Path integral approach unveils role of complex energy landscape for activated dynamics of glassy systems

Tommaso Rizzo●

ISC-CNR, UOS Rome, Università "Sapienza," Piazzale A. Moro 2, I-00185 Rome, Italy
and Dipartimento di Fisica, Università "Sapienza," Piazzale A. Moro 2, I–00185 Rome, Italy

The complex dynamics of an increasing number of systems is attributed to the emergence of a rugged energy landscape with an exponential number of metastable states. To develop this picture into a predictive dynamical theory, I discuss how to compute the exponentially small probability of a jump from one metastable state to another. This is expressed as a path integral that can be evaluated by saddle-point methods in mean-field models, leading to a boundary value problem. The resulting dynamical equations are solved numerically by means of a Newton-Krylov algorithm in the paradigmatic spherical p-spin glass model that is invoked in diverse contexts from supercooled liquids to machine-learning algorithms. I discuss the solutions in the asymptotic regime of large times and the physical implications on the nature of the ergodicity-restoring processes.

I. INTRODUCTION

The emergence of a rugged free energy landscape with many minima and saddles is a paradigm often invoked to explain complex dynamical phenomena ranging from supercooled liquids [1] to the performance of widely used algorithm in machine learning and inference [2,3]. Powerful methods developed originally in the spin-glass literature [4] allow to obtain a rather complete description of the landscape in a growing number of statistical physics model. On the debit side the resulting picture is essentially limited to *static* properties of the landscape, like the energy of the metastable states, their free energy and notably their number, the so-called configurational entropy. Thus understanding the exact way in which the landscape shapes the dynamics is largely an open problem. The question is particularly urgent in the context of supercooled liquids where the rugged landscape paradigm is at the core of the random-first-frder-transition (RFOT) theory [5,6].

More than thirty years after its formulation RFOT is still one of the major competing theories in the ongoing debate on the nature of the glass transition [7]. In a nutshell, the theory posits that the physics of supercooled liquids is the same of spin-glass (SG) models displaying one step of Parisi's replica-symmetry breaking (1RSB) [4]. The mean-field versions of these models display an ergodicity-breaking transition at a dynamical temperature T_d where the phase space splits into many metastable states that trap the dynamics; at lower temperatures, the configurational entropy, i.e., the logarithm of the number of metastable states, decreases eventually vanishing at a static temperature T_s. Ergodicity breaking between T_d and T_s is a mean-field artifact and one expects that in real systems ergodicity is restored through dropletlike excitations. Furthermore the size of these excitations must diverge as the configurational entropy vanishes leading eventually to a

genuine ergodicity-breaking transition at T_s. RFOT originated from the realization [8] that similar features had been discussed in various unrelated (and themselves controversial) theories of supercooled liquids, most notably (1) dynamics at the ergodicity-breaking transition is the same of the (avoided) mode-coupling theory (MCT) of supercooled liquids [9] and (2) within the Adams-Gibbs-Di Marzio theory [10,11], the glass transition is driven by a correlation length that diverges at the Kauzmann temperature where the configurational entropy vanishes.

Efforts to validate the theory have been driving theoretical, numerical, and experimental research for years. At the theoretical level, RFOT has been substantiated by a number of results arguing that mean-field models of supercooled liquids exhibit 1RSB [12,13] including the solution of the limit of infinite physical dimensions [1,14]. Numerically, the MCT phenomenology is well documented [15,16] as well as the increase of dynamic [17,18] and static correlation lengths [19,20]. Experimentally, the observation of a decreasing configurational entropy dates back to the 40's, while more recently RFOT has inspired measurements of nonlinear susceptibilities [21]. Furthermore the analogy with supercooled liquids has also led to the discovery that off-equilibrium relaxational dynamics of mean-field Spin-glass models display aging [22] and it is an active line of research [23,24].

In spite of this huge body of work, the consensus on the validity of the theory is still lacking. One of the problem is that RFOT-inspired literature often focuses on quantities, namely point-to-set correlation lengths and configurational entropy, whose actual relevance for the problem of the glass transition, that instead is dynamical in its essence, can also be questioned. Besides, predictions are often merely qualitatively, which is a problem given that, e.g., the observed static length increases are too modest to convince the community that they actually drive the slowing down of the dynamics. To make
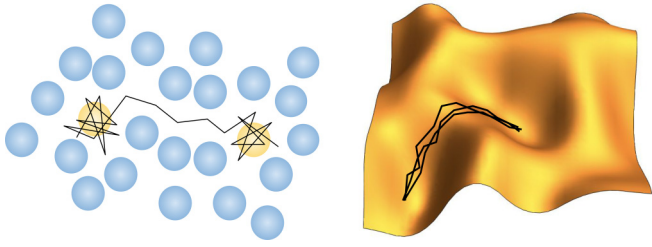
094203-1

FIG. 1. Supercooled liquids display caging and hopping in real space (left) that correspond to evolution in the rugged phase space of mean-field models (right).

progress it would be important to obtain RFOT predictions that are both *quantitative* and *dynamical*. The main challenge is that current theoretical knowledge is mostly limited to mean-field models that display ergodicity breaking at T_d: in order to recover ergodicity between T_d and T_s and obtain realistic predictions we have to go beyond mean-field, intense efforts in this line of research are currently underway [25–32]. In recent years, progress has been made for the temperature region *close* to the dynamical temperature T_d [33–35]. It is now possible to describe qualitatively and quantitatively how the ergodicity-breaking MCT transition is turned into a dynamical crossover in mean-field SG [36] (due to finite-size effects) and most importantly in some finite-dimensional models [37]. In this paper, I consider instead the region between T_d and T_s where metastable states are present and discuss how to compute the transition rate, that is *the exponentially small probability* of a jump from an equilibrium state to another occurring in a finite time. To make contact with the phenomenology of supercooled liquids, we have to remember that below the experimental MCT transition temperature a particle is trapped most of the time into a cage formed by the surrounding particles and diffusion occurs through *hopping*, i.e., sudden rare jumps from a cage to another (see Fig. 1). Mean-field models capture *caging* through the appearance of metastable states and the study of jumps in the free energy landscape initiated in this paper is essential to a quantitative description of *hopping* in real space.

A. Main results

The main results to be discussed in the paper are (i) a path integral method to compute the transition rate, (ii) a Newton-Krylov algorithm yielding the numerical solution of the corresponding dynamical equations, (iii) an asymptotic analysis of the solutions in the regime of large times, and (iv) some nontrivial implications on the ergodicity-restoring processes. At the methodological level, the problem is formulated as a path integral over Langevin dynamic trajectories that can be computed though saddle-point methods in mean-field models. The problem displays some important differences with respect to the standard relaxational dynamics [22,38,39], namely, the use of replicas and the need to explicitly handle the divergent path integral. I have focused on the paradigmatic spherical *p*-spin SG model [39,40] but the method is fairly general and the equations can be derived with some effort for other mean-field systems, e.g., supercooled liquids in large dimension [41–43]. Another more important difference follows

from to the fact that while in ordinary relaxation dynamics one only fixes the *initial* condition, in order to study activated dynamics one has to fix both the *initial* and *final* conditions. This difference manifests itself at the level of the dynamical equations: while relaxational equations display first-order time derivatives the equations obtained here are second-order. One must then solve a more difficult boundary value problem instead of a simpler initial value problem.[1] Indeed, while the relaxational dynamical equations can be solved at times $t + \Delta t$ iteratively using the solution at times $t' \leqslant t$ the activated dynamical equations must be solved simultaneously at all times, something for which no standard algorithm exists. A successful solution strategy has been developed based on three elements: (1) Newton's method on discretized equations, (2) Krylov methods with physical preconditioning to invert the Jacobian, and (3) Richardson extrapolation to reach the continuum limit. The whole technology can be again exported to other problems, with possible computational complexity issues due to the specific dynamical order parameter.

The paper is organized as follows. In the remaining of this section, I will give a compact presentation of the main results, leaving the details for the body of the paper and the appendices. In particular, in Sec. I A 1, the object I compute is introduced and its general properties discussed. Later on, I discuss the asymptotic behavior of the solutions that is interesting for a number of physical and technical reasons. From the physical point of view, the most interesting outcome of the computation is associated to the ergodicity restoring processes as discussed in Sec. I A 3. The actual dynamical equations are given later in Sec. II, their solutions will be studied in various regimes (free, ergodic, and activated). The numerical solution is challenging and will be discussed in Sec. II F. In Sec. III, I will give some concluding perspectives.

1. The transition rate

The object considered is the transition rate $T_{2\mathcal{T}}(\sigma|\tau)$ defined as the probability that the system is in configuration σ at time $t_{\text{fin}} = \mathcal{T}$ given that it was in configuration τ at time $t_{\text{in}} = -\mathcal{T}$. It is convenient to actually consider the following object that, due to detailed balance, is symmetric with respect to the exchange of σ and τ:

$$\hat{T}_{2\mathcal{T}}(\sigma, \tau) \equiv T_{2\mathcal{T}}(\sigma|\tau) e^{\frac{\beta}{2}H(\sigma) - \frac{\beta}{2}H(\tau)}. \qquad (1)$$

An integral representation of Langevin dynamics is used and σ, τ are chosen as generic *equilibrium* configurations. At low temperatures, $\hat{T}(\sigma, \tau)$ is exponentially small in the system size N in mean-field models, therefore an annealed average $P_{\text{eq}}(\sigma)P_{\text{eq}}(\tau)\hat{T}(\sigma, \tau)$ would interfere with the equilibrium measure of σ and τ and the correct procedure is to consider the quenched average

$$[\ln \hat{T}] \equiv \sum_{\sigma, \tau} P_{\text{eq}}(\sigma)P_{\text{eq}}(\tau) \ln \hat{T}(\sigma, \tau) . \qquad (2)$$

---

[1]This is also the main difference with the earlier method of Refs. [53,54] and more recently of Ref. [32] that allow to study some special activated processes through relaxational dynamics equations but cannot be used to study transition rates between generic equilibrium states.

One can resort to the replica method to eliminate the logarithm, besides, if quenched disorder is present, the corresponding averages (represented by an overline in the following) require the introduction of additional replicas of the initial and final configurations.

One can argue that the rate is self-averaging, meaning that most couples (σ, τ) display the same rate,[2] thus, given an initial configuration σ, the total number of configurations with rate equal to the typical value given by the average $\overline{[\ln \hat{T}_{2\mathcal{T}}]}$ is equal to the total number of equilibrium configurations τ, i.e., the exponential of the entropy e^S. Neglecting exponentially small corrections, the total probability of jumping to one of the equilibrium configurations is thus $e^{[\ln \hat{T}]+S}$ and must be smaller than one leading to the bound:

$$\overline{[\ln \hat{T}_{2\mathcal{T}}]} + S \leqslant 0. \tag{3}$$

Now, after a finite time a system in equilibrium will be in another equilibrium configuration *correlated* with the initial one, therefore the probability to be in a *generic* equilibrium configuration (that is uncorrelated to the initial one) must be smaller than one meaning that at any finite $\mathcal{T}$ the above bound should not be saturated. This also implies that even for $T > T_d$ if $\mathcal{T} < \infty$, one must consider the logarithm of the rate (and thus resort to the replica method with replica number $n = 0$) in order not to interfere with the equilibrium measure on the initial and final conditions. On the other hand ergodicity implies that when $\mathcal{T}$ goes to infinity the probability measure will be flat over the e^S equilibrium configurations and the average rate should become equal to e^{-S} saturating the bound.

At the mean-field level, the expression for the logarithm of the rate can be computed by the saddle-point approximation meaning that the thermodynamic limit $N \to \infty$ is always taken before the $\mathcal{T} \to \infty$ limit. Thus one can compute the quantity

$$\mathcal{A} \equiv \lim_{\mathcal{T} \to \infty} \lim_{N \to \infty} \frac{1}{N} \overline{([\ln \hat{T}_{2\mathcal{T}}] + S)}. \tag{4}$$

In 1RSB models, as discussed in the introduction, the dynamical temperature T_d marks the onset of activated dynamics and this results in the two limits ceasing to commute:

$$\mathcal{A} = 0 \quad \text{for } T \geqslant T_d, \tag{5}$$

$$\mathcal{A} < 0 \quad \text{for } T < T_d. \tag{6}$$

In the following sections, the expression for the average transition rate of the spherical p-spin model for $T > T_s$ is given. The Hamiltonian of the model is given by

$$H = \sum_{p=1}^{\infty} \sum_{i_1 < \cdots < i_p} J_{i_1 \ldots i_p} s_{i_1} \ldots s_{i_p}, \tag{7}$$

where the N spins verify the global spherical constraint $\sum_i s_i = N$, the J's are quenched Gaussian random variables

of zero mean and variance:

$$\overline{J_{i_1 \ldots i_p}^2} = \frac{\mu_p \, p!}{2N^{p-1}}. \tag{8}$$

The following function is defined for convenience:

$$f(x) \equiv \sum_{p=1}^{\infty} \mu_p \, x^p. \tag{9}$$

The expression for the rate $\overline{[\ln \hat{T}_{2\mathcal{T}}]}$ depends on six real functions $C(t, t')$, $\hat{R}(t, t')$, $\hat{X}(t, t')$, $C^{dt}(t, t')$, $\hat{R}^{dt}(t, t')$, and $\hat{X}^{dt}(t, t')$ defined for $-\mathcal{T} \leqslant t, t' \leqslant \mathcal{T}$. Two additional functions $\mu(t)$ and $\hat{\mu}(t)$ enforce the spherical constraint leading to $C(t, t) = 1$ and $\hat{R}(t, t) = 1/2$ for all t. Extremization of the expression leads to eight nonlinear integrodifferential equations (albeit two are redundant due to symmetries) that will be discussed in the next section. The physical meaning of $C(t, t')$ and $C^{dt}(t, t')$ is straightforward:

$$C(t, t') = \overline{[\langle s_i(t) s_i(t') \rangle]}, \tag{10}$$

$$C^{dt}(t, t') = \overline{[\langle s_i(t) \rangle \langle s_i(t') \rangle]}, \tag{11}$$

where the square brackets mean averages with respect to the dynamical trajectories at fixed initial and final configurations (σ, τ) and fixed disorder. Thus $C(t, t')$ is the average correlation between configurations visited by the same trajectory at times t and t' while $C^{dt}(t, t')$ is the correlation between configurations visited by *different trajectories* (hence the suffix dt). It follows that the equations must satisfy the boundary conditions $C(\pm\mathcal{T}, \pm\mathcal{T}) = C^{dt}(\pm\mathcal{T}, \pm\mathcal{T}) = 1$ and $C(\pm\mathcal{T}, \mp\mathcal{T}) = C^{dt}(\pm\mathcal{T}, \mp\mathcal{T}) = 0$.

In Sec. II D, the $T > T_d$ regime will be discussed, here one can show that in the $\mathcal{T} \to \infty$ limit the saddle-point equations admit a solution in which the six functions are expressed in terms of the equilibrium correlation $C_{\text{eq}}(0, t)$ and that the rate tends to $-S$ leading to $\mathcal{A} = 0$. This is possible because $C_{\text{eq}}(0, \infty) = 0$ and the boundary condition $C(\pm\mathcal{T}, \mp\mathcal{T}) = 0$ can be satisfied for $\mathcal{T} \to \infty$ by the equilibrium solution. This is no longer true for $T < T_d$ because equilibrium dynamics is trapped and $C_{\text{eq}}(0, \infty) \neq 0$, as a consequence for $T < T_d$ the average logarithm of the rate is smaller than $-S$ also in $\mathcal{T} \to \infty$ limit meaning that $\mathcal{A}$ becomes negative *continuously* at T_d. In Fig. 2, the numerical solution for the spherical SG model with $p = 3$ is shown for $T = 1/1.695 < T_d$ and $\mathcal{T} = 40$. The transition from the *ergodic* ($T > T_d$) to the *activated* ($T < T_d$) regime is marked by a *qualitative* change in the solutions. As it will be shown in Sec. II D in the ergodic phase at large $\mathcal{T}$ the various functions approach equilibrium time-translational-invariant forms [see Eqs. (76)–(78)] and density plots like those of Fig. 2 tend to become symmetric with respect to the $t = t'$ axes. Instead in the activated phase the density plots for $C(t, t')$ and $C^{dt}(t, t')$ display the block structure visible in Fig. 2 at all values of $\mathcal{T}$, corresponding to the fact that the system close to the initial and final time performs essentially a relaxational dynamics in the corresponding states and jumps from one state to the other for $t, t' \approx 0$. A quantity that also displays a qualitative change is the instantaneous intensive energy $e(t)$ of the system along the trajectory from one equilibrium configuration to another one. By definition we have that $e(\pm\mathcal{T})$ is equal to the equilibrium

---

[2]This can be shown computing $O(n)$ corrections to the rate where n is the replica number.
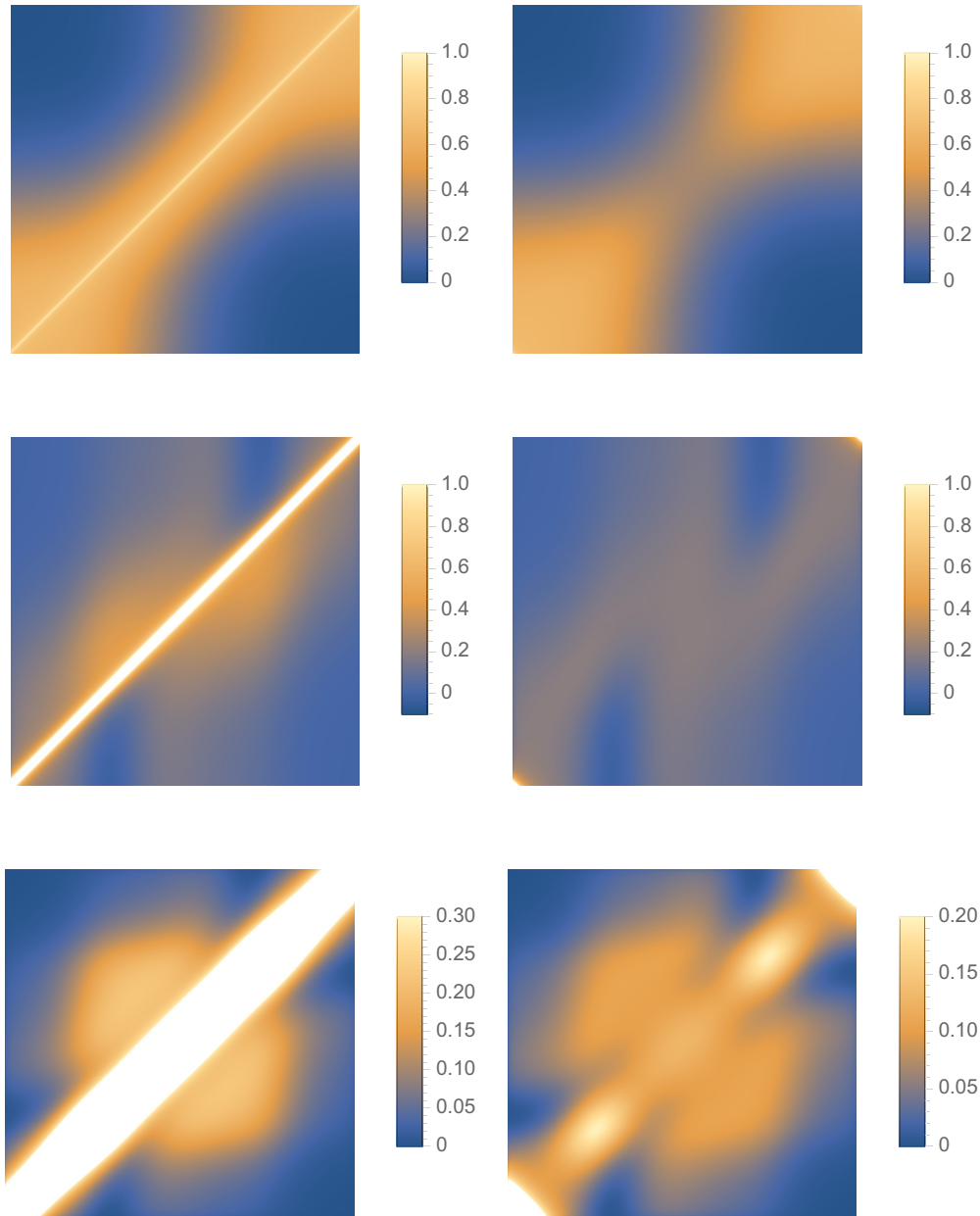
FIG. 2. Left to right, top to bottom: density plot of the six functions $C(t, t')$, $C^{dt}(t, t')$, $\mathcal{T}\hat{R}(t, t')$, $\mathcal{T}\hat{R}^{dt}(t, t')$, $\mathcal{T}^2\hat{X}(t, t')$, $\mathcal{T}^2\hat{X}^{dt}(t, t')$ for $-\mathcal{T} \leqslant t, t' \leqslant \mathcal{T}$, and $\mathcal{T} = 40$ for the spherical 3-SG model at $T = 1/1.695 < T_d$, (numerical solution with $\Delta t = \mathcal{T}/2000$). Too large values close the diagonal are not shown for clarity (white).

value e_{eq} at the corresponding temperature. In general, $e(t)$ is larger than e_{eq} at intermediate times but in the limit $\mathcal{T} \to \infty$, it tends to be equal to e_{eq} at all times for $T > T_d$, while for $T < T_d$, $e(t) > e_{eq}$ also in the large $\mathcal{T}$ limit and indeed one finds that $e(t) \approx e_u(t/\mathcal{T})$ where $e_u(x)$ is a universal function independent of $\mathcal{T}$ as shown by Fig. 3. This kind of asymptotic behavior is also shared by the solution as we will further discuss in the following.

Let us now mentions for completeness a few technical features of the problem that will be further discussed in the following sections. At all temperatures, even in the ergodic phase $T > T_d$, the solutions cannot be expressed in terms of the equilibrium correlation if $\mathcal{T}$ is *finite*, however the equations can be solved analytically at finite $\mathcal{T}$ in the free case

$(T = \infty)$ in which the system performs a Brownian motion on the $N - 1$ dimensional sphere (see Sec. II C). In general the expression for the average rate requires the computation of a path integral with an infinite normalization factor, a well-known pathology that is typically discussed in the context of the harmonic oscillator [44,45]. In the noninteracting case, one can use the harmonic oscillator formulas to derive the expression of the rate as a function of $\mathcal{T}$. Knowledge of the rate in the noninteracting case provides an alternative way to compute the rate at finite temperature without having to deal with the infinite normalization factor. Differentiating the saddle-point expression of the rate with respect to the inverse temperature β one gets indeed a finite expression that yields the correct derivative when evaluated on the solutions of the
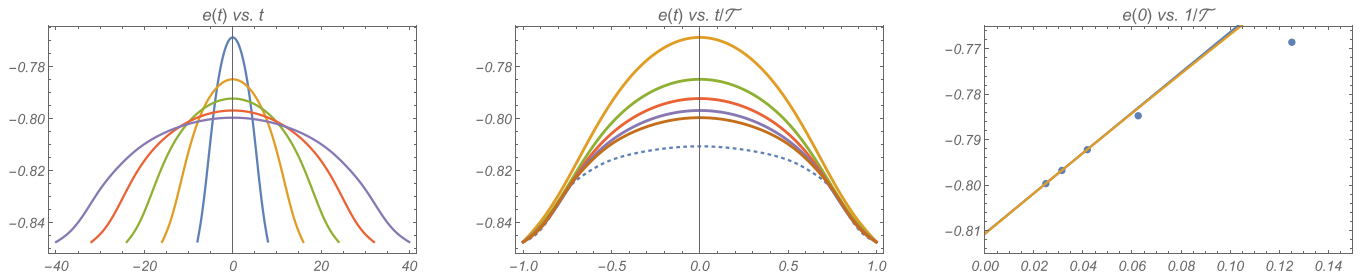
FIG. 3. Spherical 3-SG at $T = 1/1.695 < T_d$. (Left) Instantaneous energy of the trajectory for for $\mathcal{T} = 8$, 16, 24, 32, 40, each curve was obtained from finite N_t solutions through the Richardson extrapolation to $\Delta t = 0$ (see text). (Middle) Same data with times rescaled with $\mathcal{T}$. The dotted line is the $\mathcal{T} = \infty$ limit obtained by linear extrapolation in $1/\mathcal{T}$. (Right) The energy at $t = 0$ plotted vs $1/\mathcal{T}$. Convergence to to the $\mathcal{T} = \infty$ limit with $1/\mathcal{T}$ corrections is clearly visible, the continuous lines are linear fits on the last two and three points. The energy varies continuously along the trajectory from the equilibrium value $e_{eq} = -\beta/2 = -0.8475$ at the initial and final time to larger values at intermediate times. Note that the (extrapolated) maximum value $e(0) = -0.8106$ at $t = 0$ is larger than the threshold energy $e_{th} = -0.8311$, as computed from Eqs. (12) and (13) in Ref. [22].

saddle-point equations. The rate at finite $\mathcal{T}$ and β can then be obtained by integration in β starting from the exact $\beta = 0$ result.

2. Asymptotic behavior

The equations in the activated phase $T < T_d$ can only be studied numerically for finite $\mathcal{T}$. One then faces the problem of extrapolating the results to $\mathcal{T} \to \infty$ in order to avoid transient effects. The values of $\mathcal{T}$ that it was possible to reach numerically allow to identify clearly the *asymptotic behavior* for $\mathcal{T} \to \infty$ in the activated region. In particular, in Sec. II D, we will show see that for finite time differences $|t - t'| = O(1) \ll \mathcal{T}$, the functions $C(t, t')$, $\hat{R}(t, t')$, and $\hat{X}(t, t')$ verify the equilibrium relationships corresponding to fluctuation-dissipation theorem and time-translational invariance meaning that on finite timescales the trajectories are essentially equilibrium trajectories in a self-induced slowly varying field. Instead, on the large timescales $|t - t'| = O(\mathcal{T})$, the solutions approach universal functions independent of $\mathcal{T}$:

$$C(t, t') \approx C_u(t/\mathcal{T}, t'/\mathcal{T}), \tag{12}$$

$$C^{dt}(t, t') \approx C_u^{dt}(t/\mathcal{T}, t'/\mathcal{T}), \tag{13}$$

$$\hat{R}(t, t') \approx \mathcal{T}^{-1}\hat{R}_u(t/\mathcal{T}, t'/\mathcal{T}), \tag{14}$$

$$\hat{R}^{dt}(t, t') \approx \mathcal{T}^{-1}\hat{R}_u^{dt}(t/\mathcal{T}, t'/\mathcal{T}), \tag{15}$$

$$\hat{X}(t, t') \approx \mathcal{T}^{-2}\hat{X}_u(t/\mathcal{T}, t'/\mathcal{T}), \tag{16}$$

$$\hat{X}^{dt}(t, t') \approx \mathcal{T}^{-2}\hat{X}_u^{dt}(t/\mathcal{T}, t'/\mathcal{T}), \tag{17}$$

$$\mu(t) \approx \mu_u(t/\mathcal{T}), \tag{18}$$

$$\hat{\mu}(t) \approx \mathcal{T}^{-1}\hat{\mu}_u(t/\mathcal{T}). \tag{19}$$

The above behavior is clearly visible in Fig. 4 for the function $C(t, t')$. This also implies that the density plots for $C(t, t')$ and $C^{dt}(t, t')$ at different (large) value of $\mathcal{T}$ are indistinguishable. Note also that $\hat{R}(t, t')$ and $\hat{X}(t, t')$ are small for $|t - t'| = O(\mathcal{T})$ while they are finite close to the diagonal, see Fig. 2.

Plugging the above asymptotic expressions in the full dynamical equations one sees that *the universal functions solve the dynamical equations with the second derivatives dropped*, similarly to what happens for equilibrium dynamics at T_d and for off-equilibrium dynamics [22,39]. Closing those equations would allow to work directly at $\mathcal{T} \to \infty$ and is an open problem that is left for future work. At any rate inspection of the equation suggests that the solution should approach their $\mathcal{T} = \infty$ limit with $1/\mathcal{T}$ corrections. This is indeed supported by the data as can be seen considering, e.g., the instantaneous (intensive) energy $e(t)$ along a trajectory, see the right panel of Fig. 3.

One should note that the asymptotic structure of the solutions is utterly different from that of metastability in ferromagnetism. The corresponding computation describes the transition rate from the metastable minimum to the stable one and leads to an instantonic equation in which the second order derivatives is not dropped in the asymptotic limit. As a consequence even if $\mathcal{T} \to \infty$ the jump effectively occurs in time window centered around $t = 0$ that remains *finite* and does not scale with $\mathcal{T}$. Another more striking difference with metastability in ferromagnetism occurs when we consider the ergodicity-restoring processes as discussed next.

3. Ergodicity restoring processes

This work is focused on the *exponentially small* probability that the system jumps to another equilibrium state in a *finite* time, which is complementary to the problem of determining the *exponentially large* timescale τ_{erg} over which the system finds itself into another equilibrium state with *finite* probability.

Given that, according to the discussion before Eq. (3), the exponentially small probability to jump to an equilibrium state with typical rate is $p = e^{[\ln \mathcal{T}] + S}$ it is natural to expect that such a probability becomes finite on a timescale of order $1/p$. This is indeed what happens for ferromagnetism, e.g., in the Curie-Weiss model at low temperatures where two metastable states are present. The following discussion shows instead that *the connection does not hold in presence of an exponential number of metastable states, i.e., a finite configurational entropy*. In order to see this, it is convenient to introduce the
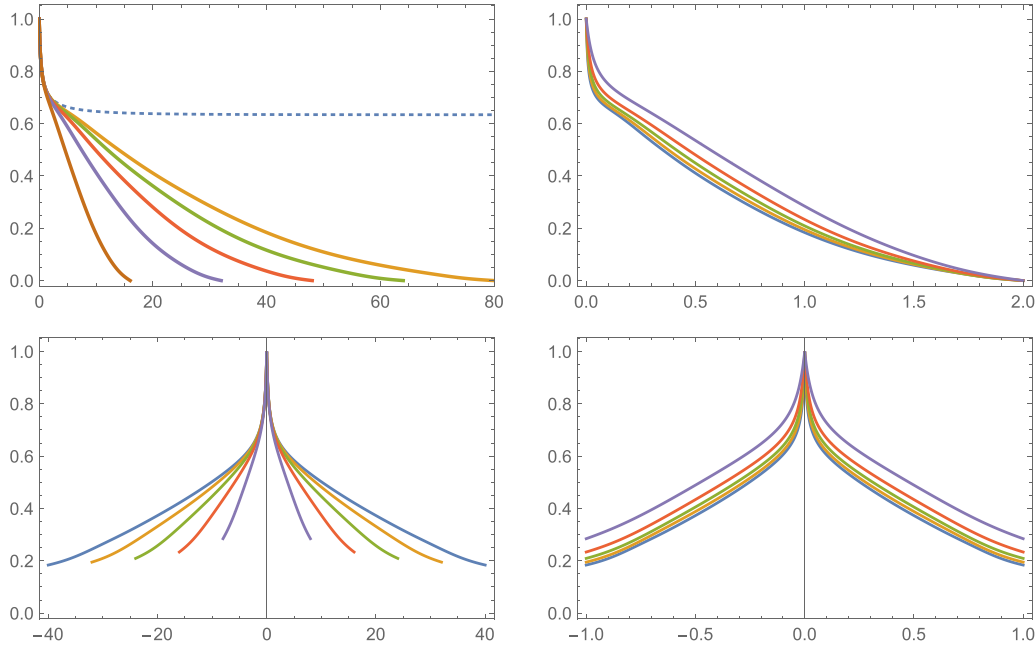
FIG. 4. Spherical 3-SG at $T = 1/1.695 < T_d$. (Top, left) $C(-\mathcal{T}, t)$ vs $t + \mathcal{T}$ for $\mathcal{T} = 8, 16, 24, 32, 40$, the dashed line is the equilibrium dynamics that is trapped inside a state and has a plateau at 0.6340. right: same vs $t/\mathcal{T}$. (Bottom, left) $C(0, t)$ vs t for $\mathcal{T} = 8, 16, 24, 32, 40$, right: same vs $t/\mathcal{T}$. Each curve was obtained from finite N_t solutions through Richardson extrapolation to $\Delta t = 0$ (see text).

average of the rate at inverse temperature β over initial and final configurations that are in equilibrium at different inverse temperatures β_{in} and β_{fin}:

$$[\ln T]_{\beta_{\mathrm{in}}, \beta, \beta_{\mathrm{fin}}} \equiv \sum_{\sigma, \tau} P_{\mathrm{eq}}^{(\beta_{\mathrm{fin}})}(\sigma) P_{\mathrm{eq}}^{(\beta_{\mathrm{in}})}(\tau) \ln T_{2\mathcal{T}}^{(\beta)}(\sigma | \tau), \quad (20)$$

where $P_{\mathrm{eq}}^{(\beta)}(\sigma)$ is the Boltzmann distribution at inverse temperature β and $T_{2\mathcal{T}}^{(\beta)}(\sigma | \tau)$ is the transition rate from configuration τ to σ due to Langevin dynamics at inverse temperature β. The quantity that generalizes $\mathcal{A}$ as given in Eq. (4) is then

$$\mathcal{A}(\beta' | \beta) \equiv \lim_{\mathcal{T} \to \infty} \lim_{N \to \infty} \frac{1}{N} ([\overline{\ln T_{2\mathcal{T}}}]_{\beta, \beta, \beta'} + S(\beta')). \quad (21)$$

An expression of the above quantity can be easily obtained in terms of second-order dynamical equations as a straightforward generalization of the case where both the initial and final configurations are at equilibrium, see Appendix E. One can thus compute the total probability of jumping from an equilibrium state to a state with a different energy $E' \neq E_{\mathrm{eq}}$, as

$$p(E') \propto \exp N \mathcal{A}(\beta' | \beta), \quad (22)$$

where naturally β' is such that E' is the equilibrium energy at the (inverse) temperature β'. An explicit computation (see Appendix E) shows that

$$\left. \frac{dp(E')}{dE'} \right|_{E' = E_{\mathrm{eq}}} = 0 \quad \text{for } T \geqslant T_d,$$

$$\left. \frac{dp(E')}{dE'} \right|_{E' = E_{\mathrm{eq}}} > 0 \quad \text{for } T < T_d.$$

The behavior for $T > T_d$ implies that the probability $p(E')$ has a maximum for $E' = E_{\mathrm{eq}}$, this was to be expected due to the

fact that $p(E_{\mathrm{eq}}) = 1$ [see Eq. (5)] saturates the bound $p(E') \leqslant 1$. Instead, for $T < T_d$, we have $p(E') \gg p$ for $E' \geqslant E_{\mathrm{eq}}$ and the system has a finite probability to jump to one of these higher energy states on a timescale $1/p(E')$ exponentially smaller than $1/p(E_{\mathrm{eq}})$:

$$1/p(E') \ll 1/p(E_{\mathrm{eq}}) \quad \text{for } T < T_d, E' > E_{\mathrm{eq}}. \quad (23)$$

On the other hand, the transition rate from a configuration with energy E' back to a configuration with energy E obeys the detailed balance condition:

$$e^{-\beta H(\tau)} T_{\Delta t}(\tau' | \tau) = e^{-\beta H(\tau')} T_{\Delta t}(\tau | \tau'), \quad (24)$$

taking into account that there are S configurations with energy E and S' configurations with energy E' we find that the probability to jump from a state with energy E' *back* to an equilibrium state is

$$p(E') e^{\Delta S - \beta \Delta E} \gg p(E'), \quad (25)$$

i.e., it is exponentially larger than $p(E')$ given that, by definition, the free energy $-S + \beta E$ has its minimum on the equilibrium states. Therefore, after the system has jumped to a state with higher energy $E' > E_{\mathrm{eq}}$, it will jump *back* to a *generic* equilibrium state on a scale exponentially smaller than $1/p(E')$, i.e., instantaneously on that scale. *This implies that an intermediate jump to one of the exponentially many metastable states with $E' > E_{\mathrm{eq}}$ provides a more efficient path for restoring ergodicity than a direct jump to another equilibrium state and thus the ergodic scale is smaller than $1/p(E_{\mathrm{eq}})$, at variance with ferromagnetism where there are only two metastable states.* It is to be expected that $p(E')$ reaches a maximum $p(E_{\mathrm{max}}) \ll 1$ at some $E_{\mathrm{max}} > E_{\mathrm{eq}}$ for $T < T_d$. The true ergodic time τ_{erg} should be identified with $1/p(E_{\mathrm{max}})$ but could be even smaller, a detailed analysis of $p(E')$ and of the correct τ_{erg} is left for future work.

4. Reproducibility

The full commented code to solve numerically the equations is provided online together with the required initialization files, see the ancillary files section of Ref. [46].

II. ANALYSIS OF THE SADDLE-POINT EQUATIONS OF THE SPHERICAL *p*-SPIN-GLASS MODEL

A. The order parameter and its meaning

In Appendices A and B, I will derive the expression for the logarithm of the rate $\mathcal{A}$ associated to Langevin dynamics of the spherical *p*-SG model in terms on an order parameter determined through saddle-point equations. The order parameter is a couple of 2×2 matrices of functions of two times t and t' on the square $-\mathcal{T} \leqslant t, t' \leqslant \mathcal{T}$:

$$C \equiv \begin{pmatrix} C(t,t') & \hat{R}_2(t,t') \\ \hat{R}_1(t,t') & \hat{X}(t,t') \end{pmatrix}, \tag{26}$$

$$C^{dt} \equiv \begin{pmatrix} C^{dt}(t,t') & \hat{R}_2^{dt}(t,t') \\ \hat{R}_1^{dt}(t,t') & \hat{X}^{dt}(t,t') \end{pmatrix}. \tag{27}$$

The physical meaning of $C(t,t')$ and $C^{dt}(t,t')$ has been discussed already: $C(t,t')$ is the correlation on the same trajectory [therefore $C(t,t) = 1$ at all times in the spherical and Ising model], while $C^{dt}(t,t')$ measures the correlations between the configurations visited by *different trajectories*. Since by definition all trajectories have the same initial and final conditions, we have

$$C^{dt}(\pm\mathcal{T}, t) = C(\pm\mathcal{T}, t). \tag{28}$$

The correlations are obviously symmetric with respect to $(t,t') \to (t',t)$:

$$C(t,t') = C(t',t), \tag{29}$$

$$C^{dt}(t,t') = C^{dt}(t',t). \tag{30}$$

Furthermore given that the measure over the trajectories is invariant under time reversal we have an additional symmetry with respect to the exchange of the initial and final configurations. Given that $t_{\text{in}} = -t_{\text{fin}}$, this symmetry translates into

$$C(t,t') = C(-t,-t'), \tag{31}$$

$$C^{dt}(t,t') = C^{dt}(-t,-t'). \tag{32}$$

For the $\hat{R}$ components of the order parameter, we have

$$\hat{R}_1(t,t') = \overline{[\langle \hat{x}_i(t) s_i(t') \rangle]}, \tag{33}$$

$$\hat{R}_1^{dt}(t,t') = \overline{[\langle \hat{x}_i(t) \rangle \langle s_i(t') \rangle]}, \tag{34}$$

where $\hat{x}$ is an auxiliary variable of the dynamics (see the Appendix). They translate into

$$\hat{R}_1(t,t') = \overline{\left[\frac{\delta \langle s_i(t') \rangle}{\beta \, \delta h(t)} \right]} - \frac{1}{2} \frac{d}{dt} C(t,t'), \tag{35}$$

$$\hat{R}_1^{dt}(t,t') = \overline{\left[\frac{\delta \ln D(\sigma,\tau)}{\beta \, \delta h(t)} \langle s_i(t') \rangle \right]} - \frac{1}{2} \frac{d}{dt} C^{dt}(t,t'). \tag{36}$$

Thus $\hat{R}_1(t,t')$ is connected to the response of the time average of the spin over trajectories to a field $h(t)$. The functions

$\hat{R}_2(t,t')$ and $\hat{R}_2^{dt}(t,t')$ are equal to the left-hand sides of the above equations with the exchange $t \leftrightarrow t'$. Thus while *neither* functions are symmetric with respect to $t \leftrightarrow t'$, they are related through two functions $\hat{R}(t,t')$ and $\hat{R}^{dt}(t,t')$ such that

$$\hat{R}_1(t,t') = \hat{R}_2(t',t) = \hat{R}(t',t), \tag{37}$$

$$\hat{R}_1^{dt}(t,t') = \hat{R}_2^{dt}(t',t) = \hat{R}^{dt}(t',t), \tag{38}$$

which implies that the matrices (26) and (27) are symmetric. On the other hand, time-reversal invariance implies that $\hat{R}_1(t,t'), \hat{R}_2(t,t'), \hat{R}_1^{dt}(t,t'), \hat{R}_2^{dt}(t,t')$, are symmetric with respect to $(t,t') \to (-t,-t')$ (because $t_{\text{in}} = -t_{\text{fin}}$ as above). For the $\hat{X}$ components, we have

$$\hat{X}(t,t') = \overline{[\langle \hat{x}_i(t) \hat{x}_i(t') \rangle]}, \tag{39}$$

$$\hat{X}^{dt}(t,t') = \overline{[\langle \hat{x}_i(t) \rangle \langle \hat{x}_i(t') \rangle]}. \tag{40}$$

The physical meaning is also associated to particular responses that takes a simple form in the ergodic phase, see Sec. II D. The above formulas imply that both $\hat{X}(t,t')$ and $\hat{X}^{dt}(t,t')$ are symmetric with respect to $t \leftrightarrow t'$ and to time reversal $(t,t') \to (-t,-t')$.

B. The saddle-point equations

The saddle-point equations can be written in a compact form which is also suitable for numerical integration considering the space of 2×2 matrices whose components are functions of two times t and t' on the square $-\mathcal{T} \leqslant t, t' \leqslant \mathcal{T}$. The generic element of this space can be written as

$$A \equiv \begin{pmatrix} C_A(t,t') & \hat{R}_{2,A}(t,t') \\ \hat{R}_{1,A}(t,t') & \hat{X}_A(t,t') \end{pmatrix}. \tag{41}$$

Given two elements A and B in the above space we have a natural definition of the product that generalizes the matrix product (it corresponds to exactly to ordinary matrix products if times are discretized). For a real function $B(x)$, we also define the elementwise function $B[A]$ such that

$$C_{B[A]}(t,t') \equiv B(C_A(t,t')),$$
$$\hat{R}_{2,B[A]}(t,t') \equiv B'(C_A(t,t')) \hat{R}_{2,A}(t,t'),$$
$$\hat{R}_{1,B[A]}(t,t') \equiv B'(C_A(t,t')) \hat{R}_{1,A}(t,t'),$$
$$\hat{X}_{B[A]}(t,t') \equiv B'(C_A(t,t')) \hat{X}_A(t,t')$$
$$\qquad + B''(C_A(t,t')) \hat{R}_{1,A}(t,t') \hat{R}_{2,A}(t,t').$$

We define also

$$M \equiv \begin{pmatrix} -\frac{1}{2\Gamma_0} \delta''(t,t') + \hat{\mu}(t)\delta(t,t') & \mu(t)\delta(t,t') \\ \mu(t)\delta(t,t') & -\frac{2}{\Gamma_0}\delta(t,t') \end{pmatrix} \tag{42}$$

and

$$T \equiv \begin{pmatrix} 0 & \delta(t,t') \\ \delta(t,t') & o \end{pmatrix}.$$

Note that due to the spherical constraint the operator M depends on two additional quantities $\mu(t)$ and $\hat{\mu}(t)$, see Appendix A 3. In order to write down the saddle-point equations it is useful to introduce two additional objects Λ and Λ^{dt} that

are also 2×2 matrices of two-time functions. Another useful quantity is

$$\delta_{\mp} \equiv \begin{pmatrix} \hat{R}_{\Lambda,1}(t,\mp\mathcal{T})C(\mp\mathcal{T},t') & \hat{R}_{\Lambda,1}(t,\mp\mathcal{T})\hat{R}_2(\mp\mathcal{T},t') \\ C_{\Lambda}(t,\mp\mathcal{T})C(\mp\mathcal{T},t') & C_{\Lambda}(t,\mp\mathcal{T})\hat{R}_2(\mp\mathcal{T},t') \end{pmatrix}.$$
(43)

With the above definitions the saddle-point equations of the spherical model derived in the Appendix take the following compact expressions:

$$\Lambda = -\frac{\beta^2}{2}f'[C], \quad \Lambda^{dt} = -\frac{\beta^2}{2}f'[C^{dt}], \quad (44)$$

$$MC + T\,\Lambda\,T\,C + (n-1)\,T\,\Lambda^{dt}\,T\,C^{dt} + \delta_- + \delta_+ = I, \quad (45)$$

$$MC^{dt} + T\,\Lambda^{dt}\,T\,C + T\,\Lambda\,T\,C^{dt} + (n-2)\,T\,\Lambda^{dt}\,T\,C^{dt}$$
$$+\delta_- + \delta_+ = 0. \quad (46)$$

The quantities $\mu(t)$ and $\hat{\mu}(t)$ are unknown and must be determined self-consistently imposing the conditions:

$$C(t,t) = 1, \quad \hat{R}(t,t) = \frac{1}{2} \quad \forall t. \quad (47)$$

Due to the presence of the operator M that contains second-order derivatives the equations must be also supplemented with the following *boundary conditions*:

$$C(\pm\mathcal{T},t') = C(t',\pm\mathcal{T}), \quad (48)$$

$$\hat{R}_2(\pm\mathcal{T},t') = \hat{R}_1(t',\pm\mathcal{T}), \quad (49)$$

$$C^{dt}(\pm\mathcal{T},t') = C^{dt}(t',\pm\mathcal{T}), \quad (50)$$

$$\hat{R}_2^{dt}(\pm\mathcal{T},t') = \hat{R}_1^{dt}(t',\pm\mathcal{T}). \quad (51)$$

Additional boundary conditions follow from the properties of the initial and final configurations:

$$C(\pm\mathcal{T},\pm\mathcal{T}) = 1, \quad C(\pm\mathcal{T},\mp\mathcal{T}) = 0,$$

$$C^{dt}(\pm\mathcal{T},\pm\mathcal{T}) = 1, \quad C^{dt}(\pm\mathcal{T},\mp\mathcal{T}) = 0.$$

The above boundary conditions are sufficient to compute the r.h.s. of the saddle-point equations (45) and (46) for a generic C and C^{dt} but given the structure of the equation and the physical meaning of the order parameter the actual solution verifies the additional symmetries discussed previously. In particular, the symmetry of the problem under the exchange between the initial and final configuration implies:

$$\mu(t) = \mu(-t), \quad \hat{\mu}(t) = \hat{\mu}(-t). \quad (52)$$

In the numerical analysis times are discretized, thus C and C^{dt} become actual matrices and the above symmetries allow a fourfold reduction of the memory required to store them.

C. The free case

In the infinite temperature limit the interaction vanishes and the system performs a free Brownian motion on the $N-1$ dimensional sphere. In this case, one can find the analytic solution of the saddle-point equations which, in turn, is useful to guess an initial solution to feed to the Newton's algorithm at finite temperature. Besides, as we discussed in the

introduction, it allows to bypass the nonrenormalizability of the expression for $\overline{[\ln \hat{T}]}$ by integrating with respect to the temperature. The saddle-point equations (45) and (46) become in the free case:

$$MC = I, \quad MC^{dt} = 0. \quad (53)$$

The first equation corresponds to

$$-2\hat{R}_1(t,t') + \mu(t)C(t,t') = 0,$$

$$-\frac{1}{2}\frac{d^2}{dt^2}C(t,t') + \mu(t)\hat{R}_1(t,t') + \hat{\mu}(t)C(t,t') = \delta(t-t')$$

$$-2\hat{X}(t,t') + \mu(t)\hat{R}_2(t,t') = \delta(t-t'),$$

$$-\frac{1}{2}\frac{d^2}{dt^2}\hat{R}_2(t,t') + \mu(t)\hat{X}(t,t') + \hat{\mu}(t)\hat{R}_2(t,t') = 0.$$

Note that the equations for C do not depend on C^{dt}. To solve them one can start noticing that the conditions (47) plugged into the first equation lead to

$$\mu(t) = 1 \quad \forall t. \quad (54)$$

Then the first equation and the symmetries $\hat{R}_2(t,t') = \hat{R}_1(t',t), C(t,t') = C(t',t)$ and the third equation lead to

$$\hat{R}_1(t,t') = \frac{1}{2}C(t,t'), \quad (55)$$

$$\hat{R}_2(t,t') = \frac{1}{2}C(t,t'), \quad (56)$$

$$\hat{X}(t,t') = -\frac{1}{2}\delta(t-t') + \frac{1}{4}C(t,t'). \quad (57)$$

As a consequence the correlation $C(t,t')$ is determined by the following equation:

$$-\frac{1}{2}\frac{d^2}{dt^2}C(t,t') + \left(\frac{1}{2} + \hat{\mu}(t)\right)C(t,t') = \delta(t-t').$$

To be solved with boundary conditions $C(\pm\mathcal{T},t) = C(t,\pm\mathcal{T})$ and $C(\pm\mathcal{T},\pm\mathcal{T}) = 1, C(\pm\mathcal{T},\mp\mathcal{T}) = 0$. The equation has the solution

$$C(t,t') = C(|t-t'|), \quad \hat{\mu}(t) = \hat{\mu}. \quad (58)$$

To determine the function $C(x)$ it is convenient to define a quantity $-\infty < a < 1$ implicitly through the equation:

$$\mathcal{T} = \frac{1}{2\sqrt{a}}\operatorname{arctanh}\sqrt{a}. \quad (59)$$

We then have, see Fig. 5:

$$\hat{\mu} = \frac{a-1}{2}, \quad (60)$$

$$C(x) = \cosh\sqrt{a}x - \frac{1}{\sqrt{a}}\sinh\sqrt{a}x. \quad (61)$$

Note that we also have $C(0) = 1$ and $\dot{C}(0) = -1$ and these are general results that hold also at finite temperature being a consequence of the presence of the delta functions and of the symmetry $C(t,t') = C(t',t)$. Let us discuss the limits of large and small $\mathcal{T}$. For large values of $\mathcal{T}$, we have

$$a \approx 1 - 4e^{-4\mathcal{T}}, \quad \hat{\mu} \approx -2e^{-4\mathcal{T}} \text{ for } \mathcal{T} \to \infty. \quad (62)$$

Note that in the large $\mathcal{T}$ limit $\hat{\mu}$ tends to zero and the function $C(x)$ tends to the equilibrium solution $C_{\text{eq}}(x) = e^{-x}$. On the
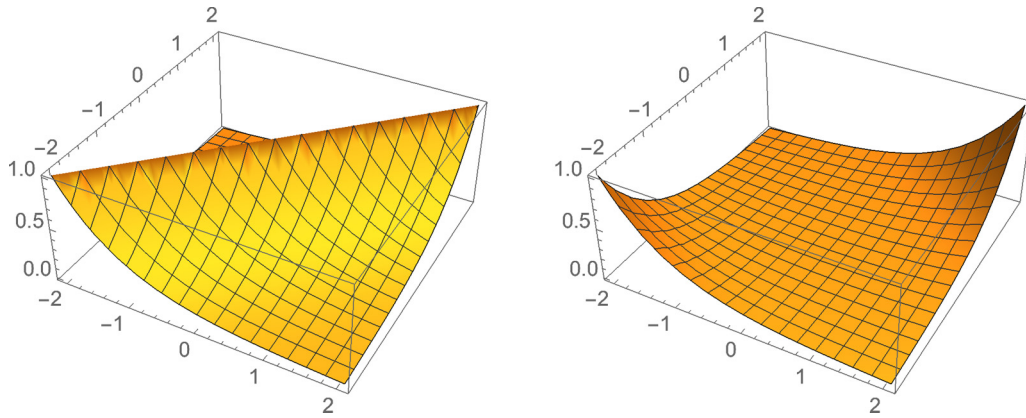
FIG. 5. (Left) The function $C(t, t')$ in the free case for $a = 0.999$ corresponding to $\mathcal{T} = 2.07442$ (see text). (Right) The function $C^{dt}(t, t')$.

other hand a becomes negative for $\mathcal{T} < 1/2$:

$$a \approx 6\left(\mathcal{T} - \frac{1}{2}\right) \quad \text{for} \quad \mathcal{T} \approx \frac{1}{2} \tag{63}$$

therefore $a^{1/2}$ is imaginary and the hyperbolic functions in (61) become ordinary trigonometric functions:

$$C(x) = \cos\sqrt{|a|}x - \frac{1}{\sqrt{|a|}} \sin\sqrt{|a|}x. \tag{64}$$

In the small $\mathcal{T}$ limit, a tends to minus infinity as

$$a \approx -\frac{\pi^2}{16\mathcal{T}^2}, \quad \hat{\mu} \approx -\frac{\pi^2}{32\mathcal{T}^2} \quad \text{for} \quad \mathcal{T} \to 0 \tag{65}$$

leading to

$$C(x) \approx \cos\frac{\pi x}{4\mathcal{T}} \text{ for } \mathcal{T} \to 0. \tag{66}$$

The above expression correctly vanishes at $x = 2\mathcal{T}$ but does not display the correct behavior $C(x) \approx 1 - x$ at small x. Indeed, it is only valid for $x = O(\mathcal{T})$ and the correct linear behavior is recovered for $x \ll \mathcal{T}$. We now turn to the equations for C^{dt}:

$$-2\hat{R}_1(t, t') + \mu(t)C^{dt}(t, t') = 0,$$

$$-\frac{1}{2}\frac{d^2}{dt^2}C^{dt}(t, t') + \mu(t)\hat{R}_1^{dt}(t, t') + \hat{\mu}(t)C^{dt}(t, t') = 0,$$

$$-2\hat{X}^{dt}(t, t') + \mu(t)\hat{R}_2^{dt}(t, t') = 0,$$

$$-\frac{1}{2}\frac{d^2}{dt^2}\hat{R}_2^{dt}(t, t') + \mu(t)\hat{X}^{dt}(t, t') + \hat{\mu}(t)\hat{R}_2^{dt}(t, t') = 0.$$

The condition $\mu(t) = 1$ derived earlier, the first equation, the symmetries $\hat{R}_2^{dt}(t, t') = \hat{R}_1^{dt}(t', t)$, $C^{dt}(t, t') = C^{dt}(t', t)$ and the third equation lead to

$$\hat{R}_1^{dt}(t, t') = \frac{1}{2}C^{dt}(t, t'), \tag{67}$$

$$\hat{R}_2^{dt}(t, t') = \frac{1}{2}C^{dt}(t, t'), \tag{68}$$

$$\hat{X}^{dt}(t, t') = \frac{1}{4}C^{dt}(t, t'), \tag{69}$$

and $C^{dt}(t, t')$ obeys the equation

$$-\frac{1}{2}\frac{d^2}{dt^2}C^{dt}(t, t') + \left(\frac{1}{2} + \hat{\mu}\right)C^{dt}(t, t') = 0. \tag{70}$$

To be solved with boundary conditions $C^{dt}(\pm\mathcal{T}, t) = C^{dt}(t, \pm\mathcal{T})$ and $C^{dt}(\pm\mathcal{T}, \pm\mathcal{T}) = 1$, $C^{dt}(\pm\mathcal{T}, \mp\mathcal{T}) = 0$. The solution reads

$$C^{dt}(t, t') = \frac{\sqrt{1-a}}{a}\cosh(\sqrt{a}(t + t'))$$
$$- \frac{1-a}{a}\cosh(\sqrt{a}(t - t')), \text{ for } \mathcal{T} > 1/2;$$

$$C^{dt}(t, t') = \frac{\sqrt{1-a}}{a}\cos(\sqrt{|a|}(t + t'))$$
$$- \frac{1-a}{a}\cos(\sqrt{|a|}(t - t')) \text{ for } 0 < \mathcal{T} < 1/2.$$

In the limit $\mathcal{T} \to \infty$, we have $C^{dt}(t, t') = 0$ for any finite t, t'. Close to the initial and finite time, we have instead

$$C^{dt}(\pm\mathcal{T} + \delta t, \pm\mathcal{T} + \delta t') = C_{\text{eq}}(|\delta t + \delta t'|) \text{ for } \mathcal{T} \to \infty.$$

In the limit $\mathcal{T} \to 0$, we obtain instead

$$C^{dt}(t, t') \approx \cos\frac{\pi(t - t')}{4\mathcal{T}} \text{ for } \mathcal{T} \to 0.$$

Note that in this limit, we have also $C^{dt}(t, t') \approx C(t, t')$ meaning that all trajectories tend to follow the same path.

The computation of the transition rate requires to treat carefully the divergences of the path integral and is given in the Appendix. The final result is

$$\frac{1}{N}\overline{[\ln\hat{T}]} = -\frac{1}{2}\ln 2\pi - \frac{1}{2} + (1 + \hat{\mu})\mathcal{T} + \frac{1}{4}\ln\frac{-\hat{\mu}}{2}.$$

From the above expression, we see that the logarithm of the rate tends to minus infinity as $\mathcal{T}$ goes to zero:

$$\frac{1}{N}\overline{[\ln\hat{T}]} \approx -\frac{\pi^2}{32\mathcal{T}}, \quad \mathcal{T} \approx 0.$$

For $\mathcal{T}$ going to infinity one can see that although $\hat{\mu} \approx -2e^{-4\mathcal{T}}$ the $O(\mathcal{T})$ divergences in the last two terms cancel and the rate has a finite limit. This limit is exactly equal to minus the entropy $S(\beta)$ of the spherical model at infinite temperature ($\beta = 0$), i.e., the logarithm of the surface of the N-dimensional sphere of radius $\sqrt{N}$:

$$\lim_{\mathcal{T} \to \infty} \frac{1}{N}\overline{[\ln\hat{T}]} = -\frac{1}{2}\ln 2\pi - \frac{1}{2} = -S(0).$$

This is expected in the $\mathcal{T} \to \infty$ limit at any finite N and implies that the two limits commute. One can show that the approach to the $\mathcal{T} \to \infty$ limit is exponential:

$$\frac{1}{N}\overline{[\ln \hat{T}]} \approx -S(0) - \frac{1}{2}e^{-4\mathcal{T}}, \quad \mathcal{T} \gg 1.$$

D. The ergodic phase

As explained in the introduction the ergodic phase defined by $T > T_d$ is characterized by the fact that the $\mathcal{T} \to \infty$ limit commutes with the $N \to \infty$ limit. In order to discuss the solutions in this regime, it is convenient to analyze first the ergodic limit at finite N. While the present formalism is fully invariant under time reversal to discuss this limit it is convenient to reintroduce the arrow of time. In the ergodic limit $\mathcal{T} \to \infty$, dynamics looses any dependence on the initial configuration and the transition rate obeys

$$\lim_{\mathcal{T} \to \infty} T_{2\mathcal{T}}(\sigma|\tau) = \frac{e^{-\beta H(\sigma)}}{Z} \quad (71)$$

that, according to the previous definitions leads to

$$\lim_{\mathcal{T} \to \infty} \hat{T}_{2\mathcal{T}}(\sigma, \tau) = \frac{1}{Z} e^{-\frac{\beta}{2}H(\sigma) - \frac{\beta}{2}H(\tau)} \quad (72)$$

and the following exact result:

$$\lim_{\mathcal{T} \to \infty} [\ln \hat{T}_{2\mathcal{T}}(\sigma, \tau)] = -\beta E - \ln Z = -S. \quad (73)$$

In the ergodic limit, we expect that trajectories close to the initial condition at time $-\mathcal{T}$ are not influenced by the fact that we are fixing the final configuration at time $\mathcal{T}$. This implies that they are typical trajectories and since the initial configuration is weighted with the equilibrium weight we expect that correlations and response are those valid at equilibrium and in particular satisfy time-translational invariance (TTI) and fluctuation-dissipation theorem (FDT). To see the implications on the functions it is convenient to start from the expressions of the six functions as averages of $s_i(t)$ and $\hat{x}_i(t)$ obtained in Sec. II A. It is then convenient to transform back to the variable $\hat{s}$ introduced in the first steps of the integral representation (see Appendix A 2) by writing

$$\hat{x}_i = \hat{s}_i(t) - \dot{s}_i(t). \quad (74)$$

It is well known the equilibrium averages of $\hat{s}_i(t)$ are associated to responses to a field at time t and since according to (71) the distribution at σ is completely independent of what happens at any finite time $t \ll \mathcal{T}$, we have for these times

$$\langle \hat{s}(t) \rangle = 0, \quad \langle \hat{s}(t)\hat{s}(t') \rangle = 0. \quad (75)$$

The above relationships plus FDT allow to derive the following expressions in the ergodic limit:

$$C(t, t') = C_{\text{eq}}(|t - t'|), \quad (76)$$

$$\hat{R}_2(t, t') = \hat{R}_1(t', t) = \text{sign}(t - t')\frac{1}{2}\frac{d}{dt'}C(t, t'), \quad (77)$$

$$\hat{X}(t, t') = -\frac{1}{4}\frac{d^2}{dt\,dt'}C(t, t'). \quad (78)$$

Note that $\hat{R}(t, t')$ is symmetric (which is not true in the activated phase) and it is equal to $1/2$ on the diagonal according to

the saddle-point equations. The expression for $\hat{X}(t, t')$ is valid also for $t = t'$ and thus it may be integrated, this is consistent with the fact that $\hat{X}(t, t')$ has a term proportional $-\delta(t - t')/2$ on the diagonal. The above properties have been derived for $t, t' \ll \mathcal{T}$, in order to study the region of times close to $\mathcal{T}$ it is better to reverse the arrow of time, following the same arguments we obtain the above equations with $(t, t') \to (-t, -t')$ and since the functions are symmetric with respect to this transformation we conclude that they are valid at all times.

At large values of $\mathcal{T}$, the correlation between different trajectories are different from zero only for times t, t' that are both close to either $-\mathcal{T}$ or $\mathcal{T}$. To determine them, we can use the following relationship :

$$\sum_\tau \frac{e^{-\beta H(\tau)}}{Z}T_{\Delta t}(\tau'|\tau)T_{\Delta t'}(\tau''|\tau) = \frac{e^{-\beta H(\tau')}}{Z}T_{\Delta t + \Delta t'}(\tau''|\tau') \quad (79)$$

that follows from the detailed balance condition:

$$e^{-\beta H(\tau)}T_{\Delta t}(\tau'|\tau) = e^{-\beta H(\tau')}T_{\Delta t}(\tau|\tau') \quad (80)$$

and from the general property:

$$\sum_\tau T_{\Delta t'}(\tau''|\tau)\,T_{\Delta t}(\tau|\tau') = T_{\Delta t + \Delta t'}(\tau''|\tau'). \quad (81)$$

Furthermore responses between different trajectories vanish because of $\langle \hat{s}(t) \rangle = 0$. In particular if we write times as $t = \mp\mathcal{T} \pm \Delta t$ ($\Delta t \geqslant 0$), we have

$$C^{dt}(t, t') = C_{\text{eq}}(\Delta t + \Delta t'), \quad (82)$$

$$\hat{R}_2^{dt}(t, t') = \hat{R}_1^{dt}(t', t) = \mp\frac{1}{2}\frac{d}{dt'}C^{dt}(t, t'), \quad (83)$$

$$\hat{X}^{dt}(t, t') = \frac{1}{4}\frac{d^2}{dt\,dt'}C^{dt}(t, t'). \quad (84)$$

Note that the expression for $\hat{R}_2^{dt}(t, t')$ changes sign depending on weather we are close to $\mp\mathcal{T}$, indeed the second case $(+\mathcal{T})$ is obtained from the transformation $(t, t') \to (-t, -t')$ applied to the first case $(-\mathcal{T})$. A solution of the saddle-point equations with the above structure can be found only if $C_{\text{eq}}(\infty) = 0$, thus the ergodic solution does not exist below T_d where $C_{\text{eq}}(\infty) = q$ in the thermodynamic limit and the condition $C(\pm\mathcal{T}, \mp\mathcal{T}) = 0$ cannot be fulfilled. The above expressions are valid for both Ising and spherical models, in addition for the spherical model, we have

$$\mu(t) = \mu_{\text{eq}}, \quad \hat{\mu}(t) = 0. \quad (85)$$

Note that all components are parameterized by the equilibrium correlation $C_{\text{eq}}(t)$, let us see how an equation for $C_{\text{eq}}(t)$ can be derived in the present context. In the formulation of the problem discussed in the Appendix, one works with objects $Q(ab)$. In the ergodic phase, one can additionally introduce a special class of objects $A(ab)$ whose components $A(ab)$ and $A^{dt}(ab)$ obey formulas (76)–(78) and (82)–(84) with $C_{\text{eq}}(t)$ replaced by some generic function $C_A(t)$. In the following, this kind of objects will be loosely referred as having an "equilibrium" structure. A simple computation shows that the equilibrium structure is preserved by the application of a function, i.e.,

$f(\boldsymbol{A}(\boldsymbol{ab}))$ has also the equilibrium structure with:

$$f(\boldsymbol{A}(\boldsymbol{ab})) \to C_{f[A]}(t) = f(C_A(t)). \tag{86}$$

Similarly, given another equilibrium object $\boldsymbol{B}(\boldsymbol{ab})$ parameterized by a function $C_B(t)$, one can show through a tedious computation that the product has also the equilibrium structure with

$$\int \boldsymbol{A}(\boldsymbol{ac})\boldsymbol{B}(\boldsymbol{cb})\boldsymbol{dc} \to C_{AB}(t),$$

$$C_{AB}(t) = C_A(0)C_B(t) - \int_0^t C_A(t-s)\frac{dC_B(s)}{ds}ds.$$

The above equation holds for $C_A(\infty) = C_B(\infty) = 0$ which is granted by the fact that we are working in zero field and $f'(0) = 0$ so that the overlap between different equilibrium states is zero. The last two equations applied to the saddle-point equation (B4) lead immediately to the equilibrium equation:

$$\dot{C}(t) = -C(t) - \frac{\beta^2}{2}\int_0^t f'(C(t-s))\dot{C}(s)\,ds, \tag{87}$$

where we have used $\mu = 1 + \frac{\beta^2}{2}f'(1)$ that follows from the condition $C(0) = -\dot{C}(0) = 1$. The above equations is usually written as [39,40,47]

$$\dot{C}(t) = -C(t) + \frac{\beta^2}{2}f'(C(t))(1 - C(t))$$
$$- \frac{\beta^2}{2}\int_0^t (f'(C(t-s)) - f'(C(t)))\frac{dC}{ds}\,ds, \tag{88}$$

where $C(\infty)$ is the solution of the equation

$$C = \frac{\beta^2}{2}f'(C)(1 - C). \tag{89}$$

This equation admits the solution $C = 0$ at all temperature but develops an additional nonzero solution at T_d specified by the condition:

$$1 = \frac{\beta_d^2}{2}(f''(C_d)(1 - C_d) - f'(C_d)). \tag{90}$$

In the pure p-spin models $f(x) = x^p$, we have

$$T_d = \sqrt{\frac{p(p-2)^{p-2}}{2(p-1)^{p-1}}}, \quad C_d(\infty) = \frac{p-2}{p-1}. \tag{91}$$

To conclude the discussion of the ergodic phase, we have to show that for $T > T_d$, the $\mathcal{T} \to \infty$ and $N \to \infty$ limits commute and we have

$$\mathcal{A} = 0 \quad \text{for } T_d < T < \infty. \tag{92}$$

In the previous section, we have shown that the relationship is satisfied for $\beta = 1/T = 0$ and thus it suffices to show that its derivative with respect to the β is also zero. In Appendix E, we show that the derivative of the rate with respect to β is $\frac{\beta}{2}\frac{d}{dn}\int d\boldsymbol{a}d\boldsymbol{b}f(\boldsymbol{Q}(\boldsymbol{a},\boldsymbol{b}))$. The above expression takes a very simple form in the ergodic phase, indeed using the above equilibrium formulas one can show that the generic formula (E11) for the integrals takes a very

simple form:

$$\int \boldsymbol{B}(\boldsymbol{ab})\boldsymbol{dadb} = (m + m')C_B(0)$$

$$+ n\left(4\int R_{2,B}(-\mathcal{T},t)dt + \int \hat{X}_B(t,t')\,dt\,dt'\right.$$

$$\left. + (n-1)\int \hat{X}_{B^{dt}}(t,t')\,dt\,dt'\right)$$

$$= (m + m' + n)C_B(0) \tag{93}$$

that leads to

$$\frac{d}{d\beta}\lim_{\mathcal{T}\to\infty}\lim_{N\to\infty}\frac{1}{N}\overline{[\ln\hat{T}]} = \frac{\beta}{2}f(1), \tag{94}$$

to be compared with

$$\frac{1}{N}\frac{dS(\beta)}{d\beta} = -\frac{\beta}{2}f(1). \tag{95}$$

We thus see that Eq. (92) is verified

$$\lim_{\mathcal{T}\to\infty}\lim_{N\to\infty}\frac{1}{N}\overline{[\ln\hat{T}]} = -S(\beta) = -S(0) + \frac{\beta^2}{4}f(1). \tag{96}$$

The above relationship is valid also for Ising systems.

E. The activated phase

In the activated phase $T_s < T < T_d$, the saddle-point equations have been solved numerically for the classic pure p-spin with $f(x) = x^3$, where $\beta_d = 1.63299$ and $\beta_s = 1.70633$ [39,40,47]. Due to the significant resources needed to solve the equations, as discussed in the next section, most of the analysis has been done at a single inverse temperature $\beta = 1.695$ for values of $\mathcal{T} = 8, 16, 24, 32, 40$. In Fig. 6, we clearly see the features anticipated in the introduction, in particular: (i) the overlap between the initial (final) configuration and the $t = 0$ intermediate configuration $C(\pm\mathcal{T}, 0)$ tends to a finite value in the $\mathcal{T} \to \infty$ limit, and (ii) the functions $\hat{R}(t, t')$, $\hat{R}^{dt}(t, t')$, $\hat{X}(t, t')$, $\hat{X}^{dt}(t, t')$ decrease with increasing $\mathcal{T}$ for $|t - t'| = O(\mathcal{T})$.

In Fig. 7, the same data are rescaled to fully demonstrate the asymptotic limit $\mathcal{T} \to \infty$ discussed in Sec. I A 2. The same-trajectory functions $\hat{R}(t, t')$ and $\hat{X}(t, t')$ deviate from this scaling in the region $|t - t'| = O(1)$ that goes to zero on the scale of the plots for $\mathcal{T} \to \infty$. In the region $|t - t'| = O(1)$ $\hat{R}(t, t')$ and $\hat{X}(t, t')$ converge to a finite limit asymptotically as shown in Fig. 8. One can also check that in that region $C(t, t')$, $\hat{R}(t, t')$, and $\hat{X}(t, t')$ satisfy the equilibrium relationships (77) and (78) consistently with the fact that the system is in equilibrium on scales much smaller than $\mathcal{T}$ in a field slowly varying on a $O(\mathcal{T})$ timescale. The auxiliary functions $\mu(t)$ and $\hat{\mu}(t)$ are shown in Fig. 9. Note that the $\mathcal{T}^{-1}$ scaling of $\hat{\mu}(t)$ is consistent with the fact that $C(t, t')$, $\hat{R}(t, t')$, and $\hat{X}(t, t')$ satisfy equilibrium relationship on the diagonal according to Eq. (85). However $\mu(t)$ changes with time implying equilibrium on finite timescales but not on the global scale $O(\mathcal{T})$. In Fig. 10, we plot the derivative of $\langle\ln\hat{T}\rangle + S$ with respect to the inverse temperature computed according to expression (E1). According to the result of the
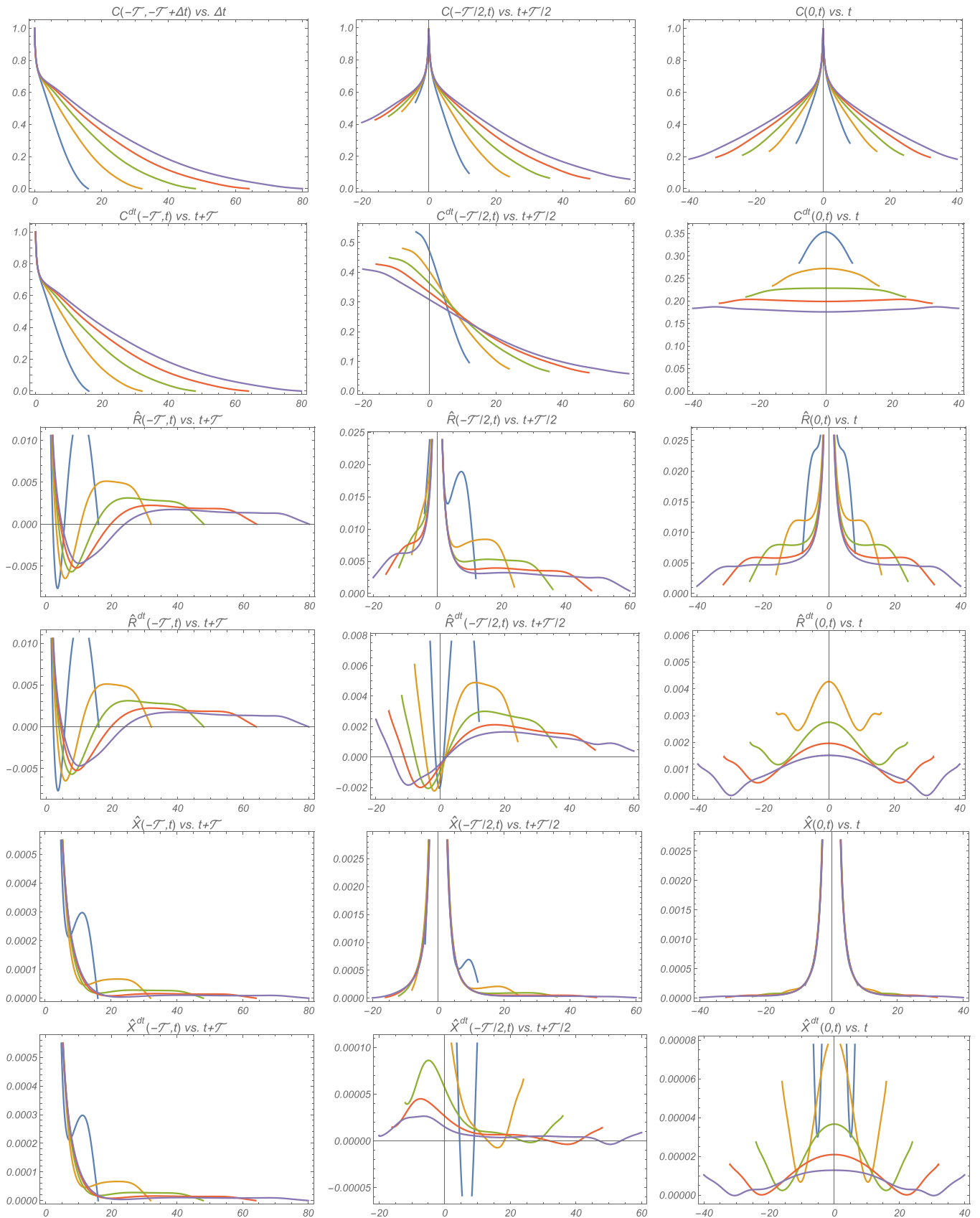
FIG. 6. Spherical p-spin glass for $\beta = 1.695$ and $p = 3$. Solutions for $\mathcal{T} = 8$ (blue), 16 (yellow), 24 (green), 32 (red), and 40 (purple) on various strips of the $-\mathcal{T} \leqslant t, t' \leqslant \mathcal{T}$ plane. The vertical scale of the $\hat{R}(t, t')$ and $\hat{X}(t, t')$ functions does not allows to see the diagonal region $t - t' = O(1)$.

previous section this quantity should go to zero in the $\mathcal{T} \to \infty$ limit for $\beta < \beta_d = 1.6329$. On the other hand, the asymptotic behavior sets in when $\mathcal{T}$ is larger than the equilibrium relaxation time that diverges as $|T - T_d|^{-\gamma}$, where $\gamma = 1.765$ is a MCT exponent [39,48,49]. Thus, at any $\mathcal{T}$, no matter how large, there is always a range of temperatures $\Delta T \propto \mathcal{T}^{-1/\gamma}$ such that for $T - T_d < \Delta T$, $\overline{\langle \ln \hat{T} \rangle} + S$ is smaller than zero. The effect decreases with increasing $\mathcal{T}$ as the figure shows but is still significant at the values of $\mathcal{T}$ that we could study.

On the other hand, the curves seems to have converged to -0.47 at $\beta = 1.695$ (as data from $\mathcal{T} = 40$ also suggest) and this result, supplemented with the information that it must be zero asymptotically at β_d, allows for a rough estimate of the integral leading to

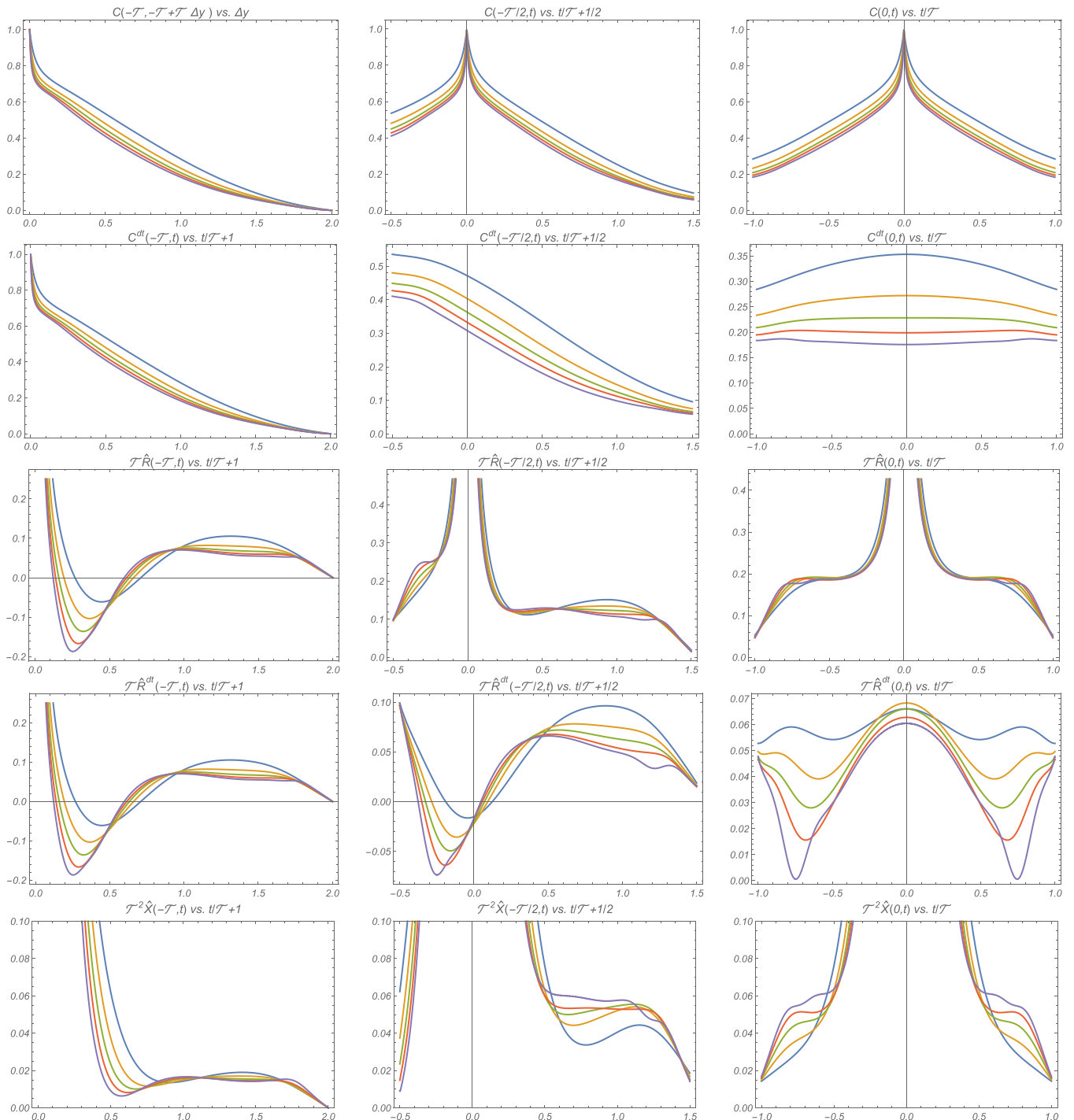$$\mathcal{A} = -0.014 \text{ for } \beta = 1.695. \tag{97}$$

FIG. 7. Spherical p-spin glass for $\beta = 1.695$ and $p = 3$. Scaled solutions for $\mathcal{T} = 8$ (blue), 16 (yellow), 24 (green), 32 (red), and 40 (purple) on various strips of the $-\mathcal{T} \leqslant t, t' \leqslant \mathcal{T}$ plane.
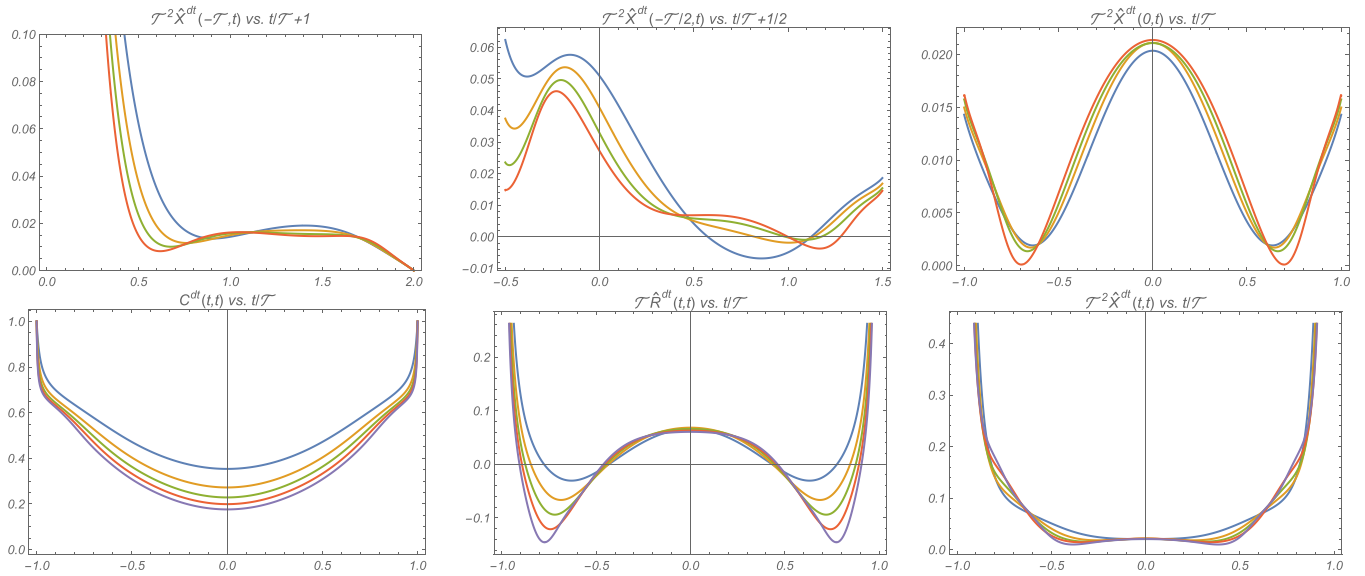
FIG. 7. *(Continued.).*

F. Numerical solution

1. Newton-Krylov methods

To solve the equations numerically time was discretized in steps $\Delta t = \mathcal{T}/N_t$ for integer N_t up to $N_t = 2000$. The six functions can then be jointly represented as two real matrices C and C^{dt} of size $2(2N_t + 1)$ so that the total number of variables is $O(32N_t^2)$. Using standard formulas for discrete integrals and second-order derivatives (see below) one obtains expressions with an $O(\Delta t^2)$ error.

It turned out that the equations can be solved by Newton's method. To initialize the algorithm, an approximate solution not too far away from the correct solution can be obtained from the analytic solutions in the free case. One can start from

$\beta = 0$ at finite $\mathcal{T}$ and switch on the temperature, then at fixed temperature, $\mathcal{T}$ can be changed changing the discretization parameter Δt at fixed N_t by small amounts. At fixed $\mathcal{T}$ and β, Δt can be reduced extrapolating the result of a coarser grid to a finer grid (larger N_t) and using it as a starting point for Newton's method at the new N_t. The main technical problem is that every iteration of Newton's method requires to invert the Jacobian of the equations, a $(32N_t^2) \times (32N_t^2)$ matrix. Even exploiting the symmetries of the problem, with current technology *exact* inversion of the Jacobian becomes unfeasible for N_t of the order 70–80 limiting the values of $\mathcal{T}$ that can be studied. N_t could instead be increased up to 2000 using an approximate method for the well-studied problem [50] of solving a very large linear system $A\,x = b$. Specifically
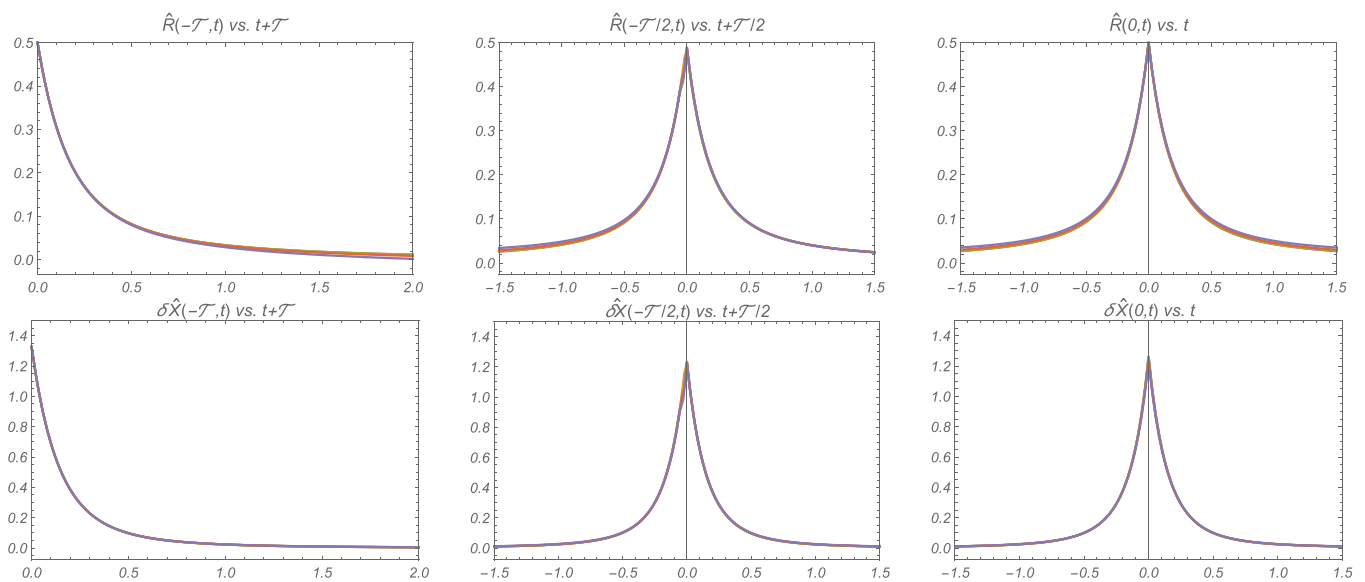

FIG. 8. Spherical p-spin glass for $\beta = 1.695$ and $p = 3$. The functions $\hat{R}(t, t')$ and $\delta \hat{X}(t, t') = \hat{X}(t, t') + \delta(t - t')/2$ for $\mathcal{T} = 8, 16, 24, 32, 40$ on various slices of the $-\mathcal{T} \leqslant t, t' \leqslant \mathcal{T}$ plane. The vertical and horizontal ranges are appropriate to visualize the diagonal region $t - t' = O(1)$. On this scale the functions for different $\mathcal{T}$'s are almost indistinguishable.
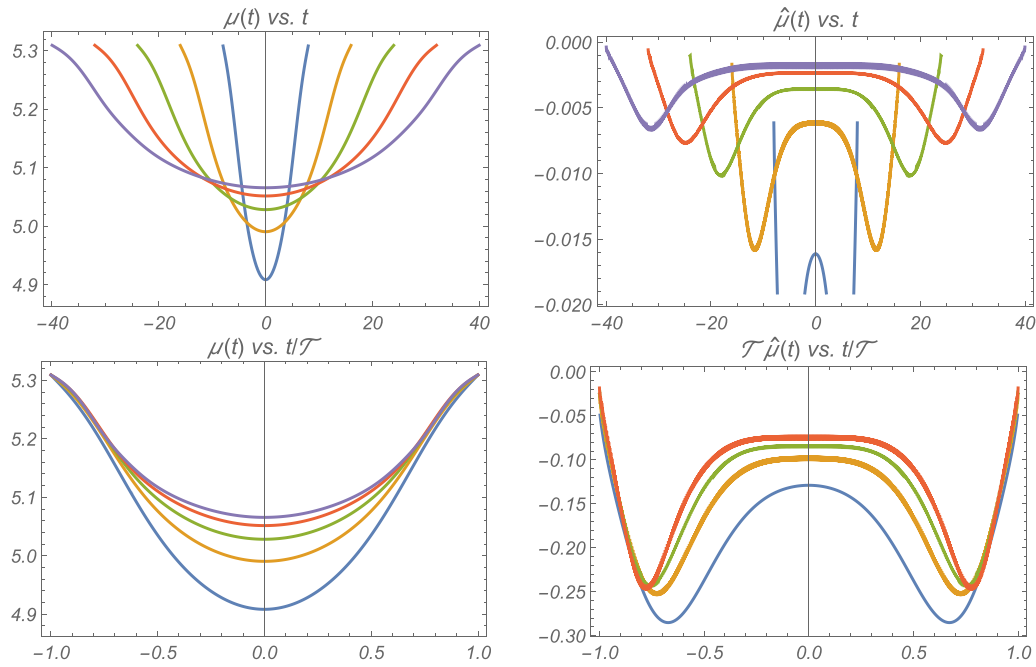
FIG. 9. Spherical p-spin glass for $\beta = 1.695$ and $p = 3$. The auxiliary functions $\mu(t)$ and $\hat\mu(t)$ for $\mathcal{T} = 8$ (blue), 16 (yellow), 24 (green), 32 (red), and 40 (purple). The $\mathcal{T} = 40$ value in the rescaled plot for $\hat\mu(t)$ has not been plot because the extrapolation from $N_t = 2000$ seems not accurate enough on that scale.

I used the generalized minimal residue (GMRES) algorithm [50,51] that requires the computation of the Krylov subspace of order k, defined by the vectors $\{b, A\,b, A^2\,b, A^3\,b, \ldots, A^k b\}$. The advantage of Krylov methods is that one always work with vectors v and has to perform matrix multiplications $A\,v$ without the need to store the full $(32N_t^2) \times (32N_t^2)$ matrix A. In GMRES, one searches for an approximate solution in the Krylov subspace of order k: an orthonormal basis is obtained

using a numerically stable Gram-Schmidt orthogonalization called Arnoldi iteration and the solution is then found by least squares minimization of the linear equations in this space. The advantage of the method is that the error decreases systematically increasing k, the drawback is that it requires to generate and store all the k vectors of the basis and k cannot be too small to obtain accurate solutions. Two important ingredients that are key to the efficiency of the procedure are precondi-
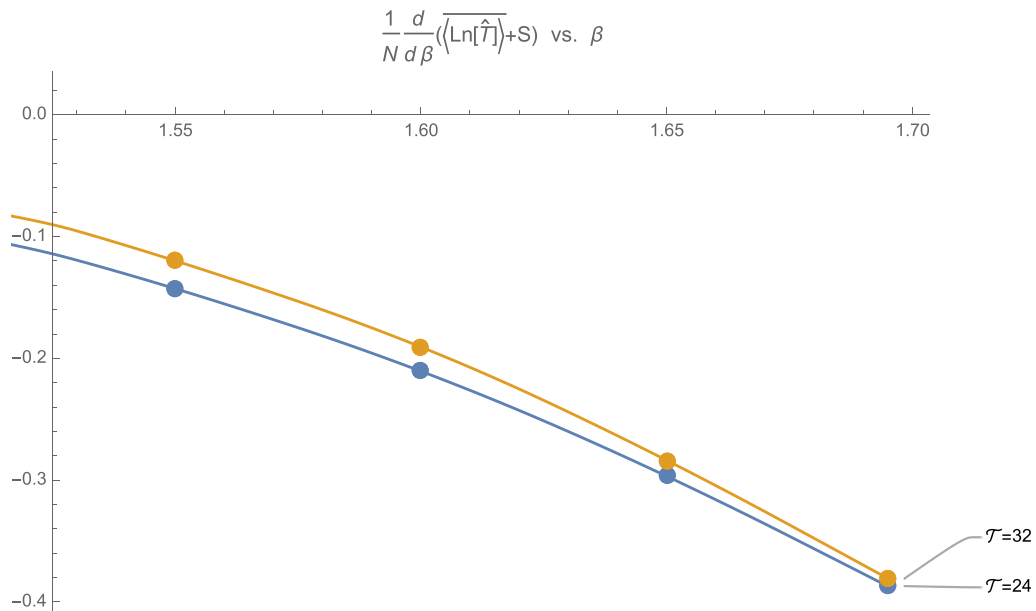

FIG. 10. Spherical p-spin glass with $p = 3$. Derivative of the total transition probability with respect to the inverse temperature for $\mathcal{T} = 24, 32$ and $\beta = 1.55, 1.6, 1.65, 1.695$, the lines are second-order interpolations.
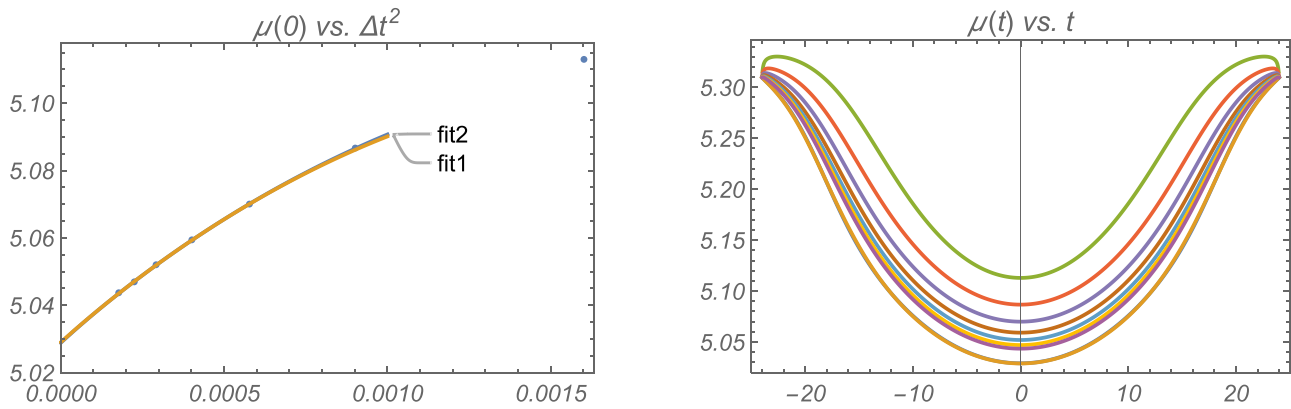
FIG. 11. Spherical p-spin glass with $p = 3$, numerical solution for $\beta = 1.695$, $\mathcal{T} = 24$, and $N_t = 600, 800, 1000, 1200, 1400, 1600, 1800$. (Left) Numerical values of $\mu(0)$, the lines are two polynomial interpolation of the form $c_0 + c_2 \, \Delta t^2 + c_3 \, \Delta t^3 + c_4 \, \Delta t^4$ over the data for $\{1200, 1400, 1600, 1800\}$ and $\{1000, 1200, 1400, 1600\}$. The polynomial extrapolations give respectively $c_0 = 5.0290$ and $c_0 = 5.0288$. (Right) From top to bottom data for increasing values of N_t (individual points are not distinguishable at the scale of the plot). At the bottom, there are two (indistinguishable) lines obtained from polynomial interpolations performed separately for each time t, one of the form $c_0 + c_2 \, \Delta t^2 + c_3 \, \Delta t^3 + c_4 \, \Delta t^4$ over the data for $\{1200, 1400, 1600, 1800\}$, and the other of the form $c_0 + c_2 \, \Delta t^2 + c_3 \, \Delta t^3$ over the data for $\{600, 800, 1000\}$.

tioning and compression to be discussed later. I wrote a code using Mathematica (retrievable in the ancillary files section of Ref. [46]) being able to reach values of $N_t = 2000$ with $k = 400$ using up to 90% memory on a cluster with 256 Giga of RAM.

2. Richardson extrapolation

Once the numerical solutions at fixed β and $\mathcal{T}$ are obtained for various N_t, a polynomial (Richardson) extrapolation is essential to reach the $\Delta t = 0$ limit and remove a few pathologies of the finite Δt solutions. Let us discuss first the order of the algorithm. For the second derivatives appearing in the equations, I have used the formula

$$f''(t) \approx \frac{f(t + \Delta t) + f(t - \Delta t) - 2 f(t)}{\Delta t}, \qquad (98)$$

where

$$\Delta t = \mathcal{T}/N_t. \qquad (99)$$

The above formula has a $O(\Delta t^2)$ error if $f(t)$ has continuous derivatives up to the third order. The integrals have been written as

$$\int_{-\mathcal{T}}^{+\mathcal{T}} f(t) dt \approx \frac{\Delta t}{2} (f(-\mathcal{T}) + f(+\mathcal{T})) + \sum_{i=-N_t+1}^{N_t-1} f(i \, \Delta t) \Delta t. \qquad (100)$$

The above trapezoidal rule also has a $O(\Delta t^2)$ error for a continuous function. As mentioned before, the full algorithm has a $O(\Delta t^2)$ error, however this is *not* trivial because the functions $C(t, t')$, $\hat{R}(t, t')$, and $\hat{X}(t, t')$ have discontinuous odd derivatives for $t = t'$. While the discontinuity of the first derivative is canceled by the delta function, the discontinuity of the third derivative leads to a $O(\Delta t)$ error in expression (98) on the diagonal $t = t'$. However, on the diagonal, the order of the equation is not $O(1)$ but $O(1/\Delta t)$ due to the presence of the delta function in Eq. (C2) and of $\hat{X}(t, t')$ (that can be written as a delta function on the diagonal plus

a regular part) in Eq. (C3), thus even if the absolute error on the diagonal is $O(\Delta t)$, the *relative* error is $O(\Delta t^2)$. In Fig. 11, it is demonstrated that the error is indeed $O(\Delta t^2)$ for $\mu(t)$ and it is shown how that the $\Delta t \to 0$ limit is safely reached by polynomial extrapolations. The figures discussed in the previous sections were all obtained by means of polynomial extrapolations on the largest N_t set available (up to $N_t = 2000$ for $\mathcal{T} = 32, 40$) using in most cases a fourth order form $c_0 + c_2 \, \Delta t^2 + c_3 \, \Delta t^3 + c_4 \, \Delta t^4$ where the vanishing of the linear term Δt was imposed. One should note that the finite N_t curves often display pathologies due to the discretisation that tend to be less severe increasing N_t. For instance, in the right of Fig. 11, we see that $\mu(t)$ at finite N_t displays cusps close to $t = \pm \mathcal{T}$ that are absent for $N_t \to \infty$. It is impressing how a simple polynomial extrapolation over N_t leads to the disappearance of these spurious features and allows to obtain accurate predictions with relatively small values of N_t that are indistinguishable from extrapolations obtained from considerably larger values of N_t. The same features are also seen in Fig. 12, in this case, corrections higher that $O(\Delta t^2)$ are so small that too high-order interpolation functions overfit that data and it is convenient to use the form $c_0 + c_2 \, \Delta t^2$. Besides we see that, even if the finite N_t results for $\hat{\mu}(t)$ have pronounced spurious cusps close to $t \pm \mathcal{T}$ and may even have the wrong sign, the extrapolations are cuspless and negative for all t.

3. Preconditioning

Efficient Krylov methods often require preconditioning which amount to reduce the span of the eivengalues of A. The equations in the form (45) and (46) are ill-conditioned because the operator M contains second-order derivatives leading to an unbounded continuous spectrum. To overcome this problem I have multiplied Eqs. (45) and (46) times M^{-1}. The corresponding Jacobian turns out to have a discrete and bounded spectrum as can be seen numerically using the fact that the Arnoldi diagonalization allows to obtain an approximate set of
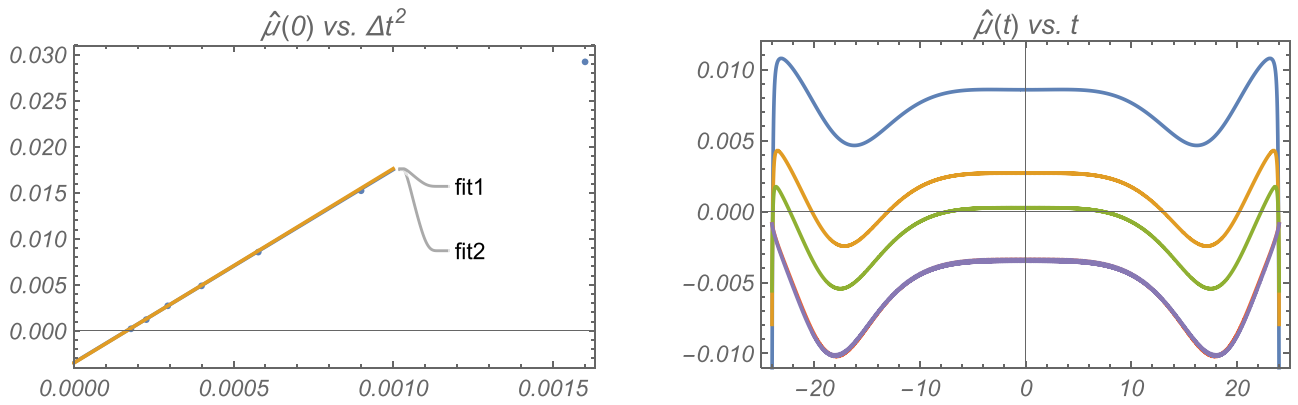
FIG. 12. Spherical p-spin glass with $p = 3$, numerical solution for $\beta = 1.695$, $\mathcal{T} = 24$, and various N_t. (Left) Numerical values of $\hat{\mu}(0)$ for $N_t = 600, 800, 1000, 1200, 1400, 1600, 1800$, the lines are two polynomial interpolation of the form $c_0 + c_2 \, \Delta t^2$ over the data for $\{1600, 1800\}$ and $\{1400, 1600\}$. (Right) From top to bottom data for $N_t = 1000$ (blue), 1400 (yellow), 1800 (green) (individual points are not distinguishable at the scale of the plot). At the bottom, there are two (indistinguishable) lines (red, purple) obtained from polynomial interpolations of the form $c_0 + c_2 \, \Delta t^2$ performed separately for each time t, using respectively $N_t = \{1200, 1400\}$ and $N_t = \{1600, 1800\}$.

eigenvalues and eigenvectors. More details on the procedure can be found in the commented MATHEMATICA codes provided in the ancillary files section of Ref. [46].

4. Compression

To compute the equations and the matrix-times-vector products $J\,c$ involving the Jacobian J, we have to store huge matrices and perform multiplications and elementwise operations, for which many highly optimized and parallelized libraries exist. One can solve Eqs. (45) and (46) for generic C and C^{dt} and verify that as expected the algorithm converges to solutions with the required symmetries discussed in Sec. II A. However, it is useful to work from the start in the subspace of solutions with the required symmetries obtaining a fourfold reductions of the memory required to store C and C^{dt} and thus the generic element of the Krylov subspace. Besides it turns out that this allows to consider a smaller Krylov subspace to obtain the same level of accuracy. Explicit use of the symmetry however requires an efficient procedure to quickly compress and decompress the large matrices C and C^{dt} and the elements of the Krylov subspace. The choice depends on the specific linear algebra programming tool used.

III. CONCLUDING PERSPECTIVES

In this work, I have shown how the transition rate can be computed in mean-field models. As stated in the introduction, the main motivation to perform such a computation is to obtain dynamical quantitative predictions not accessible through static methods. Besides the most interesting physical outcome of the computation is to shed new light on the role of the complex landscape for ergodicity restoring processes as discussed in Sec. I A 3: the conclusion is that to visit different equilibrium states it is more convenient to make intermediate jumps to off-equilibrium metastable states rather than direct jumps from one equilibrium state to another.

The methods discussed here can be applied to other mean-field systems, in the introduction, we mentioned supercooled liquids models in high dimensions but the Langevin equation can be also viewed as an algorithm in the wider family of

(stochastic) gradient-descent algorithms that are widely used in the context of machine-learning and statistical inference. The analysis of the algorithm performance in the context of high-dimensional inference has been initiated in Ref. [3] unveiling a glassy phase that limits the efficiency of the algorithm. Those results where obtained in the thermodynamic limit where the evolution of the algorithm can be associated to relaxational dynamic equations similar to those of the spherical p-spin SG model [22,39] recently reconsidered in Ref. [23]. On the hand, other realistic problems have *finite* size and thus the algorithm can overcome those thresholds albeit with larger convergence times that can be studied by the methods presented here switching from the initial value dynamical equations studied in Ref. [3] to boundary value dynamical equations. Concerning the extension to supercooled liquid in high dimensions, the equations can be derived and solved in principle by the GMRES algorithm although the actual implementation is likely to be considerably more complex than for the spherical model [43]. In this respect, it is worth noticing that for more complex problems one can use a Jacobian-free method, in which the product $J\,v$ required by Newton-Krylov methods is approximated by $E(c + \epsilon\,v) - E(c))/\epsilon$ for small ϵ and thus only the numerical computation of the equations $E(c)$ is required. Furthermore other algorithms exist that do not require to store the whole Krylov space, e.g., biconjugated gradient [50]; in general, they are less safe and controlled than GMRES, but an efficient algorithm requiring less memory would be welcomed. Note that GMRES allows to have an approximation for the spectrum of the true Jacobian, which could be useful information to devise alternative algorithms.

We have mentioned that in the *asymptotic* regime $\mathcal{T} \to \infty$ the solutions solve $\mathcal{T}$-independent scaling equations with the time derivatives dropped. In the context of off-equilibrium dynamics the same thing happens and allows to establish a remarkable connection with the equations of the static/replica framework [22]. This grants that the phase diagram and many nontrivial off-equilibrium quantities can be determined without solving explicitly the dynamical equations and it would be interesting to determine if there is also some static potential from which the results of the $\mathcal{T} \to \infty$ limit can be recovered

without solving the dynamics. A natural question is weather such a static potential is the celebrated Franz-Parisi potential [52] that develops a secondary minimum at T_d. One should indeed remember that, while the minimum is associated to equilibrium states and configurational complexity, it is not known if the potential *difference* between the secondary minimum and the maximum is in fact associated to any quantity computed by actual dynamical methods as those considered here.

An interesting open question is how one should expect the solutions to look like at T_d, in particular in the asymptotic form. Some guidance could be offered by the fact that in the $T \to T_d$ limit the results should match somehow those obtained in Refs. [33–37] by considering finite-size/finite-dimensional systems directly at $T = T_d$ (from a technical point of view the two approaches differ by which one of the limits $T \to T_d$ or $N \to \infty$ is taken first). This suggests for instance that the important physics occurs close to the plateau value of the overlap.

It should be noted that the dynamical equations have been solved by making a replica-symmetric ansatz on the n dynamical replicas introduced to compute the logarithm of the rate [see Eq. (B7) in Appendix B] and one may ask if it yields correct results or a RSB ansatz should be used. While a full answer requires a complicated analysis of the stability of the saddle-point equations, an instability could nonetheless manifest itself in some inconsistent value of some physical quantity, much as the RS entropy of the Sherrington-Kirkpatrick model becomes negative at low temperatures [4]. This is not the case as the numerical solution did not show any visible inconsistency suggesting that the RS ansatz is correct in the range of temperatures considered.

To conclude we recall that in Sec. I A 3, it was shown that the presence of many metastable states implies that the *exponentially small* probability that the system jumps to another equilibrium state in a *finite* time is not trivially related to the *exponentially large* timescale τ_{erg} over which the system jumps to another equilibrium state with *finite* probability. A question worth of further investigation is weather the latter could also be computed by similar methods.

APPENDIX A: PATH INTEGRAL EXPRESSION OF THE RATE

1. Replicas

The logarithm in the quenched average can be eliminated by the replica method

$$\sum_{\sigma,\tau} P_{\text{eq}}(\sigma) P_{\text{eq}}(\tau) \ln \hat{T}(\sigma, \tau), \tag{A1}$$

furthermore we consider systems with quenched disorder whose equilibrium properties can also be studied by the replica method, thus we introduce the following object:

$$Z_{\hat{T}}(m, m', n) \equiv \sum_{\sigma_1...\sigma_m} \sum_{\tau_1...\tau_{m'}} \exp\left[-\beta \sum_{i=1}^{m} H_J(\sigma_i) - \beta \sum_{i=1}^{m'} H_J(\tau_i)\right] \hat{T}(\sigma_1, \tau_1)^n, \tag{A2}$$

and we have

$$\sum_{\sigma,\tau} P_{\text{eq}}(\sigma) P_{\text{eq}}(\tau) \ln \hat{T}(\sigma, \tau) = \lim_{m \to 0} \lim_{m' \to 0} \lim_{n \to 0} \frac{d}{dn} \ln Z_{\hat{T}}(m, m', n). \tag{A3}$$

As usual in the context of mean-field models, the thermodynamic limit is taken before the above limits. Note that the quantity $\lim_{m \to 0, m' \to 0} Z_{\hat{T}}(m, m', n)$ at finite n allows to study the large deviations of $\ln \hat{T}(\sigma, \tau)$ and determine if it is self-averaging with respect to the equilibrium configurations σ and τ. As we will see in the next section it is possible to obtain an integral representation of the dynamics (see Ref. [44], Chap. 17):

$$\hat{T}(\sigma_1, \tau_1) = \int_{s(-\mathcal{T})=\sigma_1, s(\mathcal{T})=\tau_1} ds \exp\left[\frac{1}{2}\int d1 d2 s(1)\Gamma(1, 2)s(2) - \beta \int d1 H_J(s(1))\right]. \tag{A4}$$

Where the 1 and 2 are coordinates that collect a time variable and Fermionic variables introduced to obtain the integral representation. The expression is made of two parts: a universal dynamic one encoded by the matrix $\Gamma(1, 2)$ and an interaction part that depends on the model encoded in the Hamiltonian H_J. The above expressions leads to

$$Z_{\hat{T}}(m, m', n) = \sum_{\sigma_1...\sigma_m} \sum_{\tau_1...\tau_{m'}} \prod_{a=1}^{n}\left[\int_{s_a(-\mathcal{T})=\sigma_1, s_a(\mathcal{T})=\tau_1} [ds_a]\right] \exp\left[\sum_{a=1}^{n} \frac{1}{2}\int d1 d2 s_a(1)\Gamma(1, 2)s_a(2)\right.$$
$$\left. - \beta \sum_{i=1}^{m} H_J(\sigma_i) - \beta \sum_{i=1}^{m'} H_J(\tau_i) - \beta \sum_{a=1}^{n}\int d1 H_J(s_a(1))\right], \tag{A5}$$

this in turn can be written in a compact form as

$$Z_{\hat{T}}(m, m', n) = \int [d\mathbf{s}] \exp\left[\frac{1}{2} \int \mathbf{d1d2s(1)\Gamma(1, 2)s(2)} - \beta \int d\mathbf{1} H_J(\mathbf{s(1)})\right], \quad \text{(A6)}$$

where the bold index **1** runs over the replicas and the dynamical indexes

$$\mathbf{s(1)} = \{\sigma_1, \ldots, \sigma_m, s_1(1), \ldots, s_n(1), \tau_1, \ldots, \tau_{m'}\}, \quad \text{(A7)}$$

and we have

$$\int f(\mathbf{s(1)}) d\mathbf{1} = \sum_{i=1}^{m} f(\sigma_i) + \sum_{i=1}^{m'} f(\tau_i) + \sum_{a=1}^{n} \int d\mathbf{1} f(s_a(1)). \quad \text{(A8)}$$

The dynamical operator $\mathbf{\Gamma(1, 2)}$ is diagonal with respect to the replica indexes and associates each dynamical replica with the boundary conditions at σ_1 and τ_1. Its definition in the specific representation we will use is given in Eq. (B29). As usual the great advantage of having the above compact representation is that one can perform the average of the disorder and then perform standard manipulations yielding an expression formally identical to the one obtained in the case of a static replica computation.

2. Path integral representation of Langevin dynamics

This section discusses the path integral representation of Langevin dynamics, given in terms of a single component q. The one-dimensional case is discussed for simplicity, the generalisation to N-dimensional vectors used in the paper is straightforward. The Langevin equation reads

$$\frac{1}{\Gamma_0}\dot{q} = -\beta\frac{dH}{dq} + \xi, \quad \langle \xi(t)\xi(t')\rangle = \frac{2}{\Gamma_0}\delta(t - t') \quad \text{(A9)}$$

and we discretize it as

$$\frac{1}{\Gamma_0}\frac{q_{i+i} - q_i}{\Delta t} = -\beta\left(c\frac{dH}{dq_i} + (1 - c)\frac{dH}{dq_{i+1}}\right) + \xi_i, \quad \text{(A10)}$$

where $0 \leqslant c \leqslant 1$ is an arbitrary constant: in Itô discretization with have $c = 1$, in Stratonovich, we have $c = 1/2$. Enforcing the equations through an integral representation we can write the average over trajectories at fixed initial and final conditions as an integral

$$T(\sigma|\tau)$$

$$= \int \left(\prod_{i=1}^{N-1} dq_i\right)\left(\prod_{i=0}^{N-1} \frac{d\hat{q}_i}{2\pi}\right)$$

$$\times \exp\left[-\sum_{i=0}^{N-1}\left(\hat{q}_i E_i \Delta t - \ln\left[\frac{1}{\Gamma_0} - (1-c)\beta\frac{d^2H}{dq_{i+1}^2}\Delta t\right]\right)\right].$$

where the interval between the initial and final time is divided in N subintervals of size Δt and

$$E_i \equiv \frac{1}{\Gamma_0}\frac{q_{i+i} - q_i}{\Delta t} + \beta\left(c\frac{dH}{dq_i} + (1 - c)\frac{dH}{dq_{i+1}}\right) - \xi_i. \quad \text{(A11)}$$

The logarithm comes from the determinant of the Jacobian that gets contribution only from the diagonal since the Jacobian is a triangular matrix. Expanding the Jacobian at first order in Δt, we obtain

$$T(\sigma|\tau) = \int \left(\prod_{i=1}^{N-1} dq_i\right)\left(\prod_{i=0}^{N-1} \frac{d\hat{q}_i}{2\pi\Gamma_0}\right)\exp[-\mathcal{L}]. \quad \text{(A12)}$$

The Lagrangian reads

$$\mathcal{L} = \sum_{i=0}^{N-1}\left(\hat{q}_i E_i \Delta t - (1 - c)\beta\frac{d^2H}{dq_{i+1}^2}\Gamma_0 \Delta t\right) \quad \text{(A13)}$$

and in the continuum limit, we have

$$\mathcal{L} = \int dt\left[\frac{1}{\Gamma_0}(\hat{q}\dot{q} - \hat{q}^2) + \hat{q}\beta\frac{dH}{dq} - (1 - c)\Gamma_0\beta\frac{d^2H}{dq^2}\right]. \quad \text{(A14)}$$

Note that unexpectedly the continuum limit expression depends on the microscopic parameter c of the discretization while one would expect it to be irrelevant. This is a well-known ambiguity of path integral representation of stochastic equations. One can choose to use the Itô discretization corresponding to $c = 1$ and neglect it but it will resurface later in the computation. In the following, we prefer to keep it also to remind us that the continuum limit of stochastic equations must be taken with care as the ordinary rules of calculus (integration by parts, differentiation, chain rules) are modified. Beside we will use Hamiltonians where the interaction part is just linear (the p-spin interactions) and thus in the end we will go back to special Langevin equation for a single variable. Let us consider the symmetric rate defined as

$$\hat{T}(\sigma, \tau) \equiv T(\sigma|\tau)e^{\frac{\beta}{2}(H(\sigma)-H(\tau))}. \quad \text{(A15)}$$

In the continuum limit, we would expect the following to be an equality

$$\frac{\beta}{2}(H(\sigma) - H(\tau)) \neq \frac{\beta}{2}\int_{-\mathcal{T}}^{+\mathcal{T}} dt\,\dot{q}\frac{\partial H}{\partial q}. \quad \text{(A16)}$$

Instead in order to get the correct expression, we should go back to the discretized expression. We have

$$H(q_{i+1}) - H(q_i) = \frac{dH}{dq_i}\Delta q_i + \frac{1}{2}\frac{d^2H}{dq_i^2}\Delta q_i^2, \quad \text{(A17)}$$

then we have to use the fact that in the Lagrangian we use the following discretized definition of dH/dq:

$$c\frac{dH}{dq_i} + (1 - c)\frac{dH}{dq_{i+1}}. \quad \text{(A18)}$$

By rewriting the differential as

$$\frac{dH}{dq_i} = \left(c\frac{dH}{dq_i} + (1 - c)\frac{dH}{dq_{i+1}}\right) - (1 - c)\frac{d^2H}{dq_i^2}\Delta q_i, \quad \text{(A19)}$$

we obtain

$$dH = \frac{dH}{dq} dq + \left(c - \frac{1}{2} \right) \frac{d^2 H}{dq^2} dq^2. \qquad (A20)$$

We can see that for $c = 1$, we recover Itô's lemma while for the Stratonovich prescription $c = 1/2$, we find that the ordinary chain rule applies. The second term cannot be neglected because it gives an $O(dt)$ contribution but we can make the replacement

$$dq^2 = 2 \Gamma_0 dt \qquad (A21)$$

and obtain

$$\frac{\beta}{2} (H(\sigma) - H(\tau))$$

$$= \frac{\beta}{2} \int_{-\mathcal{T}}^{+\mathcal{T}} dt \left(\frac{dH}{dq} \dot{q} + \left(c - \frac{1}{2} \right) \frac{d^2 H}{dq^2} 2 \Gamma_0 \right).$$

The same result can also be obtained reabsorbing the term Δq^2 in the term $\dot{q}^2$ below (see e.g., Ref. [44], Sec. 4.6). Making the following change of variable

$$\hat{q} = \hat{x} + \frac{\dot{q}}{2}, \qquad (A22)$$

we finally obtain

$$\hat{\mathcal{L}}(t) = \frac{1}{\Gamma_0} \left(\frac{\dot{q}^2}{4} - \hat{x}^2 \right) + \hat{x} \beta \frac{dH}{dq} - \frac{1}{2} \Gamma_0 \beta \frac{d^2 H}{dq^2}. \quad (A23)$$

We can further integrate out the $\hat{x}$:

$$\hat{\mathcal{L}}'(t) = \frac{1}{\Gamma_0} \frac{\dot{q}^2}{4} + \frac{\Gamma_0}{4} \left(\beta \frac{dH}{dq} \right)^2 - \frac{1}{2} \Gamma_0 \beta \frac{d^2 H}{dq^2}, \quad (A24)$$

see also Ref. [44], p. 70. Note that the integration over $\hat{x}$ leads to a divergent prefactor to the above path integral:

$$\frac{1}{2\pi \Gamma_0} \int d\hat{x} e^{\frac{\hat{x}^2}{\Gamma_0}} dt = \frac{1}{2\pi \Gamma_0} \left(2\pi \frac{\Gamma_0}{2 \, dt} \right)^{1/2}, \qquad (A25)$$

which is usually buried into the expression $[dq]$ defined as

$$[dq] \equiv dq \left(\frac{1/(2\Gamma_0)}{2\pi \, dt} \right)^{N/2} \qquad (A26)$$

see Eq. (2.20) in Ref. [44].

3. Path integral representation for models with Multilinear interactions

Typical mean-field SG models have interactions that are multilinear, while the nonlinear part of the Hamiltonian is local and does not need to be decoupled in order to obtain a saddle-point expression. This allows to use simplified integral representations of the dynamics in which fewer variables are introduced with respect to the general case discussed in Ref. [44], Chap. 17. In the following, we introduce a bosonic variable η that behaves as the product of two Grassmann variables, that is we have

$$\eta^2 = 0, \quad \int d\eta = 0, \quad \int \eta \, d\eta = 1 \qquad (A27)$$

with it we define a new coordinate $1 \equiv (t, \eta)$ and field $s(1)$:

$$s_i(1) \equiv s_i(t) + \hat{x}_i(t) \eta. \qquad (A28)$$

With the above definitions the interacting part of the multilinear Hamiltonian can be written as

$$\beta \int dt \sum_i \hat{x}_i(t) \frac{dH}{ds_i} = \beta \int d1 H(s(1)). \qquad (A29)$$

A similar formulation is also useful in the spherical model. In this case, however, the nonlinear part of the Hamiltonian is due to the spherical constraint which is not local and must be treated appropriately. Let us consider the problem of the integral representation of Langevin dynamics of N real spins s_i constrained on a $(N - 1)$-dimensional surface specified by some condition $G(s) = 0$. The statics of the problem can be written as

$$\int d^N s \, \delta(G) |\nabla G| e^{-\beta H(s)}. \qquad (A30)$$

A convenient way to define Langevin dynamics on the surface is to relax the delta function replacing it with a Gaussian of infinitesimal variance ϵ. Then one have to compute the standard dynamical integral in presence of a Hamiltonian $H(x) = G^2(x)/(2N\epsilon)$. For the spherical constraint on N continuous spins s_i, we have

$$G(s) \equiv \sum_{i=1}^{N} s_i^2 - N \qquad (A31)$$

the term $|\nabla G|$ is exactly equal to N and can be ignored, For $H(x) = G^2(x)/(2N\epsilon)$, we have

$$\sum_i \hat{x}_i \beta \frac{dH}{dq_i} - \frac{1}{2} \Gamma_0 \beta \sum_i \frac{d^2 H}{dq_i^2} = \frac{\beta}{N\epsilon} G(s) \left(2 \sum_i \hat{x}_i s_i - \Gamma_0 N \right) \qquad (A32)$$

plus $o(N)$ terms. The expression can be decoupled through a Hubbard-Stratonovich transformation in terms of two additional fields $\mu(t)$ and $\tilde{\mu}(t)$ and taking the limit $\epsilon \to 0$ the quadratic part disappears leading to the following contribution to the action:

$$-\frac{1}{2} \int dt \hat{\mu}(t) \left(\sum_i s_i^2 - N \right)$$

$$-\frac{1}{2} \int dt \mu(t) \left(2 \sum_i \hat{x}_i s_i - \Gamma_0 N \right).$$

Introducing the variable $\mu(1) = \mu(t) + \hat{\mu}(t)\eta$, we can write the above term as

$$-\frac{1}{2} \int d1 \mu(1) \left(\sum_i s_i^2(1) - N \right) + \frac{1}{2} \Gamma_0 \int dt \mu(t).$$

Note that presence on the second term that is essential to get the correct saddle-point equations and the correct value of the rate.

APPENDIX B: THE p-SPIN SPHERICAL MODEL: THE ACTION AND SADDLE-POINT EQUATIONS

In the standard treatment [4], the expression of the free energy of the fully-connected spin-glass models is obtained through a number of steps. First n replicas of the model

are introduced, then the disorder average over the partition function of the replicated system is performed and this leads to an expression in which different replicas are coupled. Then a Hubbard-Stratonovich transformation is performed leading to an action depending on a single local replicated spin variable and on two $n \times n$ matrices Q_{ab} and Λ_{ab}. Finally the action has to be integrated over Q_{ab} and Λ_{ab} but the mean-field nature of the model allows to use the saddle-point method. The very same manipulations can be applied to the interaction part of expression (A6), the only difference is that instead of having spins s_a labeled by indexes $a = 1, \ldots, n$ the spins depend on a more complex index (or coordinate) **1** specified in Eq. (A7). As a result the objects Q and Λ are (formally) matrices in this more complex coordinate. In spherical models, the spin variables can be also integrated out and we obtain the following expression for $Z_{\hat{T}}$ of the p-spin spherical model:

$$
\begin{aligned}
\exp N \bigg[&\frac{\beta^2}{4} \int d\mathbf{a} d\mathbf{b} f(\mathbf{Q}(\mathbf{a}, \mathbf{b})) + \frac{1}{2} \mathbf{\Lambda}(\mathbf{a}, \mathbf{b}) \mathbf{Q}(\mathbf{a}, \mathbf{b}) \\
&- \frac{1}{2} \mathrm{Tr} \ln(\mathbf{\Gamma} + \mu + \mathbf{\Lambda}) + const. + \frac{1}{2} \int d\mathbf{a}\, \mu(\mathbf{a}) \\
&+ \sum_{a=1}^{n} \frac{\Gamma_0}{2} \int \mu_a(t) dt \bigg].
\end{aligned}
\tag{B1}
$$

The expression const. above collects a number of terms coming from the explicit Gaussian integration, it is divergent and cancels the divergences associated to the expression $\mathrm{Tr} \ln(\mathbf{\Gamma} + \mu + \mathbf{\Lambda})$. These are pathologies of the path integral representation that are fixed going back to the discrete times and will be further discussed in Appendix G. Note that the application of the standard manipulations to the form (A6) would lead to the above expression without the last term that appears instead if we want to use the simplified formulation which is suitable for multilinear interactions. The above expression has to be extremized with respect to $\mathbf{Q}(\mathbf{ab})$, $\mathbf{\Lambda}(\mathbf{ab})$, and with respect to $\boldsymbol{\mu}(\boldsymbol{a})$. Extremization with respect to $\mathbf{Q}(\mathbf{ab})$, $\mathbf{\Lambda}(\mathbf{ab})$ leads to the saddle-point equations:

$$
\mathbf{\Lambda}(\mathbf{ab}) = -\frac{\beta^2}{2} f'(\mathbf{Q}(\mathbf{ab})),
\tag{B2}
$$

$$
\mathbf{Q} = \frac{1}{\mathbf{\Gamma} + \mathbf{\Lambda} + \mu}.
\tag{B3}
$$

The last equation can be rewritten as

$$
\begin{aligned}
\mathbf{\Gamma}(\boldsymbol{ac})\mathbf{Q}(\boldsymbol{cb}) &+ \mu(\boldsymbol{a})\, \delta(\boldsymbol{ac})\mathbf{Q}(\boldsymbol{cb}) + \mathbf{\Lambda}(\boldsymbol{ac})\mathbf{Q}(\boldsymbol{cb}) \\
&= \delta(\boldsymbol{ab}),
\end{aligned}
\tag{B4}
$$

where there integration over the variable $\boldsymbol{c}$ is implicit. The above expression is extremely compact, in the following, we will see that it encodes eight integrodifferential equations.

From now on we specialize to the case of a 1RSB transition. We start noticing that the initial and final configuration are weighted with the equilibrium Gibbs measure and their properties must not depend on the dynamics. This is granted by the fact that the terms depending on the dynamics in the equations for the replicas of the initial and final configurations are $O(n)$ and disappear in the $n \to 0$ limit (that must be taken first). This means that $Q(ab)$ with a and b corresponding to

the $a, b = 1, \ldots, m + m'$ replicas associated to the equilibrium boundary conditions is the ordinary equilibrium replica matrix. For simplicity, we will work in zero field and zero random field meaning that

$$
f'(0) = 0,
\tag{B5}
$$

in this way the overlap between different equilibrium configurations is zero. Therefore above T_d the solution is $Q(ab) = \delta(ab)$ that plugged into the above equations leads to

$$
\mu = 1 + \frac{\beta^2}{2} f'(1).
\tag{B6}
$$

Below T_d the solution is actually 1RSB. The replicas determining the boundary conditions should naturally belong to different RSB blocks ensuring that we are studying the transition between different states. On the other hand the equations for the dynamical part will have a nonvanishing correlation with the remaining $x - 1$ replicas in the block of the initial configuration and with the $x - 1$ replicas in the block of the final configuration. This will lead to a correction of order $x - 1$ to the dynamical equations valid at high temperature. However, for $T_s < T < T_d$, we have exactly $x = 1$ and thus this contribution vanishes. Therefore we will safely use the high temperature replica-symmetric (RS) solution $Q(ab) = \delta(ab)$ also in the region $T_s < T < T_d$. This is consistent with the known result that annealed and quenched averages are equivalent above T_s.

The global order parameter $\boldsymbol{Q(ab)}$ can be divided into a static part corresponding to the $m + m'$ replicas, a dynamic part describing the n replicas of the dynamics and a mixed part. We have seen before that, as it should, the static part does not depend on the dynamics part due to the $n \to 0$ limit. We now focus on the dynamic part. We will make a RS ansatz on the n dynamical replicas, therefore the dynamical component of $\boldsymbol{Q(ab)}$ will be characterized by two matrices

$$
\boldsymbol{Q(ab)} = \delta_{\alpha\beta} Q(ab) + (1 - \delta_{\alpha\beta}) Q^{dt}(ab),
\tag{B7}
$$

where we have moved from the full coordinates $\boldsymbol{(ab)}$ to purely dynamical coordinates (ab) and replica coordinates $\alpha, \beta = 1, \ldots, n$. The RS ansatz implies also

$$
\mu_\alpha(t) = \mu(t), \quad \hat{\mu}_\alpha(t) = \hat{\mu}(t)
\tag{B8}
$$

and we have

$$
\boldsymbol{\Gamma(ab)} = \delta_{\alpha\beta} \Gamma(ab),
\tag{B9}
$$

$$
\boldsymbol{\delta(ab)} = \delta_{\alpha\beta} \delta(ab).
\tag{B10}
$$

The dynamical components of equations (B4) can then be rewritten in a form that can be analytically continued to real values of the replica number n

$$
\Gamma(ac)Q(cb) + \beta\mu(a)\delta(ac)Q(cb) + (\Lambda Q)^{st} = \delta(ab), \tag{B11}
$$

$$
\Gamma(ac)Q^{dt}(cb) + \beta\mu(a)\delta(ac)Q^{dt}(cb) + (\Lambda Q)^{dt} = 0, \tag{B12}
$$

where again the integration over the variable c is implicit and

$$
\begin{aligned}
(\Lambda Q)^{st} = \Lambda(ac)Q(cb) &+ (n-1)\Lambda^{dt}(ac)Q^{dt}(cb) \\
&+ \Lambda(a1)Q(1b) + \Lambda(a2)Q(2b)',
\end{aligned}
\tag{B13}
$$

$$(\Lambda Q)^{dt} = \Lambda^{dt}(ac)Q(cb) + \Lambda(ac)Q^{dt}(cb)$$
$$+ (n-2)\Lambda^{dt}(ac)Q^{dt}(cb)$$
$$+ \Lambda(a1)Q(1b) + \Lambda(a2)Q(2b), \quad \text{(B14)}$$

where 1 and 2 label the two static replicas whose configurations are chosen as initial and final condition of the dynamics at times $\mp\mathcal{T}$. Note that the corresponding terms appear when we integrate over the full coordinate c in Eq. (B4). In the above equations, we have naturally

$$\Lambda(ab) = -\frac{\beta^2}{2}f'(Q(ab)), \quad \text{(B15)}$$

$$\Lambda^{dt}(ab) = -\frac{\beta^2}{2}f'(Q^{dt}(ab)). \quad \text{(B16)}$$

$Q(ab)$ and $Q^{dt}(ab)$ can be expressed in terms of four two-time functions as

$$Q(ab) \equiv C(t_a, t_b) + \hat{R}_1(t_a, t_b)\eta_a + \hat{R}_2(t_a, t_b)\eta_b$$
$$+ \hat{X}(t_a, t_b)\eta_a\eta_b, \quad \text{(B17)}$$

$$Q^{dt}(ab) \equiv C^{dt}(t_a, t_b) + \hat{R}_1^{dt}(t_a, t_b)\eta_a + \hat{R}_2^{dt}(t_a, t_b)\eta_b$$
$$+ \hat{X}^{dt}(t_a, t_b)\eta_a\eta_b. \quad \text{(B18)}$$

The same representation can be obtained from any function $A(ab)$ as

$$A(ab) \equiv C_A(t_a, t_b) + \hat{R}_{1,A}(t_a, t_b)\eta_a$$
$$+ \hat{R}_{2,A}(t_a, t_b)\eta_b + \hat{X}_A(t_a, t_b)\eta_a\eta_b, \quad \text{(B19)}$$

from which we obtain

$$C_\Lambda(t_a, t_b) = -\frac{\beta^2}{2}f'(C(t_a, t_b)), \quad \text{(B20)}$$

$$\hat{R}_{1,\Lambda}(t_a, t_b) = -\frac{\beta^2}{2}f''(C(t_a, t_b))\,\hat{R}_1(t_a, t_b), \quad \text{(B21)}$$

$$\hat{R}_{2,\Lambda}(t_a, t_b) = -\frac{\beta^2}{2}f''(C(t_a, t_b))\,\hat{R}_2(t_a, t_b), \quad \text{(B22)}$$

$$\hat{X}_\Lambda(t_a, t_b) = -\frac{\beta^2}{2}f''(C(t_a, t_b))\,\hat{X}(t_a, t_b)$$
$$-\frac{\beta^2}{2}f'''(C(t_a, t_b))\,\hat{R}_1(t_a, t_b)\hat{R}_2(t_a, t_b), \quad \text{(B23)}$$

$$C_{\Lambda^{dt}}(t_a, t_b) = -\frac{\beta^2}{2}f'(C^{dt}(t_a, t_b)), \quad \text{(B24)}$$

$$\hat{R}_{1,\Lambda^{dt}}(t_a, t_b) = -\frac{\beta^2}{2}f''(C^{dt}(t_a, t_b))\,\hat{R}_1^{dt}(t_a, t_b), \quad \text{(B25)}$$

$$\hat{R}_{2,\Lambda^{dt}}(t_a, t_b) = -\frac{\beta^2}{2}f''(C^{dt}(t_a, t_b))\,\hat{R}_2^{dt}(t_a, t_b), \quad \text{(B26)}$$

$$\hat{X}_{\Lambda^{dt}}(t_a, t_b) = -\frac{\beta^2}{2}f''(C^{dt}(t_a, t_b))\,\hat{X}^{dt}(t_a, t_b)$$
$$-\frac{\beta^2}{2}f'''(C^{dt}(t_a, t_b))\,\hat{R}_1^{dt}(t_a, t_b)\hat{R}_2^{dt}(t_a, t_b). \quad \text{(B27)}$$

The operator $\Gamma(ab)$ is defined through

$$\frac{1}{2}\int da\,db\,\Gamma(ab)q(a)q(b) \equiv \int \frac{1}{\Gamma_0}\left(\frac{\dot{q}^2}{4} - \hat{x}^2\right)dt \quad \text{(B28)}$$

that leads to

$$\Gamma(ab) \equiv \frac{1}{\Gamma_0}\left[-\frac{1}{2}\delta''(t_a - t_b)\eta_a\,\eta_b - 2\,\delta(t_a - t_b)\right]. \quad \text{(B29)}$$

For $\Gamma_0 = 1$, we have

$$\Gamma(ac)Q(cb) = -2\hat{R}_1(t_a, t_b) - \frac{1}{2}\frac{d^2}{dt_a^2}C(t_a, t_b)\eta_a$$
$$-2\hat{X}(t_a, t_b)\eta_b - \frac{1}{2}\frac{d^2}{dt_a^2}\hat{R}_2(t_a, t_b)\eta_a\eta_b.$$

The corresponding expression in Eq. (B12) is obtained replacing $Q(ab)$ with $Q^{dt}(ab)$. The term depending on μ can be written as

$$\mu(a)Q(ab) = \mu(t_a)C(t_a, t_b) + \eta_a(\mu(t_a)\hat{R}_1(t_a, t_b)$$
$$+ \hat{\mu}(t_a)C(t_a, t_b)) + \eta_b\,\mu(t_a)\hat{R}_2(t_a, t_b)$$
$$+ \eta_a\eta_b(\hat{\mu}(t_a)\hat{R}_2(t_a, t_b) + \mu(t_a)\hat{X}(t_a, t_b)).$$

As above, the corresponding expression in Eq. (B12) is obtained replacing $Q(ab)$ with $Q^{dt}(ab)$. In order to complete the derivation of the saddle-point equations, we need the expression for a product of the form $A(ac)B(cb)$:

$$A(ac)B(cb)$$
$$= C_A(t_a, t_c)\hat{R}_{1,B}(t_c, t_b) + \hat{R}_{2,A}(t_a, t_c)C_B(t_c, t_b)$$
$$+ \eta_a(\hat{R}_{1,A}(t_a, t_c)\hat{R}_{1,B}(t_c, t_b) + \hat{X}_A(t_a, t_c)C_B(t_c, t_b))$$
$$+ \eta_b(\hat{R}_{2,A}(t_a, t_c)\hat{R}_{2,B}(t_c, t_b) + C_A(t_a, t_c)\hat{X}_B(t_c, t_b))$$
$$+ \eta_a\eta_b(\hat{R}_{1,A}(t_a, t_c)\hat{X}_B(t_c, t_b) + \hat{X}_A(t_a, t_c)\hat{R}_{2,B}(t_c, t_b)).$$
$$\text{(B30)}$$

For the contributions of the initial and final configurations in the interaction term, we have

$$\Lambda(a1)Q(1b) + \Lambda(a2)Q(2b)$$
$$= C_\Lambda(t_a, -\mathcal{T})C(-\mathcal{T}, t_b) + C_\Lambda(t_a, \mathcal{T})C(\mathcal{T}, t_b)$$
$$+ \eta_a(\hat{R}_{1,\Lambda}(t_a, -\mathcal{T})C(-\mathcal{T}, t_b) + \hat{R}_{1,\Lambda}(t_a, \mathcal{T})C(\mathcal{T}, t_b))$$
$$+ \eta_b(C_\Lambda(t_a, -\mathcal{T})\hat{R}_2(-\mathcal{T}, t_b) + C_\Lambda(t_a, \mathcal{T})\hat{R}_2(\mathcal{T}, t_b))$$
$$+ \eta_a\eta_b(\hat{R}_{1,\Lambda}(t_a, -\mathcal{T})\hat{R}_2(-\mathcal{T}, t_b)$$
$$+ \hat{R}_{1,\Lambda}(t_a, \mathcal{T})\hat{R}_2(\mathcal{T}, t_b)). \quad \text{(B31)}$$

We also have

$$\delta(ab) = (\eta_a + \eta_b)\delta(t_a - t_b). \quad \text{(B32)}$$

Collecting the various components in Eqs. (B11) and (B12), we obtain eight integrodifferential equations that will be written in expanded form in Appendix C and that were written in compact form in Sec. II B. The extremization of expression (B1) with respect to $\mu(t)$ and $\hat{\mu}(t)$ gives the following conditions:

$$C(t, t,) = 1, \quad \hat{R}_1(t, t) = \hat{R}_2(t, t') = \tfrac{1}{2}. \quad \text{(B33)}$$

Note that the last term in (B1) is essential to obtain the correct expression for $\hat{R}_1(t, t)$ and $\hat{R}_2(t, t)$.

APPENDIX C: THE EQUATIONS IN EXPANDED FORM

Expressions (45) and (46) are compact and useful for a numerical treatment. They correspond to eight integrodiffer-ential equations that we write in the following in explicit form for completeness. They have to be supplemented with the definitions (B20)–(B27) and the boundary conditions discussed in Sec. II B.

$$
0 = -2\hat{R}_1(t,t') + \mu(t)C(t,t') + \int_{-\mathcal{T}}^{+\mathcal{T}} (C_\Lambda(t,t'')\hat{R}_1(t'',t') + \hat{R}_{2,\Lambda}(t,t'')C(t'',t'))dt''
$$
$$
+ (n-1)\int_{-\mathcal{T}}^{+\mathcal{T}} \left(C_{\Lambda^{dt}}(t,t'')\hat{R}_1^{dt}(t'',t') + \hat{R}_{2,\Lambda^{dt}}(t,t'')C^{dt}(t'',t')\right)dt'' + C_\Lambda(t,-\mathcal{T})C(-\mathcal{T},t') + C_\Lambda(t,\mathcal{T})C(\mathcal{T},t'), \quad \text{(C1)}
$$

$$
\delta(t-t') = -\frac{1}{2}\frac{d^2}{dt^2}C(t,t') + \mu(t)\hat{R}_1(t,t') + \hat{\mu}(t)C(t,t') + \int_{-\mathcal{T}}^{+\mathcal{T}} (\hat{R}_{1,\Lambda}(t,t'')\hat{R}_1(t'',t') + \hat{X}_\Lambda(t,t'')C(t'',t'))dt''
$$
$$
+ (n-1)\int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{1,\Lambda^{dt}}(t,t'')\hat{R}_1^{dt}(t'',t') + \hat{X}_{\Lambda^{dt}}(t,t'')C^{dt}(t'',t')\right)dt'' + \hat{R}_{1,\Lambda}(t,-\mathcal{T})C(-\mathcal{T},t') + \hat{R}_{1,\Lambda}(t,\mathcal{T})C(\mathcal{T},t'),
$$
$$
\text{(C2)}
$$

$$
\delta(t-t') = -2\hat{X}(t,t') + \mu(t)\hat{R}_2(t,t') + \int_{-\mathcal{T}}^{+\mathcal{T}} (\hat{R}_{2,\Lambda}(t,t'')\hat{R}_2(t'',t') + C_\Lambda(t,t'')\hat{X}(t'',t'))dt''
$$
$$
+ (n-1)\int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{2,\Lambda^{dt}}(t,t'')\hat{R}_2^{dt}(t'',t') + C_{\Lambda^{dt}}(t,t'')\hat{X}^{dt}(t'',t')\right)dt'' + C_\Lambda(t,-\mathcal{T})\hat{R}_2(-\mathcal{T},t') + C_\Lambda(t,\mathcal{T})\hat{R}_2(\mathcal{T},t'),
$$
$$
\text{(C3)}
$$

$$
0 = -\frac{1}{2}\frac{d^2}{dt^2}\hat{R}_2(t,t') + \mu(t)\hat{X}(t,t') + \hat{\mu}(t)\hat{R}_2(t,t') + \int_{-\mathcal{T}}^{+\mathcal{T}} (\hat{R}_{1,\Lambda}(t,t'')\hat{X}(t'',t') + \hat{X}_\Lambda(t,t'')\hat{R}_2(t'',t'))dt''
$$
$$
+ (n-1)\int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{1,\Lambda^{dt}}(t,t'')\hat{X}^{dt}(t'',t') + \hat{X}_{\Lambda^{dt}}(t,t'')\hat{R}_2^{dt}(t'',t')\right)dt'' + \hat{R}_{1,\Lambda}(t,-\mathcal{T})\hat{R}_2(-\mathcal{T},t') + \hat{R}_{1,\Lambda}(t,\mathcal{T})\hat{R}_2(\mathcal{T},t'),
$$
$$
\text{(C4)}
$$

$$
0 = -2\hat{R}_1^{dt}(t,t') + \mu(t)C^{dt}(t,t') + \int_{-\mathcal{T}}^{+\mathcal{T}} (C_{\Lambda^{dt}}(t,t'')\hat{R}_1(t'',t') + \hat{R}_{2,\Lambda^{dt}}(t,t'')C(t'',t'))dt''
$$
$$
+ \int_{-\mathcal{T}}^{+\mathcal{T}} \left(C_\Lambda(t,t'')\hat{R}_1^{dt}(t'',t') + \hat{R}_{2,\Lambda}(t,t'')C^{dt}(t'',t')\right)dt'' + (n-2)\int_{-\mathcal{T}}^{+\mathcal{T}} \left(C_{\Lambda^{dt}}(t,t'')\hat{R}_1^{dt}(t'',t')\right.
$$
$$
\left. + \hat{R}_{2,\Lambda^{dt}}(t,t'')C^{dt}(t'',t')\right)dt'' + C_\Lambda(t,-\mathcal{T})C(-\mathcal{T},t') + C_\Lambda(t,\mathcal{T})C(\mathcal{T},t'), \quad \text{(C5)}
$$

$$
0 = -\frac{1}{2}\frac{d^2}{dt^2}C^{dt}(t,t') + \mu(t)\hat{R}_1^{dt}(t,t') + \hat{\mu}(t)C^{dt}(t,t') + \int_{-\mathcal{T}}^{+\mathcal{T}} (\hat{R}_{1,\Lambda^{dt}}(t,t'')\hat{R}_1(t'',t') + \hat{X}_{\Lambda^{dt}}(t,t'')C(t'',t'))dt''
$$
$$
+ \int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{1,\Lambda}(t,t'')\hat{R}_1^{dt}(t'',t') + \hat{X}_\Lambda(t,t'')C^{dt}(t'',t')\right)dt'' + (n-2)\int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{1,\Lambda^{dt}}(t,t'')\hat{R}_1^{dt}(t'',t')\right.
$$
$$
\left. + \hat{X}_{\Lambda^{dt}}(t,t'')C^{dt}(t'',t')\right)dt'' + \hat{R}_{1,\Lambda}(t,-\mathcal{T})C(-\mathcal{T},t') + \hat{R}_{1,\Lambda}(t,\mathcal{T})C(\mathcal{T},t'), \quad \text{(C6)}
$$

$$
0 = -2\hat{X}^{dt}(t,t') + \mu(t)\hat{R}_2^{dt}(t,t') + \int_{-\mathcal{T}}^{+\mathcal{T}} (\hat{R}_{2,\Lambda^{dt}}(t,t'')\hat{R}_2(t'',t') + C_{\Lambda^{dt}}(t,t'')\hat{X}(t'',t'))dt''
$$
$$
+ \int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{2,\Lambda}(t,t'')\hat{R}_2^{dt}(t'',t') + C_\Lambda(t,t'')\hat{X}^{dt}(t'',t')\right)dt'' + (n-2)\int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{2,\Lambda^{dt}}(t,t'')\hat{R}_2^{dt}(t'',t')\right.
$$
$$
\left. + C_{\Lambda^{dt}}(t,t'')\hat{X}^{dt}(t'',t')\right)dt'' + C_\Lambda(t,-\mathcal{T})\hat{R}_2(-\mathcal{T},t') + C_\Lambda(t,\mathcal{T})\hat{R}_2(\mathcal{T},t'), \quad \text{(C7)}
$$

$$
0 = -\frac{1}{2}\frac{d^2}{dt^2}\hat{R}_2^{dt}(t,t') + \mu(t)\hat{X}^{dt}(t,t') + \hat{\mu}(t)\hat{R}_2^{dt}(t,t') + \int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{1,\Lambda^{dt}}(t,t'')\hat{X}(t'',t') + \hat{X}_{\Lambda^{dt}}(t,t'')\hat{R}_2(t'',t')\right)dt''
$$
$$
+ \int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{1,\Lambda}(t,t'')\hat{X}^{dt}(t'',t') + \hat{X}_\Lambda(t,t'')\hat{R}_2^{dt}(t'',t')\right)dt'' + (n-2)\int_{-\mathcal{T}}^{+\mathcal{T}} \left(\hat{R}_{1,\Lambda^{dt}}(t,t'')\hat{X}^{dt}(t'',t')\right.
$$
$$
\left. + \hat{X}_{\Lambda^{dt}}(t,t'')\hat{R}_2^{dt}(t'',t')\right)dt'' + \hat{R}_{1,\Lambda}(t,-\mathcal{T})\hat{R}_2(-\mathcal{T},t') + \hat{R}_{1,\Lambda}(t,\mathcal{T})\hat{R}_2(\mathcal{T},t'). \quad \text{(C8)}
$$

APPENDIX D: THE SOLUTIONS ON THE CORNERS AND ON THE DIAGONAL

In this Appendix, we discuss a number of useful properties that follow from the equations. Equation (C3) implies that

$$\hat{X}(t, t') = -\tfrac{1}{2}\delta(t - t') + \delta\hat{X}(t, t'), \quad \text{(D1)}$$

where $\delta\hat{X}(t, t')$ is a bounded function. In the limit $t, t' \to \pm\mathcal{T}$, the interaction part in the equations greatly simplifies because we have

$$C(-\mathcal{T}, t) = C^{dt}(-\mathcal{T}, t), \quad \hat{R}_1(t, -\mathcal{T}) = \hat{R}_1^{dt}(t, -\mathcal{T}),$$

$$\delta\hat{X}(-\mathcal{T}, t) = \hat{X}^{dt}(-\mathcal{T}, t).$$

This allows to characterize the order parameters on the corners of the t, t' domain. In particular, one easily sees that for $n = 0$, the contributions of the integrals between $-\mathcal{T}$ and $\mathcal{T}$ cancel in the limit $t, t' \to -\mathcal{T}$ in all equations. In particular, Eq. (C1) leads to

$$-1 + \mu(-\mathcal{T}) - \frac{\beta^2}{2}f'(1) = 0 \to \lim_{t \to \pm\mathcal{T}} \mu(t) = \mu_{\text{eq}}. \quad \text{(D2)}$$

Thus $\mu(t)$ goes continuously to the equilibrium value as $t \to \pm\mathcal{T}$. Equations (C3) and (C7) lead to

$$\delta\hat{X}(\pm\mathcal{T}, \pm\mathcal{T}) = \hat{X}^{dt}(\pm\mathcal{T}, \pm\mathcal{T}) = \frac{\mu_{\text{eq}}}{4} = \frac{1}{4} + \frac{\beta^2}{8}f'(1). \quad \text{(D3)}$$

Similarly for $n = 0$, all terms depending on Λ cancels in the limit $t \to \mp\mathcal{T}$, $t' \to \pm\mathcal{T}$ due to the boundary conditions

$$0 = C(\pm\mathcal{T}, \mp\mathcal{T}) = C^{dt}(\pm\mathcal{T}, \mp\mathcal{T}), \quad \text{(D4)}$$

and one obtains from Eqs. (C1) and (C3):

$$0 = \hat{R}_1(\pm\mathcal{T}, \mp\mathcal{T}) = \hat{R}_1^{dt}(\pm\mathcal{T}, \mp\mathcal{T}), \quad \text{(D5)}$$

$$0 = \hat{R}_2(\pm\mathcal{T}, \mp\mathcal{T}) = \hat{R}_2^{dt}(\pm\mathcal{T}, \mp\mathcal{T}), \quad \text{(D6)}$$

$$0 = \delta\hat{X}(\pm\mathcal{T}, \mp\mathcal{T}) = \hat{X}^{dt}(\pm\mathcal{T}, \mp\mathcal{T}). \quad \text{(D7)}$$

APPENDIX E: DERIVATIVES OF THE LOGARITHM OF THE RATE

The expression (B1) of the logarithm of the rate in terms of the order parameter obtained in the path integral formulation is divergent as is shown explicitly for the free case in Appendix (G) and its computation requires to go back to the discretized case. A convenient alternative is to compute the derivative of the logarithm of the rate with respect to the temperature and integrate the result using the knowledge of the infinite temperature limit given in Sec. II C. The partial derivative with respect to the temperature of expression (B1) is simply given by $\frac{\beta}{2}\frac{d}{dn}\int d\mathbf{a}d\mathbf{b}f(\mathbf{Q}(\mathbf{a}, \mathbf{b}))$ and it coincides with the total derivative when computed on the solution of the saddle-point equations. Using the formulas for the integrals computed below (Appendix E 1), we

obtain

$$\frac{1}{N}\frac{\partial\overline{[\ln\hat{T}]}}{\partial\beta} = \frac{\beta}{2}\left(4\int_{-\mathcal{T}}^{+\mathcal{T}}\hat{R}_{1,f[C]}(t, -\mathcal{T})dt\right.$$

$$+ \int_{-\mathcal{T}}^{+\mathcal{T}}\int_{-\mathcal{T}}^{+\mathcal{T}}\hat{X}_{f[C]}(t, t')dt\,dt'$$

$$\left. + (n - 1)\int_{-\mathcal{T}}^{+\mathcal{T}}\int_{-\mathcal{T}}^{+\mathcal{T}}\hat{X}_{f[C^{dt}]}(t, t')dt\,dt'\right). \quad \text{(E1)}$$

As discussed in Sec. I A 3, it is interesting to consider a generalized $[\ln\hat{T}]$ in which the final configuration at time $+\mathcal{T}$ is selected with the Gibbs weight corresponding to a different temperature β_2. This allows to compute the total probability to jump to states with energy higher than the equilibrium one. Again the actual computation of the divergent expression can be avoided studying the derivative of $[\ln\hat{T}]$ with respect to β_2. The same argument above implies that the total derivative coincides with the partial derivative. With a computation similar to that leading to Eq. (E11), one finds

$$\frac{1}{N}\frac{\partial\overline{[\ln\hat{T}]}}{\partial\beta_2} = \frac{\beta}{2}\int_{-\mathcal{T}}^{+\mathcal{T}}\hat{R}_{1,f[C]}(t, \mathcal{T})dt. \quad \text{(E2)}$$

Note that when $\beta_2 \neq \beta$ the order-parameter functions are no longer symmetric with respect to the exchange $t \to -t$. The probability of jumping to the configurations with energy E_2 (the equilibrium energy at inverse temperature β_2) with typical rate is then

$$\exp[\beta E(\beta)/2 + \overline{[\ln\hat{T}]} - \beta E(\beta_2)/2 + S(\beta_2)] \quad \text{(E3)}$$

and the derivative of the logarithm of the above expression with respect to β_2 is

$$\frac{\beta}{2}\int_{-\mathcal{T}}^{+\mathcal{T}}\hat{R}_{1,f[C]}(t, \mathcal{T})dt - \frac{\beta}{4}f(1). \quad \text{(E4)}$$

I have not solved numerically the case $\beta_2 \neq \beta$, but the evaluation of the derivative at $\beta_2 = \beta$ allows to characterize qualitatively the total logarithm of the rate for $\beta_2 \neq \beta$ in the neighborhood of β. This quantity turns out to be negative when evaluated on the numerical solution at $\beta = 1.695$ with $\beta_2 = \beta$ meaning that the total logarithm of the rate is larger considering configurations with energies larger than equilibrium, which is the key result at bases of the discussion of Sec. I A 3.

Note that consistently in the ergodic phase expression (E4) vanishes in the $\mathcal{T} \to \infty$ limit because the solutions satisfy Eqs. (77). Instead, as we just saw, for $T < T_d$ the above quantity is negative for $\beta_2 = \beta$ also in the $\mathcal{T} \to \infty$ limit and should vanish for some $\beta_2 = \beta_{\text{max}} < \beta$ corresponding to a maximum at a higher energy (and temperature) with $\beta_{\text{max}} \to \beta$ as $T \to T_d^-$.

1. Integrals

We consider the expression of the sum of the elements of a generic object $\boldsymbol{B(ab)}$ in terms of its components,

$$\int \boldsymbol{B(ab)}d\boldsymbol{a}d\boldsymbol{b}. \quad \text{(E5)}$$

We specialize to the case in which $\boldsymbol{B}(ab)$ has the structure of the order parameter $\boldsymbol{Q}(ab)$ as discussed in Sec. B. In particular, we have for the $m + m'$ static replicas:

$$B(\alpha, \beta) = \delta_{\alpha\beta}C_B(\mathcal{T}, \mathcal{T}) \tag{E6}$$

similarly the n dynamical replicas are correlated only with the static replicas associated to the initial and final conditions and their correlations have a RS form. We define

$$\boldsymbol{A}(\boldsymbol{a}) \equiv \int \boldsymbol{B}(\boldsymbol{ba})d\boldsymbol{b}. \tag{E7}$$

For $\boldsymbol{a}$ corresponding to any of the $m + m'$ static replicas other than those fixing the initial and final conditions, we simply have

$$\boldsymbol{A}(\boldsymbol{a}) = C_B(\mathcal{T}, \mathcal{T}). \tag{E8}$$

For $\boldsymbol{a}$ corresponding to either one of the two static replicas controlling the initial and final conditions, we have

$$\boldsymbol{A}(\mathcal{T}) = \int \boldsymbol{B}(\mathcal{T}, \boldsymbol{b})d\boldsymbol{b}$$
$$= C_B(\mathcal{T}, \mathcal{T}) + C_B(\mathcal{T}, -\mathcal{T}) + n \int \hat{R}_{2,B}(\mathcal{T}, t)dt. \tag{E9}$$

For $\boldsymbol{a}$ given by the η component of one of the dynamical replicas, we have

$$\hat{a}(t) = \hat{R}_{1,B}(t, \mathcal{T}) + \hat{R}_{1,B}(t, -\mathcal{T})$$
$$+ \int \hat{X}_B(t, t')dt' + (n-1) \int \hat{X}_B^{dt}(t, t')dt'. \tag{E10}$$

Putting everything together one finds

$$\int \boldsymbol{B}(ab)d\boldsymbol{a}d\boldsymbol{b}$$
$$= \int \boldsymbol{A}(\boldsymbol{a})d\boldsymbol{a} = (m + m')C_B(\mathcal{T}, \mathcal{T})$$
$$+ n\left(\int \hat{R}_{1,B}(t, -\mathcal{T})dt + \int \hat{R}_{1,B}(t, \mathcal{T})dt \right.$$
$$+ \int \hat{R}_{2,B}(\mathcal{T}, t)dt + \int \hat{R}_{2,B}(-\mathcal{T}, t)dt$$
$$+ \left. \int \hat{X}_B(t, t')\,dt\,dt' + (n-1) \int \hat{X}_{B^{dt}}(t, t')\,dt\,dt' \right). \tag{E11}$$

APPENDIX F: THE ENERGY

The instantaneous energy on the trajectory reads

$$E(t) \equiv \sum_{p=1}^{\infty} \sum_{i_1 < \cdots < i_p} \overline{J_{i_1\ldots i_p}[\langle s_{i_1}(t)\ldots s_{i_p}(t)\rangle]}. \tag{F1}$$

Exploiting the fact that the J's are Gaussian random variables through an integration by part one easily obtains

$$e(t) \equiv \frac{E(t)}{N}$$
$$= -\frac{\beta}{2}\left(C_{f[C]}(t, -\mathcal{T}) + C_{f[C]}(t, \mathcal{T}) \right.$$

$$+ \int_{-\mathcal{T}}^{+\mathcal{T}} \hat{R}_{2,f[C]}(t, t')dt'$$
$$+ \left. (n-1) \int_{-\mathcal{T}}^{+\mathcal{T}} \hat{R}_{2,f[C^{dt}]}(t, t')dt' \right). \tag{F2}$$

In the special case of the pure p spin, i.e., $f(x) = x^p$, this leads to a simple relationship between the energy and $\mu(t)$:

$$e(t) = \frac{1}{p\beta}(1 - \mu(t)). \tag{F3}$$

The above can be shown using Eq. (C1) at equal times and using the property that

$$f'(x) + f''(x)x = pf'(x), \quad \text{for } f(x) = x^p, \tag{F4}$$

to make a connection with expression (F2).

APPENDIX G: THE TRANSITION RATE IN THE FREE CASE

In the free case, the expression for the replicated logarithm transition rate (B1) simplifies considerably due to $\beta = 0$ and $\boldsymbol{\Lambda}(ab) = 0$. In the RS ansatz, one obtains

$$n\left(\frac{1}{2} \int dt\, \hat{\mu}(t) + \frac{\Gamma_0}{2} \int \mu(t)dt \right.$$
$$+ \left. \int \frac{d\sigma}{\sqrt{2\pi}} e^{-\frac{\sigma^2}{2}} \int \frac{d\tau}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}} \ln Z(\sigma, \tau) \right),$$
$$Z(\sigma, \tau) = \int_{q(-\mathcal{T})=\sigma}^{q(\mathcal{T})=\tau} [dq] \exp\left[\int dt\left(-\frac{\dot{q}^2}{4\Gamma_0} \right. \right.$$
$$\left. \left. -\frac{\Gamma_0}{4}\mu^2 q^2 - \frac{1}{2}\hat{\mu}q^2 \right) \right]. \tag{G1}$$

We recognize the path integral representation of the Harmonic oscillator that is usually written as

$$Z(\sigma, \tau) = \int_{q(-\mathcal{T})=\sigma}^{q(\mathcal{T})=\tau} [dq] \exp\left[-\int dt\left(\frac{1}{2}m\dot{q}^2 + \frac{1}{2}m\omega^2 q^2 \right) \right]$$

with the identification $m = 1/(2\Gamma_0)$ that leads to the same $\dot{q}^2$ factor and the same $[dq]$. As discussed in classic textbooks the above path integral is ill defined. This is easily seen switching to a frequency representation where it is ultraviolet divergent as $\ln Z \propto \int^\infty dk \ln(\omega^2 + k^2)$. The actual quantity $Z(\sigma, \tau)$ is finite because the differential $[dq]$ includes a prefactor diverging as $1/\Delta t$ as we have seen in section (A 2). A careful computation leads to the following expression for $\ln Z(\sigma, \tau)$ (see Eq. (2.23) in Zinn-Justin's [44] or Eq. (13.45) in Parisi's [45]):

$$\frac{1}{2} \ln \frac{m\omega}{2\pi \sinh \omega\tau} - \frac{m\omega}{2 \sinh \omega\tau}[\cosh \omega\tau(\sigma^2 + \tau^2) - 2\sigma\tau],$$

$$m = \frac{1}{2}, \quad \omega = \sqrt{a}, \quad \tau \equiv\equiv t_{\text{fin}} - t_{\text{in}} = 2\mathcal{T},$$

where we have fixed $\Gamma_0 = 1$ and used the results $\mu = 1$ and $a \equiv 1 + 2\hat{\mu}$ obtained in Sec. II C. Performing the averages over σ and τ, we finally obtain

$$\frac{1}{N}\overline{[\ln \hat{T}]} = -\frac{1}{2}\ln 2\pi - \frac{1}{2} + (1 + \hat{\mu})\mathcal{T} + \frac{1}{4}\ln \frac{-\hat{\mu}}{2}. \tag{G2}$$

[1] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, Nat. Commun. **5**, 1 (2014).

[2] V. Ros, G. Ben Arous, G. Biroli, and C. Cammarota, Phys. Rev. X **9**, 011003 (2019).

[3] S. Sarao Mannelli, G. Biroli, C. Cammarota, F. Krzakala, P. Urbani, and L. Zdeborová, Phys. Rev. X **10**, 011057 (2020).

[4] M. Mézard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications* (World Scientific, Singapore, 1987), Vol. 9.

[5] T. R. Kirkpatrick, D. Thirumalai, and P. G. Wolynes, Phys. Rev. A **40**, 1045 (1989).

[6] P. G. Wolynes and V. Lubchenko, *Structural Glasses and Supercooled Liquids: Theory, Experiment, and Applications* (Wiley, Hoboken, New Jersey, 2012).

[7] T. Rizzo, Critical phenomena in glasses, in *Handbook of Materials Modeling: Methods: Theory and Modeling*, edited by W. Andreoni and S. Yip (Springer, Cham, 2020), pp. 183–203.

[8] T. R. Kirkpatrick and D. Thirumalai, Phys. Rev. B **36**, 5388 (1987).

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

[10] J. H. Gibbs and E. A. DiMarzio, J. Chem. Phys. **28**, 373 (1958).

[11] G. Adam and J. H. Gibbs, J. Chem. Phys. **43**, 139 (1965).

[12] M. Mézard and G. Parisi, Glasses and replicas, in *Structural Glasses and Supercooled Liquids* (John Wiley & Sons, Ltd, Hoboken New Jersey, 2012), Chap. 4, pp. 151–191.

[13] R. Monasson, Phys. Rev. Lett. **75**, 2847 (1995).

[14] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, Annu. Rev. Condens. Matter Phys. **8**, 265 (2017).

[15] W. Kob and H. C. Andersen, Phys. Rev. E **51**, 4626 (1995).

[16] W. Kob, J. Condens. Matter Phys. **11**, R85 (1999).

[17] W. Kob, C. Donati, S. J. Plimpton, P. H. Poole, and S. C. Glotzer, Phys. Rev. Lett. **79**, 2827 (1997).

[18] E. Flenner, H. Staley, and G. Szamel, Phys. Rev. Lett. **112**, 097801 (2014).

[19] L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, D. El Masri, D. L'Hôte, F. Ladieu, and M. Pierno, Science **310**, 1797 (2005).

[20] G. Biroli, J.-P. Bouchaud, A. Cavagna, T. S. Grigera, and P. Verrocchio, Nat. Phys. **4**, 771 (2008).

[21] S. Albert, T. Bauer, M. Michl, G. Biroli, J.-P. Bouchaud, A. Loidl, P. Lunkenheimer, R. Tourbot, C. Wiertel-Gasquet, and F. Ladieu, Science **352**, 1308 (2016).

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

[23] G. Folena, S. Franz, and F. Ricci-Tersenghi, Phys. Rev. X **10**, 031045 (2020).

[24] A. Altieri, G. Biroli, and C. Cammarota, J. Phys. A: Math. Theor. **53**, 375006 (2020).

[25] M. Baity-Jesi, A. Achard-de Lustrac, and G. Biroli, Phys. Rev. E **98**, 012133 (2018).

[26] M. Baity-Jesi, G. Biroli, and C. Cammarota, J. Stat. Mech. (2018) 013301.

[27] M. R. Carbone, V. Astuti, and M. Baity-Jesi, Phys. Rev. E **101**, 052304 (2020).

[28] I. Hartarsky, M. Baity-Jesi, R. Ravasio, A. Billoire, and G. Biroli, J. Stat. Mech. (2019) 093302.

[29] V. Ros, G. Biroli, and C. Cammarota, Europhys. Lett. **126**, 20003 (2019).

[30] D. A. Stariolo and L. F. Cugliandolo, Europhys. Lett. **127**, 16002 (2019).

[31] D. A. Stariolo and L. F. Cugliandolo, Phys. Rev. E **102**, 022126 (2020).

[32] V. Ros, G. Biroli, and C. Cammarota, SciPost Phys. **10**, 002 (2021).

[33] T. Rizzo, Europhys. Lett. **106**, 56003 (2014).

[34] T. Rizzo and T. Voigtmann, Europhys. Lett. **111**, 56008 (2015).

[35] T. Rizzo, Phys. Rev. B **94**, 014202 (2016).

[36] T. Rizzo, Philos. Mag. **96**, 636 (2016).

[37] T. Rizzo and T. Voigtmann, Phys. Rev. Lett. **124**, 195501 (2020).

[38] H. Sompolinsky and A. Zippelius, Phys. Rev. B **25**, 6860 (1982).

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

[40] A. Crisanti and H.-J. Sommers, Z. Phys. B: Condens. Matter **87**, 341 (1992).

[41] J. Kurchan, T. Maimbourg, and F. Zamponi, J. Stat. Mech. (2016) 033210.

[42] T. Maimbourg, J. Kurchan, and F. Zamponi, Phys. Rev. Lett. **116**, 015902 (2016).

[43] A. Manacorda, G. Schehr, and F. Zamponi, J. Chem. Phys. **152**, 164506 (2020).

[44] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Oxford Science Publications, Oxford, 2002).

[45] G. Parisi, *Statistical Field Theory* (Addison-Wesley, Reading, Massachusetts, 1988).

[46] T. Rizzo, arXiv:2012.09556.

[47] T. Castellani and A. Cavagna, J. Stat. Mech. (2005) P05012.

[48] F. Caltagirone, U. Ferrari, L. Leuzzi, G. Parisi, F. Ricci-Tersenghi, and T. Rizzo, Phys. Rev. Lett. **108**, 085702 (2012).

[49] U. Ferrari, L. Leuzzi, G. Parisi, and T. Rizzo, Phys. Rev. B **86**, 014204 (2012).

[50] Y. Saad, *Iterative Methods for Sparse Linear Systems* (SIAM, Philadelphia, 2003).

[51] Y. Saad and M. H. Schultz, SIAM J. Sci. Stat. Comput. **7**, 856 (1986).

[52] S. Franz and G. Parisi, Phys. Rev. Lett. **79**, 2486 (1997).

[53] A. V. Lopatin and L. Ioffe, Phys. Rev. B **60**, 6412 (1999).

[54] A. V. Lopatin and L. Ioffe, Phys. Rev. Lett. **84**, 4208 (2000).