aa}}}. \tag{39}$$

If we assume that the slow relaxation dynamics follows essentially the soft mode of marginal states, we expect that the larger γ_{aa}, the smaller the relative decrease of the self-overlap in the course of the dynamics.

C. SK model: Replicated free energy

Let us now turn explicitly to the SK model with the Hamiltonian

$$\mathcal{H} = -\sum_{i<j} s_i J_{ij} s_j, \tag{40}$$

where J_{ij} is a random Gaussian matrix of zero mean and variance $1/N$, coupling all N Ising spins s_i together.

Using Monasson's clone method to access higher-lying metastable states, one computes the quenched free energy of m copies of the system,

$$-\beta F_m = \lim_{n\to 0} \frac{1}{n} \log \overline{Z_J^{mn}} = \operatorname*{ext}_f [\overline{\log \mathcal{N}_J(f)} - \beta m N f]. \quad (41)$$

Z_J^{mn} is the partition function of $n \times m$ copies of a system with Hamiltonian (40). n is the number of replicas introduced to compute the quenched average, while m is the number of real copies [the Legendre conjugate of the free-energy density, see Eq. (43) below]. In order to capture marginal states, we divide the m replicas further into two groups of $m+K$ and $-K$ elements each. As in the toy model, K is sent to infinity in the end, which forces the two groups (at fixed replica index $a \in \{1,\ldots,n\}$) into marginal minimum-saddle pairs.

In the thermodynamic limit $N\to\infty$, the left-hand side of Eq. (41) yields the two-group replicated free energy,

$$-\beta m \Phi_{2G}(m) \equiv \lim_{N\to\infty} \frac{-\beta F_m}{N} = \lim_{n\to 0} \lim_{N\to\infty} \frac{1}{nN} \log \overline{Z_J^{mn}}, \quad (42)$$

while the right-hand side evaluates to

$$\frac{1}{N} \overline{\log \mathcal{N}_J(f)} - \beta m f \equiv \Sigma(f) - \beta m f, \quad (43)$$

where as usual we assume an exponential growth of the number of metastable states with system size, $\mathcal{N}_J(f) \sim \exp[N\Sigma(f)]$, the latter defining the complexity function $\Sigma(f)$. From Eq. (43), the function $\beta m \Phi_{2G}(m)$ is thus seen to be the Legendre transform of the complexity with respect to the pair of conjugated variables f and βm.

After a standard Hubbard-Stratonovich decoupling, one obtains the averaged partition function in terms of a functional integral over an $nm \times nm$ replica coupling matrix $\mathcal{Q}$,

$$\overline{Z_J^{mn}} = \int \mathcal{D}\mathcal{Q} \exp\{nN\mathcal{F}[\mathcal{Q}]\}, \quad (44)$$

$$\mathcal{F}[\mathcal{Q}] = \frac{1}{n} \log \sum_{\{s_a^i\}} \exp\left(\frac{\beta^2}{2} \sum_{ab} \sum_{ij} s_a^i \mathcal{Q}_{ab}^{ij} s_b^j\right) + \frac{\beta^2}{4}\left(m - \frac{1}{n}\mathrm{Tr}\mathcal{Q}^2\right), \quad (45)$$

where indices run through $a,b=1,\ldots,n$ and $i,j=1,\ldots,m$, respectively.

As motivated above, the two-group ansatz consists in writing the matrix $\mathcal{Q}_{ab}^{ij}$ as n^2 submatrices $\mathbf{Q}_{ab}$, each of dimension $m \times m$ of the form

$$\mathbf{Q}_{ab} = \begin{pmatrix} \overbrace{Q_{ab}^{++}}^{m+K} & \overbrace{Q_{ab}^{+-}}^{-K} \\ Q_{ab}^{+-} & Q_{ab}^{--} \end{pmatrix}, \quad (46)$$

and we adopt the parametrization (33) and (34) for the three sectors. The diagonal elements are $\mathcal{Q}_{aa}^{ii} \equiv 0$.

The partition function (44) can be evaluated through a saddle-point computation which leads to the thermodynamic potential

$$-\beta m \Phi_{2G} \equiv \lim_{n\to 0} \operatorname*{ext}_{\mathcal{Q}} \mathcal{F}[\mathcal{Q}] = \lim_{n\to 0}\Bigg[\beta\phi_m + \frac{\beta^2}{4} m(1-Q)^2 - \beta^2 A(1-Q) - \frac{\beta^2}{n}\sum_{ab}\left(\frac{1}{2}(A_{ab}^2 + Q_{ab}C_{ab}) + mQ_{ab}A_{ab} + \frac{m^2}{4}Q_{ab}^2\right)\Bigg] \quad (47)$$

with

$$\beta\phi_m \equiv \frac{1}{n} \log \sum_{\{s_a^i\}} \exp\left(\frac{\beta^2}{2}\sum_{ab}\sum_{ij} s_a^i \mathcal{Q}_{ab}^{ij} s_b^j\right). \quad (48)$$

As for the standard Parisi ("one-group") ansatz, the self-consistency conditions for the large N saddle point read

$$Q_{ab}^{\sigma\sigma'} = \langle s_a^{i\sigma} s_b^{i\sigma'} \rangle, \quad (49)$$

where $\sigma,\sigma' \in \{+,-\}$, and the average is taken over the Boltzmann factor in Eq. (48). Here an index i_+ is restricted to $[1,m+K]$, while $i_- \in [m+K+1,m]$. Solving for Q_{ab}, A_{ab}, and C_{ab} [cf. Eqs. (33) and (34)], we may also write

$$Q_{ab} = \frac{1}{4}\langle (s_a^{i_+} + s_a^{i_-})(s_b^{i_+} + s_b^{i_-}) \rangle, \quad (50)$$

$$A_{ab} = \frac{K}{2}\langle s_a^{i_+}(s_b^{i_+} - s_b^{i_-}) + (s_a^{i_+} - s_a^{i_-}) s_b^{i_-} \rangle, \quad (51)$$

$$C_{ab} = K^2 \langle (s_a^{i_+} - s_a^{i_-})(s_b^{i_+} - s_b^{i_-}) \rangle, \quad (52)$$

which resembles Eqs. (36)–(38). The detailed connection between the two sets of equations is established in Appendix A.

The log-trace term ϕ_m in Eq. (48) can be reexpressed in two equivalent ways that will be helpful to make the connection with the generalized cavity approach and the counting of TAP states, respectively. Technical details of the derivation can be found in Appendix A.

Following Parisi and Potters,[15] one obtains the expression

$$e^{n\beta\phi_m} = 2^{nm} \int_{-i\infty}^{i\infty} \prod_a \frac{dx_a}{2\pi i} \int_{-1}^{1} \prod_a \frac{dm_a}{(1-m_a^2)} \times \exp\left\{ -\sum_a \left[x_a \tanh^{-1} m_a + \frac{m}{2} \log(1-m_a^2) \right] + \beta^2 \sum_{ab} \frac{1}{2} x_a Q_{ab} x_b + \frac{1}{2} m_a C_{ab} m_b + m_a A_{ab} x_b \right\}, \quad (53)$$

where the sum over (a,b) also includes diagonal terms, $Q_{aa} \equiv Q$, $A_{aa} \equiv A$, $C_{aa} \equiv C$. We will see in Sec. V that this form also emerges from a direct counting of TAP solutions.

In this representation, the self-consistency equations (50) can be cast into the form

$$Q_{ab} = \langle m_a m_b \rangle, \quad (54)$$

$$A_{ab} + \delta_{ab}(1 - Q_{aa}) = \langle m_a(x_b - mm_b)\rangle, \tag{55}$$

$$C_{ab} - \delta_{ab}[2A_{aa} + m(1 - Q_{aa})] = \langle (x_a - mm_a)(x_b - mm_b)\rangle, \tag{56}$$

where the average $\langle\cdots\rangle$ is taken over the measure in Eq. (53).

In Appendix A we show that the field x_a is in a certain sense a twofold Hubbard Stratonovich transformation of the spin variables $S_a = \Sigma_i s_a^i$. Its average magnetization is therefore mm_a, and the terms $x_a - mm_a$ can be thought of as magnetization fluctuations in replica space. Indeed this furnishes an intuitive understanding of the off-diagonal part ($a \neq b$) of Eqs. (54)–(56) and supports our interpretation of the overlap matrices. The extra diagonal terms ($a=b$) on the left-hand side of Eqs. (55) and (56) arise due to the fragility of the marginal states, as we will see in Sec. VI from an alternative derivation.

In Appendix A, we derive the equivalent expression

$$e^{n\beta\phi_m} = \frac{1}{\sqrt{\det \mathcal{M}}} \int \prod_{a=1}^{n} \frac{dh_a dz_a}{2\pi} \times \exp\left\{ \sum_a [m \log 2\cosh(\beta h_a) + \tanh(\beta h_a)\beta z_a] - \frac{1}{2}\sum_{ab} \underline{\xi}_a^\dagger \mathcal{M}_{ab}^{-1} \underline{\xi}_a \right\}, \tag{57}$$

where $\underline{\xi}_a^\dagger \equiv (h_a, z_a)$ and $\mathcal{M}$ is the $2n \times 2n$ covariance matrix

$$\mathbf{M} = \begin{pmatrix} \overbrace{\mathbf{Q}}^{n} & \overbrace{\mathbf{A}}^{n} \\ \mathbf{A} & \mathbf{C} \end{pmatrix}. \tag{58}$$

Moreover, introducing magnetization and soft mode variables [cf. Eqs. (A17) and (A18)],

$$\tilde{m}_a = \tanh(\beta h_a), \tag{59}$$

$$\delta m_a = \frac{\beta z_a}{\cosh^2(\beta h_a)}, \tag{60}$$

the self-consistency equations (50)–(52) take the simple form

$$Q_{ab} = \langle \tilde{m}_a \tilde{m}_b \rangle, \tag{61}$$

$$A_{ab} = \langle \tilde{m}_a \delta m_b \rangle, \tag{62}$$

$$C_{ab} = \langle \delta m_a \delta m_b \rangle, \tag{63}$$

where averages are over the measure defined by the integrand in Eq. (57). We will use this form in the discussion of the quenched computation in Sec. VII.

D. Annealed approximation

In the following, we will focus on the annealed approximation, which corresponds to averaging the two-group partition function $\overline{Z_J^m}$ instead of its logarithm. This is known to be an exact description at high enough free-energy densities, $f > f^\star$,[40] cf. Eq. (69) below. Technically, this approximation corresponds to reducing the overlap matrix to its diagonal part ($a=b$) described by $\{Q, A, C\}$, and setting all off-diagonal elements to zero.

The formalism for a quenched computation with continuous RSB is reviewed and physically interpreted in Sec. VII.

Integrating out z_a in the annealed version of Eq. (57), one finds

$$e^{\beta\phi_m} = \int_{-\infty}^{\infty} \frac{dh}{\sqrt{2\pi Q}} e^{-h^2/2Q} (2\cosh\beta h)^m \times \exp\left[\frac{A}{Q}\beta h \tanh\beta h + \frac{\beta^2}{2}\frac{QC - A^2}{Q}\tanh^2\beta h \right]. \tag{64}$$

The same result is obtained by integrating out x_a in the annealed version of Eq. (53) and using the relation $m_a = \tanh\beta h_a$ between magnetization m_a and local field h_a. The integrand on the right-hand side has the interpretation of a probability distribution of local fields (up to a normalization),

$$P_{\text{ann}}(h) = \frac{1}{\mathcal{N}} \exp\left\{ m\log(2\cosh\beta h) - \frac{h^2}{2Q} + \frac{A}{Q}\beta h \tanh\beta h + \frac{\beta^2}{2}\frac{QC - A^2}{Q}\tanh^2\beta h \right\}. \tag{65}$$

The joint distribution of magnetizations $\tilde{m}$ and soft mode components δm can be similarly obtained from normalizing the replica-diagonal version of the measure in Eq. (57) and changing variables according to Eqs. (59) and (60).

In the annealed approximation, Eqs. (54)–(56) take the form (performing the Gaussian averages over x)

$$Q = \langle \tanh^2(\beta h)\rangle, \tag{66}$$

$$A + 1 - Q = -A - mQ + \frac{\langle h\tanh(\beta h)\rangle}{\beta Q}, \tag{67}$$

$$C - 2A - m(1 - Q) = m^2 Q + 2mA + \frac{A^2}{Q} - 2\left(m + \frac{A}{Q}\right) \times \frac{\langle h\tanh(\beta h)\rangle}{\beta Q} - \frac{1}{\beta^2 Q}\left(1 - \frac{\langle h^2\rangle}{Q}\right), \tag{68}$$

the average $\langle\cdots\rangle$ now denoting an integral over $P_{\text{ann}}(h)$. Equations (65)–(68) allow one to find the annealed solution $\{Q, A, C\}$ easily, e.g., by iteration.

The above equations turn out to admit two solutions,[27] only one of which (with $A, C \neq 0$) is physical, as discussed further in Secs. V and VIII.

Bray and Moore showed in Ref. 40 that this annealed solution is stable with respect to continuous replica symmetry breaking (onset of correlations between typical metastable states) as long as $f > f^\star = f(m^\star)$, where $m^\star$ satisfies the condition

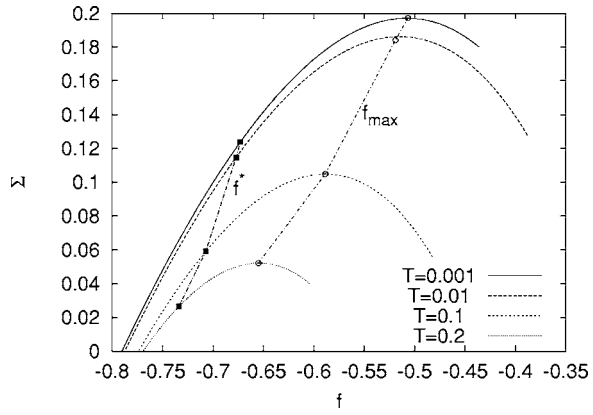

FIG. 2. Complexity curves in the annealed approximation. Only the data for $f>f^\star$ are exact. $f_{\max}$ corresponds to the states of maximal complexity at a given temperature. The high-energy part of the curves is not fully shown [they extend up to $f_{\rm hbe}$, where $\Sigma(f_{\rm hbe}=0)$].

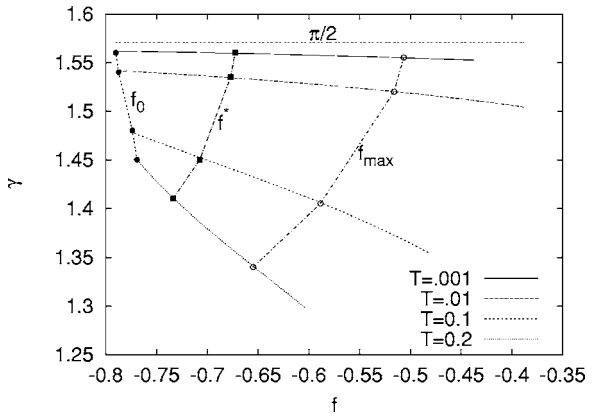

FIG. 3. Angle between soft mode and magnetization of typical marginal states. The parameter f_0 is the value of the free-energy density at which the complexity of marginal states computed in the annealed approximation goes to zero (f_0 is not equal to the true equilibrium value $f_{\rm eq}$ since the approximation breaks down below $f^\star$). $f_{\max}$ denotes the free-energy density with maximal complexity (see Fig. 2).

$$1=\beta(1-Q)+\beta\left(A+\frac{m^\star Q}{2}\right)+\beta\sqrt{Q\left(C+m^\star A+\frac{m^{\star 2}}{4}Q-2A\right)}. \tag{69}$$

E. Complexity and direction of the marginal mode

Evaluating the replicated free energy Φ_{2G} [Eq. (47)] within the annealed approximation, one eventually obtains the complexity as its Legendre transform,

$$\Sigma(f)=-\beta m\Phi_{2\rm G}(m)+m\beta f. \tag{70}$$

In the annealed approximation, the above observables read

$$f(m)=\frac{\partial m\Phi_{2\rm G}(m)}{\partial m}=-\frac{\langle\log 2\cosh\beta h\rangle}{\beta}-\frac{\beta}{4}[(1-Q)^2-2mQ^2-4AQ], \tag{71}$$

$$\Sigma(m)=\beta m^2\frac{\partial\Phi_{2\rm G}}{\partial m}=\beta\phi_m-m\langle\log 2\cosh\beta h\rangle+\frac{\beta^2}{4}[m^2Q^2-4A(1-Q)-2(A^2+QC)], \tag{72}$$

which are used to compute the complexity curves of Fig. 2.

In Fig. 3, we plot the angle γ between the soft mode and the magnetization vector in configuration space, cf. Eq. (39), versus free-energy density for various temperatures. Only the part of the curves to the right of $f^\star(T)$ is exact, whereas the low-energy part is an annealed approximation to the regime where states are correlated among each other. Nevertheless, there is a clear tendency to increasing toward orthogonality, $\gamma\to\pi/2$, both as f is approaching the lower band edge of the f interval, and as $T\to 0$. We will derive this behavior explicitly in the low-T limit in Sec. IX.

V. COUNTING TAP STATES: BREAKING OF BRST SYMMETRY AND EQUIVALENCE TO THE TWO-GROUP ANSATZ

In this section, we briefly review the counting of TAP states, defined as stationary points (minima) of the TAP free energy,

$$F_{\rm TAP}(\{m_i\})=-\frac{1}{2}\sum_{ij}J_{ij}m_im_j+\frac{1}{\beta}\sum_i f_0(q,m_i), \tag{73}$$

with

$$f_0(q,m_i)=\frac{1}{2}\ln(1-m_i^2)+m_i\tanh^{-1}m_i-\ln 2-\frac{\beta^2}{4}(1-q)^2. \tag{74}$$

In order to select states at a given free-energy density, we weight the TAP states (labeled by α) with an exponential factor,

$$\mathcal{Z}_J(\beta;m)\equiv\sum_\alpha e^{-m\beta F_{\rm TAP}(\{m_\alpha\})}. \tag{75}$$

We can rewrite this sum in the integral representation[41]

$$\mathcal{Z}_J(\beta;m)=\int\mathcal{D}m\,\mathcal{D}x\,\mathcal{D}\bar\psi\,\mathcal{D}\psi\times\exp\left\{\beta\left[-\sum_i x_i\partial_{m_i}F_{\rm TAP}(\{m_i\})+\sum_{ij}\bar\psi_i\psi_j\partial_{m_i}\partial_{m_j}F_{\rm TAP}(\{m_i\})-mF_{\rm TAP}(\{m_i\})\right]\right\}. \tag{76}$$

One then proceeds by replicating the above expression, av-

eraging over the Gaussian bond disorder, and decoupling quartic terms in $x, m, \psi, \bar{\psi}$ by Lagrange multipliers. Integrating out the fermionic fields and imposing saddle-point conditions, one obtains the expression[42]

$$\overline{\mathcal{Z}^n(\beta;m)} = \prod_a \int_{-1}^{1} \frac{dm_a}{1-m_a^2} \prod_a \int_{-i\infty}^{i\infty} \frac{dx_a}{2\pi i} \times \exp\left\{ N\left[-\sum_a x_a \tanh^{-1}(m_a) - m\sum_a f_0(Q_{aa},m_a) - \frac{\beta^2}{2}\sum_{ab} G_{ab} - \frac{\beta^2}{2}\sum_a 2A_{aa}(1-Q_{aa}) + \frac{\beta^2}{2}\sum_{ab}\left[x_a Q_{ab} x_b + x_a(2A_{ab}+mQ_{ab})m_b + m_a(C_{ab}+2mA_{ab}+m^2Q_{ab})m_b\right]\right]\right\} \tag{77}$$

with

$$G_{ab} \equiv 2(A_{ab}^2 + Q_{ab}C_{ab}) + 4mA_{ab}Q_{ab} + m^2Q_{ab}^2. \tag{78}$$

This should be optimized with respect to Q_{ab}, A_{ab}, C_{ab}, leading to the saddle-point equations

$$Q_{ab} = \langle m_a m_b \rangle, \tag{79}$$

$$A_{ab} + \delta_{ab}(1-Q_{aa}) = \langle x_a m_b \rangle, \tag{80}$$

$$C_{ab} - \delta_{ab}[2A_{aa} + m(1-Q_{aa})] = \langle x_a x_b \rangle. \tag{81}$$

As in the previous section, these equations admit two solutions, i.e., there are two possible saddle points for the integrand in Eq. (76): (i) one with $A=C=0$, conserving the BRST symmetry of the action (76), and (ii) a saddle point with $A \neq 0$, $C \neq 0$ which spontaneously breaks this symmetry. We will discuss in Sec. VIII that only the second solution is physical.

A. Equivalence with the two-group ansatz

Taking the logarithm of Eq. (77), we obtain

$$F(\beta;m) \equiv \overline{\log \mathcal{Z}_J} = \lim_{n\to 0} \frac{1}{n} \log \overline{\mathcal{Z}_J^n}(\beta;m), \tag{82}$$

where m is the factor in the weight of Eq. (75). It takes the same role as the total number of replicas in the two groups in the previous section. Again, a Legendre transform with respect to m yields the complexity function.

Comparing the thermodynamic potentials of Eq. (82) and Eq. (42) [evaluated according to Eq. (53)], we establish their complete equivalence after shifting the integration variable to $\tilde{x}_a = x_a - m m_a$ in Eq. (53). As explained in Sec. IV, this shift amounts to using natural variables for magnetization fluctuations (see also Appendix A).

From the present derivation, we obtain a better understanding of the auxiliary fields x_a in the saddle-point equations (79)–(81), which are obviously equivalent to Eqs. (54)–(56) obtained from the two-group ansatz. As one may expect from the functional integral (76), the variables x_a are prone to fluctuate most strongly along the softest mode of the TAP solutions, and this leads to a rather natural interpretation of the saddle-point equations (79)–(81) for $a \neq b$: A_{ab} and C_{ab} describe correlations between the soft mode directions and the local magnetizations, as also suggested by the two-group ansatz.

For $a=b$, extra terms arise. This is because x_a does not exactly describe soft mode directions in the same sense as m_a describes magnetizations. While this is immaterial for interstate correlations, it introduces correction terms when intrastate correlations are considered. In the next subsection, this will become clearer from a direct derivation of these contributions from generalized Ward identities.

B. BRST symmetry breaking and the generalization of Ward identities

The action in the exponent of the integrand in Eq. (76) is invariant under the fermionic BRST symmetry,[22]

$$\delta m_i = \epsilon \psi_i, \quad \delta x_i = -\epsilon \frac{m}{2} \psi_i, \quad \delta \bar{\psi}_i = -\epsilon x_i, \quad \delta \psi_i = 0. \tag{83}$$

If the dominant states are stable minima (cf. Sec. II A), the BRST symmetry of the action (76) is not broken by the saddle point Eq. (77), and Ward identities

$$\langle m_i x_i \rangle = \langle \psi_i \bar{\psi}_i \rangle, \tag{84}$$

$$\langle x_i x_i \rangle = -m \langle \psi_i \bar{\psi}_i \rangle, \tag{85}$$

hold. On the macroscopic level (upon average over sites i), these impose that the order parameters $A_{ab}=C_{ab}=0$. This can be appreciated by noting that Eqs. (84) and (85) reproduce the diagonal saddle point equations (80) and (81), considering that $N^{-1}\Sigma_i \langle \psi_i \bar{\psi}_i \rangle = 1-Q$.

If instead the majority of states are marginally stable states (i.e., minima merging with saddles of rank 1), the BRST symmetry is broken, translating into $A_{ab} \neq 0$, $C_{ab} \neq 0$. Equations (79)–(81) result from a (macroscopic) generalization of the above Ward identities. One can derive them from direct inspection of the sum over TAP states with a regularization. As we did for the toy model of Sec. III [Eq. (26)], we add a regularizing term $\exp[\lambda_X X(\{m_j\})]$ to the weight in Eq. (75). As explained in Appendix C, this procedure selects minimally stable states with a susceptibility diverging as $1/\lambda_X$ as $\lambda_X \to 0$,

$$\mathcal{Z}_J(\beta;m,\lambda_X) = \sum_{\alpha=1}^{\mathcal{N}_{\rm sol}} e^{-m\beta F_{\rm TAP}(\{m_\alpha\}) + \lambda_X X(\{m_\alpha\})}. \tag{86}$$

The quantity X is an arbitrary, extensive, symmetric function of the average site magnetizations subject only to the constraint that its gradient be nonorthogonal to the soft mode of typical marginal states,

$$\vec{\nabla}_m X(\vec{m}) \cdot \delta m \neq 0. \tag{87}$$

This regularization explicitly lifts the marginality throughout the computation. Sending the control parameter λ_X eventually to zero allows us to recover the marginal states.

When the TAP equations are perturbed by external fields, there is no guarantee that the number of solutions at a given value of the free-energy density is conserved. Indeed, only the set of TAP solutions counted by the BRST symmetric saddle point is robust to perturbations:[27] the number of solutions at any free-energy density is conserved and there is a one-to-one correspondence between states before and after the perturbation. In the case of broken BRST symmetry, the correspondence is lost and solutions can appear and disappear at any free-energy density. This fragility is related to the marginality of the states. The regularization Eq. (86) fixes this problem by lifting the marginality. As a consequence, the number of selected states is conserved under perturbations (for fixed parameters f *and* λ_X), but at the same time the BRST symmetry is explicitly broken.

We will now derive the generalizations of the Ward identities (84) and (85). Moreover, by showing their equivalence with the self-consistency Eqs. (79)–(81), we shall confirm the interpretation of the order parameters of the two-group ansatz put forward in Sec. IV B.

We consider sums over TAP states that are dominated by shallow minima due to regularization. Further, we regard the states as functions of small perturbing fields $\{h_k\}$, in the sense that $m_\alpha(\{h_k\})$ denotes the unique solution of $\partial_{m_i} F_{\mathrm{TAP}}(\{m_j\}) = h_i$ in the vicinity of the unperturbed state $m_\alpha(\{h_k=0\})$.

Since the free-energy Hessian possesses a soft mode along the direction $\delta m_i \propto \zeta_i$ (normalized by $N^{-1}\Sigma_i \zeta_i^2 = 1$), the latter dominates the response to a perturbing external field. The regularization selects states whose susceptibility along the soft mode is finite, $\chi_{\mathrm{soft}} = \Sigma_{ij} \zeta_i \chi_{ij} \zeta_j = (g\lambda_X)^{-1}$, where g is a self-averaging constant, as shown in Appendix C.

The coupling of perturbing external fields $\{h_i\}$ to the soft mode is given by their projection on the soft mode, $\Sigma_i \zeta_i h_i$. In particular, restricting to the response along the soft mode, we find

$$\frac{\partial m_i}{\partial h_k} \approx \frac{\zeta_i \zeta_k}{g \lambda_X} \tag{88}$$

and thus, in the limit of a vanishing regularizer, one finds a finite limit for the derivative,

$$\frac{1}{\beta} \lim_{\lambda_X \to 0} \frac{\partial \lambda_X X}{\partial h_k} = \frac{1}{\beta g} \left(\sum_i \frac{\partial X}{\partial m_i} \zeta_i \right) \zeta_k \equiv \delta m_k. \tag{89}$$

Note that the so-defined vector $\{\delta m_k\}$ is proportional to the soft mode $\{\zeta_k\}$ and independent of X (since the constant g is proportional to X, and only the gradient of X in the direction of $\vec{\zeta}$ matters). This same vector will appear again in the cavity method (see Sec. VI). It gives a precise meaning to the amplitude of the soft mode appearing in the heuristic derivation of the two-group ansatz, cf. Eq. (32).

With the above observation, we are able to derive the following generalized Ward identities (see Appendix B for details):

$$\langle m_i^a x_i^b \rangle = \delta_{ab} \langle \psi_i^a \bar{\psi}_i^b \rangle + \langle m_i^a \delta m_i^b \rangle, \tag{90}$$

$$\langle x_i^a x_i^b \rangle = -m \delta_{ab} \langle \psi_i^a \bar{\psi}_i^b \rangle + \langle \delta m_i^a \delta m_i^b \rangle - 2 \delta_{ab} \langle m_i^a \delta m_i^b \rangle. \tag{91}$$

These equations are indeed equivalent to Eqs. (80) and (81) provided that we make the identification $A_{ab} = \langle m_a \delta m_b \rangle$, $C_{ab} = \langle \delta m_a \delta m_b \rangle$, as suggested by our interpretation of A and C as describing correlations among soft-mode directions and magnetizations.

In the absence of soft modes, the limit $\lambda_X \to 0$ cancels the terms containing $\delta\vec{m}$, and the BRST-Ward identities [Eqs. (84) and (85)] are recovered automatically.

VI. GENERALIZED CAVITY APPROACH AND THE TWO-GROUP ANSATZ

A. Cavity with marginal states revisited

In the standard cavity approach, one writes recursion relations for the local (cavity) field h_0 acting on an added site 0 in terms of the local fields h_i acting on its k neighboring sites when site 0 is absent. Here, $k+1$ is the connectivity of the lattice. In Ising systems, the cavity field is a sum of *messages* u_i,[43]

$$h_0 = \sum_{i=1}^{k} u_i(h_i; J_{0i}), \tag{92}$$

$$u_i(h_i; J_{0i}) = \beta^{-1} \tanh^{-1}[\tanh(\beta h_i)\tanh(\beta J_{0i})], \tag{93}$$

where J_{0i} are the quenched bonds coupling the cavity spin 0 to its neighbors. The free energy gain for the addition of a spin at site 0 is

$$\exp(-\beta \Delta F) = 2 \cosh(\beta h_0) \prod_{i=1}^{k} \frac{\cosh(\beta J_{0i})}{\cosh(\beta u_i)}. \tag{94}$$

Around the considered free energy density f, the density of metastable states is assumed to grow exponentially as $\rho(F) \propto \exp[m(f)(F - F_0)]$, where $m(f) = \partial \Sigma(f)/\partial f$ is the local slope of the complexity function (in other words, m is the Legendre conjugate of f), $F = fN$, and F_0 is an arbitrary reference value (to be absorbed into the normalization constant). In order to determine the average shift $\Phi_{\mathrm{cav}}(m)$ of this distribution when a site is added, one averages the "reweighting" factor $\exp(-\beta m \Delta F)$ over cavity iterations to obtain[43] $\exp[-\beta m \Phi_{\mathrm{cav}}(m)] \equiv \langle \exp[-\beta m \Delta F] \rangle_{\{h_i, J_{0i}\}}$.

In the case of the SK model, the connectivity goes to infinity $N \equiv k+1 \to \infty$ and the single bond strength tends to zero as $J_{0i}^2 = 1/N$. The above relations thus simplify. In particular, we have $u_i \approx J_{0i} m_i$ with $m_i \equiv \tanh(\beta h_i)$. The free-energy shift due to a spin addition is

$$\exp(-\beta\Delta F) = 2\cosh(\beta h_0)\exp\left[\frac{\beta^2}{2}\sum_{i=1}^{N} J_{0i}^2(1-m_i^2)\right]$$
$$= 2\cosh(\beta h_0)\exp\left[\frac{\beta^2}{2}\langle 1-m_i^2\rangle\right], \quad (95)$$

where $\langle\ \rangle$ denotes a site average, and we have used the fact that in the large-N limit, the second term does not fluctuate.

If a marginal mode is present, this standard method fails since the addition of a spin has an anomalously strong impact on the system, rendering previous states unstable. In order to circumvent this difficulty, we regularize the problem once again by reweighting the states with a factor $\exp[\lambda_X X(\{m_i\})]$, in addition to the standard reweighting $\exp[-\beta m F(\{m_i\})]$ with respect to free energy, which selects a certain free-energy density. Eventually, we will take $\lambda_X \to 0$ to recover the marginal states. As in earlier sections, the extensive observable X is an arbitrary symmetric function of the magnetizations, subject to the requirement (87). This method was originally introduced by Rizzo.[21] Here we go beyond his analysis by including the selection of a given free-energy density, and establishing the precise connection with the two-group ansatz. Further, we provide a clear interpretation of the auxiliary cavity fields that need to be introduced.

We expect from the toy models of Sec. III that the regularization scheme will select minimally stable states with the lowest eigenvalue of their free-energy Hessian being proportional to λ_X (for a derivation in the present case, see Appendix C). The change of X upon addition of a site is most conveniently computed as a derivative of the free-energy change with respect to the field h_X conjugated to the observable X,

$$\Delta X = -\frac{d\Delta F}{dh_X}. \quad (96)$$

The regularization is implemented in the cavity approach by reweighting each cavity iteration by $\exp[\lambda_X \Delta X]$. Note that the exponent remains finite in the limit $\lambda_X \to 0$, because the susceptibility of the soft mode diverges as $1/\lambda_X$. More precisely, we have

$$\lim_{\lambda_X\to 0} \lambda_X \Delta X = -\sum_{i=0}^{N} \frac{d\Delta F}{dh_i}\lambda_X \frac{dh_i}{dh_X} = -\sum_{i=0}^{N} \frac{d\Delta F}{dh_i}\beta z_i, \quad (97)$$

where

$$z_i \equiv \beta^{-1} \lim_{\lambda_X\to 0} \lambda_X \frac{dh_i}{dh_X} \quad (98)$$

are finite fields proportional to the component Δh_i of the local field fluctuation arising from a soft mode excitation. The fields z_i are in fact independent of the choice of X, as shown in Appendix C.

Deriving the recursion relation (92) with respect to h_X, we obtain a relation for z_0,

$$z_0 = \frac{dh_0}{dh_X}$$
$$= \sum_{i=1}^{N} J_{0i}\beta(1-m_i^2)z_i \equiv \sum_{i=1}^{N} J_{0i}\delta m_i, \quad (99)$$

where we introduced the soft mode in the magnetizations,

$$\delta m_i \equiv \frac{dm_i}{dh_i} z_i = \beta(1-m_i^2)z_i. \quad (100)$$

Note the correspondence with the definition (89), which is most transparent if we notice that

$$\delta m_i = \beta^{-1} \lim_{\lambda_X\to 0} \lambda_X \frac{dm_i}{dh_X}, \quad (101)$$

using the definition (98).

For the SK model, the shift of the regularizer, Eq. (97), is readily evaluated using Eq. (95),

$$\lim_{\lambda_X\to 0} \lambda_X \Delta X = \tanh(\beta h_0)\beta z_0 - \sum_{i=1}^{N} \tanh(\beta u_i)\frac{du_i}{dh_i}\beta z_i$$
$$= \beta m_0 z_0 - \beta^2 \sum_{i=1}^{N} J_{0i}^2 m_i \beta(1-m_i^2)z_i$$
$$= \beta m_0 z_0 - \beta^2 \langle m_i \delta m_i\rangle, \quad (102)$$

where $m_0 = \tanh \beta h_0$. Similarly as in Eq. (95), we may neglect the fluctuations of the second term in the large-N limit.

B. Free energy and self-consistency equations

Putting elements together, we obtain the regularized free-energy shift $\exp[-\beta m \Phi_{\rm cav}]$ upon a site addition by averaging the two reweighting factors $\exp[-m\beta\Delta F]\exp[\lambda_X \Delta X]$ over all possible random configurations $\{h_i, z_i\}$ of the neighboring cavity fields, as well as over the random couplings J_{0i},

$$\exp[-\beta m \Phi_{\rm cav}] = \langle \exp(-\beta m \Delta F)\exp(\lambda_X \Delta X)\rangle_{\{h_i, z_i, J_{0i}\}}. \quad (103)$$

The first term has the standard interpretation of a shift of the exponential distribution of states. The second term equals 1 if the considered states are stable. In the marginal case, it captures information about the probability for site additions to render existing states unstable or to make new marginal states emerge.

We may use the central limit theorem to infer from Eqs. (92) and (99) that $h_0(\{h_i, z_i, J_{0i}\})$ and $z_0(\{h_i, z_i, J_{0i}\})$ are Gaussian variables with covariance matrix

$$\mathcal{M} = \begin{pmatrix} Q^\star & A^\star \\ A^\star & C^\star \end{pmatrix}, \quad (104)$$

where

$$Q^\star = \overline{\left(\sum_{i=1}^{N} J_{0i} m_i\right)^2} = \langle m_i^2\rangle, \quad (105)$$

$$A^\star = \overline{\sum_{i=1}^{N} J_{0i} m_i \sum_{i=1}^{N} J_{0i} \delta m_i} = \langle m_i \delta m_i \rangle = \langle m_i \beta (1 - m_i^2) z_i \rangle, \tag{106}$$

$$C^\star = \overline{\left(\sum_{i=1}^{N} J_{0i} \delta m_i \right)^2} = \langle (\delta m_i)^2 \rangle = \langle [(1 - m_i^2) \beta z_i]^2 \rangle, \tag{107}$$

and $\langle\ \rangle$ denote site averages. We can then reexpress Eq. (103) as

$$\exp[-\beta m \Phi_{\text{cav}}] \tag{108}$$

$$= \exp\left[m \frac{\beta^2}{2} (1 - Q^\star) - \beta^2 A^\star \right] \times \int \frac{dh_0 dz_0}{2\pi [\det \mathcal{M}]^{1/2}} \exp\left[-\frac{1}{2} \underline{\xi}_0^\dagger \mathcal{M}^{-1} \underline{\xi}_0 \right] \times [2 \cosh(\beta h_0)]^m \exp[\tanh(\beta h_0) \beta z_0] \tag{109}$$

with $\underline{\xi}_0^\dagger \equiv (h_0, z_0)$. The reweighting terms can alternatively be seen as describing the relative probability of a cavity configuration $\{h_i, z_i, J_{0i}\}$ to occur, given a fixed free energy after the site addition. With this interpretation in mind, the probability distribution to find local fields h_0 and soft-mode components z_0 on site can be read off from Eq. (108) as

$$P(h_0, z_0) = \mathcal{N}^{-1} \exp\left[-\frac{1}{2} \underline{\xi}_0^\dagger \mathcal{M}^{-1} \underline{\xi}_0 \right] \times [2 \cosh(\beta h_0)]^m \exp[\tanh(\beta h_0) \beta z_0], \tag{110}$$

where $\mathcal{N}$ is a normalization constant. The self-consistency of the cavity approach requires that the average correlations on site 0 are the same as on the neighboring sites, i.e.,

$$Q^\star = \langle m_i^2 \rangle = \int dh_0 dz_0 P(h_0, z_0) \tanh(\beta h_0)^2, \tag{111}$$

$$A^\star = \langle m_i \delta m_i \rangle = \int dh_0 dz_0 P(h_0, z_0) \frac{\tanh(\beta h_0) \beta z_0}{\cosh^2(\beta h_0)}, \tag{112}$$

$$C^\star = \langle \delta m_i^2 \rangle = \int dh_0 dz_0 P(h_0, z_0) \left(\frac{\beta z_0}{\cosh^2(\beta h_0)} \right)^2. \tag{113}$$

C. Equivalence between the generalized cavity method and the two-group ansatz

The cavity approach with a single reweighting $\exp[-\beta m \Delta F]$ corresponds to an annealed calculation, neglecting correlations and clustering among different states. Let us thus establish its connection with the annealed approximation in the two-group formalism.

It is straightforward to convince oneself that the above self-consistency conditions Eqs. (111)–(113) are identical to the saddle-point equations of the two-group ansatz (50) and (52), evaluated with the help of the annealed free-energy expression (57) (see Appendix A for details). In particular, we find that the two-group order parameters Q, A, and C coincide with the above $Q^\star$, $A^\star$, and $C^\star$.

We finally need to establish the precise correspondence between the regularized free-energy shift $\Phi_{\text{cav}}(m)$ computed within the generalized cavity method, and the replicated free-energy density $\Phi_{\text{2G}}(m)$ computed within the annealed two-group replica ansatz. We need to take into account that by adding a spin to an SK model with N spins, one obtains a system with slightly stronger couplings (by an average fraction of $1/2N$) than in a standard system with $N+1$ spins. This is basically equivalent to raising the inverse temperature to $\beta \to \beta' = \beta(1 + 1/2N)$ simultaneously with the spin addition.[19]

We thus expect the relationship

$$\Phi_{\text{cav}}(m) = \Phi_{\text{2G}} + \frac{1}{2} \frac{\partial(\beta \Phi_{\text{2G}})}{\partial \beta} \tag{114}$$

between the two free-energy densities. Explicit evaluation of the right-hand side using Eq. (47) indeed yields

$$-\beta m \Phi_{\text{2G}} - \frac{\beta}{2} \frac{\partial}{\partial \beta} (\beta m \Phi_{\text{2G}}) = \beta \phi_m + m \frac{\beta^2}{2} (1 - Q) - \beta^2 A, \tag{115}$$

which precisely coincides with $-\beta m \Phi_{\text{cav}}(m)$ from Eq. (108), if we recall the annealed version of Eq. (57) for ϕ_m.

We have thus proven the equivalence of the generalized cavity method and the two-group replica calculation for all free-energy densities in the annealed regime $f > f^\star$. Further, we confirmed once more the interpretation of the order parameters.

To use the cavity method beyond the annealed approximation is a rather cumbersome task; see Ref. 19. It is much easier to carry out the two-group computation with full replica symmetry breaking, even though it is less intuitive than the cavity approach. To help the reader appreciate the physical content of such a quenched calculation, we devote the following section to formalism.

VII. QUENCHED TWO-GROUP ANSATZ: FORMALISM AND INTERPRETATION

In Sec. IV D, we have seen that for free-energy densities $f > f^\star$ the annealed solution of the two-group replica approach is correct. For $f > f^\star$, however, the marginal metastable states are correlated and thus $Q_{ab}, A_{ab}, C_{ab} \neq 0$ for $a \neq b$. In analogy to the Parisi solution for the ground state of the SK model, we are looking for a hierarchical breaking of the replica symmetry encoding the assumption that marginal states are organized in an ultrametric tree in phase space: the smaller the distance on the tree, the larger the similarity between the states. In particular, we expect the off-diagonal part of the matrices Q_{ab}, A_{ab}, and C_{ab} to tend to the functions $q(x)$, $a(x)$, and $c(x)$, respectively, describing the continuous breaking of replica symmetry. We continue to denote diagonal entries as $Q_{aa} \equiv Q$, $A_{aa} \equiv A$, and $C_{aa} \equiv C$.

A. Replicated free energy

With such an ansatz for the overlap matrices, the quenched replicated free energy of m copies, Eq. (47), reads

$$-\beta m\Phi_{2G}^{qu} = \beta\phi(x=0,\mathbf{y}=(0,0)) + \frac{\beta^2}{4}m(1-Q)^2 - \frac{\beta^2}{4}m^2\left[Q^2 - \int_0^1 dx q^2(x)\right] - \frac{\beta^2}{2}[2A + A^2 + QC - 2(1-m)AQ] + \frac{\beta^2}{2}\int_0^1 dx[a^2(x) + q(x)c(x) + 2ma(x)q(x)], \tag{116}$$

where $\mathbf{y}\equiv(y_1,y_2)$. The function $\phi(x,\mathbf{y})$ is the free energy per replica of a subsystem of xm coupled replicas subject to external fields y_1,y_2 acting on the two groups of replicas [representing minima $(+)$ and saddles $(-)$] with the same and opposite sign, respectively: $y^{\pm}=y_1\pm y_2/(2K)$ [cf., the analogous expression Eq. (32) for the magnetizations]. More precisely, we define

$$\exp[x\phi(x,\mathbf{y})] = \sum_{s_a^i=\pm 1} \exp[\mathcal{H}(x,\mathbf{y},\{s_a^i\})],$$

$$\mathcal{H}(x,\mathbf{y},\{s_a^i\}) = \frac{\beta^2}{2}\sum_{a,b}^{1,x}\sum_{i,j}^{1,m} s_a^i \mathcal{Q}[x]_{ab}^{ij} s_b^j + \beta\sum_{a=1}^{x}\left[y_1\sum_{i=1}^{m} s_a^i + \frac{y_2}{2K}\left(\sum_{i^+} s_a^{i^+} - \sum_{i^-} s_a^{i^-}\right)\right]. \tag{117}$$

Here $\mathcal{Q}[x]_{ab}^{ij}$ denotes the matrix $\mathcal{Q}_{ab}^{ij}-Q^{\sigma_i\sigma_j}(x)$ restricted to a block of $(xm)\times(xm)$ replicas.

The representation (117) allows for the derivation of recursion equations for $\phi(x-dx,\mathbf{y})$ in terms of $\phi(x,\mathbf{y})$ (see, e.g., Ref. 44). In the limit of continuous overlap functions $(dx\to 0)$, they reduce to Parisi's differential equation,

$$\dot\phi = -\frac{\dot q}{2}\left[\frac{\partial^2\phi}{\partial y_1^2} + \beta x\left(\frac{\partial\phi}{\partial y_1}\right)^2\right] - \dot a\left[\frac{\partial^2\phi}{\partial y_1\partial y_2} + \beta x\frac{\partial\phi}{\partial y_1}\frac{\partial\phi}{\partial y_2}\right] - \frac{\dot c}{2}\left[\frac{\partial^2\phi}{\partial y_2^2} + \beta x\left(\frac{\partial\phi}{\partial y_2}\right)^2\right], \tag{118}$$

where a dot denotes $\partial/\partial x$. The boundary condition at $x=1$ follows from the definition (117) as

$$\phi(x=1,\mathbf{y}) = \lim_{K\to\infty}\frac{1}{\beta}\log\int dh\,dz\mu(h,z)\times\left[2\cosh\beta\left(y_1+h+\frac{y_2+z}{2K}\right)\right]^{m+K}\times\left[2\cosh\beta\left(y_1+h-\frac{y_2+z}{2K}\right)\right]^{-K}$$

$$= \frac{1}{\beta}\log\int dh\,dz\mu(h,z)[2\cosh\beta(y_1+h)]^m \times\exp[\beta(y_2+z)\tanh\beta(y_1+h)], \tag{119}$$

where

$$\mu(h,z)dhdz \equiv \mu(\underline{\xi})d^2\underline{\xi} = \frac{d^2\underline{\xi}}{2\pi\sqrt{\det\Delta}}\exp\left[-\frac{1}{2}\underline{\xi}^{\dagger}\cdot\Delta^{-1}\cdot\underline{\xi}\right], \tag{120}$$

$$\Delta \equiv \begin{pmatrix}\Delta q & \Delta a\\ \Delta a & \Delta c\end{pmatrix}, \tag{121}$$

$$\underline{\xi}^{\dagger} \equiv (h,z). \tag{122}$$

Here we introduced the notation $\Delta q\equiv Q-q(1)$, etc. for the jump between the diagonal entry and the closest off-diagonal elements of the overlaps. In general, these jumps are finite, indicating that individual marginal states are clearly distinct from their closest neighboring states in phase space. This is in contrast to the situation at $f=f_{eq}$ where the overlap of neighboring states can come arbitrarily close to the self-overlap Q.

B. Distribution of local fields y_1,y_2

Following Sommers and Dupont,[45] we also introduce the distribution $P(x,\mathbf{y})$ of local fields on the scale x requiring that

$$\langle s_{a_1}^{i_1}\cdots s_{a_r}^{i_r}\rangle = \int d\mathbf{y}P(x,\mathbf{y})\langle(s_{a_1}^{i_1}\cdots s_{a_r}^{i_r})\rangle_{\mathcal{H}(x,\mathbf{y})}$$

for all $x\leqslant\{a_1,\ldots,a_r\}\leqslant 1$. Note that here we need to keep track of the distribution of *both* fields y_1 and y_2.

In the continuous limit, the ensuing recursion relations relating P at x and $x+dx$ lead to the flow equations

$$\dot P = \frac{\dot q}{2}\left[\frac{\partial^2 P}{\partial {y_1}^2} - 2\beta x\frac{\partial}{\partial y_1}\left(P\frac{\partial\phi}{\partial y_1}\right)\right] + \frac{\dot c}{2}\left[\frac{\partial^2 P}{\partial {y_2}^2} - 2\beta x\frac{\partial}{\partial y_2}\left(P\frac{\partial\phi}{\partial y_2}\right)\right] + \dot a\left\{\frac{\partial^2 P}{\partial y_1\partial y_2} - \beta x\left[\frac{\partial}{\partial y_1}\left(P\frac{\partial\phi}{\partial y_2}\right) + \frac{\partial}{\partial y_2}\left(P\frac{\partial\phi}{\partial y_1}\right)\right]\right\}, \tag{123}$$

with the boundary condition at $x=0$,

$$P(x=0,\mathbf{y}) = \delta^{(2)}(\mathbf{y}). \tag{124}$$

The joint distribution of local fields (y_1,y_2) within a typical marginal state [with free energy density $f=f(m)$] is eventually obtained from the convolution

$$P_{qu}(\mathbf{y}) = [2\cosh(\beta y_1)]^m\exp[\beta y_2\tanh(\beta y_1)] \times\int\frac{d^2\underline{\xi}e^{-(1/2)\underline{\xi}^T\cdot\Delta^{-1}\cdot\underline{\xi}}}{2\pi\sqrt{\det\Delta}}P(1,\mathbf{y}-\underline{\xi})e^{-\beta\phi(1,\mathbf{y}-\underline{\xi})}. \tag{125}$$

The variable transformation (59) and (60), $\tilde{m}=\tanh(\beta y_1)$, and $\delta m=\beta y_2/\cosh^2(\beta y_1)$ yield the joint distribution of local magnetizations and soft-mode components,

$$P_{\mathrm{qu}}(\tilde{m},\delta m)=\int d^2\mathbf{y}P_{\mathrm{qu}}(\mathbf{y})\delta(\tilde{m}-\tanh(\beta y_1))\times\delta\left(\delta m-\frac{\beta y_2}{\cosh^2(\beta y_1)}\right), \tag{126}$$

one of the central objects of interest characterizing the metastable states.

C. Self-consistency equations for the order parameters

Using the definitions of ϕ and P, one can convince oneself that the continuous limit of the off-diagonal self-consistency equations (50)–(52) can be cast into the form

$$q(x)=\left[\left(\frac{\partial\phi(x,\mathbf{y})}{\partial y_2}\right)^2\right]_x, \tag{127}$$

$$a(x)=\left[\frac{\partial\phi(x,\mathbf{y})}{\partial y_2}\left(\frac{\partial\phi(x,\mathbf{y})}{\partial y_1}-m\frac{\partial\phi(x,\mathbf{y})}{\partial y_2}\right)\right]_x, \tag{128}$$

$$c(x)=\left[\left(\frac{\partial\phi(x,\mathbf{y})}{\partial y_1}-m\frac{\partial\phi(x,\mathbf{y})}{\partial y_2}\right)^2\right]_x, \tag{129}$$

where we have introduced the notation

$$[o(\mathbf{y})]_x\equiv\int_{-\infty}^{\infty}dy_1\int_{-\infty}^{\infty}dy_2P(x,\mathbf{y})o(\mathbf{y}) \tag{130}$$

to denote an average over the local field distribution at the scale x. Alternatively, Eqs. (127)–(129) derived from a variational formulation of the quenched problem.[29]

For the diagonal (intrastate) overlaps ($a=b$), we have

$$Q=\langle\tilde{m}^2\rangle, \tag{131}$$

$$A=\langle\tilde{m}\delta m\rangle, \tag{132}$$

$$C=\langle\delta m^2\rangle, \tag{133}$$

where averages are over the joint distribution Eq. (126). Integrating out explicitly the soft modes, this can be cast into a form similar to the annealed equations (66)–(68),

$$Q=\langle\tanh^2(\beta y_1)\rangle, \tag{134}$$

$$A+1-Q=-Q\frac{\Delta a}{\Delta q}-mQ+\frac{\langle y_1\tanh(\beta y_1)\rangle}{\beta\Delta q}, \tag{135}$$

$$C-2A-m(1-Q)=m^2Q+2mQ\frac{\Delta a}{\Delta q}+Q\left(\frac{\Delta a}{\Delta q}\right)^2-2\left(m+\frac{\Delta a}{\Delta q}\right)\frac{\langle y_1\tanh\beta(y_1)\rangle}{\beta\Delta q}-\frac{1}{\beta^2\Delta q}\left(1-\frac{\langle y_1^2\rangle}{\Delta q}\right), \tag{136}$$

where $\langle\cdots\rangle$ denotes an average over $P_{\mathrm{qu}}(y_1)\equiv\int dy_2P_{\mathrm{qu}}(y_1,y_2)$.

VIII. INTERNAL CONSISTENCY AND THERMODYNAMIC STABILITY

Sections IV, V, and VI, provide three equivalent methods to capture the properties of marginal states at a given free-energy density above $f^\star$, and the previous section extended the formalism to the low-energy regime. Yet we did not bother so far about the internal consistency and thermodynamic stability of the obtained solutions. Moreover, we merely stated the existence of marginal states without analyzing in more detail the eigenvalue spectrum of the free energy Hessian and the related question of local stability of the states. In this section we address these issues, and show how to obtain more information about the local environment of the marginal states.

Even though the analysis of the local stability and the thermodynamic consistency are two a priori very different aspects of the problem, they turn out to be closely related.

A. Internal consistency of the 2G solution

In order to understand the local free-energy landscape of a given metastable state (TAP solution), we need to characterize the free-energy Hessian of a typical local minimum, or in other words, the inverse of the susceptibility matrix, $\chi_{ij}^{-1}=\beta\partial_i\partial_jF_{\mathrm{TAP}}(\{m\})$.[46,47] The fluctuation-dissipation relation requires that the ("zero-field cooled") susceptibility, i.e., the trace of χ, be equal to $\beta(1-Q)$.[47] However, not all solutions of the TAP equations satisfy this constraint, but only those for which the inequality

$$x_P\equiv 1-\beta^2\frac{1}{N}\sum_i(1-m_i^2)^2\geqslant 0 \tag{137}$$

holds. Other solutions are unphysical.

This condition always has to be checked separately, after having obtained a self-consistent solution of site magnetizations m_i.

For $N\to\infty$, this can be rewritten as

$$x_P=1-\beta^2(1-2q+\langle\tilde{m}^4\rangle)\geqslant 0, \tag{138}$$

where the average is over the appropriate magnetization distribution [Eq. (65) with $\tilde{m}\equiv\tanh(\beta h)$ in the annealed case, and Eq. (126) in the quenched case].

In the regime $f>f^\star$, the condition (138) is satisfied as a strict inequality in all marginal states. In contrast, if one aims at describing genuinely stable minima by imposing saddle points that conserve the BRST symmetry, the condition (138) is always violated. This leads to the conclusion that in the SK model there are no genuinely stable TAP states that are not closely related to the family of dominating marginal states [see Ref. 27 (annealed case) and Refs. 29 and 48 (quenched case)]. While we know from the regularization procedure that there are actually stable states, they are always close to being marginal, sharing similar properties with the dominating marginal states.

The only thermodynamically consistent states that do not break the BRST symmetry are the states at the equilibrium free-energy density f_{eq} for which the criterion (138) is marginally satisfied as an equality.

We note that Eq. (138) is actually equivalent to the requirement of a positive *replicon* Λ_R, defined as the smallest eigenvalue characterizing the fluctuations of the replicated free energy as a function of the order parameter matrix $\mathcal{Q}_{ab}$. In fact, Λ_R turns out to be simply proportional to x_P.[49] The implications of a vanishing replicon, and the possibility of its simultaneous occurrence with the breaking of the BRST symmetry in the regime $f_{eq}<f<f^\star$, will be discussed below.

B. Free-energy Hessian

The extensive part of the spectrum of the inverse susceptibility matrix $\chi^{-1}(\{m\})$ in a generic TAP state $\{m_i\}$ was determined in Ref. 47 from an analysis of the TAP equations, neglecting terms of order $O(1/N)$. The extensive part starts off as a semicircle,

$$\rho(\lambda) = \frac{N}{\pi\sqrt{p}}\sqrt{\lambda-\lambda_0}, \quad \lambda-\lambda_0 \ll 1, \tag{139}$$

where the lower band edge λ_0, together with the resolvent $r_0=\mathrm{Tr}[(\lambda_0-\chi^{-1})^{-1}]$, follows from the solution of

$$r_0 = -f_1(r_0+\lambda_0), \tag{140}$$

$$1 = f_2(r_0+\lambda_0), \tag{141}$$

$$f_n(x) \equiv \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\left(\frac{1}{\beta(1-m_i^2)}+\beta(1-Q)+x\right)^n}. \tag{142}$$

The semicircle's amplitude is controlled by $p=f_3(r_0+\lambda_0)$. The support of the continuous part was proven to be strictly positive,[47] $\lambda_0 \geqslant 0$. However, this result, valid to leading order in N, does not exclude the presence of a subextensive number of negative eigenvalues. In fact, such regions of phase space *must* exist as guaranteed by the Morse theorem.

One can easily check that if $x_P=0$, the solution of Eqs. (140) and (141) is $\lambda_0=0$, $r_0=-\beta(1-Q)$, while $\lambda_0=0$ implies $x_P=0$ and $r_0=-\beta(1-Q)$. In particular, the vanishing of the parameter x_P and of the "band gap" λ_0 occur simultaneously:[12,47] $\lambda_0 \simeq x_P^2/(4p)$.

At finite N, the spectrum exhibits tails below the extensive band edge λ_0 that may even extend to negative eigenvalues. However, as is usual in random matrix problems, it is reasonable to assume that their density decays exponentially as $\exp(-aN^{1/6})$,[50] so that they do not survive for large N.

However, there is one eigenvalue below the gap edge that survives in the thermodynamic limit. Indeed, including the corrections of order $O(1/N)$ in the analysis of TAP states, one finds that the Hessian possesses a single isolated eigenvalue, which is not captured by the analysis to leading order. Such an eigenvalue is a common feature of many mean-field spin-glass models. To our knowledge, it was first encountered as *the longitudinal eigenvalue* in the spherical p-spin model[2] where it is positive and the corresponding states are genuine minima. The situation is presumably similar in the free-energy regime of the p-spin Ising model where states are stable.[6,7] In contrast, for the regime $f>f^\star$ of the SK model, it was proven that the isolated eigenvalue is exactly zero in the thermodynamic limit.[3,5] This provided the first evidence for the marginality of the dominant states.

An exactly vanishing eigenvalue should actually be present in any BRST-breaking states. In this paper, we used this insight as a starting point, assuming a soft eigenvalue λ_{soft}, and deriving its self-consistency in various ways in the preceding sections.

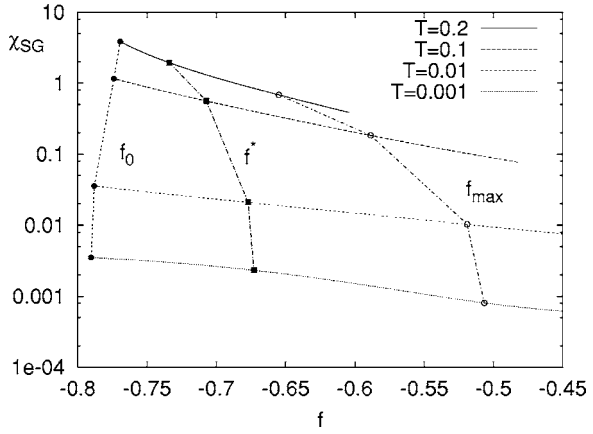

FIG. 4. Spin-glass susceptibility of marginal states, χ_{SG}, vs free-energy density f at different temperatures $T=0.2$, 0.1, 0.01, and 0.001. The susceptibility is plotted in log-scale. The physical results hold for $f>f^\star$. f_0 is the point at which Σ goes to zero in the annealed approximation.

C. Spin-glass susceptibility

The spin-glass (SG) susceptibility (see Fig. 4) is defined as

$$\chi_{SG} = \frac{1}{N}\sum_{ij}\chi_{ij}^2 = \frac{1}{N}\mathrm{Tr}\chi^2 = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{\lambda_j^2}, \tag{143}$$

where λ_j are the eigenvalues of the Hessian. It has a simple expression in terms of the above introduced parameter x_P,[12]

$$\chi_{SG} = \frac{1-x_P}{x_P}, \tag{144}$$

which is valid for physical states with $x_P \geqslant 0$.

As $x_P \to 0$ the spin-glass susceptibility diverges. As mentioned above, this happens when the gap λ_0 vanishes, i.e., when there is an accumulation of eigenvalues of the Hessian at $\lambda=0$, cf. Eq. (139). Such states are "fully marginal," in the sense that they possess an extensive number of soft modes.[51]

In marginal states with only one soft mode and a finite gap λ_0 to the continuous part of the spectrum, the spin-glass susceptibility remains finite. Indeed, the linear susceptibility diverges only along the marginal direction, which results in a nonextensive effect for χ_{SG}.[52]

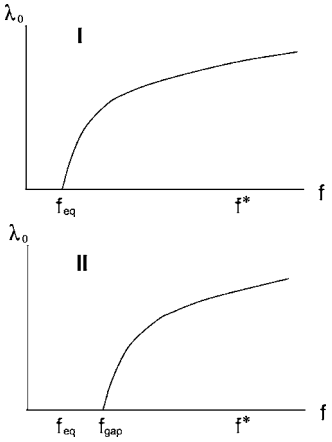

FIG. 5. Possible scenarios for the gap in the spectrum, λ_0, at free-energy densities below the threshold of validity of the annealed approximation, $f^\star$. Scenario I: the gap λ_0 goes to zero at f_{eq}, as well as the soft-mode overlaps $\{a(x),A\}$ and $\{c(x),C\}$. Scenario 2: the gap goes to zero at higher free energy (f_{gap}) implying the existence of *fully marginal states* also above the global minimum.

D. Discussion

At this stage, we only have an exact description of the free-energy regime $f>f^\star$, while an analysis of the low-energy regime down to f_{eq} requires the quenched solution of the two-group equations of Sec. VII. The latter is a technically difficult task and will be carried out elsewhere.[53] Here, we will content ourselves with a discussion of possible scenarios for the local environment of dominating states as one decreases the free energy.

The spectrum of the Hessian of a TAP state depends on its free-energy density f. As we have seen above, all states with free energies $f \geqslant f^\star$ have a finite gap $\lambda_0>0$ in the spectrum, while the thermodynamically dominant state at f_{eq} is fully marginal,[12] as implied by the vanishing of the replicon, $\lambda_0 \propto x_P^2 \propto \Lambda_R^2=0$. The latter is an interesting situation, which leads to a nontrivial dynamical behavior owing to the multitude of flat directions in phase space. It is thus an important question to establish whether full marginality is a unique property of the equilibrium state, or whether there actually exists a window of free energies $f_{eq} \leqslant f \leqslant f_{gap}$ where the dominant states are fully marginal. Here f_{gap} denotes the highest free-energy density where the gap λ_0 in the Hessian spectrum vanishes.

The following scenarios are possible in principle.

(i) Scenario I (see Fig. 5, top). In the simplest scenario, the gap λ_0 tends smoothly to zero as f tends to f_{eq} (together with the BRST breaking order parameters A and C) and thus $f_{gap}=f_{eq}$. While the soft mode tends to become orthogonal to the magnetization ($A \to 0$), its amplitude vanishes simultaneously ($C \to 0$) as the equilibrium state is approached. Under this scenario, only the states at f_{eq} are fully marginal and the formalism developed in Secs. IV and VII should provide a consistent description for the dominant states (with a finite complexity) at all free energies and temperatures.

(ii) Scenario II (see Fig. 5, bottom). Fully marginal states already occur at $f_{gap}>f_{eq}$. In this case, it is *a priori* not clear whether the notion of the (isolated) soft eigenvalue continues to make sense, since it plunges into a continuum of other eigenvalues. If, nevertheless, one of the many soft modes can be singled out (e.g., with the help of the projector term appearing to order $1/N$ in the TAP equations[3,5]), the two-group formalism may probably still describe this regime. Conversely, if one finds a thermodynamically consistent two-group solution with $\lambda_0=0$, this strongly suggests that the singled out soft mode still preserves a meaning. In this case, the solution would still be characterized by broken BRST symmetry ($C>0$), but we conjecture a vanishing order parameter $A=0$: while the magnitude of the soft mode stays finite, its direction is expected to be orthogonal to the magnetization vector in phase space, as we will argue in the next section. Further, we would expect that this branch of solutions continuously joins the BRST symmetric Parisi solution (i.e., $C \to 0$) as $f \to f_{eq}$. Such a branch would yield an extensive complexity, implying an exponential number of fully marginal states. On the other hand, the isolated soft mode might lose its meaning as the gap closes, and the two-group ansatz might have no solution with nonzero soft-mode amplitude C, suggesting that both A and C vanish at f_{gap}. With $A=C=0$, however, the description of the system reduces to the standard Parisi ("one-group," BRST symmetric) ansatz, and it has been shown in earlier studies[29,48] that no thermodynamically consistent BRST symmetric continuation of the equilibrium state at f_{eq} exists, not even within a quenched RSB computation. The only option left would then be the possibility that no states at all exist in the interval $]f_{eq},f_{gap}[$, unless a completely different replica symmetry breaking scheme is considered. This last scenario is rather unlikely, given that the absence of states between f_{eq} and f_{gap} is hardly reconcilable with numerical studies of TAP solutions,[4] which provide evidence for TAP states basically at all free energies.

IX. LOCAL FIELD DISTRIBUTION IN THE ANNEALED REGIME

In this section, we analyze in more detail the local field distribution $P_{ann}(h)$, Eq. (65), of marginal states in the free-energy regime $f \geqslant f^\star$ where the annealed description applies. We remind the reader that the distribution depends on the selected free-energy density via the Legendre transform parameter m.

A. Low-temperature analysis

The low-temperature limit of the complexity was studied long ago in Refs. 18, 54, and 55. Here, we focus on the susceptibility and relate it to the properties of $P(h)$ that are rather unexpected and, to our knowledge, have not been discussed before.

As usual, the relevant range of the Legendre parameter m scales as $m \sim T$ at low temperatures. We thus use the variable $\omega=m/T$ to obtain a sensible $T=0$ limit. We anticipate that the susceptibility behaves as $\chi=\beta(1-Q) \sim T$, similarly to what is known from Parisi's ground-state solution. This assumption will turn out to be self-consistent. The saddle-point equations (67) and (68) for the order parameters A, C suggest a low-temperature scaling

$$\lim_{\beta\to\infty} \beta A = \lim_{\beta\to\infty} \frac{\beta C}{2} \equiv -\epsilon - \frac{\omega}{2}, \tag{145}$$

where

$$\epsilon = -\frac{1}{2}\lim_{\beta\to\infty}\int_{-\infty}^{\infty} dh P_{\rm ann}(h) h \tanh(\beta h)$$

$$= -\frac{1}{2}\lim_{\beta\to\infty}\int_{-\infty}^{\infty} dh P_{\rm ann}(h)|h| \tag{146}$$

is the (ω-dependent) energy density of the selected states, as follows from the low-temperature limit of the energy $E=-1/2\Sigma_{ij}s_iJ_{ij}s_j$ and the local fields $h_i=\Sigma_j J_{ij}s_j$.

In the regime of local fields

$$T\log(1/T) \ll h \ll \beta A \approx -\epsilon - \omega/2, \tag{147}$$

the field distribution Eq. (65) takes the low-temperature form

$$P_{\rm ann}(h) = \mathcal{N}_0^{-1} \exp\left[\frac{\omega}{2}|h| - \frac{(|h|+\epsilon)^2}{2}\right], \tag{148}$$

$$\mathcal{N}_0 = \int dh \exp\left[\frac{\omega}{2}|h| - \frac{(|h|+\epsilon)^2}{2}\right]. \tag{149}$$

The self-consistency Eq. (67) then reduces to[54]

$$\epsilon = -\frac{1}{2}\int_{-\infty}^{\infty} dh|h|P_{\rm ann}(h)$$

$$= \frac{\epsilon}{2} - \frac{\omega}{4} - \frac{\exp[-\epsilon^2/2]}{\mathcal{N}_0}, \tag{150}$$

which relates ϵ and the Legendre parameter ω.

B. Soft-mode direction

As a corollary of the low-T scaling Eq. (145), we obtain the behavior of the angle between the soft mode and the magnetization as a function of (free) energy density,

$$\gamma(\epsilon;T) \approx \frac{\pi}{2} - [-\epsilon - \omega(\epsilon)/2]^{1/2}T^{1/2}, \quad \epsilon \geqslant \epsilon^\star, \tag{151}$$

where the $T=0$ limit of the annealed threshold $f^\star$ is $\epsilon^\star=-0.672$. In particular, we note that the lower the temperature, the more the soft mode tends to be perpendicular to the magnetization, cf. Fig. 3.

C. Susceptibility

In order to confirm that indeed $\chi\sim T$, we write

$$\chi = \beta(1-Q) = \int_{-\infty}^{\infty} dh P_{\rm ann}(h)\frac{\beta}{\cosh^2(\beta h)}. \tag{152}$$

The integral is dominated by $T\ll h\sim T\log(1/T)$, where the local field distribution simplifies to

$$P_{\rm ann}(h) \approx \mathcal{N}_0^{-1}\exp\left[-\frac{\beta(-\epsilon-\omega/2)}{\cosh^2(\beta h)} - \frac{\epsilon^2}{2}\right]. \tag{153}$$

Changing variables to $\xi=\beta/\cosh^2(\beta h)$, we indeed find

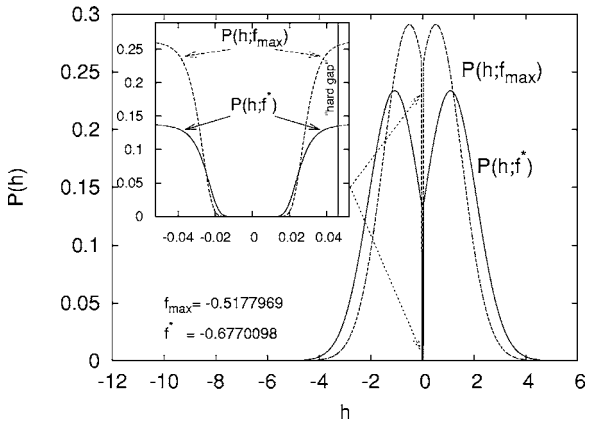

FIG. 6. Local-field distribution Eq. (153) in the annealed regime for $f_{\rm max}$ and $f^\star$ at $T=0.01$. In the inset, the region around $h=0$ is enlarged to show the nearly hard gap occurring at any f in the annealed regime. The vertical line marks $T\log T$, which is the scale $h_{\rm gap}$ at which the gap is expected to open.

$$\chi \approx T\int_0^\beta d\xi \frac{\exp[-(\epsilon-\omega/2)\xi - \epsilon^2/2]}{\mathcal{N}_0} \approx \frac{T}{2}, \tag{154}$$

where we have made use of Eq. (150). Surprisingly, the susceptibility does not depend on the energy density ϵ in the low-T limit. The origin of its linear T dependence will be discussed further below.

D. Discussion

1. Pseudogap and full marginality

The local field distribution Eq. (153) is rather peculiar (it is plotted for $T=0.01$ in Fig. 6). It differs significantly from that in the equilibrium state, both for thermally active sites (with fields of order $h\sim T$) as for larger fields of order $h\sim O(1)$.

The field distribution of the ground state is known to assume a universal low-temperature scaling form[45]

$$P(h,T) = \begin{cases} T\psi(h/T) & \text{for } h\sim T, \\ \text{cst.}|h| & \text{for } h \gg T, \end{cases} \tag{155}$$

which is related to an asymptotic fixed point in Parisi's flow equations.[56] The expression (155) implies a susceptibility linear in T, arising from $O(T^2)$ spins with a local linear response of order $1/T$.

In contrast, in the marginal states at higher energies, $f>f^\star$, the linear susceptibility, Eq. (154), is due to $O(T)$ spins with a local response of order $O(1)$ [on sites with fields $h\sim T\log(1/T)$]. On the other hand, the density of really active spins (subject to local fields of order $h\lesssim T$) is exponentially suppressed at low temperature,

$$P_{\rm ann}(h=0) \sim \exp[-\beta(-\epsilon-\omega/2)]. \tag{156}$$

The scaling behavior Eq. (155) in the ground state is closely related to its full marginality.[17,45] Indeed, full marginality requires the vanishing of the replicon—see Sec. VIII B ($\Lambda_R \propto x_P \to 0$), which in turn is equivalent to the marginality condition

$$1 = \beta^2 \int dh \frac{P_{\mathrm{qu}}(h)}{\cosh^4(\beta h)}. \tag{157}$$

Here, the local field distribution $P_{\mathrm{qu}}(h)$ is given by the quenched Parisi solution [i.e., Eq. (125) in the limit $A = a(x) = C = c(x) = \Delta q = 0$, with $y_1 \equiv h$ and y_2 being trivially integrated out]. From the equality (157), valid for all T, one infers the low-temperature scaling (155).

With the connection between full marginality and h/T scaling in mind, the lack of the latter in higher marginal states (with only one soft mode) may not be too surprising, since we know that above $f^\star$ the states are not fully marginal (the replicon Λ_R remains finite throughout). Further consequences for the regime $h = O(1)$, such as the absence of the linear pseudogap implied by Eq. (155), will be discussed in the next subsection.

There is a noteworthy aspect of the nearly hard gap in the low-field distribution Eq. (156): the exponential suppression of $P_{\mathrm{ann}}(h \lesssim T)$ (see the inset of Fig. 6) decreases with $-\epsilon - \omega/2 \approx \beta A$ and eventually would vanish at low enough energies as $A \to 0$. This is, however, preempted by the breakdown of the annealed approximation and the occurrence of full replica symmetry breaking at $f^\star$, which requires a quenched computation. Nevertheless, the above observation suggests that the vanishing of A is connected with the onset of *full* marginality, i.e., the approach to zero of the continuous spectrum of the free-energy Hessian. This is supported by a result by Parisi and Rizzo,[5] who showed that as the continuous spectrum approaches $\lambda_0 \to 0$, the (isolated) marginal mode becomes orthogonal to the magnetization (that is, $A = 0$ in our formalism). We thus conjecture that in general the vanishing of the gap in the spectrum, $\lambda_0 \propto x_p^2 = 0$, implies the vanishing of A_{ab}. With respect to the discussion of the previous Sec. VIII D, this is guaranteed to happen in scenario I, while it constitutes a nontrivial prediction for scenario II.

2. On stability and slow dynamics

The field distribution Eq. (153) differs from the equilibrium one also in the range of fields $h = O(1)$. In the limit $T \to 0$, the fraction of fields with a fixed magnitude $T \ll h \ll 1$ remains bounded from below by a finite constant, cf. Eq. (65). In particular, there is *no* linear pseudogap $P(h) \sim |h|$, in contrast to the equilibrium state, cf. Eq. (155). Strictly at $T = 0$ there is a finite probability of configurations at $h = 0$, while at finite T there is an almost hard gap on the scale $T \log T$ as discussed in the previous section.

The absence of a linear pseudogap immediately raises the question about the stability of these states. As was realized in the early days of the SK model, a linear pseudogap is the minimal suppression of the low-field distribution for a truly stable state see, e.g., the discussion in Ref. 57, a configuration with a finite density of local fields around $h = 0$ being unstable to the flipping of a finite number of spins (exactly at $T = 0$). A very similar argument led Efros and Shklovskii to infer the presence of the Coulomb gap in long-range interacting electron glasses.[58] Recently, it was recognized that both pseudogaps are related to the full marginality of the equilibrium state.[45,59,60]

The absence of a pseudogap in marginal states would seem to render them unstable with respect to a finite number of spin flips. While this is true strictly at $T = 0$,[18] the finite-temperature analysis is more subtle, because the thermodynamic limit and the $T \to 0$ limit do not commute. One can apply stability arguments analogous to those of Ref. 57 to the field distribution Eq. (153) of marginal states above $f^\star$. Such an analysis shows that at low but finite temperature, a collective flip of order $O([T \log(1/T)]^2 N)$ spins (randomly chosen among the sites with small local fields) is necessary in general to render the state unstable. This contrasts with $O(T^2 N)$ flips for states with a linear pseudogap, but no hard gap on the scale of $h \sim T$. The logarithmic enhancement of stability against *random* spin flips for $f > f^\star$ is indeed due to the presence of an almost hard gap on the scale of $h_{\mathrm{gap}} \sim T \log(1/T)$, see the inset of Fig. 6. One may therefore expect that a finite T dynamics based on random activated spin flips exhibits similarly long (if not longer) escape times as low-energy states with a linear pseudogap.

On the other hand it is clear that a collective flip of a set of spins with a significant projection onto the marginal mode will take the system immediately out of the local state since the free-energy barrier presumably decreases to zero in the thermodynamic limit, as suggested by recent simulations.[39] However, such an escape may be very difficult to realize since it amounts to flipping a large number of spins in a concerted manner. In summary, the presence of a single marginal direction does not seem to decrease the metastability of higher-lying marginal states in a dramatic way, rather their stability seems comparable to that of states closer to equilibrium.

At the present stage, the question as to the consequences of marginal states and their local environment on the dynamics is still open. Nevertheless, assuming that the dynamics is eventually dominated by the "easy" escape via the marginal mode, we may expect that the rate at which the Edwards-Anderson parameter Q decreases depends crucially on the angle γ between magnetization and soft mode, cf. Eq. (39), and slows down as $\cos(\gamma) = A/\sqrt{CQ} \to 0$ with decreasing f. This is illustrated in Fig. 7, where we plot the decreasing overlap Q as a function of the increasing angle γ.

This picture is consistent with relaxation dynamics that takes the system toward lower and lower-lying metastable states. Indeed, $Q(f)$ *decreases* as f decreases (at least in the annealed regime, see Fig. 8). At first sight, this result looks counterintuitive, being opposite to the behavior of the self-overlap in familiar one-step glasses with genuinely stable states (e.g., p-spin models below the marginality threshold) where $Q(f)$ increases as one descends to lower energies. It may be interesting to note that an increase of $Q(f)$ with decreasing f is found in the BRST-symmetric solution of Eqs. (54)–(56); see the inset of Fig. 8. However, in the SK model this solution is thermodynamically inconsistent at all free energies above f_{eq}, even within a quenched computation.[29]

The above observation suggests to think of the marginal metastable states discussed here as relatively small pockets of phase space that restrict the local magnetizations to slightly *higher* values than in more equilibrated states. Relaxation dynamics out of these shallow traps, and subsequent

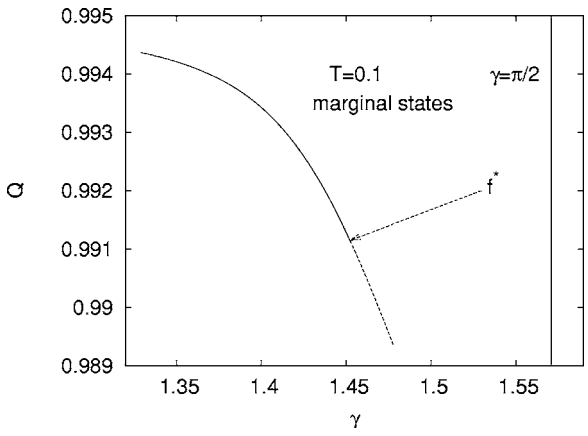

FIG. 7. Overlap Q vs angle γ at $T=0.1$. The soft mode becomes more and more orthogonal to the magnetization vector as the free energy f and the Edwards-Anderson parameter Q decrease. The vertical line indicates $\gamma=\pi/2$, which is not reached within the annealed regime. The dashed low-energy part of the curve $Q(\gamma)$ is unphysical in the annealed approximation.

descent in free energy, will allow the spins to explore more degrees of freedom and to lower the self-overlap Q. Clearly, it will be important to check such a picture in analytic studies of glassy dynamics as well as in simulations, for which the interpretation of the order parameters in terms of soft-mode correlations may provide helpful guidelines.

X. OUTLOOK AND OPEN QUESTIONS

Our study of the local landscape around a given metastable state leaves open the question as to the organization of these states within phase space. In particular, one would like to understand the typical fate of a system once it manages to escape from a local marginal trap via the soft mode. This would be an important element to a more complete understanding of relaxation dynamics as well as to avalanchelike dynamics observed when the SK glass is driven with an external field.[61]

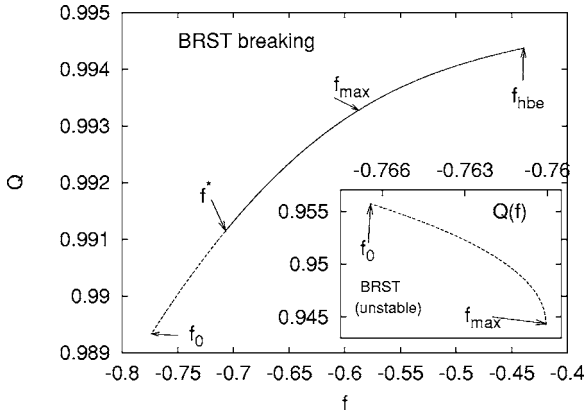

FIG. 8. Free-energy dependence of the order parameter Q in the two-group approach at $T=0.1$. Note the *decrease* of $Q(f)$ with free energy. For comparison, we also plot the (unphysical) BRST-symmetric solution where $Q(f)$ displays the opposite behavior. $f_{\max}$ denotes the value of maximum Σ, f_0, and $f_{\rm hbe}$, the point of vanishing Σ (lower band edge and upper band edge, respectively). Above $f^\star$, the annealed solution is exact.

The toy model of Sec. III was very helpful to understand the meaning of the two-group ansatz. Similar models may be guidelines as to how to use "vectorial symmetry breaking" in order to obtain physical information in complicated glasses, and to address questions such as the above. A natural example is the random field Ising model, where an extension of the two-group ansatz to three groups can be used to describe dominant Griffith singularities, and to capture rare disorder configurations that describe the emergence of metastable states in the ferromagnetic regime.[62]

A. Describing barriers with replicas

Another interesting application of the two-group ansatz might be the study of barriers. Indeed, from the solution of the toy model it is clear that replica saddle points with finite $k_2 \equiv K$ describe disorder realizations with two stable minima separated by a finite barrier of order $O(1/K)$. It may be interesting to pursue this idea and to work at fixed $K<\infty$ in order to impose a certain barrier height, just as fixing m allows one to choose the mean free-energy density of the minimum-saddle pairs. One may hope that such an approach will yield information on the properties of barriers between metastable states, which is a notoriously difficult and unresolved problem in glasses.

B. Properties of states in the correlated regime

Even though we are able to describe some key features of metastable states in the high-energy regime $f>f^\star$ where most marginal states are uncorrelated among each other, the analysis of lower-lying states in the free-energy range $f_{\rm eq}<f<f^\star$ requires full replica symmetry breaking on top of a two-group replica structure. The assumption that marginal states are clustered in a hierarchical manner in much the same way as states close to the ground-state energy suggests to look for a Parisi-type ansatz in this energy range, the physical content of which we have described in Sec. VII.

It will be important to establish how the characteristics of the local landscape (marginal mode, gap in the spectrum of the Hessian) behave as a function of free energy in this regime. In particular, one would like to know whether the breaking of BRST-symmetry persists in the whole interval, as actually suggested by the impossibility to construct a physical BRST-symmetric solution at any $f>f_{\rm eq}$.[29,48] This issue is closely related to the interesting open question: do any typical states with $f>f_{\rm eq}$ display full marginality, or is this intriguing property (implying a divergent spin glass susceptibility, critical fluctuations, and dynamics) unique to the ground state? This important issue will be clarified by the quenched solution of the two-group ansatz.[53]

It will be interesting to extend the two-group analysis to other mean-field-like models, such as random manifolds with a large number of transverse dimensions[63] or electron glasses.[59,60] In the latter problem, the presented techniques may help to address experimentally relevant questions con-

cerning the evolution of the Coulomb gap with decreasing free energy.

C. Ground-state properties in general mean-field glasses

The SK model is rather special in that its equilibrium state is always fully marginal below the glass transition. This observation, and the above discussion of the approach to full marginality as a function of free energy, raises a question about the nature of low-energy states in more general mean-field glasses. In the Ising p-spin model, it has been established[6] that above the so-called Gardner temperature T_G the states in an energy interval $[f_{\rm eq},f_G]$ including the ground state are stable minima described by (one-group) one-step RSB. At higher energies, marginal minima-saddle pairs dominate, as in the SK model (two-group one-step RSB). Precisely at $f=f_G$, where the two regimes meet, the states are fully marginal. At $T=T_G$, the energy interval of stable states shrinks to zero, $f_G \to f_{\rm eq}$, and consequently the (one-step) ground state displays an instability. What happens to the ground state at lower temperature is not known. Both the scenario of a permanently fully marginal ground state similar to the SK model, or a marginal ground state with a single soft mode (continuing the branch of solutions at higher free energies above T_G), can be imagined. This open problem is currently under investigation. Similar questions arise at phase transitions in many optimization problems (at $T=0$) on diluted lattices. In this context, it would be very interesting to investigate the physical meaning of soft modes at zero temperature.[37]

XI. CONCLUSION

In this paper, we have discussed three equivalent approaches to the description of metastable states in mean-field glasses: the replica two-group formalism, the counting of TAP solutions, and the cavity method. We have focused on the generalizations of these techniques allowing one to capture the marginal states that generically dominate the free-energy landscape at high enough free-energy densities, and probably play an important role in the dynamics of these glasses. We found the physical meaning of the additional order parameters (generalized overlaps) arising in the marginal case: they describe correlations and relative orientations between the magnetization and the soft mode of marginal states.

We show the two-group replica formalism to be an effective and compact computational tool to obtain most of the key features characterizing the local landscape of marginal states (such as the distribution of local fields, the spectrum of the free-energy Hessian, the direction of the soft mode, etc.), in particular when it comes to the description of correlated low-energy states. We have made an effort to exhibit the physical content of this ansatz in order to render this tool accessible to a wider audience. Extensions of this ansatz can be of use in other problems, e.g., for the study of barriers in glasses. Simple toy models as introduced in the beginning (Sec. III) may serve as useful guides in the development and interpretation of such techniques.

Revisiting the uncorrelated high-energy regime of metastable states where an annealed disorder average can be performed, we have discussed their properties in the light of the new interpretation of order parameters. Further investigations will be required to establish the dynamical significance of these states.

ACKNOWLEDGMENTS

The authors are indebted to C. De Dominicis, M. Mézard, and G. Parisi for stimulating discussions.

APPENDIX A: EXPRESSIONS ϕ_m

The log-trace term (48) in $-\beta m \Phi_{2G}$ is given by

$$\beta\phi_m \equiv \frac{1}{n} \log \sum_{\{s_a^i\}} \exp\left(\frac{\beta^2}{2} \sum_{ab} \sum_{ij} s_a^i Q_{ab}^{ij} s_b^j \right).$$

Here we derive two representations of this term [Eqs. (53) and (57)] in order to facilitate the connection with the counting of TAP states of Sec. V and the generalized cavity approach of Sec. VI, respectively.

Decoupling spin products in Eq. (A1) by means of Hubbard-Stratonovich transformations, one obtains the representation

$$\begin{aligned} e^{n\beta\phi_m} = K^n \sqrt{\frac{\det \Gamma_+ \det \Gamma_-}{\det Q}} \int \prod_{a=1}^{n} \frac{dg_a}{\sqrt{2\pi}} \frac{dg_a^+}{\sqrt{2\pi}} \frac{dg_a^-}{\sqrt{2\pi}} \\ \times \sum_{s_a^{i_\pm}} \exp\left[\sum_a \left(\sum_{i_+=1}^{m+K} \beta(g_a + g_a^+) s_a^{i_+} \right.\right. \\ \left.\left. + \sum_{i_-=m+K+1}^{m} \beta(g_a + g_a^-) s_a^{i_-} \right) \right] \\ \times \exp\left[-\frac{1}{2} (g_a Q_{ab}^{-1} g_b + g_a^+ \Gamma_{+,ab}^{-1} g_b^+ + g_a^- \Gamma_{-,ab}^{-1} g_b^-) \right], \end{aligned} \tag{A1}$$

with $n \times n$ matrices

$$\Gamma_{\pm,ab} \equiv \pm \frac{A_{ab}}{K} + \frac{C_{ab}}{2K^2}. \tag{A2}$$

1. Saddle point for $g_a^\pm$ for large K

For large K one can expand the Γ matrices in Eq. (A2) as

$$\Gamma_\pm^{-1} \simeq K\left(\pm A^{-1} + \frac{R}{K} \right), \tag{A3}$$

$$R = -\frac{1}{2} A^{-1} C A^{-1}. \tag{A4}$$

We isolate the terms of order $O(K)$ in the exponential of Eq. (A1), and perform the spin sums,

$$e^{n\beta\phi_m} \simeq \frac{|K|^n}{|\det A|\sqrt{\det Q}} \int \prod_{a=1}^{n} \frac{dg_a}{\sqrt{2\pi}} \frac{dg_a^+}{\sqrt{2\pi}} \frac{dg_a^-}{\sqrt{2\pi}} \times \exp\left\{ m \sum_a \ln 2\cosh[\beta(g_a + g_a^+)] + K\left[\sum_a \ln \frac{\cosh[\beta(g_a+g_a^+)]}{\cosh[\beta(g_a+g_a^-)]} - \frac{1}{2}\sum_{ab} A^{-1}_{ab}(g_a^+ g_b^+ - g_a^- g_b^-) \right] - \frac{1}{2}\sum_{ab}[g_a Q^{-1}_{ab} g_b + R_{ab}(g_a^+ g_b^+ + g_a^- g_b^-)] \right\}. \tag{A5}$$

In the limit $K \to \infty$, we can take the saddle point approximation for the term in the exponent proportional to K, which leads to the saddle point equations (with respect to g_a^+, g_a^-)

$$\tanh[\beta(g_a + g_a^+)] = -\beta^{-1} A^{-1}_{ab} g_b^+, \tag{A6}$$

$$\tanh[\beta(g_a + g_a^-)] = -\beta^{-1} A^{-1}_{ab} g_b^-, \tag{A7}$$

determining the location of the saddle point as $g_a^+ = g_a^- = g_a^*(\{g_c\})$; g_a^+ and g_a^- can then be integrated out. Decoupling the term $g_a Q^{-1}_{ab} g_b$ with a further Hubbard Stratonovich field x_a, and changing variables from g_a to

$$m_a \equiv -\beta^{-1} A^{-1}_{ab} g_b^*(\{g_a\}), \tag{A8}$$

one eventually obtains the expression [Eq. (53)]

$$e^{n\beta\phi_m} = 2^{nm} \int_{-i\infty}^{i\infty} \prod_a \frac{dx_a}{2\pi i} \int_{-1}^{1} \prod_a \frac{dm_a}{(1-m_a^2)} \times \exp\left\{ -\sum_a \left[x_a \tanh^{-1} m_a + \frac{m}{2} \log(1-m_a^2) \right] + \beta^2 \sum_{ab} \frac{1}{2} x_a Q_{ab} x_b + \frac{1}{2} m_a C_{ab} m_b + m_a A_{ab} x_b \right\}. \tag{A9}$$

Note that the sum over (a,b) also includes diagonal terms with $a=b$. The Gaussian integral over x_a can be carried out in principle. For the annealed solution, this leads to an expression as in Eq. (64) upon a change of variables $m_a = \tanh(\beta h_a)$.

2. Cumulant expansion in $g_a^\pm$

Alternatively, we may proceed from Eq. (A1) by changing integration variables to

$$l_a \equiv \frac{g_a^+ + g_a^-}{2}, \tag{A10}$$

$$h_a \equiv g_a + \frac{g_a^+ + g_a^-}{2}, \tag{A11}$$

$$z_a \equiv K(g_a^+ - g_a^-). \tag{A12}$$

The fields acting on the spins in Eq. (A1) turn into $g_a + g_a^\pm = h_a \pm z_a/2$ K, while l_a only occurs in the Gaussian weight and can be integrated out. The log-trace term then takes the simpler form

$$e^{n\beta\phi_m} = \frac{1}{\sqrt{\det \mathcal{M}}} \int \prod_{a=1}^{n} \frac{dh_a dz_a}{2\pi} \times \sum_{s_a^{i_\pm}} \exp\left[\sum_a \left(\sum_{i_+=1}^{m+K} \beta(h_a + z_a/2K) s_a^{i_+} + \sum_{i_-=m+K+1}^{m} \beta(h_a - z_a/2K) s_a^{i_-} \right) \right] \times \exp\left[-\frac{1}{2} \sum_{ab} \underline{\xi}_a^\dagger \mathcal{M}^{-1}_{ab} \underline{\xi}_b \right], \tag{A13}$$

where $\underline{\xi}_a = (h_a, z_a)$, $a = 1, \ldots, n$. The covariance matrix $\mathcal{M}$ is a $2n \times 2n$ matrix given by

$$\mathcal{M} = \begin{pmatrix} \overbrace{\mathbf{Q}}^{n} & \overbrace{\mathbf{A}}^{n} \\ \mathbf{A} & \mathbf{C} \end{pmatrix}. \tag{A14}$$

i.e., a 2×2 matrix (as in the annealed case where off-diagonal terms vanish) with $n \times n$ matrices as entries. Its determinant is that of $\mathbf{D} \equiv \mathbf{QC} - \mathbf{A}^2$, and its inverse is

$$[\mathcal{M}^{-1}]_{ab} = \sum_c [D^{-1}]_{ac} \begin{pmatrix} C_{cb} & -A_{cb} \\ -A_{cb} & Q_{cb} \end{pmatrix}. \tag{A15}$$

Carrying out the spin sums and taking the limit $K \to \infty$, one finds the contribution to the replicated free energy [Eq. (57)],

$$e^{n\beta\phi_m} = \frac{1}{\sqrt{\det \mathcal{M}}} \int \prod_{a=1}^{n} \frac{dh_a dz_a}{2\pi} \times \exp\left\{ \sum_a [m \log 2\cosh(\beta h_a) + \tanh(\beta h_a)\beta z_a] \right\} \times \exp\left[-\frac{1}{2} \sum_{ab} \underline{\xi}_a^\dagger \mathcal{M}^{-1}_{ab} \underline{\xi}_b \right]. \tag{A16}$$

The integration over z_a in the annealed version of Eq. (A16) leads again to the expression Eq. (64).

3. Self-consistency equations and equivalence between the two-group and cavity approach

With the help of the measure in Eq. (A13), it is easy to see that every spin $s^{i\sigma}$ under the averages in the two-group self-consistency equations, Eqs. (50)–(52), is replaced by $\tanh[\beta(h_a + \sigma z_a/2K)]$. In particular, in the limit $K \to \infty$ we have

$$s_a^{i\sigma} \to \tanh[\beta(h_a + \sigma z_a/2K)] \xrightarrow{K\to\infty} \tanh(\beta h_a) \equiv \tilde{m}_a, \tag{A17}$$

$$K(s_a^{i_+} - s_a^{j_-}) \to K\{\tanh[\beta(h_a + z_a/2K)] - \tanh[\beta(h_a - z_a/2K\,1\,)]\} \xrightarrow{K\to\infty} \frac{\beta z_a}{\cosh^2(\beta h_a)} \equiv \delta m_a, \tag{A18}$$

where we introduced the variables $\tilde{m}_a$ and δm_a. This leads to the self-consistency equations (61)–(64), and establishes the equivalence with the expressions from the cavity approach, Eqs. (111)–(113), on the annealed level.

APPENDIX B: GENERALIZATION OF WARD IDENTITIES

In this appendix, we derive generalizations of the BRST Ward identities, reproducing the saddle-point equations for marginal states. We consider sums over TAP states such as in Eq. (86). We can assume the states to be minima (due to regularization), which we regard as functions of external fields $\{h_k\}$, in the sense that $m_\alpha(\{h_k\})$ is the unique solution of $\partial_i F_{\mathrm{TAP}}(\{m_j\})=h_i$ in the vicinity of the unperturbed state $m_\alpha(\{h_i=0\})$.

After replicating the system n times (with $n\to 0$ eventually), we consider the identity

$$\beta\langle m_i^a x_k^b\rangle \mathcal{Z}_J^n = \sum_{\alpha=1}^{\mathcal{N}_{\mathrm{sol}}} \frac{\partial}{\partial h_k^b}\left\{ m_{i,\alpha}^a \exp \sum_{c=1}^{n} [\lambda_X X(\{m_{j,\alpha}^c\}) - \beta m F_{\mathrm{TAP}}(\{m_{j,\alpha}^c\})]\right\}, \tag{B1}$$

where the limit $\lambda_X\to 0$ will be taken at the end. The function Z_J is the TAP partition sum, which can be expressed both as the functional integral Eq. (76), including the regularizing term $\lambda_X X(\{m_i\})$ in the exponential, or as the sum over TAP states Eq. (86). The average on the left-hand side is taken over the measure defined by the action Eq. (76). Evaluating the right-hand side, we find

$$\begin{aligned}\beta\langle m_i^a x_k^b\rangle &= \frac{1}{\mathcal{Z}_J^n}\sum_{\alpha=1}^{\mathcal{N}_{\mathrm{sol}}}\left(\frac{\partial m_{i,\alpha}^a}{\partial h_k^b} + m_{i,\alpha}^a \frac{\partial \lambda_X X}{\partial h_k^b} - \beta m m_{i,\alpha}^a \frac{\partial F_{\mathrm{TAP}}}{\partial h_k^b}\right) \times \exp\left\{\sum_{c=1}^{n}[-\beta m F_{\mathrm{TAP}}(\{m_{j,\alpha}^c\}) + \lambda_X X(\{m_{j,\alpha}^c\})]\right\} \\ &= \left\langle \delta_{ab}\chi_{ik}^{ab} + m_{i,\alpha}^a \frac{\partial \lambda_X X}{\partial h_k^b}\right\rangle,\end{aligned} \tag{B2}$$

where the stationarity of F_{TAP} was used. Using the definition of the soft mode, Eq. (89), and summing Eq. (B2) over $i=k$, we establish the generalization of the first Ward identity [Eq. (90)],

$$\beta\langle m^a x^b\rangle = \delta_{ab}\langle \psi_a \overline{\psi_b}\rangle + \beta\langle m_a \delta m_b\rangle. \tag{B3}$$

The second identity follows similarly from the relation

$$\begin{aligned}\beta^2\langle x_a x_b\rangle \mathcal{Z}_J^n &= \frac{1}{N}\sum_{i,\alpha}\frac{\partial^2}{\partial h_i^a \partial h_i^b} \times \exp\left\{\sum_{c=1}^{n}[-\beta m F_{\mathrm{TAP}}(\{m_{j,\alpha}^c\}) + \lambda_X X(\{m_{j,\alpha}^c\})]\right\} \\ &= \frac{1}{N}\sum_{i,\alpha}\left(-m\beta\chi_{ii}\delta_{ab} + \beta^2 \delta m_{i,\alpha}^a \delta m_{i,\alpha}^b + \frac{\partial^2 \lambda_X X}{\partial h_i^a \partial h_i^b}\right) \times \exp\left\{\sum_{c=1}^{n}[-\beta m F_{\mathrm{TAP}}(\{m_{j,\alpha}^c\}) + \lambda_X X(\{m_{j,\alpha}^c\})]\right\} \\ &= -\beta^2 m \delta_{ab}\langle \psi_a \psi_b\rangle + \beta^2\langle \delta m_a \delta m_b\rangle + \frac{1}{N}\sum_i \left\langle \frac{\partial^2 \lambda_X X}{\partial h_i^a \partial h_i^b}\right\rangle.\end{aligned} \tag{B4}$$

The very last term can be rewritten as

$$\begin{aligned}\frac{1}{N}\sum_i \frac{\partial^2 \lambda_X X}{\partial h_i^a \partial h_i^b} &= \frac{\lambda_X}{N}\left[\delta_{ab}\sum_{i,k}\frac{\partial X}{\partial m_k^a}\frac{\partial^2 m_k^a}{(\partial h_i^a)^2} + \sum_{i,k,l}\chi_{li}^a \frac{\partial^2 X}{\partial m_k^a \partial m_l^a}\chi_{ki}^a\right] \\ &= -2\beta\delta_{ab}\frac{1}{N}\sum_{i,k}\lambda_X \frac{\partial X}{\partial m_k^a}\frac{\partial m_k^a}{\partial h_i^a} m_i^a + O\left(\frac{1}{\lambda_X N}, \lambda_X\right) \\ &= -2\beta^2 \delta_{ab}\frac{1}{N}\sum_i \langle m_i^a \delta m_i^a\rangle + O\left(\frac{1}{\lambda_X N}, \lambda_X\right),\end{aligned} \tag{B5}$$

where we have used the identity

$$\frac{\partial^2 m_k}{\partial h_i^2} = \frac{\partial^2 m_i}{\partial h_i \partial h_k} = \beta\frac{\partial(1-m_i^2)}{\partial h_k} = -2\beta m_i \frac{\partial m_k}{\partial h_i} \tag{B6}$$

and the definition Eq. (89) for the soft mode $\delta\vec{m}$. Furthermore, a spectral decomposition of the susceptibility matrix χ (as in Sec. VIII) shows that the last term of Eq. (B5), $N^{-1}\mathrm{Tr}[\chi\partial^2 X\chi]$, consists of a contribution $(\lambda_X N)^{-1}$ due to the soft mode, and a further contribution $\sim\lambda_X$. Both are negligible, as the limit $N\to\infty$ is taken before $\lambda_X\to 0$.

Finally, inserting Eq. (B5) into Eq. (B4), we find Eq. (92).

APPENDIX C: EFFECT OF REGULARIZATION IN THE PRESENCE OF MANY STATES

By introducing a regularizing weight factor $\exp(\lambda_X X)$ into the cavity formalism and the explicit sum over TAP solutions, we shift the weight slightly toward stable states. The following constructive description of the selected set of states may be helpful to understand this formal trick, even though we do not have a rigorous proof for its correctness.

The selected states will be slightly stable. However, under a small rise of the temperature they would become marginal again. We assume that the states are selected by a weight function X affecting all states in the same (self-averaging) manner [with X a symmetric function of the m_i, e.g., $X=\Sigma_i \psi(m_i)$ with arbitrary ψ]. Then most of the selected states will become marginal at the same slightly higher temperature $T_\lambda = T + \delta T_\lambda$. Likewise, almost all states that are marginal at

T_λ will adiabatically evolve into typical states selected by λX at the lower temperature T. We may thus expect the number of selected states to be given by $\exp[N\Sigma(T_\lambda, f_\lambda)]$, where f_λ is unequivocally determined by the considered free-energy density f at T. In general, one expects the complexity to decrease proportionally to the increment of temperature,

$$\Sigma(T,f) - \Sigma(T_\lambda, f_\lambda) \sim \delta T_\lambda. \tag{C1}$$

The properties of the selected states at T follow from perturbation theory around the marginal situation at (T_λ, f_λ). One finds that the temperature shift δT_λ induces a change of local magnetizations of order $\Delta m_i \sim (\delta T_\lambda)^{1/2} \zeta_i$ along the soft mode $\{\zeta_i\}$, the square root reflecting the anomalous response of the soft mode. Accordingly, the value of the regularizer changes by $\delta X = \nabla X \cdot \delta \vec{m} \sim N(\delta T_\lambda)^{1/2}$, and the soft mode acquires a finite susceptibility $\chi_{\text{soft}}^{-1} = \Sigma_{ij} \zeta_i \chi_{ij}^{-1} \zeta_j \sim (\delta T_\lambda)^{1/2}$.

The weighting procedure favors the states optimizing $\exp[N\Sigma(T_\lambda, f_\lambda) - \lambda \delta X]$, which yields $(\delta T_\lambda)^{1/2} \sim \lambda$, very similarly to the mechanism in the toy model of Sec. III. Consequently, the susceptibility along the soft mode scales as

$$\chi_{\text{soft}} \sim \frac{1}{(\delta T_\lambda)^{1/2}} \sim \frac{1}{\lambda}. \tag{C2}$$

Note that all proportionality constants in the above arguments are self-averaging constants. This implies in particular that $g \equiv (\chi_{\text{soft}} \lambda)^{-1}$ is independent of the specific metastable state.

[1] G. Parisi, in *Lecture Notes of the Les Houches Summer School*, edited by A. Bovier, F. Dunlop, A. Van Enter, F. Den Hollander, and J. Dalibard (Elsevier, Amsterdam, 2006).
[2] A. Crisanti, L. Leuzzi, and T. Rizzo, Eur. Phys. J. B **36**, 129 (2003).
[3] T. Aspelmeier, A. J. Bray, and M. A. Moore, Phys. Rev. Lett. **92**, 087203 (2004).
[4] A. Cavagna, I. Giardina, and G. Parisi, Phys. Rev. Lett. **92**, 120603 (2004).
[5] G. Parisi and T. Rizzo, J. Phys. A **37**, 7979 (2004).
[6] A. Crisanti, L. Leuzzi, and T. Rizzo, Phys. Rev. B **71**, 094202 (2005).
[7] L. Leuzzi, Prog. Theor. Phys. **157**, 94 (2005).
[8] A. Annibale, G. Gualdi, and A. Cavagna, J. Phys. A **37**, 11311 (2004).
[9] A. Crisanti and H. J. Sommers, Z. Phys. B: Condens. Matter **87**, 341 (1992).
[10] A. Crisanti and H. J. Sommers, Z. Phys. B: Condens. Matter **92**, 257 (1993).
[11] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).
[12] A. J. Bray and M. A. Moore, J. Phys. C **12**, L441 (1979).
[13] In the case of the SK ground state, it is known that the same methods as used for genuinely stable minima yield a correct description. It is unknown, however, whether this holds in general when a multitude of soft modes is present; see also the discussion in Sec. VIII D.
[14] A. J. Bray and M. A. Moore, Phys. Rev. Lett. **41**, 1068 (1978).
[15] G. Parisi and M. Potters, Europhys. Lett. **32**, 13 (1995).
[16] G. Parisi, J. Phys. A **13**, L115 (1980).
[17] D. J. Thouless, P. W. Anderson, and R. G. Palmer, Philos. Mag. **35**, 593 (1977).
[18] A. J. Bray and M. A. Moore, J. Phys. C **13**, L469 (1980).
[19] M. Mezárd, G. Parisi, and M. A. Virasoro, Europhys. Lett. **1**, 77 (1986).
[20] A. Cavagna, I. Giardina, and G. Parisi, Phys. Rev. B **71**, 024422 (2004).
[21] T. Rizzo, J. Phys. A **38**, 3287 (2005).
[22] C. Becchi, A. Stora, and R. Rouet, Commun. Math. Phys. **42**, 127 (1975); I. V. Tyutin (unpublished).
[23] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Clarendon Press, Oxford, 1989).
[24] J. Kurchan, J. Phys. A **24**, 4969 (1991).
[25] G. Parisi, J. Phys. A **13**, 1101 (1980).
[26] R. Monasson, Phys. Rev. Lett. **75**, 2847 (1995).
[27] A. Crisanti, L. Leuzzi, G. Parisi, and T. Rizzo, Phys. Rev. B **68**, 174401 (2003).
[28] G. Parisi, in *Proceedings of Les Houches 1982*, edited by Zuber and Stora (North Holland, Amsterdam, 1984).
[29] A. Crisanti, L. Leuzzi, G. Parisi, and T. Rizzo, Phys. Rev. B **70**, 064423 (2004).
[30] S. F. Edwards and P. W. Anderson, J. Phys. F: Met. Phys. **5**, 965 (1975).
[31] A. J. Bray and M. A. Moore, J. Phys. C **13**, L907 (1980).
[32] V. Dotsenko and M. Mézard, J. Phys. A **30**, 3363 (1997).
[33] V. Dotsenko and G. Parisi, J. Phys. A **25**, 3143 (1992).
[34] V. Dotsenko, J. Stat. Mech.: Theory Exp. 2006, P06003 (2006).
[35] However, we note that the case of the random field Ising model is actually more subtle. A thorough analysis shows that the description of the leading Griffith singularities requires a further generalization of the two-group ansatz.[62]
[36] Instead, saddles with $k_1 = 0$ describe the probability, P_{marg}, to find a rare random field giving rise to a secondary minimum, $P_{\text{marg}} \approx \exp[-\beta F_{k_1=0, k_2 \to \infty}]$.
[37] G. Parisi and T. Rizzo, Phys. Rev. B **72**, 184431 (2005).
[38] G. Parisi and M. Potters, J. Phys. A **28**, 5267 (1995).
[39] T. Aspelmeier, R. A. Blythe, A. J. Bray, and M. A. Moore, cond-mat/0602639 (unpublished).
[40] A. J. Bray and M. A. Moore, J. Phys. C **14**, 2629 (1981).
[41] A. Annibale, A. Cavagna, I. Giardina, G. Parisi, and E. Trevigne, J. Phys. A **36**, 10937 (2003).
[42] The order parameters introduced in Ref. 41 are related to our notation by $W_{ab} \equiv \langle x_a m_b \rangle = A_{ab} + \delta_{ab}(1 - Q_{aa})$ and $\lambda_{ab} = \beta^2/2(C_{ab} + 2mA_{ab} + m^2 Q_{ab})$.
[43] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).
[44] B. Duplantier, J. Phys. A **14**, 283 (1981).
[45] H. J. Sommers and W. Dupont, J. Phys. C **17**, 5785 (1984).
[46] T. Plefka, J. Phys. A **15**, 1971 (1982).
[47] T. Plefka, Europhys. Lett. **58**, 892 (2002).
[48] A. Crisanti, L. Leuzzi, G. Parisi, and T. Rizzo, Phys. Rev. Lett. **92**, 127203 (2004).
[49] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and*

Beyond, Lecture Notes in Physics Vol. 9 (World Scientific, Teaneck, NJ, 1987).
[50] M. Metha, *Random Matrices* (Academic, New York, 1967).
[51] For every finite ϵ, there are $O(N)$ eigenvalues $\lambda < \epsilon$.
[52] The finite size scaling of the soft eigenvalue of the Hessian is believed to behave as $\lambda_{\text{soft}} \sim N^{-1/2}$. The contribution of the marginal mode to the spin glass susceptibility, $(N\lambda_{\text{soft}}^2)^{-1}$, thus tends to a nondivergent constant.
[53] L. Leuzzi, M. Müller, and A. Crisanti (unpublished).
[54] C. De Dominicis, M. Gabay, T. Garel, and H. Orland, J. Phys. (Paris) **41**, 923 (1980).
[55] F. Tanaka and S. F. Edwards, J. Phys. F: Met. Phys. **10**, 2769 (1980).
[56] S. Pankov, Phys. Rev. Lett. **96**, 197204 (2006).
[57] P. W. Anderson, *Ill-condensed Matter*, Les Houches 1978, Session XXXI, edited by R. Balian, R. Maynard, and G. Toulouse (North Holland, Amsterdam, 1979).
[58] A. L. Efros and B. I. Shklovskii, J. Phys. C **8**, L49 (1975).
[59] M. Müller and L. B. Ioffe, Phys. Rev. Lett. **93**, 256403 (2004).
[60] M. Müller and S. Pankov (unpublished).
[61] F. Pazmandi, G. Zarand, and G. T. Zimanyi, Phys. Rev. Lett. **83**, 1034 (1999).
[62] M. Müller and A. Silva, Phys. Rev. Lett. **96**, 117202 (2006); (unpublished).
[63] M. Mézard and G. Parisi, J. Phys. I **1**, 809 (1991).