# Rethinking Mean-Field Glassy Dynamics and Its Relation with the Energy Landscape: The Surprising Case of the Spherical Mixed $p$-Spin Model

Giampaolo Folena,[1,2] Silvio Franz,[1,2] and Federico Ricci-Tersenghi[1,3]

[1]*Dipartimento di Fisica, Università "La Sapienza", P.le A. Moro 5, 00185, Rome, Italy*
[2]*Université Paris-Saclay, CNRS, LPTMS, 91405, Orsay, France*
[3]*INFN, Sezione di Roma1, and CNR-Nanotec, Rome unit, P.le A. Moro 5, 00185, Rome, Italy*

The spherical $p$-spin model is a fundamental model in statistical mechanics of a disordered system with a random first-order transition. The dynamics of this model is interesting both for the physics of glasses and for its implications on hard optimization problems. Here, we revisit the out-of-equilibrium dynamics of the spherical mixed $p$-spin model, which differs from the pure $p$-spin model by the fact that the Hamiltonian is not a homogeneous function of its variables. We consider quenches (gradient descent dynamics) starting from initial conditions thermalized in the high-temperature ergodic phase. Unexpectedly, we find that, differently from the pure $p$-spin case, the asymptotic states of the dynamics keep memory of the initial condition. The final energy is a decreasing function of the initial temperature, and the system remains correlated with the initial state. This dependence disproves the idea of a unique "threshold" energy level attracting dynamics starting from high-temperature initial conditions. Thermalization, which could be achieved, e.g., by an algorithm like simulated annealing, provides an advantage in gradient descent dynamics and, last but not least, brings mean-field models closer to real glass phenomenology, where such a dependence is observed in numerical simulations. We investigate the nature of the asymptotic dynamics, finding an aging state that relaxes towards deep, marginally stable minima. However, careful analysis rules out simple generalizations of the aging solution of the pure model. We compute the constrained complexity with the aim of connecting the asymptotic solution to the energy landscape.

Subject Areas: Complex Systems, Statistical Physics

## I. INTRODUCTION

Understanding the relation between the dynamical behavior and the underlying energy landscape is a fundamental question in the physics of glassy systems. The same question arises in a broader interdisciplinary perspective in optimization, information theory, and computer science. In these fields, the comprehension of the performances of classes of iterative algorithms represents both a practical and a fundamental theoretical problem. One notable example, among many, is provided by supervised learning, where one formalizes data learning as the problem of minimizing a loss function that depends parametrically on the data [1].

The development of the mean-field theory of spin glasses has provided unifying concepts that, starting from glassy materials, also allowed deep insight in optimization, inference, and machine learning [2].

Simple mean-field spin-glass models have been fundamental to understand the emergence of glassy dynamics within the so-called random first-order transition (RFOT) scenario [3] since they provide a concrete realization of the phenomena observed in more realistic glass formers. Many recent reviews exist [4–6] that discuss in detail the RFOT and the simple spin-glass models related to it. The same mean-field spin-glass theory has recently been used to address fundamental problems in the theory of computation [2] and to model the jamming state via high-dimensional interacting spheres [7], thus showing that key concepts of this theory have a broad range of applications.

Theoretical progress has been paralleled by the invention of powerful optimization algorithms—such as simulated annealing [8], parallel tempering [9], and quantum annealing [10]—that mimic physical dynamics, or message passing algorithms, such as survey propagation [11] and approximate message passing [12], which are based on the mean-field equations for spin glasses.

The spin-glass model that, more than any other, has shaped our theoretical ideas about the relation between the energy landscape and the relaxation dynamics is the spherical $p$-spin model [13]. This model is a system of spherical spins that interact through a disordered and

long-range $p$-body Hamiltonian. Despite its simplicity and its abstract nature, the model is important. In a nutshell, its very explicit exact solution provides the complete description of glassy phenomena in the mean field, including the long-time dynamics [14,15], thermodynamics [13], and the structure of metastable states [16]. Arguably, this model makes the RFOT more explicit and makes it easier to produce theoretical predictions [17]. Thanks to these virtues, the same system has been proposed as a simplified model to rationalize the properties of the loss function and the dynamics of learning in deep networks [18] and in problems involving tensors such as tensor completion and tensor principal component analysis [19,20]. Last but not least, the model has been widely used in the mathematical physics community to understand spin-glass mean-field theory in rigorous terms [21].

The picture of the glassy dynamics that emerges from the $p$-spin model is very attractive. Below the temperature $T_{MCT}$, identified as an ideal mode-coupling transition [22], the dynamics following a quench from a random initial condition falls in an aging state [15]. In this state, despite the lack of equilibration, any memory of the initial state is lost and the energy of the system tends to a temperature-dependent threshold value extensively higher than the equilibrium one. A crucial prediction, which has deep consequences both for glassy physics and for algorithms, is that the same asymptotic energy is reached if instead of starting from a completely random condition the system is quenched from an initial condition well thermalized at some temperature above $T_{MCT}$ (see Fig. 1, left diagram).
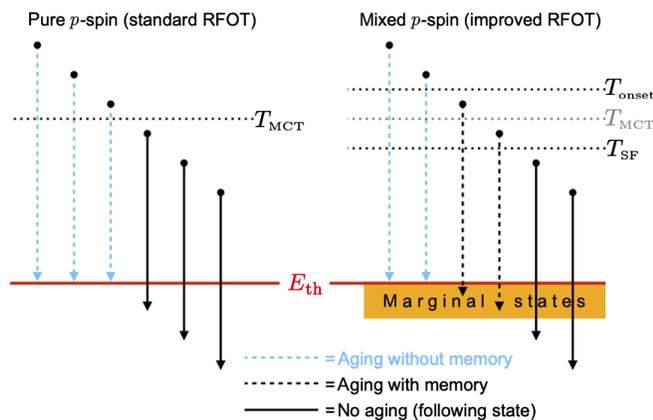


FIG. 1. Schematic representation of the different behavior of gradient descent dynamics in pure and mixed $p$-spin models. In the pure model (left diagram), all initial temperatures higher than $T_{MCT}$ lead to the same aging dynamics at the threshold energy, while starting below $T_{MCT}$ the dynamics is trapped within a state: Here, $T_{MCT}$ is the only relevant dynamical critical temperature. In the more generic mixed model (right diagram), two dynamical critical temperatures exist, $T_{onset}$ and $T_{SF}$, such that, starting from a temperature in between, the dynamics tends to a marginal state depending on the initial configuration while aging and keeping memory of the short-time evolution.

In the standard aging solution, the property that the asymptotic dynamics forgets the initial condition, as well as any configuration reached at finite times, is called weak ergodicity breaking [23]. It is essential to the solution of the long-time relaxation dynamics [15] because it allows us to decouple the dynamics at short and long times. It is believed to hold in any mean-field spin-glass model, although very recent results are raising some doubt [24].

In the context of glasses, the independence on the initial condition contrasts with simulations of realistic glass-formers. In these systems, one can define a temperature value $T_{onset}$ well in the liquid region, where relaxation crosses over from exponential to stretched exponential. Around this temperature and below, the states reached after quenching, deemed inherent structures (IS), have energy that is sensibly dependent on the starting temperature [25,26].

This mismatch has often been considered as a failure of mean-field theory to describe an important aspect of finite-dimensional glassy dynamics, the main argument being based on the assumption that the phenomenon is due to activated processes that are absent in mean-field models, where barriers are divergent with the system size. The net result is the lack of a theoretical comprehension of dynamical onset within RFOT.

For algorithms, the above picture rules out simulated annealing as a valid optimization protocol in systems with an ideal RFOT transition. Any time spent in the ergodic phase is irrelevant for optimization: In the long run, the system gets stuck at the universal threshold energy level. By extension, it has been hypothesized that algorithms running in a polynomial fail to optimize below threshold.

Unfortunately, despite its popularity, the $p$-spin model possesses a symmetry that makes it singular among mean-field spin glasses. Its Hamiltonian is a homogeneous function of its variables. While it is very well known that this homothetic symmetry puts in a one-to-one correspondence finite-temperature metastable state with energy minima, the independence of the low-temperature asymptotic energy on the initial condition is generally believed to be a generic feature of fully connected models.

In this paper, we present analytical evidence that this belief needs to be revised. As soon as we move to the mixed $p$-spin model [27,28] where the Hamiltonian is no longer homogeneous, the physical picture changes drastically. This model shares with the simpler pure $p$-spin model the property that one can write closed dynamical mean-field equations that can be integrated numerically with high precision at finite time and analyzed analytically in the asymptotic regimes.

Our main findings are schematically summarized in Fig. 1. In the mixed model, an onset temperature exists, such that below $T_{onset}$ the system never forgets the initial configuration, and it relaxes to marginal states whose energies depend on the initial temperature. If we define

the threshold energy as the value reached after a quench from a completely random point (infinite temperature), we find that in this model, the relaxation dynamics can go below threshold while aging and keeping memory of the short-time evolution. This picture is richer than the one that is usually associated with a RFOT, thus providing support for an *improved* RFOT already at the mean-field level.

(i) For glasses, we have a simple mean-field model, more generic than the pure $p$-spin model, allowing us to naturally incorporate the onset of glassy dynamics in the unifying theoretical framework of RFOT. This central result shows that, differently from what is commonly believed, genuinely finite-dimensional phenomena such as activation—but also interface dynamics or strongly spatially heterogeneous fluctuations—need not necessarily be invoked to explain the appearance of the dynamical onset.

(ii) For algorithms, we have that even in cases where the asymptotic energy turns out to be larger than the ground-state energy, thermalization in the high-temperature region is helpful. Relaxation dynamics can go below the threshold energy even without jumping over barriers. Thus, the present model gives rigorous support to the empirical observations that annealing works and suggests explanations that do not require activated processes.

Unfortunately, in the interesting regime where the energy goes below threshold, the asymptotic dynamics is not described by a simple generalization of the Cugliandolo-Kurchan aging solution [15]. The relation between the dynamics and landscape in mean-field models of glasses is more complex and also harder to solve than previously thought.

The paper is organized as follows. The next section briefly notes the salient features of mean-field glasses. In Sec. III, we define the model. In Sec. IV, we discuss the properties of the landscape and the number of stationary points of the energy surface of the model. Section V presents the equations describing the off-equilibrium dynamics and the effect of a nonrandom initial condition. Section VI constitutes the core of the paper, where we discuss the energy of inherent structures, the puzzling properties of the zero-temperature dynamics, and its relation to the landscape. The results are extended to relaxation in the presence of a small thermal noise in Sec. VII. We discuss an approximated dynamical solution and the emergence of an onset temperature in Sec. VIII. We comment on the results and their consequences in Sec. IX. The paper comprises a large number of Appendixes devoted to the discussion of more technical aspects of the work that support and complement the main text. In Appendix A, we present some formal solutions of the long-time asymptotic aging regime and its connections with the effective potential theory. An approximate solution to the asymptotic dynamics is analyzed in Appendix B. We present the computation of the constrained

complexity of energy minima in Appendix C. In Appendix D, we discuss finite-temperature dynamics in equilibrium. Appendix E deals with the properties of the response function at short times in the aging regimes. Finally, in Appendix F, we describe the numerical solution of the dynamical equations.

## II. SETTING THE STAGE

A widely used statistical characterization of the energy landscape sampled by the glassy dynamics is the numerical study of the inherent structures (ISs), defined as the local minima of the energy potential reached by a steepest descent procedure starting from well-thermalized initial conditions at temperature $T$ [25,29–32]. The thermalization temperature of the initial configuration is sometimes called the parent temperature, but here, when there is no ambiguity, we simply call it temperature. In simulations of glass-forming liquids, it is observed that starting from thermalized configurations at temperature $T$, the energy of the corresponding ISs concentrates around a value $E_{IS}(T)$ that depends on $T$ only below a characteristic onset temperature $T_{onset}$. Remarkably, this temperature is higher than the estimated "mode-coupling temperature" $T_{MCT}$, where mode-coupling theory would predict structural arrest. Similar effects are typical in finite-dimensional disordered systems and are also observed, e.g., in 3D Heisenberg spin glasses [33]. Here, simulated annealing is effective: In order to minimize the energy, equilibrating at lower and lower temperatures is a better strategy than a sudden quench from a high-energy state.

As stressed in the previous section, according to the common belief, the temperature dependence of the IS energy is a consequence of activation and should not appear in models with long-range interactions: Initial conditions thermalized above $T_{MCT}$ would all have the same value of the IS energy; any memory of the initial condition would be lost after a long enough time.

However, this belief is based on the solution of the spherical *pure $p$-spin* model [13], which, in a nutshell, gives the following dynamical picture: (1) From the point of view of equilibrium dynamics, the model provides an exact realization of the mode-coupling theory scenario. One finds dynamical freezing at a temperature $T_{MCT}$, where the free energy does not present singularities [14]. (2) Randomly chosen initial conditions evolved according to the Langevin dynamics at temperature $T_f < T_{MCT}$ fail to equilibrate. They fall in an aging state, which fails to attain a time-translation-invariant long-time regime and present a nontrivial response to perturbations [15]. (3) Any memory of the initial condition is lost in the aging state; the choice of an initial condition thermalized at temperatures $T_{in} > T_{MCT}$ would not change the asymptotic of a dynamics in the presence of a bath at temperature $T_f < T_{MCT}$. There is a

universal threshold energy $E_{th}$ that attracts the zero-temperature dynamics of any such initial conditions, i.e., $E_{IS}(T) = E_{th}$ for all $T > T_{MCT}$. (4) Initial conditions thermalized at $T < T_{MCT}$ live in metastable states that can be followed down in temperature [27,34,35]. In that case, one reaches an energy $E_0(T)$, which is monotonous in $T$ and smaller than the threshold value $E_{th}$. Unfortunately, such initial conditions cannot be generated in a finite time.

Such a detailed dynamical picture is deeply rooted in the structure and organization of the stationary points of the energy and the free-energy landscapes [36]. (A) For $E < E_{th}$, almost all stationary points of the energy are isolated and stable minima. The stability gap of the minima (i.e., the minimum eigenvalue of the Hessian) is uniquely a function of the energy. Saddle points are exponentially suppressed and have only a finite number of unstable directions. (B) For $E > E_{th}$, almost all stationary points are saddles. (C) Minima with $E = E_{th}$ are marginal. Dynamics is attracted by these threshold states, no matter what the thermalization effort is at $T > T_{MCT}$. It is clear that simulated annealing is of no use for such a system. (D) Below-threshold energy minima are separated by extensive barriers. There is a one-to-one correspondence between energy minima and finite-temperature metastable states (minima of the Thouless-Anderson-Palmer free-energy [16]). Metastable states can be followed up and down in temperature.

The general picture above, both for the dynamics and for the structure of the landscape, is not restricted to the spherical $p$-spin model; it is generic to mean-field models with a discontinuous glass transition (one-step replica symmetry breaking, or 1RSB in spin-glass jargon) and has been confirmed by the exact description of glasses of particles in the limit of infinite dimensions [37]. However, there are important aspects of the physics of the spherical $p$-spin that are deeply model-specific. Its Hamiltonian is homogeneous and does not allow for bifurcations or merging of metastable states while changing the temperature. The energy landscape is simpler than more generic mean-field models. It is well known that in other mean-field modes like the Ising $p$-spin model, the Potts glass or even spherical models with mixtures of multibody interactions [27], when metastable states are followed down in temperature, bifurcations are encountered. This result leads to the glass-to-glass Gardner phase transition, which has received much attention recently [38].

The study of metastable states below such a transition reveals puzzling aspects. The mean-field solutions describing the evolution of states in temperature may unexpectedly disappear [39] as the temperature is lowered. This anomalous behavior has not received a satisfactory explanation, and its implications are not very well studied. In this paper, we reexamine dynamics in models where these phenomena occur and challenge crucial aspects of the $p$-spin dynamical picture: the emergence of a universal dynamical threshold and the loss of memory of high-temperature initial conditions.

## III. MODEL DEFINITION

We concentrate on the mixed $p$-spin spherical model. This model, while presenting a more generic phenomenology, keeps much of the analytical simplicity of the pure spherical $p$-spin. In the mixed model, states bifurcate, and the above-mentioned anomalous behavior occurs. The Hamiltonian of the mixed $p$-spin model is just a sum of $p$-spin interacting terms with different $p$ values,

$$H_J[\sigma] = -\sum_p \sqrt{a_p} \sum_{i_1 < \ldots < i_p}^{N} J_{i_1 \ldots i_p}^{(p)} \sigma_{i_1} \ldots \sigma_{i_p}. \quad (1)$$

Choosing the couplings $J_{i_1 \ldots i_p}^{(p)}$ as independent random Gaussian variables of zero mean and variance, $\mathbb{E}[(J_{i_1 \ldots i_p}^{(p)})^2] = (p!/2N^{p-1})$, and the coefficients $a_p \geq 0$ with $\sum_p a_p < \infty$, we find that $H_J$ is a random Gaussian function on the $N$-dimensional sphere ($\sum_i \sigma_i^2 = N$) with covariance

$$\overline{H[\sigma]H[\tau]} = Nf(q_{\sigma\tau}) \quad \text{where}$$

$$f(q) = \frac{1}{2}\sum_p a_p q^p \quad \text{and} \quad q_{\sigma\tau} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i \tau_i.$$

We choose $f(q)$ as a polynomial such that the model has a RFOT in thermodynamics, associated with an ideal mode-coupling transition in equilibrium dynamics. In addition, we exclude from our analysis models that can present a Gardner transition [38] to a continuous replica symmetry-breaking state. It is well known that a sufficient condition for this transition is that the function $f$ is such that $1/\sqrt{f''(q)}$ is a convex function of $q$, a property that we assume throughout the paper. The pure $p$-spin model corresponds to the case of a single monomial $f(q) = \frac{1}{2}q^p$; we study a more generic class of polynomials $f$ where at least two coefficients are positive (mixed $p$-spin model). While we keep the function $f$ generic in formulas, all of our numerical results are presented for $f(q) = \frac{1}{2}(q^3 + q^4)$ and compared with the pure case with $p = 3$. In other examples, we checked that all of the results we get in this particular case are typical of general mixed models.

To investigate how much of the above picture based on the solution of the pure $p$-spin model is generic, we perform a high-precision integration of the dynamical equations describing the evolution of the system in the thermodynamic limit, starting from a thermalized initial condition at temperature $T$ and subsequently undergoing a gradient descent dynamics. Our main finding is that the above picture is not generic, and inhomogeneous models show several unexpected features. The value of the asymptotic energy predicted under the assumption of aging with

loss of memory is only found starting from high enough temperatures. We find strong evidence for the existence of an onset temperature $T_{\text{onset}} > T_{\text{MCT}}$ below which the energy of the IS depends on the initial temperature $T$ and goes below the threshold value.

## IV. ENERGY LANDSCAPE

In order to put our dynamical results in the right framework, we first discuss the structure and organization of the stationary point of the mixed Hamiltonian. The stationary points of the Hamiltonian $H[\sigma]$ on the sphere $\sum_i \sigma_i^2 = N$ verify

$$\frac{\partial H[\sigma]}{\partial \sigma_i} + \mu \sigma_i = H'_i + \mu \sigma_i = 0. \tag{2}$$

The parameter $\mu$, which we call the radial reaction force, is a Lagrange multiplier that ensures the spherical constraint. In any stationary point, it takes the value

$$\mu = -\frac{1}{N} \sum_i \sigma_i H'_i. \tag{3}$$

The radial reaction is directly related to the nature and the stability of the stationary points; in fact, the Hessian matrix (intended to be restricted to fluctuations on the sphere) reads

$$M_{ij} = \frac{\partial H[\sigma]}{\partial \sigma_i \partial \sigma_j} + \mu \delta_{ij} = H''_{ij} + \mu \delta_{ij}. \tag{4}$$

It is well known [36] and rigorously proven [40] that $H''_{ij}$ is a random matrix that belongs to the Gaussian orthogonal ensemble (GOE) with variance $\text{Var}[H''_{ij}] = (1/N)f''(1)$. Thus, the matrix $M$ has a semicircular spectral distribution $\rho(\lambda) = \sqrt{(\lambda - \mu)^2 - 4f''(1)}/(\pi\sqrt{f''(1)})$, which has a lower band edge at $\lambda_{\min} = \mu - 2\sqrt{f''(1)}$. We notice that the value of the radial reaction determines the nature of the stationary points [41]. It is natural to define a marginal value of the radial reaction $\mu_{\text{mg}} = 2\sqrt{f''(1)}$ that separates minima from saddles with an extensive number of negative directions. Stationary points of $H$ are overwhelmingly minima if $\mu > \mu_{\text{mg}}$, and they are saddles or maxima if $\mu < \mu_{\text{mg}}$. Notice that all marginal minima with vanishing spectral gaps lie on the manifold specified by $\mu(\sigma) = \mu_{\text{mg}}$. In the pure model, at any stationary point, the radial reaction and the energy verify $\mu = -pE$. In the mixed model, there is no such one-to-one relation between $\mu$ and $E$ in stationary points. This case can be seen by computing the number of stationary points in Eq. (2), for which both $E$ and $\mu$ are fixed, $\mathcal{N}(E, \mu) = e^{N\Sigma(E,\mu)}$.

The computation of the complexity $\Sigma(E, \mu)$ was performed in Ref. [42], and we quote the result here [for
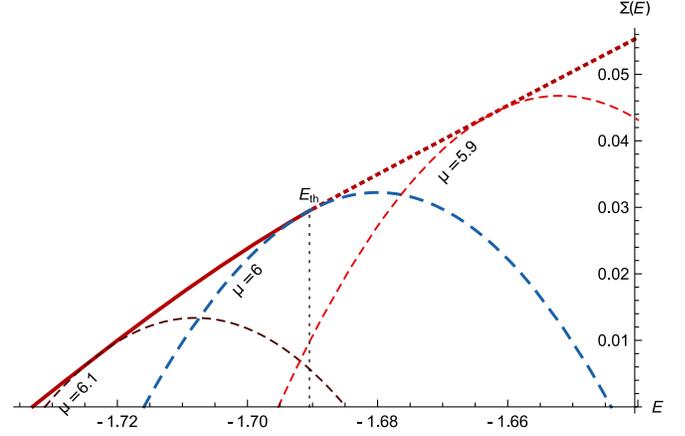


FIG. 2. Complexity $\Sigma(E, \mu)$ of the $(3 + 4)$-spin model as a function of the energy $E$ for different values of $\mu$. The solid line ($E < E_{\text{th}}$) represents the complexity of the dominant stable minima ($\mu > \mu_{\text{mg}} = 6$); its continuation above $E_{\text{th}}$ (short-dashed line, $\mu < \mu_{\text{mg}} = 6$) represents the complexity of the dominant saddles. The parabolic long-dashed lines below represent the complexity fixing the radial reaction: $\mu = 6$ corresponds to marginal minima (dashed blue line), $\mu > 6$ to stable minima, and $\mu < 6$ to saddles. For the same value of $E$, stationary points of different nature coexist, and more importantly, marginal states ($\mu = 6$) exist in a broad energy range $E \in [-1.716, -1.644]$.

completeness, a physicist's derivation and the explicit formula for $D(\mu)$ is provided in Appendix C],

$$\Sigma(E, \mu) = D(\mu) - \frac{E^2 f''(1) + (E + \mu)^2 f'(1)}{2\{f(1)[f''(1) + f'(1)] - f'(1)^2\}}. \tag{5}$$

The above expression holds when it returns a positive value.

Fixing $\mu$, the complexity $\Sigma(E, \mu)$ is a parabola in $E$ with a curvature that is finite for mixed models [see Fig. 2, where we show the complexity for the $(3 + 4)$-spin model] and that tends to infinity in the pure limit. For $\mu < \mu_{\max}$, with $\mu_{\max}$ implicitly defined by $0 = \Sigma(E_{\max}, \mu_{\max}) = \partial_E \Sigma(E_{\max}, \mu_{\max})$ with $E_{\max}$ the location of the parabola maximum, there is a finite interval of energies $[E_-(\mu), E_+(\mu)]$ such that, within the interval, $\Sigma(E, \mu) > 0$. Most importantly, this is true for the threshold value of the radial reaction $\mu_{\text{mg}}$. Seen from the point of view of $\mu$, the organization of the stationary points resembles the one of the pure model; however, from the point of view of the energy, there are deep differences. In the pure model, for a given level of energy, the index of any stationary point [43] differs, at most, by a finite number from the index of dominating ones. In the mixed models, for a given energy, stationary points coexist with macroscopically different indices: i.e., their spectra are shifted by a finite amount. In particular, as it is apparent in Fig. 2, there is a whole interval in energy where marginally stable states exist. A unique

relation between $E$ and $\mu$ only holds for the exponentially dominating stationary points.

## V. DYNAMICAL EQUATIONS IN THE OUT-OF-EQUILIBRIUM REGIME

Given the Hamiltonian in Eq. (1), we consider the following Langevin dynamics:

$$\partial_t \sigma_i(t) = -\mu(t)\sigma_i(t) - \frac{\partial H}{\partial \sigma_i}(t) + \xi_i(t),$$

$$\langle \xi_i(t)\xi_j(t')\rangle = 2T_f \delta_{ij}\delta(t-t'),$$

$$P[\sigma(0)] = \frac{e^{-\beta H(\sigma(0))}}{Z(\beta)}, \qquad (6)$$

where $T_f$ is the temperature of the thermal bath (later set to zero) and $\beta = 1/T$ is the inverse temperature at which the

initial condition is equilibrated. The time-dependent radial reaction force $\mu(t)$ is the Lagrange multiplier that constrains the dynamics on the sphere, and at time $t$, it takes the value

$$\mu(t) = T_f - \frac{1}{N}\sum_i \sigma_i(t)\frac{\partial H}{\partial \sigma_i}(t). \qquad (7)$$

Taking the thermodynamic limit, $N \to \infty$, at finite times, Eq. (6) implies closed integrodifferential equations [14,44,45] for the correlation $C(t,t') \equiv \langle \sigma_i(t)\sigma_i(t')\rangle$ and for the response $R(t,t') \equiv [\partial\langle\sigma_i(t)\rangle/\partial\xi_i(t')]$ that vanishes for $t < t'$ and verifies $R(t^+,t) = 1$. For an initial temperatures $T$ greater than the Kauzmann temperature of the thermodynamic phase transition $T_K$, the dynamical equations, valid for $t > t'$, read [35]

$$\partial_t C(t,t') = -\mu(t)C(t,t') + \int_0^t ds f''(C(t,s))R(t,s)C(s,t') + \int_0^{t'} ds f'(C(t,s))R(t',s) + \beta f'(C(t,0))C(t',0),$$

$$\partial_t R(t,t') = -\mu(t)R(t,t') + \int_{t'}^t ds f''(C(t,s))R(t,s)R(s,t'),$$

$$\mu(t) = T_f + \int_0^t ds f''(C(t,s))R(t,s)C(t,s) + \int_0^t ds f'(C(t,s))R(t,s) + \beta f'(C(t,0))C(t,0),$$

with energy given by $E(t) = -\int_0^t f'(C(t,s))R(t,s)ds - \beta f(C(t,0)).$ \qquad (8)

The initial condition enters the above equations in a rather simple way. For $\beta = \beta_f$, one gets equilibrium dynamics, whose MCT transition is illustrated in Appendix D. For $\beta = 0$, the equations reduce to the usual form that is valid for an uncorrelated initial condition. In this case, for $T_f < T_{\mathrm{MCT}}$, the equations admit a Cugliandolo-Kurchan weak ergodicity-breaking aging solution completely analogous to the one of the pure model. It is possible to show that in this solution, for $T_f \to 0$, the radial reaction $\mu(t)$ tends to the marginal value $\mu_{\mathrm{mg}}$ for $t \to \infty$, while the energy tends to the threshold value $E_{\mathrm{th}} = \{-f'(1)^2 + f(1)[f'(1) + f''(1)]/f'(1)\sqrt{f''(1)}\}$. In correspondence with this value of energy, the complexity $\Sigma(E_{\mathrm{th}},\mu)$, as a function of $\mu$, is maximum for $\mu = \mu_{\mathrm{mg}}$. In other words, $E_{\mathrm{th}}$ is not the energy of the most numerous marginal states but the energy where marginal states exponentially dominate the complexity.

It is possible to see that for $T < \infty$, the same aging solution solves the dynamical equations in the long-time limit, if one assumes that memory of the initial condition is lost and $C(t,0) \to 0$ at a large time. If this is the correct description for all $T > T_{\mathrm{MCT}}$, then the asymptotic energy would be independent of $T$ in this domain.

We study the dynamical equations in Eq. (8) by numerical integration. The simplest possibility, which we

adopt here, is a fixed time step $\Delta t$, first-order Euler discretization algorithm that gives very reliable results at short times [46] and allows for reliable extrapolations in the $\Delta t \to 0$ limit (see Appendix F). Other algorithms with variable time steps have been proposed and used in the literature [47,48]. Unfortunately, in the case of a mixed $p$-spin model with a correlated initial condition ($\beta > 0$), these algorithms appear to be unstable at short times and do not allow any improvement over the simplest one. With the Euler integration scheme and a maximum step of $\Delta t = 0.1$, we reach times of order $10^3$. These timescales are often not large enough to allow for a simple naive extrapolation to the infinite-time limit. Nonetheless, under some conditions, they will allow us to make clear claims on the large-time dynamics. In this paper, we concentrate on the case of zero-temperature dynamics, where the energy monotonically decreases and, once below threshold, it cannot increase again. However, we show in Sec. VII that this $T_f = 0$ result extends to a range of positive temperatures, $T_f > 0$.

## VI. AFFINITIES AND DIVERGENCES BETWEEN MIXED AND PURE MODELS

Here, we discuss the core question of this paper and compare the energy reached by the $T_f = 0$ gradient descent
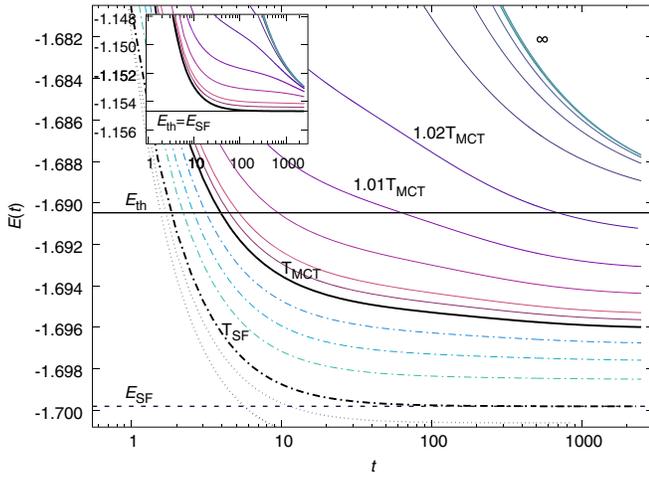
FIG. 3. Energy relaxation in the $(3+4)$-spin model, starting from different temperatures $T$ and quenching to zero temperature. Solid lines are quenches from the ergodic phase $(T \geq T_{\mathrm{MCT}} = 0.805166)$; from bottom to top, $T/T_{\mathrm{MCT}} = 1$, 1.001, 1.002, 1.005, 1.01, 1.02, 1.05, 1.1, 1.2, 1.5, 2, 4, 10, $\infty$. Dashed-dotted lines correspond to quenches from temperatures $T_{\mathrm{SF}} = 0.9914 T_{\mathrm{MCT}} = 0.798 \leq T < T_{\mathrm{MCT}}$; from top to bottom, $T/T_{\mathrm{MCT}} = 0.998$, 0.996, 0.994, 0.992, and at $T_{\mathrm{SF}}$. Dotted lines correspond to $T < T_{\mathrm{SF}}$ $(T/T_{\mathrm{MCT}} = 0.99, 0.988)$. We clearly see that the energy goes below threshold for temperatures that are small enough, although larger than $T_{\mathrm{MCT}}$. Inset: Same dynamics for the three-spin model $(T \geq T_{\mathrm{MCT}})$; the energy never goes below threshold.

dynamics from different initial conditions to the threshold energy that can be computed from the asymptotic solution with $\beta = 0$ [15] or by computation of the complexity of minima [16]. For the $3+4$ model, the value of the threshold energy is $E_{\mathrm{th}} = -71/42 \simeq -1.6905$. In Fig. 3, we show the curves of the energy as a function of time for different $T$, ranging from $T = \infty$ to $T = 0.796$, the dynamical transition temperature of the model being $T_{\mathrm{MCT}} = 0.805166$ (see Appendix D). Data are obtained from an integration with step $\Delta t = 0.1$, reaching the maximum time of 2500. This step size gives a relative integration error smaller than $10^{-5}$ (see detailed discussion in Appendix F).

On the timescale we can reach and for $T \gtrsim 1$, the energy seems to have reached an asymptotic behavior well fitted by a power law, $E(t) = E_{\mathrm{th}} + a/t^{\gamma}$, with $a$ slightly dependent on temperature and $\gamma = 0.66 \pm 0.01$, making us confident that $E_{\mathrm{th}}$ is the asymptotic value of the energy in this range of temperatures. On the other hand, if $T$ is small enough but still larger than $T_{\mathrm{MCT}}$, the curves of the energy go below $E_{\mathrm{th}}$. For comparison, in the upper inset we present the same curves for the pure model with $p = 3$. In that case, the energy manifestly tends to the threshold value for all $T \geq T_{\mathrm{MCT}}$. These results suggest a scenario with a new dynamical transition in the mixed model, with an onset temperature $T_{\mathrm{onset}}$ separating a memoryless phase for

$T > T_{\mathrm{onset}}$ from a phase with memory at $T < T_{\mathrm{onset}}$. It is also interesting to look at $T < T_{\mathrm{MCT}}$. In our data, we do not observe anything special happening at $T_{\mathrm{MCT}}$; it is only below a state-following temperature $T_{\mathrm{SF}} \approx 0.798$ that we observe fast relaxation and the energy decaying exponentially to its asymptotic value. We can predict the asymptotic energy in this region through a quasistatic solution of the dynamical equations ("state following" with memory of the initial condition). In this region, the dynamics just consists in a simple relaxation inside a stable energy minimum, which conserves its identity as a metastable ergodic component when the temperature is changed. In principle, the state-following procedure could also be used above $T_{\mathrm{SF}}$, but we know, in that regime, it provides just an approximate solution that eventually disappears if the temperature is too high [39]. We discuss this solution in more detail in Appendixes A and B. In the next section, we show that a small temperature $T_f$ in the dynamical equations does not change this behavior.

In order to understand the dynamical mechanisms that allow us to beat the threshold, we look at the behavior of the response and correlations. The first quantity that we study is the correlation with the initial state, $C(t, 0)$. According to the usual weak ergodicity-breaking scenario, this quantity should vanish at large $t$. Both in the mixed and in the pure model, we observe that the relaxation of $C(t, 0)$ is slower and slower as $T$ decreases. As for the energy, we can identify three temperature regimes for $T > T_K$: (I) a high-temperature standard aging regime $T > T_{\mathrm{onset}}$, where memory of the initial condition is lost and $C(t, 0) \to 0$; (II) an intermediate *hic sunt leones* regime $T_{\mathrm{SF}} < T < T_{\mathrm{onset}}$, where the relaxation is slow, any extrapolation is difficult, but evidence suggests aging with memory is taking place, i.e., $q_{12} = \lim_{t \to \infty} C(t, 0) > 0$; (III) a low-temperature state-following regime $T < T_{\mathrm{SF}}$, where $C(t, 0)$ relaxes exponentially fast to $q_{12} > 0$.

While no aging is found in regime III, in regimes I and II, we find aging that is qualitatively similar in the two cases. In Fig. 4, we show $C(t, s)$ as a function of $t - s$ for several
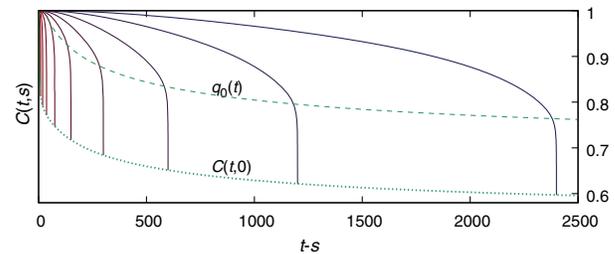


FIG. 4. Correlation function $C(t, s)$ for initial temperature $T = 0.813 = 1.01 T_{\mathrm{MCT}}$ as a function of $t - s$ for fixed values of $t = 2400 \times 2^{-n}$, with $n = 0, \ldots, 9$ (from right to left). The correlation shows clear aging features, becoming slower for larger times. After the slow decay, at a well-defined value $q_0(t)$ (dashed line), the correlation displays a sudden drop to $C(t, 0)$ (dotted line).
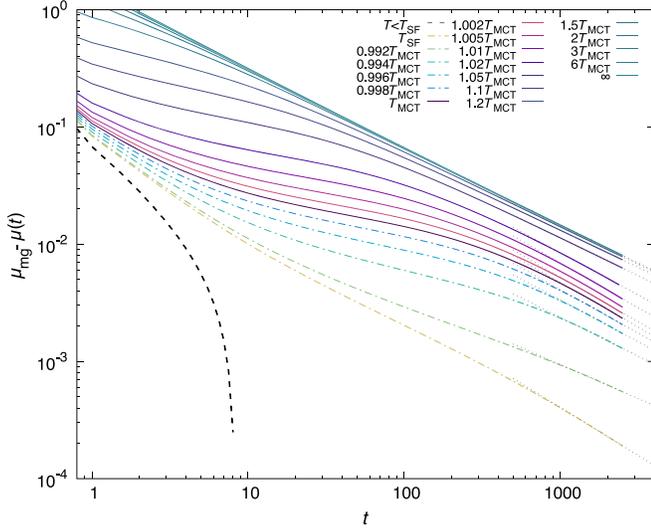
FIG. 5. Decay of the radial reaction with time in the $(3+4)$-spin model on a double logarithmic scale. Its asymptotic value is compatible with $\mu_{mg} = 6$ for $T \geq T_{SF}$ and goes above $\mu_{mg} = 6$ for $T < T_{SF}$. This result proves that dynamics tends towards marginal minima for $T > T_{SF}$.

times $t$ and initial temperature $T = 1.01T_{MCT}$. The curves show a clear aging behavior, with the dynamics decorrelating more slowly as time passes, followed by a fast drop of the correlation for small values of $s$ to the value $C(t, 0)$.

Even though the energy goes below threshold, we find compelling evidence that the dynamics remains critical and approaches asymptotically, marginally, stable minima. We reach this conclusion by studying the asymptotic behavior of the radial reaction $\mu(t)$, which, for $t \to \infty$, determines the spectral gap of the asymptotically visited states. In Fig. 5, we can clearly see that both in region I and II, the radial reaction tends to the marginal value $\mu_{mg} \equiv 2\sqrt{f''(1)}$ of marginal minima. In region III, by contrast, the system reaches stable minima and $\mu(\infty) > \mu_{mg}$.

Aging behavior is often qualified by studying the relation between the response and correlation [49]. Figure 6 shows the parametric plot of the integrated response $\chi(t, s) \equiv \int_s^t R(t, u) du$ versus $C(t, s)$ as a function of $s$ for various values of $t$. In zero-temperature dynamics, in the large-time limit, the integrated response has a jump in $C = 1$ to the intrastate susceptibility $\chi_{EA}$. In a given minimum,

$$\chi_{EA} = \int d\lambda \frac{\rho(\lambda)}{\lambda} = \frac{\mu + \sqrt{\mu^2 - 4f''(1)}}{2f''(1)}, \quad (9)$$

and in marginal minima, $\mu = \mu_{mg}$ implies $\chi_{EA} = \chi_{mg} \equiv 1/\sqrt{f''(1)}$. The curves in Fig. 6 show the formation of a jump in $C = 1$ that is very well compatible with this value.
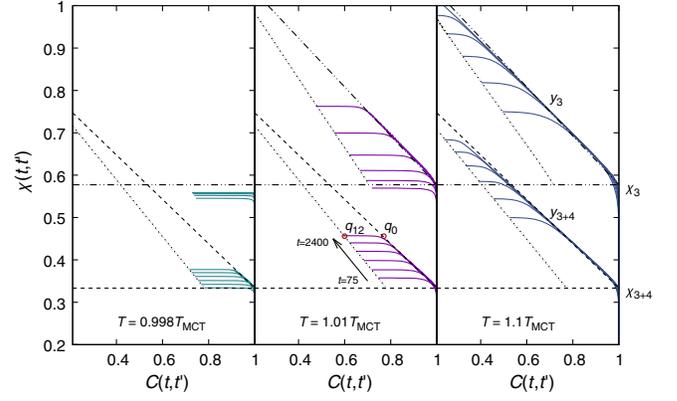


FIG. 6. Integrated response $\chi(t, t')$ versus correlation $C(t, t')$ for three different initial temperatures $T/T_{MCT} = 0.998, 1.01, 1.1$. The lines are plotted by fixing (from bottom to top) $t = 75, 150, 300, 600, 1200, 2400$ and changing $t'$. In all the panels, upper data are for the three-spin model, and lower data are for the $(3+4)$-spin model. The dash-dotted/dotted lines indicate the susceptibility $\chi_{mg}$ and the fluctuation dissipation ratio $y_0$ predicted by the memoryless solution. It is interesting to note that in the $(3+4)$-spin model, the dynamics does not have any appreciable change at $T_{MCT}$, while in the three-spin model, starting below $T_{MCT}$, the system remains trapped in a state.

The memoryless solution [15] predicts a linear behavior for $C < 1$ with the slope given by the "fluctuation-dissipation ratio" $y_0 \equiv \sqrt{f''(1)}/f'(1) - \chi_{mg}$.

On the timescales we can access, both the pure and the mixed models show, for all the values of $T > T_{SF}$, after the jump to $\chi \simeq \chi_{mg}$, an approximately linear part. The linear behavior continues until a rather well-defined time-dependent value of the correlation, $q_0(t) > C(t, 0)$, where the response essentially stops increasing. The value $q_0(t)$ identified in Fig. 4 marks the value that separates slow dynamics, where an effective temperature emerges, from fast partial decorrelation from the initial condition, during which the system does not respond (see Appendix E). In the memoryless phase, $T > T_{onset}$, both $q_0(t)$ and $C(t, 0)$ go to zero for $t \to \infty$, while they tend to nonzero limits in the aging-with-memory phase.

Having established that the system relaxes towards marginal minima, the question is which marginal minima are selected. In an attempt to address this question, we study the constrained complexity [50] $\Sigma(E, \mu, q_{12}; T)$ of energy minima of fixed radial reaction $\mu$, energy $E$, and correlation $q_{12}$ from a reference configuration thermalized at temperature $T$. The quantity for $q_{12} = 0$ reduces to the one in Eq. (5) studied in Sec. IV. The detailed computation of this constrained complexity is presented in Appendix C. The computation reveals that for $q_{12} > 0$, we have qualitatively the same picture illustrated in Fig. 2, but with $\Sigma(E, \mu, q_{12}; T) < \Sigma(E, \mu)$, with $\Sigma(E, \mu, q_{12}; T)$ monotonically decreasing in $q_{12}$. One can define a generalized threshold energy that depends on $T$ and $q_{12}$, $E_{th}(q_{12}, T)$,

where marginal states dominate the constrained complexity. It turns out that $E_{th}(q_{12}, T)$ is a decreasing function of $q_{12}$ and $T$. Unfortunately, as explained in detail in Appendix C, while the resulting $E_{th}(q_{12}, T)$ has the right qualitative behavior, the corresponding minima have features that are incompatible with the actual attractors of the dynamics. The complete static characterization of the attractors of the dynamics remains an important open problem.

## VII. EXTENDING RESULTS TO FINITE-TEMPERATURE RELAXATION DYNAMICS

As we said in Sec. V, for $T_f = 0$, the energy monotonically decreases, while if the bath temperature $T_f$ is positive, this is not necessarily the case. If we integrate the dynamical equations for a small temperature $T_f$, we observe that the finite-time energy is continuous, $E(t, T_f) = E(t, T_f = 0) + o(T_f)$. We show that, for small enough $T_f$, this implies that $\lim_{t \to \infty} E(t, T_f) < E_{th}(T_f)$.

At $T_f > 0$, the Lyapunov function of the Langevin dynamics, which decreases monotonically in time, is the free-energy functional [51]:

$$F(t) = \int d\sigma P_t[\sigma] H[\sigma] + T_f \int d\sigma P_t[\sigma] \log P_t[\sigma]$$
$$\equiv E(t) - T_f S(t),$$

where $P_t[\sigma]$ is the probability distribution of configurations at time $t$ induced by the initial condition and the dynamics.

In principle, the energy function $E(t)$ could be non-monotonous in time for $T_f > 0$ because an entropy increase $\Delta S(t) = S_\infty - S(t) > 0$ can induce an energy increase $\Delta E > 0$. However, since the free energy is a decreasing function of time, we have $\Delta F = \Delta E - T_f \Delta S < 0$, and the energy increase is upper-bounded by $\Delta E < T_f \Delta S$. This quantity tends to zero for $T_f \to 0$. Consequently, if the temperature is small enough, the energy $E(t)$, which is below threshold by an $O(1)$ amount, cannot be pushed back to the threshold value.

In Fig. 7, we show the energy reached at several times $t \in \{125, 250, 500, 1000\}$ during the integration of the dynamical equations as a function of the temperature $T_f$. The topmost solid line is the analytic prediction for the threshold energy as a function of $T_f$. We observe that the dynamical energy lies below the threshold energy and shows a continuous and nicely linear behavior in $T_f$. The weak time dependence of the data shown in Fig. 7 suggests the energy is close to its asymptotic value. Even insisting that an eventual entropy increase may induce an energy increase at later times, this increase would tend to zero for small $T_f$, and thus, by continuity, there would be a finite $T_f$
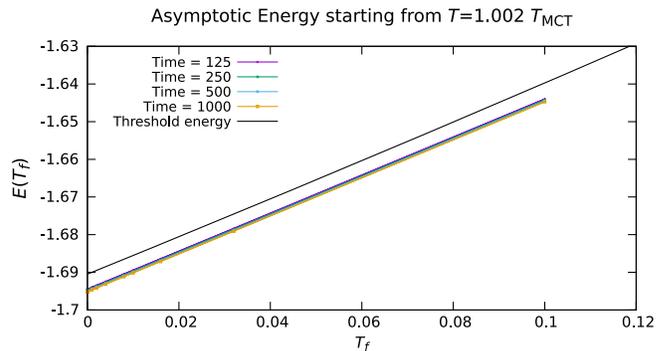


FIG. 7. The dependence of the dynamical energy of the $(3 + 4)$-spin model as a function of the temperature $T_f$. Data measured at different times are smooth and linear in $T_f$ and clearly below the threshold energy (shown by the upper line). At small enough $T_f$, such an energy difference cannot be compensated by an entropy increase, as discussed in the text.

range where the energy would certainly remain below the threshold value even at an infinite time.

## VIII. APPROXIMATE ASYMPTOTIC SOLUTION TO THE DYNAMICS

The dependence of the final energy on $T$ in regime II witnesses a form of memory of the initial condition. In contrast, thermalized configurations lie in basins of attraction of different marginal states. The asymptotic ansatz of Cugliandolo and Kurchan has been generalized in Ref. [27] by supposing "aging within a metabasin" and a nonvanishing correlation with the initial state $\lim_{t \to \infty} C(t, 0) = q_{12} > 0$. Unfortunately, the resulting equations for $\chi$, $y$, $q_{12}$, $q_0$ only have an aging solution that is different from the "amnesic" one ($q_{12} = q_0 = 0$) in the interval $T_{SF} \leq T \leq T_0 = 0.803 < T_{MCT}$. Even when the solution exists, the values of the various parameters do not match the observations; in particular, one finds $y \approx 0.3$, while in dynamics, $y \approx 0.52$. This aging solution terminates at the temperature $T_{SF}$, where $q_0$ tends to 1. Below that temperature, there is a stable state-following solution with no aging, $q_{12} > 0$, $\mu > \mu_{mg}$ and $\chi < \chi_{mg}$ (see Appendix A for a detailed derivation). This solution correctly reproduces the asymptotic values of the energy, correlation, and radial reaction of the large-time dynamics for temperatures $T \lesssim T_{SF}$. Despite all the efforts, we have been unable to find a different asymptotic dynamical ansatz describing aging with memory in the *hic sunt leones* regime II and/or to find in the numerics an indication of where the memory of the initial condition could affect the asymptotic solution. The lack of simple aging solutions is certainly related, and equally paradoxical, to the phenomenon of a lost solution in the state-following procedure of Ref. [39].

Insisting on using a 1RSB dynamical ansatz with $\chi_{EA} = \chi_{mg}$, $\mu = \mu_{mg}$, and $y \simeq y_0$, which is supported by
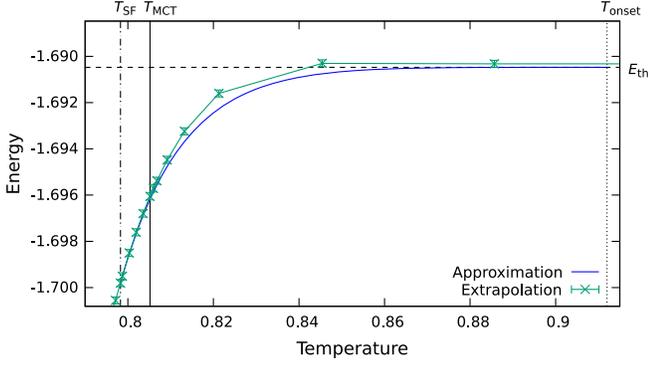
FIG. 8. Numerically extrapolated asymptotic energy (data with errors) and approximated asymptotic solution of dynamical equations (solid line).

the numerical solution, one can easily derive the following relation:

$$y_0 f'(q_0) q_0 = \beta f'(q_{12}) q_{12}. \qquad (10)$$

A second relation can be obtained from the observation that $\chi_{\text{tot}}(t) \equiv \int_0^t ds R(t, s) \approx \chi_{\text{mg}} + y_0(1 - q_0(t))$ depends approximately linearly on $C(t, 0)$ even at finite times (see dotted lines in Fig. 6). Moreover, this relation does not present an appreciable temperature dependence in the range $T_{\text{SF}} \leq T \lesssim 1$ (data shown in Appendix B) and implies a linear relation between $q_{12}$ and $q_0$ that can be estimated from two analytically known limits: For $T$ large enough, $q_{12} = q_0 = 0$, while at $T = T_{\text{SF}}$, we have $q_0 = 1$ and $q_{12,\text{SF}} = \sqrt{1 - f'(1)/f''(1)}$. Assuming the relation $q_{12} = q_{12,\text{SF}} q_0$, Eq. (10) admits a solution with $q_0 > 0$ if $T < T_{\text{onset}} \equiv q_{12,\text{SF}}^k / y_0$, where $k$ is defined by $f(q) \propto q^k$ for $q \to 0$ (in the $3 + 4$-model $T_{\text{onset}} = 0.91$). This approximated solution gives an asymptotic energy close to the one extrapolated from the numerical integration (see Fig. 8). Even though this is not an exact solution of the asymptotic equations, it gives us a strong indication that the passage from memoryless aging to aging with memory could be marked by a phase transition. In addition, it shows that the exact solution cannot be too far from a 1RSB weak ergodicity-breaking solution inside a basin.

## IX. DISCUSSION AND PERSPECTIVES

Solving the out-of-equilibrium relaxation dynamics in the spherical mixed $p$-spin model, starting from a thermalized configuration at temperature $T$, we have uncovered several interesting and unexpected features. While the known solution of the relaxation dynamics in the pure $p$-spin model suggested the existence of a unique dynamical phase transition at $T_{\text{MCT}}$, where ergodicity breaks down, and of a unique threshold energy where the dynamics relaxes below $T_{\text{MCT}}$, our results about the mixed $p$-spin model reveal a much richer scenario.
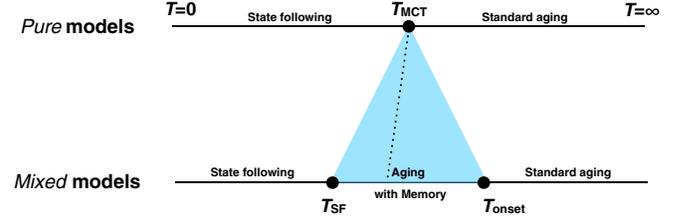


FIG. 9. Schematic phase diagram change when going from pure to mixed $p$-spin models. In general, the mode-coupling temperature $T_{\text{MCT}}$ (dashed line) is not a phase boundary.

Our main result is summarized in the schematic picture in Fig. 9: In the mixed $p$-spin model, there is an entire temperature range around $T_{\text{MCT}}$ (especially above it) where the final energy depends on $T$ and memory of the initial condition is kept. A temperature $T_{\text{onset}}$ marks the onset of this effect. This behavior resembles realistic glass formers but was unexpected in mean-field models because such a phenomenon was ascribed to activated processes, which are absent in the out-of-equilibrium relaxation of mean-field models. Our results show, contrary to common wisdom, that the dependence on the initial configuration can arise even in long-range models, following a purely relaxation dynamics.

From the numerical solution of the dynamical mean-field equations, this dynamical onset of memory may be a crossover or a new genuine phase transition. We have found an approximate solution of the relaxation dynamics at large times that describes the onset as a phase transition and provides an analytic prediction for the onset temperature. Thus, we finally have a solvable (although approximate) model whose predictions can be compared with numerical experiments.

Another piece of common wisdom that the present work demystifies is the mantra "relaxation dynamics in mean-field models relaxes to the threshold energy where the most numerous metastable states are found." Calling threshold energy the energy where the relaxation dynamics converge if starting from a random configuration ($T = \infty$), we have shown that (i) metastable states of an infinite lifetime (i.e., locally stable states) also exist above the threshold energy and (ii) the most numerous states are above the threshold energy. The threshold energy can thus be understood as the energy value where the measure is dominated by marginal states, which can be computed via the unconstrained complexity. Above the threshold energy, an exponential number of marginal states already exist, but the measure over stationary points is dominated by saddles. The same argument is trivialized in the pure $p$-spin model since, in the latter, marginal states exist only at the threshold energy; thus, assuming convergence to marginal states, we can uniquely determine the corresponding threshold energy.

We have shown evidence that the correct way to understand the large-time limit of the relaxation dynamics is to assume that such a dynamical out-of-equilibrium process

always tends to marginal states. For $T < T_{\mathrm{onset}}$, the persistent memory of the initial configuration suggests that we compute a constrained complexity of these marginal states that might be correlated with the initial condition. We have performed this task with different levels of accuracy and success.

For an initial temperature below $T_{\mathrm{SF}}$, which is strictly smaller than $T_{\mathrm{MCT}}$ in the mixed $p$-spin model, the starting configuration lies in a well-defined metastable state that can be followed down to zero temperature. In this situation, the dynamics never enters an aging regime, and its asymptotic behavior can be computed from a standard state-following computation.

The most interesting regime is the one reached with an initial temperature $T_{\mathrm{SF}} < T < T_{\mathrm{onset}}$. The relaxation process falls asymptotically in an aging regime, but it remains correlated with the initial configuration. In this regime, the constrained complexity, while confirming the existence of marginal states below threshold and correlated with the initial condition, fails to identify the ones chosen by the dynamics.

We notice that the mode-coupling transition temperature lies in this regime, $T_{\mathrm{SF}} < T_{\mathrm{MCT}} < T_{\mathrm{onset}}$, but plays no role at all in the out-of-equilibrium relaxation, which is controlled by $T_{\mathrm{onset}}$ and $T_{\mathrm{SF}}$. The central role $T_{\mathrm{MCT}}$ that has played until now is due to the accidental equality $T_{\mathrm{SF}} = T_{\mathrm{MCT}} = T_{\mathrm{onset}}$ holding in the pure $p$-spin model. All our results remain true for dynamics performed at a small temperature $T_f$.

We believe that, contrary to the pure $p$-spin model, the mixed $p$-spin model we studied is generic enough that the properties we uncovered should typically hold in models with a mean-field RFOT, such as Potts glasses [52], Ising $p$-spins [53–55], and, importantly, glasses of spherical particles in the infinite-dimensional limit [37].

In conclusion, on the basis of the above results, we have to revisit our common beliefs about the relaxation dynamics in $p$-spin models and, more generically, in models presenting a mean-field RFOT. The physical picture that has been a central paradigm of our scientific community, where a simple connection between the relaxation dynamics and the complexity of metastable states exists, is unfortunately valid only in the case of the pure $p$-spin model and is false in more general models.

In pure $p$-spin models, strong symmetries lead to the equalities $T_{\mathrm{SF}} = T_{\mathrm{onset}} = T_{\mathrm{MCT}}$, and these make many of the interesting phenomena disappear that we described in this work and that have misled us for two decades with a too-simplistic connection between the dynamics and energy landscape. Mixed models appear to be well suited to describe the dynamical onset of glassy phenomenology; they have much more complex energy landscapes, which deserve a more thorough investigation in the future. The exact solution to the asymptotic dynamics remains unknown, and the connection between asymptotic

dynamics and the energy landscape is still very open in spherical mixed $p$-spin models. Further research is needed in two directions: finding algorithms capable of reaching large times in solving the dynamical equations, and establishing new theoretical ideas on possible structures of the aging with memory to understand the asymptotic aging regime.

The richer scenario we have uncovered in this study is likely to also have an impact on the many applications in machine learning we discussed in the Introduction. Given the simplest connection between relaxation dynamics and the energy landscape, it is clear that more in-depth studies of the energy landscape will be needed in order to predict the performance of algorithms (a first example can be found in Ref. [56]), especially those that do not start from a random configuration but use some smart initialization to improve their performance [57].

## APPENDIX A: ASYMPTOTIC SOLUTIONS

In Ref. [27], it was shown that the asymptotic solution to the dynamical equations, assuming a simple aging with memory within a state with a unique effective temperature, has parameters $\chi$, $q_0$, $q_{12}$, and $y$ that satisfy the same equations that can be obtained by extremizing the Franz-Parisi (FP) potential computed in the one-step replica symmetry-breaking (1RSB) scenario, provided that the parameter $y$ is fixed by a marginality condition.

In our case, the 1RSB FP potential has to be computed at zero temperature ($T_f = 0$), while the reference configuration is in equilibrium at temperature $T = 1/\beta$,

$$
\begin{aligned}
&-2V_{1\mathrm{RSB}}(q_{12}, \chi, q_0, y) \\
&= \chi f'(1) + y(f(1) - f(q_0)) \\
&\quad + \frac{1}{y}\log\left(\frac{\chi + y(1 - q_0)}{\chi}\right) + \frac{q_0 - q_{12}^2}{\chi + y(1 - q_0)} + 2\beta f(q_{12}).
\end{aligned}
$$
(A1)

The saddle-point equations and the marginality condition thus read

$$\partial_\chi V_{1\mathrm{RSB}} = 0 \Rightarrow \chi(1 - q_{12}^2) + y(1 - q_0)^2$$
$$- \chi(\chi + y(1 - q_0))^2 f'(1) = 0$$
$$\partial_{q_0} V_{1\mathrm{RSB}} = 0 \Rightarrow q_0 - q_{12}^2 - (\chi + y(1 - q_0))^2 f'(q_0) = 0$$
$$\partial_{q_{12}} V_{1\mathrm{RSB}} = 0 \Rightarrow q_{12} - \beta(\chi + y(1 - q_0))f'(q_{12}) = 0$$
$$\text{marginality} \Rightarrow \chi^2 f''(1) = 1. \tag{A2}$$

We notice *en passant* that Eqs. (A2) do not always select the minima of the FP potential because, in order to do that, one needs to set to zero the total derivative with respect to $q_{12}$ and not the partial derivative. The energy and the radial reaction are given, respectively, by

$$E = -\chi f'(1) - y(f(1) - f(q_0)) - \beta f(q_{12}),$$
$$\mu = \chi f''(1) + (\chi + y)f'(1) - y q_0 f'(q_0) + \beta q_{12} f'(q_{12}). \tag{A3}$$

Equations (A2) always admit the solution with $q_{12} = q_0 = 0$, representing the memoryless or "amnesic" aging solution with parameters

$$\chi = \chi_{\mathrm{mg}} \equiv \frac{1}{\sqrt{f''(1)}},$$
$$y = y_0 \equiv \frac{f''(1) - f'(1)}{f'(1)\sqrt{f''(1)}} = \frac{\sqrt{f''(1)}}{f'(1)} - \chi_{\mathrm{mg}},$$

and, consequently, the energy and radial reaction are given by

$$E = E_{\mathrm{th}} \equiv -\chi_{\mathrm{mg}} f'(1) - y_0 f(1), \qquad \mu = \mu_{\mathrm{mg}} \equiv 2\sqrt{f''(1)}.$$

For the $(3 + 4)$-spin model, the numerical values for the above parameters are $\chi_{\mathrm{mg}} = 1/3$, $y_0 = 11/21 \simeq 0.52381$, $E_{\mathrm{th}} = -71/42 \simeq -1.69048$, and $\mu_{\mathrm{mg}} = 6$.

A nontrivial aging-with-memory solution with strictly positive values for $q_{12}$ and $q_0$, disconnected from the amnesic aging solution, exists only for temperatures below $T_0$. Moreover, at an even lower temperature $T_{\mathrm{SF}}$, we have that $q_0 \to 1$, and this solution becomes replica symmetric. In Fig. 10, we report the values of $q_0$, $q_{12}$, and $y$ in the aging solution with memory. In the $(3 + 4)$-spin model, the existence domain for this solution is upper bounded by $T_0 = 0.8031557$, where the solution disappears by a square-root singularity, and lower bounded by $T_{\mathrm{SF}} = 0.7982754$, where $q_0 \to 1$ and the solution becomes replica symmetric (RS). At $T = T_{\mathrm{SF}}$, the parameters extremizing the potential can be computed analytically. The value of $q_{12}$ is more easily obtained from the RS solution (see below), while the value of $y$ can be obtained from the third-order expansion in $\varepsilon = 1 - q_0$ of the difference between the two equations in (A2), thus obtaining
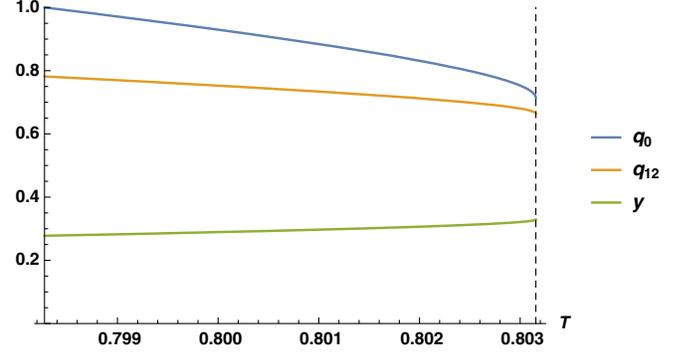


FIG. 10. Aging solution with memory of the initial configuration as predicted by the derivative of the FP potential for the $(3 + 4)$-spin model: $q_0$ (upper line), $q_{12}$ (central line), and $y$ (lower line) as a function of $T$. The temperature in the plot is in the range between $T_{\mathrm{SF}} = 0.7982754$ and $T_0 = 0.8031557$ (marked by a dashed vertical line). Notice that in this solution, we have $y \approx 0.3$, which is quite far from the value $y \approx 0.52$ that we observe in the numerical solution to the dynamical equations. Also notice that $q_{12}$ and $q_0$ have finite values when the solution appears at $T_0$.

$$y(T_{\mathrm{SF}}) = \frac{f'''(1)}{2[f''(1)]^{3/2}}. \tag{A4}$$

The numerical integration of the off-equilibrium dynamics does not give any indication in favor of this solution; in particular, for the $(3 + 4)$ model, this solution has $y \approx 0.3$ in the whole range of validity, which is incompatible with the value $y \approx 0.52$ that we obtain from the numerical solution of the dynamics. Even if we believe that the dynamics would slowly cross over to this solution on timescales that we cannot observe numerically, this would not solve the puzzle of the behavior of the asymptotic energy going below threshold for $T > T_0$. In this solution, the value of $q_0$ tends to 1 at temperature $T_{\mathrm{SF}}$, well above the temperature of the static (Kauzmann) phase transition of the model. The solution above is the only exact aging solution that we found; we searched for, without success, more complex solutions with more than one effective temperature.

When $q_0 \to 1$, the 1RSB potential reduces to the RS potential,

$$-2V_{\mathrm{RS}}(q_{12}, \chi) = \chi f'(1) + \frac{1 - q_{12}^2}{\chi} + 2\beta f(q_{12}), \tag{A5}$$

and the saddle-point equations read

$$1 - q_{12}^2 - \chi^2 f'(1) = 0$$
$$q_{12} - \chi\beta f'(q_{12}) = 0, \tag{A6}$$

with the energy taking the value $E = -\chi f'(1) - \beta f(q_{12})$. In the $3 + 4$ model, the energy of this solution at $T_{\mathrm{SF}}$ is equal to $E_{\mathrm{SF}} = -1.69981$ and decreases below. The susceptibility $\chi$
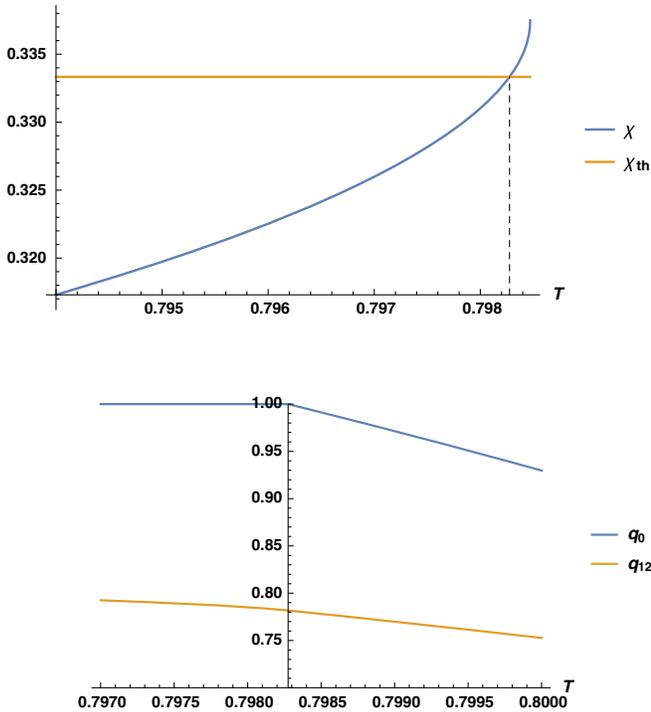
FIG. 11.    Upper panel: Susceptibility $\chi$ in the RS state following the solution for the $(3+4)$-spin model. Here, $T_{SF} = 0.7982756$ is defined by the condition $\chi = \chi_{mg}$, marked by a vertical dashed line. Lower panel: Since the vertical axis is located exactly at $T = T_{SF}$, values for $q_0$ and $q_{12}$ on the right are from the RSB aging solution, while those on the left are from the RS state-following solution.

takes the marginal value $\chi_{mg}$ at $T_{SF}$ and becomes smaller than this value at lower $T$, as shown in the upper panel of Fig. 11. Our estimate $T_{SF} = 0.7982756$ comes from exactly imposing $\chi = \chi_{mg}$ in this RS solution. In the lower panel of Fig. 11, we plot the values of $q_0$ and $q_{12}$ obtained in the RS state-following solution on the left of the vertical axis, which is located exactly at $T = T_{SF}$. On the right of the vertical axis, we show the corresponding values computed in the aging RSB solution (the same as plotted in Fig. 10). We observe that $q_{12}$ is a smooth function at $T_{SF}$. Its value at $T_{SF}$ can be computed analytically from the first equation in Eq. (A6) by imposing $\chi = \chi_{mg}$, leading to

$$q_{12,SF} = \sqrt{1 - \frac{f'(1)}{f''(1)}}. \qquad (A7)$$

For $T \leq T_{SF}$, the FP potential has a secondary minimum described by a replica symmetric ansatz. Thus, we are effectively describing the quenching process from temperature $T$ to zero temperature as a state-following process [39]: The observation that, both at the starting and ending temperatures, the state we are following is well described by a replica symmetric ansatz is evidence that static-dynamics

equivalence holds in this case. This is indeed what we observe by comparing the asymptotic dynamics obtained by numerically integrating the dynamical equations to the values derived here from the thermodynamic FP potential. For $T \leq T_{SF}$, aging disappears, and one finds a simple agingless relaxation within a state described by the parameters computed in the secondary minimum of the FP potential.

Given the difficulty in finding a 1RSB solution by solving the equations for the asymptotic dynamics, one may wonder whether a solution with more steps of RSB would be helpful. The thermodynamics of the model, even under the constraint of a fixed overlap with the initial configuration, is exactly solved by the 1RSB ansatz. Thus, a higher-order RSB solution is unlikely to be needed, and, as shown in Appendix B, the actual asymptotic dynamics looks very similar to the one of a 1RSB model. Nonetheless, we cannot exclude such a possibility: Let us just stress that, if different RSB ansatz are required for the thermodynamic and the dynamic solutions, the statics-dynamics connection that has been used so much in the field of spin glasses [58–60] would be strongly weakened.

## APPENDIX B: SEMIEMPIRICAL RELATIONS TO DESCRIBE LARGE-TIME DYNAMICS

We have seen that there is no simple aging solution to the asymptotic equations with $q_{12} > 0$ that is consistent with the finite-time numerical integration of the dynamical equations. Nonetheless, the FD plot in the mixed $(3+4)$-spin model does not differ in any appreciable way from the usual ones of the pure $p$-spin model with a single effective temperature and a value of $y$ equal to the one that holds for the $\beta = 0$ initial condition. Representing the data in the *hic sunt leones* region II, as much as possible, with a single-effective-temperature dynamical ansatz, we study what kind of relation can be derived between the overlaps $q_0$ and $q_{12}$. We suppose that $\mu$ and $\chi$ take the marginal values $\mu_{mg} \equiv 2\sqrt{f''(1)}$ and $\chi_{mg} \equiv 1/\sqrt{f''(1)}$. Moreover, we assume that the FD slope in the range $C(t, t') \in [q_0(t), 1]$ is independent of the initial temperature; thus, it is equal to the value obtained with $\beta = 0$, i.e., $y_0 \equiv \sqrt{f''(1)}/f'(1) - 1/\sqrt{f''(1)}$. Finally, we assume the response is null in the range $C(t, t') \in [C(t, 0), q_0(t)]$, where the correlation decays extremely fast.

Within this 1RSB ansatz, the asymptotic value for the radial reaction is given by

$$\mu = \chi f''(1) + (\chi + y)f'(1) - yf'(q_0)q_0 + \beta f'(q_{12})q_{12}$$

and by imposing $\mu = \mu_{mg}$, $\chi = \chi_{mg}$ and $y = y_0$, we get the relation

$$y_0 f'(q_0)q_0 = \beta f'(q_{12})q_{12}, \qquad (B1)$$
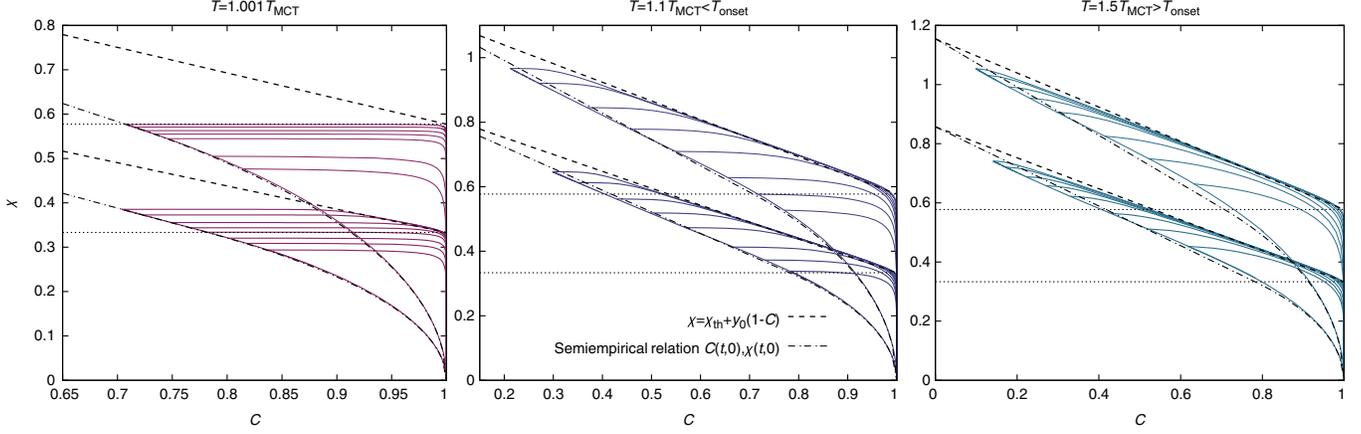
FIG. 12. FD curves $\chi(t,s)$ versus $C(t,s)$, each one with a different largest time $t$, plotted parametrically by varying the smallest time $s$. In each panel, upper data are for the pure three-spin model, while lower data are for the mixed $(3+4)$-spin model. Notice that the three panels have different scales, but the dashed and dash-dotted lines are the same in all panels (see text for details).

which must hold in the asymptotic solution. We notice at this point that using Eq. (A3), the energy can be written as $E = E_{\text{th}} + \Delta E$, with $\Delta E = yf(q_0) - \beta f(q_{12})$. Only in the pure model does Eq. (B1) imply $\Delta E = 0$; in all the other cases, it gives a nonvanishing $\Delta E$.

A second relation between $q_0$ and $q_{12}$ can be derived from the observation of the FD plots in Fig. 12, where each curve is a parametric plot of $\chi(t,t')$ versus $C(t,t')$ at fixed $t$, varying $t'$. We notice that, below the onset temperature, the points $(C(t,0), \chi(t,0))$ closely follow the dash-dotted curve that we derive analytically. It is important to remark that the dash-dotted curve is the same in all the panels of Fig. 12, so it is a temperature-independent relation. In addition, the dashed line is temperature independent and corresponds to the FD relation in the aging solution with $\beta = 0$, which is a line of slope $-y_0$ connecting the points $(0, \chi_{\text{mg}} + y_0)$ and $(1, \chi_{\text{mg}})$.

In order to obtain the dash-dotted line, we assume that in the aging regime [i.e., for $\chi(t,0) \geq \chi_{\text{mg}}$], the relation between $C(t,0)$ and $\chi(t,0)$ is linear, while in the regime $\chi(t,0) < \chi_{\text{mg}}$, the dynamics asymptotically explores a state; thus, we assume the relation holds within the RS state-following solution. The latter can be easily derived from the first equation in Eq. (A6),

$$\chi = \sqrt{\frac{1 - q_{12}^2}{f'(1)}}, \qquad (B2)$$

which thus holds for $q_{12} \in [q_{12,\text{SF}}, 1]$, while the linear part has slope $-y_0/q_{12,\text{SF}}$ and passes through the points $(0, \chi_{\text{mg}} + y_0)$ and $(q_{12,\text{SF}}, \chi_{\text{mg}})$. We notice that the dash-dotted curve has a continuous first derivative at the point $(q_{12,\text{SF}}, \chi_{\text{mg}})$, as can be easily checked by taking derivatives.

Our asymptotic ansatz thus implies a very simple relation between the overlaps describing the asymptotic aging regime, namely,

$$q_{12} = q_{12,\text{SF}} q_0, \qquad (B3)$$

and plugging this relation into Eq. (B1), it is easy to find that a solution with $q_{12} > 0$ can exist only if

$$T < T_{\text{onset}} \equiv \frac{q_{12,\text{SF}}^k}{y_0} = \frac{f'(1)[f''(1) - f'(1)]^{\frac{k}{2}-1}}{f''(1)^{\frac{k-1}{2}}}, \qquad (B4)$$

where $k$ is defined by $f(q) \propto q^k$ for $q \to 0$, i.e., the smallest power appearing in $f$ (in the $3 + 4$ model, $T_{\text{onset}} = 0.91$).

Even though this is not an exact solution of the asymptotic equations, it is a strong indication that there is a phase transition between a memoryless phase where dynamics decorrelates from the initial condition and falls over the "usual" threshold states with $E = E_{\text{th}}$ and a phase where aging takes place in a confined space, with an asymptotic energy below threshold and depending on $T$.

In Fig. 13, we report even stronger evidence that this approximate solution provides a very good description of
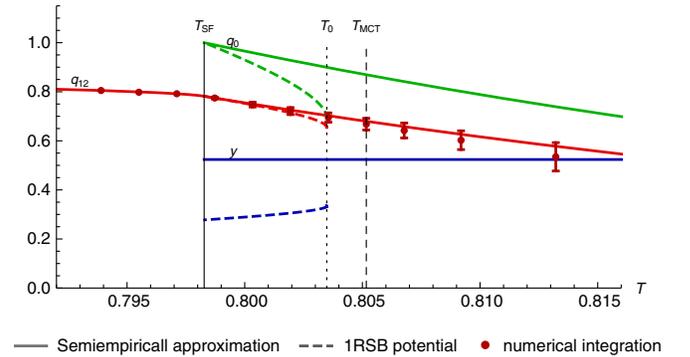


FIG. 13. Comparison between asymptotic aging parameters extrapolated from the numerical solution of the dynamical equations (data with errors), the analytical prediction obtained via the semiempirical approximation derived in Appendix B (solid curves), and the standard RSB aging solution derived in Appendix A (dashed curves).
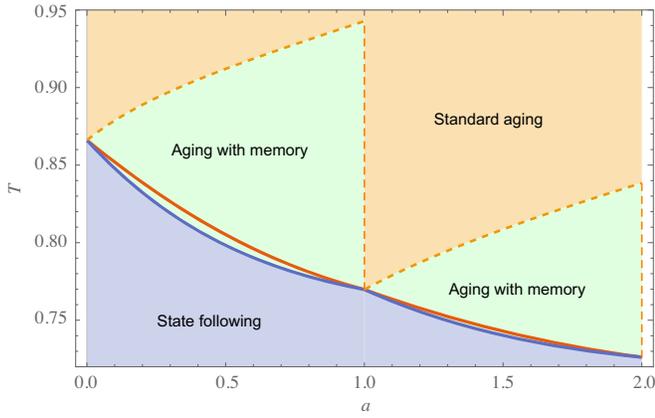
FIG. 14. Approximate phase diagram for the mixed model defined in the text, interpolating between pure $p$-spin models with $p = 3 + a$ when $a$ is an integer. The three different regimes are delimited by $T_{\text{onset}}$ (orange dashed line) and $T_{\text{SF}}$ (blue solid line). In addition, we plot the mode-coupling temperature $T_{\text{MCT}}$ (red solid line). Note that $T_{\text{onset}}$ is shown with a dashed line to stress the fact that it is the result of an approximation. Notice its discontinuous behavior around the pure models.

the asymptotic dynamics obtained by numerically integrating the dynamical equations. We plot numerically extrapolated values as data points with errors; solid lines are predictions from the approximate solution presented in this Appendix, and dashed lines correspond to the standard 1RSB aging solution discussed in the previous Appendix. There is no need to comment on which one is better in describing the asymptotic dynamics. Given that the 1RSB solution has three main defects: it is not available above $T_0$, it predicts a slightly wrong $q_{12}$ and a completely wrong $y$.

Since the onset temperature in Eq. (B4) depends only on the smallest-degree monomial in $f(q)$, one may wonder what the different effects are of perturbing a pure model with a smaller- or larger-degree polynomial. In order to answer this interesting question, let us study a mixed $p$-spin model that interpolates between pure $p$-spin models by just changing a single parameter $a$,

$$f(q) = \begin{cases} (1-a)q^3 + aq^4 & \text{if } 0 \le a \le 1 \\ (2-a)q^4 + (a-1)q^5 & \text{if } 1 < a \le 2. \end{cases} \quad \text{(B5)}$$

According to this definition, $f(q)$ is at most a binomial; it recovers a pure $p$-spin model for $a = 0, 1, 2$, while it reduces to our reference mixed model for $a = 1/2$.

Given the formula in Eq. (B4) for the onset temperature, we expect a different behavior when a pure $p$-spin model, e.g., the four-spin model ($a = 1$), is perturbed by a smaller-degree ($a = 1 - \varepsilon$) or larger-degree ($a = 1 + \varepsilon$) term. This behavior is clearly visible in Fig. 14, where we draw the phase diagram in the $(a, T)$ plane according to the approximation discussed in this Appendix.

The nonanalytic behavior of the onset temperature $T_{\text{onset}}$ around the pure models may appear strongly counterintuitive, as one would expect $T_{\text{onset}}$ to reduce to $T_{\text{MCT}}$ in the
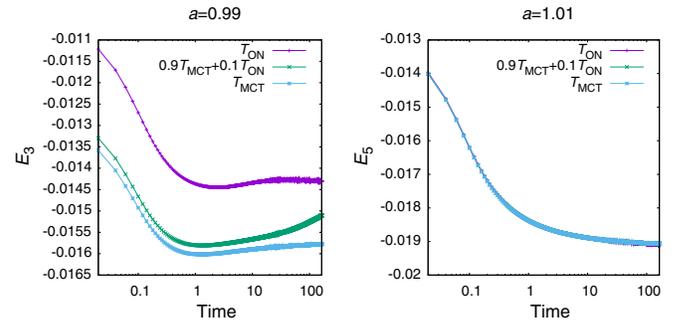


FIG. 15. Four-spin model slightly perturbed with three-spin terms ($a = 0.99$, left panel) and with five-spin terms ($a = 1.01$, right panel). The time evolution of the perturbing term energy makes evident the different dynamical behavior: For $a = 0.99$, we see the effects of the onset of glassy dynamics, which are not visible for $a = 1.01$, thus supporting the singular $T_{\text{onset}}$ of Fig. 14.

pure case. Such a limit is obtained only when the perturbing term is of higher order, e.g., in the limit $a \to 1^+$ in the present case, $T_{\text{onset}} \to T_{\text{MCT}}$ as shown in Fig. 14. On the contrary, when the perturbing term is of lower order, e.g., for $a \to 1^-$ in the present case, we have a discontinuity in $T_{\text{onset}}$ (see Fig. 14). Such a discontinuity could be a genuine effect or an artefact of our approximation.

To prove that the singular behavior in $T_{\text{onset}}$ is not incompatible with the physics of the system, let us show that perturbing a pure $p$-spin model with a higher- or lower-order term produces different effects on the relaxation dynamics. In Fig. 15, we plot the relaxation of the perturbing term in the energy as a function of time for $a = 0.99$ (left panel) and $a = 1.01$ (right panel). For $a = 1.01$, the relaxation of the perturbing term is extremely similar to the one of the pure $p$-spin model; conversely, for
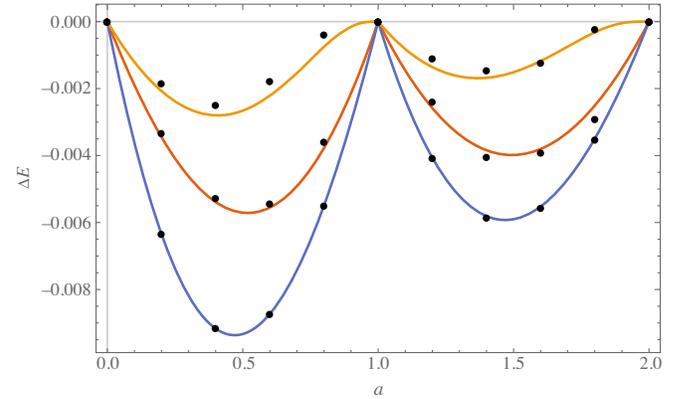


FIG. 16. The energy excess $\Delta E = y f(q_0) - \beta f(q_{12})$ as a function of $a$, for the approximate solution of the same models as in Fig. 14. The three curves correspond to the $a$-dependent temperatures $T_{\text{SF}}$ (blue), $T_{\text{MCT}}$ (red), and $0.9 T_{\text{MCT}} + 0.1 T_{\text{onset}}$ (orange). We see that $\Delta E \to 0$ in the limit of pure models with no discontinuity, while $\Delta E < 0$ in general mixed models. The points are extrapolated values from the numerical integration of the dynamical equations; the comparison shows good agreement at all temperatures.
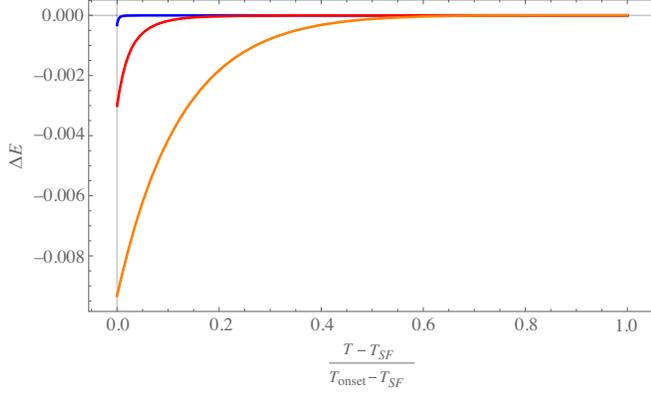
FIG. 17. The energy excess $\Delta E = y f(q_0) - \beta f(q_{12})$ as a function of the reduced temperature $(T - T_{SF})/(T_{onset} - T_{SF})$ for $a = 0.5$ (orange), 0.9 (red), and 0.99 (blue), from bottom to top. According to the phase diagram shown in Fig. 14, the normalizing factor $(T_{onset} - T_{SF})$ remains finite in the $a \to 1^-$ limit, but $\Delta E \to 0$ in that limit. For $a \lesssim 1$, $\Delta E$ is different from zero only very close to $T_{SF}$.

$a = 0.99$, we observe an undershoot and a clear dependence of the asymptotic energy on the temperature. This dynamical effect, clearly visible even if the perturbing term is tiny, could be due to the existence of an onset temperature $T_{onset}$ larger than $T_{MCT}$ in the limit $a \to 1^-$ as suggested by our approximation.

The simplest explanation for the undershoot in the energy relaxation observed for $a = 0.99$ in Fig. 15 can be provided by assuming that lower-order interactions in the energy are satisfied (i.e., minimized) sooner during the relaxation, and once satisfied, they become a sort of pinning field for the relaxation of the remaining energy terms. For this reason, even a tiny fraction of three-spin terms would have a visible effect in the relaxation of a four-spin model, with the energy $E_3$ becoming too small too soon and eventually increasing again when the four-spin part is satisfied.

Nonetheless, we show in Fig. 16 that the nonanalytic behavior of $T_{onset}$ does not correspond to a discontinuity in

the approximated value of $\Delta E$, which tends to zero in the limit of a pure model from both sides (as long as $T > T_{MCT}$). In Fig. 17, we show, in particular, the behavior approaching $a = 1$ from below. We notice that, while $T_{onset} - T_{SF}$ remains finite (as shown in Fig. 14), the energy excess $\Delta E$ vanishes in the whole interval $[T_{onset}, T_{SF}]$ when $a \to 1^-$. This result means that, in such a limit, $T_{onset}$ cannot be easily detected from the energy relaxation.

## APPENDIX C: COUNTING THE MINIMA

In this Appendix, we try to answer the question of whether or not the attractors of the dynamics can be well described in terms of typical marginal saddles and minima that lie close to the initial configuration. Let us consider the stationary points of the Hamiltonian $H[\sigma]$ on the sphere $\sum_i \sigma_i^2 = N$:

$$H_i' + \mu \sigma_i = 0. \tag{C1}$$

As in dynamics, the radial reaction $\mu$ takes, in any stationary point, the value

$$\mu = -\frac{1}{N} \sum_i \sigma_i H_i'. \tag{C2}$$

We classify the stationary points according to their energy $E = (1/N)H[\sigma]$ and the value of the radial reaction $\mu$.

Differently from pure models where $\mu = pE$, the relation between $E$ and $\mu$ here is not univocal, and stationary points are found in a whole region of the $(E, \mu)$ plane.

We have seen that for $T > T_{SF}$, dynamics is attracted by some family of marginal minima. In order to characterize these minima, we count the number of stationary points of the Hamiltonian $H[\sigma]$ with fixed energy $E$ and radial reaction $\mu$ that lie at a fixed overlap $\sigma \cdot \sigma_0 = N q_{12}$ from a reference configuration $\sigma_0$, weighted with a Gibbs measure $e^{-\beta H[\sigma_0]}$. Since the complexity, i.e., the logarithm of their number, is self-averaging, we write

$$\Sigma(E, \mu, q_{12}, \beta) = \int_{S_N} \mathcal{D}\sigma_0 \frac{e^{-\beta H[\sigma_0]}}{Z_\beta} \log\left( \int_S \mathcal{D}\sigma \delta(N q_{12} - \sigma \cdot \sigma_0) \delta(NE - H) \delta(\mu \sigma + H') |\det(\mu I + H'')| \right). \tag{C3}$$

The computation of $\Sigma$ is standard and can be performed in several steps. First of all, since the matrix $H''$ is a GOE random matrix, the distribution of eigenvalues of $\mu I + H''$ is self-averaging and is a shifted semicircular $\rho(\lambda) = \{2/[\pi 4f''(1)]\}\sqrt{4f''(1) - (\lambda - \mu)^2}$. The modulus of the determinant $|\det(\mu I + H'')|$ is self-averaging, and its logarithm reads

$$D(\mu) = \frac{1}{N} \log|\det(\mu I + H)|$$

$$= \mathbf{Re}\left[ \frac{1}{4f''(1)} \left( \mu^2 - \mu\sqrt{\mu^2 - 4f''(1)} - 2f''(1)\left( -2\log\left( \sqrt{\mu^2 - 4f''(1)} + \mu \right) + 1 + \log(4) \right) + \mu^2 \right) \right], \tag{C4}$$

which only depends on $\mu$. The imaginary part of this expression is the proportion of negative eigenvalues.

To evaluate the remaining terms, we use replicas and write $\Sigma \equiv \overline{\log(\mathcal{N})} = \lim_{n\to 0}[(\overline{\mathcal{N}^n} - 1)/n]$. We concentrate on the case with temperatures greater than the static transition temperature $(T > T_K)$ of the model, where the partition function appearing in the denominator of Eq. (C3) is self-averaging and takes its annealed value $\overline{Z_\beta} = e^{(N/2)\beta^2 f(1)}$. One can then average over the disorder and the configuration $\sigma_0$ at the same time. Opening the delta function in the Fourier basis,

$$\Sigma(E, \mu, p, \beta) = \lim_{n\to 0}\frac{1}{n}\left(e^{-\frac{1}{2}\beta^2 f(1)}\int \mathcal{D}s\, e^{\sum_a^n N(i\hat{\beta}_a E - i\hat{\sigma}_a \cdot \sigma_a \mu)}\delta(Nq_{12} - \sigma_a \cdot \sigma_0)\overline{e^{-\beta H_0}e^{\sum_a^n(i\hat{\beta}_a + i\hat{\sigma}_a \cdot \nabla)H_a}}\right) + ND(\mu), \qquad \text{(C5)}$$

where $\int \mathcal{D}s = \int_S \mathcal{D}\sigma_0 \prod_a (\int_S \mathcal{D}\sigma_a \int \mathcal{D}\hat{\sigma}_a \int \hat{\beta}_a)$. Since the disorder is Gaussian,

$$\overline{e^{-\beta H_0}e^{\sum_a^n(i\hat{\beta}_a + i\hat{\sigma}_a \cdot \nabla)H_a}} = e^{\frac{1}{2}(\beta^2 f(\sigma_0 \cdot \sigma_0/N) + 2\beta \sum_a(i\hat{\beta}_a + i\hat{\sigma}_a \cdot \nabla^a)f(\sigma_a \cdot \sigma_0/N) + \sum_{ab}(i\hat{\beta}_a + i\hat{\sigma}_a \cdot \nabla^a)(i\hat{\beta}_b + i\hat{\sigma}_b \cdot \tilde{\nabla}^b)f(\sigma_a \cdot \tilde{\sigma}_b/N)|_{\tilde{\sigma}\to\sigma})}.$$

Now, we define overlap variables $NQ_{ab} = \sigma_a \cdot \sigma_b$, $N\chi_{ab} = i\sigma_a \cdot \hat{\sigma}_b$, and $NR_{ab} = -\hat{\sigma}_a \cdot \hat{\sigma}_b$, and the overlaps with the reference configuration $Nq_{12} = \sigma_a \cdot \sigma_0$, $N\chi_p = i\hat{\sigma}_a \cdot \sigma_0$. This change of variables defines a matrix

$$\mathcal{Q} \equiv \begin{pmatrix} 1 & q_{12} & -i\chi_p \\ q_{12} & Q_{ab} & -i\chi_{ab} \\ -i\chi_p & -i\chi_{ab} & -R_{ab} \end{pmatrix},$$

where $Q_{aa} = 1$ due to the spherical constraint. In addition, from the equivalence between replicas, we fix $i\hat{\beta}_a = y$ and $\chi_{aa} = \chi \; \forall \; a$. With this change of variables, Eq. (C5) becomes

$$\Sigma = (yE - \mu\chi) + \beta(yf(q_{12}) + \chi_p f'(q_{12})) + \lim_{n\to 0}\frac{1}{n}\left(\frac{1}{2}\log(\det \mathcal{Q})\right) + D(\mu)$$

$$+ \lim_{n\to 0}\frac{1}{n}\left(\frac{1}{2}\sum_{ab}[y^2 f(Q_{ab}) + 2yf'(Q_{ab})\chi_{ab} + f'(Q_{ab})R_{ab} + f''(Q_{ab})(\chi_{ab})^2]\right), \qquad \text{(C6)}$$

where $\frac{1}{2}\log(\det \mathcal{Q})$ is the volume factor that comes from the change of variables from spins to overlaps. To get the leading $N$ contribution, we must extremize with respect to all the overlap parameters $\mathcal{Q}$ and $y$. We notice that a further simplification to the expression (C6) can be obtained by first extremizing with respect to $R_{ab}$. Assuming a replica symmetric ansatz for the overlap matrices $Q$ and $\chi$, i.e., $Q = \delta_{ab} + (1 - \delta_{ab})q_0$ and $\chi_{ab} = \delta_{ab}\chi + (1 - \delta_{ab})\chi_1$, we get, in the limit $n \to 0$,

$$\Sigma(E, \mu, q_{12}, \beta; y, \chi, \chi_1, \chi_p, q_0) = +yE - \mu\chi + \beta[yf(q_{12}) + \chi_p f'(q_{12})]$$

$$+ \frac{1}{2}[y^2(f(1) - f(q_0)) + 2y(f'(1)\chi + f'(q_0)\chi_1) + \mathcal{R} + (\chi^2 f''(1) - \chi_1^2 f''(q_0))]$$

$$+ \frac{1}{2}\left[\log(1 - q_0) + \frac{q_0 - q_{12}^2}{1 - q_0} - \log(f'(1) - f'(q_0)) - \frac{f'(q_0)}{f'(1) - f'(q_0)}\right] + D(\mu), \qquad \text{(C7)}$$

where

$$\mathcal{R} \equiv 1 + f'(1)\left(\frac{\chi - \chi_1}{1 - q_0}\left((\chi - \chi_1) - \frac{q_0 - q_{12}^2}{1 - q_0} + 2(\chi_1 - p\chi_p)\right) + \chi_p^2\right) - f'(q_0)\left(\frac{\chi - \chi_1}{1 - q_0}\left(-\frac{q_0 - q_{12}^2}{1 - q_0} + 2(\chi_1 - p\chi_p)\right) + \chi_p^2\right). \qquad \text{(C8)}$$

This result can be extremized explicitly with respect to $y, \chi, \chi_1, \chi_p$, while $q_0$ extremization has to be done numerically. For $q_{12} = 0$, the solution is $q_0 = 0$, and we recover the expression found in Ref. [42],

$$\Sigma(E, \mu) = \max \left\{ 0, \text{Re} \left( -\frac{(E^2(f''(1) + f'(1)) + 2E\mu f'(1) + f(1)\mu^2)}{2f(1)(f''(1) + f'(1)) - f'(1)^2} \right. \right.$$
$$\left. \left. + \frac{\mu}{\sqrt{\mu^2 - 4f''(1)} + \mu} + \log\left(\sqrt{\mu^2 - 4f''(1)} + \mu\right) + \frac{1}{2}\log\left(\frac{1}{f'(1)}\right) - \log(2) \right) \right\}. \tag{C9}$$

This unconstrained complexity is plotted in Fig. 2. For any value of $\mu$, $\Sigma(E, \mu)$ has a concave parabolic shape as a function of the energy, and many values of the energy are possible for the same $\mu$ value. In particular, this case is true for $\mu = \mu_{\text{mg}}$: There is not a unique threshold energy (at variance to pure $p$-spin models) but a whole interval for which $\Sigma(E, \mu_{\text{mg}}) > 0$. We define the dominating stationary points at a given value of the energy as the ones that maximize $\Sigma$ as a function of $\mu$. The threshold energy as defined by dynamics corresponds to the values of the energy that separates minima from saddles on the dominating line. Notice that this value does not correspond to the most numerous marginal minima, which occur for $\mu = \mu_{\text{mg}}$ and $E > E_{\text{th}}$, but to the point where the most numerous stationary saddles become minima. This observation sheds some light on the dynamics from a random initial condition; minima are not accessible at levels of the energy $E > E_{\text{th}}$, where saddles dominate the landscape, even if many minima are present.

Extending the concept of threshold energy to the case $q_{12} > 0$ is straightforward. As shown in Fig. 18, for any
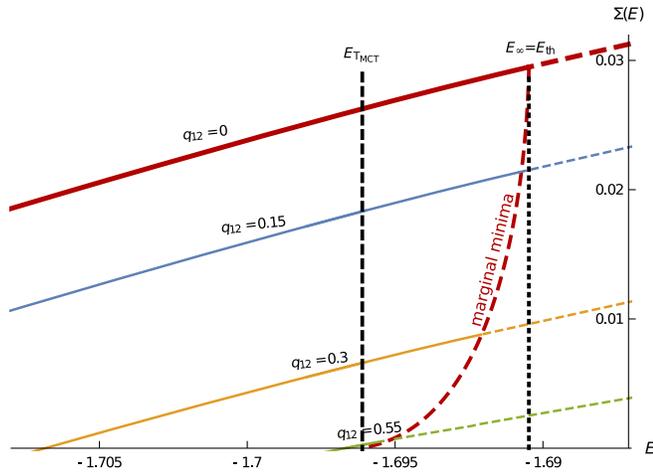


FIG. 18. Constrained complexity at an overlap $q_{12}$ from a reference configuration sampled at temperature $T = T_{\text{MCT}}$. Vertical lines mark energy values $E_\infty$ (dotted line on the right) and $E_{T_{\text{MCT}}}$ (dashed line on the left), corresponding to extrapolated asymptotic energies reached by the dynamics starting, respectively, from $T = \infty$ and $T = T_{\text{MCT}}$. From a random configuration with $T = \infty$, the dynamics goes to the energy level $E_{\text{th}}$, where the dominant stationary points are marginal minima, while starting near $T_{\text{MCT}}$, the dynamics goes below $E_{\text{th}}$. If $q_{12} > 0$, the energy where marginal minima dominate decreases, and we represent it with a bold dashed red curve.

fixed value of $q_{12}$, we can plot the curve corresponding to dominating minima. This curve ends at an energy value that we call $E_{\text{th}}(q_{12})$, which is the best candidate for the asymptotic energy in the relaxation dynamics. Indeed, for $E > E_{\text{th}}(q_{12})$, dominating stationary points are saddles, and with high probability, the dynamics will be able to relax further until the energy level $E_{\text{th}}(q_{12})$ is reached. However, even assuming the dynamics converges to $E_{\text{th}}(q_{12})$, we are still left with an unknown: the value of $q_{12}$, i.e., the asymptotic value for $C(t, 0)$. This problem is the same one the authors of Ref. [50] faced after computing the constrained complexity for the free energy at a nonzero temperature.

In Fig. 19, we plot the quenched constrained complexity of dominant states in the $(3 + 4)$-spin model, computed with $T = T_{\text{MCT}}$ (left panel) and $T = 1.02T_{\text{MCT}}$ (right panel). Please notice that the choice of the temperature is not crucial: As long as $T \in (T_{\text{SF}}, T_{\text{onset}})$, the plot would be very similar. The almost-horizontal red curve in the plot represents the threshold energy $E_{\text{th}}(q_{12})$ defined above. We notice that the range of energies with a positive complexity is very large compared to variations in $E_{\text{th}}$. The green ellipsis represents our best estimate for the large-time limit of the actual dynamics solved numerically: While the energy can be very well estimated, the limit of $C(t, 0)$ is plagued by a large uncertainty due to its slow convergence (until now, we have not attempted such an extrapolation; however, here we want to be more speculative, and we take the risk). From the plot, one is tempted to conjecture that the dynamics always converges to marginal states with a threshold energy $E_{\text{th}}(q_{12})$. However, we have not found any principle to fix the value of $q_{12}$ solely from the complexity curve, and further studies are needed to better match the large-time limit of the dynamics to the energy landscape.

Although Fig. 19 may suggest a relation between dynamics and the generalized threshold energies, a more careful analysis reveals its limitations. Assuming that at large times the relaxation dynamics converges to the manifold of marginal states belonging to the curve $E_{\text{th}}(q_{12})$, one could estimate the point reached by the dynamics by extrapolating the asymptotic energy $E_\infty(T)$ and estimating $q_{12}$ from the equality $E_{\text{th}}(q_{12}) = E_\infty$. Thus, having fixed the values of $E$ and $q_{12}$, one can proceed by estimating the remaining parameters of the asymptotic aging dynamics, $q_0$ and $y = \partial_E \Sigma$, from the saddle-point equations used to compute the quenched complexity. The result of this computation is shown in Fig. 20 with solid lines and compared to the (very uncertain) extrapolation of
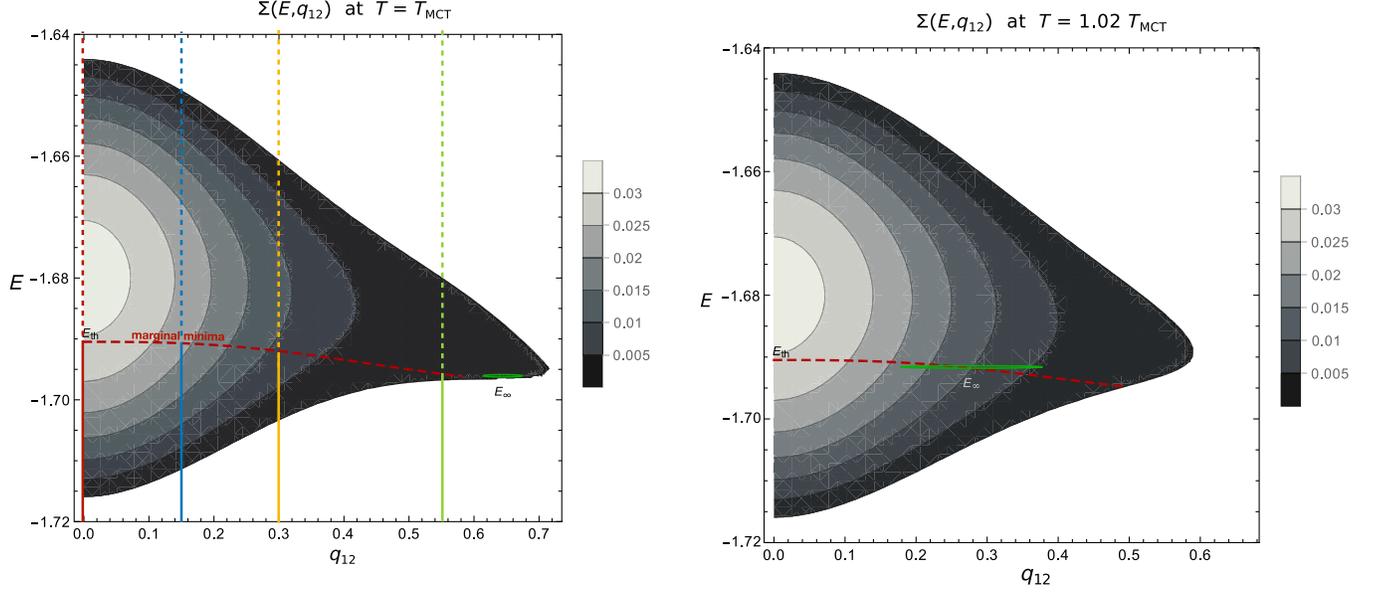
FIG. 19.   Quenched complexity of dominant states with energy $E$ and overlap $q_{12}$ with respect to an equilibrium configuration at temperature $T = T_{MCT}$ (left panel) and $T = 1.02 T_{MCT}$ (right panel). In the left panel, the four colored vertical lines highlight the complexities shown in Fig. 18. The choice of the temperature is not crucial as long as $T \in (T_{SF}, T_{onset})$. The almost horizontal red curves mark the threshold energy $E_{th}(q_{12})$, and the green ellipsis is our best estimate for the large-time limit of the relaxation dynamics obtained from the numerical integration.

$C(\infty, 0)$, shown by red points with errors. We clearly see that, while the estimate of $q_{12}$ is compatible with the actual dynamics, the other two parameters are far from the values measured in the numerical solution to the dynamics. Indeed, $q_0$ becomes smaller than $q_{12}$; however, in the actual dynamics, the inequality $q_{12} < q_0$ is always satisfied, and $y$ becomes much smaller than $y_0$ (marked by a dotted horizontal line in Fig. 20), which is a good descriptor of the actual aging in the whole temperature range studied. Thus, the present computation of the quenched complexity
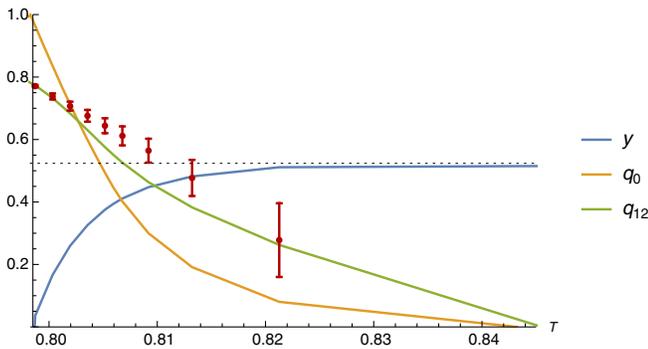


FIG. 20.   Assuming the dynamics relaxes on the marginal manifold with energy $E_{th}(q_{12})$ and fixing the energy from the large-time extrapolation of the numerical data, which in turn fixes the value for $q_{12}$, we can compute analytical values for the remaining aging parameters, namely, $q_0$ and $y = \partial_E \Sigma$. While the estimate of $q_{12}$ is compatible with the large-time extrapolation of $C(t, 0)$ (red points with error in the figure), the other aging parameters are far from those measured in the actual dynamics.

of marginal minima of the energy function in mixed $p$-spin models does not allow us to identify the attractors of the dynamics. More work will be required to connect the large-time aging dynamics to the properties of the energy landscape.

## APPENDIX D: EQUILIBRIUM DYNAMICS

The onset temperature $T_{onset}$ marks the point where the landscape influences gradient descent dynamics. In realistic model glasses, it has been found that the same temperature marks the onset of nonexponential relaxation, thus making a direct relation between equilibrium dynamics and the energy landscape [25]. Therefore, it is natural to search for a signature of $T_{onset}$ in equilibrium dynamics.

The aim of this Appendix is to briefly review the equilibrium dynamics at finite temperature and show that, while the usual MCT-like transition takes place at $T_{MCT}$, no anomalies around $T_{onset}$ are found. The equilibrium dynamical equation for the correlation function reads

$$\frac{dC(t)}{dt} = -TC(t) + \beta \int_0^t ds \, f'(C(t-s)) \frac{dC(s)}{ds}. \quad (D1)$$

This equation is readily obtained by imposing $T_f = T$ in Eq. (8), together with time-translation invariance in the correlation function $C(t, s) = C(t - s)$ and the fluctuation-dissipation relation $R(t) = -\beta\{[dC(t)]/dt\}$. At high temperatures, in the ergodic phase, $C(t)$ vanishes at large times. The dynamical MCT-like ergodicity-breaking
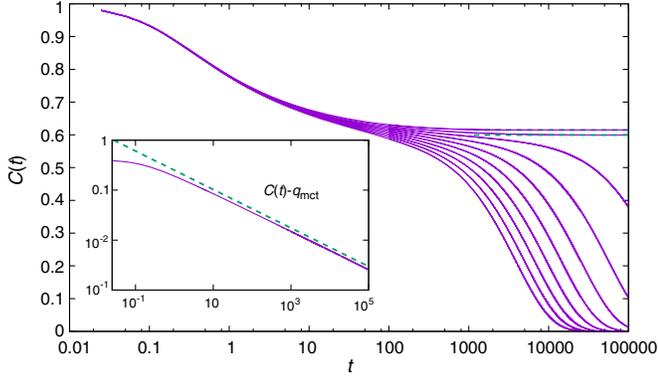
FIG. 21. Main panel: Equilibrium correlation function as a function of time on a log-linear scale at different temperatures from $T = 0.804$ to $T = 0.813$ in steps $\Delta T = 0.001$ (from top to bottom). All but the first two curves are for $T > T_{\mathrm{MCT}} = 0.805166$. We clearly see the formation of a plateau as $T_{\mathrm{MCT}}$ is approached from above. The two curves below $T_{\mathrm{MCT}}$ correspond, respectively, to $q_1 = 0.599574$ and $q_1 = 0.614868$, which we show with dashed lines. Inset: Correlation $C(t) - q_1$ at the critical temperature $T_{\mathrm{MCT}}$ on a log-log scale. For comparison, the dashed line represents the expected behavior, a power law with exponent $a = -0.38504$.

transition is signaled by the appearance of a nonzero limit $q_1 = \lim_{t\to\infty} C(t)$, with $q_1$ discontinuous at the transition. Evaluating Eq. (D1) at large times, one finds that $q_1$ is a solution of

$$\beta^2 f'(q_1)(1 - q_1) = q_1. \tag{D2}$$

This equation coincides with the equation defining the nontrivial minimum of the equal-temperature effective potential discussed in detail in Appendix A.

The transition temperature, the highest temperature for which Eq. (D2) has a nonzero solution, satisfies $\beta^2 f''(q_1)(1 - q_1)^2 = 1$. Standard mode-coupling analysis of the dynamics [22] describes the approach to the transition in terms of the formation of a plateau in the correlation function as a function of log time. This analysis, in particular, shows that the approach to $q_1$ at $T_{\mathrm{MCT}}$ follows a power law with an exponent $a$, which is determined by $\Gamma(1 - a)^2/\Gamma(1 - 2a) = (T/2)f'''[q]/f''[q]^{3/2}$. For the mixed model $f(q) = \frac{1}{2}(q^3 + q^4)$, the transition temperature is $T_{\mathrm{MCT}} = 21/(2\sqrt{37\sqrt{37} - 55}) \approx 0.805166$, corresponding to $q_1 = (1 + \sqrt{37})/12 \approx 0.59023$ and $a = -0.38504$. In Fig. 21, we display data from the numerical integration of Eq. (D1) with a simple first-order Euler algorithm similar to the one we use in the off-equilibrium case. The curves show that theoretical expectations are met: A transition occurs at $T_{\mathrm{MCT}}$ according to the expected pattern, and below $T_{\mathrm{MCT}}$, the system is in a nonergodic state.

We then investigate the temperature range around $T_{\mathrm{onset}}$ to see if this temperature corresponds to a qualitative
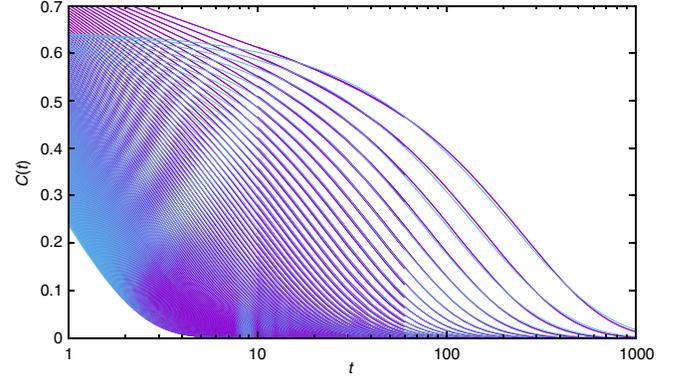


FIG. 22. Stretched exponential fit to the equilibrium correlation function $C(t)$ for temperatures in the range $T \in [0.84, 2]$ in steps $\Delta T = 0.01$. The fit has acceptable quality if the temperature is not too close to $T_{\mathrm{MCT}}$.

change in the equilibrium dynamics. Following Ref. [25], we fit the correlation function at intermediate times with a stretched exponential $f(t) = a \exp(-(t/\tau)^{\hat\beta})$ and study the behavior of the stretching exponent $\hat\beta$ as a function of the temperature [61]. In Ref. [25], for Lennard-Jones liquids, $\hat\beta$ was shown to become sensibly different from 1 around $T_{\mathrm{onset}}$. While the long-time behavior of the correlation function is a pure exponential, we can fit the decay at intermediate times. In Fig. 22, we see that the stretched exponential form provides a decent fit if we do not get too close to $T_{\mathrm{MCT}}$.

We plot the stretching exponent $\hat\beta$ for the mixed model as a function of the temperature in Fig. 23, together with the same exponent in the pure model. In both cases, we observe that while $\hat\beta \approx 1$ at high enough temperatures, it starts to decrease from that value as temperature is lowered.
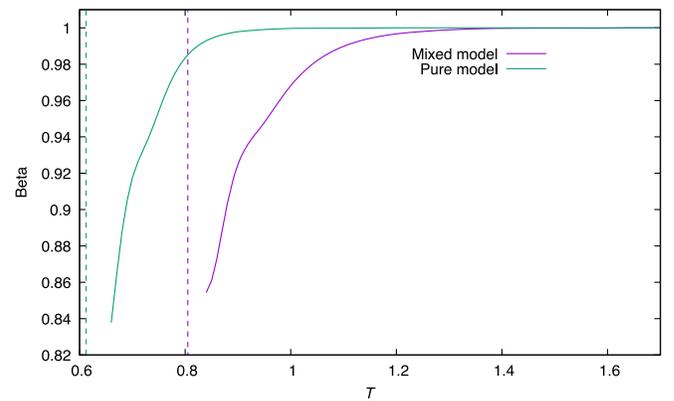


FIG. 23. The stretching exponent $\hat\beta$ as a function of temperature in the mixed $(3 + 4)$-spin model and in the pure three-spin model. The vertical lines correspond to the transition temperatures $T_{\mathrm{MCT}}$ of the two models. We see no qualitative differences between the two cases. Moreover, in the case of the mixed model, deviations from pure exponential behavior are seen much above the onset temperature $T_{\mathrm{onset}} \simeq 0.91$.

We could not find any qualitative difference here in the behavior of the pure and the mixed model, and we observe that in the mixed model, the stretching starts at $T \simeq 1.2$, well above the estimated value of $T_{\text{onset}} \simeq 0.91$. In conclusion, we do not see any signature of $T_{\text{onset}}$ in the equilibrium dynamics of the mixed model.

## APPENDIX E: RESPONSE AT SHORT TIMES

A natural hypothesis about the memory effect in region II is that it could be associated with an increased response $R(t, s)$ to perturbations applied at short times $s$. Here, we show that short-time dynamics provides hints against this hypothesis. Let us then compare the response $R(t, s)$ starting from infinite temperature and the one for an initial temperature in the region II. In Fig. 24, we plot the response as a function of the time difference $t - s$ starting from infinite temperature and from a temperature $T = 1.01T_{\text{MCT}}$. We clearly see that starting from finite temperature, the response at short times $s$ is much smaller than the one starting from infinite temperature. It is also interesting to study the behavior in time of $R(t, 0)$, which we plot in Fig. 25. It is well known that starting from a uniformly distributed initial condition, one finds that the response to the initial condition coincides with the remanent magnetization, namely, $R(t, 0) = C(t, 0)$ [46]. For finite $T$, this relation does not hold; in fact, it is possible to show that $R(t, 0) = C(t, 0) - \beta \int_0^t ds\, f'(C(0, s))R(t, s)$. We see that a finite limiting value for $C(t, 0)$ does not necessarily imply a finite response to the initial condition. Our data show that the two terms largely compensate each other; the larger $\beta$, the larger the compensation effect. Both for the mixed and the pure model, the lower the starting temperature, the lower $R(t, 0)$. The decay of $R(t, 0)$ becomes exponential only for $T = T_{\text{SF}}$. In the pure model, this case coincides with $T_{\text{MCT}}$, while in the mixed model, $T_{\text{SF}} < T_{\text{MCT}}$.
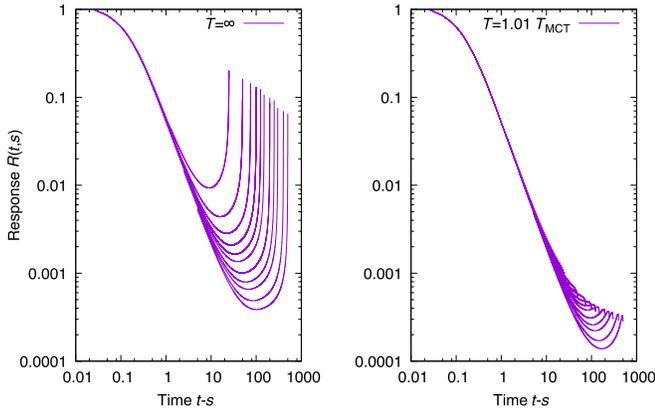


FIG. 24. Response function $R(t, s)$ with $T = \infty$ (left panel) and $T = 1.01T_{\text{MCT}}$ (right panel), as a function of the time difference $t - s$ for $t = 25, 50, 75, 100, 125, 150, 200, 250, 300, 400,$ and $500$. It is evident that the response is much smaller starting from finite temperature.
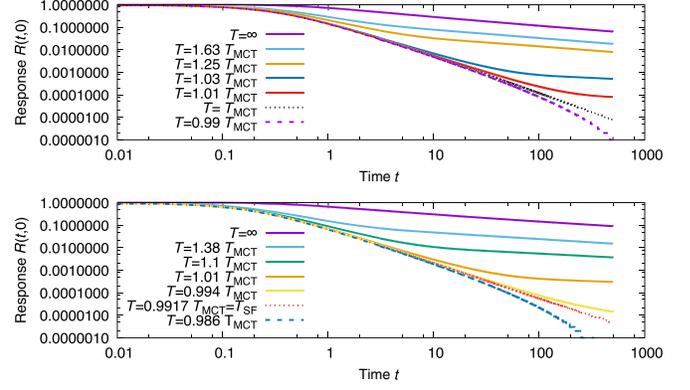


FIG. 25. The response to the initial condition $R(t, 0)$ for various initial temperatures in the pure three-spin model (upper panel) and in the mixed $3 + 4$-spin model (lower panel). The qualitative behavior is very similar; the higher the temperature, the higher the values of $R(t, 0)$. For high temperature, we see a slow decay of $R(t, 0)$, possibly a power law. At low temperature, below $T_{\text{MCT}}$ in the pure model and below $T_{\text{SF}}$ in the mixed model, the decay is exponential.

## APPENDIX F: NUMERICAL EXTRAPOLATIONS

The dynamical equations for the correlation and the response functions, written in Eq. (4), were integrated via a simple Euler algorithm with a fixed integration step $\Delta t$. In this way, we get extremely precise results at finite time, but the times we can reach are limited. We tried more sophisticated integration schemes [47], where the integration steps are increased during the evolution; unfortunately, as soon as a mixture is used, we have met numerical instabilities at very short times. In practice, only the pure $p$-spin model seems to allow those integration schemes to work.

The results presented in the main text have been obtained by integrating with an optimal integration step $\Delta t = 0.1$, which allows us to reach the largest times without facing any numerical instability (consider that at short times the differential equations we are solving have a natural timescale of order 1, so $\Delta t$ cannot be much larger than the value we used).

The integration error in the Euler algorithm is linear in the integration step $\Delta t$, and we are interested in understanding how much physical quantities computed with $\Delta t = 0.1$ differ from the corresponding $\Delta t \to 0$ limit. To this purpose, we study the relative errors in one-time quantities defined as follows:

$$\Delta_r x(t) = \frac{x(t; \Delta t) - x(t; \Delta t = 0.1)}{x(t; \Delta t)}. \tag{F1}$$

By definition, this relative error is zero for $\Delta t = 0.1$, it should be linear in $\Delta t$ for small enough $\Delta t$, and, in the limit $\Delta t \to 0$, it provides the relative error from using an
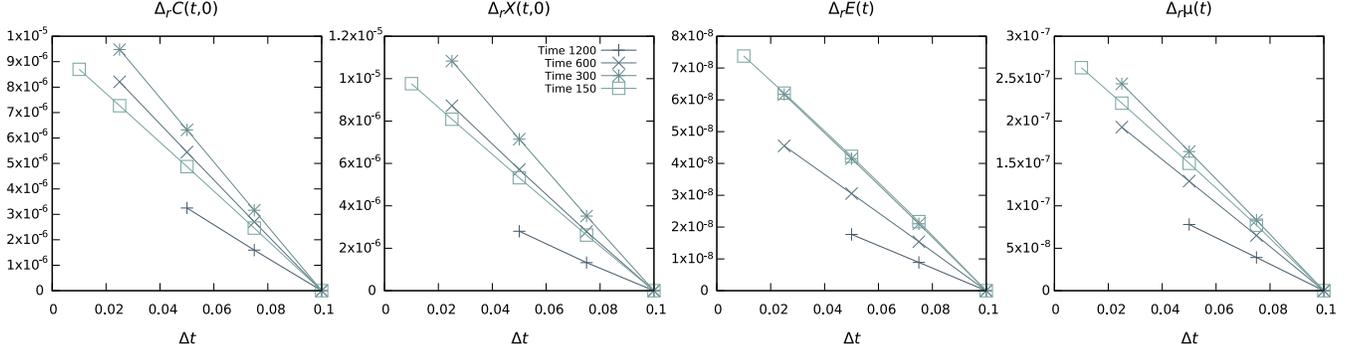
FIG. 26. Relative errors, as defined in Eq. (F1), for $C(t,0)$, $\chi(t,0)$, $E(t)$ and $\mu(t)$, obtained in the integration of the dynamical equations for the $(3+4)$ spin with $T = T_{MCT}$. The dependence is linear in the integration step $\Delta t$, and the resulting integration error in using $\Delta t = 0.1$ is very small compared to the precision needed in the asymptotic evaluation of the observables.

integration step $\Delta t = 0.1$ instead of the exact integration ($\Delta t \to 0$).

In Fig. 26, we show the relative errors in some one-time quantities, $C(t,0)$, $\chi(t,0)$, $E(t)$, and $\mu(t)$, obtained in the integration of the dynamical equations for the $(3+4)$ spin with $T = T_{MCT}$. We clearly see that, at any time $t$, the relative error scales roughly linearly with $\Delta t$. For any time $t$, the relative errors are extremely small: of order $10^{-5}$ for correlation and response, and of order $10^{-7}$ for the energy and radial reaction. Moreover, the coefficient of the linear relation does not seem to grow indefinitely with time, but it actually shows a maximum around $t = O(100)$.

In order to better appreciate the nonmonotonic behavior of the relative errors, we plot in Fig. 27 the integration error in the energy obtained starting from $T = T_{MCT}$ as a function of time. The fact that relative errors decrease at very large times is probably related to another important observation: Trajectories integrated with different $\Delta t$ may differ sensibly at short times, but at large times, they seem to converge to the same asymptotic solution and are thus
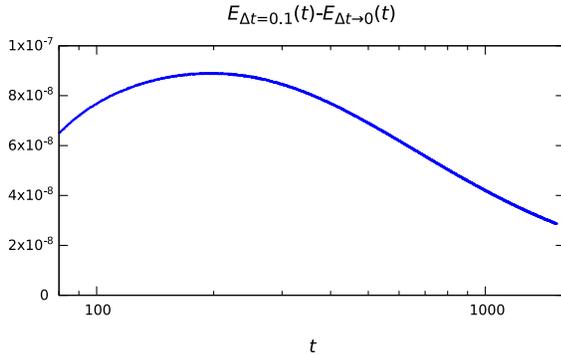


FIG. 27. Relative errors in the integration of the dynamical equations for the $(3+4)$ spin, which show a maximum at a finite time. Here, we plot the relative error in $E(t)$ during the integration with $T = T_{MCT}$. We notice that the maximum relative error is much smaller, by several orders of magnitude, than the physical effect we seek (energy going under the threshold).

very attractive and stable (this result may be the reason why integration errors at very large times are smaller than those at intermediate times). The results of the integration shown in the present work are very reliable and stable.

Despite the high precision at finite times, the extrapolation to large $t$ of physical quantities, notably $\mu(t)$, $E(t)$, or $C(t,0)$, is delicate. While $C(t,0)$ decays very slowly and we do not attempt any extrapolation in the $t \to \infty$ limit, both $\mu(t)$ and $E(t)$ converge fast enough, and their large-time limits can be estimated.

We already noticed in Fig. 5 that for $T \geq T_{SF}$, the radial reaction $\mu(t)$ is perfectly compatible with an extrapolation to the marginal value $\mu_{mg} = 6$. At very high temperatures, the decay is a simple power law, while at lower temperatures, we observe preasymptotic effects and an eventual crossover to the asymptotic decay. The latter seems to be characterized by an exponent that is roughly temperature independent.

Given the critical character of the dynamics in the aging regimes I and II, we have estimated the asymptotic energy $E_\infty(T)$ via power-law fits to $E(t)$ data, and we show the results in Fig. 28. While at high temperatures the data follow a nice power law in time for a couple of decades, at lower temperatures the presence of the preasymptotic regime and the crossover to the asymptotic one makes it very hard to assess the reliability of the value of $E_\infty(T)$ that we extrapolate in this way.

However, comparing the data shown in Fig. 5 for $\mu(t)$ and in Fig. 28 for $E(t)$, we notice a very similar behavior (including the crossover), and thus it could be useful to parametrically plot $E(t)$ versus $\mu(t)$, varying $t$, to see whether a better estimation of $E_\infty(T)$ could be obtained. This method is used in Fig. 29. The behavior of the energy as a function of the radial reaction is very smooth, practically linear at high temperatures and close to quadratic approaching $T_{SF}$. We can then fit the relation $E(\mu)$ and obtain a very good estimate of $E_\infty(T)$ by just assuming that $\mu(t)$ converges to the marginal value $\mu_{mg}$ in the large-time limit. We have interpolated the data via the function
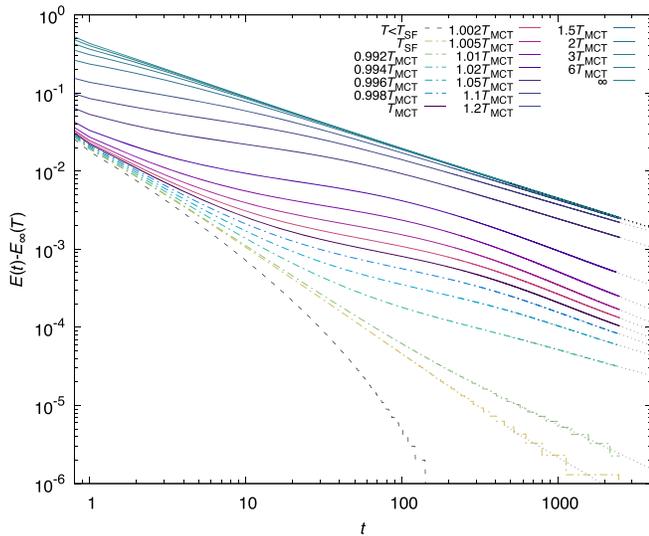
FIG. 28. Energy decay with time in the $(3+4)$-spin model on a double-logarithmic scale. The asymptotic value $E_\infty(T)$ has been estimated via power-law fits that provide a good description of the asymptotic decay for $T \geq T_{SF}$.

$$E(\mu) = E_\infty + A(\mu_{mg} - \mu)^\alpha, \qquad (F2)$$

reporting the best fits using dotted curves in Fig. 29.

In Fig. 30, we show the two estimates of $E_\infty(T)$ obtained from the procedures just described: the power-law fit of energy versus time and the power-law fit of energy versus radial reaction $\mu_{mg} - \mu(t)$. We see that they are perfectly compatible.

Although we do not estimate the large-time limit of $C(t,0)$ directly by extrapolating in time (in Appendix B, an estimate is provided based on a different approach), we



FIG. 30. Comparison of two different extrapolations of the asymptotic energies $E_\infty(T)$, obtained by fitting versus $t$ and versus $\mu_{mg} - \mu(t)$.

believe it is useful to show the $C(t,0)$ curves so that the readers can decide for themselves. In Fig. 31, the data for $C(t,0)$ are shown, measured both in the pure three-spin model and in the mixed $(3+4)$-spin model. We have decided to show these data as a function of the scaling variable $t^{-1/3}$ that describes the decay quite well starting from a random configuration ($T = \infty$). *En passant*, we notice that the exact value for this decay is not known. We see in Fig. 31 that, while data for the three-spin model can be easily extrapolated to zero in the large-time limit, the data for the $(3+4)$-spin model seem much more compatible with a nonzero limit $\lim_{t\to\infty} C(t,0)$ when the temperature gets close to $T_{MCT}$. An alternative representation of the same data is provided in Fig. 32.



FIG. 29. Parametric plot of $E(t)$ versus $\mu_{mg} - \mu(t)$, varying $t$, which allows for more reliable extrapolations. The dotted lines are the best power-law fits in the form $E = E_\infty(T) + A(T)(\mu_{mg} - \mu)^{\alpha(T)}$ used to estimate the asymptotic energies.
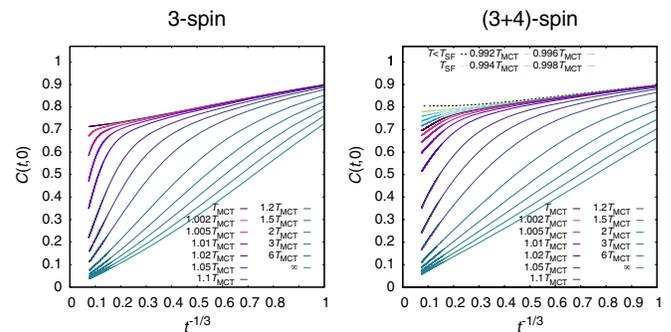


FIG. 31. Decay of the correlation with the initial configuration $C(t,0)$ vs $t^{-1/3}$ in the three-spin model (upper panel) and the $(3+4)$-spin model (lower panel). Our choice of using the $t^{-1/3}$ variable is to make the relaxation from $T = \infty$ as linear as possible [notice, however, that the exact value for the decay exponent of $C(t,0)$ is unknown even in this case]. While data for the three-spin model can be easily extrapolated to zero in the large-time limit, the data for the $(3+4)$-spin model seem much more compatible with a nonzero limit $q_{12} = \lim_{t\to\infty} C(t,0)$ when the temperature gets close to $T_{MCT}$.
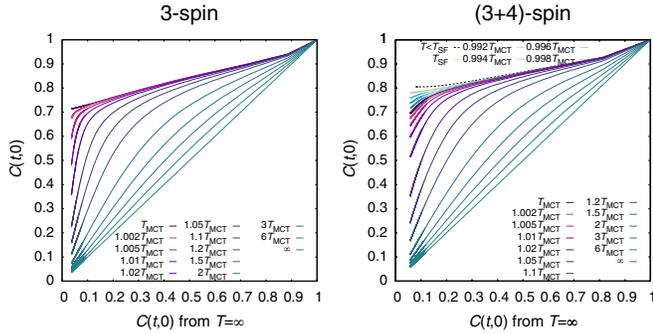
FIG. 32. Same data for the decay of the correlation with the initial configuration $C(t, 0)$ as in Fig. 31, but here instead of the time, we use the $C(t, 0)$ measured starting from a random configuration ($T = \infty$) on the $x$ axis. As in Fig. 31, the data for the three-spin model extrapolate to zero, while the data for the $(3 + 4)$-spin model seem much more compatible with a nonzero limit $q_{12}$ when the temperature gets close to $T_{\text{MCT}}$.

The reason why we have not been able to integrate the dynamical equations for a longer time resides in the difficulties in using integration algorithms with adaptive integration steps. These kinds of adaptive algorithms have been used in the past to integrate the dynamical equations in the case of the pure $p$-spin model starting from random initial conditions [47,48]. Unfortunately, we found that these adaptive algorithms are not robust to changes in the details of the dynamics, and their stability strictly depends on the regime under consideration. At each contraction of the grid, some information on the correlation and response functions computed in the previous steps is lost, which does not seem to be relevant if one starts from a random state and memory of the initial conditions is lost in the asymptotic state. However, in the mixed $p$-spin model, where memory of the initial configuration crucially affects the large-time behavior, the error induced by the time-step adaptation gives rise to strong numerical instabilities. Increasing the grid size pushes the instability to later times, but unfortunately, even the largest grid size we could use ($8192 \times 8192$) did not allow substantial improvements with respect to the simple Euler algorithm.

[1] J. Friedman, T. Hastie, and R. Tibshirani, *The Elements of Statistical Learning* Springer Series in Statistics Vol. 1 (Springer, New York, 2001).

[2] M. Mezard, M. Mezard, and A. Montanari, *Information, Physics, and Computation* (Oxford University, New York, 2009).

[3] T. R. Kirkpatrick, D. Thirumalai, and P. G. Wolynes, *Scaling Concepts for the Dynamics of Viscous Liquids Near an Ideal Glassy State*, Phys. Rev. A **40**, 1045 (1989).

[4] A. Cavagna, *Supercooled Liquids for Pedestrians*, Phys. Rep. **476**, 51 (2009).

[5] L. Berthier and G. Biroli, *Theoretical Perspective on the Glass Transition and Amorphous Materials*, Rev. Mod. Phys. **83**, 587 (2011).

[6] G. Biroli and J.-P. Bouchaud, *The Random First-Order Transition Theory of Glasses: A Critical Assessment*, in *Structural Glasses and Supercooled Liquids: Theory, Experiment, and Applications* (Wiley Online Library, 2012), pp. 31–113 https://doi.org/10.1002/9781118202470.ch2.

[7] G. Parisi, P. Urbani, and F. Zamponi, *Theory of Simple Glasses: Exact Solutions in Infinite Dimensions* (Cambridge University, Cambridge, England, 2020).

[8] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Optimization by Simulated Annealing*, Science **220**, 671 (1983).

[9] D. J. Earl and M. W. Deem, *Parallel Tempering: Theory, Applications, and New Perspectives*, Phys. Chem. Chem. Phys. **7**, 3910 (2005).

[10] A. Finnila, M. Gomez, C. Sebenik, C. Stenson, and J. Doll, *Quantum Annealing: A New Method for Minimizing Multidimensional Functions*, Chem. Phys. Lett. **219**, 343 (1994).

[11] M. Mézard, G. Parisi, and R. Zecchina, *Analytic and Algorithmic Solution of Random Satisfiability Problems*, Science **297**, 812 (2002).

[12] D. L. Donoho, A. Maleki, and A. Montanari, *Message Passing Algorithms for Compressed Sensing: I. Motivation and Construction*, in *2010 IEEE Information Theory Workshop on Information Theory* (IEEE, New York, 2010), pp. 1–5.

[13] A. Crisanti and H.-J. Sommers, *The Spherical p-Spin Interaction Spin Glass Model: The Statics*, Z. Phys. B **87**, 341 (1992).

[14] A. Crisanti, H. Horner, and H.-J. Sommers, *The Spherical p-Spin Interaction Spin-Glass Model*, Z. Phys. B **92**, 257 (1993).

[15] L. F. Cugliandolo and J. Kurchan, *Analytical Solution of the Off-Equilibrium Dynamics of a Long Range Spin-Glass Model*, Phys. Rev. Lett. **71**, 173 (1993).

[16] A. Crisanti and H.-J. Sommers, *Thouless-Anderson-Palmer Approach to the Spherical p-Spin Spin Glass Model*, J. Phys. I **5**, 805 (1995).

[17] T. Castellani and A. Cavagna, *Spin-Glass Theory for Pedestrians*, J. Stat. Mech. (2005) P05012.

[18] A. Choromanska, M. Henaff, M. Mathieu, G. Ben Arous, and Y. LeCun, *The Loss Surfaces of Multilayer Networks*, in *Artificial Intelligence and Statistics* (Microtome Publishing, 2015), pp. 192–204.

[19] A. Montanari and E. Richard, *A Statistical Model for Tensor PCA*, in *Adv. Neural Inf. Process. Syst.* (Neural Information Processing Systems Foundation, Inc., 2014), pp. 2897–2905.

[20] G. Ben Arous, R. Gheissari, and A. Jagannath, *Algorithmic Thresholds for Tensor PCA*, arXiv:1808.00921.

[21] A. Auffinger and W.-K. Chen, *On the Energy Landscape of Spherical Spin Glasses*, Adv. Math. **330**, 553 (2018); E. Subag, *The Geometry of the Gibbs Measure of Pure Spherical Spin Glasses*, Inventiones Mathematicae **210**, 135 (2017); A. Auffinger and G. Ben Arous, *Complexity of Random Smooth Functions on the High-Dimensional Sphere*, Ann. Prob. **41**, 4214 (2013) and references therein.

[22] W. Götze, *Complex Dynamics of Glass-Forming Liquids: A Mode-Coupling Theory* (Oxford University, New York, 2008), Vol. 143.

[23] J.-P. Bouchaud, *Weak Ergodicity Breaking and Aging in Disordered Systems*, J. Phys. I **2**, 1705 (1992).

[24] M. Bernaschi, A. Billoire, A. Maiorano, G. Parisi, and F. Ricci-Tersenghi, *Strong Ergodicity Breaking in Aging of Mean Field Spin Glasses*, arXiv:1906.11195.

[25] S. Sastry, P. G. Debenedetti, and F. H. Stillinger, *Signatures of Distinct Dynamical Regimes in the Energy Landscape of a Glass-Forming Liquid*, Nature (London) **393**, 554 (1998).

[26] A. Banerjee, M. K. Nandi, S. Sastry, and S. M. Bhattacharyya, *Determination of Onset Temperature from the Entropy for Fragile to Strong Liquids*, J. Chem. Phys. **147**, 024504 (2017).

[27] A. Barrat, S. Franz, and G. Parisi, *Temperature Evolution and Bifurcations of Metastable States in Mean-Field Spin Glasses, with Connections with Structural Glasses*, J. Phys. A **30**, 5593 (1997).

[28] A. Crisanti and L. Leuzzi, *Spherical $2 + p$ Spin-Glass Model: An Exactly Solvable Model for Glass to Spin-Glass Transition*, Phys. Rev. Lett. **93**, 217203 (2004).

[29] F. Sciortino, W. Kob, and P. Tartaglia, *Inherent Structure Entropy of Supercooled Liquids*, Phys. Rev. Lett. **83**, 3214 (1999).

[30] E. Lerner and E. Bouchbinder, *Effect of Instantaneous and Continuous Quenches on the Density of Vibrational Modes in Model Glasses*, Phys. Rev. E **96**, 020104 (2017).

[31] C. Rainone, E. Bouchbinder, and E. Lerner, *Pinching a Glass Reveals Key Properties of Its Soft Spots*, Proc. Natl. Acad. Sci. U.S.A. **117**, 5228 (2020).

[32] L. Wang, A. Ninarello, P. Guan, L. Berthier, G. Szamel, and E. Flenner, *Low-Frequency Vibrational Modes of Stable Glasses*, Nat. Commun. **10**, 1 (2019).

[33] M. Baity-Jesi and V. Martín-Mayor, *Precursors of the Spin Glass Transition in Three Dimensions*, J. Stat. Mech. (2019) 084016.

[34] S. Franz and G. Parisi, *Recipes for Metastable States in Spin Glasses*, J. Phys. I **5**, 1401 (1995).

[35] A. Barrat, R. Burioni, and M. Mzard, *Dynamics within Metastable States in a Mean-Field Spin Glass*, J. Phys. A **29**, L81 (1996).

[36] A. Cavagna, I. Giardina, and G. Parisi, *On the Stationary Points of the TAP Free Energy*, Phys. Rev. B **57**, 11251 (1998).

[37] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, *Fractal Free Energy Landscapes in Structural Glasses*, Nat. Commun. **5**, 3725 (2014).

[38] L. Berthier, G. Biroli, P. Charbonneau, E. I. Corwin, S. Franz, and F. Zamponi, *Gardner Physics in Amorphous Solids and Beyond*, J. Chem. Phys. **151**, 010901 (2019).

[39] Y. Sun, A. Crisanti, F. Krzakala, L. Leuzzi, and L. Zdeborová, *Following States in Temperature in the Spherical s + p-spin Glass Model*, J. Stat. Mech. (2012) P07002.

[40] A. Auffinger, G. Ben Arous, and J. Cerny, *Random Matrices and Complexity of Spin Glasses*, Commun. Pure Appl. Math. **66**, 165 (2013).

[41] As is well known in random matrix theory, while macroscopic fluctuations of the spectral bulk are suppressed by factors scaling exponentially in $N^2$, the probability of having a finite number of eigenvalues at finite distance from the spectral support is exponentially small in $N$. As in the case of the pure model, given the value of $\mu$ such that the spectral support is positive, besides the dominating minima, we have saddle points with a finite number of negative directions.

[42] G. Ben Arous, E. Subag, and O. Zeitouni, *Geometry and Temperature Chaos in Mixed Spherical Spin Glasses at Low Temperature: The Perturbative Regime*, Commun. Pure Appl. Math. **73**, 1732 (2018).

[43] The index counts the number of negative eigenvalues of the Hessian at a stationary point.

[44] H. Sompolinsky and A. Zippelius, *Relaxational Dynamics of the Edwards-Anderson Model and the Mean-Field Theory of Spin-Glasses*, Phys. Rev. B **25**, 6860 (1982).

[45] M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications*, World Scientific Lecture Notes in Physics (World Scientific, Singapore, 1987).

[46] S. Franz and M. Mezard, *On Mean Field Glassy Dynamics out of Equilibrium*, Physica A **210**, 48 (1994).

[47] B. Kim and A. Latz, *The Dynamics of the Spherical p-Spin Model: From Microscopic to Asymptotics*, Europhys. Lett. **53**, 660 (2001).

[48] L. Berthier, G. Biroli, J.-P. Bouchaud, W. Kob, K. Miyazaki, and D. Reichman, *Spontaneous and Induced Dynamic Correlations in Glass-Formers II: Model Calculations and Comparison to Numerical Simulations*, J. Chem. Phys. **126**, 184504 (2007).

[49] L. F. Cugliandolo, J. Kurchan, and L. Peliti, *Energy Flow, Partial Equilibration and Effective Temperatures in Systems with Slow Dynamics*, Phys. Rev. E **55**, 3898 (1997).

[50] B. Capone, T. Castellani, I. Giardina, and F. Ricci-Tersenghi, *Off-Equilibrium Confined Dynamics in a Glassy System with Level-Crossing States*, Phys. Rev. B **74**, 144301 (2006).

[51] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics* (Springer Science & Business Media, New York, 2012), Vol. 31.

[52] D. J. Gross, I. Kanter, and H. Sompolinsky, *Mean-Field Theory of the Potts Glass*, Phys. Rev. Lett. **55**, 304 (1985).

[53] E. Gardner, *Spin Glasses with p-Spin Interactions*, Nucl. Phys. **B257**, 747 (1985).

[54] A. Montanari and F. Ricci-Tersenghi, *On the Nature of the Low-Temperature Phase in Discontinuous Mean-Field Spin Glasses*, Eur. Phys. J. B **33**, 339 (2003).

[55] A. Montanari and F. Ricci-Tersenghi, *Cooling-Schedule Dependence of the Dynamics of Mean-Field Glasses*, Phys. Rev. B **70**, 134406 (2004).

[56] S. S. Mannelli, G. Biroli, C. Cammarota, F. Krzakala, and L. Zdeborová, *Who Is Afraid of Big Bad Minima? Analysis of Gradient-Flow in Spiked Matrix-Tensor Models*, in *Adv. Neural Inf. Process. Syst.* (Neural Information Processing Systems Foundation, Inc., 2019), pp. 8676–8686.

[57] Y. M. Lu and G. Li, *Phase Transitions of Spectral Initialization for High-Dimensional Nonconvex Estimation*, arXiv:1702.06435.

[58] L. F. Cugliandolo, J. Kurchan, and L. Peliti, *Energy Flow, Partial Equilibration, and Effective Temperatures in Systems with Slow Dynamics*, Phys. Rev. E **55**, 3898 (1997).

[59] S. Franz, M. Mézard, G. Parisi, and L. Peliti, *Measuring Equilibrium Properties in Aging Systems*, Phys. Rev. Lett. **81**, 1758 (1998).

[60] M. Baity-Jesi, E. Calore, A. Cruz, L. A. Fernandez, J. M. Gil-Narvión, A. Gordillo-Guerrero, D. Iñiguez, A. Maiorano, E. Marinari, V. Martin-Mayor *et al.*, *A Statics-Dynamics Equivalence through the Fluctuation-Dissipation Ratio Provides a Window into the Spin-Glass Phase from Nonequilibrium Measurements*, Proc. Natl. Acad. Sci. U.S.A. **114**, 1838 (2017).

[61] Instead of the standard symbol $\beta$, we use $\hat{\beta}$ to avoid confusion with the inverse temperature.