

**Weierstraß-Institut**  
**für Angewandte Analysis und Stochastik**  
**Leibniz-Institut im Forschungsverbund Berlin e. V.**

Preprint

ISSN 0946 – 8633

**Extremal aging for trap models**

Onur Gün

submitted: December 4, 2013

Weierstrass Institute  
Mohrenstr. 39  
10117 Berlin  
Germany  
E-Mail: onur.guen@wias-berlin.de

No. 1888  
Berlin 2013



---

2010 *Mathematics Subject Classification.* 82C44, 82D30, 60G70.

*Key words and phrases.* Random walk, random environment, REM, dynamics of spin glasses, aging, extremal processes.

Edited by  
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)  
Leibniz-Institut im Forschungsverbund Berlin e. V.  
Mohrenstraße 39  
10117 Berlin  
Germany

Fax: +49 30 20372-303  
E-Mail: [preprint@wias-berlin.de](mailto:preprint@wias-berlin.de)  
World Wide Web: <http://www.wias-berlin.de/>

ABSTRACT. In the seminal work [5], Ben Arous and Černý give a general characterization of aging for trap models in terms of  $\alpha$ -stable subordinators with  $\alpha \in (0, 1)$ . Some of the important examples that fall into this universality class are Random Hopping Time (RHT) dynamics of Random Energy Model (REM) and  $p$ -spin models observed on exponential time scales. In this paper, we explain a different aging mechanism in terms of *extremal processes* that can be seen as the extension of  $\alpha$ -stable aging to the case  $\alpha = 0$ . We apply this mechanism to the RHT dynamics of the REM for a wide range of temperature and time scales. The other examples that exhibit extremal aging include the Sherrington Kirkpatrick (SK) model and  $p$ -spin models [6, 9], and biased random walk on critical Galton-Watson trees conditioned to survive [11].

## 1. INTRODUCTION

A striking feature common to many disordered systems is that in an out-of-equilibrium phase convergence to equilibrium is very slow, and it gets slower as the system gets ‘older’. This phenomena is usually called *aging* and has been studied extensively in physics community, both experimentally and theoretically. *Spin glasses* constitute an important class of disordered systems where aging occurs, and the dynamics of mean field spin glasses is one of the focal points of this article.

On the theoretical side, *trap models* have become the central paradigm to study aging properties of disordered systems. They are simple enough to track analytically and many examples exhibit aging that can be established in a mathematically rigorous way. The general description of trap models is the following. Let  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  be a countable, connected graph where  $\mathcal{V}$  and  $\mathcal{E}$  denote the set of vertices and edges, respectively. Consider a random collection of positive numbers  $\tau = \{\tau_x : x \in \mathcal{V}\}$ . A *trap model*,  $(X(t) : t \geq 0)$ , is a continuous time Markov process on  $\mathcal{V}$  whose dynamics is described as follows: at a vertex  $x \in \mathcal{V}$ , it waits for an exponential time with mean  $\tau_x$  and then moves to one of the neighbors of  $x$  uniform at random. Here, one can view a vertex  $x$  as a *trap* and  $\tau_x$  as the *depth* of the trap  $x$ . Let  $(Y(i) : i \in \mathbb{N}_0)$  denote the discrete time simple random walk on  $\mathcal{V}$ . Another view of trap models is that  $X$  follows  $Y$  that gets trapped at vertices, collecting exponential random variables, and hence, is a time change of it. This time-change process  $S$ , called the *clock-process*, is defined for  $k \in \mathbb{N}$  by

$$S(k) = \sum_{i=0}^{k-1} \tau_{Y(i)} e_i, \quad t \geq 0 \quad (1.1)$$

where  $(e_i : i \in \mathbb{N}_0)$  is a collection of i.i.d. exponential mean-one random variables. The clock process and the Markov chain  $Y$  completely describes  $X$  via

$$X(t) = Y(S^{\leftarrow}(t)) \quad (1.2)$$

where  $S^{\leftarrow}$  is the generalized right inverse of  $S$ .

One central question in the study of the trap models is the analysis of the clock process, more precisely, the influence of the random trapping landscape. If the trapping landscape is very heterogenous on certain large times, one expects that the main contribution to the clock process come from few ‘deep traps’. Moreover, if the graph is transient enough on these time scales, the walk, upon leaving a deep trap, will get thoroughly lost before finding a new deep trap, and as a result, the jumps of the clock process will be i.i.d. with the completely annealed distribution as the common distribution. Summation of random variables being dominated by few large ones signifies a stable convergence for the clock process. In the situations just described, the properly rescaled clock process converges to an  $\alpha$ -stable subordinator with  $0 < \alpha < 1$ . From this, one can easily deduce that the system exhibits aging, namely, the two-time correlation function

$$R(t_1, t_2) := \mathbb{P}(X_n(t_1) = X_n(t_2) | \tau) \quad (1.3)$$

can be approximated, for large  $t_1$  and  $t_2$  corresponding time scales of observation, by the probability that an  $\alpha$ -stable subordinator jumps over the interval  $[t_1, t_2]$ . The latter is a function of the ratio  $t_1/t_2$  and is given by the classical arcsine law. The existence of a correlation function that depends only on the trajectory of  $X$  on a time interval  $[t_1, t_2]$  whose large time limit is a non-trivial function of the ratio  $t_1/t_2$  is usually described as aging both in mathematics and physics literature.

The picture described in the above paragraph has been made rigorous in [5] in the context of sequence of diverging finite graphs. In this article, too, we will study a growing sequence of finite graphs. The reason to study trap models on finite graphs is twofold: Firstly, it allows one to employ potential theoretical methods strongly. Secondly, our main motivation, the mean field spin glasses, is in this setup.

Our main goal is to understand the one end of the spectrum of the  $\alpha$ -stable aging, namely, the case  $\alpha = 0$ . More precisely, we are interested in the situations where as the graphs grow, the heavy-tail index  $\alpha$  converges to 0. In this case, heterogeneity becomes even stronger in the sense that the main contribution to the clock process comes from the ‘deepest’ of the deep traps the walk can find. Hence, the limiting clock process has a structure of a ‘record’ process, namely, it is an *extremal process*. More precisely, after a linear rescaling, contributions from deep traps still grow on an exponential scale. As a result, firstly, the maximal term dominates, and secondly, one has to perform a further non-linear scaling, coherent with case of sum of i.i.d. random variables with slowly varying tails (see e.g. [12, 14]).

In the spirit of [5] we will give a set of conditions on the trapping landscape and the potential theoretical properties of the graph which will ensure that the clock process can be approximated, for appropriate large time scales, by an extremal process. Next, we will describe two additional conditions that guarantees that the two-time correlation function in (1.3) can be approximated using extremal processes. Here, in order to get a non-trivial limit, one has to slightly enlarge the ratio of the two times with the volume, due to the non-linear scaling of clock process. We have called this type of aging *extremal aging*.

Let us now discuss the trap model dynamics of spin glass models, also known as Random Hopping Time (RHT) dynamics. Our focus is on mean field models. The simplest mean field spin glass model is Random Energy Model (REM) introduced by Derrida in [13], where there are no correlation between energies. The first aging results for REM were proved in [3] and [4] for time scales only slightly shorter than equilibration times, using renewal techniques. Later, [5] proved that the dynamics of REM model ages with the arcsine law as the limiting aging function. The time scales where REM exhibits aging were later extended in [10]. In the second part of this article we extend these results to include extremal aging. We will let the temperature vary with the volume of the system in order to get a richer picture and prove that the dynamics of REM exhibits extremal aging for a wide range of temperature and time scales. More precisely, extremal aging occurs for fixed positive temperatures and subexponential time scales (in dimension of the system); for temperatures vanishing with the volume and for exponential, subexponential and superexponential time scales; and, for temperatures diverging with the volume and for (very) subexponential time scales. These results also signify that the aging is related to how transient the relaxation of the system is when far from equilibrium and it might have little to do with the actual equilibrium properties of the system. The occurrence of aging even in infinite temperature is a strong demonstration of this fact.

Let us mention the results on correlated mean field spin glass models. The arcsine law as an aging scheme, surprisingly, proved to be true even for some correlated spin glasses. In [2], authors proved that the same aging mechanism essentially holds true for short enough time scales for the  $p$ -spin models. Later, finer results on the aging of  $p$ -spin models were achieved in [8].

Our original motivation to study the  $\alpha = 0$  case, or rather the dynamics of mean field spin glasses on subexponential time scales, stemmed from the aim of extending the REM universality to the dynamics

of the Sherrington Kirkpatrick (SK) model. The results on the statistical properties of the SK model (see [1]) indicated that in order to not feel the correlations in the model, one has to investigate the dynamics on subexponential time scales. The dynamics on such time scales are in the category of extremal aging. In [6] we proved that on subexponential time scales the clock process of the SK model and  $p$ -spin SK models converge to extremal processes and these systems exhibit extremal aging. Finer analysis was carried out later in [9].

Recently, extremal aging has been observed for a type of model different than spin glass models. In [11], it was proved that biased random walk on critical Galton-Watson trees conditioned to survive exhibits extremal aging.

The rest of the article is organized as follows: In Section 2 we describe precisely the models we study, give a set of conditions on the trapping landscape and potential theoretical properties of the graph, and prove that these conditions imply the convergence of the rescaled clock process to an extremal process. In Section 3 we set two additional conditions that leads to extremal aging. In Section 4 we prove our results on the dynamics of REM.

## 2. CONVERGENCE OF THE CLOCK PROCESS TO AN EXTREMAL PROCESS.

We start this section by introducing precisely the type of trap models we will study. Let  $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n)$ ,  $n \in \mathbb{N}$ , be a sequence of finite, connected graphs with  $\mathcal{V}_n$  and  $\mathcal{E}_n$  denoting the set of vertices and edges, respectively. We use the notation  $x \sim y$  for an edge  $(x, y) \in \mathcal{E}_n$ . For each  $n \in \mathbb{N}$  and vertex  $x \in \mathcal{V}_n$  we attach a positive number  $\tau_x$  which constitutes the "depth" of the trap at site  $x$ . We denote the collection of depths by  $\tau_n = \{\tau_x : x \in \mathcal{V}_n\}$ . We will call  $\tau_n$  the 'trapping landscape' or 'environment' and later we will choose  $\tau_n$  random. Given the trapping landscape  $\tau_n$  we define a continuous time Markov process  $\{X_n(t) : t \geq 0\}$  on  $\mathcal{V}_n$  whose transition rates are given by  $w_{xy}^{\tau_n} = \mathbb{1}\{x \sim y\} / (d_x \tau_x)$  where  $d_x = \#\{y \in \mathcal{V}_n : x \sim y\}$  is the degree of the vertex  $x$ . In other words, at a vertex  $x$ , the Markov process  $X_n$  waits an exponential time with mean  $\tau_x$  and then it moves to one of its neighbors chosen uniformly at random. We denote by  $\mathbb{P}_x$  and  $\mathbb{E}_x$  the distribution and expectation of  $X_n$  starting from  $x \in \mathcal{V}_n$ . We will always start  $X_n$  from an arbitrary but fixed vertex we denote by 0 that does not depend on  $\tau_n$  and write for simplicity  $\mathbb{P} = \mathbb{P}_0$  and  $\mathbb{E} = \mathbb{E}_0$ . Note that, although  $X_n$  depends on  $\tau_n$  we suppressed it in the notation.

For each  $n \in \mathbb{N}$  we take the trapping landscapes  $\tau_n \in (0, \infty)^{\mathcal{V}_n}$  random with  $P_n$  and  $E_n$  denoting its distribution and expectation, respectively. We embed all the random variables  $\tau_n$ ,  $n \in \mathbb{N}$ , independently into a common probability space and  $P$  and  $E$  stands for the distribution and expectation of this probability space. For the events that happens  $P$  almost surely, we will simply say 'for a.s. random environment  $\tau$ '.

For any  $n \in \mathbb{N}$ , let  $\{Y_n(i) : i \in \mathbb{N}_0\}$  be the simple random walk on the graph  $\mathcal{G}_n$ , that is, the discrete time Markov chain on  $\mathcal{V}_n$  with the transition probabilities  $p_{xy} = \mathbb{1}\{x \sim y\} / d_x$  and we set the starting point  $Y_n(0) = 0$ . We assume that the distribution of  $Y_n$  is defined in the probability space  $\mathbb{P}_0$ . We define the clock process  $S_n$  by setting  $S_n(u) = 0$  for  $u \in [0, 1)$  and for  $u \in [1, \infty)$

$$S_n(u) := \sum_{i=0}^{\lfloor u \rfloor - 1} \tau_{Y_n(i)} e_i \quad (2.1)$$

where  $\{e_i : i \in \mathbb{N}\}$  is an i.i.d. collection of exponential mean-one random variables independent from anything else. In words, for  $k \in \mathbb{N}$ ,  $S_n(k)$  is the time it takes for  $X_n$  to make  $k$  jumps. Clearly we have,

$$X_n(t) = Y_n(k) \quad \text{if } S_n(k) \leq t < S_n(k+1). \quad (2.2)$$

For  $\tau_n$  fixed,  $S_n$  is a random variable taking values in  $D([0, \infty))$ , the space of càdlàg function on  $[0, \infty)$ . Similarly, for any fixed  $T > 0$  the restriction of  $S_n$  to  $[0, T]$  is a random variable on  $D([0, T])$ , the space of càdlàg function on  $[0, T]$ .

We need now further notation. Let  $T_n$  be a stopping time for the chain  $Y_n$ . We define  $G_{T_n}^n(x, y)$ ,  $x, y \in \mathcal{V}_n$  to be the Green's function of  $Y_n$ , that is, the mean number of times  $Y_n$  visits  $y$  before  $T_n$ , started from  $x$ :

$$G_{T_n}^n(x, y) := \mathbb{E}_x \left[ \sum_{i=0}^{T_n-1} \mathbb{1}\{Y_n(i) = y\} \right]. \quad (2.3)$$

For  $A \subseteq \mathcal{V}_n$  we define the first hitting time of  $A$  by

$$H_n(A) := \inf\{i \geq 0 : Y_n(i) \in A\}. \quad (2.4)$$

For ease of notation we write  $G_A^n$  for  $G_{H_n(A)}^n$ . Finally, we say that a random subset  $A \subseteq \mathcal{V}_n$  is a percolation cloud with density  $\rho \in (0, 1)$  if  $x \in A$  with probability  $\rho$ , independently of all other vertices.

Presently, we set 4 conditions that are about the trapping landscape and the properties of the walk  $Y_n$  on  $\mathcal{V}_n$ . This set of very general, potential theoretical conditions will be used to prove our main result.

The first condition tells that a certain density of traps exceed a depth scale in a very heterogeneous way.

**Condition A:** For any  $n \in \mathbb{N}$  let  $\tau_n$  be i.i.d. in  $x \in \mathcal{V}_n$ . There exists a *depth rate scale*  $\alpha_n$ , a *depth scale*  $g_n$  and a *density scale*  $b_n$  with  $g_n \rightarrow \infty$  and  $\alpha_n, b_n \rightarrow 0$  as  $n \rightarrow \infty$  such that

$$b_n^{-1} P_n(\tau_x \geq u^{1/\alpha_n} g_n) \xrightarrow{n \rightarrow \infty} 1/u \quad (2.5)$$

uniformly in  $u$  on all compact subsets of  $(0, \infty)$ . Moreover, there exists a constant  $C$  such that for all  $u > 0$  and  $d > 0$ , for all  $n \in \mathbb{N}$

$$P_n(\tau_x \geq u^{1/\alpha_n} d g_n) \leq \frac{C b_n}{u d^{\alpha_n}}. \quad (2.6)$$

The next two conditions concern the potential theoretical properties of the graph.

**Condition B:** Let  $b_n$  be as in Condition A. Let  $A_n, n \in \mathbb{N}$  be a sequence of percolation clouds on  $\mathcal{V}_n$  with densities  $\rho b_n$  where  $\rho \in (0, \infty)$ . There exists a scale  $f_n$  with  $f_n \rightarrow \infty$  as  $n \rightarrow \infty$  and a constant  $\mathcal{K}_G$  independent of  $\rho$  such that for a.s. sequence  $A_n$

$$\max_{x \in A_n} |f_n^{-1} G_{A_n \setminus \{x\}}^n(x, x) - \mathcal{K}_G| \xrightarrow{n \rightarrow \infty} 0. \quad (2.7)$$

**Condition C:** Let  $A_n$  be as in Condition B. There exists a constant  $\mathcal{K}_r > 0$  such that for all  $s > 0$  and a.s. sequence  $A_n$

$$\max_{x \in A_n \cup \{0\}} \left| \mathbb{E}_x \left[ \exp\left(-\frac{s}{r_n} H_n(A_n \setminus \{x\})\right) \right] - \frac{\mathcal{K}_r \rho}{s + \mathcal{K}_r \rho} \right| \xrightarrow{n \rightarrow \infty} 0, \quad (2.8)$$

where  $r_n = f_n/b_n$ .

The last condition contains technical restrictions.

**Condition D:**

- (i) There exists a sequence of positive numbers  $\lambda_n$  and a positive constant  $\mathcal{K}_s$  such that for all  $T > 0$  and  $n$  large enough

$$\sum_{x \in \mathcal{V}_n} (e^{\lambda_n G_{T r_n}(0, x)} - 1) \leq \mathcal{K}_s \lambda_n T r_n \quad (2.9)$$

and  $\sum_{n=1}^{\infty} \exp(-c\lambda_n f_n) < \infty$  for some  $c > 0$ .

(ii)  $\alpha_n \log(f_n) \rightarrow 0$  as  $n \rightarrow \infty$ .

Now we introduce formally the extremal processes. Let  $F$  be a probability distribution function on  $(-\infty, \infty)$ . For  $l \in \mathbb{N}$  and  $0 \leq t_1 \leq \dots \leq t_l$  define the finite dimensional distributions

$$F_{t_1, \dots, t_l}(x_1, \dots, x_l) = F^{t_1} \left( \bigwedge_{i=1}^l x_i \right) F^{t_2 - t_1} \left( \bigwedge_{i=2}^l x_i \right) \dots F^{t_l - t_{l-1}} \left( \bigwedge_{i=l}^l x_l \right) \quad (2.10)$$

where  $\bigwedge$  stands for the minimum. The family of finite dimensional distributions defined by (2.10) is consistent and thus, by Kolmogorov extension theorem there exists a continuous time stochastic process  $W = (W(t) : t \geq 0)$  with these finite dimensional distributions.  $W$  is called an extremal process generated by  $F$  or  $F$ -extremal process.

We give another description of an extremal process. Let  $F$  be as in the previous paragraph. Assume that  $F$  is a continuous distribution with  $\text{supp}(F) = \mathbb{R}$ . Let  $N$  denote a Poisson Random Measure (PRM) on  $[0, \infty) \times (0, \infty)$  with mean measure  $dt \times \nu(dx)$  where  $\nu(x, \infty) = -\log F(x)$ . Let us denote by  $(t_j, \xi_j)$  the marks of  $N$ . Then if we define

$$W(t) = \max_{t_i \leq t} \xi_i, \quad t \geq 0 \quad (2.11)$$

$W$  is an  $F$ -extremal process. It is enough to check that  $W$  satisfies (2.10) for any  $0 \leq t_1 \leq \dots \leq t_l$  and for any  $x_1 \leq x_2 \leq \dots \leq x_l$ . By the continuity of  $F$  and the independence properties of a PRM we get

$$\begin{aligned} & P(W(t_1) \leq x_1, \dots, W(t_l) \leq x_l) \\ &= P(N([0, t_1] \times [x_1, \infty)) = 0) \dots P(N([t_{l-1}, t_l] \times [x_l, \infty)) = 0) \\ &= e^{-t_1 \nu(x_1, \infty)} \dots e^{-(t_l - t_{l-1}) \nu(x_l, \infty)} \\ &= F^{t_1}(x_1) \dots F^{t_l - t_{l-1}}(x_l). \end{aligned} \quad (2.12)$$

For details about extremal processes we refer readers to [15].

For our convergence results of clock processes we use two different topologies on  $D([0, T])$ , namely,  $M_1$  and  $J_1$  topologies. We will indicate the topology of weak convergence by simply writing ‘in  $D([0, T], J_1)$ ’ and ‘in  $D([0, T], M_1)$ ’. The essential difference between these two topologies is that while  $M_1$  topology allows approximating processes make several jumps in short intervals of time to produce one bigger jump of the limiting process, while  $J_1$  does not. See [16] for detailed information on these topologies.

Below is our main result about the convergence of clock processes to an extremal process.

**Theorem 1.** *Assume Conditions A-D are satisfied. Set  $t_n = f_n g_n$ . Then for a.s. random environment  $\tau$ , for any  $T > 0$*

$$\left( \frac{S_n(\cdot, r_n)}{t_n} \right)^{\alpha_n} \xrightarrow{n \rightarrow \infty} W(\cdot) \text{ in } D([0, T]; M_1) \quad (2.13)$$

where  $W$  is the extremal process generated by the distribution function  $F(x) = e^{-1/x}$ . Moreover, the above convergence holds in the stronger topology  $D([0, T]; J_1)$  if  $f(n) = 1$  and  $\mathcal{K}_G = 1$ .

**Remark 2.1.** *Theorem 1 can be stated in a more general way. Namely, Condition A can be generalized such that (2.5) is replaced by*

$$b_n P_n(\tau_x \geq L^{\leftarrow}(u h_n)) \xrightarrow{n \rightarrow \infty} \frac{1}{u} \quad (2.14)$$

and (2.6) by

$$P_n(\tau_x \geq L_n^{\leftarrow}(ud^{\alpha_n} h_n)) \leq \frac{Cb_n}{ud^{\alpha_n}} \quad (2.15)$$

where  $h_n$  is a diverging scale. Here, for each  $n \in \mathbb{N}$ ,  $L_n$  is positive, as  $u \rightarrow \infty$ ,  $L_n(u) \rightarrow \infty$  and for any  $\lambda > 0$

$$\frac{L_n(\lambda u)}{L_n(u)} \longrightarrow \lambda^{\alpha_n} \quad (2.16)$$

in a mildly uniform way in  $n$ . Also, in this setup one can have  $\alpha_n = 0$  which would mean that  $L_n$  is a slowly varying function for such an  $n$ . Then (2.13) in Theorem 1 becomes

$$\frac{L_n(S_n(\cdot r_n)/f_n)}{h_n} \Longrightarrow W(\cdot). \quad (2.17)$$

Since such a setup makes the notation very difficult and does not bring a new conceptual insight we preferred to use the current setup. Finally, note that, choosing  $L_n(u) = u^{\alpha_n}$  and  $g_n = h_n^{1/\alpha_n}$  gives Theorem 1.

In the rest of the current section we prove Theorem 1. We start by defining set of deep traps, very deep traps and shallow traps determined by the depth scale and the depth rate scale  $\alpha_n$  and  $g_n$ : for  $0 < \epsilon < M$

$$\begin{aligned} T_\epsilon^M(n) &:= \{x \in \mathcal{V} : \epsilon^{1/\alpha_n} g_n \leq \tau_x \leq M^{1/\alpha_n} g_n\} && \text{deep traps,} \\ T_M(n) &:= \{x \in \mathcal{V} : \tau_x > M^{1/\alpha_n} g_n\} && \text{very deep traps,} \\ T^\epsilon(n) &:= \{x \in \mathcal{V} : \tau_x < \epsilon^{1/\alpha_n} g_n\} && \text{shallow traps.} \end{aligned} \quad (2.18)$$

Let  $d_n(j)$  be sequence of times where a deep trap different from the last deep trap visited is found. We set  $d_n(0) = 0$  and for  $j \in \mathbb{N}$  define recursively

$$d_n(j) := \min \{i > d_n(j-1) : Y_n(i) \in T_\epsilon^M(n) \setminus \{Y_n(d_n(j-1))\}\}. \quad (2.19)$$

We define the process  $(U_n(j) : j \in \mathbb{N}_0)$

$$U_n(j) := Y_n(d_n(j)), \quad (2.20)$$

and  $\zeta_n$  as the last time the random walk finds a deep trap before  $Tr_n$ :

$$\zeta_n(T) := \max \{j : d_n(j) \leq Tr_n\}. \quad (2.21)$$

Let  $s_n(j)$  be the time  $X_n$  spends at  $U_n(j)$  between the first time it visits  $U_n(j)$  until it finds a different deep trap, that is,

$$s_n(j) = \sum_{i=d_n(j)}^{d_n(j+1)} e_i \tau_{Y_n(i)} \mathbb{1} \{Y_n(i) = U_n(j)\}. \quad (2.22)$$

We call  $s_n(j)$  the score of  $U_n(j)$ . Note that, for given environment  $\tau_n$  and  $U_n(j)$ , the expectation of  $s_n(j)$  over all the other random sources is

$$\tau_{U_n(j)} G_{T_\epsilon^M(n) \setminus \{U_n(j)\}}^n(U_n(j), U_n(j)). \quad (2.23)$$

Finally, we define  $m_n$ , the record process of  $s_n$ , for we expect it to be a good approximation of the clock process. For technical convenience we set  $m_n(0) = 0$  and  $m_n(1) = S_n(d_n(1))$ , and define for  $j \geq 2$

$$m_n(j) := \max_{i=1, \dots, j-1} s_n(i). \quad (2.24)$$

As a first step of the proof, we need to control the distribution of depths of deep traps visited. We introduce the following notation:

$$\rho_a^b := a^{-1} - b^{-1}, \quad 0 < a < b. \quad (2.25)$$

We need the following lemma which is from [5], stated in a slightly different way.



**Lemma 2.1.** (Lemma 2.5 on page 304 in [5]) Recall that  $A_n$  is a percolation cloud with density  $\rho b_n$  and assume Condition C. Let  $A_n^1$  and  $A_n^2$  be such that  $A_n^1 \cup A_n^2 = A_n$  and  $A_n^1 \cap A_n^2 = \emptyset$  and

$$\lim_{n \rightarrow \infty} \frac{|A_n^1|}{|A_n|} = \frac{\rho_1}{\rho}, \quad \lim_{n \rightarrow \infty} \frac{|A_n^2|}{|A_n|} = \frac{\rho_2}{\rho} \quad \text{with} \quad \rho_1 + \rho_2 = 1. \quad (2.26)$$

Then

$$\lim_{n \rightarrow \infty} \max_{x \in A_n \cup \{0\}} \left| \mathbb{P}_x \left( H_n(A_n^1 \setminus \{x\}) < H_n(A_n^2 \setminus \{x\}) \right) - \frac{\rho_1}{\rho} \right| = 0. \quad (2.27)$$

Let  $(\sigma_\epsilon^M(j) : j \in \mathbb{N})$  be an i.i.d. sequence of random variables with distribution

$$P(\sigma_\epsilon^M(j) \geq u) = \frac{\rho_u^M}{\rho_\epsilon^M}. \quad (2.28)$$

**Proposition 2.1.** The collection  $((\tau_{U_n(j)}/g(n))^{\alpha_n} : j \in \mathbb{N})$  converges weakly to  $(\sigma_\epsilon^M(j) : j \in \mathbb{N})$  as  $n \rightarrow \infty$ .

*Proof.* Since  $(U_n(j) : j \in \mathbb{N})$  is a Markov sequence, it is enough to show that for any  $j \in \mathbb{N}$ , as  $n \rightarrow \infty$

$$\mathbb{P} \left( (\tau_{U_n(j)}/g(n))^{\alpha_n} \geq u | U_n(j-1) \right) \longrightarrow \frac{\rho_u^\epsilon}{\rho_\epsilon^M}. \quad (2.29)$$

We have

$$\mathbb{P} \left( (\tau_{U_n(j)}/g(n))^{\alpha_n} \geq u | U_n(j-1) \right) = \mathbb{P}_x \left( H_n(T_u^M(n) \setminus \{x\}) < H_n(T_\epsilon^u(n) \setminus \{x\}) \right) \quad (2.30)$$

where  $x = U_n(j-1)$ . Note that  $T_\epsilon^M(n) = T_u^M(n) \cup T_\epsilon^u(n)$  and  $T_u^M(n) \cap T_\epsilon^u(n) = \emptyset$ . Moreover, by Condition A

$$\lim_{n \rightarrow \infty} \frac{|T_u^M(n)|}{|T_\epsilon^M(n)|} = \frac{\rho_u^M}{\rho_\epsilon^M} \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{|T_\epsilon^u(n)|}{|T_\epsilon^M(n)|} = \frac{\rho_\epsilon^u}{\rho_\epsilon^M}. \quad (2.31)$$

Hence, (2.30), Condition A and Lemma 2.1 finish the proof.  $\square$

**Lemma 2.2.** The collection  $((s_n(j)/t_n)^{\alpha_n} : j \in \mathbb{N})$  converges weakly to  $(\sigma_\epsilon^M(j) : j \in \mathbb{N})$  as  $n \rightarrow \infty$ .

*Proof.* Using the Markov property of the random walk we see that

$$\begin{aligned} & \mathbb{P} \left( (s_n(j)/t_n)^{\alpha_n} \geq u | s_n(1), \dots, s_n(j-1) \right) \\ &= \int_\epsilon^M \mathbb{P} \left( s_n(j) \geq u^{1/\alpha_n} t_n | \tau_{U_n(j)} = v^{1/\alpha_n} g_n, s_n(1), \dots, s_n(j-1) \right) \\ & \quad \times \mathbb{P} \left( (\tau_{U_n(j)}/g_n)^{\alpha_n} \in dv | s_n(1), \dots, s_n(j-1) \right). \end{aligned} \quad (2.32)$$

Note that the first term inside the integral in the above display is equal to

$$\mathbb{P} \left( s_n(j) \geq u^{1/\alpha_n} t_n | \tau_{U_n(j)} = v^{1/\alpha_n} g_n \right) \leq \left( \frac{u}{v} \right)^{1/\alpha_n} \frac{g_n}{t_n} G_{T_\epsilon^M(n) \setminus \{U_n(j)\}}^n(U_n(j), U_n(j)) \quad (2.33)$$

where we use Chebyshev's inequality and (2.23) for the upper bound. By Condition B we have  $G_{T_\epsilon^M(n) \setminus \{x\}}^n(x, x) = \mathcal{K}_G f_n(1 + o(1))$  as  $n \rightarrow \infty$  and the error term is uniformly bounded in  $x \in T_\epsilon^M(n)$ . Hence, recalling that  $f_n g_n = t_n$ , if  $u > v$  we have

$$\mathbb{P} \left( s_n(j) \geq u^{1/\alpha_n} t_n | \tau_{U_n(j)} = v^{1/\alpha_n} g_n \right) \longrightarrow 0. \quad (2.34)$$

Since  $\sum_{i=d_n(j)}^{d_n(j+1)} \mathbb{1} \{Y_n(i) = U_n(j)\}$  is always at least one and  $t_n = f_n g_n$ , using Condition D part (ii) one we have for  $u < v$

$$\mathbb{P} \left( s_n(j) \geq u^{1/\alpha_n} t_n | \tau_{U_n(j)} = v^{1/\alpha_n} g_n \right) \geq \mathbb{P}(e_1 v^{1/\alpha_n} \geq u^{1/\alpha_n} f_n) \longrightarrow 1. \quad (2.35)$$

Using (2.34) and (2.35) in (2.32), than applying Proposition 2.1 finishes the proof.  $\square$

We next show that the contribution from the shallow traps are negligible.

**Proposition 2.2.** *For any  $a, d > 0$  and  $T > 0$ , for any  $\delta > 0$  given,  $\tau$  a.s. for  $n$  large enough*

$$\mathbb{E} \left[ \sum_{i=0}^{Tr_n} \tau_{Y_n(i)} e_i \mathbb{1} \{Y_n(i) \in T^a(n)\} \mid \tau \right] \leq (a+d)^{1/\alpha_n} t_n. \quad (2.36)$$

*Proof.* We first prove that  $\tau$  a.s. for  $n$  large enough, for all  $j \in \mathbb{N}$

$$\mathbb{E} \left[ \sum_{i=1}^{Tr_n} e_i \tau_{Y_n(i)} \mathbb{1} \{Y_n(i) \in T_{a2^{-j\alpha_n}}^{a2^{-j\alpha_n} + \alpha_n}(n)\} \mid \tau \right] \leq (a+d/2)^{1/\alpha_n} t_n 2^{j\alpha_n - j + 1}. \quad (2.37)$$

By Condition A we have for all  $n \in \mathbb{N}$  and  $j \in \mathbb{N}$

$$p_{n,j} := P \left( \mathbb{1} \{x \in T_{a2^{-j\alpha_n}}^{a2^{-j\alpha_n} + \alpha_n}(n)\} \right) \leq \frac{C2^{j\alpha_n}}{a} b_n. \quad (2.38)$$

We have

$$\begin{aligned} & P \left( \mathbb{E} \left[ \sum_{i=1}^{Tr_n} e_i \tau_{Y_n(i)} \mathbb{1} \{Y_n(i) \in T_{a2^{-j\alpha_n}}^{a2^{-j\alpha_n} + \alpha_n}(n)\} \mid \tau \right] \geq (a+d/2)^{1/\alpha_n} t_n 2^{j\alpha_n - j + 1} \right) \\ &= P \left( \sum_{x \in \mathcal{V}_n} G_{Tr_n}(0, x) \tau_x \mathbb{1} \{x \in T_{a2^{-j\alpha_n}}^{a2^{-j\alpha_n} + \alpha_n}(n)\} \geq (a+d/2)^{1/\alpha_n} t_n 2^{j\alpha_n - j + 1} \right) \\ &\leq P \left( \sum_{x \in \mathcal{V}_n} G_{Tr_n}(0, x) \mathbb{1} \{x \in T_{a2^{-j\alpha_n}}^{a2^{-j\alpha_n} + \alpha_n}(n)\} \geq \left( \frac{a+d/2}{a} \right)^{1/\alpha_n} 2^{j\alpha_n} f_n \right) \end{aligned} \quad (2.39)$$

Using exponential Chebyshev inequality with  $\lambda_n$  from Condition D, (2.38) and the bound  $\log(1+x) \leq x$ , for  $x \geq 0$ ; we conclude that (2.39) is bounded above by

$$\exp \left( -\lambda_n \left( \frac{a+d/2}{a} \right)^{1/\alpha_n} 2^{j\alpha_n} f_n + \frac{C2^{j\alpha_n}}{a} b_n \sum_{x \in \mathcal{V}_n} (e^{\lambda_n G_{Tr_n}(0,x)} - 1) \right). \quad (2.40)$$

Now using Condition D and that  $b_n r_n = f_n$ , we arrive at that (2.40) is bounded above by

$$\exp \left( -2^{j\alpha_n} \lambda_n f_n D_n \right) \quad (2.41)$$

for all  $n$  large enough, for all  $j \in \mathbb{N}$ , where

$$D_n := \left\{ \left( \frac{a+d/2}{a} \right)^{1/\alpha_n} - \frac{CK_s T}{a} \right\}. \quad (2.42)$$

We have for all  $n$

$$\sum_{j=1}^{\infty} \exp \left( -2^{j\alpha_n} \lambda_n f_n D_n \right) \leq \exp \left( -2^{\alpha_n} \lambda_n f_n D_n \right) + \frac{\exp \left( -2^{\alpha_n} \lambda_n f_n D_n \right)}{\alpha_n (\log 2) 2^{\alpha_n} \lambda_n f_n D_n} \quad (2.43)$$

Since both  $\alpha_n D_n \rightarrow \infty$  and  $D_n \rightarrow \infty$  as  $n \rightarrow \infty$ ; by the part of Condition D that states that  $\sum_{n=1}^{\infty} \exp(-c\lambda_n f(n)) < \infty$  for some  $c > 0$ , we can conclude that the right hand side of (2.43) is summable in  $n$ . Hence, Borel-Cantelli lemma yields (2.37). Finally, summing over  $j$  finishes the proof of Lemma 2.2.  $\square$

The next proposition shows that the sum of scores is dominated by the largest score.

**Proposition 2.3.** For any  $\delta > 0$  given, for any  $\epsilon$  small and  $M$  large enough,  $\tau$  a.s. there exists a constant  $K \geq 1$  such that for  $n$  large enough

$$\mathbb{P}(S_n(d_n(j)) \leq Km_n(j), \quad \forall j \text{ s.t. } d_n(j) \leq Tr_n) \geq 1 - \delta. \quad (2.44)$$

*Proof.* Recall that  $\zeta_n(T)$  is the last time the random walk finds a deep trap before  $Tr_n$ . Note that for  $j = 0, 1$  we have by definition  $S_n(d_n(j)) = m_n(j)$ . Hence, we assume that  $\zeta_n(T) \geq 2$  and prove that

$$\mathbb{P}(S_n(d_n(j)) \leq Km_n(j), \quad \forall j = 2, \dots, \zeta_n(T) | \tau) \geq 1 - \delta. \quad (2.45)$$

We define the sequence of events that the random walk cannot find a very deep trap before time  $Tr_n$ :

$$I_n^1 := \{H_n(T_M(n)) \geq Tr_n\}. \quad (2.46)$$

We can use Condition C with  $\rho = 1/M$  to conclude that for  $M$  large enough

$$\mathbb{P}(I_n^1 | \tau) \geq 1 - \delta/4. \quad (2.47)$$

For  $d > 0$  define the sequence of events

$$I_n^2 := \{s_n(1) \geq (\epsilon + d)^{1/\alpha_n} t_n\}. \quad (2.48)$$

Using Proposition 2.1, for  $d$  small enough,  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(I_n^2 | \tau) \geq 1 - \delta/4. \quad (2.49)$$

We define another sequence of events

$$I_n^3 := \left\{ \sum_{i=0}^{Tr_n} e_i \tau_{Y_n(i)} \mathbb{1} \{Y_n(i) \in T^\epsilon(n)\} \leq (\epsilon + d/2)^{1/\alpha_n} t_n \right\}. \quad (2.50)$$

By Proposition 2.2 we have  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(I_n^3 | \tau) \geq 1 - \delta/4. \quad (2.51)$$

Finally, defining

$$I_n^4 := \{\zeta_n(T) \leq K - 1\}, \quad (2.52)$$

using Condition C, for  $K$  large enough,  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(I_n^4 | \tau) \geq 1 - \delta/4. \quad (2.53)$$

For  $d$  and  $M$  chosen as above, for  $j = 2, \dots, \zeta_n(T)$ , we partition the sum  $S_n(d_n(j))$  as follows

$$S_n(d_n(j)) = \sum_{i=0}^{d_n(j)-1} e_i \tau_{Y_n(i)} \mathbb{1} \{Y_n(i) \in T^\epsilon(n)\} + \sum_{i=0}^{d_n(j)-1} e_i \tau_{Y_n(i)} \mathbb{1} \{Y_n(i) \in T_M(n)\} + \sum_{k=1}^j s_n(k). \quad (2.54)$$

Recall that  $m_n(j) = \max_{i=1, \dots, j-1} s_n(i)$ . Let  $I_n := I_n^1 \cap I_n^2 \cap I_n^3 \cap I_n^4$ . We have  $\tau$  a.s. for  $n$  large enough  $\mathbb{P}(I_n) \geq 1 - \delta$ . On the event  $I_n$ , the first term on the right hand side of (2.54) is bounded above by  $m_n(j)$  for any  $j \geq 1$  using (2.49) and (2.50); the second term is 0 since on  $I_n^1$  no deep trap has been found; the third term is bounded by  $K - 1$  since on  $I_n^4$   $\zeta_n(T) \leq K - 1$ . Hence, on  $I_n$ ,  $S_n(d_n(j))$  is bounded  $Km_n(j)$  and we are finished with the proof.  $\square$

**Proposition 2.4.** For any  $\delta > 0$  given, for  $\epsilon$  small and  $M$  large enough,  $\tau$  a.s. for large enough  $n$

$$\mathbb{P} \left( \max_{j \leq \zeta_n(T)} \left| \left( \frac{S_n(d_n(j))}{t_n} \right)^{\alpha_n} - \left( \frac{m_n(j)}{t_n} \right)^{\alpha_n} \right| \geq \delta \mid \tau \right) \leq \delta. \quad (2.55)$$

*Proof.* By definition we have  $m_n(j) \leq S_n(d_n(j))$  for any  $j$ . Using Proposition 2.3 we can find a positive constant  $K$  s.t.  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(S_n(d_n(j)) \leq Km_n(j), \forall j = 1, \dots, \zeta_n(T) | \tau) \geq 1 - \delta/4. \quad (2.56)$$

For  $K$  as above, using Lemma 2.2 we can choose a small enough  $d$  so that  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(m_n(j) \in [(\epsilon - d)^{1/\alpha_n} t_n, (M + d)^{1/\alpha_n} t_n], \forall j = 1, \dots, K) \geq 1 - \delta/4. \quad (2.57)$$

Finally, for  $K$  and  $d$  as above, for  $n$  large enough

$$|K^{\alpha_n} - 1|(M + 2d) \leq \delta/2. \quad (2.58)$$

We denote by  $I_n$  the intersection of the events inside the probability displays in (2.56) and (2.57). Then using (2.56), (2.57) and (2.58) we get

$$\begin{aligned} & \mathbb{P}\left(\max_{j \leq \zeta_n(T)} \left| \left( \frac{S_n(d_n(j))}{t_n} \right)^{\alpha_n} - \left( \frac{m_n(j)}{t_n} \right)^{\alpha_n} \right| \geq \delta \mid \tau\right) \\ & \leq \delta/2 + \mathbb{P}\left(\max_{j \leq \zeta_n(T)} \left| \left( \frac{Km_n(j)}{t_n} \right)^{\alpha_n} - \left( \frac{m_n(j)}{t_n} \right)^{\alpha_n} \right| \geq \delta \mid I_n, \tau\right) = \delta/2. \end{aligned} \quad (2.59)$$

□

*Proof of Theorem 1.* We start the proof with the proof of convergence of the finite dimensional distributions. Let  $0 = t_0 \leq t_1 \leq t_2 \leq \dots \leq t_k \leq T$  and  $0 \leq x_1 \leq x_2 \leq \dots \leq x_k$  be given. Consider the random variables  $\zeta_n(t_i), i = 1, \dots, k$ . For convenience, we denote by  $N(\lambda)$ , a Poisson random variable with mean  $\lambda\rho_\epsilon^M$ . Let  $\{N(t_i - t_{i-1}) : i = 1, \dots, k\}$  be an independent collection. By Condition C, we have as  $\tau$  a.s.  $n \rightarrow \infty$

$$\{\zeta_n(t_i) - \zeta_n(t_{i-1}) : i = 1, \dots, k\} \implies \{N(t_i - t_{i-1}) : i = 1, \dots, k\} \quad (2.60)$$

where  $\implies$  stands for convergence in distribution. For ease of notation we defined the rescaled clock process

$$\bar{S}_n(t) := \left( \frac{S_n(tr_n)}{t_n} \right)^{\alpha_n}. \quad (2.61)$$

By Proposition 2.4 and using Proposition 2.2 as in the proof of Proposition 2.4, we can conclude that  $\tau$  a.s. the sequence of events that for all  $i = 0, 1, \dots, k$

$$\left( \frac{m_n(\zeta_n(t_i))}{t_n} \right)^{\alpha_n} - \delta \leq \bar{S}_n(t_i) \leq \left( \frac{m_n(\zeta_n(t_i))}{t_n} \right)^{\alpha_n} + \delta \quad (2.62)$$

has probability larger than  $1 - \delta$  for all  $n$  large enough. Hence, the sequence of finite dimensional distributions

$$\mathbb{P}\left(\bar{S}_n(t_1) \leq x_1, \dots, \bar{S}_n(t_k) \leq x_k \mid \tau\right) \quad (2.63)$$

is bounded above by

$$\mathbb{P}\left(\left(\frac{m_n(\zeta_n(t_1))}{t_n}\right)^{\alpha_n} \leq x_1 + \delta, \dots, \left(\frac{m_n(\zeta_n(t_k))}{t_n}\right)^{\alpha_n} \leq x_k + \delta \mid \tau\right) \quad (2.64)$$

and below by

$$\mathbb{P}\left(\left(\frac{m_n(\zeta_n(t_1))}{t_n}\right)^{\alpha_n} \leq x_1 - \delta, \dots, \left(\frac{m_n(\zeta_n(t_k))}{t_n}\right)^{\alpha_n} \leq x_k - \delta \mid \tau\right). \quad (2.65)$$

We prove only the upper bound for the lower bound can be achieved similarly. By Lemma 2.2 and (2.60) (also recall (2.28)),  $\tau$  a.s. as  $n \rightarrow \infty$ , the sequence of probability terms in (2.64) converges to

$$\mathbb{E}\left[\left(\rho_\epsilon^{x_1 + \delta} / \rho_\epsilon^M\right)^{N(t_1)}\right] \mathbb{E}\left[\left(\rho_\epsilon^{x_2 + \delta} / \rho_\epsilon^M\right)^{N(t_2 - t_1)}\right] \dots \mathbb{E}\left[\left(\rho_\epsilon^{x_k + \delta} / \rho_\epsilon^M\right)^{N(t_k - t_{k-1})}\right]. \quad (2.66)$$

A simple calculation yields that for any  $x, \lambda \geq 0$

$$\mathbb{E} \left[ \left( \frac{\rho_\epsilon^x}{\rho_\epsilon^M} \right)^{N(\lambda)} \right] = \exp(-\lambda \rho_x^M). \quad (2.67)$$

Hence, (2.66) is equal to

$$\exp \left( -\frac{t_1}{x_1 + \delta} \right) \exp \left( -\frac{t_2 - t_1}{x_2 + \delta} \right) \cdots \exp \left( -\frac{t_k - t_{k-1}}{x_k + \delta} \right). \quad (2.68)$$

Finally, letting  $\delta \rightarrow 0$  finishes the proof of the convergence of finite dimensional distributions.

For tightness characterizations we need the following definitions:

$$\begin{aligned} w_f(\delta) &= \sup \left\{ \min \left( |f(t) - f(t_1)|, |f(t_2) - f(t)| \right) : t_1 \leq t \leq t_2 \leq T, t_2 - t_1 \leq \delta \right\}, \\ w'_f(\delta) &= \sup \left\{ \inf_{\alpha \in [0,1]} |f(t) - (\alpha f(t_1) + (1 - \alpha)f(t_2))| : t_1 \leq t \leq t_2 \leq T, t_2 - t_1 \leq \delta \right\}, \\ v_f(t, \delta) &= \sup \{ |f(t_1) - f(t_2)| : t_1, t_2 \in [0, T] \cap (t - \delta, t + \delta) \}. \end{aligned}$$

Following is from Theorem 12.12.3 of [16] and Theorem 15.3 of [7].

**Theorem 2.** *The sequence of probability measures  $\{P_n\}$  on  $D([0, T])$  is tight in the  $M_1$ -topology if*

(i) *For each positive  $\varepsilon$  there exists a  $c$  such that*

$$P_n[h : \|h\|_\infty \geq c] \leq \varepsilon, \quad n \geq 1 \quad (2.69)$$

(ii) *For each  $\varepsilon > 0$  and  $\eta > 0$ , there exists a  $\delta, 0 < \delta < T$ , and an integer  $n_0$  such that*

$$P_n[h : w'_h(\delta) \geq \eta] \leq \varepsilon, \quad n \geq n_0 \quad (2.70)$$

and

$$P_n[h : v_h(0, \delta) \geq \eta] \leq \varepsilon \text{ and } P_n[h : v_h(T, \delta) \geq \eta] \leq \varepsilon, \quad n \geq n_0 \quad (2.71)$$

Moreover, the same is true for  $J_1$  topology with  $w'_h(\delta)$  in (2.70) replaced by  $w_h(\delta)$ .

For the first claim in Theorem 1, we check tightness in  $M_1$  topology, using Theorem 2. Since  $\bar{S}_n(t)$  is non-decreasing in  $t$ , to check condition (i) it is enough to check that  $\bar{S}_n(T)$  is tight. In this case, the convergence of fixed time distribution gives the desired result.

Since for monotone functions  $w'_h(\delta)$  is 0 in part (ii) of Theorem 2 we only need to control  $v_{\bar{S}_n}(0, \delta)$  and  $v_{\bar{S}_n}(T, \delta)$ . Again, using the monotonicity, controlling  $v_{\bar{S}_n}(0, \delta)$  boils down to check that  $\mathbb{P}[\bar{S}_n(\delta) \geq \eta] \leq \varepsilon$  for small enough  $\delta$  and large enough  $n$ . By convergence of the fixed time distribution it is enough to check  $\mathbb{P}(W(\delta) \geq \eta/2) \leq \varepsilon/2$ . Since  $\mathbb{P}(W(\delta) \geq \eta/2) = 1 - e^{-2\delta/\eta}$ , this claim is true for small enough  $\delta$ . Similarly, to control  $v_{\bar{S}_n}(T, \delta)$  it is enough to find  $\eta$  small enough so that

$$\mathbb{P}[W(T) - W(T - \delta) \geq \eta] \leq \varepsilon/2. \quad (2.72)$$

Observe that by (2.12)

$$\mathbb{P}[W(T) - W(T - \delta) = 0] = \frac{T - \delta}{T}, \quad (2.73)$$

then

$$\mathbb{P}[W(T) - W(T - \delta) \geq \eta] \leq 1 - \mathbb{P}[W(T) - W(T - \delta) = 0] = \frac{\delta}{T}.$$

Hence, (2.72) follows by taking  $\delta \leq T\varepsilon/4$ . Hence, we are finished with the proof of the first part of Theorem 1.

Now we assume that  $f_n = 1$  and  $\mathcal{K}_G = 1$  and prove the tightness of  $\bar{S}_n$  in  $J_1$  topology. We only need to check that (2.70) with  $w_{\bar{S}_n}$ . It is enough to show that  $\tau$  a.s. as  $n \rightarrow \infty$

$$\max_{x \in T_\epsilon^M(n)} \mathbb{P}_x(Y_n(i) \notin T_\epsilon^M(n), i = 1, \dots, \delta r_n) \leq \varepsilon \quad (2.74)$$

for  $\delta$  small enough. Since Condition B is satisfied with  $f_n = \mathcal{K}_G = 1$  we have

$$\max_{x \in T_\epsilon^M(n)} \mathbb{P}_x(H'_n(x) < H_n(T_\epsilon^M(n) \setminus \{x\})) \longrightarrow 0 \quad (2.75)$$

where

$$H'_n(x) := \min\{i \geq 1 : Y_n(i) = x\} \quad (2.76)$$

Combining this with Condition C yields (2.74). Hence, we have proved the  $J_1$  convergence.  $\square$

### 3. EXTREMAL AGING

Next, we give two extra conditions that ensure extremal aging occurs. For this purpose we need to control the random walk between record sites. First, we define the sequence of times when a new record site is found. We define  $q_n(1) = 1$  and for  $j \geq 2$

$$q_n(j) = \min\{i > q_n(j-1) : m_n(i) > m_n(q_n(j-1))\}. \quad (3.1)$$

Hence,  $q_n(j)$ th score is greater than all the scores before it. To keep track of the times when the random walk visits a record site we introduce

$$k_n(j) = d_n(q_n(j)), \quad j \in \mathbb{N}. \quad (3.2)$$

Finally, we define the process  $V_n(j)$  that records the trajectory of  $Y_n$  restricted to deep traps whose score is a record:

$$V_n(j) = U_n(q_n(j)). \quad (3.3)$$

We fix numbers  $a$  and  $b$  with  $0 < a < b$ . We set  $\xi_n = Tr_n$ , the number of steps of the random walk that we observe. Later we will choose  $T$  large enough so that the clock process reaches the level  $a^{1/\alpha_n} t_n$ . The first condition is that in the time between the record site  $V_n(j)$  is found and the next record site  $V_n(j+1)$  is found, with a high probability the trap model is at  $V_n(j)$ . More precisely, let  $t_n$  be a deterministic sequence of times satisfying  $(a/2)^{1/\alpha_n} t_n \leq t'_n \leq b^{1/\alpha_n} t_n$  and let  $\delta > 0$ . We define  $j_n \in \mathbb{N}$  by

$$\left( \frac{S_n(k_n(j_n))}{t_n} \right) + \delta \leq \left( \frac{t'_n}{t_n} \right)^{\alpha_n} \leq \left( \frac{S_n(k_n(j_n+1))}{t_n} \right) - \delta \quad (3.4)$$

and  $j_n = \infty$  if the above inequalities are not satisfied by an integer. Let  $A_n(\delta)$  be the event

$$A_n(\delta) := \{0 < j_n < \zeta_n(T)\}. \quad (3.5)$$

**Condition 1:** For any  $\delta > 0$  it is possible to choose  $\epsilon$  small and  $M$  large enough so that  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(X_n(t'_n) = V_n(j_n) | A_n(\delta), \tau) \geq 1 - \delta. \quad (3.6)$$

The second condition states that there are no repetitions among record sites.

**Condition 2:** For any  $\epsilon$  and  $M$ ,  $\tau$  a.s.

$$\lim_{n \rightarrow \infty} \mathbb{P}(\exists i, j \text{ s.t. } i \neq j, q_n(i), q_n(j) \leq \zeta_n(T), V_n(i) = V_n(j) | \tau) = 0. \quad (3.7)$$

As discussed before, our choice of the two-time correlation function  $R_n$  is

$$R_n(t_1, t_2) := \mathbb{P}(X_n(t_1) = X_n(t_2) | \tau_n) \quad (3.8)$$

Now we are ready to state our extremal aging result.

**Theorem 3.** Assume that Conditions A-D and 1-2 hold and let  $0 < a < b$ . Then  $\tau$  a.s.

$$\lim_{n \rightarrow \infty} R_n(a^{1/\alpha_n} t_n, b^{1/\alpha_n} t_n | \tau) = \frac{a}{b}. \quad (3.9)$$

*Proof of Theorem 3.* We first calculate the probability that the record process jumps over an interval. We define the range of  $m_n$  by

$$\mathcal{R}(m_n) := \left\{ \left( \frac{m_n(j)}{t_n} \right)^{\alpha_n} : j \in \mathbb{N} \right\}. \quad (3.10)$$

Note that,  $m_n$  depends on  $\epsilon$  and  $M$  for the choice of  $T_\epsilon^M(n)$ .

**Lemma 3.1.** *For  $\epsilon < a < b < M$ ,  $\tau$  a.s.*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \mathcal{R}(m_n) \cap [a, b] = \emptyset \right) = \frac{\rho_b^M}{\rho_a^M}. \quad (3.11)$$

*Proof.* By definition we have  $m_n(1) = S_n(d_n(1))$ . Since  $\epsilon < a$ , by Lemma 2.2, as  $n \rightarrow \infty$  the probability that  $m_n(1) \geq a^{1/\alpha_n} t_n$  vanishes. Hence, using the weak convergence result in Lemma 2.2 and the fact that the distribution of  $\sigma_\epsilon^M$  has no atoms

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \mathcal{R}(m_n) \cap [a, b] = \emptyset \right) = \mathbb{P} \left( \left\{ \max_{i=1, \dots, j} \sigma_\epsilon^M(i) : j \in \mathbb{N} \right\} \cap [a, b] = \emptyset \right) \quad (3.12)$$

Since  $(\sigma_\epsilon^M(i) : i \in \mathbb{N})$  is an i.i.d. sequence we have (3.12) equal to

$$\begin{aligned} & \mathbb{P} \left( \sigma_\epsilon^M(1) \geq b \right) + \sum_{j=0}^{\infty} \mathbb{P} \left( \max_{i=1, \dots, j} \sigma_\epsilon^M(i) \leq a \right) \mathbb{P} \left( \sigma_\epsilon^M(j+1) \geq b \right) \\ &= \sum_{j=0}^{\infty} \mathbb{P} \left( \sigma_\epsilon^M(1) \leq a \right)^j \mathbb{P} \left( \sigma_\epsilon^M(1) \geq b \right) = \frac{\mathbb{P} \left( \sigma_\epsilon^M(1) \geq b \right)}{\mathbb{P} \left( \sigma_\epsilon^M(1) \geq a \right)} = \frac{\rho_b^M}{\rho_a^M}. \end{aligned} \quad (3.13)$$

□

Let  $\delta > 0$  be small enough so that  $\epsilon < a - 2\delta < b + \delta < M$ . Define the sequence of events  $I_n^1, I_n^2$  and  $I_n^3$  as follows:

$$\begin{aligned} I_n^1 &:= \left\{ \text{dist}(a, \mathcal{R}(m_n)) \leq \delta \text{ or } \text{dist}(b, \mathcal{R}(m_n)) \leq \delta \right\}, \\ I_n^2 &:= \left\{ \text{dist}(a, \mathcal{R}(m_n)) \geq \delta, \text{dist}(b, \mathcal{R}(m_n)) \geq \delta \text{ and } [a + \delta, b - \delta] \cap \mathcal{R}(m_n) \neq \emptyset \right\}, \\ I_n^3 &:= \left\{ [a - \delta, b + \delta] \cap \mathcal{R}(m_n) = \emptyset \right\} \end{aligned} \quad (3.14)$$

We also define  $F_n := \{m_n(\zeta_n(T)) \geq b^{1/\alpha_n} t_n\}$ . Since a.s.  $W(t) \rightarrow \infty$  as  $t \rightarrow \infty$ , using Theorem 1 we can choose  $T > 0$  large enough so that  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(F_n | \tau) \geq 1 - \delta/4. \quad (3.15)$$

Finally, let us define the sequence of events we want to approximate

$$G_n := \left\{ X_n(a^{1/\alpha_n} t_n) = X_n(b^{1/\alpha_n} t_n) | \tau \right\}. \quad (3.16)$$

We want to show that  $G_n$  can be well approximated by  $I_n^3$ . Since the jumps of  $m_n$  have continuous distribution we have  $(I_n^3)^c = I_n^1 \cup I_n^2$  and  $(I_n^2)^c \subset I_n^1 \cup I_n^3$ . This yields to

$$\mathbb{P}(G_n \cap I_n^3) \leq \mathbb{P}(G_n) \leq \mathbb{P}(I_n^1) + \mathbb{P}(I_n^3) + \mathbb{P}(G_n \cap I_n^2). \quad (3.17)$$

On  $F_n$  and  $I_n^2$  there exist  $j_1$  and  $j_2$  with  $0 < j_1 < j_2$  and  $q_n(j_1), q_n(j_2) < \zeta_n(T)$  such that

$$\begin{aligned} \left( \frac{m_n(q_n(j_1))}{t_n} \right)^{\alpha_n} + \delta &\leq a \leq \left( \frac{m_n(q_n(j_1 + 1))}{t_n} \right)^{\alpha_n} - \delta \\ \left( \frac{m_n(q_n(j_2))}{t_n} \right)^{\alpha_n} + \delta &\leq a \leq \left( \frac{m_n(q_n(j_2 + 1))}{t_n} \right)^{\alpha_n} - \delta. \end{aligned} \quad (3.18)$$

Using Proposition 2.4 this yields to

$$\begin{aligned} \left(\frac{S_n(q_n(j_1))}{t_n}\right)^{\alpha_n} + \delta/2 &\leq a \leq \left(\frac{S_n(q_n(j_1+1))}{t_n}\right)^{\alpha_n} - \delta/2 \\ \left(\frac{S_n(q_n(j_2))}{t_n}\right)^{\alpha_n} + \delta/2 &\leq a \leq \left(\frac{S_n(q_n(j_2+1))}{t_n}\right)^{\alpha_n} - \delta/2. \end{aligned} \quad (3.19)$$

Hence, by Condition 1 we have  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(X_n(a^{1/\alpha_n}) = V_n(j_1), X_n(b^{1/\alpha_n}) = V_n(j_2) | I_n^2, F_n, \tau) \geq 1 - \delta. \quad (3.20)$$

Combining this with Condition 2 and (3.15) we get

$$\mathbb{P}(G_n \cap I_n^2) \leq \delta. \quad (3.21)$$

Similarly on  $F_n \cap I_n^3$  there exists a  $j_1$  such that

$$\left(\frac{m_n(q_n(j_1))}{t_n}\right)^{\alpha_n} + \delta \leq a \leq b \leq \left(\frac{m_n(q_n(j_1+1))}{t_n}\right)^{\alpha_n} - \delta. \quad (3.22)$$

Hence, by Condition 1 we have  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(X_n(a^{1/\alpha_n} t_n) = X_n(b^{1/\alpha_n} t_n) = V_n(j_1) | \tau) \geq 1 - \delta, \quad (3.23)$$

and since  $\mathbb{P}(G_n) \geq \mathbb{P}(X_n(a^{1/\alpha_n} t_n) = X_n(b^{1/\alpha_n} t_n) = V_n(j_1) | \tau)$  we get

$$\mathbb{P}(I_n^3) - \delta \leq \mathbb{P}(I_n^3 \cap G_n). \quad (3.24)$$

By Lemma 3.1 we have  $\mathbb{P}(I_n^1) \rightarrow 0$  and  $\mathbb{P}(I_n^3) \rightarrow \rho_b^M / \rho_a^M$  as  $\delta \rightarrow 0$ . Finally, taking  $M \rightarrow \infty$  and  $\delta \rightarrow 0$  finishes the proof.  $\square$

#### 4. EXTREMAL AGING FOR RANDOM ENERGY MODEL

The Random Energy Model was first introduced by Bernard Derrida in [13] as an exactly solvable mean field spin glass model. The state space is  $\mathcal{V}_n = \{-1, +1\}^n$ , the  $n$  dimensional hypercube. To each configuration  $\sigma \in \mathcal{V}_n$  is attached a random Hamiltonian  $H_n(\sigma)$ . The choice for the energy landscape in REM is that of i.i.d. Gaussians with mean zero and variance  $n$ . More precisely,  $H_n(\sigma) = -\sqrt{n}Z_\sigma$  where

$$\left\{ Z_\sigma : \sigma \in \mathcal{V}_n \right\} \quad (4.1)$$

is an i.i.d. sequence of standard Gaussian random variables. Let  $\beta > 0$  be the inverse temperature (later we will let  $\beta$  vary with the dimension  $n$ ), then the Gibbs measure at inverse temperature  $\beta$  is given by

$$\mu_n(\sigma) = \frac{1}{\mathcal{Z}_{n,\beta}} e^{-\beta H_n(\sigma)} = \frac{1}{\mathcal{Z}_{n,\beta}} e^{\beta\sqrt{n}Z_\sigma}, \quad \sigma \in \mathcal{V}_n \quad (4.2)$$

where  $\mathcal{Z}_{n,\beta}$  is the usual partition function.

We will study the trap model dynamics of REM. In our setting, the graph is  $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n)$  where the set of edges  $\mathcal{E}_n$  is given by

$$\mathcal{E}_n = \left\{ (\sigma, \sigma') : \frac{1}{2} \sum_{i=1}^n |\sigma_i - \sigma'_i| = 1 \right\}. \quad (4.3)$$

In other words, two configurations in  $\mathcal{V}_n$  are neighbors if they differ only at one spin. For the trapping landscape  $\{\tau_\sigma : \sigma \in \mathcal{V}_n\}$  we choose the Gibbs waits

$$\tau_\sigma := \exp(-\beta H_n(\sigma)) = \exp(\beta\sqrt{n}Z_\sigma), \quad \sigma \in \mathcal{V}_n. \quad (4.4)$$



As before we denote the corresponding trap model by  $(X_n(t) : t \geq 0)$ . It is trivial that the Gibbs measure of REM is the unique invariant measure of  $X_n$ . We want to note there that in the literature this type of dynamics is sometimes called Random Hopping Time dynamics.

We will study the dynamics of REM model in the following context. We let the temperature vary by the volume of the system, hence, here after we parameterize the inverse temperature by  $n$  and replace  $\beta$  with  $\beta_n$  in the above equations. Let us rewrite the trapping landscape in this new setup

$$\tau := \left\{ \exp(\beta_n \sqrt{n} Z_\sigma) : \sigma \in \mathcal{V}_n \right\} \quad (4.5)$$

where the sequence  $\{Z_\sigma : \sigma \in \mathcal{V}_n\}$  is as in (4.1). As before, we denote by  $Y_n$  the simple random walk on the hypercube  $\mathcal{V}_n$  and the clock process by  $S_n$ .

Next we choose the time scales that we will observe the dynamics at. We set first our depth rate scale  $\alpha_n$ ; as in Section 2,  $\alpha_n \rightarrow 0$  as  $n \rightarrow \infty$ . We consider the following scales:

$$g_n = t_n = \exp(\alpha_n \beta_n^2 n), \quad b_n^{-1} = r_n = \alpha_n \beta_n \sqrt{2\pi n} \exp(\alpha_n^2 \beta_n^2 n / 2), \quad f_n = 1. \quad (4.6)$$

We will always assume that

$$\limsup_{n \rightarrow \infty} \alpha_n \beta_n < \sqrt{2 \log 2}. \quad (4.7)$$

Furthermore, we will assume that  $\alpha_n$  and  $\beta_n$  are chosen such that

$$n \log n \ll r_n \quad \text{as } n \rightarrow \infty. \quad (4.8)$$

The following theorem describes the extremal aging for the dynamics of REM.

**Theorem 4.** *For a.s. random environment  $\tau$*

(i) *for any  $T > 0$ , as  $n \rightarrow \infty$*

$$\left( \frac{S_n(\cdot \cdot r_n)}{t_n} \right)^{\alpha_n} \implies W(\cdot) \quad \text{in } D([0, T], J_1) \quad (4.9)$$

*where  $W$  is the extremal process generated by the distribution function  $F(x) = e^{-1/x}$ ,*

(ii) *for any  $0 < a < b$  as  $n \rightarrow \infty$*

$$R_n \left( a^{1/\alpha_n} t_n, b^{1/\alpha_n} t_n \mid \tau \right) \longrightarrow \frac{a}{b}. \quad (4.10)$$

**Remark 4.1.** *Note that the ratio of the two times  $t_n a^{1/\alpha_n}$  and  $t_n b^{1/\alpha_n}$  diverges with  $n$  but the logarithmic ratio  $\log(t_n a^{1/\alpha_n}) / \log(t_n b^{1/\alpha_n})$  converges to 1 as  $n \rightarrow \infty$ , due to (4.8). Hence, we can think of the extremal aging result of Theorem 4 as “just before aging”.*

**Remark 4.2.** *Let us describe the cases that Theorem 4 covers. For constant temperature case  $\beta_n \equiv \beta > 0$ , the time scales in Theorem 4 consist of subexponential time scales (in  $n$ ). For the case that  $1 \ll \beta_n$ , that is, zero temperature dynamics, depending on  $\beta_n$  and  $\alpha_n$ , it covers superexponential, exponential and subexponential time scales. Finally, for  $\beta_n \ll 1$ , time scales covered are subexponential.*

*Proof of Theorem 4.* In order to prove Theorem 4 we will check Conditions A-D and Conditions 1-2; and to achieve convergence in  $J_1$  we will check Condition B with  $\mathcal{K}_G = 1$ .

**Condition A:** It is well-known for a standard Gaussian random variable  $Z$  that

$$P(Z \geq u) = \frac{1}{u\sqrt{2\pi}} e^{-u^2/2} (1 + o(1)) \quad \text{as } u \rightarrow \infty, \quad (4.11)$$

and

$$P(Z \geq u) \leq \frac{1}{u\sqrt{2\pi}} e^{-u^2/2}, \quad \forall u > 0. \quad (4.12)$$

By (4.8) we have  $1 \ll \alpha_n \beta_n \sqrt{n}$ . Hence, using (4.5) and (4.11)

$$\begin{aligned} & (\alpha_n \beta_n \sqrt{2\pi n}) e^{\alpha_n^2 \beta_n^2 n/2} P(\tau_x \geq u^{1/\alpha_n} e^{\alpha_n \beta_n^2 n}) \\ &= (\alpha_n \beta_n \sqrt{2\pi n}) e^{\alpha_n^2 \beta_n^2 n/2} P\left(Z \geq \alpha_n \beta_n \sqrt{n} + \frac{\log u}{\alpha_n \beta_n \sqrt{n}}\right) \xrightarrow{n \rightarrow \infty} \frac{1}{u}. \end{aligned} \quad (4.13)$$

This proves the first part of Condition A. The second part of Condition A follows trivially from (4.12) and a calculation similar to the above.

**Condition C:** For proving Condition C we use the following theorem from [10]:

**Theorem 5.** (Theorem 1 in [10]) Let  $\bar{m}(n)$  be such that

$$n \log n \ll \bar{m}(n) \ll 2^n (\log n)^{-1}, \quad (4.14)$$

and let  $A_n$  be a sequence of percolation clouds on  $\mathcal{V}_n$  with densities  $\bar{m}(n)^{-1}$ . Then, for all  $a > 0$ ,

$$\lim_{n \rightarrow \infty} \max_{x \in \mathcal{V}_n} \left| P_x \left( H_n(A_n \setminus \{x\}) \geq a \bar{m}(n) \right) - \exp(-a) \right| = 0. \quad (4.15)$$

In our case, Condition C is equivalent to the result of the above theorem when  $\bar{m}(n) = (\rho b_n)^{-1}$ ,  $\rho > 0$ . We have in (4.6)  $b_n^{-1} = \alpha_n \beta_n \sqrt{2\pi n} \exp(\alpha_n^2 \beta_n^2 n/2)$  and, thus, (4.7) and (4.8) yield  $n \log n \ll (\rho b_n)^{-1} \ll 2^n (\log n)^{-1}$ . Hence, we can apply the theorem and verify Condition C.

**Condition B:** Recall the notation  $\xi_n = Tr_n$ . It is trivial that for any  $k_1, k_2$  with  $k_1 \leq k_2$ , and for all  $x, y \in \mathcal{V}_n$ ,  $G_{k_1}^n(x, y) \leq G_{k_2}^n(x, y)$ . Also, for any  $k \in \mathbb{N}$  and  $x \in \mathcal{V}_n$ ,  $G_k^n(x, x) \geq 1$  by definition. Hence, using part (a) of Lemma 3.9 in [5] we conclude that

$$\limsup_{n \rightarrow \infty} G_{\xi_n \log \xi_n}^n(0, 0) = 1. \quad (4.16)$$

Now we claim that uniformly for  $x \in T_\epsilon^M(n)$  that

$$\lim_{n \rightarrow \infty} G_{T_\epsilon^M(n) \setminus \{x\}}^n(x, x) = 1, \quad (4.17)$$

which is enough to check Condition B with  $f_n = 1$  and  $\mathcal{K}_G = 1$ . Recalling that  $H'_n(x) = \min\{i \geq 1 : Y_n(i) = x\}$  we have

$$G_{\xi_n \log \xi_n}^n(0, 0) = \left( P_x(H_n(T_\epsilon^M(n) \setminus \{x\}) < H'_n(x)) \right)^{-1}. \quad (4.18)$$

We have

$$P_x(H_n(T_\epsilon^M(n) \setminus \{x\}) \geq H'_n(x)) \leq P_x(H'_n(x) \leq \xi_n \log \xi_n) + P_x(H_n(T_\epsilon^M(n) \setminus \{x\}) \geq \xi_n \log \xi_n). \quad (4.19)$$

The first term on the right hand side of the above display is obviously independent of  $x$  and converges to 0 by (4.16); using Condition C, that is, the fact that uniformly for  $x \in T_\epsilon^M(n)$ ,  $H_n(T_\epsilon^M(n) \setminus \{x\})/\xi_n$  is asymptotically an exponential random variables, we can conclude that the second term also vanishes uniformly for  $x \in \mathcal{V}_n$ . Hence, we have proved (4.17).

**Condition D:** By Lemma 3.9 in [5] we have for all  $x \neq 0$ , for all  $n$  large enough

$$G_{\xi_n}^n(0, x) \leq C/n \quad (4.20)$$

for some positive constant  $C > 0$ . By (4.8) we can choose a sequence  $\lambda_n$  such that

$$\lambda_n \ll n, \quad \lambda_n \ll \alpha_n^2 \beta_n^2 n \quad (4.21)$$

and  $\sum_n \exp(-\lambda_n) < \infty$  (recall that  $f_n = 1$ ). For such  $\lambda_n$ , using (4.20) we have  $\lambda_n G_{\xi_n(0,x)} \ll 1$  uniformly for all  $x \in \mathcal{V}_n \setminus \{x\}$ . Combining this with (4.16) we get that for some positive constant  $K$ , for all  $n$  large enough

$$\sum_{x \in \mathcal{V}_n} (e^{\lambda_n G_{\xi_n}^n(0,x)} - 1) \leq e^{2\lambda_n} + \sum_{x \in \mathcal{V}_n \setminus \{x\}} (e^{\lambda_n G_{\xi_n}^n(0,x)} - 1) \leq e^{2\lambda_n} + \lambda_n T r_n \leq K T \lambda_n T r_n \quad (4.22)$$

where in the last step we used that  $\lambda_n \ll \alpha_n^2 \beta_n^2 n$ . Hence, we have checked the first part of Condition D. The second part of Condition D is trivial since  $f_n = 1$ .

**Condition 1:** Let  $t'_n$  be a deterministic sequence of times satisfying (3.4) and  $A_n(\delta)$  be defined as (3.5). We check that Condition 1 is satisfied uniformly for  $0 < a < b$ , conditioned on  $m_n(q_n(j_n)) = a^{1/\alpha_n} t_n$  and  $m_n(q_n(j_n + 1)) = b^{1/\alpha_n} t_n$ . For  $\epsilon > 0$  small enough so that  $a + 2\epsilon < b - 2\epsilon$  we define the sequence of events:

$$B_n := \{ \tau_{Y_n(i)} \leq (b - 2\epsilon)^{1/\alpha_n} t_n : \forall i = k_n(j_n), \dots, k_n(j_n + 1) - 1 \}. \quad (4.23)$$

By Condition C and Proposition 2.1 we can choose  $\epsilon$  small enough so that  $\tau$  a.s. for  $n$  large enough we have

$$\mathbb{P}(B_n | \tau) \geq 1 - \delta/4. \quad (4.24)$$

By Proposition 2.4,  $\tau$  a.s. for  $n$  large enough

$$\mathbb{P}(S_n(k_n(j_n + 1)) - S_n(k_n(j_n)) \geq (b - \epsilon/2)^{1/\alpha_n} t_n - (a + \epsilon)^{1/\alpha_n} t_n | \tau) \geq 1 - \delta/4. \quad (4.25)$$

Note that  $(b - \epsilon/2)^{1/\alpha_n} t_n - (a + \epsilon)^{1/\alpha_n} t_n \gg (b - 2\epsilon)^{1/\alpha_n} t_n$  as  $n \rightarrow \infty$ . Also, by Proposition 2.2, on  $I_n$ , the contribution from traps between  $k_n(j_n)$  and  $k_n(j_n + 1)$  is smaller than  $(b - \epsilon)^{1/\alpha_n} t_n$  with a probability larger than  $1 - \delta/4$ . Hence, conditioned on  $I_n$ ,  $A_n(\delta)$  and the sequence of events inside the probability term in the last display, the only way  $X_n(t'_n) \neq V_n(j_n)$  is if the random walk comes back to  $V_n(j_n)$  before visiting  $V_n(j_n + 1)$ . However, by (4.16) this probability goes to 0. Thus, Condition 1 is satisfied.

**Condition 2:** By Condition C, the time for  $Y_n$  to visit a deep trap divided by  $r_n$  is approximately exponentially distributed. Hence, by (4.16), we have that the probability of  $Y_n$  revisits a deep trap converges to 0. Thus, Condition 2 follows. □

## REFERENCES

- [1] Gérard Ben Arous, Véronique Gaynard, and Alexey Kuptsov. A new REM conjecture. In *In and out of equilibrium. 2*, volume 60 of *Progr. Probab.*, pages 59–96. Birkhäuser, Basel, 2008.
- [2] Gérard Ben Arous, Anton Bovier, and Jiří Černý. Universality of the REM for dynamics of mean-field spin glasses. *Comm. Math. Phys.*, 282(3):663–695, 2008.
- [3] Gérard Ben Arous, Anton Bovier, and Véronique Gaynard. Glauber dynamics of the random energy model. I. Metastable motion on the extreme states. *Comm. Math. Phys.*, 235(3):379–425, 2003.
- [4] Gérard Ben Arous, Anton Bovier, and Véronique Gaynard. Glauber dynamics of the random energy model. II. Aging below the critical temperature. *Comm. Math. Phys.*, 236(1):1–54, 2003.
- [5] Gérard Ben Arous and Jiří Černý. The arcsine law as a universal aging scheme for trap models. *Comm. Pure Appl. Math.*, 61(3):289–329, 2008.
- [6] Gérard Ben Arous and Onur Gün. Universality and extremal aging for dynamics of spin glasses on subexponential time scales. *Comm. Pure Appl. Math.*, 65(1):77–127, 2012.
- [7] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.
- [8] Anton Bovier and Véronique Gaynard. Convergence of clock processes in random environments and ageing in the  $p$ -spin SK model. *Ann. Probab.*, 41(2):817–847, 2013.
- [9] Anton Bovier, Véronique Gaynard, and Adéla Švejda. Convergence to extremal processes in random environments and extremal ageing in SK models. *Probab. Theory Related Fields*, 157(1-2):251–283, 2013.

- [10] Jiří Černý and Véronique Gayrard. Hitting time of large subsets of the hypercube. *Random Structures Algorithms*, 33(2):252–267, 2008.
- [11] D. A. Croydon, A. Fribergh, and T. Kumagai. Biased random walk on critical Galton–Watson trees conditioned to survive. *Probab. Theory Related Fields*, 157(1-2):453–507, 2013.
- [12] D. A. Darling. The influence of the maximum term in the addition of independent random variables. *Trans. Amer. Math. Soc.*, 73:95–107, 1952.
- [13] Bernard Derrida. Random-energy model: an exactly solvable model of disordered systems. *Phys. Rev. B (3)*, 24(5):2613–2626, 1981.
- [14] Yuji Kasahara. A limit theorem for sums of i.i.d. random variables with slowly varying tail probability. *J. Math. Kyoto Univ.*, 26(3):437–443, 1986.
- [15] Sidney I. Resnick. *Extreme values, regular variation and point processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original.
- [16] Ward Whitt. *Stochastic-process limits*. Springer Series in Operations Research. Springer-Verlag, New York, 2002. An introduction to stochastic-process limits and their application to queues.