Competition between energy- and entropy-driven activation in glasses

Matthew R. Carbone and Marco Baity-Jesi

Phys. Rev. E **106**, 024603 — Published 2 August 2022

DOI: 10.1103/PhysRevE.106.024603

Competition between Energy- and Entropy-Driven Activation in Glasses

Matthew R. Carbone[1] and Marco Baity-Jesi[2,*]

[1] *Computational Science Initiative, Brookhaven National Laboratory, Upton, New York 11973, USA*
[2] *Eawag, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland*

(Dated: July 19, 2022)

In simplified models of glasses we clarify the existence of two different kinds of coexisting activated dynamics, with one of the two dominating over the other. One is the energy barrier hopping that is typically used to understand activation, and the other, which we call entropic activation, is driven by the scarcity of convenient directions in phase space. When entropic activation dominates, the height of the energy barriers is no longer the primary factor governing the system's slowdown. In our analysis, dominance of one mechanism over the other depends on temperature and the shape of the density of states. We also find that at low temperatures a phase transition between the two kinds of activation can occur. Our observations are used to provide a scenario that can harmonize the facilitation and thermodynamic pictures of the slowdown of glasses into a single description.

I. INTRODUCTION

Glasses are inherently slow systems. Their slowness can be captured by mean-field (MF) theory, which recently motivated a series of breakthroughs that allowed for a deeper understanding their sluggish behavior [1–3]. However, MF theory predicts divergences of the relaxation time that do not occur in real systems, because it does not capture relaxation mechanisms that appear in low-dimensions. These are generically called *activation*, and they are most often pictured as the hopping of energy barriers [4, 5]: since in MF the barriers diverge with the system size N, a simple argument is that activation in MF cannot occur because barriers cannot be hopped [6].

However, activation can be studied in MF models by restricting to finite-N and long times [7, 8]. Several works focused on comparing the dynamics of simple MF models, such as the Random-Energy Model (REM) [9], to the Trap Model (TM) [10–12], to establish whether the barrier hopping dynamics can be assimilated to jumping between traps with a fixed threshold energy [13–21]. Other works studied the saddles connecting minima [22–25], extensions of the Franz-Parisi potential [26], or path-integral approaches to study the dynamics between different minima [27].

Comparisons with the TM were also performed in models with a trivial landscape, such as the Step [28] or Funnel models [29], where it was shown that entropic effects can lead to Trap-like activation, if instead of considering basins in phase space we construct them dynamically. Indications of entropic effects in long-time dynamics was also found in less idealized systems, such as the p-spin model [30, 31], finite-connectivity Step models [32], or even $3D$ Lennard-Jones mixtures between the dynamical temperature T_d and the onset temperature T_o [33] and experimental metallic glasses [34]. In these works, however, the underlying framework is either that activation only exists as barrier hopping [4, 11]; either that it is an entropic effect that can be assimilated to barrier hopping [28, 29]; or it is a transient behavior which eventually turns into hopping [7, 32, 35]. This is hard to reconcile with other pictures of the dynamical arrest of glasses, such as facilitation [36, 37], that argue that the landscape (energy barriers) is not crucial to explain the slowness of glasses, which should instead be attributed to kinetic constraints.

Here, we clarify the nature of these entropic effects, showing how under the right lens they can be used to unify the landscape with the facilitation pictures. We take a paradigmatic model of glasses, the REM [9], and show that both energy- and entropy-driven activation mechanisms coexist. When energy-driven activation dominates, the system's slowness is driven by the energy barrier separating basins, while when entropic activation dominates, the height of the barrier becomes unimportant, and the slowness is instead driven by the scarcity of convenient directions. In our analysis, the dominance of one mechanism or the other depends on the density of states $\rho(E)$, and on the temperature $k_\mathrm{B}T \equiv \beta^{-1}$, which is the control parameter of a non-equilibrium phase transition between the two different activated regimes (we set $k_\mathrm{B} = 1$ in this work).

II. MODELS

In the REM, we have a system with N spins, $s_i = \pm 1$. The dynamics takes place by flipping a single spin at a time: from any configuration one can reach N neighboring states. The energy of a configuration is independent of the configuration itself, and is drawn from a probability distribution $\rho(E)$. Two choices of $\rho(E)$ are used in literature. The initial formulation of the REM used a Gaussian energy distribution [9],

$$\rho_g(E) = \frac{1}{\sqrt{2\pi N}} e^{-\frac{E^2}{2N}}, \tag{1}$$

* marco.baityjesi@eawag.ch

while more recent efforts also considered the Exponential REM (EREM) [20], which has

$$\rho_e(E) = \beta_c e^{\beta_c E}\Theta(-E)\,, \tag{2}$$

where we set $\beta_c = 1$. These models have a transition from a paramagnetic to a glassy phase at $\beta_g = \sqrt{2\ln 2}$ and $\beta_e = \beta_c$, respectively.

III. CHARACTERISTIC ENERGIES

A. Threshold Energy

In both models, we can define a threshold energy, $E_{\rm th}$, analogous to that of the p-spin model [38], above which no barriers are typically found. We calculate it as the energy for which the probability that a neighboring configuration has a lower energy is $1/N$ [20, 39],

$$\frac{1}{N} = \int_{-\infty}^{E_{\rm th}} \rho(E)dE\,. \tag{3}$$

From Eq. (3) one obtains that, to leading order, the threshold energies of REM and EREM are

$$E^g_{\rm th} = -\sqrt{2N\ln N}\,, \tag{4}$$

$$E^e_{\rm th} = -\frac{1}{\beta_c}\ln N\,. \tag{5}$$

This definition ensures that there are asymptotically no barriers above $E_{\rm th}$, and descends from the mutual independence of neighboring configurations in the (E)REM.

After long times, the system typically finds itself in a configuration with an extensively deep energy [20, 39], which we call trap (or basin). To transition from one trap/basin to another, the system needs to climb to the threshold. Energetic activated dynamics is mainly driven by the jumps on these energy barriers, in a manner that is analogous to what happens in the Trap Model, in which the time spent in a trap of energy E follows and Arrhenius law, $\tau \sim \exp[\beta(E_{\rm th} - E)]$. This was indeed found to be the case, by looking at the limiting values of the aging functions and comparing them with the predictions of the TM [18, 20].

B. Attractor Energy

We can also define another characteristic energy, which stems from toy models which represent a purely entropic kind of activation [28, 29, 40]. This quantity was called threshold energy in Refs. 28 and 29. As we will show, it is quantitatively and qualitatively different from the threshold energy, so we henceforth refer to it as the *attractor energy*, $E_{\rm a}$. This quantity was first defined in Ref. 41, as the energy such that the energy of the next step is, on average, neither higher nor lower. However, we find it more intuitive to work with the definition given in Ref. 28, according to which $E_{\rm a}$ is the energy at which the probability $P_\uparrow(E)$ of increasing the energy at the next step is equal to the probability $P_\downarrow(E)$ of decreasing it:

$$P_\uparrow(E_{\rm a}) \equiv P_\downarrow(E_{\rm a})\,. \tag{6}$$

$P_\uparrow(E)$ (or $P_\downarrow(E)$) is calculated by considering the probability of finding neighbors with a higher (or lower) energy E', and the transition rate $w(E \to E')$ telling whether a move towards such an energy is accepted. We consider Monte Carlo Metropolis dynamics, $w(E \to E') = \min(1, \exp(-\beta(E' - E)))$. As for the distribution of neighboring energies, it is particularly easy to calculate in the (E)REM, because the energy distribution of the neighbors is the same as that of the whole system. Therefore. one has

$$P_\downarrow(E) = \int_{-\infty}^{E} dE'\rho(E')w(E \to E') = \int_{-\infty}^{E} dE'\rho(E')\,, \tag{7}$$

$$P_\uparrow(E,\beta) = \int_{E}^{\infty} dE'\rho(E')w(E \to E') = \int_{E}^{\infty} dE'\rho(E')e^{-\beta(E'-E)}\,, \tag{8}$$

where we made it explicit that $P_\uparrow$ depends on β. We note that, since the configurations in the (E)REM are i.i.d., $P_\downarrow(E)$ is the cumulative distribution of $\rho(E)$, so we can write $P_\downarrow(E_{\rm th}) = \frac{1}{N}$.

Since $P_\uparrow(E,\beta)$ varies with β while $P_\downarrow(E)$ does not, also $E_{\rm a}$ must vary with temperature. For the Gaussian REM, we find that the attractor energy is

$$E^g_{\rm a} = -\frac{N\beta}{2}\,, \tag{9}$$

whereas for the EREM it is

$$E^e_{\rm a} = \begin{cases} \frac{1}{\beta-\beta_c}\ln\left(\frac{2\beta_c-\beta}{\beta_c}\right) & ; \quad \text{if } 0 < \beta < 2\beta_c\,, \\ -\infty\,; & \text{if } \beta > 2\beta_c\,. \end{cases} \tag{10}$$

The attractor energy for the EREM is the same that is found in Ref. 28 in a model which corresponds to a global

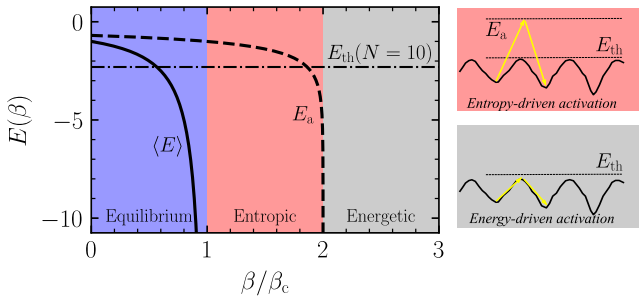

FIG. 1. **Left**: Phase diagram of the EREM, with the associated characteristic energies and dynamical regimes. For $\beta < \beta_{\rm c}$ we have the equilibrium phase. The solid line is the equilibrium energy $\langle E\rangle$. At $\beta = 2\beta_{\rm c}$ we have the transition from the entropically to the energetically activated regime. The dashed line is the attractor energy. The horizontal dashed-dotted line represents $E_{\rm th}(N=10)$. $E_{\rm th}$ exists $\forall\beta$ but is only relevant for $\beta > 2\beta_{\rm c}$. **Right**: two diagrams showing how activated dynamics qualitatively takes place in each of the regimes.

dynamics in the EREM. [1] This is due to the fact that in the EREM the distribution of neighboring energies (local) is equal to $\rho_e(E)$ (global).

Note that $E_{\rm a}$ is not directly related to the equilibrium energy $\langle E\rangle$. For example, in the EREM, there is no equilibrium energy from $\beta \geq \beta_{\rm c}$ (the equilibrium energy diverges to $-\infty$ at $\beta = \beta_{\rm c}$), [2] but $E_{\rm a}^e$ still is finite for $\beta_{\rm c} \leq \beta < 2\beta_{\rm c}$. We depict this in Fig. 1. In other words, the system still is *attracted* towards finite energies even though the equilibrium energy is divergingly negative. In fact, it is straightforward to check that when $E < E_{\rm a}$ (or $E > E_{\rm a}$) the next configuration during the dynamics is more likely to have a higher (or lower) energy. Intuitively, what happens is that even when neighbors with lower energy exist, they are to hard to find, and the system will find it more convenient to just increase the energy. However, the system will immediately abandon high-energy configurations, while it will stay a very long time in the low-energy ones.

IV. ENERGY- AND ENTROPY-DRIVEN ACTIVATION

A. Energy-driven

We call *energy-driven* the typical thermally activated dynamics consisting of hopping barriers with an Arrhenius rate. Since in the (E)REM the overwhelming majority of the barriers are at $E = E_{\rm th}$ (see *e.g.* Ref. 21), this kind of activation resembles the dynamics of the Trap Model [10, 11, 18, 20].

B. Entropy-driven

Contrary to $E_{\rm th}$, $E_{\rm a}$ can also exist in systems with a single energy basin, such as the Step [28] or the Funnel model [29]. In these systems, despite the absence of energy-driven activation, the dynamics is a renewal process, provided that one identifies the traps dynamically, as all the configurations visited while $E < E_{\rm a}$ [28]. With this construction, the relationship between aging functions and trapping time distributions is the one predicted by the TM [28], and the typical time scales grow exponentially in N [29]. In other words, $E_{\rm a}$ identifies an activated dynamics which is *not* driven by the height of the energy barriers (as there are none), but which shares many signatures of energy-driven activation. This kind of activated dynamics is instead driven by the scarcity of convenient directions, and was therefore called entropy-driven [28].

C. Interplay between the two mechanisms

In the (E)REM, we can study the interplay between these two mechanisms. Since $E_{\rm th}$ is the minimal height at which the system must go in order to leave a trap, the entropic mechanism is not expected to play a role when $E_{\rm a} < E_{\rm th}$. We see that, for sufficiently large sizes, in the REM $E_{\rm a}^g(\beta) < E_{\rm th}^g\ \forall\beta$. Therefore, we expect that activation in the REM is purely energy-driven.

In contrast, in the EREM we have different behaviors depending on β. When $\beta > 2\beta_{\rm c}$, we have $E_{\rm a} < E_{\rm th}$, so the slow dynamics should be energy-driven. When instead $\beta_{\rm c} < \beta < 2\beta_{\rm c}$ (excluding a range of β close to $2\beta_{\rm c}$ which shrinks with increasing N), we have $E_{\rm a} > E_{\rm th}$. This indicates that even when the system manages escaping a trap, reaching $E_{\rm th}$, it will likely keep going up in energy, attracted towards $E_{\rm a}$. Thus, the height of the barrier, $(E_{\rm th} - E)$, is not that important. The reason for this is that even though in this regime there are directions in phase space which would decrease the energy, they are too rare, and the system would rather keep increasing its energy than "invest time" looking for a descending direction. In Fig. 1 we show a diagram of how activation takes place in each of the two out-of-equilibrium phases.

Since activated dynamics is relevant in the limit of large but finite N, we can work out the transition inverse temperatures β_* by setting $E_{\rm a}(\beta_*, N) = E_{\rm th}(N)$. This gives us the N-dependent transitions

$$\beta_*^g = 2\sqrt{\frac{2\ln N}{N}} \overset{N\to\infty}{\longrightarrow} 0\,, \tag{11}$$

$$\beta_*^e = \beta_{\rm c}\frac{2\ln N + W(-\frac{\ln N}{N})}{\ln N} \overset{N\to\infty}{\longrightarrow} 2\beta_{\rm c}\,, \tag{12}$$

[1] By global dynamics we mean that any number of spins is flipped at each time step, in contrast with the single-flip dynamics considered in this paper. With global dynamics local minima disappear.

[2] In the EREM we have $\langle E\rangle = \frac{\int_{-\infty}^0 e^{(\beta_{\rm c}-\beta)E}EdE}{\int_{-\infty}^0 e^{(\beta_{\rm c}-\beta)E}dE} = -\frac{\beta_{\rm c}}{\beta_{\rm c}-\beta}$.

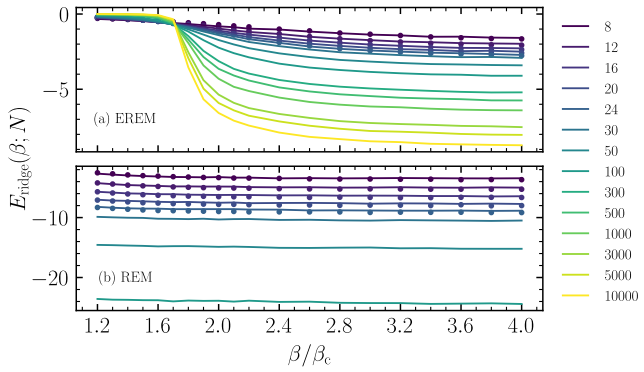

FIG. 2. Median $E_{\rm ridge}(\beta;N)$ in the EREM (a) and REM (b). Different lines stand for different system sizes. Here and in Fig. 3, results from memory-1 dynamics are shown in solid lines, and results from full-memory dynamics ($N < 30$) are shown as markers.

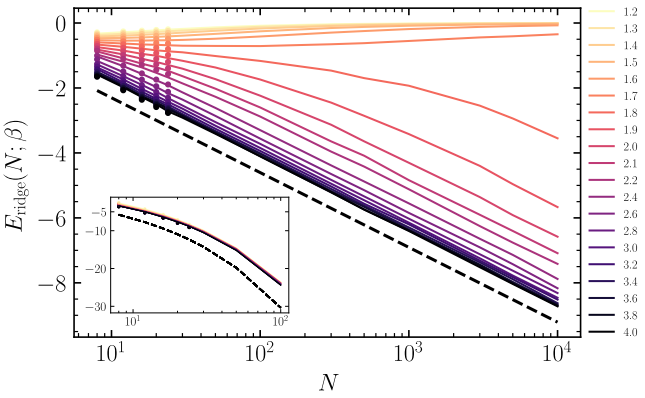

FIG. 3. Median $E_{\rm ridge}$ as a function of N in the EREM and REM (inset). Different lines correspond to different values of β. The black dashed line represents $E_{\rm th}(N)$ for each model, according to Eqs. (4) and (5).

where $W(z)$ is the Lambert function. In the Gaussian REM the entropic phase disappears for increasing N. In the EREM, instead, the transition stays at a finite temperature, and we have entropy-driven activation at low β, and energy-driven activation at high β.

V. RIDGE ENERGY AND PHASE TRANSITION

We run Monte Carlo simulations (details in App. A) to verify this transition from an energy-driven phase at high β, to an entropy-driven phase at lower β. Since REM and EREM do not allow for exact simulations at large system sizes, for sizes $N \geq 30$ we rely on modified dynamics, where every time that a new configuration is visited, all its neighbors' energies, except that of the last-visited configuration, are drawn anew. We call this *memory-1 dynamics*, and elaborate more on it in App. B.

Our intent is now to relate this transition to physical observables. We measure the ridge energy, $E_{\rm ridge}$, defined as the highest-reached energy during each basin transition, *i.e.* in each of the time intervals during which $E(t) > E_{\rm th}$. With energy-driven activation, we expect that $E_{\rm ridge}(\beta;N)$ will stay close to $E_{\rm th}(N)$, while for entropy-driven activation it will overshoot to higher values. In Fig. 2 we plot the median $E_{\rm ridge}$ as a function of β, for EREM and REM. While in the REM the ridge energies decrease steadily $\forall\beta$, in the EREM we see two distinct entropy-driven (low β) and energy-driven (high β) phases.

As we show in Fig. 3, the ridge energy scales with N in the way we expect. In the EREM at low β we are in the entropic phase, and $E_{\rm ridge} \sim -1 \sim E_{\rm a}$. In the energy-driven phase, instead, $E_{\rm ridge} \sim -\ln(N) \sim E_{\rm th}$. In the REM, where we only have energy-driven activation, the ridge energy follows $E_{\rm th}$ at every β.

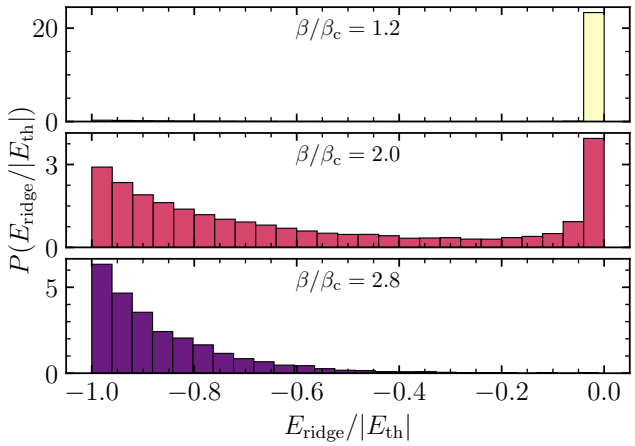

FIG. 4. The density of ridge energies in the EREM model, using memory-1 dynamics, for $N = 10000$.

Through the lens of the median $E_{\rm ridge}$, the transition appears at β slightly smaller than 2, which we understand through two observations. First, β_*^e [Eq. (12)] indicates when $P_\uparrow > P_\downarrow$. However, in order to reach $E_{\rm a}$ from $E_{\rm th}$, the system needs to go through a large number of steps (*i.e.* growing with N) with ascending energy. If $P_\uparrow$ is only slightly larger than $P_\downarrow$ (which is what happens at β slightly lower than 2), this is not enough to accept a sufficiently large number of steps to go all the way up to energies of order 1. Second, this transition from energy- to entropy-driven activation has features of a first-order phase transition. We see this from Fig. 4, where we plot the distribution $P(\frac{E_{\rm ridge}}{|E_{\rm th}|})$ at three different temperatures. At low β it is peaked around 0, while at high β it is peaked around -1, as expected. Around the transition, instead, we see the characteristic two-peak structure of first-order phase transitions.

VI. CONCLUSIONS

A. Summary

We showed that the "threshold energy" calculated dynamically in models with a trivial landscape [28, 29] does not correspond to the typical landscape-based definition of threshold energy, and decided to rename it as *attractor energy*. Threshold and attractor energies define two different kinds of activated dynamics, entropy- and energy-driven, which can coexist, though typically one dominates over the other. Energy-driven activation corresponds to the typical picture of Arrhenius-like basin hopping, whereas entropy-driven activation is not driven by barrier heights, but rather by the scarcity of convenient directions. While at sufficiently low temperatures energetic activation always exists in landscapes with multiple minima, the existence of a higher-T phase where entropy-driven activation dominates depends on the shape of $\rho(E)$. In this entropic activation phase, the attractor energy $E_{\rm a}$, towards which the system is regularly driven, is higher than the threshold energy $E_{\rm th}$, the height of the barriers loses relevance and we have a breakdown of the Arrhenius behavior.

B. Reconsidering the importance of saddles

One consequence is that the study of the transition paths of glasses by searching the lowest saddles, a technically daunting task [22–25], is potentially not informative for the dynamics at temperatures around the glass transition. Calculating the attractor energy, with its comparison to the threshold, is simpler and has the potential of unlocking the true activated nature of the dynamics.

C. Rationalizing observations in other systems

Additionally, entropic activation can elegantly explain an apparently puzzling result recently found by T. Rizzo [27], who calculated, in the spherical p-spin model, the path from one equilibrium low-temperature configuration to another. He found that the maximum energy reached during this transition is considerably higher than $E_{\rm th}$ even though the largest barriers should not exceed $E_{\rm th}$. We see that the same phenomenon occurs in the EREM, in a similar setting (we followed the system from one trap at $E < E_{\rm th}$ to another one), where we showed that the system is pushed towards a higher energy $E_{\rm a}$ by the scarcity of paths with energy close to $E_{\rm th}$. We can thus attribute this behavior to the same entropic effects, and conjecture that at lower temperatures the paths will be close to $E_{\rm th}$ (as per the transition that we found in the EREM).

This is also consistent with recent numerical observations by Stariolo and Cugliandolo in the discrete p-spin model, that the trapping time distributions seem to follow the Step-model predictions better than those of the Trap Model [30, 31]. In particular, they define the traps dynamically by taking, instead of the position of the saddle, the highest point reached during the dynamics. Therefore, they are calculating the traps through the attractor instead of the threshold energy, which explains the observed Step-like behavior. Furthermore, they find that the energy of the ridge is larger than $E_{\rm th}$, which is an indication that the lowest-energy path is not used, just as we find here.

Also in 3D systems, such a Lennard-Jones mixtures, although activated dynamics takes place between $T_{\rm o}$ and $T_{\rm d}$ [4, 5, 42], it was recently shown that it is not dominantly of an energy-driven kind, since, for example, the system is moving at energies significantly higher than the ridges separating metabasins [33]: the height of the barriers separating basins does not play a crucial role in this regime.

D. The nature of activation in different kinds of models

A typical intuition of why MF results do not apply in low dimensions is that the energy barriers diverge with the system size. However, in systems like hard spheres there are no energy barriers, so it seems like beyond-MF effects are not attributed to the right mechanism. It was suggested in several occasions that the low-D correspondent of a MF model can be non-trivial [43–45]. Our analysis suggests that we should often think of activation less as the hopping of energy barriers, and more as a search for convenient directions, which require a collective cooperative behavior that is hard to obtain by randomly moving particles. We should therefore regard activation as a process beyond MF, not because the barriers are diverging, but rather because it involves processes (energetic *and* entropic) which take place on time scales $t \gg N$ (usually $t \sim e^N$).

E. Connection to Facilitation Theory

The dynamical facilitation picture shows that a strong glass-like slowdown can appear in a trivial landscape (without barriers), with the dynamics being slowed down by dynamical constraints [36, 46]. These dynamical constraints are localized in space, so we do not aim at a 1:1 correspondence with a mean-field model (as the (E)REM), but we do note that the entropic barriers that lead to the entropic phase in the EREM are (soft) kinetic constraints. In fact, they mark directions in phase space along which motion is suppressed (with high probability a spin cannot be flipped), and they also appear in the absence of potential energy barriers [28, 29, 41, 47, 48]. As an extra, here, kinetic constraints are not the only slow dynamics mechanism, but they act concurrently with energetic activation; in addition to the fast mechanism of

diffusion towards lower energies. In other words, the phenomenological ingredients of facilitation are present, together with additional mechanisms that should appear in the description of a glass.

Since the dynamical constraints, as well as activation, appear at T_o [4, 36], we suppose a correspondence with the thermodynamic transition in the EREM (at $\beta = \beta_c$). Cooling further, one observes entropy-driven activated dynamics between T_o and T_d [33], and at T_d the dynamics is deemed to become energy-driven [49]. Also in the EREM, entropic activation stays dominant until the transition to the energetic phase at $\beta = 2\beta_c$, suggesting a correspondence between T_d and the entropic-energetic transition we observed in the EREM. However, in order to confirm these speculations we need to be able to observe entropic activation in more complex models.

F. Verifying entropic activation in more complex models

In simple models such as the EREM or the Funnel model, the entropic phase is present both for local and global phase space dynamics [29]. However, this is not necessarily true for more complex models such as the p-spin. A starting point which would allow to study correlated energy levels without changing the overall $\rho(E)$ would be the Correlated REM [21] and the Number Partitioning Problem (NPP) [50]. These two models have different kinds of correlations since, in the first, the basins are smooth, while in the NPP the traps are anti-correlated with their neighbors. In particular, in the NPP it was observed that, at $\beta = 2\beta_c$, there is a transition within the glass phase, associated with a violation of the fluctuation-dissipation theorem, with relevant differences between local and global dynamics. The relationships between our findings and those of Ref. 50 could be studied analytically within the framework of Ref. 51, which studies the Step model in the limit of slowly decorrelating variables.

ACKNOWLEDGMENTS

We thank A. Altieri, V. Astuti, F. Ricci-Tersenghi, T. Rizzo, P. Sollich and D. Tapias for interesting conversations. This work was funded by the Simons Foundation for the collaboration Cracking the Glass Problem (No. 454951 to D.R. Reichman). M. R. C. acknowledges the following support: This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship under Award Number DE-FG02-97ER25308.

Appendix A: Numerical details

1. Monte Carlo Simulations

Our simulations are performed with the Metropolis Monte Carlo algorithm [52]. Depending on the value of β, we use two equivalent implementations, which we call "standard" and "Gillespie". The standard dynamics is a textbook Monte Carlo simulation with Metropolis acceptance criteria (see e.g. Ref. 53 for a detailed description). The run time of the standard procedure is tightly bound, with t_{max} time steps. However, at large β, the rejection rate of standard dynamics is high and it may take many steps for a new configuration to be accepted. We therefore use the Gillespie procedure (which is formally equivalent) is much more efficient. The Gillespie algorithm is a rejectionless method, that computes the time that the system spends in a given configuration, and transitions without rejection to one of the neighbors, according to how probable it is to transition to each neighbor. For a more detailed and didactic explanation of the Gillespie algorithm we refer the reader to Ref. 54. While the Gillespie method this is efficient at large β, it is extremely inefficient at small β. Thus, we set a cutoff of $\beta/\beta_c = 2.3$, such that when $\beta/\beta_c < 2.3$, dynamics are run using the standard procedure, and Gillespie otherwise. We also use $t_{max} = 10^7$ for all calculations.

Note that the standard and Gillespie procedures are formally equivalent [54]. To demonstrate this empirically, we show the mean energy $E(t)$ for various choices of N and β (with $\beta_c = 1$) in the EREM, in Fig. 5.

2. Inherent Structures

To ensure that we only measure transitions between different traps, with memory dynamics we measure the inherent structure (IS) before and after the transition, and keep the transition only if they are different. To measure the IS we take the steepest descent path towards the nearest minimum. Since the system is discrete, we define steepest descent as a trajectory which at each step goes to its lowest-energy neighbor, until a minimum is reached.

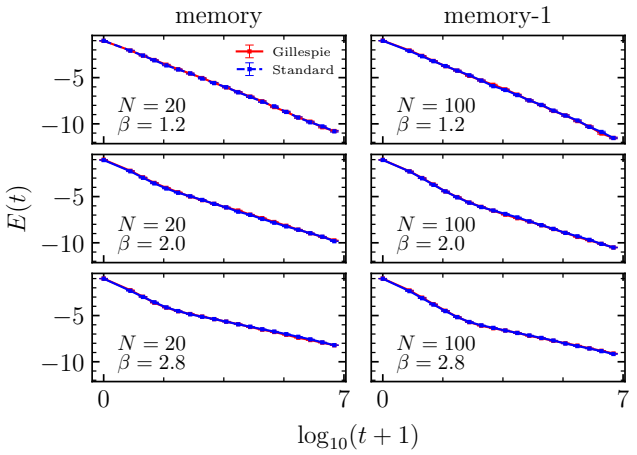

FIG. 5. Energy as a function of time, in simulations with Gillespie (red) and regular Monte Carlo (blue) dynamics. Each row is a different inverse temperature. On the first column we have the full memory dynamics, while on the second we have the memory-1 dynamics (App. B).

3. Threshold Energy

As already pointed out in Ref. 20, the subleading corrections in the threshold energy of the REM are large (around 13% in the largest sizes). Therefore, in our REM simulations we calculated the threshold energy numerically. For the EREM simulations, we used Eq. (4).

Appendix B: Memory and Memory-1 dynamics

In REM and EREM, the energy of each configuration is a fixed random variable. This means that, for N spins, there are 2^N energies sampled from $\rho(E)$, each of which is permanently paired to some configuration. Therefore, in order to perform a long simulation in these models, we need to store the energy of all the 2^N states, to ensure that if a configuration is visited twice its energy has not changed. Storing 2^N double precision floating point numbers is expensive in terms of memory, and limits the largest system sizes that we can simulate. This is why, in order to simulate $N \geq 30$, instead of storing the all the 2^N energies, we only stored the last visited one, and sampled anew the remaining $N-1$ neighbors. We call this dynamics *memory-1*, in contrast with the *memory* dynamics which stores all the energies throughout the whole simulation.

This simplification neglects loops in the dynamics, which for large N are arguably rare, and does not allow the system to directly return to configurations visited more than one step earlier. The latter can be seen as an advantage, since we want to wash back and forth motion out of the dynamics we are measuring [4, 20]. Since the REM and EREM dynamics is a renewal process [20], we can expect that anyhow after some time the previously visited phase space should be forgotten.

An additional difference between memory and memory-1 dynamics is that the latter does not suffer from finite size effects descending from the phase space being of limited size: with memory dynamics there exists a lowest-reachable energy, while with memory-1 it is always possible to reach a lower energy. In other words, memory-1 dynamics suffer less from finite-size effects than the exact dynamics, and in any case this kind of effects does not affect the calculation of E_{ridge}. In Fig. 6 we show the comparison between the two dynamics for varying β and N. At all temperatures, the difference between the two dynamics decreases as N grows. Both dynamics present finite-size effects, which decrease as the system becomes larger.

[1] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, Nature Communications **5**, 3725 (2014), arXiv:1404.6809.
[2] T. Maimbourg, J. Kurchan, and F. Zamponi, Physical review letters **116**, 015902 (2016).
[3] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, Annual Review of Condensed Matter Physics **8**, 265 (2017).
[4] B. Doliwa and A. Heuer, Physical Review E **67**, 031506 (2003), arxiv:cond-mat/0209139.
[5] B. Doliwa and A. Heuer, Physical Review E **67**, 030501(R) (2003).
[6] F. Arceri, F. P. Landes, L. Berthier, and G. Biroli, "Glasses and aging: a statistical mechanics perspective," (2020), arXiv:2006.09725.
[7] A. Crisanti and F. Ritort, EPL (Europhysics Letters) **52**, 640 (2000).
[8] A. Crisanti and F. Ritort, EPL (Europhysics Letters) **51**, 147 (2000).
[9] B. Derrida, Phys. Rev. Lett. **45**, 79 (1980).
[10] J. C. Dyre, Phys. Rev. Lett. **58**, 792 (1987).
[11] J. Bouchaud, J. Phys. I France **2**, 1705 (1992).
[12] J.-P. Bouchaud and D.S. Dean, J. Phys. I France **5**, 265 (1995), arXiv:cond-mat/9410022.
[13] G. Ben Arous, A. Bovier, and V. Gayrard, Phys. Rev. Lett. **88**, 087201 (2002).
[14] G. Ben-Arous, A. Bovier, and J. Černỳ, Communications in Mathematical Physics **282**, 663 (2008).
[15] V. Gayrard, Annales Henri Poincaré **17**, 537 (2016).
[16] V. Gayrard and O. Gün, Markov Processes and Related Fields **22**, 165 (2016), arXiv:1501.02986.
[17] J. Černý and T. Wassmer, Probab. Theory Relat. Fields **167**, 253 (2017).
[18] V. Gayrard, Probab. Theory Relat. Fields (2018), 10.1007/s00440-018-0873-6, arXiv:1602.06081.
[19] V. Gayrard and L. Hartung, in *Statistical Mechanics of Classical and Disordered Systems*, edited by V. Gayrard, L.-P. Arguin, N. Kistler, and I. Kourkova (Springer

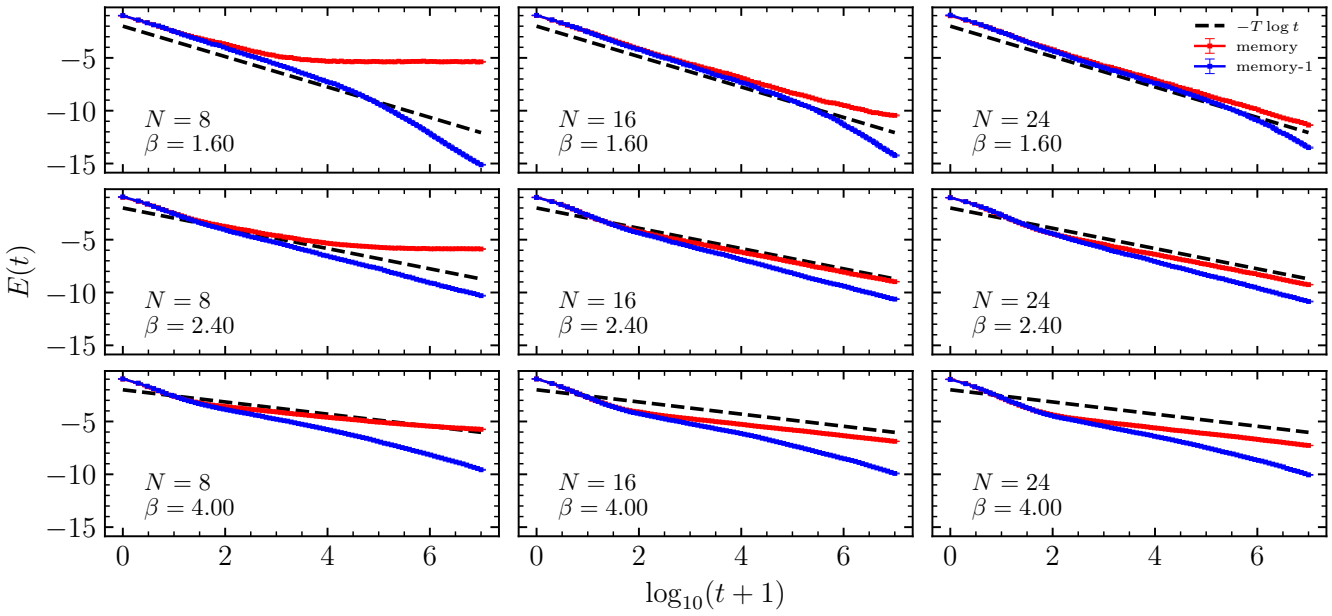

FIG. 6. Energy as a function of time for memory and memory-1 dynamics. The dashed curve is the slope $-T\ln(t)$ that one would expect in the infinite-size limit.

International Publishing, Cham, 2019) pp. 111–170, arXiv:1903.04446.

[20] M. Baity-Jesi, G. Biroli, and C. Cammarota, J. Stat. Mech.: Theory Exp , 013301 (2018), arXiv:1708.03268.

[21] M. Baity-Jesi, A. Achard-deLustrac, and G. Biroli, Phys. Rev. E **98**, 012133 (2018), arXiv:1805.04581.

[22] V. Ros, G. Ben Arous, G. Biroli, and C. Cammarota, Physical Review X **9**, 011003 (2019), arXiv:1804.02686.

[23] V. Ros, G. Biroli, and C. Cammarota, EPL (Europhysics Letters) **126**, 20003 (2019), arXiv:1809.05440.

[24] V. Ros, Journal of Physics A: Mathematical and Theoretical **53**, 125002 (2020), arXiv:1912.04708.

[25] V. Ros, G. Biroli, and C. Cammarota, SciPost Physics **10**, 002 (2021), 10.21468/SciPostPhys.10.1.002.

[26] S. Franz and J. Rocchi, Journal of Physics A: Mathematical and Theoretical **53**, 485002 (2020), arXiv:1909.13044.

[27] T. Rizzo, Physical Review B **104**, 094203 (2021), arXiv:2012.09556.

[28] C. Cammarota and E. Marinari, Phys. Rev. E **92**, 010301(R) (2015), arXiv:1410.2116.

[29] M. R. Carbone, V. Astuti, and M. Baity-Jesi, Phys. Rev. E **101**, 052304 (2020), arXiv:2001.02567.

[30] D. A. Stariolo and L. F. Cugliandolo, EPL (Europhysics Letters) **127**, 16002 (2019), arXiv:1904.10731.

[31] D. A. Stariolo and L. F. Cugliandolo, Phys. Rev. E **102**, 022126 (2020), arXiv:2004.09410.

[32] D. Tapias, E. Paprotzki, and P. Sollich, Journal of Statistical Mechanics: Theory and Experiment **2020**, 093302 (2020), arXiv:2005.04994.

[33] M. Baity-Jesi, G. Biroli, and D. R. Reichman, The European Physical Journal E **44**, 1 (2021), arXiv:2103.07211.

[34] L. Song, W. Xu, J. Huo, F. Li, L.-M. Wang, M. D. Ediger, and J.-Q. Wang, Physical Review Letters **125**, 135501 (2020).

[35] D. Tapias and P. Sollich, arXiv:2201.03012 (2022).

[36] D. Chandler and J. P. Garrahan, Annual review of physical chemistry **61**, 191 (2010).

[37] C. P. Royall, F. Turci, and T. Speck, The Journal of Chemical Physics **153**, 090901 (2020), arXiv:2003.03700.

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

[39] I. Hartarsky, M. Baity-Jesi, R. Ravasio, A. Billoire, and G. Biroli, J. Stat. Mech.: Theory Exp **2019**, 093302 (2019), arXiv:1904.08024.

[40] A. Barrat and M. Mézard, Journal de Physique I **5**, 941 (1995), cond-mat/9504088.

[41] E. M. Bertin, J. Phys. A **36**, 10683 (2003), arXiv:cond-mat/0305538.

[42] B. Doliwa and A. Heuer, Phys. Rev. Lett. **91**, 235501 (2003).

[43] T. R. Kirkpatrick, D. Thirumalai, and P. G. Wolynes, Physical Review A **40**, 1045 (1989).

[44] M. Baity-Jesi and V. Martín-Mayor, Journal of Statistical Mechanics: Theory and Experiment **2019**, 084016 (2019), arXiv:1901.05581.

[45] G. Folena, S. Franz, and F. Ricci-Tersenghi, Phys. Rev. X **10**, 031045 (2020), arXiv:1903.01421.

[46] T. Speck, Journal of Statistical Mechanics: Theory and Experiment **2019**, 084015 (2019), arXiv:1902.07768.

[47] E. Bertin, Journal of Physics A: Mathematical and Theoretical **43**, 345002 (2010), arXiv:1004.1597.

[48] E. Bertin, Phys. Rev. E **85**, 042104 (2012), arXiv:1112.0189.

[49] T. B. Schrøder, S. Sastry, J. C. Dyre, and S. C. Glotzer, The Journal of Chemical Physics **112**, 9834 (2000), https://doi.org/10.1063/1.481621.

[50] I. Junier and J. Kurchan, J. Phys. A **37**, 3945 (2004), arXiv:cond-mat/0311158v1.

[51] P. Sollich, Journal of Physics A: Mathematical and General **39**, 2573 (2006), arXiv:cond-mat/0601007.

[52] N. Metropolis and S. Ulam, Journal of the American statistical association **44**, 335 (1949).

[53] W. Krauth, *Statistical Mechanics: Algorithms and Computations* (Oxford University Press, Oxford, 2006).

[54] R. G. Margiotta, R. Kühn, and P. Sollich, Journal of Physics A: Mathematical and Theoretical **51**, 294001 (2018).