Maximum-energy records in glassy landscapes

Ivailo Hartarsky[*1], Marco Baity-Jesi[2], Riccardo Ravasio[3], Alain Billoire[4], and
Giulio Biroli[5]

[1]Département Mathématiques et Applications, École Normale Supérieure, CNRS, PSL Research
University, 75005 Paris, France
[2]Chemistry Department, Columbia University, 3000 Broadway, 10027 New York (NY), USA
[3]Institute of Physics, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
[4]Institut de Physique Théorique, Université Paris Saclay, CNRS, CEA, F-91191, Gif-sur-Yvette, France
[5]Laboratoire de Physique Statistique, École Normale Supérieure, CNRS, PSL Research University, 75005
Paris, France

June 15, 2019

Abstract

We study the evolution of the maximum energy $E_{\max}(t)$ reached between time 0 and time
t in the dynamics of simple models with glassy energy landscapes, after an instant quench
from infinite temperature to a target temperature T. Through a detailed description of the
activated dynamics, we are able to describe the evolution of $E_{\max}(t)$ from short times, through
the aging regime, until after equilibrium is reached, thus providing a detailed description of
the long-time dynamics. Finally, we compare our findings with numerical simulations of the
p-spin glass and show how the maximum energy record can be used to identify the threshold
energy in this model.

Keywords: Activated dynamics; energy landscapes; extreme value statistics; slow relaxation,
glassy dynamics, aging; spin glasses

Contents

---

[*]ivailo.hartarsky@ens.fr

1 Introduction

Glass formers incur a sensational dynamical slowdown as the temperature is decreased [1]. At high temperatures, this slowdown is well explained by mean field (MF) theory [2, 3], which predicts a breaking of ergodicity at a temperature T_{d}: at $T < T_{\mathrm{d}}$ the typical configurations are confined in local minima of the energy landscape. Since, in MF, exiting these local minima requires overcoming barriers of height diverging with the system size N, in the thermodynamic limit the system remains stuck in a portion of phase space.

Because in non-MF systems the energy barriers are not diverging, real low-dimensional glasses are able to escape local minima of the energy and ergodicity is restored. Nevertheless, the dynamics below T_{d} is extremely slow, since overcoming a barrier of height Δ requires an exponentially large time τ, given by the Arrhenius law, $\tau \propto \exp(\Delta/k_{\mathrm{B}}T)$. This extremely slow low-temperature dynamical regime, consisting in jumps over energy barriers, is often referred to as (thermal) *activation* or activated dynamics.

Activated dynamics can be studied in MF models by considering exponentially large times in N [4, 5, 6], which restricts numerical simulations to very small system sizes. From the analytical point of view, the situation is even harder, because the thermodynamic limit cannot be taken, since once N is sent to infinity it becomes impossible to treat times of order $\exp(N)$. A consequence is that, to our knowledge, there are very few exact analytical calculations for the activated regime in canonical glass models.

A successful alternative analytical framework for the description of activated dynamics is the Trap Model (TM) [7, 8], a toy model that describes the simplest case of activation, in the absence of any other dynamical mechanism, where local minima (traps) are exited by reaching a fixed threshold energy E_{th}. A detailed non-trivial analysis has been made for the TM [8, 9, 10], and a series of works that culminated recently has shown that for large enough system sizes the TM correctly describes the activated long-time dynamics of a paradigmatic glass system, the Random Energy Model (REM) [11, 12, 13, 14, 5].

In this paper we go beyond these works in two main ways. Firstly, we thoroughly characterise the dynamics of the REM at *all* time scales. Secondly, we unveil an interesting behaviour of the maximum reached energy $E_{\mathrm{max}}(t)$, which can be related to static quantities such as the threshold energy E_{th}. Our findings are then tested by numerical simulations in the *p*-spin model, which provides a more complex case of glassy dynamics than the TM and REM.

Our approach is based on record statistics, which addresses questions on the extreme events arising in series of random variables [15, 16]. The simplest case of tracking the maximum in a series of independent random variables is well understood, but as soon as some dependence is added to the problem, it becomes less trivial, and rigorous results are only known for a very limited number of problems, such as random walks [16] (see [17] for a review) and eigenvalue statistics [18, 19, 20] (see [21] for a review).

We focus on three paradigmatic models of mean-field spin glasses, the Trap Model (TM), the Random Energy Model (REM), and the p-spin model, with particular emphasis on the REM. We study the evolution of the maximum energy, $E_{\max}(t)$, reached by time t after an instant quench from an infinite to a small temperature T.

The TM consists of a random walk in a fully-connected graph. The graph is equipped with a random potential and the transition rates do not depend on the destination energy. For this model we find that $E_{\max}(t)$ has a simple behaviour at all times.

In the REM, the random walk is on a more sparse graph and the transition rates depend on the difference between the starting and destination energy. These features lead to a rich non-trivial behaviour of $E_{\max}(t)$, in which we are able to identify several different regimes for different lapses of time from the quench (see Sec. 4 for a summary and Tab. 1 for more detailed results). Following $E_{\max}(t)$ opens a window on exponentially large time scales which are usually hard to study.

Finally, we make a digression to the p-spin model, whose landscape is more complex than the one of the REM. This is due to an explicit correlation among the energy levels, which are no longer independent identically distributed (i.i.d.) random variables. We show numerically that even in this case the maximum energy reached grows logarithmically. We also argue that the maximal energy $E_{\max}(t/2, t)$, reached in the time interval $[t/2, t]$, can be used to identify the threshold energy E_{th} (i.e. the highest energy at which local minima are found) in a finite size system. We further show numerically that this estimate converges smoothly to the analytical prediction in the thermodynamic limit.

In section 2 we define the models we study. In section 3 we describe the evolution of the maximum energy reached in the TM, whereas in section 4 it is analysed in the REM. In section 5 we discuss the p-spin model and we show how the maximum energy reached can be used to determine the threshold energy. Finally, we give some concluding remarks in section 6. In appendix A we recall some standard extreme-value statistics results, and in appendix B we take into account a number of subleading corrections which allow for a quantitative comparison of our results with simulations on very small systems.

2 Models

In this section we describe the three models treated in this paper.

2.1 Trap model

In the TM [7, 8, 22], M states lie on a fully connected graph, and their energies are drawn independently from an exponential distribution

$$\rho_{\exp}(E) = \alpha e^{\alpha E} \Theta(-E) \,, \tag{1}$$

where $\Theta(\cdot)$ is the Heaviside step function and $\alpha > 0$ is a parameter that will be set to 1.

The transition rate, $q_{i,j}$, from a state i to a state j does not depend on the target state j:

$$q_{i,j} = \frac{1}{M} e^{\beta E(i)} \,, \tag{2}$$

where $\beta = 1/T$ is the inverse temperature and $E(i)$ is the energy of state i. This represents a system in which leaving a state (or *trap*) requires reaching a threshold energy $E_{\mathrm{th}} = 0$. Once the threshold is reached, the whole space of states becomes accessible. This model has weak ergodicity breaking at temperatures $T \leq T_{\mathrm{c}} = \frac{1}{\alpha}$ [8] (see also paragraph 3.1).

Summarizing, in the TM the energies of different states are independent, the space of states is fully connected, and the dynamics depends only on the state issuing the transition and not

on the target state, implying that there is a constant threshold energy $E_{\mathrm{th}} = 0$ separating the configurations.

Gaussian Trap model The TM can be defined analogously for a Gaussian energy distribution

$$\rho_{\mathrm{gauss}}(E) = \frac{1}{\sqrt{2\pi N}} e^{-\frac{E^2}{2N}} ,\tag{3}$$

setting $M = 2^N$. In this case, also positive energies can be reached, but since $E_{\mathrm{th}} = 0$ the dynamics leaves them quickly with a rate given by equation (2).

2.2   Random Energy Model

The REM describes a system of N binary spins [23, 24]. As a consequence, the total number of states is $M = 2^N$, and from each one it is possible to reach N new states by flipping a single spin. The energy change induced by flipping any of the spins is assumed to be so drastic that the energy levels are independent from each other. The energy E of a state is drawn from the probability distribution ρ_{gauss} (see Eq. (3)).

The statics of the REM can be solved exactly. There is a disordered phase for temperature $T > T_{\mathrm{c}} = \frac{1}{\sqrt{2\ln(2)}}$, and a glassy phase for $T < T_{\mathrm{c}}$ [23, 24]. [1] Concerning the dynamics of the REM, we will focus on Monte Carlo (MC) Metropolis dynamics. More precisely, we consider that spins flip one at a time and we set the transition rate from state i to state j, differing by a single spin, to be

$$q_{i,j} = \frac{1}{N} \min\left(1, e^{\beta(E(i)-E(j))}\right) .\tag{4}$$

In a nutshell, the REM can be seen as a TM with (i) a hypercubic space of configurations (instead of fully connected), and (ii) with more physical dynamics to move in it. In the limit of very long times and large system sizes, the dynamics of the REM is qualitatively equivalent to the one of the TM, with a threshold energy $E_{\mathrm{th}} = -\sqrt{2N\ln(N)}$ [5, 14].

2.3   The p-spin model

The p-spin model represents a system of N binary spins $\sigma_i = \pm 1$ [23]. At variance with the TM and the REM, where the energies are i.i.d., the energy depends on the microscopic configuration of the system through the Hamiltonian

$$H = -\sum_{i_1 < \ldots < i_p} J_{i_1,\ldots,i_p} \sigma_{i_1} \ldots \sigma_{i_p} ,\tag{5}$$

in which the interactions couple all the possible groups of p different spins. The bonds $J_{i_1,\ldots,i_p}$, associated to a p-tuple of spins, are independent and usually follow a Gaussian or bimodal distribution. We choose the latter, with $J_{i_1,\ldots,i_p} = \pm\sqrt{p!/N^{p-1}}$ with probability one half. An alternative way to define the previous Hamiltonian is as a Gaussian random field defined on the N-dimensional hypercube (in this case the couplings are Gaussian random variables). Its mean is zero, whereas its covariance in the large N limit is $\overline{H(\{\sigma_i\})H(\{\sigma_i'\})} \sim Nq^p$ where $q = \sum_i \sigma_i \sigma_i'/N$.

When p is finite, there is a general agreement that all models with $p \geq 3$ fall in the same universality class [25]. Instead, if p is diverging, the way limits are taken becomes important. On the one hand, if one studies the short-time limit, sending first $N \to \infty$, and only later $p, t \to \infty$, the p-spin reduces to the REM [23] (which corresponds to naively saying that in the large N limit the covariance tends to $\delta_{\{\sigma_i\}\{\sigma_i'\}}$). On the other hand, when studying the long-time regime of activated dynamics, one needs to study large but finite N at every time t (recall that by definition $p \leq N$), so the REM cannot be recovered [26].

We will focus on $p = 3$, in the limit of large but finite N, where activated processes become possible [4]. As for the dynamics, we use the single-spin MC Metropolis dynamics, described in Eq. (4).

---

[1] The critical temperature T_{c} is different from the one given in Refs. [23, 24], because we define the model with variance $\mathrm{Var}(E) = N$, instead of $\mathrm{Var}(E) = N/2$.

3 Energy maxima in the Trap Model

3.1 Qualitative description of the dynamics in the TM

We consider a typical non-equilibrium protocol: an instant quench from infinite temperature to a target temperature T. In the TM, the system is always stuck in a trap, where it will remain for a time $\tau \propto e^{-\beta E}$ (Eq. (2)), before emerging to the threshold to transition to the next trap. We call τ the trapping time.

One can easily see that the distribution of trapping times is heavy-tailed with density

$$\psi(\tau) \propto \tau^{-(\alpha T+1)} \,, \tag{6}$$

which identifies a critical temperature at $T_{\rm c} \equiv \frac{1}{\alpha}$ [8]. When $T < T_{\rm c} \equiv \frac{1}{\alpha}$ the average trapping time is infinite and the total waiting time is of the order of the maximal waiting time in a single state [8]. Consequently, the process waits in the deepest trap it has seen for about as much time as the system has spent in the heat bath, implying

$$E(t) \approx E_{\rm min}(t) \sim -T \ln t \,. \tag{7}$$

Eventually, the system is able to reach $E_{\rm th}$ and fall back into shallower traps, until it finds an even deeper trap than the one before, where it spends a time longer than previously elapsed time. As a consequence, as time passes, the energy decreases gradually, and the dynamics becomes increasingly slower. This dependence on how long the system has been in the bath is called *aging* and results in an increasing time correlation of the system. In finite systems aging goes on until equilibrium is reached.

3.2 Energy Maxima in the Exponential Trap Model

For our consideration of the maximum energy, $E_{\rm max}(t)$, reached after a time t we will restrict ourselves to the low-temperature TM, where aging occurs. In other words, we assume that $T < T_{\rm c}$, so that (7) holds. Then $E_{\rm max}(t)$ can be calculated through extremum statistics. Recalling that each new trap is chosen independently, the number of traps, n_t, that the system will have visited satisfies

$$n_t \int_{-\infty}^{E_{\rm min}(t)} \rho_{\rm exp}(E) dE \sim 1 \,, \tag{8}$$

where $E_{\rm min}(t)$ is the minimum energy reached until time t. Therefore, $n_t \sim e^{-\alpha E_{\rm min}(t)} \sim e^{-\alpha E(t)}$, since in the TM the current energy is almost always the lowest energy reached (i.e. $E(t) \approx E_{\rm min}(t)$).

The number of visited traps needs to satisfy an analogous relation for the highest energy reached

$$n_t \int_{E_{\rm max}(t)}^{0} \rho_{\rm exp}(E) dE \sim 1 \,, \tag{9}$$

that yields $n_t \sim -1/E_{\rm max}(t)$.

Equating the two expressions for n_t, and using (7), one obtains the behaviour

$$E_{\rm max}(t) \sim \ln \left(1 - t^{-\frac{T}{T_{\rm c}}}\right) \sim -t^{-\frac{T}{T_{\rm c}}} \,, \tag{10}$$

so in the exponential TM the largest energy reached approaches its extremal value $E = 0$ as a power law. [2]

Note that with respect to all the models defined below, the exponential TM is special (and a bit pathological) since it has a cut-off in the energy distribution at the most numerous energies, i.e. it does not have any rare high energy. In consequence, its behaviour is atypical with respect to all other cases considered below.

---

[2]The high-temperature case, $T > T_{\rm c}$, can be treated similarly, yielding $E_{\rm max}(t) \sim -1/t$.

3.3 Gaussian Trap Model

In the Gaussian TM, since $\rho(E)$ is symmetric, the maximum and minimum records $E_{\max}(t)$ and $E_{\min}(t)$ follow the same law, with opposite signs. This is because, from any configuration i, all the configurations are reached with equal probability, and what changes is only the amount of time the system remains in them. Therefore, a configuration with $E < E_{\min}(t)$ is reached with the same probability as a configuration with energy $E > -E_{\min}(t)$. Since, by Eq. (7), in the TM $E(t) \sim E_{\min}(t) \sim -T \ln t$, we can deduce that

$$E_{\max}(t) \sim T \ln t. \tag{11}$$

Eq. (7) still holds for the Gaussian TM in its aging regime, $t < \exp(\beta \beta_i N)$, where $\beta_i = \min(\beta, \beta_c)$ and $\beta_c = \sqrt{2 \ln 2}$. If $\beta < \beta_c$ (i.e. in the high-temperature phase), a supplementary 'post-equilibration' regime develops for $\exp(\beta^2 N) < t < \exp(N(\beta_c^2 + \beta^2)/2)$, governed by

$$E_{\max}(t) = \sqrt{2N(\ln t - \beta^2 N/2)}. \tag{12}$$

In both cases, the evolution of the maximum energy eventually ends when $E_{\max}$ reaches the maximum possible value $N\sqrt{2 \ln 2}$. As we shall discuss a similar argument for the REM in Sec.4.2.6 and since there is no difference with the minimal energy, we omit the details here. However, let us note that depending on the temperature there are two different behaviours of $E_{\max}(t)/N$ as a function of $\ln t/N$. For $\beta > \beta_c$ there is a single regime, linear in $\ln t$ (11), while otherwise there are two different ones – a linear law (11) followed by a square-root dependency (12).

4 Energy maxima in the REM

We now turn to the Metropolis dynamics on the REM. As in the previous section we examine quenches from infinite temperature to a target temperature T. Throughout this section we consider the thermodynamics large-N limit and systematically discard subleading terms. Higher-order corrections for finite systems are discussed in App.B.

Short summary of the maximum-energy records in the REM. We are able identify and study a large number of different regimes in the dynamics, which are summarised in quantitative terms in Tab. 1 at the end of Sec. 4.

The energy of the initial state is typically of order $\sqrt{N}$ (Sec. 4.2.1). In the initial stages of the evolution, the system quickly falls into a local minimum (Secs. 4.2.2 and 4.2.3). Since it is not necessary to reach energies of order $\sqrt{N}$ to escape the local minima, $E_{\max}(t)$ stays unchanged and only starts growing once a time of order $t_3 = e^{\sqrt{2\beta^3 N \sqrt{2N \ln N}}}$ has passed (Eq. (16), Sec. 4.2.4).

At larger times, and, in particular, those exponentially large in N, $E_{\max}(t)$ incurs a further slow down, because not only new records are hardly accepted, but they are also harder to find (Sec. 4.2.5). Even after equilibrium has been reached, $E_{\max}(t)$ keeps growing (Sec. 4.2.6), and saturates only when the global maximum of the system has been visited (Sec. 4.2.7).

4.1 Basic features of the landscape and the dynamics of the REM

In the REM, most energy levels are concentrated around zero at a typical distance of order $\sqrt{N}$. At variance with the TM, the phase space has a structure, and each state has N neighbours whose energies, as we show in Sec. A.1, Eq. (34), typically lie in the interval

$$I \equiv \left[-\sqrt{2N \ln N}, \sqrt{2N \ln N}\right]. \tag{13}$$

The majority of states have at least one lower neighbour, and the dynamics spends little time there, since the Metropolis update rule [Eq. (4)] privileges energy descent. As a consequence, the states in which most of the time is spent are those without a lower neighbour, which have energy $\inf I = -\sqrt{2N \ln N}$ or lower. Minima with energy $E \sim -N$ can be seen as the equivalent of the TM traps, since $N \gg \sqrt{2N \ln N}$ (the separation between these minima and the threshold diverges with N). Since the energy of such states is much lower than the typical energies, to leading order the time spent in these low-energy minima configurations scales as $\tau \sim e^{-\beta E}$ [Eq. (4)].

4.2   Energy maximum evolution

In this section, as we did for the TM, we discuss the evolution of the typical maximum energy $E_{\max}(t)$ in the REM in the limit of very large system sizes. The final results will be summarised in Table 1.

When dealing with exponentially large time scales, a central assumption of our computation is that, each time the dynamics leaves a state, the system becomes independent from its past. In other words, the neighbours are drawn anew, so that returns to a recently visited configuration are not taken into account. This hypothesis is supported by an argument in Ref. [5], stating that even though, in the actual dynamics, some configurations are revisited, this happens with a sub-exponential rate. In consequence, this should not affect exponential time scales.

4.2.1   Initial condition

Before the quench the temperature is infinite, so $E_{\max}(0)$ is drawn from $\rho_{\text{gauss}}(E)$. Therefore, its intensive value is zero in the limit of large system sizes: $E_{\max}(0)/N \sim 1/\sqrt{N} \overset{N \to \infty}{\longrightarrow} 0$.

4.2.2   Drift towards the first trap

At the beginning of the evolution, the system quickly falls into a first local minimum of the energy. See Fig. 1 (left part) for a schematic description of this and the following regimes.

In fact, with probability $1 - \frac{1}{N+1}$, the initial state has at least one lower neighbour. The rate $q_{i,j}$ [Eq. (4)] privileges energy descent, and on time-scales of order one the energy will typically immediately decrease since climbing up moves will be discarded (their rate is exponentially small in $\sqrt{N}$). The transition toward the lowest neighbour, which is at energy inf I, has a rate $q_{i,j} \sim 1/N$, hence a local minimum will be reached in a time of order $t_1 = N$ (a configuration at energy inf I is a local minimum with probability of order 1). The number of transitions required to reach such state is of order $\ln(N)$. [3]

During this regime, before the first trap is reached, $E_{\max}$ maintains its initial value of order $\sqrt{N}$, because the energy is decreasing.

4.2.3   Stay in the first trap

Once the system reaches the first local minimum, which has energy $E \approx \inf I$ (that of a typical lowest neighbour), the energy difference with its neighbours is of order $\delta = \sqrt{\frac{N}{\ln N}}$ (see App. A.1). Therefore, the system remains in the first trap and does not move at all for a time of order $t_2 = \exp\left(\beta\sqrt{\frac{N}{\ln N}}\right)$. Consequently, $E_{\max}(t)$ does not change either.

4.2.4   Reaching common energies

We now estimate the time scale t_3 required, once the dynamics left the first trap, for the energy to become of order $\sqrt{N}$. Such energy needs to be reached for a new record $E_{\max}(t) > E(0) \sim \sqrt{N}$ to be hit.

After t_2, the system starts visiting a series of local minima of the energy. When the dynamics leaves a local minimum, it does so by jumping to its lowest-energy neighbour, which, to leading order, also has energy inf I. In order to reach energies of order $\sqrt{N}$, a large number of local minima needs to be visited.

The typical trapping time in such a minimum with energy E, is

$$\tau \sim N e^{\beta(\inf I - E)} \tag{14}$$

and the rate of jumping to a configuration of energy of order $\sim \sqrt{N}$ from the local minimum is

$$q_{\text{common}} \sim e^{\beta\left(E - \mathcal{O}(\sqrt{N})\right)}, \tag{15}$$

---

[3]At each transition, the number of lower neighbours is typically divided by two, so it will take $\mathcal{O}(\log_2(N))$ steps to reach a local minimum.
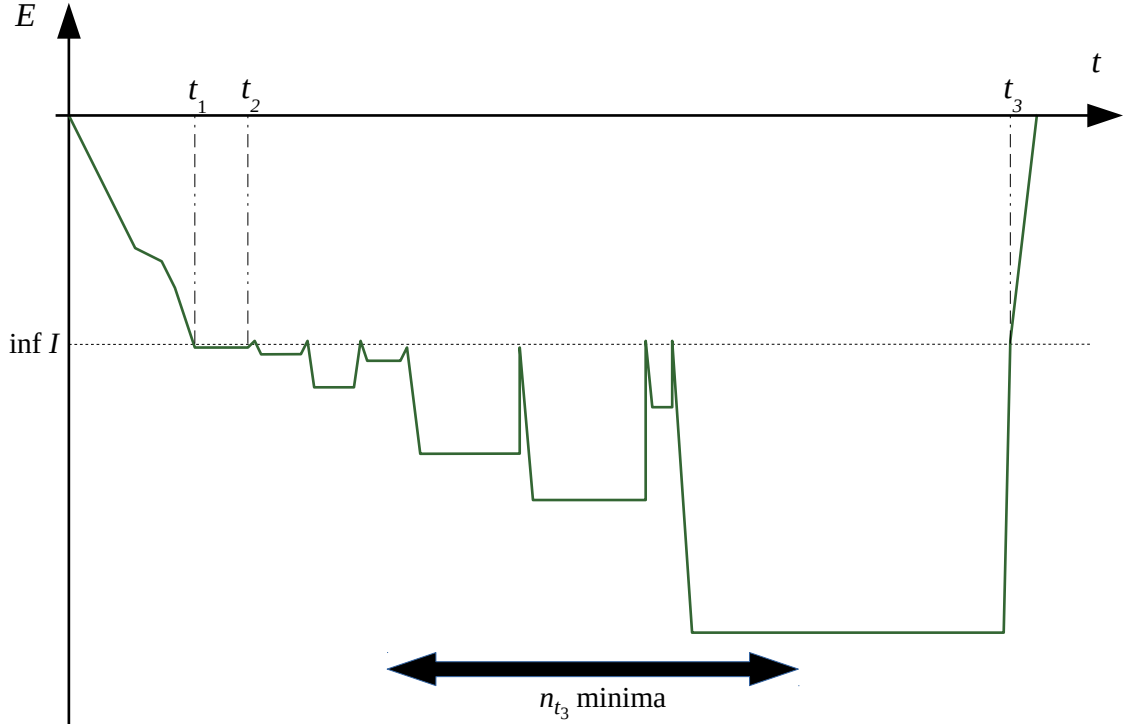
Figure 1: A schematic representation of the dynamics of the REM in the initial stages of the dynamics. In a time of order $t_1 = N$ the system falls in the first local energy minimum, whose energy is typically $\inf I$ [Eq. (13)]. It takes a time of order $t_2 = \exp(\sqrt{N/\ln N})$ to leave that state. After that, the system will visit a large number, n_{t_3}, of minima before it can reach energies of order $\sqrt{N}$, after a time $t_3 = e^{\beta E_1}$ [Eq. (16)]. During all this time $E_{\max}(t) = E(0)$. The times and energies in the diagram are not drawn to scale.

where $\mathcal{O}(\sqrt{N})$ indicates a quantity of order $\sqrt{N}$. The probability of jumping in a time τ to a configuration of energy $\sim \sqrt{N}$ is then given by

$$\tau q_{\mathrm{com}} \sim e^{\beta(\inf I - \mathcal{O}(\sqrt{N}))} \approx e^{\beta \inf I} .$$

For such a transition to be likely, the number of visited traps n_{t_3} needs to be of order $e^{-\beta \inf I}$, since any site usually has neighbours at energy $\sim \sqrt{N}$. The lowest of the n_t traps determines $E_{\min}(t)$. Thus, $E_{\min}(t)$ can be calculated as the minimum among Nn_t Gaussians (the factor N stems from the fact that we are dealing with minima of N Gaussians, not arbitrary Gaussians) with variance N. Therefore, through known extremum statistics results (see App. A.1, Eq. (36)), one obtains

$$- E_{\min}(t_3) \sim \sqrt{2N \ln{(Nn_{t_3})}} \approx \sqrt{2\beta N \sqrt{2N \ln N}} =: E_1. \tag{16}$$

By using Eq. (7), which is valid also in the REM [5], we obtain $t_3 = e^{\beta E_1}$. Until t_3, $E_{\max}(t)$ typically still maintains its initial value.

4.2.5 Aging

After t_3, $E_{\max}(t)$ starts increasing. We start by considering the aging stage, during which Eq. 7 is valid. We proceed analogously to Sec. 4.2.4 with an additional ingredient: high-energy records are harder to reach not only because transitions to them are rarely accepted – as it was also previously – but also because record-breaking states are rarely found among the neighbours of a given configuration.

The probability that a trap has a neighbour of energy at least E with probability $p_{\text{find}}(E) \sim e^{-\frac{E^2}{2N}}$. Thus, when exiting a trap, a neighbour of energy E is chosen over the lowest one with probability

$$p_{\text{find}}(E)p_{\text{accept}}(E) \sim e^{\beta(\inf I - E) - \frac{E^2}{2N}}. \tag{17}$$

Then, in order to hit a new record, the number of visited traps should be

$$n_t \sim p_{\text{find}}^{-1}(E)p_{\text{accept}}^{-1}(E). \tag{18}$$

Again, by using extremum statistics (App. A.1, Eq. (36)), one has

$$n_t \sim e^{\frac{E_{\min}(t)^2}{2N}} \tag{19}$$

Combining this with Eqs. (17) and (18) gives

$$-E_{\min}(t) \approx \sqrt{2\beta N(E_{\max}(t) - \inf I) + E_{\max}(t)^2}. \tag{20}$$

Plugging Eqs. (7) and (13) into Eq. (20), we obtain

$$E_{\max}(t) = \beta N \left(-1 + \sqrt{1 - \frac{2}{\beta}\sqrt{\frac{2\ln N}{N}} + \left(\frac{\ln t}{\beta^2 N}\right)^2} \right). \tag{21}$$

We can express Eq. (21) in simpler forms by considering explicitly the different values of $\ln t$.

First records   We first consider the times for which $\ln t$ is of order βE_1 (but greater than it, so as to have $t > t_3$). Then, by substituting Eq. (16), Eq. (21) becomes

$$E_{\max}(t) = \sqrt{2N\ln N}\left(\left(\frac{\ln t}{\beta E_1}\right)^2 - 1\right). \tag{22}$$

Intermediate regime   We next assume that $\ln t$ is much larger than βE_1, but much smaller than $\beta^2 N$. Then Eq. (21) takes the particularly simple form

$$E_{\max}(t) = \frac{\beta N}{2}\left(\frac{\ln t}{\beta^2 N}\right)^2. \tag{23}$$

Exponential scales   Finally, we have $\ln t \sim \beta^2 N$. Then Eq. (21) reduces to

$$E_{\max}(t) = \beta N \left(-1 + \sqrt{1 + \left(\frac{\ln t}{\beta^2 N}\right)^2} \right). \tag{24}$$

In order to determine time t_4 at which aging ends, we consider the mean energy, which is equal to $E_{\min}(t)$ as long as aging lasts. Then aging ends with $E_{\min}(t_4) = -\frac{\ln t_4}{\beta}$ equal to the equilibrium mean energy, itself given either by the global minimum of the energy or by $\langle E \rangle = \frac{\int E e^{-\beta E}\rho_{\text{gauss}}(E)\mathrm{d}E}{\int e^{-\beta E}\rho_{\text{gauss}}(E)\mathrm{d}E} = -\beta N$. Hence, by App. A.1, (33)

$$t_4 = e^{\beta\beta_{\text{i}}N}, \tag{25}$$

where we introduced $\beta_{\text{i}} \equiv \min(\beta, \beta_c)$.

We stress that this regime describes the aging occurring at exponential time scales, that are generally regarded as the relevant ones for activated dynamics. In this regime, $E_{\max}(t)$ grows as $\ln t$ (and so does $-E_{\min}(t)$) until the system has eventually equilibrated. Yet, since $E_{\max}(t)$ grows slower than $E_{\min}(t)$, it will keep increasing even after $E_{\min}(t)$ saturates.

4.2.6 Equilibrium regime

After t_4, $E_{\max}(t)$ continues to evolve even though the system has already reached macroscopic equilibrium. In this regime Eqs. (17) and (18) still hold, whereas Eq. (7) fails, making $E_{\min}(t)$ irrelevant for the calculation of $E_{\max}(t)$.

The calculation that led to Eq. (25) also gives that the mean time spent per state $\frac{t}{n_t}$ stations at the end of aging (the system reaches macroscopic equilibrium) at the value $\frac{t_4}{n_{t_4}}$. In other words, $\frac{t}{n_t} = \frac{t_4}{n_{t_4}}$ for all $t > t_4$.

Using Eqs. (17), (18) and (25) together with the fact that $\inf I$ is negligible compared to $E_{\max}(t)$, we get

$$t = n_t \frac{t_4}{n_{t_4}} = n_t e^{\beta \beta_{\mathrm{i}} N} e^{-N \frac{\beta_{\mathrm{i}}^2}{2}} = e^{\beta \beta_{\mathrm{i}} N + \beta E_{\max}(t) - N \frac{\beta_{\mathrm{i}}^2}{2} + \frac{E_{\max}(t)^2}{2N}} \,.$$

Thus,

$$E_{\max}(t) = \beta N \left(-1 + \sqrt{ \left(1 - \frac{\beta_{\mathrm{i}}}{\beta} \right)^2 + 2 \frac{\ln t}{\beta^2 N} } \right) \,. \tag{26}$$

Eq. (26) is valid until the global maximum $E_{\max}(\infty) = \beta_c N$ (see App. A.1, Eq. (33)) is reached. This means that it remains valid for

$$\ln t \leq (2\beta - \beta_{\mathrm{i}} + \beta_c)(\beta_c + \beta_{\mathrm{i}}) \frac{N}{2} = \begin{cases} N \frac{(\beta + \beta_c)^2}{2} & \text{if } \beta < \beta_c \\ 2N \beta_c \beta & \text{if } \beta > \beta_c \,. \end{cases} \tag{27}$$

Inequality (27) defines the typical time t_5 required to reach the global maximum

$$t_5 = e^{(2\beta - \beta_{\mathrm{i}} + \beta_c)(\beta_c + \beta_{\mathrm{i}}) \frac{N}{2}} \,. \tag{28}$$

4.2.7 Saturation

Naturally, after t_5, that is after the global maximum has been attained, the maximum energy cannot change any more, so thereafter

$$E_{\max}(t) = \beta_c N \,. \tag{29}$$

4.3 Remarks

In our discussion, we treated the typical maximum of the energy $E_{\max}(t)$. We expect that when taking the thermodynamic limit there will be a concentration of the measure around a central value, so the same results should hold when replacing typical $E_{\max}(t)$ by the expectation $\mathbb{E}[E_{\max}(t)]$. In App. B, we describe the subleading corrections that need to be taken into account in order to study finite systems, and tackle a few of them. The different regimes of $E(t)$ and $E_{\max}(t)$ (at leading order as $N \to \infty$) are summarised in Tab. 1.

Relevant time scales for the glassy activated dynamics stem naturally from our analysis. The longest time scale is exponentially large in the system size. Thus, we can define a rescaled time $\theta = \ln t / N$, which is useful to express Eqs. (24), (26) and (29) in a meaningful way. The intensive maximum energy for these time scales is plotted in Fig. 2–left, with a comparison with numerics in small systems. In the thermodynamic limit, the piecewise concatenation of these regimes is continuous.[4]

In units of θ, all the previous time regimes collapse to zero in the thermodynamic limit. For the regime at which the first records are observed, the relevant rescaled time is $\theta' = \ln t / (\beta E_1(N))$, whereas the energy should be rescaled by a factor $1/\sqrt{2N \ln N}$ [see Eq. (22)]. The resulting rescaled curve is plotted in the inset of Fig. 2–left.

In Fig. 2–right, we show that $E_{\max}(t)$ keeps growing even after the system has thermalised, and that the equilibration time, t_4, extracted from $E_{\max}(t)$ (in the thermodynamic limit) corresponds roughly to the time at which $E(t)$ reaches its equilibrium value (in a system of size $N = 12$).

---

[4] For $\beta < \beta_c$ its derivative is also continuous at $\theta = \beta \beta_{\mathrm{i}}$, corresponding to t_4.

Regime	Time scale	$E(t)/N$	$E_{\max}(t)/N$
Initial	$t = 0$	$\sim 1/\sqrt{N}$	$E(0) \sim 1/\sqrt{N}$
Drift	$\ln t < \log_2 N$	decreasing	$E(0) \sim 1/\sqrt{N}$
First trap	$\ln t < \beta\sqrt{\frac{N}{\ln N}}$	$\sqrt{2\ln N}/N$	$E(0) \sim 1/\sqrt{N}$
Common	$\ln t < \beta E_1$	$-T\ln t/N$	$E(0) \sim 1/\sqrt{N}$
First records	$\ln t \sim \beta E_1$	$-T\ln t/N$	$\sqrt{\frac{2\ln N}{N}}\left(\left(\frac{\ln t}{\beta E_1}\right)^2 - 1\right)$
Intermediate	$\beta E_1 \ll \ln t \ll \beta^2 N$	$-T\ln t/N$	$\frac{\beta}{2}\left(\frac{\ln t}{\beta^2 N}\right)^2$
Exponential	$\ln t < \beta\beta_\mathrm{i} N$	$-T\ln t/N$	$\beta\left(-1+\sqrt{1+\left(\frac{\ln t}{\beta^2 N}\right)^2}\right)$
Equilibrium	$\ln t < (2\beta - \beta_\mathrm{i} + \beta_\mathrm{c})(\beta_\mathrm{c} + \beta_\mathrm{i})\frac{N}{2}$	$-\beta_\mathrm{i}$	$\beta\left(-1+\sqrt{\left(1-\frac{\beta_\mathrm{i}}{\beta}\right)^2 + 2\frac{\ln t}{\beta^2 N}}\right)$
Saturation	any larger time scale	$-\beta_\mathrm{i}$	β_c

Table 1: Summary of the regimes of the dynamics evolution in the REM in chronological order. For each regime (first column) we depict the time of validity, the typical intensive energy one finds at time t, $E(t)/N$, and the typical intensive record of the maximum energy, $E_{\max}(t)/N$. The energy $E_1 = \sqrt{2\beta N\sqrt{2N\ln N}}$ is defined in Eq. (16), $\beta_\mathrm{c} = 1/T_\mathrm{c} = \sqrt{2\ln 2}$ is the inverse critical temperature, and $\beta_\mathrm{i} = \min(\beta_\mathrm{c}, \beta)$.

5   Energy maxima and threshold energy in the p-spin model

A natural question is how the behaviour of $E_{\max}(t)$ extends to more complex glassy systems, and how it can be used to extract useful information on them. We consider the p-spin model, since in the limit $p \to \infty$ the thermodynamics (though not the long-time activated dynamics [26]) is the same as the one of the REM.

In the p-spin model, there is a threshold energy E_{th} over which there are no energy minima [27], as it also happens in the TM and the REM. Yet, at variance with the TM and the REM, where $\lim_{N\to\infty}(E_{\mathrm{th}}/N) = 0$, in the p-spin model E_{th} is extensively negative (i.e. $\lim_{N\to\infty}(E_{\mathrm{th}}/N) < 0$).

Let us now take into account the maximum energy $E_{\max}(t/2, t)$ reached in the time interval $[t/2, t]$. Taking the maximum in a dilating time interval is useful to identify the separation of time scales that arises in the dynamics of glassy systems. In the thermodynamic limit, the dynamics of the p-spin model on time scales not diverging with N is known [28]: the energy decreases monotonically towards the threshold energy. Moreover, in this regime, the energy as a function of time has fluctuations of order $\sqrt{N}$. In consequence, $E_{\max}(t/2, t)$ is equal to the energy at time $t/2$ and approaches E_{th} for t large but not diverging with N. On times diverging with N the dynamics becomes activated (on time-scales exponentially large in N), the intensive energy goes below E_{th}, and the energy starts to have rare high excursions that increase the value of $E_{\max}(t/2, t)$ (see also Fig. 5 in App. C). Hence, one can expect that $E_{\max}(t/2, t)$ grows close to logarithmically, as in record-breaking dynamics for i.i.d. random variables (and as it happens for $E_{\max}(t)$ in the REM during the aging regime, see table 1).

The minimum value reached by $E_{\max}(t/2, t)$ can be used to define the value of E_{th} in numerical simulations. This method adds to the usual procedures of calculating the threshold, and has the advantage of not requiring extrapolations to infinite times [28], nor the computation of the complexity of minima [29]. In Fig. 3 we show $E_{\max}(t/2, t)$ both in the REM and in the p-spin model (with $p = 3$), for different system sizes. As expected, in both cases the curve reaches a minimum which we identify as E_{th}. Note that, in the REM, $\frac{E_{\mathrm{th}}}{N}$ grows as N increases whereas in the 3-spin model it decreases, in agreement with the fact that the intensive threshold energy is zero in the former model and negative in the latter. In the inset of Fig. 3, we show that, in the p-spin model, the finite-size threshold energy obtained through this procedure is controlled by the
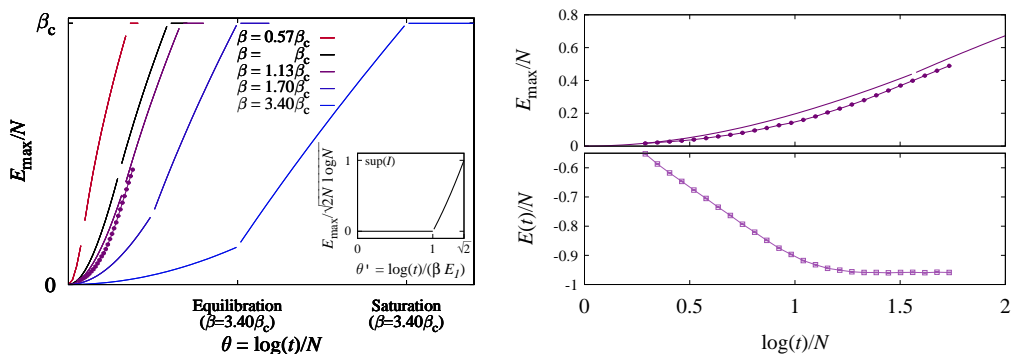
Figure 2: Intensive maximum energy, $E_{\max}$, in the REM as a function of the rescaled time $\theta = \ln t/N$. **Left**: Plots for several values of the inverse temperature β. The y axis is linear, in units of β_c. The curves are split in three, in order to stress the presence of three subsequent regimes: Aging (Sec. 4.2.5), Equilibrium (Sec. 4.2.6) and Saturation (Sec. 4.2.7). The tics on the x axis show the beginning of these regimes for the highest β, corresponding to $T = 0.25$. We also show data from simulations at $\beta = 1.13\beta_c$, for systems of size $N = 12$, which agrees well with our $N = \infty$ predictions. Finite-size corrections are discussed in App. B. In the **inset** we show $E_{\max}$ for shorter time scales θ' (see main text). **Right**: The **top** figure depicts the same curves shown in the left panel, for $\beta = 1.13\beta_c$. The solid line is the result of our calculation (in the thermodynamic limit), and the points are from runs with $N = 12$ (error bars are smaller than the points). The separation between the continuous lines emphasises the end of the aging regime according to our calculation [Eq. (25)]. The **bottom** figure shows the energy as a function of time for comparison. Numerical data here and in Fig. 3 are averaged over disorder realizations.

$N^{-1/2}$ fluctuations, and converges to its analytical value in the thermodynamic limit [30].[5] From the figure one can also see that, as argued in the previous paragraph, the relevant time scales for the growth of $E_{\max}(t/2, t)$ are exponential.

6   Conclusions

By counting the number of visited traps, and noting that in TM-like models almost all the time is spent in the deepest trap, we were able to calculate the evolution of the maximum energy reached after a time t, $E_{\max}(t)$. Our method, which applies record statistics to glasses, allowed us to add to the very reduced number of analytical calculations of activation in glasses (see also [31, 14]). It has the advantage of transparently treating the dynamics, providing new insight on the long-time behaviour of glassy systems. Further, we showed how the behaviour of the maximum energy record naturally reveals the different time scales involved in glassy relaxation, and how it can be used to identify a crucial static quantity, the threshold energy E_{th}. Calculating E_{th} by identifying the minimum of $E_{\max}(t/2, t)$ is arguably simpler than calculating the complexity of the minima [29], and does not require any kind of extrapolations to infinite times as required by traditional methods [28].

The relationship between E_{th} and $E_{\max}(t/2, t)$ relies on the assumption that it is necessary to reach E_{th} to connect low-lying energy minima. Therefore, its success in the p-spin model is a strong suggestion that there are no low-energy paths connecting minima in the p-spin.

Our results also connect to recent phenomenological works, suggesting to rationalise the dynamical slowdown of low-temperature glasses as a record-breaking process over larger and larger domains [32, 33, 34]. Our mean-field calculations could then account for this process in regions smaller than the correlation length, where the system is effectively fully-connected.

---

[5]This same scaling is not as clean in the REM, probably because the system sizes are too small. In App. B we argue that the asymptotic limit is reached for $N \sim 10^3$.
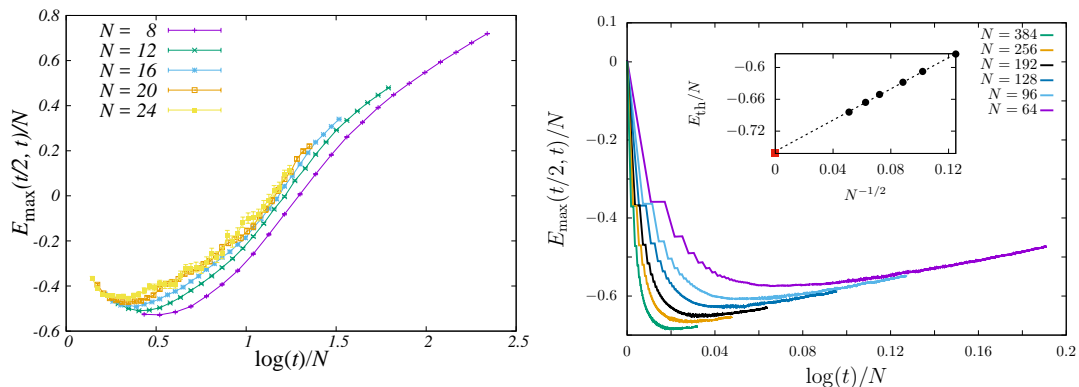
Figure 3: Intensive maximum energy $E_{\max}(t/2, t)/N$ reached in the time interval $[t/2, t]$, for different system sizes N. On the **left**, we show data for the REM at $T = 0.75$. Decreasing the temperature (not shown), the bottom of the curve becomes progressively flatter, but the lowest point stays approximately at the same height. On the **right**, we show data for the 3-spin model at $T = 0.54$. In both cases the curves seem to converge to a master curve proportional to $\ln t$. **Inset**: Scaling behaviour of the size-dependent threshold energy $E_{\mathrm{th}}(N)$ (defined as the minimum of $E_{\max}(t)$) in the 3-spin model. The prediction from Ref. [30], valid for infinitely large systems, $E_{\mathrm{th}} = -0.762$, is also shown (red square). The dotted line represents the fit $E_{\mathrm{th}}(N) = aN^{-1/2} + b$. The coefficients of the fit are $a = 1.47(1)$ and $b = -0.758(1)$.

Acknowledgements

A   Typical neighbours' energies

In this section we recall some useful results from classical extremum statistics that we use thorough in our paper. For a more detailed discussion see, for example, Refs. [15, 16].

A.1   Asymptotic bounds

We would like to determine the typical value of the maximum, M_n, out of n i.i.d. random variables drawn from distribution $\rho_{\text{gauss}}(E)$ [Eq. (3)]. We do this for arbitrary $n = n(N) \to \infty$, as different applications will be needed.

Let X be a random variable drawn from the distribution ρ_{gauss}. Its cumulative distribution $F(x)$ is the probability that X is smaller than x, $F(x) = \mathbb{P}(X \leq x)$. We consider i.i.d. random variables, so the probability that the maximum of n tries is smaller than some constant a is $\mathbb{P}(M_n \leq a) = (F(a))^n$.

We can find a location-scale transformation to show that M_n asymptotically follows the Gumbel distribution $\mathscr{L}_{\text{GUM}}((-\infty, x]) = e^{-e^{-x}}$. [6] Set

$$a_n = \sqrt{\frac{N}{2 \ln n}}, \; b_n = \sqrt{2N \ln n} \left(1 - \frac{\ln \ln n + \ln(4\pi)}{4 \ln n} \right). \tag{30}$$

Then integrating by parts we get

$$1 - F(b_n + xa_n) = \frac{1}{\sqrt{2\pi N}} \int_{b_n + xa_n}^{\infty} e^{-\frac{y^2}{2N}} \mathrm{d}y \overset{n \to \infty}{\sim} \sqrt{\frac{N}{2\pi}} e^{-\frac{(b_n + xa_n)^2}{2N}} \frac{1}{b_n + xa_n}$$

$$\sim \frac{1}{\sqrt{4\pi \ln n}} \exp\left(-x - \frac{b_n^2}{2N} \right) \sim \frac{e^{-x}}{n}. \tag{31}$$

By isolating $F(b_n + xa_n)$ and recalling that $\mathbb{P}(M_n \leq a) = (F(a))^n$, one gets that

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x \right) \overset{n \to \infty}{\longrightarrow} \mathscr{L}_{\text{GUM}}. \tag{32}$$

Explicit cases The two particular cases relevant for our purposes are $n = 2^N$ and $n = N$, which respectively indicate the global energy maximum of the system (there are 2^N sites in the system), and the highest-energy neighbour (each site has N neighbours). For the global maximum we inject $n = 2^N$ in the expression (30) of b_n and obtain

$$E_{\max}(\infty) = -E_{\min}(\infty) = M_{2^N} = \beta_{\text{c}} N + \mathcal{O}(\ln N). \tag{33}$$

In the second case, $n = N$, we have

$$-\inf I = \sup I = M_N = \sqrt{2N \ln N} \left(1 + \mathcal{O}\left(\frac{\ln \ln N}{\ln N} \right) \right). \tag{34}$$

Moreover, to the leading order Eq. (32) gives

$$n \approx e^{\frac{M_n^2}{2N}} \tag{35}$$

or equivalently

$$M_n \approx \sqrt{2N \ln n}. \tag{36}$$

It is worth noting that, according to Eq. (32), the energy difference between a local minimum and its lowest neighbour is of order a_N (see also [16, Sec. 3.1.2]).

Finally, note that due to the concentration of M_n around b_n, for large systems one can treat both the global maximum of the energies and I as deterministic.

---

[6]Mind that the Gumbel distribution is not centred: its mean is the Euler-Mascheroni constant γ.

A.2   Finite N

When dealing with small systems (see App. B) we need to go beyond the asymptotic bounds. To get the exact expression for $\sup(I) = M_N$ we can use again the fact that $\mathbb{P}(M_n \leq x) = (F(x))^n$. From there we can extract the distribution of M_n, and calculate its average, which results in

$$\sup I = -\inf I = \sqrt{\frac{N}{2\pi}} \int_{-\infty}^{+\infty} Nx(1 - \mathrm{erfc}(x))^{N-1} e^{-\frac{x^2}{2}} \, \mathrm{d}x \,, \tag{37}$$

where

$$\mathrm{erfc}(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{u^2}{2}} \, \mathrm{d}u \,. \tag{38}$$

For example, for $N = 20$, this gives $\sup I \approx 8.34$.

B   Corrections for small systems

In this appendix we take into account subleading terms of $E_{\max}(t)$, so as to make quantitative predictions also for small system sizes. The general reasoning will remain the same as in Sec. 4, but we will pay particular attention to the fact that $-\inf I$ is not much smaller than N and we will no longer use the asymptotic value of $\inf I$, in favour of Eq. (37).

B.1   The initial condition

It is convenient to study not only $\mathbb{E}[E_{\max}(t)]$ – the expectation over different instances of the REM – but also the expectation given the starting energy $\mathbb{E}[E_{\max}(t)|E(0)]$. A particular case that will be useful in the following, is when the dynamics starts at the threshold energy,

$$\mathcal{E}(t) := \mathbb{E}[E_{\max}(t)|E(0) \approx \inf I] \,. \tag{39}$$

From $\mathcal{E}(t)$ we can compute $E_{\max}(t)$ for any starting condition. If $E(0) < \inf I$, we have

$$\mathbb{E}[E_{\max}(t)|E(0)] = \begin{cases} E(0) & \text{if } t \ll e^{\beta(\inf I - E(0))} \\ \mathcal{E}(t - e^{\beta(\inf I - E(0))}) \approx \mathcal{E}(t) & \text{if } t \gg e^{\beta(\inf I - E(0))} \,, \end{cases} \tag{40}$$

whereas if $E(0) > \inf I$ and $t \gg N$, we get

$$\mathbb{E}[E_{\max}(t)|E(0)] = \max(E(0), \mathcal{E}(t)) \,. \tag{41}$$

$\mathcal{E}(0)$ can be regarded as the energy typically reached after at time of order $t_1 = N$ (see Sec. 4). Hence, studying $\mathcal{E}(t)$ is equivalent to looking at $\mathbb{E}[E_{\max}(t_1, t)]$, so disregarding the initial part of the dynamics (until reaching the first local minimum). This time delay is small compared to the other time scales in play.

B.2   Corrections

B.2.1   General corrections

The main finite-size corrections we treat are the following:

- Finite-size corrections to extreme value results.

- Typical depth of traps.

- Alternative paths for reaching high energies.

Further corrections that we do not take into account are stressed in Sec. B.3.

Finer analysis of traps and their barriers Consider a local minimum of the energy landscape and denote its energy by E. Then E is simply the minimum of $N + 1$ energies from ρ_{gauss}, which is close to $\inf I$. However, attention should be paid to the energy of its lowest neighbour, which is not simply $\inf I$, since it is conditioned to be at energy higher than E. The average energy $\phi(E)$ of the lowest neighbour of a configuration with energy E should then be rewritten as

$$\phi(E) = \frac{\sqrt{N} \int_{\frac{E}{\sqrt{N}}}^{\infty} x(1 - \text{erfc}(x))^{N-1} e^{-\frac{x^2}{2}} \, \mathrm{d}x}{\int_{\frac{E}{\sqrt{N}}}^{\infty} (1 - \text{erfc}(x))^{N-1} e^{-\frac{x^2}{2}} \, \mathrm{d}x} \, . \tag{42}$$

The function $\phi(E)$ goes down from an initial value a bit higher than $\inf I$, when $E = \inf I$, to $\inf I$, when E is a few standard deviations of $\inf I$ lower than it. In particular, by the end of aging we can safely write $\phi(\mathcal{E}_{\min}(t)) = \inf I$, since the system will have reached the lowest-energy configuration.

Using Eq. (42), the exit time of a trap, that corresponds approximately to the mean time required to go to the lowest neighbour, is rewritten as

$$t_{\text{exit}} = N e^{\beta(\phi(E) - E)} \, . \tag{43}$$

Aging regime During the aging stage, since most of the time is spent in the lowest-energy state visited so far, the total elapsed time can be expressed as

$$t \approx t_{\text{exit}}^{(\text{aging})} \approx N e^{\beta[\phi(\mathcal{E}_{\min}(t)) - \mathcal{E}_{\min}(t)]} \, . \tag{44}$$

Hence, taking into account the relations in App. A.1, the aging now ends after a time of order

$$t_4' = N e^{\beta(\inf I + \beta_i N)} \left(\sqrt{2\pi N} \beta_i e^{-\gamma} \right)^{-\frac{\beta}{\beta_i}} \, , \tag{45}$$

where we also centred the Gumbel distribution, which gives the γ term, and used $\phi(\mathcal{E}_{\min}(t_4')) \approx \inf I$. Retaining just the main order gives

$$t_4^{(0)} = e^{\beta(\inf I + \beta_i N)} \, . \tag{46}$$

Time at lowest neighbours The time spent at lowest neighbours of traps is about Nn_t, since, due to the Metropolis rate (4), the time spent before leaving them is of order N. Recall that if one draws Nn_t energies from ρ_{gauss}, their minimum is distributed as the minimum of n_t lowest neighbours (see also Sec. 4.2.4). Then, according to Eqs. (30) and (32) in Sec. A.1, defining $y = \ln(Nn_t)$, one has

$$\mathcal{E}_{\min}(t) = \sqrt{2Ny} \left(1 - \frac{\ln y + \ln(4\pi) - 2\gamma}{4y} \right) \, .$$

The second term on the right hand side is small, so it can be treated as a perturbation. To the first order, one obtains

$$\ln(Nn_t) = \frac{\mathcal{E}_{\min}^2(t)}{2N} \left(1 + N \frac{\ln \frac{\mathcal{E}_{\min}^2(t)}{2N} + \ln(4\pi) - 2\gamma}{\mathcal{E}_{\min}^2(t)} \right) \, . \tag{47}$$

Records via other neighbours As most of the time is spent at minima, a new record E for the maximum energy can be attained either transiting directly from the minimum, E_{trap}, or via one of its neighbours. Let us denote by $A(E)$ the probability that E is reached from the trap or its neighbours divided by the probability that it is reached directly from the trap.

The probability of reaching E directly from the trap is given by the time spent there,

$$N \exp(\beta(\phi(E_{\text{trap}}) - E_{\text{trap}})),$$

times the transition rate $\exp(\beta(E_{\text{trap}} - E))/N$. Let us consider the lowest neighbour, $\phi(E_{\text{trap}})$, which typically has two neighbours (including the trap) with lower energy. It has exactly half

16

the transition probability to E as directly from the trap, because the less time spent in it, $N/2$, cancels out with the higher transition rate towards E. The second-lowest energy has typically three lower neighbours, so by the same reasoning it contributes as a third of the direct transition from the trap. Continuing this way we get a total transition rate of H_n times the one directly from the trap, where H_n is the harmonic partial sum.

However, this reasoning only concerns the neighbours of the trap with energies lower than E. In the general case we get

$$A(E) \approx \max \left(\ln \left(N \operatorname{erfc} \left(\frac{-E}{\sqrt{N}} \right) \right), 1 \right), \tag{48}$$

where we approximate H_n by $\ln n$ and recall the fact that transiting directly from the trap is always possible to avoid errors for E close to $\inf I$.

B.2.2 Delayed maximum

Long-time aging regime Note that most of the lowest neighbours of visited traps are at energy $\phi(\inf I)$. Then, due to the previous observations $\mathcal{E}(t)$ should satisfy, analogously to Sec. 4.2.5,

$$\begin{aligned} 1 &\simeq N n_t \int_{\mathcal{E}(t)}^{\infty} A(E) q_{\phi(\inf I), E} N \rho_{\text{gauss}}(E) \mathrm{d}E \\ &\approx e^{-\gamma} \sqrt{\frac{2\pi \mathcal{E}_{\min}^2(t)}{N}} e^{\frac{\mathcal{E}_{\min}^2(t)}{2N}} A(\mathcal{E}(t)) \int_{\mathcal{E}(t)}^{\infty} \frac{1}{N} e^{\beta(\phi(\inf I) - E)} N \rho_{\text{gauss}}(E) \mathrm{d}E \\ &\approx e^{-\gamma} A(\mathcal{E}(t)) \frac{\sqrt{N} |\mathcal{E}_{\min}(t)|}{(\mathcal{E}(t) + \beta N) \sqrt{2\pi}} \sqrt{\frac{2\pi}{N}} e^{\frac{\mathcal{E}_{\min}^2(t)}{2N}} e^{\beta(\phi(\inf I) - \mathcal{E}(t))} e^{-\frac{\mathcal{E}^2(t)}{2N}} \\ &\approx e^{-\gamma} B(\mathcal{E}(t)) \frac{|\mathcal{E}_{\min}(t)|}{\beta N} e^{\frac{\mathcal{E}_{\min}^2(t)}{2N}} e^{\beta(\phi(\inf I) - \mathcal{E}(t))} e^{-\frac{\mathcal{E}^2(t)}{2N}}, \end{aligned} \tag{49}$$

where $q_{i,j}$ is the Metropolis rate from state i to state j, and in the last step we set

$$B(E) := \frac{A(E)}{1 + \frac{E}{\beta N}}. \tag{50}$$

Notice that the subexponential terms are only very small perturbations, so one can still solve the transcendental equation. Then, combining (44) and (49) yields to the leading order

$$\mathcal{E}^{(0)}(t) \approx -\beta N + \sqrt{(\beta N + \inf I)^2 - 2\frac{1}{\beta} \inf I \ln t + \left(\frac{1}{\beta} \ln t \right)^2}. \tag{51}$$

At this point, one can use Eqs. (50) and (51) in the bottom line of Eq. (49), to approximate the solution recursively, obtaining

$$\begin{aligned} \mathcal{E}(t) \approx -\beta N + \Big(& \beta^2 N^2 + 2\beta N \phi(\inf I) + (\mathcal{E}_{\min}(t))^2 \\ & + 2N \ln \frac{\mathcal{E}_{\min}(t)}{\beta N} + 2N \ln B \left(\mathcal{E}^{(0)}(t) \right) - 2N\gamma \Big)^{\frac{1}{2}}. \end{aligned} \tag{52}$$

where we used that $\phi(\mathcal{E}_{\min}) = \inf I$, so this expression can only be valid deep in the aging regime.

Equilibrium regime After the end of aging ($t > t_4'$, Eq. (45)) we use that the mean time spent per state stations at its value at the end of aging $\frac{t_4'}{n_{t_4'}}$. Also, at the end of aging one has

$Nn_{t'_4} \approx e^{\frac{N\beta_i{}^2}{2}}$. Hence, following Eq. (45),

$$
1 \simeq t\frac{Nn_{t'_4}}{t'_4} \int_{\mathcal{E}(t)}^{\infty} A(E)q_{\phi(\inf I),E}N\rho_{\text{gauss}}(E)\mathrm{d}E \approx
$$

$$
\approx \frac{t}{N}e^{-\beta(\beta_i N+\inf I)}\left(\sqrt{2\pi N}\beta_i e^{-\gamma}\right)^{\frac{\beta}{\beta_i}}e^{\frac{N\beta_i^2}{2}}A(\mathcal{E}(t))\cdot \int_{\mathcal{E}(t)}^{\infty} e^{\beta(\phi(\inf I)-E)}\rho_{\text{gauss}}(E)\mathrm{d}E \approx
$$

$$
\approx \frac{\sqrt{N}A(\mathcal{E}(t))}{(\mathcal{E}(t)+\beta N)\sqrt{2\pi}}\left(\sqrt{2\pi N}\beta_i e^{-\gamma}\right)^{\frac{\beta}{\beta_i}}\frac{t}{N}e^{-\beta\beta_i N}e^{\frac{N\beta_i^2}{2}}\cdot e^{-\beta\mathcal{E}(t)}e^{-\frac{\mathcal{E}^2(t)}{2N}}e^{\beta(\phi(\inf I)-\inf I)} \approx
$$

$$
\approx C(\mathcal{E}(t))\left(\sqrt{2\pi N}\beta_i\right)^{\frac{\beta}{\beta_i}-1}e^{-\frac{\gamma\beta}{\beta_i}}\frac{t}{N}e^{-\beta\beta_i N}e^{\frac{N\beta_i^2}{2}}e^{-\beta\mathcal{E}(t)}e^{-\frac{\mathcal{E}^2(t)}{2N}}e^{\beta(\phi(\inf I)-\inf I)},
$$
(53)

where in the last step we set

$$
C(E) := A(E)\frac{\beta_i}{\beta+\frac{E}{N}}.
$$
(54)

Like in the previous paragraph, neglecting subexponential factors and solving (53) for $\mathcal{E}(t)$ gives the main order

$$
\frac{\mathcal{E}^{(0)}(t)}{N} = -\beta + \sqrt{(\beta-\beta_i)^2+\frac{2}{N}\ln t},
$$
(55)

which is (accidentally) the same as (26). Then the final result is

$$
\frac{\mathcal{E}(t)}{N} = -\beta + \left((\beta-\beta_i)^2+\frac{2}{N}\ln\frac{t}{N}+\frac{2}{N}\ln C\left(\mathcal{E}^{(0)}(t)\right)\right.
$$

$$
\left.+\frac{\beta-\beta_i}{N\beta_i}\ln\left(2\pi N\beta_i{}^2\right)-\frac{2\gamma\beta}{N\beta_i}+2\frac{\beta}{N}(\phi(\inf I)-\inf I)\right)^{\frac{1}{2}}.
$$
(56)

We should note that the piecewise concatenation of these functions is not expected to give good results around t'_4. That is because the right hand side of (53) should actually be the same replacing t by $t-t'_4$ and adding the right hand side of (49). However, the transition between the two expressions happens very quickly – just within a few times t'_4, which is nearly invisible on the semi-logarithmic plot. What is more, the expression should be further smoothed by the randomness of t'_4, which we took for constant.

B.2.3 Original maximum

Now that we know $\mathcal{E}(t)$, we can use it to deduce $\mathbb{E}[E_{\max}(t)|E(0)]$ and then obtain $\mathbb{E}[E_{\max}(t)]$ by integrating over $E(0)$. Using (40) we get

$$
\mathbb{E}[E_{\max}(t)] = \mathbb{E}[\mathbb{E}[E_{\max}(t)|E(0)]] \approx
$$

$$
\approx \int_{\inf I-T\ln t}^{\infty}\max(\mathcal{E}(t),E)\rho_{\text{gauss}}(E)\mathrm{d}E + \int_{-\infty}^{\inf I-T\ln t}E\rho_{\text{gauss}}(E)\mathrm{d}E =
$$

$$
= \left(\text{erfc}\left(\frac{\inf I-T\ln t}{\sqrt{N}}\right)-\text{erfc}\left(\frac{\mathcal{E}(t)}{\sqrt{N}}\right)\right)\mathcal{E}(t)+\int_{|\mathcal{E}(t)|}^{T\ln t-\inf I}E\rho_{\text{gauss}}(E)\mathrm{d}E \approx
$$

$$
\approx \left(1-\text{erfc}\left(\frac{\mathcal{E}(t)}{\sqrt{N}}\right)\right)\mathcal{E}(t)+\sqrt{\frac{N}{2\pi}}e^{-\frac{\mathcal{E}(t)^2}{2N}},
$$

i.e.

$$
\mathbb{E}[E_{\max}(t)] \approx \left(1-\text{erfc}\left(\frac{\mathcal{E}(t)}{\sqrt{N}}\right)\right)\mathcal{E}(t)+\sqrt{\frac{N}{2\pi}}e^{-\frac{\mathcal{E}(t)^2}{2N}},
$$
(57)

which can be made explicit by plugging (45), (52) and (56) in it. This expression approaches $\mathcal{E}(t)$ for large t and remains small as long as $\mathcal{E}(t)<0$.

18

B.3 Outlook on the finite-N corrections

We analysed preasymptotic corrections that should be taken into account in order to compare our predictions with numerical simulations. Because these corrections come from different sources, it was not possible quantify their magnitude as some order of N. As we show at the end of this section, the corrections we accounted for are sufficient for a good match with numerical simulations. Still, it should be noted that there are several effects which we did not account for.

Difference between $E_{\min}(t)$ and $E(t)$. An important source of error are the deviations from the theory of $E_{\min}(t)$, which were observed in [5]. Since our calculations are based on $E_{\min}(t)$, those are likely to have an impact here, too.

Asymptotic statistics. Eqs. (49) and (53) are only valid up to a factor of order 1 on the left hand side. Moreover, we arbitrarily used mean values for both $\inf I$ and the Gumbel distribution, which should give errors on the effective value adapted to our observable of the order of their standard deviations.

Time shift. The correction introduced by using ϕ [Eq. (42)] should also slightly impact the end of aging [Eq. (45)] and thereby the whole post-aging regime. The same holds for the use of the asymptotic relation $E_{\min}(\infty) \approx -\beta_c N$.

Returns. By considering the energies without memory of the trajectory's history, we disregarded returns to recently visited traps, which dominate the dynamics at subexponential time scales [5]. Those should effectively increase the exit times from traps.

By comparing with numerical simulations (Fig. 4–left), one sees that the effects we neglected seem sufficiently unimportant. Theoretical and numerical curves follow the same trend without using any free parameter.

Convergence to the asymptotic behaviour In figure Fig. 4–right we show the convergence of the finite-size $E_{\max}(t)$ to its asymptotic limit, which is different depending on the regime. For long times, $E_{\max}(t)/N$ becomes smaller with increasing N, and saturates at around $N = 1000$. At short times, the trend is inverted, and finite-size effects are sizeable for even larger system sizes. This slow speed of convergence explains why it is necessary to take into account finite-size corrections when comparing to numerical simulations.
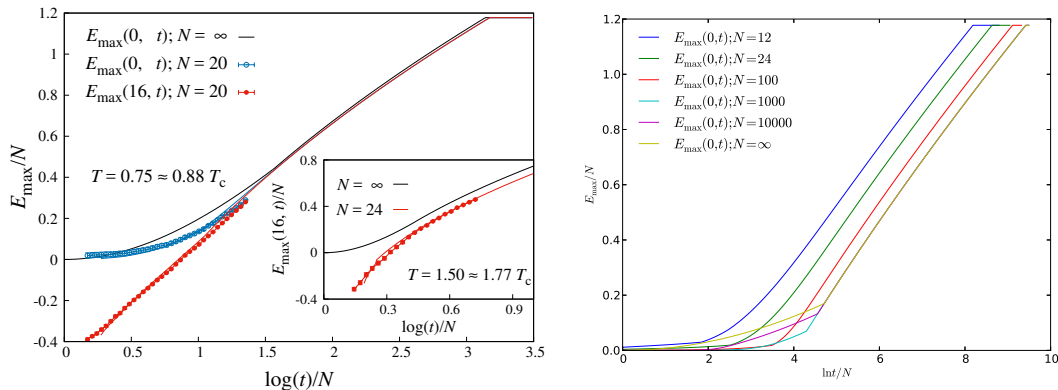

Figure 4: Finite-size predictions for the intensive maximum energy, $E_{\max}(t)/N$, in the REM as a function of the rescaled time $\theta = \ln t/N$. **Left:** Comparison between asymptotic solution, finite-size predictions and numerical simulations. We show analytical curve and numerical points for both $E_{\max}(0,t)$ and $E_{\max}(16,t)$. We show $T = 0.75 < T_c$ in the main figure, and $T = 1.50 > T_c$ in the **inset**. **Right:** Convergence of the finite-size predictions for $E_{\max}(0,t)$ to the asymptotic solution at $T = 0.25 \approx 0.29 T_c$. Note that the convergence is non-monotone and very slow, as we are deep in the glassy state. Moreover, the convergence after the equilibration time is faster.

C Records in a single trajectory

Discussing the evolution of the maximum in a single trajectory can simplify the concepts that we exposed throughout the article. In Fig. 5 we show the evolution of the maximum energy reached
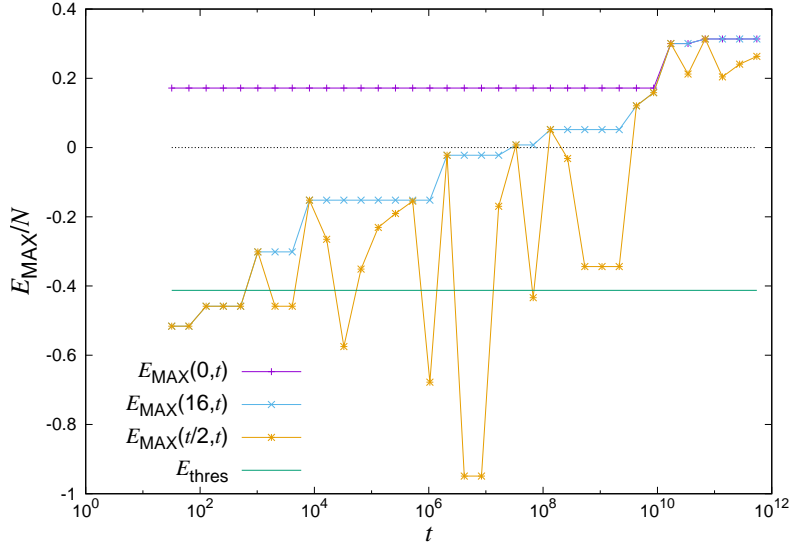

Figure 5: A random sample of the REM Metropolis dynamics, with $N = 20, T = 0.75$. We show the maximal attained energy, $E_{\max}(0, t)$, the one disregarding the first 16 time steps, $E_{\max}(16, t)$, and the one between $\frac{t}{2}$ and t, $E_{\max}(t/2, t)$. The green horizontal line represents the threshold energy E_{th} given by Eq. (37). The black dashed horizontal line is a reference for $E = 0$.

in a single trajectory. We use three different indicators: $E_{\max}(0, t)$ is the maximum energy since the beginning of the run, $E_{\max}(16, t)$ is a delayed maximum in which we wait 16 time steps before starting to record the maxima, in order to wait for the system to fall in the first local minimum, and $E_{\max}(t/2, t)$ is the maximum energy in the time window $[t/2, t]$. The latter has the advantage of showing very well the separation of time scales that occurs during the dynamics, since it takes into account increasingly larger time windows and excludes configurations visited long ago. In the run depicted, the initial state has $E(0) \approx 0.2N$, so $E_{\max}(0, t)$ sticks to that value for a long time. Instead, $E_{\max}(16, t)$ starts from a negative value, since in 16 time steps the system had the time to explore a portion of the landscape, lowering its energy. After around 10^{10} MC steps $E_{\max}(16, t)$ reaches $E(0)$, and, from that moment on, $E_{\max}(16, t) = E_{\max}(0, t)$. This is what would be the time t_3, marking the beginning of the First Records regime.

$E_{\max}(t/2, t)$ is the only indicator that can also decrease. It is equal to $E_{\max}(16, t)$ either at the beginning of the run, when the system is stuck in the initial trap, or when new records are hit. As stated previously, $E_{\max}(t/2, t)$ is a good separator of time scales. In fact, several times during the aging stage $E_{\max}(t/2, t)$ drops below inf I, indicating that the dynamics has accepted no move and has remained in the current trap for the time interval considered.

References

[1] P. G. Debenedetti and F. H. Stillinger. Supercooled liquids and the glass transition. *Nature*, 410:259–267, 2001.

[2] A. Cavagna. Supercooled liquids for pedestrians. *Phys. Rep.*, 476(4):51 – 124, 2009.

[3] T. Maimbourg, J. Kurchan, and F. Zamponi. Solution of the dynamics of liquids in the large-dimensional limit. *Phys. Rev. Lett.*, 116(1):015902, 2016.

[4] A. Crisanti and F. Ritort. Activated processes and inherent structure dynamics of finite-size mean-field models for glasses. *EPL (Europhysics Letters)*, 52(6):640, 2000.

[5] M. Baity-Jesi, G. Biroli, and C. Cammarota. Activated aging dynamics and effective trap model description in the random energy model. *J. Stat. Mech. Theory Exp.*, (1):013301, 2018.

[6] D. A. Stariolo and L. F. Cugliandolo. Relaxation dynamics of the Ising p-spin disordered model with finite number of variables. *arXiv:1904.10731*, 2019.

[7] J. C. Dyre. Master-equation appoach to the glass transition. *Phys. Rev. Lett.*, 58:792–795, 1987.

[8] J.-P. Bouchaud. Weak ergodicity breaking and aging in disordered systems. *J. Phys. I France*, 2:1705–1713, 1992.

[9] J.-P. Bouchaud and D. S. Dean. Aging on Parisi's tree. *J. Phys. I France*, 5(3):265–286, 1995.

[10] J. C. Dyre. Energy master equation: A low-temperature approximation to Bässler's random-walk model. *Phys. Rev. B*, 51:12276–12294, 1995.

[11] G. Ben-Arous, A. Bovier, and V. Gayrard. Aging in the Random Energy Model. *Phys. Rev. Lett.*, 88:087201, 2002.

[12] G. Ben-Arous, A. Bovier, and J. Černỳ. Universality of the REM for dynamics of mean-field spin glasses. *Comm. Math. Phys.*, 282(3):663–695, 2008.

[13] V. Gayrard. Convergence of clock processes and aging in metropolis dynamics of a truncated REM. *Ann. Henri Poincaré*, 17(3):537–614, 2016.

[14] V. Gayrard. Aging in Metropolis dynamics of the REM: a proof. *Probab. Theory Related Fields*, 2018.

[15] E. J. Gumbel. *Statistics of Extremes*. Columbia University Press, New York, 1958.

[16] G. Schehr and S. N. Majumdar. Exact record and order statistics of random walks via first-passage ideas. In R. Metzler, G. Oshanin, and S. Redner, editors, *First-passage phenomena and their applications*, pages 226–251. World Scientific, 2013.

[17] C. Godrèche, S. N. Majumdar, and G. Schehr. Record statistics of a strongly correlated time series: random walks and Lévy flights. *J. Phys. A*, 50(33):333001, 64, 2017.

[18] I. M. Johnstone. Multivariate analysis and Jacobi ensembles: Largest eigenvalue, Tracy–Widom limits and rates of convergence. *Ann. Statist.*, 36(6):2638–2716, 12 2008.

[19] C. A. Tracy and H. Widom. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.*, 159(1):151–174, 1994.

[20] C. A. Tracy and H. Widom. On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.*, 177(3):727–754, 1996.

[21] S. N. Majumdar and G. Schehr. Top eigenvalue of a random matrix: large deviations and third order phase transition. *J. Stat. Mech. Theory Exp.*, (1):P01012, 31, 2014.

[22] C. Monthus and J.-P. Bouchaud. Models of traps and glass phenomenology. *J. Phys. A*, 29(14):3847, 1996.

[23] B. Derrida. Random-energy model: Limit of a family of disordered models. *Phys. Rev. Lett.*, 45:79–82, 1980.

[24] B. Derrida. Random-energy model: An exactly solvable model of disordered systems. *Phys. Rev. B*, 24:2613–2626, Sep 1981.

[25] A. Billoire, L. Giomi, and E. Marinari. The mean-field infinite range $p = 3$ spin glass: Equilibrium landscape and correlation time scales. *EPL (Europhysics Letters)*, 71(5):824, 2005.

[26] M. Baity-Jesi, A. Achard-de Lustrac, and G. Biroli. Activated dynamics: an intermediate model between REM and p-spin. *Phys. Rev. E*, 98:012133, 2018.

[27] T. Castellani and A. Cavagna. Spin-glass theory for pedestrians. *J. Stat. Mech.*, 2005:P05012, 2005.

[28] L. F. Cugliandolo and J. Kurchan. Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model. *Phys. Rev. Lett.*, 71:173–176, 1993.

[29] A. Crisanti and H.-J. Sommers. Thouless-Anderson-Palmer approach to the spherical p-spin spin glass model. *J. Phys. I France*, 5(7):805–813, 1995.

[30] T. Rizzo. Replica-symmetry-breaking transitions and off-equilibrium dynamics. *Phys. Rev. E*, 88:032135, Sep 2013.

[31] J. D. Bryngelson and P. G. Wolynes. Intermediates and barrier crossing in a random energy model (with applications to protein folding). *J. Phys. Chem.*, 93(19):6902–6915, 1989.

[32] S. Boettcher and P. Sibani. Comparing extremal and thermal explorations of energy landscapes. *Eur. Phys. J. B*, 44(3):317–326, 2005.

[33] S. Boettcher and P. Sibani. Ageing in dense colloids as diffusion in the logarithm of time. *J. Phys. Condens. Matter.*, 23(6):065103, 2011.

[34] D. M. Robe, S. Boettcher, P. Sibani, and P. Yunker. Record dynamics: Direct experimental evidence from jammed colloids. *EPL (Europhysics Letters)*, 116(3):38003, 2016.