Derivation of the spin-glass order parameter from stochastic thermodynamics

A. Crisanti,[1] M. Picco,[2] and F. Ritort[3,4]

[1]*Dipartimento di Fisica, Università di Roma "La Sapienza", and Istituto dei Sistemi Complessi ISC-CNR, P.le Aldo Moro 2, 00185 Roma, Italy*
[2]*Sorbonne Université, CNRS, Université Pierre et Marie Curie, LPTHE, F-75005 Paris, France*
[3]*Small Biosystems Lab, Condensed Matter Physics Department, Universitat de Barcelona, C/ Martí i Franquès, E-08028, Barcelona, Spain*
[4]*CIBER-BBN Center for Bioengineering, Biomaterials and Nanomedicine, Instituto de Salud Carlos III, Madrid, Spain*

A fluctuation relation is derived to extract the order parameter function $q(x)$ in weakly ergodic systems. The relation is based on measuring and classifying entropy production fluctuations according to the value of the overlap q between configurations. For a fixed value of q, entropy production fluctuations are Gaussian distributed allowing us to derive the quasi-FDT so characteristic of aging systems. The theory is validated by extracting the $q(x)$ in various types of glassy models. It might be generally applicable to other nonequilibrium systems and experimental small systems.

I. INTRODUCTION

Dynamically frustrated systems (DFSs) include a wide category of soft matter and solid-state systems (e.g., polymers, gels, superconductors, magnetic alloys, biomolecules, etc.) exhibiting strong nonequilibrium effects characterized by slow relaxational kinetics, aging, and memory effects [1–3]. Such a rich phenomenology has been predicted by weak ergodicity-breaking models, yet resists interpretation in terms of equilibrium-based statistical thermodynamics approaches. Experimentally, a spatial correlation length mildly growing in time spanning a few intermolecular distances has been identified in DFS suggesting that some kind of ordering takes place [4]. The order parameter in DFS, originally introduced in the context of spin glasses as a time-dependent correlation function [5], still eludes direct experimental measurement. In contrast to homogeneous systems where one or a few single order parameters quantify the degree of macroscopic order, in DFS a whole function $q(x)$ is required. Roughly speaking, $q(x)$ quantifies the degree of correlation q between different states that share a common fraction x of thermalized degrees of freedom in a corrugated free-energy landscape with many states separated by high free-energy barriers [6]. The mathematical description of such a landscape is often known as the replica-symmetry-breaking solution of the mean-field spin glass [7]. A key aspect of $q(x)$ is the possibility to define it from statics and dynamics reflecting the fact that some kind of thermalization, characteristic of equilibrium systems, still operates under weakly ergodic conditions [8]. In fact, DFS partially equilibrate among the many states comprised by the $q(x)$; however, they do so only for finite observation times because they intermittently change state during the relaxation process. For systems with a nontrivial $q(x)$, its inverse function $x(q)$ has been interpreted in terms of frozen, i.e., not relaxed, degrees of freedom [9,10] with $P(q) = dx(q)/dq$ positively defined. In spin-glass theory $P(q)$ is defined as the probability of states having overlap q, dx being the probability measure for the states. To extract the equilibrium $x(q)$ requires monitoring the overlap between the state of the system at time t and an initially equilibrated

state and calculate the overlap distribution $P(q)$ (Fig. 1). However, in DFS such measurement is hard because one cannot equilibrate nor efficiently sample the different states.

A different route to extract $q(x)$ is to use a modified version of the fluctuation-dissipation theorem (referred to as quasi-FDT [11–14]) to measure time-dependent correlations and responses and extracting the so-called fluctuation dissipation ratio (FDR) $\tilde{x}(q)$ [Eq. (6) below]. The dynamical FDR $\tilde{x}(q)$ has been shown to be equal to the $x(q)$ from statics [15], however its experimental measurement remains difficult [16]. Here we present a novel approach to measure $q(x)$ in weakly ergodic systems based on noise measurements. A first step in this direction was presented a few years ago [17] by deriving a fluctuation relation (FR) for the aging state. Here we make a step further and introduce a FR for entropy production fluctuations at fixed overlap q from which we derive the quasi-FDT. A novel mathematical relation for the order parameter $x(q)$ from entropy production fluctuations is also obtained [cf. Eqs. (2) and (8) below].

II. FLUCTUATION RELATION AND THE GAUSSIAN UNIVERSALITY HYPOTHESIS

The key idea of the approach is to measure the entropy produced by an external perturbation of strength h that is applied to the system at a given time t_w during the relaxational process (e.g., following a T quench). The field h is conjugated to the macroscopic observable M that is experimentally accessible by following its time evolution, $M(t)$, up to a maximal time. The experiment is then repeated many times and the probability distribution $P_{t_w,q}(\Delta S) = \langle\delta(\Delta S - \Delta S_{t_w,t})\rangle_{q=C(t,t_w)}$ with $\Delta S_{t_w,t} = h\Delta M/T$ and $\Delta M = M(t) - M(t_w)$ is measured. The average $\langle\ldots\rangle_{q=C(t,t_w)}$ is constrained to those paths with $C(t,t_w) = q$ for $t > t_w$, an average that we will refer to as q statistics (or simply q stat). In Ref. [17] we proved that the unconstrained distribution $P_{t_w}(\Delta S)$, defined as the probability of observing the value $\Delta S_{t_w,t}$ after t_w (i.e., without classifying states) satisfies a FR with a large deviation parameter x_{t_w}. In the q-stat scheme we can introduce the large deviation function

(a) **(b)**

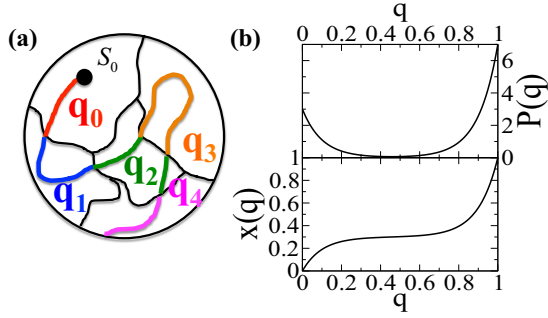

FIG. 1. The order parameter $q(x)$. (a) Illustration of a path in phase space. The time evolution of correlations $q_i = C(S_0, S_i)$ between different states S_i and state S_0 at t_w is collected to build a histogram $P(q) = \langle \delta(q - q_i) \rangle$ where $\langle \ldots \rangle$ denotes the average over paths [panel (b), top]. $x(q)$ is related to $P(q)$, $P(q) = dx(q)/dq$ [panel (b), bottom].

$x_{t_w}(q)$ through the following FR:

$$x_{t_w}(q) = \frac{k_B}{\Delta S} \ln \left(\frac{P_{t_w, q}(\Delta S)}{P_{t_w, q}(-\Delta S)} \right), \quad (1)$$

with k_B being the Boltzmann constant. In a weak ergodicity-breaking scenario, q is the appropriate parameter to partition phase space into partially equilibrated regions. We hypothesize that $P_{t_w, q}(\Delta S)$ is Gaussian for a system with a sufficiently large number of degrees of freedom. We will call this the Gaussian universality hypothesis (GUH) for $P_{t_w, q}(\Delta S)$. Its simplicity resonates with its importance: GUH provides an operational definition of partially equilibrated states in nonequilibrium weakly ergodic systems. Its range of validity might go beyond DFS, from noninteracting systems to nonequilibrium steady states. Note that the validity of the GUH is restricted to q-stat classification and should not hold for other types of classification among paths, such as fixing the time t after t_w at which $\Delta S_{t_w, t}$ is measured (referred to as t stat) or in the procedure of Ref. [17] (where no classification is made). In fact, in these latter cases the intermittent dynamics mixes states with different values of q and $x(q)$ and partial equilibration does not hold anymore. In what follows we show that such q-stat classification of partially equilibrated states and the GUH imply the validity of the quasi-FDT in DFS [Eq. (6)].

Equation (1) has a simple solution for a Gaussian $P_{t_w, q}(\Delta S)$ of mean $\langle \Delta S \rangle$ and variance $\sigma_{\Delta S}^2$,

$$x_{t_w}(q) = \frac{2k_B \langle \Delta S \rangle}{\sigma_{\Delta S}^2}. \quad (2)$$

The advantage of Eq. (2) with respect to Eq. (1) is clear. To measure $x_{t_w}(q)$ one does not need to sample the leftmost tails of $P_{t_w, q}(\Delta S)$ (i.e., the $\Delta S < 0$ events) but just the mean and variance of the distribution; a more amenable task in experiments and simulations.

From Eq. (2) we now derive the quasi-FDT. In the linear-response regime we have,

$$\langle \Delta S \rangle = \frac{h \langle \Delta M \rangle}{T} = \frac{h}{T} \left[h \int_{t_w}^{t} ds\, G(t, s) + O(h^3) \right], \quad (3)$$

where we have used the response function definition, $G(t, s) = \frac{\delta \langle M(t) \rangle}{\delta h(s)}$. Moreover,

$$\sigma_{\Delta S}^2 = \left(\frac{h}{T} \right)^2 [\tilde{C}(t, t) + \tilde{C}(t_w, t_w) - 2\tilde{C}(t, t_w)]$$

$$= \left(\frac{h}{T} \right)^2 [\tilde{C}_0(t, t) + \tilde{C}_0(t_w, t_w) - 2\tilde{C}_0(t, t_w)] + O(h^2), \quad (4)$$

with $\tilde{C}(t, s) = \langle M(t) M(s) \rangle$ being the two-times correlation function of the macroscopic observable M in the presence of a field h (the subscript 0 denotes the same correlation at $h = 0$). Substituting Eqs. (3) and (4) into Eq. (2) we obtain,

$$x_{t_w}(q) = \frac{2k_B T \int_{t_w}^{t} ds\, G(t, s)}{\tilde{C}_0(t, t) + \tilde{C}_0(t_w, t_w) - 2\tilde{C}_0(t, t_w)} + O(h^2). \quad (5)$$

Let us now consider the quasi-FDT given by [13,14]

$$k_B T G(t, s) = \theta(t - s)\tilde{x}[\tilde{C}_0(t, s)]\partial_s \tilde{C}_0(t, s), \quad (6)$$

with $\tilde{x}[\tilde{C}_0]$ the fluctuation-dissipation ratio (FDR) which depends on times t, s only through the value of $\tilde{C}_0(t, s)$. Next we derive a relation between $\tilde{x}(\tilde{C}_0)$ (6) and $x_{t_w}(q)$ (1). Inserting Eq. (6) into Eq. (5) we get to first order in h,

$$x_{t_w}(q) = \frac{2 \int_{\tilde{C}_0(t, t_w)}^{\tilde{C}_0(t, t)} d\tilde{q}\, \tilde{x}(\tilde{q})}{\tilde{C}_0(t, t) + \tilde{C}_0(t_w, t_w) - 2\tilde{C}_0(t, t_w)} + O(h^2), \quad (7)$$

with $\tilde{x}(\tilde{q} = \tilde{C}_0)$ the FDR in Eq. (6). If we now assume $\tilde{C}_0(t_w, t_w) = \tilde{C}_0(t, t) = N$ (e.g., for Ising-type systems with N the total number of spins) and define $C(t, s) = \tilde{C}(t, s)/\tilde{C}(s, s) = \tilde{C}(t, s)/N$, then we have $q = C_0(t, t_w)$ and

$$x_{t_w}(q) = \frac{1}{1 - q} \int_{q}^{1} dq'\, \tilde{x}(q') + O(h^2). \quad (8)$$

By taking the derivative of Eq. (8) with respect to q simple algebra leads to

$$\tilde{x}(q) = x_{t_w}(q) - (1 - q)\frac{dx_{t_w}(q)}{dq} + O(h^2), \quad (9)$$

directly relating $\tilde{x}(q)$ and $x_{t_w}(q)$. Note that Eq. (8) can be rewritten as $\int_{q}^{1} dq'\, \tilde{x}(q') = x_{t_w}(q)(1 - q) + O(h^2)$ meaning that, if we neglect $O(h^2)$ corrections, $\tilde{x}(q)$ equals the local slope of the curve $x_{t_w}(q)(1 - q)$ versus $1 - q$.

Equation (2) provides a way to directly extract the static quantities $x(q)$ and $q(x)$ solely from noise $\Delta M = M(t) - M(t_w)$ measurements. The easiest procedure is to collect statistics of the stochastic variable $\Delta S = h\Delta M/T$ for a fixed value of q, apply Eq. (2) and extract the local slope of the curve $x_{t_w}(q)(1 - q)$ versus $1 - q$. Note that this is the equivalent of the quasi-FDT representation (5) where one plots the susceptibility versus the correlation [13,14]. Alternatively, one can calculate the derivative of the average entropy production with respect to its variance,

$$x(q) = 2k_B \frac{\partial \langle \Delta S \rangle}{\partial \sigma_{\Delta S}^2} = \frac{2k_B T}{h} \frac{\partial \langle \Delta M \rangle}{\partial \sigma_{\Delta M}^2}. \quad (10)$$
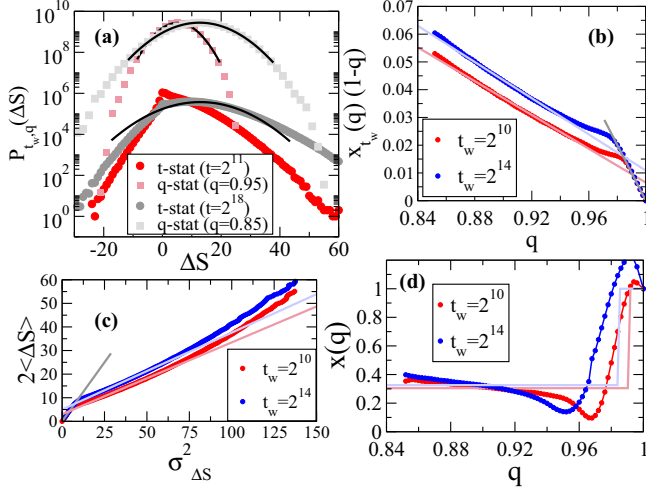
FIG. 2. Numerical test for ROM and fit to the 1RSB model. (a) q-stat and t-stat distributions for $t_w = 2^{10}$ and values of t and q such that $C(t,t_w) \simeq q$. Cases are $C(t = 2^{11}, t_w) = 0.95$ (red), $C(t = 2^{18}, t_w) = 0.85$ (gray). Black lines are Gaussian fits that are good only for q stat. (b) $(1-q)x_{t_w}(q)$ versus q from Eq. (2). Results have been fit to the 1RSB model: $(1-q)x_{t_w}(q) = 1-q$ for $q \geqslant q_{EA}$ (gray line) and $1 - q_{EA} + m(q_{EA} - q)$ for $q \leqslant q_{EA}$ with q_{EA} and m fitting parameters. We get $q_{EA} = 0.99$, $m = 0.30$ ($t_w = 2^{10}$, light red) and $q_{EA} = 0.98$, $m = 0.33$ ($t_w = 2^{14}$, light blue). (c) Alternative method based on Eq. (10) with fits taken from panel (b). (d) $x(q)$ obtained from the results in panel (b) by using Eq. (9) and compared with the 1RSB model (light-color lines). Panels (c) and (d) use the same color code as in panel (b). Simulations with $T = 0.2$, $N = 1000$, $h = 0.1$ and $10M$ quenches with one sample per quench.

III. SIMULATION TESTS

To test the validity of the theory we carried out extensive numerical simulations of three types of DFS. The first example is the random orthogonal model (ROM), a mean-field spin-glass model with one-step replica symmetry breaking (1RSB), defined by [14]

$$\mathcal{H} = - \sum_{1 \leqslant i < j \leqslant N} J_{ij} \sigma_i \sigma_j, \tag{11}$$

where $\sigma_i = \pm 1$ are Ising spins and $J_{ij} = J_{ji}$ are quenched Gaussian variables of zero mean and variance $1/N$ satisfying $\sum_k J_{ik} J_{kj} = 16 \delta_{ij}$, with $J_{ii} = 0$. The system is perturbed by applying a uniform magnetic field of strength h conjugated to the total magnetization $M(t) = \sum_i \sigma_i(t)$.

The ROM has a dynamical transition at $T_d = 0.536$ where the order parameter jumps from 0 above T_d to $q_{EA}(T_d) = 0.96$. Below T_d the system is dynamically confined into one of the exponentially many metastable states with $q(x) = q_{EA}(T)\theta(x - m(T))$. Both $q_{EA}(T)$ and $m(T)$ monotonically decrease with T, indicating the gradual freezing of degrees of freedom as $T \to 0$. Results for the ROM are shown in Figure 2 after quenching the system from $T = \infty$ to $T = 0.2$ for two values of t_w. As shown in Fig. 2(a) $P(\Delta S)$ for q stat are Gaussian whereas they are not for t stat [17]. Figure 2(b) plots $x_{t_w}(q)(1-q)$ versus $1-q$ [from Eq. (2)] fit to the theoretical prediction (8) with $\tilde{x}(q)$ the inverse of $q(x) = q_{EA}(T)\theta(x - m(T))$, i.e., $x(q \leqslant q_{EA}) = m$; $x(q > q_{EA}) = 1$.
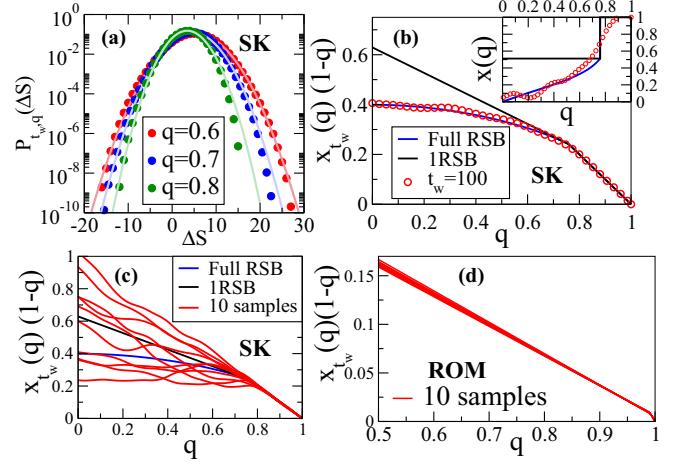

FIG. 3. Numerical test in the SK model and self-averaging behavior. (a) q-stat distributions for $t_w = 100$ for three values of q such that $C(t,t_w) \simeq q$. Light-color lines are Gaussian fits. (b) $(1-q)x_{t_w}(q)$ and $x(q)$ (inset) versus q from Eq. (2) for $t_w = 100$ (red) averaged over 10 samples. Results are compared with exact 1RSB and full RSB models [18] showing excellent agreement with the latter. (c) The same as in panel (b) but for the 10 samples showing large sample-to-sample fluctuations. (d) As a comparison we show the same as in panel (c) for the ROM where sample-to-sample fluctuations are small. Details of the fit are given in the caption of Fig. 2(b). Simulations for SK were done with $T = 0.3$, $N = 300$, $h = 0.1$.

Figure 2(c) shows the construction based on Eq. (10) where each point in the plot was obtained by extracting the mean and variance of $P_{t_w,q}(\Delta M)$ for a given value of q. Finally, Fig. 2(d) shows the $x(q)$ obtained by any of the two methods.

Our second example is the Sherrington–Kirkpatrick (SK) mean-field model described by Eq. (11) with uncorrelated Gaussian couplings of zero mean and variance $1/N$, which exhibits full RSB. Figure 3(a) shows again that $P(\Delta S)$ for q stat are Gaussian. From Eq. (2) we extract the $x_{t_w}(q)(1-q)$ which, averaged over 10 samples, excellently fit the full RSB theoretical prediction [18] [Fig. 3(b), main panel and inset]. A main feature of spin-glass theory is the presence of sample-to
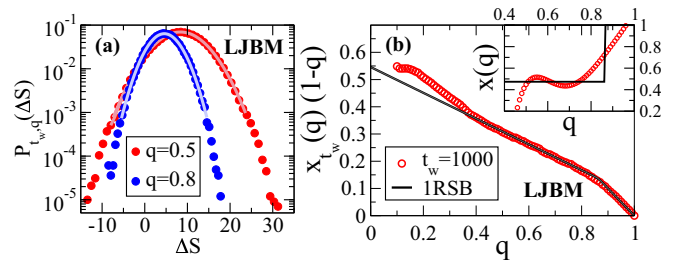

FIG. 4. Numerical test in the BMLJ model. (a) q-stat distributions for $t_w = 1000$. Light-color lines are Gaussian fits. (b) $(1-q)x_{t_w}(q)$ and $x(q)$ (inset) versus q from Eq. (2) for $t_w = 1000$ (red) compared with 1RSB with parameters taken from FD plots [17]. Energy and length parameters of the BMLJ (12) are measured in units of σ_{AA} and ϵ_{AA}: $\epsilon_{BB} = 0.5$, $\epsilon_{AB} = 1.5$, $\sigma_{BB} = 0.88$, and $\sigma_{AB} = 0.80$, and are taken to prevent crystallization [19]. The system has a reduced density $\rho = 1.2$ and exhibits a glass transition at $T_g \simeq 0.435$. Simulation parameters: $T = 0.3$, $N = 500$, $h = 0.1$, 20k runs.

sample fluctuations which are large for SK [Fig. 3(c)] as compared with ROM [Fig. 3(d)].

The $x(q)$ describes the statistics of the free-energy landscape of DFS at the level of individual samples, being also applicable to systems without quenched disorder. This is shown in Fig. 4 where we extract the $x(q)$ for the 80 : 20 binary mixture of type A and B particles interacting via a Lennard–Jones pair potential (BMLJ):

$$V_{\alpha\beta}(r) = 4\epsilon_{\alpha\beta}\left[\left(\frac{\sigma_{\alpha\beta}}{r}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r}\right)^{6}\right], \qquad (12)$$

where α, $\beta = A$, B; r is the distance between the two particles; and the parameters $\sigma_{\alpha\beta}$, $\epsilon_{\alpha\beta}$ stand for the effective diameters and well depths between species α, β. Remarkably, the $x(q)$ in the BMLJ [Fig. 4(b), inset] shows combined features of 1RSB (plateau at $q \lesssim 0.8$) and full RSB ($q \gtrsim 0.8$).

IV. CONCLUSIONS

We have introduced a FR to extract the order parameter $q(x)$ in DFS based on measuring entropy production fluctuations at a fixed value of the overlap q. Equation (1) and the GUH should be seen as key features of a weak-ergodicity-breaking scenario where q-stat classification of paths provides a operational definition of partially equilibrated states in rugged free-energy landscapes with quasi-FDT dynamics [cf. Eq. (6)]. The approach has been validated for three kinds of glassy systems being potentially applicable to other models and scenarios such as the low-temperature Gardner-like RSB phase. Initially

discovered in the Ising p-spin glass [20], the Gardner transition has been shown to describe the spectrum of low-energy excitations in jamming systems [21,22]. Moreover, it should be applicable to small systems [23], such as biomolecules [24] or microspheres trapped in random media [25], where thermal forces induce measurable energy fluctuations. A main difficulty in experiments is the measurement of the overlap between microscopic configurations, $q = x(t)x(s)$ where x stands for a reaction coordinate. Recent experimental studies [26] have shown the way to extract such quantities by measuring and testing the validity of the quasi-FDT in microspheres and single molecules mechanically driven to nonequilibrium steady states using steerable optical traps. Such measurements and the identification of suitable molecular systems exhibiting slow relaxational behavior reminiscent of DFS remains an exciting route to pursue in the near future. Nearly half a century ago, while mean-theory of the spin glass and replica symmetry breaking were developed, it was expected that mean-field results could ultimately be applied to finite-dimensional systems. The present work will hopefully shed light into this fundamental and still-unanswered question.

[1] N. Kawashima and H. Rieger, *Frustrated Spin Systems*, edited by H. Diep (World Scientific, Singapore, 2005), p. 491.

[2] A. Cavagna, Phys. Rep. **476**, 51 (2009).

[3] L. Berthier and G. Biroli, Rev. Mod. Phys. **83**, 587 (2011).

[4] L. Berthier *et al.*, Science **310**, 1797 (2005).

[5] S. F. Edwards and P. W. Anderson, J. Phys. F: Met. Phys. **5**, 965 (1975).

[6] G. Parisi, Phys. Rev. Lett. **50**, 1946 (1983).

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

[8] S. Franz and M. A. Virasoro, J. Phys. A: Math. Gen. **33**, 891 (2000).

[9] L. F. Cugliandolo, J. Kurchan, and L. Peliti, Phys. Rev. E **55**, 3898 (1997).

[10] A. Crisanti and C. De Dominicis, J. Phys. A: Math. Theor. **44**, 115006 (2011).

[11] A. Crisanti, H. Horner, and H.-J. Sommers, Z. Phys. B: Condens. Matter **92**, 257 (1993).

[12] L. F. Cugliandolo and J. Kurchan, Phys. Rev. Lett. **71**, 173 (1993).

[13] L. F. Cugliandolo, in *Slow Relaxations and Nonequilibrium Dynamics in Condensed Matter*, edited by J.-L. Barrat, M. Feiglman, J. Kurchan, and J. Dalibard (Springer, Berlin, 2004), pp. 367–521.

[14] A. Crisanti and F. Ritort, J. Phys. A: Math. Gen. **36**, R181 (2003).

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

[16] D. Hérisson and M. Ocio, Phys. Rev. Lett. **88**, 257202 (2002).

[17] A. Crisanti, M. Picco, and F. Ritort, Phys. Rev. Lett. **110**, 080601 (2013).

[18] A. Crisanti and T. Rizzo, Phys. Rev. E **65**, 046137 (2002).

[19] W. Kob and H. C. Andersen, Phys. Rev. Lett. **73**, 1376 (1994).

[20] E. Gardner, Nucl. Phys. B **257**, 747 (1985).

[21] M. Wyart, Phys. Rev. Lett. **109**, 125502 (2012).

[22] P. Charbonneau *et al.*, Nat. Commun. **5**, 3725 (2014).

[23] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).

[24] J. Camunas-Soler, A. Alemany, and F. Ritort, Science **355**, 412 (2017).

[25] J. R. Gomez-Solano, A. Petrosyan, and S. Ciliberto, Phys. Rev. Lett. **106**, 200602 (2011).

[26] E. Dieterich *et al.*, Nat. Phys. **11**, 971 (2015).