Random-Field Ising like effective theory of the glass transition: I Mean-Field Models

Giulio Biroli[*]

IPhT, CEA/DSM-CNRS/URA 2306, CEA Saclay, F-91191 Gif-sur-Yvette Cedex, France
Laboratoire de Physique Statistique, École Normale Supérieure, CNRS, France

Chiara Cammarota[†]

Department of Mathematics, King's College London, Strand, London WC2R 2LS, UK

Gilles Tarjus[‡] and Marco Tarzia[§]

LPTMC, CNRS-UMR 7600, Sorbonne Université, 4 Pl. Jussieu, F-75005 Paris, France
(Dated: July 18, 2018)

In this paper and in the companion one[1] we address the problem of identifying the effective theory that describes the statistics of the fluctuations of what is thought to be the relevant order parameter for glassy systems—the overlap field with an equilibrium reference configuration—close to the putative thermodynamic glass transition. Our starting point is the mean-field theory of glass formation which relies on the existence of a complex free-energy landscape with a multitude of metastable states. In this paper, we focus on archetypal mean-field models possessing this type of free-energy landscape and set up the framework to determine the exact effective theory. We show that the effective theory at the mean-field level is generically of the random-field + random-bond Ising type. We also discuss what are the main issues concerning the extension of our result to finite-dimensional systems. This extension is addressed in detail in the companion paper.

PACS numbers:

Contents

I. INTRODUCTION

Developing a proper treatment of nonperturbative fluctuations is one of the most difficult methodological issues one can encounter when facing a physical problem. In the past, this difficulty has been circumvented mainly in two ways: either by making use of very clever assumptions on the physics of the problem at hand or by finding a mapping that transforms the original strong-coupling model with nonperturbative fluctuations (that one cannot solve) in a weak-coupling model with perturbative fluctuations (that one can treat easily). Examples of the first case are provided by several variational wave-functions discovered along the years in condensed-matter physics (*e.g.*, the Laughlin wave-function for the fractional quantum Hall effect[2]). Examples of the latter are provided by the study of the low-temperature behavior of systems that are characterized by dilute nonperturbative excitations, such as the XY model that can be mapped onto a dilute Coulomb gas of vortices.[3] Actually, in all of these situations one solves the problem by *avoiding* to directly tackle nonperturbative fluctuations and by instead finding a suitable short-cut. There are however cases, as for example the glass transition of supercooled liquids,[4–9] where these two approaches seem to fail: No dual weak-coupling system can be identified, no clear-cut assumptions to simplify the problem can be made. Two theoretical approaches of glass formation, the dynamical-facilitation theory[10] and the approach based on geometrical frustration and avoided criticality,[11] provide valuable attempts to identify and address the source of relevant nonperturbative fluctuations in glass-forming liquids, but they remain at present not fully satisfactory.

Given this situation, an alternative route is to start from an established mean-field description and to incorporate, up to some finite length scale, the fluctuations of the identified order parameter in an *effective theory*. The hope is to derive an effective theory that (i) encompasses the main physical ingredients while leaving out inessential ones and (ii) contains nonperturbative fluctuations that can be handled in a more tractable manner than in the original problem. The aim of this work is to perform the first steps toward such an effective theory of the glass transition.

Our starting point is the mean-field theory of glasses[12,13] which has recently gained momentum through the solution of the hard-sphere glass in infinite dimensions.[14,15] The associated scenario relies on the existence, below a critical temperature associated with a dynamical transition, of a complex free-energy landscape with a multitude of metastable states that is characterized by an extensive configurational entropy. An ideal thermodynamic glass transition, known as a random first-order transition (RFOT), takes place when the configurational entropy becomes subextensive.[12] The relevant order parameter is then provided by the similarity or overlap between equilibrium liquid configurations.[16–18] However, this mean-field scenario appears fragile to the introduction of fluctuations,[19,20] and the very notion of metastable states is well-defined only when fluctuations are absent, as in a mean-field approximation, or suppressed, as in a small system.[21]

Developing an effective theory of glass-forming systems directly formulated in terms of what is thought to be the physically relevant local order parameter, the overlap with an equilibrium reference configuration, seems a valuable task for several conceptual and technical reasons:

(a) It provides a more intuitive description of the glass transition and, most importantly, allows one to circumvent the explicit description in terms of metastable states.

(b) The problem of handling in a fully satisfactory way the large-scale physics described by the replica field theory suggested by mean-field models remains very challenging, despite some recent theoretical progress (*e.g.*, instantons calculations[22], Kac analysis[23], and real space RG approaches[24–26]). This is partly due to the complicated replica matrix structure of the overlap fields. Focusing only on some of the overlaps, namely those involving the equilibrium reference configuration, while integrating out all the others naturally leads to a scalar field theory in the presence of quenched disorder. The latter is *a priori* much easier to handle than the original theory and can be studied by using powerful tools of statistical physics (*e.g.*, large-scale numerical simulations, nonperturbative functional renormalization group, etc.).

(c) Promoting the order parameter to a fully fluctuating field is a way to study fluctuations and correlations beyond mean-field theory. This provides a proper description of all large-scale and nonperturbative fluctuations and thereby allows one to assess the nature of the critical points and identify the mechanisms that could possibly destroy or alter the glass transition (RFOT) in finite dimensions.[27]

This program has already been partly achieved. A mapping to an effective theory akin to the Ising model in a random field (RFIM)[28] has been derived near (but below) the dynamical transition of the mean-field theory[29] as well as near the critical points appearing in an extended phase diagram in the presence of additional sources or pinning

fields.[30–32] In this work we focus on the more challenging problem of establishing an effective theory in the vicinity of the putative thermodynamic glass transition.

In this first paper, we consider two archetypal mean-field models for glass formation, *i.e.*, the Random Energy Model[33] (REM) and its Kac-like generalization to a finite number (2^M) of states per site[34] (which we call in the following the 2^M-KREM) on a fully connected lattice. In the REM case we show that the statistics of the thermal fluctuations of the global overlap with an equilibrium configuration is *exactly* described by an Ising variable $\sigma = \pm 1$ subjected to a random field, *i.e* a 0-dimensional RFIM,

$$\beta \mathcal{H}_{\text{eff}} = \mathcal{S}_0 - (H + \delta h)\sigma \,, \tag{1}$$

where the two values of σ correspond to one or zero overlap, H is a temperature-dependent uniform field of order N that corresponds to the configurational entropy and vanishes at the RFOT temperature T_K, and δh a random field of zero mean and fluctuations of order $\sqrt{N}$, where $N \to \infty$ is the logarithm of the number of states. A similar (but richer) result can be obtained for the fully connected 2^M-KREM, for which one can show that the effective theory for the overlap profile near T_K corresponds to a fully connected random-bond+random-field Ising model with multi-body interactions and higher order random terms:

$$\beta \mathcal{H}_{\text{eff}} = \mathcal{S}_0 - \sum_i (H + \delta h_i)\sigma^i - \frac{1}{2}\sum_{i\neq j}\Big(\frac{J_2}{N} + \frac{\delta J_{2,ij}}{\sqrt{N}}\Big)\sigma^i\sigma^j - \frac{J_3}{3!N^2}\sum_{i,j,k\neq}\sigma^i\sigma^j\sigma^k - \frac{J_4}{4!N^3}\sum_{i,j,k,l\neq}\sigma^i\sigma^j\sigma^k\sigma^l - \cdots \,. \tag{2}$$

The 2 possible states $\sigma^i = \pm 1$ correspond to a low- and a high-overlap with a reference equilibrium configuration on a given site i; the field H plays the role of the configurational entropy, vanishing at T_K, δJ_{ij} and δh_i are random variables with zero mean; $J_2 > 0$ is a ferromagnetic coupling and J_3 and J_4 are 3- and 4-body interactions (whose sign can depend on the parameters of the original microscopic model, *e.g.*, the number of states). The ellipses denote multi-body interactions beyond the 4-body one and higher-order random terms which have been omitted. The coupling constants and the variance of the random terms can be, at least in principle, computed exactly.

The REM and the fully connected 2^M-KREM can of course be exactly solved, with no need to go through a mapping onto an effective Hamiltonian, but the present treatment illustrates how an effective Ising theory with quenched disorder emerges and this sets the stage for studying finite-dimensional glass-formers. The latter, including glass-forming liquids, will be the focus of the companion paper[1]. In this case additional approximations are required but the output will again be a description of the glass transition in terms of an Ising model in an external field with random-field and random-bond disorder and long-range competing multi-body interactions.

The rest of the paper is organized as follows. In Sec. II we review the situations where the random-field Ising model in one form or another appears in the theory of glass-forming liquids and we provide a general intuitive argument for why this is so. The following section, Sec. III, is devoted to the derivation of the effective theory for the overlap with a reference equilibrium configuration in two mean-field models of structural glasses, the REM and its generalization to a finite number of states, the 2^M-KREM. We focus on the region around the putative thermodynamic glass transition (RFOT). In Sec. IV we first illustrate the difficulties that one encounters when trying to generalize the procedure developed for the mean-field models to finite-dimensional glass-forming systems. (This generalization will be the topic of the companion paper.[1]) We next discuss some of the properties of the quenched disorder appearing in the effective theory. Finally, some concluding remarks are given in Sec. V. Most of the technical details of the calculations are presented in two Appendices.

II.   RFIM-LIKE CRITICALITY IN GLASS-FORMING LIQUIDS

A.   Known results

This is not the first time that an Ising model in a random field appears in the context of supercooled liquids. Actually, the idea of mapping the glass transition onto a magnetic system with quenched disorder was put forward and analyzed for the first time in Ref. [35]. Several recent analytical and numerical investigations strongly support the idea that the effective theory which describes the thermal fluctuations of the overlap with an equilibrium configuration in glassy systems is provided by an Ising model in the presence of quenched disorder. Below we present a list of the main known results, which are pictorially summarized in Fig. 1.

From the analysis of the perturbation theory of the replica field description it was first shown in Ref. [29] that the critical fluctuations of the overlap close to the dynamical (mode-coupling-like) transition, in the so-called β-regime just below the transition, are in the same universality class as those found at the spinodal point of the (standard, short-range) RFIM. Both types of singularities, the dynamical transition and the spinodal, can only be present when

activated events such as nucleation are not taken into account. This connection was further examined in Ref. [36], where the spinodal of the RFIM was studied at zero temperature, thereby eliminating all thermal fluctuations.

In the past few years some effort has been devoted to analyze the universality class of the critical points that can be induced in glassy systems by the presence of suitable constraints. The first such case that was studied corresponds to introducing an additional attractive coupling ϵ to a reference equilibrium configuration of the system which in effect acts as a source linearly coupled to the overlap with this reference configuration. Within the mean-field theory, the glass transition found at $T = T_K$ and $\varepsilon = 0$ transforms into a line of first-order transition in the (T-ϵ) plane, which ends in a critical point at $T = T_c$ and $\varepsilon = \varepsilon_c > 0$, as illustrated in left panel of Fig. 1. This feature is a key prediction of the mean-field/RFOT theory. It is found in mean-field disordered spin models,[18] and evidence for it has been observed in computer simulations of 3-dimensional atomistic models.[17,18,38–42] Through a thorough analysis of the soft modes emerging at the terminal critical point in the replica field theory and of the resulting properties of the perturbation theory,[30] it was established that this critical point belongs to the universality class of the standard RFIM in finite dimensions. We also found the same result independently, by using an approach based on an expansion in increasing number of free replica sums.[32] Of course this is valid if the transition is not destroyed by the disorder, but numerical indications that the RFIM critical behavior can indeed be found in a 3-dimensional glass-forming liquid model has been recently obtained.[40] (Finally, a further link between the physics of RFIM and supercooled liquids comes from the study of fluctuations of amorphous interfaces in 3-dimensional liquid models[43] whose statistical properties have also been investigated through an expansion in free replica sums.[44])

Surprisingly enough, the relevance of this RFIM-like criticality to glass-forming systems also appears in the context of plaquette spin models, usually taken as an illustration of the dynamical-facilitation theory of glass formation.[45,46] In a series of papers[47–49] Garrahan and coworkers analyzed plaquette spin models in dimensions $d = 2$ (the "triangular plaquette model") and $d = 3$ (the "square pyramid model"), focusing on the thermodynamic behavior in the presence of an attractive coupling ε. For the $3d$ square pyramid model the authors presented strong numerical evidence in favor of the existence of a transition line in the (T-ε) plane, terminating in a critical point whose universal properties are those of the $3d$ RFIM. On the other hand, no such transition was found in $d = 2$, in agreement with the fact that the lower critical dimension of the RFIM is precisely $d = 2$. The role of short- versus long-range fluctuations of the overlap was also studied in these models by means of Bethe-lattice calculations.[50]

Another procedure to constrain the system toward a reference configuration (referred to as the "pinned particles" method) is to freeze the positions of a randomly chosen fraction c of the particles to the values they have in a given equilibrium configuration.[51] According to the mean-field/RFOT theory, the constraint induces a line $c_K(T)$ of thermodynamic glass transition (RFOT) and a line of dynamical (mode-coupling-like) glass transition $c_d(T)$ in the (T-c) plane, as illustrated in the right panel of Fig. 1. Differently from the case of the ε-coupling where a nonzero ϵ transforms the thermodynamic (RFOT) glass transition into a conventional first-order transition (albeit in the presence of a random field[32]), in the pinned-particle case the thermodynamic glass transition keeps its glassy RFOT character all along the line. The RFOT line and the line of dynamical transitions merge in a critical endpoint. This scenario is realized in mean-field models of glass-forming systems,[51–53] and its relevance for finite dimensions is supported by calculations based on a Migdal-Kadanoff real-space renormalization group (RG)[51] and numerical results.[37,38,54] In Ref. [31] it was established that, just like for the case of the ϵ-coupling, the critical endpoint in the (T-c) plane is in the same universality class as the critical point of the RFIM, also in agreement with the real-space RG results of Ref. [51]. Moreover, the mode-coupling theory (MCT) predicts several kinds of critical dynamical behavior for randomly pinned systems.[55–57] Along the dynamical transition line, $c_d(T)$, the transition remains of A_2 type in the MCT terminology,[7]) until the terminal point is reached, where the singularity becomes of A_3 type. The dynamical behavior at the A_3 critical endpoint displays activated dynamical scaling, a characteristic property of the critical dynamical behavior of the RFIM.[58]

Finally, the dynamics of the kinetically constrained Fredrickson-Andersen model[59] was analyzed on a Bethe lattice, showing the presence of a dynamical transition whose finite size-scaling is consistent with that of the RFIM.[60] The Fredrickson-Andersen model on a Bethe lattice was also studied in the presence of a random pinning.[61] Strong evidence was then found for the existence of a line $c_d(T)$ of dynamical glass transitions with the characteristic A_2 MCT singularity ending in a critical point with an A_3 MCT singularity related to the dynamical behavior of the RFIM (see above).

The mapping of the properties of glass-forming systems to those of the RFIM obtained so far has two main limitations. The first one is that it concerns the fluctuations of the overlap field but not directly the dynamical behavior, *i.e.* there is no direct connection between the *dynamics* of the RFIM and that of supercooled liquids. Recently, Rizzo[62] went beyond the static analysis to obtain a dynamical stochastic equation, called Stochastic-Beta-Relaxation equation, involving the effect a random field. This provides a theory of dynamical fluctuations in finite dimensions close to the avoided dynamical transition (MCT crossover) but is far from providing a full dynamical description of the approach to the physical glass transition. The second limitation is that the mapping, summarized in Fig. 1, has so far left aside the more interesting, and more challenging, case that corresponds to the situation in
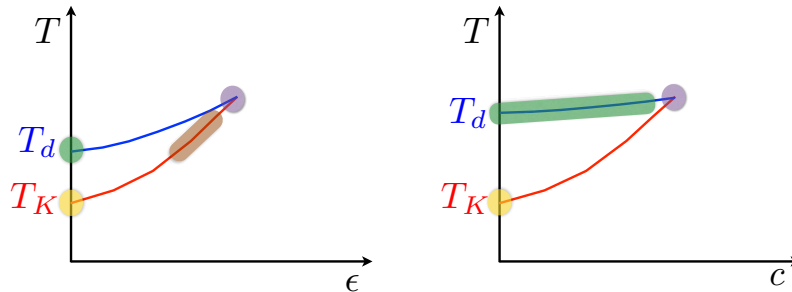
FIG. 1: Sketch of the phase diagram of constrained glass-forming systems predicted by the mean-field/RFOT theory in the $(T\text{-}\epsilon)$ plane (left) and in the $(T\text{-}c)$ plane (right), showing the regions where RFIM-like universality classes of the critical points in finite dimensions (see text for a more detailed description). The dynamical (MCT) glass transitions are in the same universality class as the spinodal point of the RFIM[29] (green). The terminal critical points in the $(T\text{-}\epsilon)$ plane and in the $(T\text{-}c)$ plane belong to the universality class of the RFIM[30,32,51] (violet). The first-order transition line in the $(T\text{-}\epsilon)$ plane is described by a first-order transition in the presence of a random field[32] (brown). The situation in the absence of coupling ($\epsilon = 0$, $c = 0$) and close to the putative RFOT T_K (yellow) has been investigated for the first time numerically in Ref. [35] and is the focus of the present paper. Note that one expects that the putative line of RFOT in the $(T\text{-}c)$ plane is described by a similar effective theory as that near T_K.

the absence of coupling ($\varepsilon = 0$, $c = 0$) and close to T_K, where a thermodynamic glass transition of RFOT type is predicted at the mean-field level. Whereas we do not address the first point in this work, we do consider the second issue of the mapping in the vicinity of the putative thermodynamic glass transition.

B. Self-induced disorder in the overlap field theory

An intuitive argument explaining why the effective theory describing the local fluctuations of the overlap order parameter in glass-forming systems is in the class of an Ising model in a random field relies on the concept of "self-induced disorder".[29] In fact, the object which plays the role of a (large deviation) Landau-like functional for the chosen order parameter is the average free-energy cost that is necessary to keep the system at an overlap $p(\mathbf{x})$ with a reference equilibrium configuration $\mathcal{C}_{\text{eq}}$. To be more concrete, take a glass-former described by configurations $\mathcal{C}$ and a Hamiltonian $\mathcal{H}[\mathcal{C}]$. Consider then a reference equilibrium configuration $\mathcal{C}_{\text{eq}}$, which is taken from the equilibrium Gibbs distribution, $\mathcal{P}(\mathcal{C}_{\text{eq}}) = e^{-\beta\mathcal{H}(\mathcal{C}_{\text{eq}})}/Z$, and denote the overlap at point $\mathbf{x}$ between a configuration $\mathcal{C}$ and the reference configuration as $Q_{\mathbf{x}}(\mathcal{C}, \mathcal{C}_{\text{eq}}) = \delta_{\mathcal{C}, \mathcal{C}_{\text{eq}}}$. (For a liquid formed by N particles one needs to introduce a smoothing function $f(y)$ with a short range of the order of the cage size corresponding to the typical extent of the vibrational motions, i.e., $Q_{\mathbf{x}}[\hat{\rho}(\mathcal{C}), \hat{\rho}(\mathcal{C}_{\text{eq}})] = \int d\mathbf{y} f(y)[\hat{\rho}(\mathbf{x} + \mathbf{y}|\mathcal{C})\hat{\rho}(\mathbf{x} - \mathbf{y}|\mathcal{C}_{\text{eq}}) - \rho^2]$, where $\hat{\rho}(\mathbf{x}|\mathcal{C})$ is the microscopic density at point $\mathbf{x}$ for a configuration $\mathcal{C}$ of the liquid and $\rho = N/V$ is the average density: see the companion paper.[1])

One can now define an overlap field $p(\mathbf{x})$ and introduce an effective Hamiltonian or action for this field, which is the large-deviation functional describing the probability to observe a certain profile of the overlap field:

$$\mathcal{S}[p|\mathcal{C}_{\text{eq}}] \equiv -\ln\left(\mathcal{P}[p|\mathcal{C}_{\text{eq}}]\right) = -\ln\left[\frac{1}{Z}\sum_{\mathcal{C}} e^{-\beta\mathcal{H}(\mathcal{C})}\,\delta\left[p(\mathbf{x}) - Q_{\mathbf{x}}(\mathcal{C}, \mathcal{C}_{\text{eq}})\right]\right]. \tag{3}$$

For a uniform overlap $p(\mathbf{x}) = p$ and in the mean-field limit, the action $\overline{\mathcal{S}[p|\mathcal{C}_{\text{eq}}]}$, averaged over all different choices of the equilibrium configuration, becomes the Franz-Parisi potential $V(p)$.[16–18] This object encodes in a compact way the properties of the complex free-energy landscape of glassy systems. Between T_d and T_K it exhibits an absolute minimum in $p \simeq 0$ and a secondary minimum in $p_\star$, which corresponds to the overlap for a typical metastable state sampled at equilibrium (also called non-ergodicity parameter and related to the Debye-Waller factor) and whose height difference with the value at the stable minimum corresponds to the configurational entropy s_c. Qualitatively, $V(p)$ exactly behaves as the Landau free-energy of a φ^4 scalar field theory in the presence of a negative external magnetic field H, which describes a first-order transition from a negative to a positive magnetization at $H = 0$. Pushing the analogy with magnetic systems a step further,[35] the overlap order parameter p plays the role of the magnetization m, the configurational entropy s_c is the counterpart of (minus) the external magnetic field H. Furthermore, a surface tension-like term γ, related to the height of the barrier between the two minima and called amorphous surface tension in the context of supercooled liquids, is proportional to the ferromagnetic coupling J. Finally, the thermodynamic
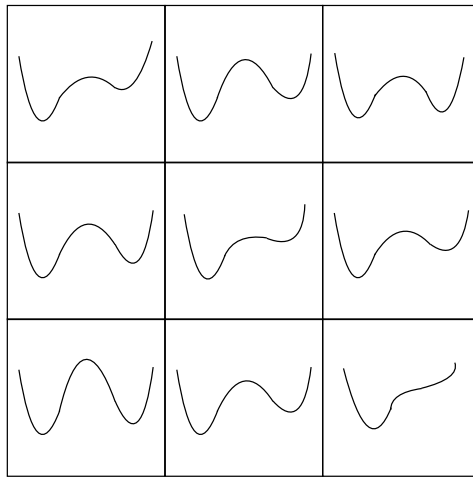
FIG. 2: Sketch of the distribution of local Franz-Parisi potentials when the system is divided in cubic boxes larger than the size of the particles but smaller than the point-to-set correlation length. This illustrates the local fluctuations of the configurational entropy and of the surface tension.

glass transition at T_K, at which the two minima in $p = 0$ and $p = p_\star$ have the same free-energy, corresponds to the first-order transition in $H = 0$, whereas the dynamical glass transition at T_d coincides with the spinodal of the positively magnetized state.

With this in mind, one can now argue that the reference equilibrium configuration acts as a quenched disorder. Indeed, although after averaging over $\mathcal{C}_{\rm eq}$ the global Franz-Parisi potential becomes independent of the reference configuration, its local properties still depend on the choice of $\mathcal{C}_{\rm eq}$ due to the density fluctuations of the reference configuration. Imagine coarse-graining the system on a scale that is larger than the microscopic scale (*i.e.,* the size of the particles or the lattice spacing) but smaller than the point-to-set length (above which metastability and configurational entropy are no longer well defined in finite dimensions[24,63]) by dividing the sample in cubic boxes as sketched in Fig. 2 and computing the Franz-Parisi potential in each of these finite-size boxes considered as independent one from another. One would then observe fluctuations of the shape of the Franz-Parisi potential from one box to another, due to the local density fluctuations of the reference configuration (see Fig. 2 for an illustration). This results in fluctuations of the height of the secondary minimum (*i.e.,* of the configurational entropy s_c akin to a magnetic field) and of the height of the barrier between the two minima (*i.e.,* of the surface tension γ akin to a ferromagnetic coupling) among boxes. At a coarse-grained level, this naturally leads to a description in terms of a scalar φ^4 effective theory with quenched disorder in the form of a random field and a random bond. (Note that the above procedure can also be operationally implemented in computer simulations of glass-forming liquid models: This will be further discussed in the companion paper.[1])

If one now tries to extend these phenomenological arguments to the vicinity of the putative thermodynamic glass (RFOT) transition at T_K, one realizes that the different boxes may become strongly correlated due to the presence of a diverging point-to-set correlation length.[63] Showing that the description based on an effective random-field + random-bond 2-state Ising-like theory is not jeopardised by these long-range correlations is a challenge. The results of the present paper and of its companion one[1] suggest that this description continues to hold, but that the presence of a diverging point-to-set correlation length generically leads to the emergence of additional features such as multi-body interactions.

III. EFFECTIVE THEORY FOR MEAN-FIELD MODELS OF STRUCTURAL GLASSES

In the first part of this paper we focus on two archetypal mean-field models for the glass formation, *i.e.,* the Random Energy Model (REM)[33] and the fully connected version of its Kac-like generalization to a finite number (2^M) of states[34] (called in the following the 2^M-KREM), and we work out, essentially exactly, the effective 2-state random Hamiltonian that describes the fluctuations of the overlap with a reference equilibrium configuration.

A. Replicas and expansion in cumulants

Because of the reference configuration, $\mathcal{C}_{\mathrm{eq}}$, the action $\mathcal{S}[p(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}]$, introduced in Eq. (3), describes a generic scalar field theory in the presence of quenched disorder. In order to analyze it and understand in more detail what kind of disorder is generated by $\mathcal{C}_{\mathrm{eq}}$ one can study the cumulants of $\mathcal{S}$ by considering replicas of the original system. As known in the context of the critical behavior of the RFIM,[64,65] $\exp(-\mathcal{S}_{\mathrm{rep}}[\{p_a(\mathbf{x})\}]) = \overline{\exp(-\sum_{a=1}^{n} \mathcal{S}[p_a(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}])}$ generates the cumulants of the action $\mathcal{S}[p(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}]$ through an expansion in increasing number of free replica sums:

$$\mathcal{S}_{\mathrm{rep}}[\{p_a(\mathbf{x})\}] = \sum_{a=1}^{n} \mathcal{S}_1[p_a(\mathbf{x})] - \frac{1}{2} \sum_{a,b=1}^{n} \mathcal{S}_2[p_a(\mathbf{x}), p_b(\mathbf{x})] + \frac{1}{3!} \sum_{a,b,c=1}^{n} \mathcal{S}_3[p_a(\mathbf{x}), p_b(\mathbf{x}), p_c(\mathbf{x})] + \dots, \qquad (4)$$

where $\mathcal{S}_l[p_1, \dots, p_l]$ is the l-th cumulant of $\mathcal{S}[p(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}]$,

$$\begin{aligned}
\mathcal{S}_1[p(\mathbf{x})] &= \overline{\mathcal{S}[p(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}]}, \\
\mathcal{S}_2[p_1(\mathbf{x}), p_2(\mathbf{x})] &= \overline{\mathcal{S}[p_1(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}]\mathcal{S}[p_2(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}]} - \overline{\mathcal{S}[p_1(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}]}\,\overline{\mathcal{S}[p_2(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}]},
\end{aligned} \qquad (5)$$

etc.

To compute the replicated action $\mathcal{S}_{\mathrm{rep}}[\{p_a(\mathbf{x})\}]$ one needs to perform the average over the reference configuration. This is achieved by introducing $n+1$ replicas, which will be identified by Greek letters $\alpha = 0, 1, \dots, n$ (whereas Roman letters $a = 1, \dots, n$ will still be used for replicas from 1 to n only):

$$\begin{aligned}
e^{-\mathcal{S}_{\mathrm{rep}}[\{p_a(\mathbf{x})\}]} &= \frac{1}{Z} \sum_{\mathcal{C}_{\mathrm{eq}}, \mathcal{C}_a} e^{-\beta \mathcal{H}(\mathcal{C}_{\mathrm{eq}})} e^{-\beta \sum_{a=1}^{n} \mathcal{H}(\mathcal{C}_a)} \prod_{a=1}^{n} \delta\left[p_a(\mathbf{x}) - Q_{\mathbf{x}}(\mathcal{C}_a, \mathcal{C}_{\mathrm{eq}})\right] \\
&= \frac{1}{Z} \sum_{\mathcal{C}_\alpha} e^{-\beta \sum_{a=1}^{n} \mathcal{H}(\mathcal{C}_\alpha)} \prod_{a=1}^{n} \delta\left[p_a(\mathbf{x}) - Q_{\mathbf{x}}(\mathcal{C}_a, \mathcal{C}_{\mathrm{eq}})\right],
\end{aligned} \qquad (6)$$

where $\mathcal{C}_0 \equiv \mathcal{C}_{\mathrm{eq}}$.

B. An illustrative toy model: the REM

The REM is the simplest mean-field system displaying a thermodynamic glass transition (RFOT), and it therefore represents a natural first benchmark for our analysis. Interestingly, the effective theory for this model can be worked out without resorting to any approximation.

The REM, which was introduced by Derrida,[33] is a disordered spin model defined as follows: The energies $E(\mathcal{C})$ of the configurations $\mathcal{C} = \{S_1, \dots, S_N\}$, where $S_i = \pm 1$, are independent and identically distributed (i.i.d.) random variables sampled from a Gaussian distribution with zero mean and variance equal to $N/2$. The REM displays a RFOT at $T_K = 1/(2\sqrt{\ln 2})$. Below this temperature the system freezes in the lowest available states whereas above it the configurational entropy is positive. (The temperature T_d of the dynamical glass transition is infinite in this model.) The main reason for the simplicity of the REM lies in the fact that "states" and "configurations" coincide, i.e., the intra-state entropy is zero. In consequence, the overlap p takes only the values one and zero, which respectively correspond to the two replicas being in the same or in different states/configurations.

Since this model has explicit quenched disorder (the random energies), we have to perform an additional average over the distribution of this disorder. Hence, Eq. (6) now reads

$$e^{-\mathcal{S}_{\mathrm{rep}}[\{p_a\}]} = \frac{1}{\overline{Z}} \overline{\sum_{\mathcal{C}_0, \dots, \mathcal{C}_n} \exp\left(-\beta \sum_\alpha E(\mathcal{C}_\alpha)\right) \prod_{a=1}^{n} \delta_{p_a, \delta_{\mathcal{C}_0, \mathcal{C}_a}}} = \frac{1}{\overline{Z}} \sum_{\mathcal{C}_0, \dots, \mathcal{C}_n} \exp\left(\frac{\beta^2 N}{4} \sum_{\alpha, \beta} \delta_{\mathcal{C}_\alpha, \mathcal{C}_\beta}\right) \prod_{a=1}^{n} \delta_{p_a, \delta_{\mathcal{C}_0, \mathcal{C}_a}}, \qquad (7)$$

where we have used the "annealed approximation", exact only for $T \geq T_K$,[66] in which one does not need to introduce another set of replicas to handle the average over the random energies. By using the fact that two replicas having an overlap one with the reference configuration also have a mutual overlap equal to one, it is straightforward to obtain that the term in the argument of the exponential in Eq. (7) can be rewritten as

$$\sum_{\alpha, \beta} \delta_{\mathcal{C}_\alpha, \mathcal{C}_\beta} = 1 + 2 \sum_{a=1}^{n} \delta_{\mathcal{C}_a, \mathcal{C}_0} + \sum_{a,b} \delta_{\mathcal{C}_a, \mathcal{C}_b} = 1 + 2 \sum_{a=1}^{n} p_a + \sum_{a,b} p_a p_b + \sum_{a,b}^{\star} \delta_{\mathcal{C}_a, \mathcal{C}_b} = \left(1 + \sum_a p_a\right)^2 + \sum_{a,b}^{\star} \delta_{\mathcal{C}_a, \mathcal{C}_b},$$

where $\sum_{a,b}^{\star}$ denotes a sum that only runs over replicas a and b having a zero overlap with the reference configuration. Using this result we can rewrite the partition function as

$$e^{-\mathcal{S}_{\mathrm{rep}}[\{p_a\}]} \propto \sum_{\mathcal{C}_0} e^{\frac{\beta^2 N}{4}\left(1+\sum_a p_a\right)^2} \sum_{\mathcal{C}_1,\ldots,\mathcal{C}_n}^{\star} \exp\left(\frac{\beta^2 N}{4}\sum_{a,b}^{\star}\delta_{\mathcal{C}_a,\mathcal{C}_b}\right), \tag{8}$$

where again the star in the sum over configurations means that one has to sum only over replicas having a zero overlap with the reference configuration (and thus only over $2^N - 1$ configurations different from $\mathcal{C}_0$). The term between parentheses is nothing else than the replicated partition function of a REM with a number of replicas equal to $n' = n - \sum_a p_a$ (and $2^N - 1$ available configurations). one can use the replica method to compute it. Since $0 \leq n' \leq n \to 0$ and we consider $T \geq T_K$, the replicated partition function appearing in Eq. (8) can be obtained as[34]

$$\sum_{\mathcal{C}_1,\ldots,\mathcal{C}_n}^{\star} \exp\left(\frac{\beta^2 N}{4}\sum_{a,b}^{\star}\delta_{\mathcal{C}_a,\mathcal{C}_b}\right) \approx \exp\left[\left(n-\sum_a p_a\right)\left(N\ln 2 + \frac{\beta^2 N}{4}\right)\right],$$

which coincides with the annealed approximation. After collecting all these results together we find the following expression for the replicated action:

$$\mathcal{S}_{\mathrm{rep}}[\{p_a\}] = -(n+1)\left(N\ln 2 + \frac{\beta^2 N}{4}\right) + \left(N\ln 2 - \frac{\beta^2 N}{4}\right)\sum_a p_a - \frac{\beta^2 N}{4}\left(\sum_a p_a\right)^2. \tag{9}$$

This can be directly interpreted as the replicated action for a two-state variable $p = 0, 1$, or equivalently as the replicated action for the Hamiltonian of an Ising variable $\sigma = 2p - 1$ coupled to a Gaussian random magnetic field,

$$\beta\mathcal{H}_{\mathrm{eff}} = \mathrm{cst} - (\mu + \delta\mu)p = \mathcal{S}_0 - (H + \delta h)\sigma, \tag{10}$$

with $H = \mu/2 = -N(4\ln 2 - \beta^2)/8$, $\overline{\delta h} = 0$, $\overline{\delta h^2} = \overline{\delta\mu^2}/4 = N\beta^2/8$, and $\mathcal{S}_0 = \mathrm{cst} - \delta h$, which is the result already given in Eq. (1) of the Introduction. (By enforcing $H = 0$ one recovers the value of the critical temperature of the REM, as expected.) Thus, the statistics of the fluctuations of the overlap with a reference configuration, which is described by $\mathcal{S}[p(\mathbf{x})|\mathcal{C}_{\mathrm{eq}}]$, is the same as that of an Ising (or discrete global overlap) variable that is subjected to a field with an average value of the order of N, favoring the $\sigma = -1$ (zero overlap) state and vanishing at the transition temperature, and with fluctuations of the order of $\sqrt{N}$. (The role of the random energy $\mathcal{S}_0$ is to ensure that the variance of the effective Hamiltonian is equal to zero in the zero-overlap, or $\sigma = -1$, state.)

In conclusion we have found that the theory describing the overlap fluctuations of the REM is a 0-dimensional RFIM. We did not attempt to generalize the computation for temperatures below T_K but from known results on the REM, we expect to find a disordered action corresponding to an Ising-like variable coupled to an external random field whose typical strength is of the order of one. The analysis performed in this section shows that without any approximation the RFIM naturally emerges in the study of glassy systems for temperatures below T_d.

C. Effective theory for the fully connected 2^M-KREM

In this section we consider a nontrivial, but still exactly solvable, generalization of the REM introduced for the first time in Ref. [34], the 2^M-KREM on a fully connected lattice. We apply the strategy outlined in Sec. III A to obtain the (quasi) exact effective theory that describes the statistics of the fluctuations of the overlap profile with an equilibrium reference configuration between T_d and T_K. This theory will turn out to be given by a fully connected random field + random bond Ising model with multi-body interactions [see Eq. (2)].

The model is defined as follows: Given N sites, on each site i there are 2^M configurations, $\mathcal{C}_i = \{1, \cdots, 2^M\}$, and on each link (i, j) we define i.i.d. Gaussian random energies $E_{ij} = E(\mathcal{C}_i, \mathcal{C}_j)$ with $\overline{E_{ij}(\mathcal{C}_i, \mathcal{C}_j)} = 0$ and $\overline{E_{ij}(\mathcal{C}_i, \mathcal{C}_j)E_{ij}(\mathcal{C}_i', \mathcal{C}_j')} = M\delta_{\mathcal{C}_i,\mathcal{C}_i'}\delta_{\mathcal{C}_j,\mathcal{C}_j'}$. The Hamiltonian of the model is simply given by

$$\mathcal{H} = \frac{1}{2\sqrt{N}}\sum_{i\neq j} E_{ij}(\mathcal{C}_i, \mathcal{C}_j).$$

The more standard mean-field result corresponds to $M \to \infty$ (M plays the same role as N in the simple REM discussed above). The model can be solved exactly by using replicas, as shown in Appendix A 1, and the thermodynamic glass transition (RFOT) taking place at T_K can be obtained within a 1-step replica-symmetry-breaking (1-RSB) ansatz.[67]

To construct the effective theory we consider $n+1$ replicas of the system and compute the replicated action for a fixed overlap configuration $\{p_a^i\}$ of the replicas $a = 1, \ldots, n$ with a given equilibrium reference configuration $\{\mathcal{C}_i^0\}$. Note that $p_a^i = 1$ only if $\mathcal{C}_a^a = \mathcal{C}_i^0$ and is zero otherwise. As already mentioned, we consider the temperature range $T_d \leq T \leq T_K$, where we can use the annealed approximation to perform the average over the random energies. (We stress that it is crucial on the other hand that the average over the quenched disorder represented by the reference configuration is performed exactly.) The replicated action then reads

$$
e^{-\mathcal{S}_{\mathrm{rep}}[\{p_a^i\}]} = \frac{1}{\overline{Z}} \overline{\sum_{\{\mathcal{C}_i^\alpha\}} e^{-\frac{\beta}{2\sqrt{N}} \sum_{i \neq j, \alpha} E_{ij}(\mathcal{C}_i^\alpha, \mathcal{C}_j^\alpha)} \prod_{a,i} \delta_{p_a^i, \delta_{\mathcal{C}_i^0, \mathcal{C}_i^a}}} = \frac{1}{\overline{Z}} \sum_{\{\mathcal{C}_i^\alpha\}} e^{\frac{\beta^2 M}{8N} \sum_{i \neq j} \sum_{\alpha, \beta = 0}^n \delta_{\mathcal{C}_i^\alpha, \mathcal{C}_i^\beta} \delta_{\mathcal{C}_j^\alpha, \mathcal{C}_j^\beta}} \prod_{a,i} \delta_{p_a^i, \delta_{\mathcal{C}_i^0, \mathcal{C}_i^a}} . \quad (11)
$$

The Kronecker δ's in the exponential of the above expression can be rewritten in terms of the overlap variables as

$$
\sum_{\alpha, \beta = 0}^n \delta_{\mathcal{C}_i^\alpha, \mathcal{C}_i^\beta} \delta_{\mathcal{C}_j^\alpha, \mathcal{C}_j^\beta} = 1 + n + 2 \sum_{a=1}^n p_a^i p_a^j + \sum_{a \neq b = 1}^n \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \delta_{\mathcal{C}_j^a, \mathcal{C}_j^b} . \quad (12)
$$

We note that if $\mathcal{C}_i^a = \mathcal{C}_i^0$ and $\mathcal{C}_i^b = \mathcal{C}_i^0$ (i.e., $p_a^i = p_b^i = 1$), then $\mathcal{C}_i^b = \mathcal{C}_i^a$. Similarly, if $\mathcal{C}_i^a = \mathcal{C}_i^0$ and $\mathcal{C}_i^b \neq \mathcal{C}_i^0$ (i.e., $p_a^i = 1$ and $p_b^i = 0$), then $\mathcal{C}_i^b \neq \mathcal{C}_i^a$. The same is true, of course, if $\mathcal{C}_i^a \neq \mathcal{C}_i^0$ and $\mathcal{C}_i^b = \mathcal{C}_i^0$. The only undetermined case corresponds to $\mathcal{C}_i^a \neq \mathcal{C}_i^0$ and $\mathcal{C}_i^b \neq \mathcal{C}_i^0$.

As discussed above, the expansion of $\mathcal{S}_{\mathrm{rep}}[\{p_a^i\}]$ in an increasing number of unrestricted sums over replicas, Eq. (4), generates the cumulants of the effective disordered Hamiltonian describing the fluctuations of the overlap with a reference configuration. Below, we compute the first and second cumulants of such an effective Hamiltonian, which correspond to the 1-replica and 2-replica components of the replicated action [see Eq. (5)].

1. First cumulant of the effective disordered Hamiltonian

Let us first focus on the first cumulant $\mathcal{S}_1[\{p^i\}]$. From Eq. (4) one realizes that the simplest way of computing it is to set all replica fields equal, $p_a^i = p^i \; \forall a = 1, \cdots, n$ and $\forall i$, keep only the term of order n in the expression of $\mathcal{S}_{\mathrm{rep}}[\{p_a^i\}]$, and take the limit $n \to 0$ in the end, as in the standard replica trick.[32] After averaging over the random energies and the reference configuration, all the sites become equivalent and $\mathcal{S}_1[\{p^i\}]$ can only be a function of $c = (1/N) \sum_i p^i$, which coincides with the global mean overlap with the reference configuration. Therefore we will use $\mathcal{S}_1(c)$ in place of $\mathcal{S}_1[\{p^i\}]$ in what follows. This implies that scanning over all the possible configurations of the overlap profile $\{p^i\}$ is the exact analogue of setting the overlap with the reference configuration for all replicas to be 1 on the first cN sites (i.e., $p_a^i = 1$ for $i = 1, \ldots, cN$ and $\forall a$) and 0 on all the other $(1-c)N$ sites (i.e., $p_a^i = 0$ for $i = cN + 1, \ldots, N$ and $\forall a$).

Although this procedure resembles that of the random pinning,[51,53] it is different in that on the sites where $p^i = 0$ the replicas *cannot* be in the same configuration as the reference one. It is also different from the computation of the Franz-Parisi potential[16–18] discussed in Sec. II B. The reason is that in the latter case one sums over all configurations in which all replicas have the same fixed global overlap with the reference configurations ($\sum_i p_a^i = cN \; \forall a$) but with different replicas having in general different overlap profiles ($p_a^i \neq p_b^i$), whereas in the present procedure one restricts the sum to configurations in which *all* the n replicas are constrained to have the *same* specific overlap profile with $\{\mathcal{C}_i^0\}$ ($p_a^i = p^i \; \forall a$, such that $\sum_i p^i = cN$).

Our basic idea is to evaluate $\mathcal{S}_1(c)$ by expanding it for small c around $c = 0$, $\mathcal{S}_1(c) = \sum_q \mathcal{S}_1^{(q)}(0) c^q/q!$, and, since any power of c can be re-expressed in terms of effective interactions among the p^i's, e.g., $c^q = (1/N^{q-1}) \sum_{i_1, \ldots, i_q} p^{i_1} \cdots p^{i_q}$, one can re-interpret the expansion of the 1-replica component $\mathcal{S}_1(c)$ as the average part of an effective diosrdered Hamiltonian with multi-body interactions of the form

$$
\mathcal{S}_1(c) = \mathrm{cst} - \mu \sum_i p^i - \frac{w_2}{2N} \sum_{i \neq j} p^i p^j - \frac{w_3}{3! N^2} \sum_{i,j,k \neq} p^i p^j p^k - \frac{w_4}{4! N^3} \sum_{i,j,k,l \neq} p^i p^j p^k p^l + \ldots, \quad (13)
$$

with $\mu = -\mathcal{S}_1^{(1)}(0)$, and $w_q = -\mathcal{S}_1^{(q)}(0)$. Since we are interested here in obtaining the effective Hamiltonian for $T \geq T_K$, the typical equilibrium configurations of the overlap are expected to have a small number of sites where $p^i = 1$. We thus look for an effective Hamiltonian which is accurate for small c and we expect that the first few coefficients of the expansion of $\mathcal{S}_1(c)$ are sufficient to reproduce its behavior correctly. (Again, reconstructing the behavior of $\mathcal{S}_1(c)$ at large c is less important, since configurations with c close to 1 are very rare for $T > T_K$.) Note that the strategy presented here to compute $\mathcal{S}_1(c)$ can in principle be straightforwardly applied to any mean-field

model in the same "universality class" as the REM and the 2^M-KREM, with a complex free-energy landscape and a thermodynamic glass transition (RFOT), possibly with some minor and model-dependent modifications.

For the chosen overlap profile, Eq. (12) becomes (after dropping sub-extensive terms)

$$\sum_{i\neq j}\left[1+n+2np^ip^j+\sum_{a\neq b}\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}\delta_{\mathcal{C}_j^a,\mathcal{C}_j^b}\right]\approx(1+n)N^2+n(n+1)c^2N^2+2cN\sum_{a\neq b}\sum_i^\star\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}+\sum_{a\neq b}\left(\sum_i^\star\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}\right)^2,$$

where the sum $\sum_i^\star$ represents the sum over the sites $i=cN+1,\dots,N$ where $p^i=0$. Inserting this expression into Eq. (11) yields

$$e^{-n\mathcal{S}_1(c)}=\frac{e^{\frac{N\beta^2M}{8}\left[1+n+n(n+1)c^2\right]}}{\overline{Z}}\sum_{\{\mathcal{C}_i^\alpha\}}e^{\frac{N\beta^2M}{8}\left[2c\sum_{a\neq b}\frac{1}{N}\sum_i^\star\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}+\sum_{a\neq b}\left(\frac{1}{N}\sum_i^\star\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}\right)^2\right]}\prod_{a,i}\delta_{p^i,\delta_{\mathcal{C}_i^0,\mathcal{C}_i^a}}.$$

On the first cN sites, $p^i=1$ and, accordingly, $\mathcal{C}_i^a=\mathcal{C}_i^0$ for all a. Since the sum over the reference configuration $\mathcal{C}_i^0$ simply gives 2^{NM}, we thus obtain

$$e^{-n\mathcal{S}_1(c)}=\frac{e^{NM\left\{\ln 2+\frac{\beta^2}{8}\left[1+n+n(n+1)c^2\right]\right\}}}{\overline{Z}}\sum_{\{\mathcal{C}_i^a\}_\star}e^{\frac{\beta^2M}{4}\sum_{a<b}\left[2c\sum_i^\star\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}+\frac{1}{N}\left(\sum_i^\star\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}\right)^2\right]},$$

where the trace $\sum_{\{\mathcal{C}_i^a\}_\star}$ represents the sum over all the 2^M-1 configurations $\mathcal{C}_i^a$ different from the reference one on the $(c-1)N$ sites where $p^i=0$.

One can now introduce the overlaps q_{ab} by performing $n(n-1)/2$ Hubbard-Stratonovich transformations (all the details of the calculations are reported in Appendix A 2 a). We posit a replica-symmetric (RS) ansatz for the overlap matrix, $q_{ab}=q_0$, which is, again, expected to be justified for $T\geq T_K$ and for small c, as it is for instance for the Franz-Parisi potential[18] where the RS ansatz is appropriate for small and large enough values of the overlap (*i.e.*, of c), whereas one needs to use a 1-RSB ansatz for intermediate values of c, around the barrier. This is quite clear on physical grounds, as increasing c effectively reduces the configurational entropy that is accessible to the constrained system, thereby inducing a 1-RSB glass transition at moderately large values of the overlap, whereas for larger values of c the system is constrained to be in same configuration as the reference one, which corresponds to a RS ansatz. In the following we will be mostly interested in the lowest-order coefficients of the expansion of $\mathcal{S}_1(c)$ in powers of c, which give the dominant effective interactions in Eq. (13). The values of these coefficients can therefore be computed by means of a RS ansatz. Furthermore, our goal here is not to obtain the complete analytic expression of $\mathcal{S}_1(c)$, but just to illustrate the general strategy allowing us to compute it. Computing $\mathcal{S}_1(c)$ within a 1-RSB ansatz is certainly doable, but would just make the calculation much more cumbersome and involved without changing the general picture.

In the thermodynamic limit ($N\to\infty$), by using the saddle-point method, we finally derive (see Appendix A 2 a for details)

$$\frac{\mathcal{S}_1(c)}{N}=-\frac{\beta^2M}{8}(1+c^2)-\frac{\beta^2Mq_0^2}{8}+\frac{(1-c)\beta^2M(q_0+c)}{4}-(1-c)\overline{\ln\left[\sum_\mathcal{C}^\star e^{\sqrt{\frac{\beta^2M(q_0+c)}{2}}z_\mathcal{C}}\right]},\qquad(14)$$

where the overlap q_0 satisfies the following self-consistent equation:

$$q_0=(1-c)\left\{1-\sqrt{\frac{2}{\beta^2M(q_0+c)}}\overline{\left(\frac{\sum_\mathcal{C}^\star z_\mathcal{C}\, e^{\sqrt{\frac{\beta^2M(q_0+c)}{2}}z_\mathcal{C}}}{\sum_\mathcal{C}^\star e^{\sqrt{\frac{\beta^2M(q_0+c)}{2}}z_\mathcal{C}}}\right)}\right\}.\qquad(15)$$

The averages of the form $\overline{[f(\vec{z}_\mathcal{C})]}$ appearing in the above expressions are defined over a Gaussian measure, $\overline{[f(\vec{z}_\mathcal{C})]}\equiv\int\prod_\mathcal{C}^\star\left[\frac{dz_\mathcal{C}}{\sqrt{2\pi}}e^{-z_\mathcal{C}^2/2}\right]f(\vec{z}_\mathcal{C})$, with $\vec{z}_\mathcal{C}$ being a (2^M-1)-dimensional vector. The solution of such a saddle-point equation can be expanded in powers of c as $q_0\approx q_{0,0}+q_{0,1}c+q_{0,2}c^2+\dots$, which, when inserted back into Eq. (14), allows one to obtain the exact coefficients of the expansion of $\mathcal{S}_1(c)$.

We illustrate the output by providing explicit analytical expressions for the chemical potential μ and the coupling constants w_q appearing in Eq. (13) in large-M limit. In this limit, q_0 is of order of $1/2^M$. In consequence, expanding
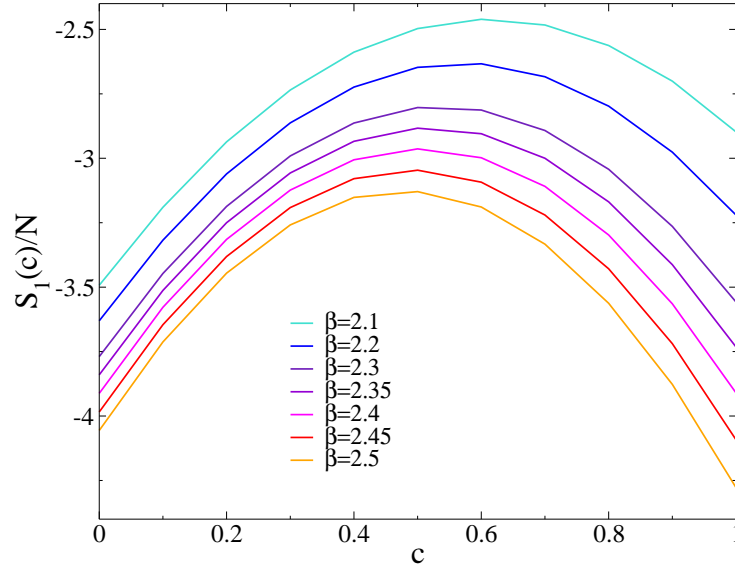
FIG. 3: 1-replica part $\mathcal{S}_1(c)$ of the replicated effective action as a function of the overlap c with a reference equilibrium configuration for the fully connected 2^M-KREM with $M = 3$ [Eqs. (14) and (15)]. Several values of the temperature, $T \geq T_K \approx 0.4$ (*i.e.*, $\beta \leq \beta_K \approx 2.5$), are shown.

the exponentials in Eqs. (14) and (15) up to the eighth order in $\sqrt{M\beta^2(q_0 + c)/2}$ gives Eq. (A7) of Appendix A 2 a, which, when inserted into Eq. (14), leads to the following expressions of μ and of w_q up to the fourth order:

$$\text{cst} = -\frac{M\beta^2}{8} - M\ln 2$$
$$\mu = -M\ln 2 - \frac{M\beta^2}{2^{M+2}}$$
$$w_2 = \frac{M\beta^2}{4} - \frac{M\beta^2(M\beta^2 - 4)}{2^{M+3}} \tag{16}$$
$$w_3 = -\frac{M^2\beta^4(M\beta^2 - 6)}{2^{M+4}},$$
$$w_4 = -\frac{M^3\beta^6(M\beta^2 - 8)}{2^{M+5}}.$$

When $M \to \infty$ all the interactions beyond the pairwise one vanish (as in the REM case: see above). Note also that all the coupling constants w_q seem to decrease as the temperature is increased, in agreement with physical intuition.

For finite values of M, another strategy to determine the coefficients of the effective Hamiltonian consists instead in solving Eq. (15) numerically for several values of c and β, inserting the result into Eq. (14), and fitting the function $\mathcal{S}_1(c)$ so obtained by a polynomial function of c. This yields the values of the coefficients μ and w_q of the effective Hamiltonian, as well as their temperature dependence. The procedure is illustrated in Fig. 3 for $M = 3$, where $\mathcal{S}_1(c)$ is plotted for several values of β for $T \gtrsim T_K$. Fitting these curves with polynomials of c of degree 4 provides a numerical estimate of μ, w_2, w_3, and w_4 for different temperatures.

2. Second cumulant of the effective disordered Hamiltonian

We now turn to the computation of the second cumulant, *i.e.*, the 2-replica part of the replicated action defined in Eq. (11). To do this we divide the n constrained replicas into two groups of n_1 and n_2 replicas respectively. The most generic overlap profile can be obtained by dividing the sites into four groups, denoted (1), (2), (12), and (0), such that on the $c_1 N$ sites belonging to the group (1) $p_1^i = 1$ and $p_2^i = 0$, on the $c_2 N$ sites belonging to the group (2) $p_1^i = 0$ and $p_2^i = 1$, on the $c_{12} N$ sites belonging to the group (12) $p_1^i = p_2^i = 1$, and on the $c_0 N$ sites belonging to the group (0) $p_1^i = p_2^i = 0$ (with $c_0 = 1 - c_1 - c_2 - c_{12}$). The second cumulant can be computed by keeping only the terms of order $n_1 n_2$ in the expression of the replicated action, and taking the limit $n_1, n_2 \to 0$ [see Eq. (4) and

Ref. [32]]. The key observation is again that the second cumulant can only be a function of c_1, c_2, and c_{12}, and thus of the global overlaps [see Eq. (18)].

Our general strategy is similar to that used for the first cumulant:

1. We first perform the standard Hubbard-Stratonovich transformations that allow us to decouple the sites of the fully connected lattice via three overlap matrices, which correspond to the overlaps between two different constrained replicas belonging to the first group, $q_{ab}^{[1]}$, to the second group, $q_{ab}^{[2]}$, and to the two different groups, $q_{ab}^{[12]}$.

2. We then posit a RS ansatz for the overlaps, $q_{ab}^{[1]} = q_1 \ \forall a \neq b$, $q_{ab}^{[2]} = q_2 \ \forall a \neq b$, and $q_{ab}^{[12]} = q_{12} \ \forall a, b$, which is expected to be justified for $T \geq T_K$ and for small enough c_1, c_2, and c_{12} , and we perform the trace over the configurations.

3. In the thermodynamic limit the saddle-point method yields the expression of the replicated free energy in terms of the overlaps [Eqs. (A8) and (A11)], with the latter obeying three self-consistent equations [Eqs. (A12) and (A13)].

4. In order to compute the terms of order $n_1 n_2$ of the replicated action, which yields the second cumulant of the effective Hamiltonian, we expand the RS overlaps as

$$
\begin{aligned}
q_1 &\approx q_1^{[0,0]} + n_2 q_1^{[0,1]} + O(n_1, n_2^2, n_1 n_2) \,, \\
q_2 &\approx q_2^{[0,0]} + n_1 q_2^{[1,0]} + O(n_2, n_1^2, n_1 n_2) \,, \\
q_{12} &\approx q_{12}^{[0,0]} + O(n_1, n_2) \,.
\end{aligned}
\tag{17}
$$

Inserting this expansion into the saddle-point equations, Eqs. (A12) and (A13), allows us to obtain $q_1^{[0,0]}$, $q_1^{[0,1]}$, $q_2^{[0,0]}$, $q_2^{[1,0]}$, and $q_{12}^{[0,0]}$ which, when inserted into the expression of the replicated action, Eq. (A11), finally leads to the second cumulant as a function of the concentrations c_1, c_2, and c_{12}.

5. In practice, we are interested in the expansion of $\mathcal{S}_2(c_1, c_2, c_{12})$ only up to the second order in the concentrations of the different kinds of sites, c_1, c_2 and c_{12}, which corresponds to the most relevant random terms. In fact, it is easy to realize that any power of the concentrations can be re-expressed as effective random terms in the expression of the second cumulant through

$$
c_1 = \frac{1}{N} \sum_i p_1^i (1 - p_2^i) \,, \qquad c_2 = \frac{1}{N} \sum_i p_2^i (1 - p_1^i) \,, \qquad c_{12} = \frac{1}{N} \sum_i p_1^i p_2^i \,.
\tag{18}
$$

The calculations, although conceptually simple, are long and tedious. In consequence, for the sake of the clarity of the presentation, we give the explicit expression of the second cumulant in the large-M limit only, up to the leading order in $1/2^M$, and we defer all the details to Appendix A 2 b:

$$
\mathcal{S}_2[\{p_1^i, p_2^i\}] \approx \frac{M\beta^2}{4} \left(1 + \frac{M\beta^2 + 4}{2^{M+1}} \right) \frac{1}{N} \sum_{i \neq j} p_1^i p_2^i p_1^j p_2^j + \frac{M\beta^2}{2^{M+1}} \left[1 - \frac{1}{N} \sum_i \left(p_1^i + p_2^i \right) \right] \sum_j p_1^j p_2^j \,.
\tag{19}
$$

3. The effective disordered Hamiltonian

Following Sec. III A, the expansion in replica sums of the replicated action is equivalent to the expansion in cumulants of a disordered Hamiltonian $\mathcal{H}_{\text{eff}}[p] \equiv \mathcal{S}[p(\mathbf{x})|\mathcal{C}_{\text{eq}}]$, with the identification $\mathcal{S}_1[p_a] = \overline{\beta \mathcal{H}_{\text{eff}}[p_a]}$, $\mathcal{S}_2[p_a, p_b] = \overline{\beta \mathcal{H}_{\text{eff}}[p_a] \beta \mathcal{H}_{\text{eff}}[p_b]} - \overline{\beta \mathcal{H}_{\text{eff}}[p_a]} \ \overline{\beta \mathcal{H}_{\text{eff}}[p_b]}$, etc., where the overline now denotes an average over the effective quenched disorder. The form of the effective disordered Hamiltonian that is able to reproduce the 1- and 2-replica parts derived above reads

$$
\beta \mathcal{H}_{\text{eff}} = \text{cst} - \sum_i (\mu + \delta\mu_i) p^i - \frac{1}{2} \sum_{i \neq j} \left(\frac{w_2}{N} + \frac{\delta w_{2,ij}}{\sqrt{N}} \right) p^i p^j - \frac{w_3}{3! N^2} \sum_{i,j,k \neq} p^i p^j p^k - \frac{w_4}{4! N^3} \sum_{i,j,k,l \neq} p^i p^j p^k p^l + \dots \,,
$$

where the chemical potential μ and the couplings w_q are given in Eq. (16), and the covariances of the quenched random variables $\delta\mu_i$, $\delta w_{2,ij}$ have to be chosen in order to reproduce the expression of $\mathcal{S}_2$ in Eq. (19), namely,

$$
\begin{aligned}
\overline{\delta\mu_i \delta\mu_j} &= \frac{M\beta^2}{2^{M+1}} \delta_{ij} \,, \\
\overline{\delta w_{2,ij} \delta w_{2,kl}} &= \frac{M\beta^2}{2} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \,, \\
\overline{\delta\mu_i \delta w_{2,jk}} &= -\frac{M\beta^2}{2^M} \frac{\delta_{ij} + \delta_{ik}}{\sqrt{N}} \,,
\end{aligned}
\tag{20}
$$

and $\overline{\delta\mu_i} = \overline{\delta w_{2,ij}} = 0$.

Going as before from overlap variables to spin variables, $p^i = (1+\sigma^i)/2$, one finally obtains the effective random-field + random-bond fully connected Ising model $\beta\mathcal{H}_{\text{eff}}[\sigma]$ given in Eq. (2). The explicit expressions of the external field and the coupling constants for $M \gg 1$ are

$$
\begin{aligned}
H &= \left[\frac{\mu}{2} + \frac{w_2}{4} + \frac{w_3}{16} + \frac{w_4}{96} + \dots\right] \approx \frac{1}{2}\left[\frac{M\beta^2}{8} - M\ln 2 - \frac{M^2\beta^4}{3\cdot 2^{M+9}}\left(24 + 4M\beta^2 + M^2\beta^4\right) + \dots\right] \\
J_2 &= \frac{w_2}{4} + \frac{w_3}{16} + \frac{w_4}{32} + \dots \approx \frac{1}{16}\left[M\beta^2 + \frac{M\beta^2}{2^{M+6}}\left(128 + 16M\beta^2 - M^2\beta^4\right) + \cdots\right] \\
J_3 &= \frac{w_3}{8} + \frac{w_4}{16} + \dots \approx \frac{M^2\beta^4(24 + 4M\beta^2 - M^2\beta^4 + \cdots)}{2^{M+9}}, \\
J_4 &= \frac{w_4}{16} + \dots \approx \frac{M^3\beta^6(8 - M\beta^2 + \cdots)}{2^{M+9}},
\end{aligned}
\tag{21}
$$

and the random fields and random couplings are characterized by

$$
\begin{aligned}
\overline{\delta h_i \delta h_j} &\approx \frac{M\beta^2}{32}\left(1 - 2^{2-M}\right)\delta_{ij} + \frac{M\beta^2}{32N}\left(1 - 2^{3-M}\right), \\
\overline{\delta J_{2,ij}\delta J_{2,kl}} &\approx \frac{M\beta^2}{32}\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right), \\
\overline{\delta h_i \delta J_{2,jk}} &\approx \frac{M\beta^2}{32}\left(1 + 2^{2-M}\right)\frac{\delta_{ij} + \delta_{ik}}{\sqrt{N}},
\end{aligned}
\tag{22}
$$

and $\overline{\delta h_i} = \overline{\delta J_{2,ij}} = 0$.

Note that compared to the REM studied in the preceding section, the introduction of a finite number of states, 2^M, leads to additional multi-body interactions and additional random terms. The additional multi-body interactions vanish in the limit $M \to \infty$. In this limit the cumulants of the random variables simplify, but one nonetheless remains with both random fields and random bonds. This difference between 2^M-KREM and REM stems from the fact that the random energies are defined on the links of the lattice in the former and on the sites in the latter.

4. Comparison with the exact solution

The main result of this section is that the effective theory describing the probability distribution of the thermal fluctuations of the overlap with an equilibrium configuration for the fully connected 2^M-KREM corresponds to a random-field + random-bond fully connected Ising model with multi-body interactions. In practice, we need to truncate the number of multi-body interactions (*e.g.*, up to 4-body terms, as done above) and truncate as well the expansion in cumulants of the random variables (*e.g.*, keeping only the second cumulants). In order to test the quantitative accuracy of the truncated effective theory we have computed its prediction for the mean overlap $\langle p \rangle$ with the reference configuration [which in the Ising model is simply related to the magnetization m through $\langle p \rangle = (1+m)/2$] as a function of temperature and compared it to the exact result derive in Appendix A 1. (To further simplify the computation we have dropped the cross-correlations between the random fields and the random bonds and we have neglected the off-diagonal term of the random-field distribution: Details can be found in Appendix B.) The comparison for the case $M = 3$ is displayed in Fig. 4. One can see that there is a very good agreement, which thus shows that neglecting higher-order interactions and cumulants is not only qualitatively and but also quantitatively justified.

Note that, since we have used the annealed approximation to perform the average over the random energies in the derivation of the effective theory, our results are in principle only valid on the high-temperature side of the RFOT, $T \geq T_K$, including the RFOT itself. Our procedure could be extended to the low-temperature side by performing a quenched average over the random energies. Although standard, the computation becomes long and tedious in this case. Moreover, the aim of our analysis is not to provide an accurate determination of the numerical values of the various parameters entering the effective disordered Hamiltonian deep in the low-temperature glass phase; it is instead to show how and why such an effective description emerges in a general, transparent, and robust way, and to analyze the properties of the transition point. For these reasons the results derived from the effective disordered Ising Hamiltonian in the region $T < T_K$ (T_K corresponds to a first-order transition with a jump of the magnetization in the effective description) are simply obtained with approximate coupling constants and covariances determined through the annealed approximation. The agreement with the exact solution is nonetheless quite good.

For finite M the expansion of the first and second cumulants $\mathcal{S}_1(c)$ and $\mathcal{S}_2(c_1, c_2, c_{12})$ in powers of c, c_1, c_2, and c_{12} generates multi-body interactions and random terms to all orders (higher-order cumulants are present as well).
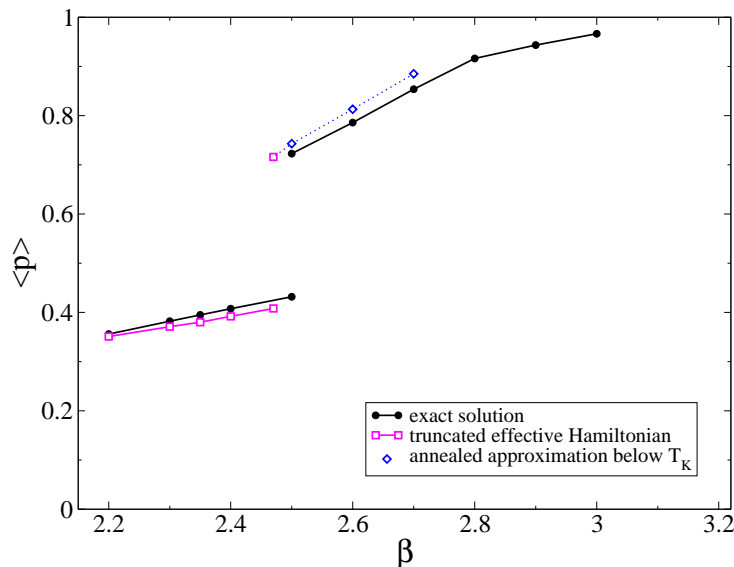
FIG. 4: Comparison between the (truncated) effective theory and the exact result for the fully connected 2^M-KREM with $M = 3$: Mean overlap $\langle p \rangle$ with a reference equilibrium configuration versus the inverse temperature β. The black curve (circles) represents the exact solution [Eqs. (A4) and (A6)], whereas the red curve (squares) corresponds to the prediction of the effective random-field + random-bond Ising Hamiltonian, with the coupling constants of the multi-body interactions (up to 4-body) numerically extracted from Fig. 3 as discussed in the text and the covariances of the random variables given in Eq. (22). The counterpart of the overlaps in the disordered Ising model is the magnetization $(1 + m)/2$, where $m = \overline{\langle \sigma_i \rangle}$ [see Appendix B and Eqs. (B2)]. The effective disordered theory has been derived by using the annealed approximation to average over the random energies, which is in principle justified only above T_K and at T_K (which includes the jump of $\langle p \rangle$). The blue curve (diamonds) shows the prediction of the effective theory below T_K with an additional approximation, i.e., when the coupling constants and the covariances are obtained by continuing the results obtained via the annealed approximation in the low-temperature phase.

Still, the dominant term that controls the transition in the effective theory is expected to be given by the competition between the ferromagnetic tendency of the interactions and the fluctuations of the random fields, and therefore to display RFIM-like behavior. The effect of the random bonds could become important if $\sqrt{\overline{\delta J_{ij}^2}} \gg J_2$, as this could generate a spin-glass-like behavior, as advocated in Ref. [27]. Yet, this possibility is excluded for the structural-glass model considered here for any $M > 1$, since $\sqrt{\overline{\delta J_{ij}^2}}/J_2 \sim 1/\sqrt{M}$.)

We stress a key difference between the 2^M-KREM with M finite and the REM. Despite being a fully connected mean-field model, the former is indeed such that the fluctuations associated with the effective disorder coming from the reference configuration give a contribution to the thermodynamics that is of the same order as that of the the average part. The difference of course disappears when $M \to \infty$, and disorder-related fluctuations are then subdominant as they scale as $\sqrt{M}$ for the 2^M-KREM and $\sqrt{N}$ for the REM whereas the average contributions scale as M and N, respectively.

IV. DISCUSSION

A. Beyond mean-field: Illustration of the difficulties

We now want to illustrate some of the difficulties that one encounters when trying to generalize the procedure developed in the preceding sections to finite-dimensional glass-forming systems close to the putative thermodynamic glass transition. To make the presentation more concrete we focus on a paradigmatic spin model of structural glass, the spherical p-spin model and consider its Kac extension that allows one to go one step beyond the conventional mean-field limit by taking into account spatially heterogeneous solutions.[23,68]

The model is defined by the following Hamiltonian:

$$\mathcal{H} = -\sum_{i_1\cdots i_p \in \Lambda} J_{i_1\cdots i_p}\sigma_{i_1}\cdots\sigma_{i_p} \tag{23}$$

where the spin variables satisfy the spherical constraint and the coupling constants $J_{i_1\cdots i_p}$ are i.i.d. Gaussian variables with variance

$$\overline{J^2_{i_1\cdots i_p}} = \frac{1}{r_0^{pd}}\sum_{k\in\Lambda}\psi\left(\frac{|k-i_1|}{r_0}\right)\cdots\psi\left(\frac{|k-i_p|}{r_0}\right), \tag{24}$$

where the function $\psi(r)$ is well-behaved and decays on a scale of order O(1). The Kac limit consists in considering the limit of a large interaction range, $r_0 \to \infty$.

The action $\mathcal{S}[\{p_a, q_{ab}\}]$ for the overlaps $\{p_a\}$ between the reference and the "constrained" replicas and for the overlaps $\{q_{ab}\}$ among the constrained replicas can then be obtained by standard methods and is given at large distance by[23]

$$\mathcal{S}[\{q_{\alpha\beta}\}] \approx r_0^d \int_x \left\{\frac{\beta c}{2}\sum_{\alpha\beta\neq}(\partial_x q_{\alpha\beta}(x))^2 - \frac{\beta^2}{4}\sum_{\alpha\beta\neq}q_{\alpha\beta}(x)^p - \frac{1}{2}\mathrm{Tr}\log[I + U(\{q_{\alpha\beta}(x)\})]\right\}, \tag{25}$$

where I is the identity and U an $(n+1)\times(n+1)$ matrix with all diagonal elements equal to 0, $U_{0a} = q_{0a} = p_a$, $U_{a0} = q_{a0} = p_a$, and $U_{ab} = q_{ab}$ for $a \neq b$. As before, Greek letters are used for the $n+1$ copies of the original system, including the reference equilibrium configuration $\alpha = 0$, whereas Latin ones are reserved for the n replicas other than the reference one. The action for the overlap fields p_a is then obtained by integrating out the overlaps q_{ab},

$$e^{-S_{\mathrm{rep}}[\{p_a\}]} \propto \int\prod_{ab\neq}\mathcal{D}q_{ab}\,e^{-\mathcal{S}[\{p_a,q_{ab}\}]}. \tag{26}$$

where, in the Kac limit, the integral over the q_{ab}'s can be performed via a saddle-point calculation, i.e.,

$$S_{\mathrm{rep}}[\{p_a\}] = \mathrm{cst} + \mathcal{S}[\{p_a, q_{ab}^*\}], \tag{27}$$

with

$$\left.\frac{\partial\mathcal{S}[\{p_a, q_{ab}\}]}{\partial q_{ab}(x)}\right|^* = 0. \tag{28}$$

Above the thermodynamic glass transition at T_K, the first cumulant is then given by (see Ref. [32] for a detailed derivation)

$$\mathcal{S}_1[p_1] = r_0^d\int_x\left\{\beta c(\partial_x p_1(x))^2 - \frac{\beta c}{2}(\partial_x q^*(x))^2 - \frac{\beta^2}{4}\left[2p_1(x)^p - q^*(x)^p\right] + \frac{p_1(x)^2 - q^*(x)}{2[1-q^*(x)]} - \frac{1}{2}\log[1-q^*(x)]\right\}, \tag{29}$$

where $q^*(x)$ satisfies the following saddle-point equation:

$$\beta c\partial_x^2 q^*(x) + \frac{\beta^2 p}{4}q^*(x)^{p-1} = \frac{1}{2}\frac{q^*(x) - p_1(x)^2}{[1-q^*(x)]^2}. \tag{30}$$

Unfortunately, the same letter p is used to denote the number of spins involved in the interactions in Eq. (23) (this is a widespread notation that it would awkward to change) and the overlap with the reference configuration. To avoid too much confusion, our convention is that p always comes with a replica index, p_a, when it refers to an overlap.

The solution of Eq. (30) is in general far from being trivial. When the thermodynamics of the system is dominated by specific uniform profiles $p_a(x) = p_a$ and smooth variations around them (e.g., in the Kac limit, $r_0 \to \infty$), one can first solve Eq. (30) for uniform p_1 and q and then consider the first nonzero gradient corrections about the uniform solution. More generally, for a large but finite r_0, gradient expansions may not be enough and one needs in principle to take into account the contribution coming from nonuniform profiles $p_a(x)$. However, even assuming that one can use the saddle-point equation (30) to compute the integral over the q_{ab}'s, one immediately sees that solving Eq. (30) for a generic nonuniform $p_1(x)$ is an impossible task. Approximations are therefore required.

A chief obstacle to devising a simple approximation for solving Eq. (30) comes from the presence of specific spatial correlations that arise in the form of point-to-set correlations. At high enough temperature, but still below the

dynamical transition temperature T_d, these correlations are short-ranged and one can therefore proceed as in the Kac limit by considering uniform configurations and smooth variations that can be described by an expansion in spatial derivatives. On the contrary when one approaches T_K one anticipates long-ranged, possibly diverging, point-to-set correlations. Consider then for instance a configuration of $p_1(x)$ that is zero almost everywhere except for a finite density of localized regions of space where it has a high value corresponding to the glassy metastable minimum. One expects that $q(x)$ will also be equal to the metastable high-overlap value in the these localized regions. What is then the value of $q(x)$ in the regions where $p_1(x) = 0$? If there were no point-to-set correlations, $q(x)$ would be zero. However, as soon as the typical distance between high-overlap regions becomes less than the point-to-set correlation length, the constraint due to the high-overlap regions will suddenly force the overlap $q(x)$ in the rest of the system to take a nontrivial value distinct from zero. At the mean-field T_K (in the Kac limit) the point-to-set correlation length is infinite[68] and even a very dilute concentration of localized high-overlap regions will induce a nontrivial finite features in the profile $q(x)$ everywhere. This example shows how nonperturbative changes of $q(x)$ can be generated by minute changes in the profile of $p_1(x)$, leading as a result to intrinsically nonlocal and long-ranged contributions to $\mathcal{S}_1[p_1]$. This is the essence of the difficulty associated with deriving an effective theory for the p_a's specifically when the system is near or below the mean-field thermodynamic glass transition. We will address this issue in the following paper.[1]

B. Specific features of the effective disorder

Another point which is worth discussing concerns the properties of the effective disorder found when mapping the statistics of the fluctuations of the overlaps $\{p_a\}$ onto a disordered Ising model. To illustrate this we again consider the Kac spherical p-spin model and we focus on the second cumulant (2-replica action) $\mathcal{S}_2[p_1, p_2]$. In the limit $r_0 \to \infty$, we find after some manipulations (see Ref. [32] for a detailed derivation) that

$$\mathcal{S}_2[p_1, p_2] = \int_x \left\{ -\beta c (\partial_x q_{12}(x))^2 + \frac{\beta c}{2}\left[\left(\partial_x q_1^{[0,1]}(x)\right)^2 + \left(\partial_x q_2^{[1,0]}(x)\right)^2 \right] + \frac{\beta^2}{2}[q_{12}(x)]^p - \frac{[q_{12}(x) - p_1(x)p_2(x)]^2}{2[1-q_1(x)][1-q_2(x)]} \right\},$$
(31)

where $q_{12}(p_1, p_2)$, $q_1(p_1)$, $q_2(p_2)$, $q_1^{[0,1]}(p_1, p_2)$, and $q_2^{[1,0]}(p_1, p_2)$ satisfy the following saddle-point equations:

$$2\beta c \partial_x^2 q_{12}(x) + \frac{\beta^2 p}{2}[q_{12}]^{p-1} = \frac{q_{12} - p_1 p_2}{(1-q_1)(1-q_2)},$$

$$\beta c \partial_x^2 q_a(x) + \frac{\beta^2 p}{4} q_a^{p-1} = \frac{1}{2} \frac{q_a - p_a^2}{[1-q_a]^2},$$
(32)

$$-\beta c \partial_x^2 q_a^{[\cdot,\cdot]_a} + \left[p(p-1)\frac{\beta^2}{4} q_a^{p-2} - \frac{1 - 2p_a^2 + q_a}{2(1-q_a)^3} \right] q_a^{[\cdot,\cdot]_a} = -\frac{(q_{12} - p_a p_b)^2}{2(1-q_a)^2(1-q_b)},$$

with $a = 1, 2$, $[\cdot, \cdot]_1 = [0, 1]$ and $[\cdot, \cdot]_2 = [1, 0]$.

As already discussed, in the Kac limit on which we focus here, one only needs to consider uniform profiles $p_a(x) = p_a$ and smooth variations around them. One can then first solve the saddle-point equations in Eq. (32) for uniform p_1 and p_2 and then calculate the first nonzero gradient correction about the uniform solution. In the following we focus on the local part of the second cumulant $\mathcal{S}_2$ which is obtained by considering uniform overlaps, as it is sufficient to illustrate our point.

The local part of the functional, $\mathcal{S}_2(p_1, p_2)/N$, is simply related to the second cumulant of a (delta correlated in space) random potential. Its second derivative $\Delta_2(p_1, p_2) = \partial_{p_1}\partial_{p_2}[\mathcal{S}_2(p_1, p_2)/N]$ then represents the second cumulant of the derivative of the random potential, i.e., the variance of an effective random force or random source conjugate to the overlap field. (After passing to the magnetic representation in terms of the magnetizations $m_a = 2p_a - 1$, Δ_2 represents, up to a factor 4, the variance of the effective random field.) It is instructive to analyze the shape of $\mathcal{S}_2(p_1, p_2)/N$ and $\Delta_2(p_1, p_2)$. Note first that $\mathcal{S}_2 = 0$ when either $p_1 = 0$ or $p_2 = 0$, which implies that the effective disorder vanishes in the liquid phase, as one could anticipate on the basis of intuitive arguments. For the Kac spherical p-spin model one also has $\Delta_2(p_1, p_2) = 0$ when either $p_1 = 0$ or $p_2 = 0$. For illustration, we plot $\Delta_2(p_1, p_2)$ when $p_1 = p_2$ in Fig. 5. It is zero in $p_1 = 0$, as announced, and grows as p_1 increases to be strictly positive at the value $p_\star$ of the glassy metastable minimum.

The property that the variance of the effective random field $\Delta_2(p_1, p_2) = 0$ when $p_1 = 0$ or $p_2 = 0$ is not true, however, for the REM: From Eq. (9), one instead obtains a constant variance $\Delta_2(p_1, p_2) = \beta^2 N/4$. In fact having a zero or nonzero value of Δ_2 in the liquid minimum is related to the details of the microscopic description. It is indeed easy to see that a disordered model with quenched disorder in the couplings only, as the Kac spherical p-spin model,
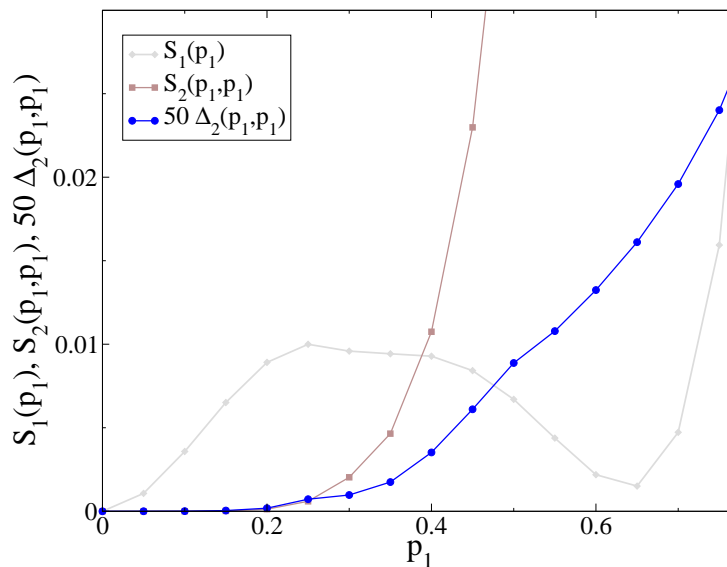
FIG. 5: Local part of the variance of the effective random field, $\Delta_2(p_1, p_2) = \partial_{p_1}\partial_{p_2} S_2(p_1, p_2)$, and of the 2-replica action, $S_2(p_1, p_2)/N$, for equal arguments $p_1 = p_2$ as a function of p_1 (blue curve, circles) for the p-spin model in the Kac limit with $p = 3$ and $\beta = 1.7$ (such that $\beta_d < \beta < \beta_K$). Δ_2 is zero in the liquid minimum ($p_1 = 0$), grows as p_1 is increased, so that it is strictly positive at the metastable glassy minimum. $S_2(p_1, p_1)$ behaves similarly. We also display the local part of the first cumulant $S_1(p_1)/N$ given in Eq. (29), to show the position of the secondary minimum around $p_\star \approx 0.65$: grey curve and diamonds (note that the region around the barrier is not properly described by the present RS solution but this is irrelevant for our illustrative purpose). Δ_2 has been multiplied by 50 to plot the three curves on the same scale.

leads to an effective theory with $\Delta_2 = 0$ in the liquid, whereas if some local (on-site) quenched disorder is present at the microscopic level, the corresponding effective theory is characterized by $S_2 = 0$ but $\Delta_2 > 0$ for either p_1 or p_2 equal to zero. There is actually a subtlety when considering glassy models in which the overlap degrees of freedom p_a^i are hard binary variables taking only values 0 and 1, as in the case of the REM and of the 2^M-KREM (but not of the Kac spherical p-spin model). What is, for instance, the counterpart for hard binary variables of a model with continuous variables and disorder in the interactions such that $S_2(p_1, p_2)/N = (\Delta/4)p_1^2 p_2^2$? Simply replacing the p_a's by hard variables (on a lattice) with $p_a^2 = p_a$ yields $S_2(p_1, p_2)/N = (\Delta/4)p_1 p_2$. Whereas $\Delta_2(p_1, p_2) = \Delta p_1 p_2$ in the former case and vanishes when either p_1 or p_2 equals zero, it is equal to a constant, $\Delta_2(p_1, p_2) = \Delta/4$, in the latter case. For hard binary variables on a lattice the only way that the local part of the second cumulant S_2 does not simply reduce to a term proportional to $p_1 p_2$ is that in the Ising (magnetic) representation, in addition to a local random-field term, random-bond disorder with cross-correlations with the random field is present. This is indeed the case for the 2^M-KREM, see Eq. (19).[69]

V.   CONCLUDING REMARKS

In this paper we have presented the derivation of a 2-state random effective theory which describes the fluctuations of what is thought to be the relevant order parameter for glassy systems, *i.e.,* the overlap field with a reference equilibrium configuration, close to the putative thermodynamic glass transition temperature. We have focused on archetypal mean-field models for the glass transition, in particular the Random Energy Model (REM)[33] and its version with a finite number of states[34] on a fully connected lattice (2^M-KREM). The effective Hamiltonian for mean-field models can in principle be worked out without resorting to any approximation. We have shown that the statistics of the fluctuations of the overlap with a reference configuration for the REM are described by an Ising variable $\sigma = \pm 1$ (corresponding to high and low overlap with the reference configuration) subjected to a random field whose average is of order N, and vanishes exactly at T_K, and whose fluctuations are of order $\sqrt{N}$. The effective theory for the fully connected 2^M-KREM is richer, and it is given by a random-bond + random-field Ising model with multi-body interactions and higher-order random terms.

We argue that the mapping is very general and should apply (possibly with some minor model-dependent adjustments) to any mean-field glassy model in the same "universality class" with a complex free-energy landscape appearing between an upper dynamical glass transition and a lower thermodynamical (RFOT) glass transition. In fact, we have

shown that for mean-field models the first cumulant of the effective action can only be a function of the global overlap with the reference configuration, and its shape shown in Fig. 5 (although distinct) is similar to that of the Franz-Parisi potential.[16–18] In order to obtain the average part of the effective Hamiltonian one thus only needs to expand the first cumulant in powers of the global overlap and re-express the resulting terms as effective multi-body interactions. The same strategy can be straightforwardly generalized to the second (and higher) cumulant, which yields the fluctuations of the effective Hamiltonian. Although conceptually simple, the calculations are somehow long and tedious, except for the REM. In the case of the 2^M-KREM we have derived for illustration the analytic expressions of the effective coupling constants and of the second cumulants of the random terms when $M \gg 1$ and we have given a numerical recipe to compute them when M is finite.

It is worth stressing that while the effective disorder is found to be subdominant in the $N \to \infty$ limit for the REM, the fluctuations associated with the effective disorder in the case of the fully connected 2^M-KREM for finite M are, in the thermodynamic limit, of the same order as the average contribution. Contrary to a naive expectation, the effective disorder is relevant for mean-field fully connected models: The choice of the reference configuration leads to site-dependent fluctuations of the "local" effective configurational entropy and of the "local" effective surface tension (as sketched in fig. 2) which give a contribution to the thermodynamics at the leading order. One needs to integrate over the fluctuations of the local overlap. In order to do it properly, taking into account the effective disorder is crucial. This also highlights the main difference between $\mathcal{S}_1(c)$ and the Franz-Parisi potential: The former is the average of the action for the instantaneous local fluctuations of the overlap, whereas the latter is the thermodynamic potential associated to the global overlap.

We finally reiterate that the interest of deriving the effective theory for mean-field models of structural glasses which can of course be exactly solved by other means is twofold:

(i) It shows that an effective description in terms of a random-field + random-bond Ising model naturally emerges in a transparent and general way.

(ii) It justifies it on a quantitative basis. In fact, by allowing a direct comparison with exact results, it justifies the truncation of the effective theory to a limited number of multi-body interactions (typically, up to 4-body), a limited order of cumulants (typically, up to the second one) and of quenched random terms (typically, random fields and random bonds). This serves as a guide for the investigation of finite-dimensional glass-forming systems. In a renormalization-group perspective, all higher-order terms will then anyhow be generated by the further renormalization of the effective theory to obtain the full solution of the thermodynamics of the overlaps. Once the effective theory is established, this final step can be achieved by using all powerful nonperturbative means at our disposal, such as large-scale numerical simulations[70] or the functional renormalization group.[64,65]

(iii) It is relevant for real finite-dimensional liquids, where on a scale much larger than the microscopic length but still much smaller than the point-to-set correlation length the mean-field description is still expected to retain some validity. One can then construct the effective theory on this scale by taking the mean-field result as a starting point. This would lead to an effective Hamiltonian of the form given in Eq. (2). In order to derive the proper effective theory for finite-dimensional glass-formers near the putative thermodynamic glass transition, one then has to take into account the role of correlations on the scale of the (diverging) point-to-set length. As already stressed, this is the main issue to be solved to go beyond the mean-field description. We will tackle it in the following paper.[1]

Appendix A: The fully connected Kac-like REM with a finite number of states: Exact solution, Effective theory, and Variational approximation

This appendix is devoted to the analysis of the Kac-like version of the REM with 2^M states (the 2^M-KREM) on a fully connected lattice. The model, first introduced in Ref. [34], is defined as follows: We consider N sites and define a state variable $\mathcal{C}_i$ on each site i which can can take 2^M possible values, $\mathcal{C}_i = 1, \ldots, 2^M$. For each pair of sites (i,j) we define the couplings $E_{ij}(\mathcal{C}_i, \mathcal{C}_j)$, which are i.i.d. Gaussian random variables such that $\overline{E_{ij}(\mathcal{C}_i, \mathcal{C}_j)} = 0$ and $\overline{E_{ij}(\mathcal{C}_i, \mathcal{C}_j) E_{ij}(\mathcal{C}'_i, \mathcal{C}'_j)} = M \delta_{\mathcal{C}_i, \mathcal{C}'_i} \delta_{\mathcal{C}_j, \mathcal{C}'_j}$. The Hamiltonian of the system is then given by $\mathcal{H} = 1/(2\sqrt{N}) \sum_{i \neq j} E_{ij}(\mathcal{C}_i, \mathcal{C}_j)$. In the following, we begin by working out the exact solution of the model by using the standard replica approach and a 1-RSB ansatz.

1. Exact solution using the standard replica approach

In order to compute the free energy of the system we use the replica trick,

$$\overline{Z^n} = \overline{\sum_{\{\mathcal{C}_i^\alpha\}} \exp\left(-\frac{\beta}{2\sqrt{N}}\sum_{i\neq j,\alpha} E_{ij}(\mathcal{C}_i^\alpha,\mathcal{C}_j^\alpha)\right)} = e^{nNM\beta^2/8}\sum_{\{\mathcal{C}_i^\alpha\}}\exp\left[\frac{M\beta^2}{8N}\sum_{\alpha\neq\beta}\left(\sum_i \delta_{\mathcal{C}_i^\alpha,\mathcal{C}_i^\beta}\right)^2\right]. \tag{A1}$$

A simple calculation shows that in the Kac limit ($M\to\infty$) the model has a (RFOT) glass transition at an inverse temperature $\beta_K = \sqrt{8\ln 2}$ (see Sec. III B). We expect that at finite M the transition, if present, will be located at a lower temperature. After performing $n(n-1)/2$ Hubbard-Stratonovich trasformations Eq. (A1) can be rewritten as

$$\overline{Z^n} = e^{nNM\beta^2/8}\left(\frac{NM\beta^2}{8\pi}\right)^{n(n-1)/2}\int\prod_{\alpha<\beta}\mathrm{d}q_{\alpha\beta}\, e^{-NA[q_{\alpha\beta}]},$$

where

$$A[q_{\alpha\beta}] = \frac{M\beta^2}{8}\sum_{\alpha\neq\beta}q_{\alpha\beta}^2 - \ln Z_1[q_{\alpha\beta}], \qquad \text{with} \qquad Z_1[q_{\alpha\beta}] = \sum_{\{\mathcal{C}^\alpha\}}e^{\frac{M\beta^2}{4}\sum_{\alpha\neq\beta}\delta_{\mathcal{C}^\alpha,\mathcal{C}^\beta}q_{\alpha\beta}}. \tag{A2}$$

The saddle-point equations trivially give $q_{\alpha\beta} = \langle\delta_{\mathcal{C}^\alpha,\mathcal{C}^\beta}\rangle_1$, where the average is computed with the single-site Hamiltonian $\mathcal{H}_1$ defined from the single-site partition function in Eq. (A2) by $Z_1[q_{\alpha\beta}] = \mathrm{Tr}\, e^{-\mathcal{H}_1}$. Note that in the following we will make repeated use of the following identity:

$$\delta_{\mathcal{C}^\alpha,\mathcal{C}^\beta} = \sum_{\mathcal{C}=1}^{2^M}\delta_{\mathcal{C}^\alpha,\mathcal{C}}\,\delta_{\mathcal{C}^\beta,\mathcal{C}}. \tag{A3}$$

a. The Replica Symmetric solution

We first consider a replica symmetric (RS) ansaz for the matrix $q_{\alpha\beta}$. Using Eq. (A3) the single-site Hamiltonian $\mathcal{H}_1$ can be rewritten as

$$\mathcal{H}_1 = n\frac{M\beta^2 q_0}{4} - \frac{M\beta^2 q_0}{4}\sum_{\mathcal{C}}\sum_{\alpha,\beta}\delta_{\mathcal{C}^\alpha,\mathcal{C}}\,\delta_{\mathcal{C}^\beta,\mathcal{C}}.$$

We now introduce 2^M gaussian integrals to decouple the sum over replicas,

$$e^{\frac{M\beta^2 q_0}{4}\sum_{\mathcal{C}}\sum_{\alpha,\beta}\delta_{\mathcal{C}^\alpha,\mathcal{C}}\,\delta_{\mathcal{C}^\beta,\mathcal{C}}} = \int\prod_{\mathcal{C}=1}^{2^M}\left[\frac{\mathrm{d}z_{\mathcal{C}}}{\sqrt{2\pi}}e^{-z_{\mathcal{C}}^2/2}\right]\exp\left(\sqrt{\frac{M\beta^2 q_0}{2}}\sum_{\mathcal{C}}\sum_\alpha\delta_{\mathcal{C}^\alpha,\mathcal{C}}\,z_{\mathcal{C}}\right).$$

The n replicas are now totally decoupled and the trace over $\{\mathcal{C}^\alpha\}$ is given by the product of n independent traces over single replicas,

$$Z_1[q_0] = e^{-n\frac{M\beta^2 q_0}{4}}\int\prod_{\mathcal{C}=1}^{2^M}\left[\frac{\mathrm{d}z_{\mathcal{C}}}{\sqrt{2\pi}}e^{-z_{\mathcal{C}}^2/2}\right]\left[\tilde{Z}(q_0,\{z_{\mathcal{C}}\})\right]^n, \qquad \text{where} \qquad \tilde{Z}(q_0,\{z_{\mathcal{C}}\}) = \sum_{\mathcal{C}}\exp\left(\sqrt{\frac{M\beta^2 q_0}{2}}z_{\mathcal{C}}\right).$$

In the $n\to 0$ limit one has

$$-\ln Z_1[q_0] = n\frac{M\beta^2 q_0}{4} - n\int\prod_{\mathcal{C}=1}^{2^M}\left[\frac{\mathrm{d}z_{\mathcal{C}}}{\sqrt{2\pi}}e^{-z_{\mathcal{C}}^2/2}\right]\ln\tilde{Z}(q_0,\{z_{\mathcal{C}}\}),$$

and the RS free energy per site reads in the thermodynamic limit

$$f(q_0) = -\frac{M\beta^2}{8} - \frac{M\beta^2}{8}q_0^2 + \frac{M\beta^2}{4}q_0 - \int\prod_{\mathcal{C}=1}^{2^M}\mathcal{D}z_{\mathcal{C}}\ln\tilde{Z}(q_0,\{z_{\mathcal{C}}\}),$$

where $\mathcal{D}x = e^{-x^2/2}dx/\sqrt{2\pi}$. By taking the derivative of the free energy with respect to q_0 we obtain the following saddle point equation:

$$q_0 = 1 - \sqrt{\frac{2}{M\beta^2 q_0}} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_\mathcal{C} \left[\sum_\mathcal{C} z_\mathcal{C}\, e^{\sqrt{\frac{M\beta^2 q_0}{2}} z_\mathcal{C}} \middle/ \sum_\mathcal{C} e^{\sqrt{\frac{M\beta^2 q_0}{2}} z_\mathcal{C}} \right], \tag{A4}$$

which can be easily solved numerically.

b.   The 1-RSB solution

In the following we introduce a 1-step replica-symmetry breaking (1-RSB) ansatz[67] for the matrix $q_{\alpha\beta}$ by considering n/m blocks of m replicas such that $q_{\alpha\beta} = q_1$ if (α, β) belong to the same block and $q_{\alpha\beta} = q_0$ if (α, β) belong to different blocks. From Eq. (A3) the single-site Hamiltonian $\mathcal{H}_1$ can be rewritten as

$$\mathcal{H}_1 = n\frac{M\beta^2 q_1}{4} - \frac{M\beta^2 q_0}{4} \sum_\mathcal{C} \sum_{\alpha,\beta} \delta_{\mathcal{C}^\alpha,\mathcal{C}}\, \delta_{\mathcal{C}^\beta,\mathcal{C}} - \frac{M\beta^2(q_1-q_0)}{4} \sum_\mathcal{C} \sum_{\alpha,\beta}^{\star} \delta_{\mathcal{C}^\alpha,\mathcal{C}}\, \delta_{\mathcal{C}^\beta,\mathcal{C}},$$

where $\sum^{\star}$ is the sum over all possible couples (α,β) belonging to the same block. We introduce 2^M gaussian integrals to decouple the first sum over replicas in the above expression. The n/m blocks of replicas are now totally decoupled and the trace over $\{\mathcal{C}^\alpha\}$ is given by the product of n/m independent traces over the replica indices of each block. We thus find

$$Z_1[q_1, q_0, m] = e^{-n\frac{M\beta^2 q_1}{4}} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_\mathcal{C} \left[Z_{\text{block}}(q_1, q_0, m, \{z_\mathcal{C}\}) \right]^{\frac{n}{m}},$$

where

$$Z_{\text{block}}(q_1, q_0, m, \{z_\mathcal{C}\}) = \sum_{\{\mathcal{C}^\alpha\}^{\star}} \exp\left(\sqrt{\frac{M\beta^2 q_0}{2}} \sum_\mathcal{C} \sum_\alpha^{\star} \delta_{\mathcal{C}^\alpha,\mathcal{C}}\, z_\mathcal{C} + \frac{M\beta^2(q_1-q_0)}{4} \sum_\mathcal{C} \sum_{\alpha,\beta}^{\star} \delta_{\mathcal{C}^\alpha,\mathcal{C}}\, \delta_{\mathcal{C}^\beta,\mathcal{C}} \right). \tag{A5}$$

The trace over $\{\mathcal{C}^\alpha\}^{\star}$ involves only one block of replicas and $\alpha = 1,\ldots,m$. In the limit $n \to 0$, $[Z_{\text{block}}(q_1, q_0, m, \{z_\mathcal{C}\})]^{n/m} \approx 1 + (n/m) \ln Z_{\text{block}}(q_1, q_0, m, \{z_\mathcal{C}\})$, yielding

$$-\ln Z_1[q_1, q_0, m] = n\frac{M\beta^2 q_1}{4} - \frac{n}{m} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_\mathcal{C} \ln Z_{\text{block}}(q_1, q_0, m, \{z_\mathcal{C}\}).$$

The second sum in the right-hand side of Eq. (A5) can again be decoupled by introducing 2^M additional gaussian integrals,

$$Z_{\text{block}}(q_1, q_0, m, \{z_\mathcal{C}\}) = \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_\mathcal{C} \left[\tilde{Z}(q_1, q_0, m, \{z_\mathcal{C}\}, \{w_\mathcal{C}\}) \right]^m,$$

where

$$\tilde{Z}(q_1, q_0, m, \{z_\mathcal{C}\}, \{w_\mathcal{C}\}) = \sum_\mathcal{C} \exp\left(\sqrt{\frac{M\beta^2 q_0}{2}} z_\mathcal{C} + \sqrt{\frac{M\beta^2(q_1-q_0)}{2}} w_\mathcal{C} \right).$$

Finally, we obtain the free energy per site in the thermodynamic limit as

$$f(q_1, q_0, m) = -\frac{M\beta^2}{8} + \frac{M\beta^2}{8}\left[(m-1)q_1^2 - mq_0^2 \right] + \frac{M\beta^2}{4} q_1$$

$$- \frac{1}{m} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_\mathcal{C} \ln\left\{ \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_\mathcal{C} \left[\tilde{Z}(q_1, q_0, m, \{z_\mathcal{C}\}, \{w_\mathcal{C}\}) \right]^m \right\}.$$

The saddle-point equations are obtained by imposing that the derivatives of the free energy with respect to q_1, q_0, and m vanish. Taking the derivative with respect to q_1 leads to

$$(m-1)q_1 + 1 - \sqrt{\frac{2}{M\beta^2(q_1-q_0)}} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_{\mathcal{C}} \times$$

$$\left\{ \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_{\mathcal{C}} \left[[\tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\})]^{m-1} \left[\sum_{\mathcal{C}} w_{\mathcal{C}} \exp\left(\sqrt{\frac{M\beta^2 q_0}{2}} z_{\mathcal{C}} + \sqrt{\frac{M\beta^2(q_1-q_0)}{2}} w_{\mathcal{C}} \right) \right] \right] \right.$$

$$\left. \Big/ \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_{\mathcal{C}} \left[\tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\}) \right]^m \right\} = 0 \,.$$

The derivative with respect to q_0 gives

$$-mq_0 - \sqrt{\frac{2}{M\beta^2}} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_{\mathcal{C}} \times$$

$$\left\{ \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_{\mathcal{C}} \left[[\tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\})]^{m-1} \left[\sum_{\mathcal{C}} \left(\frac{z_{\mathcal{C}}}{\sqrt{q_0}} - \frac{w_{\mathcal{C}}}{\sqrt{q_1-q_0}} \right) \exp\left(\sqrt{\frac{M\beta^2 q_0}{2}} z_{\mathcal{C}} + \sqrt{\frac{M\beta^2(q_1-q_0)}{2}} w_{\mathcal{C}} \right) \right] \right] \right.$$

$$\left. \Big/ \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_{\mathcal{C}} \left[\tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\}) \right]^m \right\} = 0 \,.$$

Note that for $m=1$ this equation gives back Eq. (A4). Finally, the derivative with respect to m gives

$$\frac{M\beta^2}{8}(q_1^2 - q_0^2) + \frac{1}{m^2} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_{\mathcal{C}} \ln\left\{ \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_{\mathcal{C}} \left[\tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\}) \right]^m \right\}$$

$$- \frac{1}{m} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_{\mathcal{C}} \left\{ \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_{\mathcal{C}} \left[[\tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\})]^m \ln \tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\}) \right] \right. \tag{A6}$$

$$\left. \Big/ \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_{\mathcal{C}} \left[\tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\}) \right]^m \right\} = 0 \,.$$

In order to find the thermodynamic glass transition, one thus needs to solve numerically Eq. (A4), which yields the value of q_0, and Eq. (A6) for $m=1$, which gives the value of q_1 such that $m=1$ is an extremum of the free-energy, and finally check whether $q_1 \neq q_0$. After some simple algebra Eq. (A6) for $m=1$ can be rewritten in a simpler form:

$$q_1^2 - q_0^2 + 2(q_1-q_0) + \frac{8}{M\beta^2} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_{\mathcal{C}} \ln\left(\sum_{\mathcal{C}} e^{\sqrt{\frac{M\beta^2 q_0}{2}} z_{\mathcal{C}}} \right)$$

$$- \frac{8 e^{-M\beta^2(q_1-q_0)/4}}{M\beta^2} \int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}z_{\mathcal{C}} \frac{\int \prod_{\mathcal{C}=1}^{2^M} \mathcal{D}w_{\mathcal{C}} \left[\tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\}) \ln \tilde{Z}(q_1,q_0,m,\{z_{\mathcal{C}}\},\{w_{\mathcal{C}}\}) \right]}{\sum_{\mathcal{C}} e^{\sqrt{\frac{M\beta^2 q_0}{2}} z_{\mathcal{C}}}} = 0 \,.$$

The numerical solutions of the 1-RSB equations for $M=3$ corresponds to the black curve (circles) in Fig. 4, showing a transition (RFOT) for $\beta_K \approx 2.5$.

2. Construction of the effective theory

In the following we apply the procedure described in Sec. III A to construct the effective theory of the model. To this aim we consider $n+1$ replicas of the system and compute the replicated action for a fixed overlap field $\{p_a^i\}$ of the replicas $a = 1,\ldots,n$ with a given reference configuration $\{\mathcal{C}_i^0\}$. Note that $p_a^i = 1$ only if $\mathcal{C}_i^a = \mathcal{C}_i^0$ and is zero otherwise. As already mentioned, we will consider the temperature range $T_d \leq T \leq T_K$, where we can use the

annealed approximation to average over the random energies E_{ij}. The replicated action is given in Eq. (11) of the main text, where the Kronecker δ's in the exponential can be rewritten in terms of the overlap variables as in Eq. (12). As discussed in the main text, if $\mathcal{C}_i^a = \mathcal{C}_i^0$ and $\mathcal{C}_i^b = \mathcal{C}_i^0$ (i.e., $p_a^i = p_b^i = 1$), then $\mathcal{C}_i^b = \mathcal{C}_i^a$. Similarly, if $\mathcal{C}_i^a = \mathcal{C}_i^0$ and $\mathcal{C}_i^b \neq \mathcal{C}_i^0$ (i.e., $p_a^i = 1$ and $p_b^i = 0$) then $\mathcal{C}_i^b \neq \mathcal{C}_i^a$. The same is true, of course, if $\mathcal{C}_i^a \neq \mathcal{C}_i^0$ and $\mathcal{C}_i^b = \mathcal{C}_i^0$. The only undetermined case corresponds to $\mathcal{C}_i^a \neq \mathcal{C}_i^0$ and $\mathcal{C}_i^b \neq \mathcal{C}_i^0$.

a. The average effective action: First cumulant

We first focus on the first cumulant (1-replica action) $\mathcal{S}_1[\{p^i\}]$. It is then sufficient to set all replica fields equal, $p_a^i = p^i \; \forall a = 1, \cdots, n$ and $\forall i$, keep only the term of order n in the expression of $\mathcal{S}_{\mathrm{rep}}[\{p_a^i\}]$, and take the limit $n \to 0$ in the end, as in the standard replica trick. In order to do this we set the overlap profile with the reference configuration for all replicas to be 1 on the first cN sites (i.e., $p_a^i = 1$ for $i = 1, \ldots, cN$, $\forall a$) and 0 on all the other $(1-c)N$ sites (i.e., $p_a^i = 0$ for $i = cN + 1, \ldots, N$, $\forall a$).

For the chosen overlap profile Eq. (12) becomes

$$\sum_{i \neq j} \left[1 + n + 2np^i p^j + \sum_{a \neq b} \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \delta_{\mathcal{C}_j^a, \mathcal{C}_j^b} \right] = (1+n)N(N-1) + 2ncN(cN-1) + n(n-1)cN(cN-1)$$

$$+ 2cN \sum_{a \neq b} \sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} + \sum_{a \neq b} \left[\left(\sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \right)^2 - \sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \right]$$

$$\approx (1+n)N^2 + n(n+1)c^2 N^2 + 2cN \sum_{a \neq b} \sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} + \sum_{a \neq b} \left(\sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \right)^2,$$

where in going to the last line we have thrown away all the sub-extensive diagonal ($i = j$) terms. The sum $\sum_i^\star$ represents the sum over the sites $i = cN + 1, \ldots, N$ where $p^i = 0$. Inserting this expression into Eq. (11) yields

$$e^{-n\mathcal{S}_1(c)} = \frac{e^{\frac{N\beta^2 M}{8} \left[1 + n + n(n+1)c^2 \right]}}{\overline{Z}} \sum_{\{\mathcal{C}_i^\alpha\}} e^{\frac{N\beta^2 M}{8} \left[2c \sum_{a \neq b} \frac{1}{N} \sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} + \sum_{a \neq b} \left(\frac{1}{N} \sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \right)^2 \right]} \prod_{a,i} \delta_{p_a^i, \delta_{\mathcal{C}_i^0, \mathcal{C}_i^a}}.$$

The sum over the reference configuration $\mathcal{C}_i^0$ simply gives 2^{NM}. On the first cN sites we have that $p^i = 1$ and then $\mathcal{C}_i^a = \mathcal{C}_i^0$ for all a. We thus obtain

$$e^{-n\mathcal{S}_1(c)} = \frac{e^{NM\left\{ \ln 2 + \frac{\beta^2}{8} \left[1 + n + n(n+1)c^2 \right] \right\}}}{\overline{Z}} \sum_{\{\mathcal{C}_i^a\}_\star} e^{\frac{\beta^2 M}{4} \sum_{a<b} \left[2c \sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} + \frac{1}{N} \left(\sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \right)^2 \right]},$$

where the trace $\sum_{\{\mathcal{C}_i^a\}_\star}$ represents the sum over all the $2^M - 1$ configurations $\mathcal{C}_i^a$ that are different from the reference one on the $(c-1)N$ sites where $p^i = 0$. One can now introduce the overlaps q_{ab} by performing the usual Hubbard-Stratonovich transformations,

$$e^{\frac{\beta^2 M}{4N} \sum_{a<b} \left(\sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \right)^2} = \left(\frac{N\beta^2 M}{4\pi} \right)^{n(n-1)/4} \int \prod_{a<b} \mathrm{d}q_{ab} \, e^{-\frac{N\beta^2 M}{4} \sum_{a<b} q_{ab}^2 + \frac{\beta^2 M}{2} \sum_{a<b} q_{ab} \sum_i^\star \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b}}.$$

At this point one can easily compute the trace over the configurations $\sum_{\{\mathcal{C}_i^a\}_\star}$, thanks to the fact that the sites are decoupled (the annealed partition function of the model in the denominator only yields an unimportant constant term):

$$e^{-n\mathcal{S}_1(c)} = e^{\frac{nN\beta^2 M}{8} \left[1 + (n+1)c^2 \right]} \left(\frac{N\beta^2 M}{4\pi} \right)^{n(n-1)/4} \int \prod_{a<b} \mathrm{d}q_{ab} \, e^{-NA[q_{ab}]},$$

$$A[q_{ab}] = \frac{M\beta^2}{4} \sum_{a<b} q_{ab}^2 - (1-c) \ln Z_1,$$

$$Z_1 = \sum_{\{\mathcal{C}^a\}_\star} e^{\frac{\beta^2 M}{2} \sum_{a<b} (q_{ab}+c) \delta_{\mathcal{C}^a, \mathcal{C}^b}}.$$

In the thermodynamic limit $(N \to \infty)$ the integral in the above expression can be performed via a saddle-point method, which gives $q_{ab} = (1-c)\langle \delta_{\mathcal{C}^a,\mathcal{C}^b}\rangle_1$, the average $\langle\cdots\rangle_1$ being performed with the single-site Hamiltonian $-\mathcal{H}_1 = \sum_{a<b}(q_{ab}+c)\delta_{\mathcal{C}^a,\mathcal{C}^b}$.

We now introduce a RS ansatz for the overlap matrix, $q_{ab} = q_0$. Using the identity (A3), $\delta_{\mathcal{C}^a,\mathcal{C}^b} = \sum_{\mathcal{C}}^{\star} \delta_{\mathcal{C}^a,\mathcal{C}}\,\delta_{\mathcal{C}^b,\mathcal{C}}$, the single-site Hamiltonian can be rewritten as

$$-\mathcal{H}_1 = -\frac{n\beta^2 M}{4}(q_0+c) + \frac{\beta^2 M}{4}(q_0+c)\sum_{\mathcal{C}}^{\star}\left(\sum_a \delta_{\mathcal{C}^a,\mathcal{C}}\right)^2,$$

The replicas can again be decoupled via Hubbard-Stratonovich transformations, yielding

$$\ln Z_1 = -\frac{n\beta^2 M}{4}(q_0+c) + \ln\int\prod_{\mathcal{C}}^{\star}\frac{\mathrm{d}z_{\mathcal{C}}}{\sqrt{2\pi}}e^{-\frac{z_{\mathcal{C}}^2}{2}}\left[\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_0+c)}{2}}z_{\mathcal{C}}}\right]^n.$$

In the $n \to 0$ limit one thus has

$$A[q_0] = \frac{n(n-1)\beta^2 M q_0^2}{8} + \frac{n(1-c)\beta^2 M(q_0+c)}{4} - n(1-c)\overline{\ln\left[\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_0+c)}{2}}z_{\mathcal{C}}}\right]},$$

where the average $\overline{[f(\vec{z}_{\mathcal{C}})]}$ is defined over the gaussian measure $\overline{[f(\vec{z}_{\mathcal{C}})]} \equiv \int\prod_{\mathcal{C}}^{\star}\left[\frac{\mathrm{d}z_{\mathcal{C}}}{\sqrt{2\pi}}e^{-z_{\mathcal{C}}^2/2}\right]f(\vec{z}_{\mathcal{C}})$. Putting all the results together in the $n \to 0$ limit (and neglecting subleading terms in the $N \to \infty$ limit), one gets Eq. (14) given in the main text, where the overlap q_0 must satisfy the self-consistent equation (15). The solution of such a saddle-point equation can be developed in powers of c as $q_0 \approx q_{0,0} + cq_{0,1} + c^2 q_{0,2} + \ldots$, which, when inserted back into Eqs. (15) and (14), allows one to obtain the exact expansion of $\mathcal{S}_1(c)$ in powers of c. In order to provide analytic expressions of the coupling constants K_n appearing in Eq. (13), we perform the expansion of $\mathcal{S}_1(c)$ in powers of c when $M \gg 1$. After expanding the exponentials of Eqs. (14) and (15) up to the eighth order in $\sqrt{M\beta^2(q_0+c)/2}$, we obtain

$$q_0 \approx \frac{1}{2^M} + \frac{M\beta^2-2}{2^{M+1}}c + \frac{M\beta^2(M\beta^2-4)}{2^{M+3}}c^2 + \frac{M^2\beta^4(M\beta^2-6)}{3\cdot 2^{M+4}}c^3 - \frac{M^3\beta^6}{3\cdot 2^{M+4}}c^4 + \ldots. \tag{A7}$$

Inserting this expression into Eq. (14) and re-expressing the powers of c (up to fourth order) as effective one-, two-, three-, and four-body interactions, we find an effective Hamiltonian of the form given in Eq. (13) with parameters given in Eq. (16).

b. Fluctuations of the effective action: Second cumulant

We now turn to the computation of the second cumulant (2-replica action) $\mathcal{S}_2[\{p_1^i, p_2^i\}]$. To do this we divide the n constrained replicas into two groups of n_1 and n_2 replicas respectively. The most generic overlap profile can be obtained by dividing the sites in four groups, denoted (1), (2), (12), and (0), such that on the $c_1 N$ sites belonging to the group (1) $p_1^i = 1$ and $p_2^i = 0$, on the $c_2 N$ sites belonging to the group (2) $p_1^i = 0$ and $p_2^i = 1$, on the $c_{12}N$ sites belonging to the group (12) $p_1^i = p_2^i = 1$, and on the $c_0 N$ sites belonging to the group (0) $p_1^i = p_2^i = 0$ (with $c_0 = 1 - c_1 - c_2 - c_{12}$). The second cumulant can be computed by keeping only the terms of order $n_1 n_2$ in the expression of the replicated action, and taking the limit $n_1, n_2 \to 0$ [see Eq. (4)]. The first terms of the right-hand side of Eq. (12) thus become (neglecting subleading corrections)

$$\sum_{i\neq j}\left[1 + n_1 + n_2 + 2\sum_a p_a^i p_a^j\right] \approx N^2\left[1 + n_1 + n_2 + 2n_1(c_1+c_{12})^2 + 2n_2(c_2+c_{12})^2\right].$$

On the other hand, the second term of the right-hand side of Eq. (12) reads (neglecting again subleading corrections)

$$\sum_{i\neq j}\sum_{a\neq b}\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}\delta_{\mathcal{C}_j^a,\mathcal{C}_j^b} \approx N^2\Big[(c_1+c_{12})^2 n_1(n_1-1) + (c_2+c_{12})^2 n_2(n_2-1) + 2c_{12}^2 n_1 n_2\Big]$$

$$+ 2(c_1+c_{12})N\sum_{a\neq b}^{(1)}\sum_i^{(2)+(0)}\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b} + 2(c_2+c_{12})N\sum_{a\neq b}^{(2)}\sum_i^{(1)+(0)}\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b} + 4c_{12}N\sum_a^{(1)}\sum_b^{(2)}\sum_i^{(0)}\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}$$

$$+ \sum_{a\neq b}^{(1)}\left(\sum_i^{(2)+(0)}\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}\right)^2 + \sum_{a\neq b}^{(2)}\left(\sum_i^{(1)+(0)}\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}\right)^2 + 2\sum_a^{(1)}\sum_b^{(2)}\left(\sum_i^{(0)}\delta_{\mathcal{C}_i^a,\mathcal{C}_i^b}\right)^2,$$

where $\sum_a^{(1),(2)}$ denotes the sum over the constrained replicas belonging, respectively, to the first group $(a = 1, \ldots, n_1)$ or to the second group $(a = 1 + n_1, \ldots, n_1 + n_2)$, while $\sum_i^{(s)}$ denotes the sum over i belonging to the s-th group of sites (with $s = 1$, 2, 0, or 12). We can now perform the sum over the reference configuration $\mathcal{C}_i^0$, which simply gives 2^{MN}. On the $c_1 N$ sites belonging to the group (1), $p_a^i = 1$ for all a in the first group of replicas and $p_a^i = 0$ for all a in the second group. Hence $\mathcal{C}_i^a = \mathcal{C}_i^0$ for all $a = 1, \ldots, n_1$, whereas $\mathcal{C}_i^a \neq \mathcal{C}_i^0$ for all $a = 1 + n_1, \ldots, n_1 + n_2$. Similarly, on the $c_2 N$ sites belonging to the group (2), $p_a^i = 0$ for all a in the first group of replicas and $p_a^i = 1$ for all a in the second group. Hence $\mathcal{C}_i^a \neq \mathcal{C}_i^0$ for all $a = 1, \ldots, n_1$, whereas $\mathcal{C}_i^a = \mathcal{C}_i^0$ for all $a = 1 + n_1, \ldots, n_1 + n_2$. On the $c_{12} N$ sites belonging to the group (12), $p_a^i = 1$ for all replicas. Thus $\mathcal{C}_i^a = \mathcal{C}_i^0$ for all $a = 1, \ldots, n_1 + n_2$. Finally, on the $c_0 N$ sites belonging to the group (0), $p_a^i = 0$ for all replicas. Thus $\mathcal{C}_i^a \neq \mathcal{C}_i^0$ for all $a = 1, \ldots, n_1 + n_2$. In consequence, the trace over the configuration $\sum_{\{\mathcal{C}_i^a\}}$ consists in summing over all possible $2^M - 1$ configurations different from the reference one for a belonging to the second group of replicas on the sites of group (1), to the first group of replicas on the sites of group (2), and on all replicas on the sites of group (0). In the following we will denote this sum as $\sum_{\{\mathcal{C}_i^a\}_\star}$.

At this point we introduce several overlaps, $q_{ab}^{[1]}$, $q_{ab}^{[2]}$, $q_{ab}^{[12]}$, via the usual Hubbard-Stratonovich transformations that allow one to decouple different sites,

$$e^{\frac{\beta^2 M}{4N} \sum_{a<b}^{(1)} \left(\sum_i^{(2)+(0)} \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \right)^2} = \left(\frac{N\beta^2 M}{4\pi} \right)^{n_1(n_1-1)/4} \int \prod_{a<b}^{(1)} dq_{ab}^{[1]} \, e^{-\frac{N\beta^2 M}{4} \sum_{a<b}^{(1)} \left(q_{ab}^{[1]} \right)^2 + \frac{\beta^2 M}{2} \sum_{a<b}^{(1)} q_{ab}^{[1]} \sum_i^{(2)+(0)} \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b}},$$

$$e^{\frac{\beta^2 M}{4N} \sum_{a<b}^{(2)} \left(\sum_i^{(1)+(0)} \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \right)^2} = \left(\frac{N\beta^2 M}{4\pi} \right)^{n_2(n_2-1)/4} \int \prod_{a<b}^{(2)} dq_{ab}^{[2]} \, e^{-\frac{N\beta^2 M}{4} \sum_{a<b}^{(2)} \left(q_{ab}^{[2]} \right)^2 + \frac{\beta^2 M}{2} \sum_{a<b}^{(2)} q_{ab}^{[2]} \sum_i^{(1)+(0)} \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b}},$$

$$e^{\frac{\beta^2 M}{4N} \sum_a^{(1)} \sum_b^{(2)} \left(\sum_i^{(0)} \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b} \right)^2} = \left(\frac{N\beta^2 M}{4\pi} \right)^{n_1 n_2/2} \int \prod_{a \in (1); b \in (2)} dq_{ab}^{[12]} \, e^{-\frac{N\beta^2 M}{4} \sum_a^{(1)} \sum_b^{(2)} \left(q_{ab}^{[12]} \right)^2 + \frac{\beta^2 M}{2} \sum_a^{(1)} \sum_b^{(2)} q_{ab}^{[12]} \sum_i^{(0)} \delta_{\mathcal{C}_i^a, \mathcal{C}_i^b}}.$$

Neglecting all the subleading and irrelevant terms, we can now rewrite the replicated action in the following way:

$$e^{-\mathcal{S}_{\rm rep}[c_1, c_2, c_{12}]} = e^{\frac{N\beta^2 M}{8} \left[n_1 + n_2 + n_1(n_1+1)(c_1+c_{12})^2 + n_2(n_2+1)(c_2+c_{12})^2 + 2c_{12}^2 n_1 n_2 \right]}$$

$$\times \int \prod_{a<b}^{(1)} dq_{ab}^{[1]} \prod_{a<b}^{(2)} dq_{ab}^{[2]} \prod_{a \in (1); b \in (2)} dq_{ab}^{[12]} \, e^{-NA[q_{ab}^{[1]}, q_{ab}^{[2]}, q_{ab}^{[12]}]},$$

$$A[q_{ab}^{[1]}, q_{ab}^{[2]}, q_{ab}^{[12]}] = \frac{M\beta^2}{4} \left[\sum_{a<b}^{(1)} \left(q_{ab}^{[1]} \right)^2 + \sum_{a<b}^{(2)} \left(q_{ab}^{[2]} \right)^2 + \sum_a^{(1)} \sum_b^{(2)} \left(q_{ab}^{[12]} \right)^2 \right] - c_1 \ln Z_1 - c_2 \ln Z_2 - (1 - c_1 - c_2 - c_{12}) \ln Z_0,$$

$$Z_1 = \sum_{\{\mathcal{C}^a\}_\star} e^{\frac{\beta^2 M}{2} \sum_{a<b}^{(2)} (q_{ab}^{[2]} + c_2 + c_{12}) \delta_{\mathcal{C}^a, \mathcal{C}^b}},$$

$$Z_2 = \sum_{\{\mathcal{C}^a\}_\star} e^{\frac{\beta^2 M}{2} \sum_{a<b}^{(1)} (q_{ab}^{[1]} + c_1 + c_{12}) \delta_{\mathcal{C}^a, \mathcal{C}^b}},$$

$$Z_0 = \sum_{\{\mathcal{C}^a\}_\star} e^{\frac{\beta^2 M}{2} \left[\sum_{a<b}^{(1)} (q_{ab}^{[1]} + c_1 + c_{12}) \delta_{\mathcal{C}^a, \mathcal{C}^b} + \sum_{a<b}^{(2)} (q_{ab}^{[2]} + c_2 + c_{12}) \delta_{\mathcal{C}^a, \mathcal{C}^b} + \sum_a^{(1)} \sum_b^{(2)} (q_{ab}^{[12]} + c_{12}) \delta_{\mathcal{C}^a, \mathcal{C}^b} \right]}.$$

$$(A8)$$

The integrals over the overlaps can be performed at the saddle point in the limit $N \to \infty$, which gives

$$q_{a_1 b_1}^{[1]} = c_2 \langle \delta_{\mathcal{C}^{a_1}, \mathcal{C}^{b_1}} \rangle_2 + (1 - c_1 - c_2 - c_{12}) \langle \delta_{\mathcal{C}^{a_1}, \mathcal{C}^{b_1}} \rangle_0$$

$$q_{a_2 b_2}^{[2]} = c_1 \langle \delta_{\mathcal{C}^{a_2}, \mathcal{C}^{b_2}} \rangle_1 + (1 - c_1 - c_2 - c_{12}) \langle \delta_{\mathcal{C}^{a_2}, \mathcal{C}^{b_2}} \rangle_0$$

$$q_{a_1 b_2}^{[12]} = (1 - c_1 - c_2 - c_{12}) \langle \delta_{\mathcal{C}^{a_1}, \mathcal{C}^{b_2}} \rangle_0,$$

where the indices a_1 and b_1 (resp., a_2 and b_2) belong to the first (resp., second) group of replicas (*i.e.*, $a_1, b_1 = 1, \ldots, n_1$ and $a_2, b_2 = 1 + n_1, \ldots, n_1 + n_2$), and the averages are performed over the single-site Hamiltonians $\mathcal{H}_1$, $\mathcal{H}_2$, and $\mathcal{H}_0$, corresponding to (minus) the arguments of the exponentials appearing in the expressions of Z_0, Z_1, and Z_2 in Eq. (A8). We now introduce a RS ansatz for the overlaps, $q_{ab}^{[1]} = q_1$, $q_{ab}^{[2]} = q_2$, $q_{ab}^{[12]} = q_{12}$, $\forall a, b$, which is justified for $T \geq T_K$, at least for small c_1, c_2 and c_{12}. Using once more the identity in Eq. (A3), $\delta_{\mathcal{C}^a, \mathcal{C}^b} = \sum_{\mathcal{C}}^\star \delta_{\mathcal{C}^a, \mathcal{C}} \delta_{\mathcal{C}^b, \mathcal{C}}$, the single-site

Hamiltonians can be rewritten as

$$-\mathcal{H}_1 = -\frac{n_2\beta^2 M}{4}(q_2 + c_2 + c_{12}) + \frac{\beta^2 M}{4}(q_2 + c_2 + c_{12}) \sum_{\mathcal{C}}^{\star} \left(\sum_{a}^{(2)} \delta_{\mathcal{C}^a,\mathcal{C}} \right)^2 ,$$

$$-\mathcal{H}_2 = -\frac{n_1\beta^2 M}{4}(q_1 + c_1 + c_{12}) + \frac{\beta^2 M}{4}(q_1 + c_1 + c_{12}) \sum_{\mathcal{C}}^{\star} \left(\sum_{a}^{(1)} \delta_{\mathcal{C}^a,\mathcal{C}} \right)^2 ,$$

$$-\mathcal{H}_0 = -\mathcal{H}_1 - \mathcal{H}_2 + \frac{\beta^2 M}{2}(q_{12} + c_{12}) \sum_{\mathcal{C}}^{\star} \sum_{a}^{(1)} \delta_{\mathcal{C}^a,\mathcal{C}} \sum_{b}^{(2)} \delta_{\mathcal{C}^b,\mathcal{C}} .$$

The replicas can again be decoupled via Hubbard-Stratonovich transformations, yielding (in the limit $n_1, n_2 \to 0$ and keeping only terms up to second order in the number of replicas)

$$
\ln Z_1 = -\frac{n_2\beta^2 M}{4}(q_2 + c_2 + c_{12}) + n_2 \overline{\left[\ln \left(\sum_{\mathcal{C}}^{\star} e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}} z_{\mathcal{C}}} \right) \right]}
$$

$$
+ \frac{n_2^2}{2} \left\{ \overline{\left[\ln \left(\sum_{\mathcal{C}}^{\star} e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}} z_{\mathcal{C}}} \right) \right]^2} - \left[\overline{\ln \left(\sum_{\mathcal{C}}^{\star} e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}} z_{\mathcal{C}}} \right)} \right]^2 \right\} ,
$$

$$
\ln Z_2 = -\frac{n_1\beta^2 M}{4}(q_1 + c_1 + c_{12}) + n_1 \overline{\left[\ln \left(\sum_{\mathcal{C}}^{\star} e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}} z_{\mathcal{C}}} \right) \right]}
$$

$$
+ \frac{n_1^2}{2} \left\{ \overline{\left[\ln \left(\sum_{\mathcal{C}}^{\star} e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}} z_{\mathcal{C}}} \right) \right]^2} - \left[\overline{\ln \left(\sum_{\mathcal{C}}^{\star} e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}} z_{\mathcal{C}}} \right)} \right]^2 \right\} ,
$$

(A9)

where, as before, the averages $\overline{[f(\vec{z}_{\mathcal{C}})]}$ are defined over the gaussian measure $\overline{[f(\vec{z}_{\mathcal{C}})]} \equiv \int \prod_{\mathcal{C}}^{\star} \left[\frac{dz_{\mathcal{C}}}{\sqrt{2\pi}} e^{-z_{\mathcal{C}}^2/2} \right] f(\vec{z}_{\mathcal{C}})$.

The computation of Z_0 is slightly more involved. After introducing $2(2^M - 1)$ δ-functions enforcing $x_{\mathcal{C}} = i\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}} \sum_a^{(1)} \delta_{\mathcal{C}^a,\mathcal{C}}$ and $y_{\mathcal{C}} = i\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}} \sum_a^{(2)} \delta_{\mathcal{C}^a \mathcal{C}}$ in the RS ansatz, Z_0 becomes

$$
Z_0 = e^{-\frac{\beta^2 M}{4}[n_2(q_2+c_2+c_{12})+n_1(q_1+c_1+c_{12})]}
$$

$$
\times \sum_{\{\mathcal{C}^a\}_{\star}} \prod_{\mathcal{C}}^{\star} \left\{ \int_{-\infty}^{\infty} dx_{\mathcal{C}}\, \delta \left(i\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}} \sum_a^{(1)} \delta_{\mathcal{C}^a,\mathcal{C}} - x_{\mathcal{C}} \right) \int_{-\infty}^{\infty} dy_{\mathcal{C}}\, \delta \left(i\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}} \sum_a^{(2)} \delta_{\mathcal{C}^a,\mathcal{C}} - y_{\mathcal{C}} \right) \right.
$$

$$
\left. \times \exp\left(-\frac{x_{\mathcal{C}}^2}{2} - \frac{y_{\mathcal{C}}^2}{2} - \frac{q_{12}+c_{12}}{\sqrt{(q_1+c_1+c_{12})(q_2+c_2+c_{12})}} x_{\mathcal{C}} y_{\mathcal{C}} \right) \right\} .
$$

Using the integral representation of the δ-function, $\delta(x - x_0) = \int_{-\infty}^{+\infty} d\hat{x}\, e^{-i\hat{x}(x-x_0)}$, integrating over $x_{\mathcal{C}}$ and $y_{\mathcal{C}}$, and performing the trace over configurations, one then easily finds

$$
Z_0 = e^{-\frac{\beta^2 M}{4}[n_2(q_2+c_2+c_{12})+n_1(q_1+c_1+c_{12})]} \int \prod_{\mathcal{C}}^{\star} \left[d\hat{x}_{\mathcal{C}}\, d\hat{y}_{\mathcal{C}} \frac{2\pi}{\sqrt{1-\gamma^2}} e^{-\frac{1}{1-\gamma^2}\left(\frac{\hat{x}_{\mathcal{C}}^2}{2} + \frac{\hat{y}_{\mathcal{C}}^2}{2} - \gamma \hat{x}_{\mathcal{C}} \hat{y}_{\mathcal{C}} \right)} \right]
$$

$$
\times \left(\sum_{\mathcal{C}}^{\star} e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}} \hat{x}_{\mathcal{C}}} \right)^{n_1} \left(\sum_{\mathcal{C}}^{\star} e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}} \hat{y}_{\mathcal{C}}} \right)^{n_2} ,
$$

where

$$
\gamma = \frac{q_{12}+c_{12}}{\sqrt{(q_1+c_1+c_{12})(q_2+c_2+c_{12})}} .
$$

Note that γ must be less then one (i.e., $q_{12} < q_1 + c_1$ and $q_{12} < q_2 + c_2$) for the Gaussian integrals to be well defined. We will find at the end of the computation that this is indeed the case. Expanding the logarithm of Z_0 in powers of

n_1 and n_2 and keeping only terms up to second order we obtain

$$\ln Z_0 = -\frac{\beta^2 M}{4}\left[n_2(q_2+c_2+c_{12})+n_1(q_1+c_1+c_{12})\right]+\ln(4\pi^2)$$
$$+n_1\overline{\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}}\hat{x}_{\mathcal{C}}}\right)\right]}+n_2\overline{\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}}\hat{y}_{\mathcal{C}}}\right)\right]}$$
$$+\frac{n_1^2}{2}\left\{\overline{\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}}\hat{x}_{\mathcal{C}}}\right)\right]^2}-\left[\overline{\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}}\hat{x}_{\mathcal{C}}}\right)\right]}\right]^2\right\}$$
$$+\frac{n_2^2}{2}\left\{\overline{\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}}\hat{y}_{\mathcal{C}}}\right)\right]^2}-\left[\overline{\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}}\hat{y}_{\mathcal{C}}}\right)\right]}\right]^2\right\} \tag{A10}$$
$$+n_1 n_2\left\{\overline{\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}}\hat{x}_{\mathcal{C}}}\right)\right]\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}}\hat{y}_{\mathcal{C}}}\right)\right]}^{(\star)}\right.$$
$$\left.-\overline{\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}}\hat{x}_{\mathcal{C}}}\right)\right]}\;\overline{\left[\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}}\hat{y}_{\mathcal{C}}}\right)\right]}\right\},$$

where the average $\overline{[g(\vec{\hat{x}}_{\mathcal{C}},\vec{\hat{y}}_{\mathcal{C}})]}^{(\star)}$ is defined over the gaussian measure:

$$\overline{[g(\vec{\hat{x}}_{\mathcal{C}},\vec{\hat{y}}_{\mathcal{C}})]}^{(\star)}\equiv\int\prod_{\mathcal{C}}^{\star}\left[d\hat{x}_{\mathcal{C}}\,d\hat{y}_{\mathcal{C}}\frac{1}{2\pi\sqrt{1-\gamma^2}}\,e^{-\frac{1}{1-\gamma^2}\left(\frac{\hat{x}_{\mathcal{C}}^2}{2}+\frac{\hat{y}_{\mathcal{C}}^2}{2}-\gamma\hat{x}_{\mathcal{C}}\hat{y}_{\mathcal{C}}\right)}\right]g(\vec{\hat{x}}_{\mathcal{C}},\vec{\hat{y}}_{\mathcal{C}}).$$

At this point we should find the saddle-point expressions of q_1, q_2 and q_{12} that extremize $A[q_{ab}^{[1]},q_{ab}^{[2]},q_{ab}^{[12]}]$, Eq. (A8), insert these expressions back into Eqs. (A8), (A9), and (A10), and finally determine $\mathcal{S}_{\text{rep}}(c_1,c_2,c_{12})$. It is however important to remember that in order to obtain the second cumulant of the effective Hamiltonian we do not need the whole expression of $\mathcal{S}_{\text{rep}}(c_1,c_2,c_{12})$, but only the terms of order $n_1 n_2$. It is then convenient to expand the saddle-point solutions of the overlaps in powers of n_1 and n_2. To compute the second cumulant we will only need to expand the overlap as in Eq. (17) of the main text. It is also convenient to define the following functions:

$$L_1(\vec{z}_{\mathcal{C}})\equiv\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}}z_{\mathcal{C}}}\right),\qquad L_2(\vec{z}_{\mathcal{C}})\equiv\ln\left(\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}}z_{\mathcal{C}}}\right),$$

$$K_1(\vec{z}_{\mathcal{C}})\equiv\frac{\sum_{\mathcal{C}}^{\star}z_{\mathcal{C}}\,e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}}z_{\mathcal{C}}}}{\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_1+c_1+c_{12})}{2}}z_{\mathcal{C}}}},\qquad K_2(\vec{z}_{\mathcal{C}})\equiv\frac{\sum_{\mathcal{C}}^{\star}z_{\mathcal{C}}\,e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}}z_{\mathcal{C}}}}{\sum_{\mathcal{C}}^{\star}e^{\sqrt{\frac{\beta^2 M(q_2+c_2+c_{12})}{2}}z_{\mathcal{C}}}},$$

such that $\quad\dfrac{dL_{1,2}(\vec{z}_{\mathcal{C}})}{dq_{1,2}}=\dfrac{1}{2}\sqrt{\dfrac{\beta^2 M}{2(q_{1,2}+c_{1,2}+c_{12})}}K_{1,2}(\vec{z}_{\mathcal{C}}).$

In terms of these functions, we get

$$A[q_1,q_2,q_{12}]\approx\frac{\beta^2 M}{4}\left[-n_1\frac{q_1^2}{2}-n_2\frac{q_2^2}{2}+n_1 n_2 q_{12}^2+n_1(1-c_1-c_{12})(q_1+c_1+c_{12})+n_2(1-c_2-c_{12})(q_2+c_2+c_{12})\right]$$
$$-n_1(1-c_1-c_{12})\overline{L_1(\vec{z}_{\mathcal{C}})}-n_2(1-c_2-c_{12})\overline{L_2(\vec{z}_{\mathcal{C}})}$$
$$-n_1 n_2(1-c_1-c_2-c_{12})\left(\overline{L_1(\vec{x}_{\mathcal{C}})L_2(\vec{y}_{\mathcal{C}})}^{(\star)}-\overline{L_1(\vec{z}_{\mathcal{C}})}\;\overline{L_2(\vec{z}_{\mathcal{C}})}\right)$$

$$\tag{A11}$$

The extremization of $A[q_1, q_2, q_{12}]$ with respect to q_1 gives

$$
q_1 = (1 - c_1 - c_{12}) \left[1 - \sqrt{\frac{2}{\beta^2 M (q_1 + c_1 + c_{12})}} \overline{K_1(\vec{z}_{\mathcal{C}})} \right]
$$

$$
- n_2 (1 - c_1 - c_2 - c_{12}) \left\{ \frac{2\gamma^2}{\beta^2 M (q_1 + c_1 + c_{12})(1 - \gamma^2)^2} \overline{\left[\sum_{\mathcal{C}}^{\star} \left(x_{\mathcal{C}}^2 + y_{\mathcal{C}}^2 - \frac{1 + \gamma^2}{\gamma} x_{\mathcal{C}} y_{\mathcal{C}} \right) - N(1 - \gamma^2) \right] L_1(\vec{x}_{\mathcal{C}}) L_2(\vec{y}_{\mathcal{C}})}^{(\star)} \right.
$$

$$
\left. + \sqrt{\frac{2}{\beta^2 M (q_1 + c_1 + c_{12})}} \left(\overline{K_1(\vec{x}_{\mathcal{C}}) L_2(\vec{y}_{\mathcal{C}})}^{(\star)} - \overline{K_1(\vec{z}_{\mathcal{C}})}\, \overline{L_2(\vec{z}_{\mathcal{C}})} \right) \right\} .
$$

(A12)

We obtain the same equation for q_2 by changing all indices $1 \leftrightarrow 2$. Finally, the saddle-point equation for q_{12} reads

$$
q_{12} = (1 - c_1 - c_2 - c_{12}) \frac{2\gamma^2}{\beta^2 M (q_{12} + c_{12})(1 - \gamma^2)^2} \overline{\left[N(1 - \gamma^2) - \sum_{\mathcal{C}}^{\star} \left(x_{\mathcal{C}}^2 + y_{\mathcal{C}}^2 - \frac{1 + \gamma^2}{\gamma} x_{\mathcal{C}} y_{\mathcal{C}} \right) \right] L_1(\vec{x}_{\mathcal{C}}) L_2(\vec{y}_{\mathcal{C}})}^{(\star)} . \quad (A13)
$$

Inserting the expansion (17) into Eqs. (A12) and (A13) allows us to obtain $q_1^{[0,0]}$, $q_1^{[0,1]}$, $q_2^{[0,0]}$, $q_2^{[1,0]}$, and $q_{12}^{[0,0]}$ which, once inserted into Eq. (A11), finally yield the second cumulant. As for the computation of the first cumulant, we show explicitly how this can be done for $M \gg 1$ and we keep only terms to second order in c_1, c_2, and c_{12}.

After expanding the functions $K_{1,2}(\vec{z}_{\mathcal{C}})$ in powers of $\epsilon_{1,2} = \sqrt{\frac{M \beta^2 (q_{1,2} + c_{1,2} + c_{12})}{2}}$ up to the sixth order, one obtains

$$
q_{1,2}^{[0,0]} \approx \frac{1}{2^M} + \frac{M\beta^2 - 2}{2^{M+1}} (c_{1,2} + c_{12}) + \frac{M\beta^2 (M\beta^2 - 4)}{2^{M+3}} (c_{1,2} + c_{12})^2 + \dots ,
$$

which, of course, coincides with the first two terms of Eq. (A7) with $c \to c_{1,2} + c_{12}$. Inserting these solutions into Eq. (A13) leads to the expression of the saddle-point value of q_{12} in powers of the concentrations (as above, we only consider the leading terms in $1/2^M$):

$$
q_{12}^{[0,0]} \approx \frac{1}{2^M} - \frac{1}{2^M}(c_1 + c_2) + \frac{M\beta^2 - 2}{2^{M+1}} c_{12} + \frac{M\beta^2}{2^{2M+1}}(c_1 + c_2)^2 + \frac{M\beta^2(M\beta^2 - 4)}{2^{M+3}} c_{12}^2 - \frac{M\beta^2}{2^{M+1}}(c_1 + c_2)c_{12} .
$$

From the above results we self-consistently find that $\gamma \geq 0$. The expressions of $q_{1,2}^{[0,0]}$ and $q_{12}^{[0,0]}$, when inserted into Eq. (A12), yield the corrections of order n_2 (resp. n_1) to the saddle-point value of q_1 (resp. q_2), which actually turns out to be very small for large M. Up to the leading terms in 2^M, we find

$$
q_1^{[0,1]} \approx \frac{M^2\beta^4}{2^{4M+2}} - \frac{M^2\beta^4}{2^{4M+2}}(c_1 + c_2) + \frac{M^2\beta^4}{2^{3M+1}} c_{12} - \frac{M^2\beta^4}{2^{3M+1}}(c_1 + c_2)c_{12} + \frac{M^2\beta^4}{2^{2M+2}} c_{12}^2 + \frac{M^2\beta^4(M\beta^2 - 1)}{2^{5M+3}}(c_1 + c_2)c_1 .
$$

An analogous expression for $q_2^{[1,0]}$ is obtained by changing $1 \leftrightarrow 2$. Finally, collecting all these results together into Eqs. (A8) and using Eq. (4) allows us to obtain the expression of the second cumulant of the effective Hamiltonian (up to the second order in the concentrations c_1, c_2, c_{12}):

$$
\frac{\mathcal{S}_2[c_1, c_2, c_{12}]}{N} = -\frac{1}{N} \lim_{n_1, n_2 \to 0} \frac{\mathcal{S}_{\rm rep}[c_1, c_2, c_{12}]}{n_1 n_2} = \frac{\beta^2 M}{4} c_{12}^2 - \lim_{n_1, n_2 \to 0} \frac{A[q_1, q_2, q_{12}]}{n_1 n_2}
$$

$$
\approx \frac{M\beta^2}{2^{2M+2}} - \frac{M\beta^2}{2^{2M+1}}(c_1 + c_2) + \frac{M\beta^2}{2^M + 1} c_{12} + \frac{M\beta^2}{2^{2M+2}}(c_1 + c_2)^2
$$

$$
+ \frac{M\beta^2}{4}\left(1 + \frac{M\beta^2 - 4}{2^{M+1}}\right) c_{12}^2 + \frac{M\beta^2}{2^{2M+1}} c_1 c_2 - \frac{M\beta^2}{2^{M+1}}(c_1 + c_2)c_{12}
$$

$$
\approx \frac{M\beta^2}{4}\left(1 + \frac{M\beta^2 - 4}{2^{M+1}}\right) c_{12}^2 + \frac{M\beta^2}{2^{M+1}}(1 - c_1 - c_2)c_{12} ,
$$

where in the last line we have only kept terms up to $O(1/2^M)$. By re-expressing the concentrations via Eq. (18), we finally obtain the second cumulant of the effective action given in Eq. (19) the main text.

Appendix B: Exact solution of the fully connected random-field + random-bond effective Ising model

The fully connected random-field + random-bond Ising model defined by the Hamiltonian in Eq. (2) can be solved exactly. We drop the spin-independent random term and we first simplify the problem by neglecting the cross-correlations between the random fields and the random bonds (see below and the companion paper[1] for a test of this approximation) as well as the off-diagonal part of the random field correlation. As a result, on has to consider Gaussian distributed random variables with

$$
\begin{aligned}
\overline{\delta h_i \delta h_j}^{(0)} &= \Delta_h \delta_{ij}\,, \\
\overline{\delta J_{2,ij} \delta J_{2,kl}}^{(0)} &= \Delta_J (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\,.
\end{aligned}
\tag{B1}
$$

By using the replica trick and performing Hubbard-Stratonovich transformations, the partition function of the model can be written as

$$
\overline{Z^n} = e^{\frac{nN}{2}\left(\frac{\Delta_J}{2}+\Delta_h\right)} \left(\frac{N\Delta_J}{2\pi}\right)^{n(n-1)/2} \int_{-\infty}^{+\infty} \prod_a \mathrm{d}m_a \int_{-i\infty}^{+i\infty} \prod_a \mathrm{d}\mu_a \int_{-\infty}^{+\infty} \prod_{a\neq b} \mathrm{d}q_{ab}\, e^{NF[\{m_a,\mu_a,q_{ab}\}]}\,,
$$

with

$$
F[\{m_a,\mu_a,q_{ab}\}] = \sum_a \left[(H-\mu)m_a + \frac{J_2}{2}m_a^2 + \frac{J_3}{3!}m_a^3 + \frac{J_4}{4!}m_a^4\right] - \frac{\Delta_J}{4}\sum_{a\neq b} q_{ab}^2 + \ln Z_1\,.
$$

The partition function of the single-site problem reads

$$
Z_1 = \sum_{\{\sigma_a\}} \exp\left[\sum_a \mu_a \sigma_a + \frac{1}{2}\sum_{a\neq b}(\Delta_J q_{ab} + \Delta_h)\sigma_a \sigma_b\right]\,.
$$

The saddle-point equations then provide

$$
\begin{aligned}
m_a &= \langle \sigma_a \rangle \\
q_{ab} &= \langle \sigma_a \sigma_b \rangle\,, \\
\mu_a &= H + J_2 m_a + \frac{J_3}{2}m_a^2 + \frac{J_4}{6}m_a^3\,,
\end{aligned}
$$

where the averages are computed by using the single-site Hamiltonian. Taking the replica-symmetric (RS) ansatz, $m_a = m$, $q_{ab} = q$, $\mu_a = \mu$, we find

$$
Z_1 = e^{-\frac{n}{2}(\Delta_J q + \Delta_h)} \int \mathcal{D}z \left[2\cosh\left(z\sqrt{\Delta_J q + \Delta_h} + \mu\right)\right]^n\,,
$$

where $\mathcal{D}z = e^{-z^2/2}\mathrm{d}z/\sqrt{2\pi}$. To the leading order in N in the $n \to 0$ limit, the free energy per spin then reads

$$
f(m,q,\mu) = (H-\mu)m + \frac{J_2}{2}m^2 + \frac{J_3}{3!}m^3 + \frac{J_4}{4!}m^4 + \frac{\Delta_J}{4}q^2 - \frac{\Delta_J}{2}q + \int \mathcal{D}z \ln 2\cosh\left(z\sqrt{\Delta_J q + \Delta_h} + \mu\right)\,.
$$

After taking the derivatives with respect to m, μ, and q, we obtain the following self-consistent equations:

$$
\begin{aligned}
m &= \int \mathcal{D}z \tanh\left(J_2 m + \frac{J_3}{2}m^2 + \frac{J_4}{6}m^3 + H + z\sqrt{\Delta_J q + \Delta_h}\right)\,, \\
q &= 1 - \frac{1}{\sqrt{\Delta_J q + \Delta_h}} \int \mathcal{D}z\, z \tanh\left(J_2 m + \frac{J_3}{2}m^2 + \frac{J_4}{6}m^3 + H + z\sqrt{\Delta_J q + \Delta_h}\right)\,,
\end{aligned}
\tag{B2}
$$

which can be easily solved numerically.

The result found by using the approximate effective theory [*i.e.*, solving Eqs. (B2) by using the effective coupling constants and variances of the random terms given in Eqs. (21) and (22)] is plotted in Fig. 4 for $M = 3$ (red curve, squares). It shows a good quantitative agreement with the exact solution.

Finally, one could wonder whether the higher-order correlations of the distributions of the random bonds and random fields play an important role. In order to check this, we have repeated the calculation, taking now into

account more terms of the disorder distributions, namely the correlation between random fields on different sites and the correlation between random fields and random bonds [see Eq. (22)]:

$$\overline{\delta h_i \delta J_{2,kj}} = \kappa \frac{\delta_{ij} + \delta_{ik}}{\sqrt{N}} \ .$$

In this case the saddle-point equations read

$$
\begin{aligned}
m &= \int \mathcal{D}z \tanh\left[(J_2 + 2\kappa)m + \frac{J_3}{2}m^2 + \frac{J_4}{6}m^3 + H - \kappa q + z\sqrt{\Delta_J q + \Delta_h + 2\kappa m}\right] , \\
q &= 1 - \frac{1}{\sqrt{\Delta_J q + \Delta_h + 2\kappa m}} \int \mathcal{D}z\, z \tanh\left[(J_2 + 2\kappa)m + \frac{J_3}{2}m^2 + \frac{J_4}{6}m^3 + H - \kappa q + z\sqrt{\Delta_J q + \Delta_h + 2\kappa m}\right] .
\end{aligned}
\tag{B3}
$$

We have solved these equations numerically for $\kappa = M\beta^2(1 + 2^{2-M})/32$ and found no significant difference with respect to the case in which these higher-order correlations of the disorder distribution are neglected. (We will come back to this point in the companion paper[1], see, *e.g.*, Fig. 5.)

---

* Electronic address: giulio.biroli@cea.fr
† Electronic address: chiara.cammarota@kcl.ac.uk
‡ Electronic address: tarjus@lptmc.jussieu.fr
§ Electronic address: tarzia@lptmc.jussieu.fr
[1] G. Biroli, C. Cammarota, G. Tarjus, and M. Tarzia, following paper.
[2] R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).
[3] J. M. Kosterlitz, J. Phys. C: Solid State Phys., **7** 1046 (1974).
[4] *Structural Glasses and Supercooled Liquids*, Eds: P.G. Wolynes and V. Lubchenko, Wiley (2012).
[5] L. Berthier and G. Biroli, Rev. Mod. Phys. **83**, 587 (2011).
[6] A. Cavagna, Phys. Rep. **476**, 51 (2009).
[7] W. Götze, *Liquids, Freezing and the Glass Transition*, edited by J. P. Hansen, D. Levesque, J. Zinn-Justin, Les Houches. Session LI, 1989 (North-Holland, Amsterdam), 1991;
[8] P. G. Debenedetti, *Metastable Liquids* (Princeton University Press, Princeton), 1996. J. Dyre, Rev. Mod. Phys. **78**, 953 (2006).
[9] G. Tarjus, in *Dynamical Heterogeneities in Glasses, Colloids, and Granular Media*, Eds: L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, and W. van Saarloos (Oxford University Press, New York, 2011).
[10] D. Chandler and J. P. Garrahan, Annu. Rev. Phys. Chem. **61**, 191 (2010).
[11] G. Tarjus, S. A. Kivelson, Z. Nussinov, and P. Viot, J. Phys. Condens. Matter **17**, R1143 (2005).
[12] T. R. Kirkpatrick, D. Thirumalai, and P. G. Wolynes, Phys. rev. A **40**, 1045 (1989).
[13] V. Lubchenko and P. G. Wolynes, Annu. Rev. Phys. Chem. **58**, 235 (2007).
[14] J. Kurchan, G. Parisi, and F. Zamponi, J. Stat. Mech. P10012 (2012); J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, J. Phys. Chem. B **117**, 12979 (2013); P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, J. Stat. Mech. P10009 (2014); Annu. Rev. Condens. Matter Phys. **8**, 265 (2017).
[15] J. Kurchan, T. Maimbourg, and F. Zamponi, Phys. Rev. Lett. **116**, 015902 (2016).
[16] S. Franz and G. Parisi, J. Phys. (Paris) I **5**, 1401 (1995).
[17] S. Franz, M. Cardenas, and G. Parisi, J. Phys. A **31**, L163 (1998); M. Cardenas, S. Franz, and G. Parisi, J. Chem. Phys. **110**, 1726 (1999).
[18] S. Franz and G. Parisi, Phys. Rev. Lett. **79**, 2486 (1997); S. Franz and G. Parisi, Physica (Amstardam) **261A**, 317 (1998).
[19] G. Biroli and J. P. Bouchaud, in *Structural Glasses and Supercooled Liquids*, Eds: P.G. Wolynes and V. Lubchenko, (Wiley, Singapore, 2012), Chap. 2.
[20] C. Cammarota, G. Biroli, M. Tarzia, G. Tarjus, Phys. Rev. B **87**, 064202 (2013).
[21] C. Rulquin, P. Urbani, G. Biroli, G. Tarjus, and M. Tarzia, J. Stat. Mech. 023209 (2016).
[22] M. Dzero, J. Schmalian, and P. G. Wolynes, Phys. Rev. B **72**, 100201 (2005); M. Dzero, J. Schmalian, and P. G. Wolynes, Phys. Rev. B **80**, 024204 (2009).
[23] S. Franz, J. Stat. Mech. P04001 (2005); S. Franz, Europhys. Lett. **73**, 492 (2006).
[24] C. Cammarota, G. Biroli, M. Tarzia, and G. Tarjus, Phys. Rev. Lett. **106** 115705 (2011).
[25] M. Castellana, A. Decelle, S. Franz, M. Mézard, and G. Parisi, Phys. Rev. Lett. **104**, 127206 (2010).
[26] M. C. Angelini and G. Biroli, Proc. Natl. Acad. Sci., 201613126 (2017).
[27] M. A. Moore and J. Yeo, Phys. Rev. Lett. **96**, 095701 (2006); J. Yeo, M. A. Moore, Phys. Rev. B **85**, 100405(R) (2012).
[28] For a review, see T. Nattermann, *Spin glasses and random fields* (World scientific, Singapore, 1998).
[29] S. Franz, G. Parisi, F. Ricci-Tersenghi, T. Rizzo, Eur. Phys. J. E **34**, 102 (2011).
[30] S. Franz, G. Parisi, JSTAT, P11012 (2013).

[31] S. Franz, G. Parisi, and F. Ricci-Tersenghi, J. Stat. Mech. L02001 (2013).

[32] G. Biroli, C. Cammarota, G. Tarjus, M. Tarzia, Phys. Rev. Lett. **112**, 175701 (2014).

[33] B. Derrida, Phys. Rev. B **24**, 2613 (1981).

[34] S. Franz, G. Parisi, F. Ricci-Tersenghi, Journal of Physics A: Math. Gen. **41**, 324011 (2008).

[35] J. D. Stevenson, A. M. Walczak, R. W. Hall, and P. G. Wolynes, J. Chem. Phys. **129**, 194505 (2008).

[36] S. K. Nandi, G. Biroli, and G. Tarjus, Phys. Rev. Lett. **116**, 145701 (2016).

[37] M. Ozawa, W. Kob, A. Ikeda, and K. Miyazaki, PNAS (2015): 201500730.

[38] W. Kob and L. Berthier, Phys. Rev. Lett. **110**, 245702 (2013).

[39] L. Berthier, Phys. Rev. E **88**, 022313 (2013).

[40] L. Berthier and R. L. Jack, Phys. Rev. Lett. **114**, 205701 (2015).

[41] G. Parisi and B. Seoane, Phys. Rev. E **89**, 022309 (2014).

[42] L. Berthier, P. Charbonneau, D. Coslovich, A. Ninarello, M. Ozawa, and S. Yaida, Proc. Natl. Acad. Sci U.S.A. **114**, 11356 (2017).

[43] C. Cammarota, A. Cavagna, G. Gradenigo, T.S. Grigera and P. Verrocchio J. Chem. Phys. **131**, 194901 (2009).

[44] G. Biroli and C. Cammarota, Phys. Rev. X **7**, 011011 (2017).

[45] J. P. Garrahan, J. Phys.: Condens. Matter **14**, 1571 (2002).

[46] F. Ritort and P. Sollich, Adv. Phys. **52**, 219 (2003).

[47] J. P. Garrahan, Phys. Rev. E **89**, 030301 (2014).

[48] R. L. Jack and J. P. Garrahan, Phys. Rev. Lett. **116**, 055702 (2016).

[49] R. M. Turner, R. L. Jack and J. P. Garrahan, Phys. Rev. E **92**, 022115 (2015).

[50] G. Biroli, C. Rulquin, G. Tarjus, and M. Tarzia, SciPost Phys. **1**, 007 (2016).

[51] C. Cammarota, G. Biroli, J. Chem. Phys. **138**, 12A547 (2013); G. Biroli and C. Cammarota, Proc. Natl. Acad. Sci. U.S.A. **109**, 8850 (2012).

[52] F. Ricci-Tersenghi and G. Semerjian, J. Stat. Mech. P09001 (2009).

[53] C. Cammarota, Europhys. Lett. **101**, 56001 (2013).

[54] L. Berthier and W. Kob, Phys. Rev. E **85**, 011102 (2012).

[55] C. Cammarota and G. Biroli, Europhys. Lett. **98**, 16011 (2012).

[56] V. Krakoviack, Phys. Rev. Lett. **94**, 065703 (2005).

[57] G. Szamel and E. Flenner, Europhys. Lett. **101** 66005 (2013).

[58] S. K. Nandi, G. Biroli, J.-P. Bouchaud, K. Miyazaki, and D. R. Reichman, Phys. Rev. Lett. **113** 245701 (2014).

[59] G. H. Fredrickson and H. C. Andersen, Phys. Rev. Lett. **53**, 1244 (1984).

[60] S. Franz and M. Sellitto, J. Stat. Mech. P02025 (2013).

[61] H. Ikeda, K. Miyazaki, and G. Biroli, EPL **116**, 56004 (2016).

[62] T. Rizzo, Physical Review B, **94** 014202 (2016).

[63] J.-P. Bouchaud and G. Biroli, J. Chem. Phys. **121**, 7347 (2004); G. Biroli, J.-P. Bouchaud, A. Cavagna, T. S. Grigera, and P. Verrocchio, Nat. Phys. **4**, 771 (2008).

[64] G. Tarjus and M. Tissier, Phys. Rev. Lett. **93**, 267008 (2004); G. Tarjus and M. Tissier, Phys. Rev. B **78**, 024203 (2008).

[65] M. Tissier and G. Tarjus, Phys. Rev. B **85**, 104202 (2012); Phys. Rev. B **85**, 104203 (2012).

[66] Actually, for values of the overlap close to the barrier one, a quenched computation is needed, but even quantitatively, on the basis of previous results, one does not expect significant changes.

[67] M. Mézard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1988).

[68] S. Franz and A. Montanari, J. Phys. A **40**, F251 (2007).

[69] Note that in the magnetic representation in terms of (hard) Ising variables, the role of the random energy $\mathcal{S}_0$ is to ensure that the variance $\mathcal{S}_2$ of the effective Hamiltonian is equal to zero in the zero-overlap (liquid) state.

[70] A. A. Middleton and D. S. Fisher, Phys. Rev. B **65**, 134411 (2002); N. G. Fytas and V. Martín-Mayor, Phys. Rev. Lett. **110**, 227201 (2013); Phys. Rev. Lett. **111**, 019903 (2013).