Dynamics of trap models

Gérard Ben Arous   Jiří Černý

March 6, 2006

Contents

1 Introduction

These notes cover one of the topics of the class given in the Les Houches Summer School "Mathematical statistical physics" in July 2005. The lectures tried to give a summary of the recent mathematical results about the long-time behaviour of dynamics of (mean-field) spin-glasses and other disordered media (expanding on the review [Ben02]). We have chosen here to restrict the scope of these notes to the dynamics of trap models only, but to cover this topic in somewhat more depth.

Let us begin by setting the stage of this long review about the trap models by going back to one of the motivations behind their introduction by Bouchaud [Bou92], i.e. dynamics of spin-glasses, which is indeed a very good field to get an accurate sample of possible and generic long-time phenomena (as aging, memory, rejuvenation, failure of the fluctuation-dissipation theorem; see [BCKM98] for a global review).

This class of problems can be roughly described as follows. Let Γ (a compact metric space) be the state space for spins and ν be a probability measure on Γ. Typically, in the discrete (or Ising) spins context $\Gamma = \{-1, 1\}$ and $\nu = (\delta_1 + \delta_{-1})/2$. In the continuous or soft spin context $\Gamma = I$, a compact interval of the real line, and $\nu(dx) = Z^{-1}e^{-U(x)}dx$, where $U(x)$ is the "one-body potential". For each configuration of the spin system, i.e. for each $\boldsymbol{x} = (x_1, ..., x_n) \in \Gamma^n$ one defines a random Hamiltonian, $H_{\boldsymbol{J}}^n(\boldsymbol{x})$, as a function of the configuration $\boldsymbol{x}$ and of an exterior source of randomness $\boldsymbol{J}$, i.e. a random variable defined on another probability space. The Gibbs measure at inverse temperature β is then defined on the configuration space Γ^n by

$$\mu_{\beta,\boldsymbol{J}}^n(\mathrm{d}\boldsymbol{x}) = \frac{1}{Z_{\boldsymbol{J}}^n} \exp\big(-\beta H_{\boldsymbol{J}}^n(\boldsymbol{x})\big)\nu(\mathrm{d}\boldsymbol{x}). \tag{1.1}$$

The statics problem amounts to understanding the large n behaviour of these measures for various classes of random Hamiltonians ([Tal03] is a recent and beautiful book on the mathematical results pertaining to these equilibrium problems). The dynamics question consists of understanding the behaviour of Markovian processes on the configuration space Γ^n, for which the Gibbs measure is invariant and even reversible, in the limit of large systems (large n) and long times, either when the randomness $\boldsymbol{J}$ is fixed (the quenched case) or when it is averaged (often called the annealed case in the mathematics literature, but not in the physics papers). These dynamics are typically Glauber dynamics for the discrete spin setting, or Langevin dynamics for continuous spins.

Defining precisely what we mean here by large system size n and long time t is a very important question, and very different results can be expected for various time scales $t(n)$ as functions of the size of the system. The very wide range of time (and energy or space) scales present in the dynamics of these random media is the main interest and difficulty of these questions (in our view). At one end of the spectrum of time scales one could take first the limit when n goes to infinity and then t to infinity. This is the shortest possible long-time scale, much too short typically to allow any escape from metastable states since the energy barriers the system can cross are not allowed to diverge. This short time scale is well understood for dynamics of various spin-glass models (and related models) in the physics literature, in particular for the paradigmatic Langevin dynamics of spherical

2

p-spin models of spin-glasses, mainly through the equations derived by Cugliandolo and Kurchan [CK93] (see also [CHS93]). The fact that the results given by this short-time limit are seen as correct for models with a continuous replica symmetry breaking (like the Sherrington-Kirkpatrick model) is one of the many bewildering predictions made by the physicists. In the Les Houches lectures we covered some of the recent mathematical results about this short-time scale for Langevin dynamics obtained in collaboration with Alice Guionnet and Amir Dembo ([BDG06, BDG01, BG97, Gui97]). For lack of space and in order to keep a better focus we will not touch this topic here at all.

On the contrary we will be interested in the other end of the range of time scales, i.e. time scales $t_w(n)$ depending on the system size in such a way that they allow for the escape from the deep metastable states. This is where the introduction of the phenomenological trap models by Bouchaud becomes meaningful. We will now try to explain the relevance of these models in this setting, although in necessarily rather imprecise terms.

At low temperature the Gibbs measure $\mu_{\beta,\boldsymbol{J}}^n$ should be essentially carried by a small part of the configuration space, the "deep valleys" of the random landscape, i.e. the regions of low energy, (the "lumps" of p-spin models for instance). So that the dynamics should spend most of the time in these regions, which thus becomes very sticky "attractors" or "traps". The trap models ignore the details of the dynamics inside these sticky regions. They only keep the statistics of the height of the barriers that the system must cross before leaving these regions, and therefore, the statistics of the trapping times (i.e. the times needed to escape them). Moreover, the trap models keep the structure of the possible routes from one of these traps to the others. The dynamics is, therefore, reduced to its caricature: it lives only on the graph whose vertices are all the relevant attractors and the edges are the pairs of communicating attractors (see Section 2 for precise definition).

It is a non-trivial matter to prove that these phenomenological models could be of any relevance for the original problems. In fact, in the first introduction [Bou92]of this model a large complete graph was supposed to be a good ansatz for the simplest model of a mean-field spin-glass, i.e. Derrida's Random Energy Model [Der81]. Proving rigorously that Bouchaud's ansatz or "phenomenological" model was indeed a very good approximation for metastability and aging questions is quite delicate, and was done only recently, initially in [BBG03a, BBG03b]. Later we realised ([BČ06a]) that this very simple Bouchaud ansatz for the REM was in fact even better, in the sense that the range of time scales where it is a reliable approximation is very wide.

We now believe that this relevance is even much wider ("universal", if we dare) in the following sense: some of the lessons learnt from Bouchaud's picture in the REM should be relevant for very wide classes of mean-field spin-glasses in appropriate time scales. The art is in choosing these time scales long enough so that the (usually sparse) deep traps can be found and thus some trapping can take place, but short enough so that the deepest traps are not yet relevant and thus the equilibrium (which is of course heavily model-dependent) is not yet sampled by the dynamics. This belief has propped us into understanding more deeply the trap models in the larger possible generality of the graph structure and of the time scales involved. We have indeed found a very universal picture

3

valid for all the examples we have studied, except for the very particular one-dimensional Bouchaud trap model which belongs to another class, as we will see below.

We want to emphasise here that the dynamics of the mean-field spin-glasses is far from being the only motivation that makes the study of trap models worthwhile, see e.g. [BB03] for references of applications to fragile glasses, soft glassy and granular materials, and pinning of extended defects.

Let us now describe what these notes contain in more detail. We start by giving in Section 2 the definition of the Bouchaud trap model for a general graph and a general "depth" landscape. We then study, in Section 3, the very specific one-dimensional case (the graph here is $\mathbb{Z}$) in much detail. The results and the methods are different from all other situations we study. We rely essentially on the scaling limit introduced by Fontes, Isopi and Newman in [FIN02]. This scaling limit is an interesting self-similar singular diffusion, which gives quite easily results about aging, subaging and the "environment seen from the particle". We then go, in Section 4, to the d-dimensional case and show that the essence of the results is pretty insensitive to the dimension (as long as $d \geq 2$, with some important subtleties for the most difficult case, i.e. $d = 2$). In particular we also give a scaling limit, quite different from the one-dimensional case. We show that the properly-rescaled "internal clock" of the dynamics converges to an α-stable subordinator and that the process itself when properly rescaled converges to a "fractional-kinetics" type of dynamics ([Zas02]) which is simply the time change of a d-dimensional Brownian Motion by an independent process, the inverse of an α-stable subordinator. This process is also a self-similar continuous process, but it is no longer Markovian. In fact, this scaling-limit result is in some sense a (non-trivial) triviality result. It says that the Bouchaud trap model has the same scaling limit as a Continuous Time Random Walk à la Montroll-Weiss [MW65]. The aging results are then seen as a direct consequence of the generalised arcsine law for stable subordinators.

This picture (valid for all $d \geq 2$) is also naturally valid for the infinite-dimensional (or mean-field) case, i.e. for large complete graphs which we study in Section 5. Thus, we see that the critical mean-field dimension is 2 (in fact, we do not really guess what could happen for dimensions between 1 and 2, but it could be an interesting project to look at these models on say deterministic fractals with spectral dimension between 1 and 2). For large complete graphs, which is a very easy case, we choose to give a new proof instead of following the well-established route using renewal arguments (as in [BD95] and in [BBG03b]). This proof is slightly longer but illustrates, in this simple context, the strategy that we follow in other more difficult cases. An advantage of this line of proof is worth mentioning: we get aging results in longer time scales than usual.

We then use the intuition we hope to have given in the easy Section 5.1 to explain in Section 5.2 a general (universal?) scheme for aging based on the same arguments. We isolate the arguments needed for the proof given for the complete graph to work in general. This boils down to six technical conditions under very general circumstances. We then show how, under usual circumstances, these conditions can be reduced to basic potential-theoretic conditions for the standard random walk and random subsets of the graph. This general scheme is shown to be applicable not only to the cases we already

know (i.e. the Bouchaud model on $\mathbb{Z}^d$ with $d \geq 2$, or the large complete graphs), but also to dynamics of the Random Energy Model (in a wide range of time scales, shorter than the one given in [BBG03b], including some above the critical temperature) and also to very long time scales for large boxes in finite dimensions (with periodic boundary conditions).

2 Definition of the Bouchaud trap model

We define here a general class of reversible Markov chains on graphs, which were introduced by Bouchaud [Bou92] in order to give an effective model for trapping phenomena. The precise definition of these Markov chains necessitates three ingredients: a graph G, a trapping landscape $\boldsymbol{\tau}$ and an extra real parameter a.

We start with the graph G, $G = (\mathcal{V}, \mathcal{E})$, with the set of vertices $\mathcal{V}$ and the set of edges $\mathcal{E}$. We suppose that G is non-oriented and connected; G could be finite or infinite.

We then introduce the trapping landscape $\boldsymbol{\tau}$. For each vertex x, τ_x is a positive real number which is referred to as the depth of the trap at x. $\boldsymbol{\tau}$ can also be seen as a (positive) measure on $\mathcal{V}$,

$$\boldsymbol{\tau} = \sum_{x \in \mathcal{V}} \tau_x \delta_x. \tag{2.1}$$

Finally, we define the continuous time Markov chain $X(t)$ on $\mathcal{V}$ by its jump rates w_{xy},

$$w_{xy} = \begin{cases} \nu \tau_x^{-(1-a)} \tau_y^a, & \text{if } \langle x, y \rangle \in \mathcal{E}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

Hence, the generator of the chain is

$$Lf(x) = \sum_{y : \langle x, y \rangle \in \mathcal{E}} w_{xy}\big(f(y) - f(x)\big). \tag{2.3}$$

Here, the linear scaling factor ν defines a time unit and is irrelevant for the dynamical properties. Its value in these notes varies for different graphs, mostly for technical convenience. The parameter $a \in [0, 1]$ characterises the "symmetry" or "locality" of the dynamics; its role will be explained later. The initial state of the process will be also specified later, usually we set $X(0) = \mathbf{0}$, where $\mathbf{0}$ is an arbitrary fixed vertex of the graph.

Definition 2.1. Given a graph $G = (\mathcal{V}, \mathcal{E})$, a trapping landscape $\boldsymbol{\tau}$ and a real constant $a \in [0, 1]$, we define *the Bouchaud trap model*, $\mathsf{BTM}(G, \boldsymbol{\tau}, a)$ as the Markov chain $X(t)$ on $\mathcal{V}$ whose dynamics is given by (2.2) and (2.3).

In the original introduction of the model [Bou92], the trapping landscape $\boldsymbol{\tau}$ is given by a non-normalised Gibbs measure

$$\boldsymbol{\tau} = \sum_{x \in \mathcal{V}} \tau_x \delta_x = \sum_{x \in \mathcal{V}} e^{-\beta E_x} \delta_x, \tag{2.4}$$

5

where $\beta > 0$ is the inverse temperature and E_x is seen as the energy at x. The rates w_{xy} can be then expressed using the random variables E_x instead of τ_x,

$$w_{xy} = \nu \exp\left\{\beta\big((1-a)E_x - aE_y\big)\right\}, \qquad \text{if } \langle x,y\rangle \in \mathcal{E}. \tag{2.5}$$

It is easy to check that $\boldsymbol{\tau}$ is a reversible measure for the Markov chain $X(t)$; the detailed balance condition is easily verified:

$$\tau_x w_{xy} = \nu \tau_x^a \tau_y^a = \tau_y w_{yx}. \tag{2.6}$$

Let us give some intuition about the dynamics of the chain X. Consider the embedded discrete-time Markov chain $Y(n)$:

$$X(t) = Y(n) \qquad \text{for all } S(n) \le t < S(n+1), \tag{2.7}$$

where $S(0) = 0$ and $S(n)$ is the time of the nth jump of $X(\cdot)$.

If $a = 0$, the process X is particularly simple. Its jumping rates w_{xy} do not depend on the depth of the target vertex y. X waits at the vertex x an exponentially distributed time with mean $\tau_x(\nu d_x)^{-1}$, where d_x is the degree of x in G. After this time, it jumps to one of the neighbours of x chosen uniformly at random. Hence, the embedded discrete-time Markov chain $Y(n)$ is a simple random walk on the graph G and $X(t)$ is its time change. The $a = 0$ dynamics is sometimes referred to as *Random Hopping Times* (RHT) dynamics.

If $a \neq 0$, the jumping rates w_{xy} depend on the target vertex. The process therefore does not jump uniformly to all neighbours of x. $Y(n)$ is no longer a simple random walk but a kind of discrete Random Walk in Random Environment. To observe the effects of $a > 0$ it is useful to consider a particular relatively deep trap x with much shallower neighbours. In this case, as a increases, the mean waiting time at x decreases. On the other hand, if X is located at some of the neighbours of x, then it is attracted by the deep trap x since w_{yx} is relatively large. Hence, as a increases, the process X stays at x a shorter time, but, after leaving it, it has larger probability to return there. We will see later that these two competing phenomena might exactly cancel in the long-time behaviour of well-chosen characteristics of the Markov chain.

We are not interested here in a natural line of questions which would be to find the best conditions under which the trapping mechanism is not crucial, and the BTM behaves as a simple random walk. On the contrary, we want to see how the trapping landscape can have a strong influence on the long time behaviour. Obviously, this can happen only if this trapping landscape is strongly inhomogeneous.

Strong inhomogeneity can be easily achieved in the class of random landscapes with heavy tails, the essential hypothesis being that the expectation of the depth should be infinite. One of the assumptions we will use is therefore:

Assumption 2.2. The depths τ_x are positive i.i.d. random variables belonging to the domain of the attraction of the totally asymmetric α-stable law with $\alpha \in (0,1)$. This means that there exists a slowly varying function L (i.e., for all $s > 0$ $\lim_{u\to\infty} L(us)/L(u) = 1$), such that

$$\mathbb{P}[\tau_x \ge u] = u^{-\alpha}L(u). \tag{2.8}$$

Sometimes, to avoid unnecessary technical difficulties, we use the stronger assumption:

Assumption 2.3. The depths τ_x are positive i.i.d. random variables satisfying

$$\lim_{u \to \infty} u^\alpha \mathbb{P}[\tau_x \geq u] = 1, \qquad \alpha \in (0, 1). \tag{2.9}$$

These assumptions are satisfied at low temperature for the standard choice of the statistical physics literature: for $-E_x$ being an i.i.d. collection of exponentially distributed random variables with mean 1. The depth of the traps then satisfies

$$\mathbb{P}[\tau_x \geq u] = \mathbb{P}[e^{-\beta E_x} \geq u] = u^{-1/\beta}, \qquad u \geq 1. \tag{2.10}$$

Hence, if $\beta > 1$, then Assumption 2.3 is satisfied with $\alpha = 1/\beta$, and the expected value of the depth diverges.

2.1 Examples of trap models

Specifically, we will consider the following models:

1. The Bouchaud model on $\mathbb{Z}$. Here G is the integer lattice, $\mathbb{Z}$, with the nearest neighbour edges. The trapping landscape will be as in Assumption 2.2. We will report in Section 3 on work by [FIN02, BČ05, Čer06].

2. The BTM on $\mathbb{Z}^d$, $d > 1$, with a random landscape as in Assumption 2.3. The main results about this case are contained in [BČM06, Čer03, BČ06b].

We will also deal with a generalisation of the former setting, i.e., with a sequence of Bouchaud trap models, $\{\mathsf{BTM}(G_n, \boldsymbol{\tau}_n, a) : n \in \mathbb{N}\}$. We will then consider different time scales depending on n and write $X_n(t)$ for the Markov chain on the level n. In this case the law of τ_x may depend on n. Then the requirement $\mathbb{E}[\tau_x] = \infty$ is not necessary, as we will see.

3. The BTM in a large box in $\mathbb{Z}^d$ with periodic boundary condition, here $G_n = \mathbb{Z}^d/n\mathbb{Z}^d$ is the torus of size n [BČ06a].

4. The BTM on a large complete graph. Here we consider a sequence of complete graphs with n vertices. This is the model that was originally proposed in [Bou92].

5. The Random Energy Model (REM) dynamics. We deal here with a sequence of n-dimensional hypercubes $G_n = \{-1, 1\}^n$. The landscape will be given by normally distributed (centred, with variance n) energies E_x. This is the case where τ's are not heavy-tailed.

We will see that all the previous cases with exception of the one-dimensional lattice behave very similarly.

2.2 Natural questions on trap models

We will be interested in the long-time behaviour of the BTM. There are several natural questions to ask in order to quantify the influence of trapping. The most important question we will address in these notes is the question of aging (at different time scales). Let us now define what we will call aging and subaging here. We will consider a time interval $[t_w, t_w + t]$, where the *waiting time* t_w (or the age of the system) as well as the length t of the time window (or duration of the observation) will grow to infinity. We will then consider various two-time functions, say $C(t_w, t_w + t)$, which depend on the trajectory of the Bouchaud trap model in the time window $[t_w, t_w + t]$. We will say that there is aging for these two-time functions iff

$$\lim_{t_w \to \infty} C(t_w, t_w + \theta t_w) \tag{2.11}$$

exists and is non-trivial. We will call this limit, $C(\theta)$, the aging function. We will say that there is subaging with exponent $\gamma < 1$ iff

$$\lim_{t_w \to \infty} C(t_w, t_w + \theta t_w^\gamma) \tag{2.12}$$

exists and is again non-trivial.

We need now to define good two-time functions in order to be able to deal with aging for BTM. The following functions are mostly studied:

(a) The probability that, conditionally on $\boldsymbol{\tau}$, the process does not jump during the specified time interval $[t_w, t_w + t]$,

$$\Pi(t_w, t_w + t; \boldsymbol{\tau}) = \mathbb{P}\big[X(t') = X(t_w)\, \forall t' \in [t_w, t_w + t]\big|\boldsymbol{\tau}\big]. \tag{2.13}$$

(b) The probability that the system is in the same trap at both times t_w and $t_w + t$,

$$R(t_w, t_w + t; \boldsymbol{\tau}) = \mathbb{P}[X(t_w) = X(t_w + t)|\boldsymbol{\tau}]. \tag{2.14}$$

(c) And finally, the quantity

$$R^q(t_w, t_w + t; \boldsymbol{\tau}) = \mathbb{E}\Big[\sum_{x \in \mathcal{V}}[\mathbb{P}(X(t_w + t) = x|\boldsymbol{\tau}, X(t_w))]^2\Big|\boldsymbol{\tau}\Big], \tag{2.15}$$

which is the probability that two independent walkers will be at the same site after time $t + t_w$ if they were at the same site at time t_w, averaged over the distribution of the common starting point $X(t_w)$.

These quantities are random objects: they still depend on the randomness of the trapping landscape $\boldsymbol{\tau}$. They are usually called *quenched* two-time functions. In addition to the quenched two-time functions, we will also consider their average over the random landscape. We define the *averaged* two-time functions:

$$\begin{aligned}
\Pi(t_w, t_w + t) &= \mathbb{P}\big[X(t') = X(t_w)\, \forall t' \in [t_w, t_w + t]\big], \\
R(t_w, t + t_w) &= \mathbb{P}[X(t_w) = X(t_w + t)], \\
R^q(t_w, t_w + t) &= \mathbb{E}\Big[\sum_{x \in \mathcal{V}}[\mathbb{P}(X(t_w + t) = x|\boldsymbol{\tau}, X(t_w))]^2\Big].
\end{aligned} \tag{2.16}$$

We will state aging results for both quenched and averaged two-time functions for the various Bouchaud trap models given in Section 2.1. We will strive to get the widest possible range of time scales where aging occurs.

Even though our main motivation was the study of aging, there are many other questions about the long-time behaviour of the BTM which are of interest. For instance:

- The behaviour of the environment seen from the particle. Here the prominent feature of this environment seen from the position at time t is simply the depth of the trap $\tau_{X(t)}$ where the process is located. We will give limit theorems for this quantity, which are crucial for most of the aging results.

- Nature of the spectrum of the Markov chain close to its edge. Naturally, the long time behaviour of $X(t)$ can be understood from the edge of the spectrum of the generator L. This question deserves further study (see [BF05, BF06] and also [MB97]). We intend to address this question for BTM in finite dimensions in a forthcoming work.

- Anomalous diffusion. In the case where graph is $\mathbb{Z}^d$, can we see that $X(t)$ is slow: for instance that $\mathbb{E}[X(t)^2] \ll t$? Or get the tail behaviour of $|X(t)|$?

- Scaling limit. Again in the case of $\mathbb{Z}^d$, is there a scaling limit for the process $X(t)$, i.e. a way to normalise space and time so that $X^\varepsilon(t) = \varepsilon X(t/h(\varepsilon))$ converges to a process on $\mathbb{R}^d$ which we can describe.

2.3 References

The physics literature on trap models is so abundant that we cannot try to be exhaustive. For earlier references on finite-dimensional questions see [Mac85, BG90] where the anomalous character of the diffusion is given as well as a scaling limit. For aging questions on large complete graphs and relation to spin-glass dynamics see [Bou92, BD95], see also [BCKM98] for a more global picture. For aging questions for the finite-dimensional model see [MB96, RMB00, RMB01, BB03] among many other studies.

We will give references to mathematical papers at the end of every section, when necessary.

3 The one-dimensional trap model

We will consider in this section $\mathsf{BTM}(\mathbb{Z}, \boldsymbol{\tau}, a)$ on the one-dimensional lattice, $\mathbb{Z}$, with nearest-neighbour edges. The random depths, τ_x, will be taken to be i.i.d., in the domain of the attraction of an α-stable law, $\alpha < 1$, as in Assumption 2.2. There are several reasons why this particular graph should be treated apart. First, as usual, the one-dimensional model is easier to study. Second, as we have already mentioned, the one-dimensional BTM has some specific features that distinguish it from all other cases presented later in these notes.

Another distinguishing feature (of technical character) is that, at present, the one-dimensional BTM is the only case where the asymmetric variant ($a > 0$) has been rigorously studied. For technical convenience we choose here $\nu = \nu_a = \mathbb{E}[\tau_0^{-a}]^2/2$, that is we set

$$w_{xy} = \frac{1}{2}\mathbb{E}[\tau_0^{-a}]^2 \tau_x^{-(1-a)} \tau_y^a, \qquad \text{if } |x - y| = 1. \tag{3.1}$$

We set $X(0) = 0$. We suppose that ν_a is finite for all $a \in [0, 1]$, this is obviously the case if, e.g., $\tau_0 > c$ a.s. for some $c > 0$.

3.1 The Fontes-Isopi-Newman singular diffusion

The most useful feature of the one-dimensional BTM is that we can identify its scaling limit as an interesting one-dimensional singular diffusion in random environment introduced by Fontes, Isopi and Newman [FIN02].

Definition 3.1 (The F.I.N. diffusion). Let (x_i, v_i) be an inhomogeneous Poisson point process on $\mathbb{R} \times (0, \infty)$ with intensity measure $dx\, \alpha v^{-1-\alpha} dv$. Define the random discrete measure $\rho = \sum_i v_i \delta_{x_i}$. We call ρ the random environment. Conditionally on ρ, we define the *F.I.N. diffusion* $Z(s)$ as a diffusion process (with $Z(0) = 0$) that can be expressed as a time change of a standard one-dimensional Brownian motion $B(t)$ with the speed measure ρ, as follows [IM65]: Denoting by $\ell(t, y)$ the local time of the standard Brownian motion $B(t)$ at y, we define

$$\phi_\rho(t) = \int_{\mathbb{R}} \ell(t, y)\rho(dy) \tag{3.2}$$

and its generalised right-continuous inverse

$$\psi_\rho(s) = \inf\{t > 0 : \phi_\rho(t) > s\}. \tag{3.3}$$

Then $Z(s) = B(\psi_\rho(s))$.

The following proposition lists some of the properties of the diffusion Z and the measure ρ that may be of interest.

Proposition 3.2. *(i) The intensity measure of the Poisson point process is non-integrable at $v = 0$, therefore the set of all atoms of ρ is a.s. dense in $\mathbb{R}$.*

(ii) Conditionally on ρ, the distribution of $Z(t)$, $t > 0$, is a discrete probability measure $\nu_t^\rho = \sum_i w_i(t)\delta_{x_i}$, with the same set of atoms as ρ.

(iii) The diffusion Z has continuous sample paths.

(iv) Define $p_\rho(t, x_i) = \mathbb{P}[Z(t) = x_i|\rho]/v_i$ for all atoms x_i of ρ and $t > 0$. The function p_ρ has a unique jointly continuous extension to $(0, \infty) \times \mathbb{R}$. Moreover, p_ρ satisfies the following equation

$$\frac{\partial}{\partial t}p_\rho(t, x) = \frac{\partial^2}{\partial\rho\partial x}p_\rho(t, x). \tag{3.4}$$

The singular differential operator $\partial^2 f/\partial\rho\partial x$ is defined (see, e.g., [DM76, KW82]) by $h = \partial^2 f/\partial\rho\partial x$ if for some $c_1, c_2 \in \mathbb{R}$

$$f(x) = c_1 + \int_0^x \left(c_2 + \int_0^u h(v)\rho(\mathrm{d}v) \right) \mathrm{d}u. \tag{3.5}$$

(v) The diffusion Z and its speed measure ρ are self-similar: for all $\lambda > 0$, $t > 0$ and $x \in \mathbb{R}$

$$\rho([0, x]) \stackrel{\text{law}}{=} \lambda^{-1/\alpha}\rho([0, \lambda x]) \quad and \quad Z(t) \stackrel{\text{law}}{=} \lambda^{-1}Z(t\lambda^{(1+\alpha)/\alpha}). \tag{3.6}$$

Therefore, the diffusion Z is anomalous.

(vi) There exist constants C, c such that for all x and $t > 0$

$$\mathbb{P}\big[|Z(t)| \geq x\big] \leq C\exp\Big[-c\Big(\frac{x}{t^{\frac{\alpha}{1+\alpha}}}\Big)^{1+\alpha}\Big]. \tag{3.7}$$

Proof. Statement (i) is trivial, (ii) is proved in [FIN02]. Claim (iii) follows from (i), the continuity of sample paths of B, and the definition of Z. (iv) is a non-trivial claim of the theory of quasi-diffusions, see above references. The first part of (v) is a direct consequence of the definition of ρ. The second part then follows from the first one and from the well-known scaling relations for the Brownian motion B and its local time:

$$B(t) \stackrel{\text{law}}{=} \lambda^{-1}B(\lambda^2 t) \qquad and \qquad \ell(t, x) \stackrel{\text{law}}{=} \lambda^{-1}\ell(\lambda^2 t, \lambda x). \tag{3.8}$$

The last claim is proved in [Čer06]. $\qquad\square$

Remark. The scale of the upper bound in (vi) is probably optimal. The corresponding lower bound was however never proved. The numerical simulations and non-rigorous arguments in [BB03] however support this conjecture, and give even exact values for constants C and c.

3.2 The scaling limit

We now explain how the F.I.N. diffusion appears as a scaling limit of the BTM. For all $\varepsilon \in (0, 1)$ we consider the rescaled process

$$X^\varepsilon(t) = \varepsilon X(t/\varepsilon c_\varepsilon), \tag{3.9}$$

where

$$c_\varepsilon = \big(\inf[t \geq 0 : \mathbb{P}(\tau_0 > t) \leq \varepsilon]\big)^{-1}. \tag{3.10}$$

It follows from Assumption 2.2 that there is a slowly varying function L' such that $c_\varepsilon = \varepsilon^{1/\alpha}L'(1/\varepsilon)$. We will also consider the following rescaled landscapes,

$$\boldsymbol{\tau}^\varepsilon(\mathrm{d}x) = c_\varepsilon \sum_{y \in \mathbb{Z}} \tau_y \delta_{\varepsilon y}(\mathrm{d}x) =: \sum_{y \in \mathbb{Z}} \tau_{\varepsilon y}^\varepsilon \delta_{\varepsilon y}(\mathrm{d}x). \tag{3.11}$$

It is not so difficult to see that the distribution of $\boldsymbol{\tau}^\varepsilon$ converges to the distribution of ρ as $\varepsilon \to 0$. The next proposition states that it is possible to construct a coupling between different scales such that the convergence becomes almost sure.

Proposition 3.3 (Existence of coupling). *There exists a family of measures $\bar{\boldsymbol{\tau}}^\varepsilon$ and processes $\bar{X}^\varepsilon$ constructed on the same probability space as the measure ρ and the Brownian motion B such that*

(i) *For all $\varepsilon > 0$, $\bar{\boldsymbol{\tau}}^\varepsilon$ has the same distribution as $\boldsymbol{\tau}^\varepsilon$.*

(ii) *The measures $\bar{\boldsymbol{\tau}}^\varepsilon$ converge to ρ vaguely and in the point-process sense, ρ-a.s.*

(iii) *$\bar{X}^\varepsilon$ can be expressed as a time(-scale) change of the Brownian motion B with the speed measure $\bar{\boldsymbol{\tau}}^\varepsilon$ (see Section 3.2.1). It has the same distribution as X^ε.*

We can now state the principal theorem of this section. It was proved in [FIN02] for $a = 0$ and in [BČ05] for $a > 0$.

Theorem 3.4 (Scaling limit of the one-dimensional BTM). *As $\varepsilon \to 0$, for every fixed $t > 0$ and all $a \in [0,1]$, the distribution of $\big(\bar{X}^\varepsilon(t), \bar{\tau}^\varepsilon_{\bar{X}^\varepsilon(t)}\big)$ converges weakly and in the point-process sense to the distribution of $\big(Z(t), \rho(\{Z(t)\})\big)$, ρ-a.s.*

The notion of convergence in the point-process sense used in the theorem was introduced in [FIN02]. It is used here because we want to deal with quantities like $\mathbb{P}[X^\varepsilon(t) = X^\varepsilon(t')]$ to prove aging. The usual weak or vague convergences of measures are insensitive to such kind of quantities. This notion of convergence is defined by

Definition 3.5 (Point-process convergence). Given a family ν, ν^ε, $\varepsilon > 0$, of locally finite measures on $\mathbb{R}$, we say that ν^ε converges *in the point process sense* to ν, and write $\nu^\varepsilon \overset{pp}{\to} \nu$, as $\varepsilon \to 0$, provided the following holds: If the atoms of ν, ν^ε are, respectively, at the distinct locations y_i, $y^\varepsilon_{i'}$ with weights w_i, $w^\varepsilon_{i'}$, then the subsets of $U^\varepsilon \equiv \cup_{i'}\{(y^\varepsilon_{i'}, w^\varepsilon_{i'})\}$ of $\mathbb{R} \times (0, \infty)$ converge to $U \equiv \cup_i\{(y_i, w_i)\}$ as $\varepsilon \to 0$ in the sense that for any open O, whose closure is a compact subset of $\mathbb{R} \times (0, \infty)$ such that its boundary contains no points of U, the number of points $|U^\varepsilon \cap O|$ in $U^\varepsilon \cap O$ is finite and equals $|U \cap O|$ for all ε small enough.

Remark. The convergence of $\bar{\tau}^\varepsilon_{\bar{X}^\varepsilon(t)}$ gives a description of the environment seen by the particle. More explicitly, the distribution of the normalised depth of the trap where X is located at large time $c_\varepsilon \tau_{X(t/\varepsilon c_\varepsilon)}$ converges to the distribution of $\rho(\{Z(t)\})$.

We now sketch the three main tools that are used in the proof of Theorem 3.4.

3.2.1 Time-scale change of Brownian motion

To better understand how (Z, ρ) arises as the scaling limit of $(X, \boldsymbol{\tau})$, one should use the fact that not only diffusions, but also nearest-neighbour random walks in dimension one, can be expressed as time(-scale) change of the Brownian motion. The scale change is necessary only if $a \neq 0$, because the process $X(t)$ does not jump left or right with equal probabilities.

We first define the time-scale change. Consider a locally-finite, discrete, non-random measure

$$\mu(\mathrm{d}x) = \sum_i w_i \delta_{y_i}(\mathrm{d}x), \tag{3.12}$$

which has atoms with weights w_i at positions y_i. The measure μ will be referred to as the *speed measure*. Let S be a strictly increasing function defined on the set $\{y_i\}$. We call such S the *scaling function*. Let us introduce slightly nonstandard notation $S \circ \mu$ for the "scaled measure"

$$(S \circ \mu)(\mathrm{d}x) = \sum_i w_i \delta_{S(y_i)}(\mathrm{d}x). \tag{3.13}$$

Similarly as in definition of Z, we define the function

$$\phi(\mu, S)(t) = \int_{\mathbb{R}} \ell(t, y)(S \circ \mu)(\mathrm{d}y) \tag{3.14}$$

and the stopping time $\psi(\mu, S)(s)$ as the first time when $\phi(\mu, S)(t) = s$. The function $\phi(\mu, S)(t)$ is a nondecreasing, continuous function, and $\psi(\mu, S)(s)$ is its generalised right continuous inverse. It is an easy corollary of the results of [Sto63] that the process

$$X(\mu, S)(t) := S^{-1}(B(\psi(\mu, S)(t))) \tag{3.15}$$

is a nearest-neighbour random walk on the set of atoms of μ. Moreover, every nearest-neighbour random walk on a countable, nowhere-dense subset of $\mathbb{R}$ satisfying some mild conditions on transition probabilities can be expressed in this way. We call the process $X(\mu, S)$ the *time-scale change* of the Brownian motion. If $S = \mathrm{Id}$, the identity mapping, we speak only about the time change.

The following proposition summarises the properties of $X(\mu, S)$ if the set of atoms of μ has no accumulation point. In this case we can suppose that the locations of atoms y_i satisfy $y_i < y_j$ if $i < j$.

Proposition 3.6 (Stone, [Sto63])**.** *The process $X(\mu, S)(t)$ is a nearest-neighbour random walk on the set $\{y_i\}$ of atoms of μ. The waiting time in the state y_i is exponentially distributed with mean*

$$2w_i \frac{(S(y_{i+1}) - S(y_i))(S(y_i) - S(y_{i-1}))}{S(y_{i+1}) - S(y_{i-1})}. \tag{3.16}$$

After leaving state y_i, $X(\mu, S)$ enters states y_{i-1} and y_{i+1} with respective probabilities

$$\frac{S(y_{i+1}) - S(y_i)}{S(y_{i+1}) - S(y_{i-1})} \quad and \quad \frac{S(y_i) - S(y_{i-1})}{S(y_{i+1}) - S(y_{i-1})}. \tag{3.17}$$

Using this proposition it is possible to express the processes $X^\varepsilon(t)$ (see (3.9)) as a time-scale change of the Brownian motion B. It is not surprise that $\boldsymbol{\tau}^\varepsilon$ should be chosen as the speed measures. The scaling function is defined by

$$S(x) = \begin{cases} \sum_{y=0}^{x-1} r_y, & \text{if } x \geq 0, \\ -\sum_{y=x}^{-1} r_y, & \text{otherwise}, \end{cases} \tag{3.18}$$

13

where

$$r_x = \frac{1}{2} \nu_a^{-1} \tau_x^{-a} \tau_{x+1}^{-a}. \tag{3.19}$$

Observe that ν_a was chosen in such way that $\mathbb{E}[r_x] = 1$. If $a = 0$, S is the identity mapping on $\mathbb{Z}$, there is no scale change in this case. Define further $S^\varepsilon(\cdot) = \varepsilon S(\varepsilon^{-1} \cdot)$. It is easy to check, using Proposition 3.6, that the processes $X(\boldsymbol{\tau}^\varepsilon, S^\varepsilon)$ have the same distribution as X^ε.

It is convenient to introduce processes $\mathcal{X}^\varepsilon(t)$ that are the time change of the Brownian motion with speed measures $S^\varepsilon \circ \boldsymbol{\tau}^\varepsilon$. Namely,

$$\mathcal{X}^\varepsilon(t) = X(S^\varepsilon \circ \boldsymbol{\tau}^\varepsilon, \mathrm{Id})(t). \tag{3.20}$$

The processes $\mathcal{X}^\varepsilon$ are related to X^ε by $X^\varepsilon(t) = (S^\varepsilon)^{-1}(\mathcal{X}^\varepsilon(t))$.

3.2.2 Convergence of the fixed-time distributions

We have expressed the processes $\mathcal{X}^\varepsilon$ as the time change of the Brownian motion with the speed measure $S^\varepsilon \circ \boldsymbol{\tau}^\varepsilon$. We want to show that $\mathcal{X}^\varepsilon$ and mainly X^ε converge to Z. As stated in the following important theorem, it is sufficient to check the convergence of the speed measures to prove the convergence of fixed time distributions. Observe that the theorem deals only with non-random measures.

Theorem 3.7 ([Sto63, FIN02]). *Let μ^ε, μ be a collection of* non-random *locally-finite measures, and let $\mathcal{Y}^\varepsilon$, $\mathcal{Y}$ be defined by*

$$\mathcal{Y}^\varepsilon(t) = X(\mu^\varepsilon, \mathrm{Id})(t) \quad and \quad \mathcal{Y}(t) = X(\mu, \mathrm{Id})(t). \tag{3.21}$$

For any deterministic $t_0 > 0$, let ν^ε denote the distribution of $\mathcal{Y}^\varepsilon(t_0)$ and ν denote the distribution of $\mathcal{Y}(t_0)$. Suppose that

$$\mu^\varepsilon \xrightarrow{v} \mu \quad and \quad \mu^\varepsilon \xrightarrow{pp} \mu \quad as \quad \varepsilon \to 0. \tag{3.22}$$

Then, as $\varepsilon \to 0$,

$$\nu^\varepsilon \xrightarrow{v} \nu \quad and \quad \nu^\varepsilon \xrightarrow{pp} \nu. \tag{3.23}$$

(Here $\xrightarrow{v}$ stands for the vague convergence.)

3.2.3 A coupling for walks on different scales.

The major pitfall of the preceding theorem is that it works only with sequences of deterministic speed measures. We want, however, to consider random speed measures $\boldsymbol{\tau}^\varepsilon$. As we have already remarked, it is not difficult to see that $\boldsymbol{\tau}^\varepsilon$ converge to ρ vaguely in distribution. However, it is not enough to make an application of Theorem 3.7 possible. Here the coupling whose existence is stated in Proposition 3.3(i) comes into play. It allows to replace the convergence in distribution by the almost sure convergence. Then it is possible to apply Theorem 3.7. Let us construct this coupling.

Consider a two-sided Lévy process (or α-stable subordinator) $U(x)$, $x \in \mathbb{R}$, $U(0) = 0$, with stationary and independent increments and cadlag paths defined by

$$\mathbb{E}\big[e^{-\lambda(U(x+x_0)-U(x_0))}\big] = \exp\Big[x\alpha \int_0^\infty (e^{-\lambda w} - 1)w^{-1-\alpha}\mathrm{d}w\Big]. \tag{3.24}$$

Let $\bar{\rho}$ be the random Lebesgue-Stieltjes measure on $\mathbb{R}$ associated to U, $\bar{\rho}(a, b] = U(b) - U(a)$. It is a known fact that $\bar{\rho}(\mathrm{d}x) = \sum_j v_j \delta_{x_j}(\mathrm{d}x)$, where (x_j, v_j) is an inhomogeneous Poisson point process with density $\mathrm{d}x\,\alpha v^{-1-\alpha}\mathrm{d}v$, which means that $\bar{\rho}$ has the same distribution as ρ.

For each fixed $\varepsilon > 0$, we will now define the sequence of i.i.d. random variables τ_x^ε such that τ_x^ε's are functions of U and have the same distribution as τ_0. Let $G : [0, \infty) \mapsto [0, \infty)$ be such that

$$\mathbb{P}(U(1) > G(x)) = \mathbb{P}(\tau_0 > x). \tag{3.25}$$

It is well defined since $U(1)$ has continuous distribution, it is nondecreasing and right-continuous, and hence has nondecreasing right-continuous generalised inverse G^{-1}.

Lemma 3.8. *Let*

$$\tau_x^\varepsilon := G^{-1}\big(\varepsilon^{-1/\alpha}(U(\varepsilon(x+1)) - U(\varepsilon x))\big). \tag{3.26}$$

Then for any $\varepsilon > 0$, the τ_x^ε are i.i.d. with the same law as τ_0.

Proof. By stationarity and independence of increments of U it is sufficient to show $\mathbb{P}(\tau_0^\varepsilon > t) = \mathbb{P}(\tau_0 > t)$. However,

$$\mathbb{P}(\tau_0^\varepsilon > t) = \mathbb{P}(U(\varepsilon) > \varepsilon^{1/\alpha}G(t)) \tag{3.27}$$

by the definitions of τ_0^ε and G. The result then follows from (3.25) and the scaling invariance of U: $U(\varepsilon) \overset{\text{law}}{=} \varepsilon^{1/\alpha}U(1)$. $\qquad\square$

Let us now define the random speed measures $\bar{\boldsymbol{\tau}}^\varepsilon$ using the collections $\{\tau_x^\varepsilon\}$ from the previous lemma,

$$\bar{\boldsymbol{\tau}}^\varepsilon(\mathrm{d}x) = \sum_{i \in \mathbb{Z}} c_\varepsilon \tau_i^\varepsilon \delta_{\varepsilon i}(\mathrm{d}x). \tag{3.28}$$

Finally, using τ_x^ε instead of τ_x, we define the scaling functions $\bar{S}^\varepsilon$ similarly as in (3.18) and (3.19). The process $\bar{X}^\varepsilon$ is then given by $\bar{X}^\varepsilon = X(\bar{\boldsymbol{\tau}}^\varepsilon, \bar{S}^\varepsilon)$, and the construction of the coupling from Proposition 3.3 is finished.

3.2.4 Scaling limit

Using the three tools introduced above, we can now sketch the proof of Theorem 3.4. Actually, not many steps remain.

First, to prove the convergence of $\mathcal{X}^\varepsilon$ to Z it is sufficient to verify the a.s. convergence of the speed measures $\bar{S}^\varepsilon \circ \bar{\boldsymbol{\tau}}^\varepsilon$ to $\bar{\rho}$ and then apply Theorem 3.7. The proof of this convergence is not difficult, however, slightly lengthy. It can be found in [FIN02] and [BČ05].

Proposition 3.9. *Let $\bar{\tau}^\varepsilon$ and $\bar{\rho}$ be defined as above. Then*

$$\bar{S}^\varepsilon \circ \bar{\tau}^\varepsilon \xrightarrow{v} \bar{\rho} \quad and \quad \bar{S}^\varepsilon \circ \bar{\tau}^\varepsilon \xrightarrow{pp} \bar{\rho} \quad as \ \varepsilon \to 0, \quad \bar{\rho}\text{-}a.s. \tag{3.29}$$

Finally, to pass from the convergence of $\mathcal{X}^\varepsilon$ to the convergence of X^ε it is necessary to control the scaling functions $\bar{S}^\varepsilon$.

Lemma 3.10. *As $\varepsilon \to 0$ we have*

$$\bar{S}^\varepsilon(\varepsilon\lfloor\varepsilon^{-1}y\rfloor) \to y, \qquad \bar{\rho}\text{-}a.s., \tag{3.30}$$

uniformly on compact intervals.

Observe that this lemma also implies that the embedded discrete-time random walk Y converges, after a renormalisation, to the Brownian motion, independently of the value of a. This is valid also if $a > 0$ and the discrete-time embedded process Y is not a simple random walk but a random walk in random environment.

Since Lemma 3.10 is one of the key parts of the proof of the scaling limit for $a \neq 0$ we prove it here.

Proof of Lemma 3.10. We consider only $y > 0$. The proof for $y < 0$ is very similar. By definition of $\bar{S}^\varepsilon$ we have $\varepsilon\bar{S}^\varepsilon(\lfloor\varepsilon^{-1}y\rfloor) = \varepsilon \sum_{j=0}^{\lfloor\varepsilon^{-1}y\rfloor-1} \bar{r}_j^\varepsilon$, where for fixed ε the sequence $\bar{r}_i^\varepsilon$ is an ergodic sequence of bounded positive random variables. Moreover, $\bar{r}_i^\varepsilon$ is independent of all $\bar{r}_j^\varepsilon$ with $j \notin \{i-1, i, i+1\}$. The $\bar{\rho}$-a.s. convergence for fixed y is then a consequence of the strong law of large numbers for triangular arrays. Note that this law of large numbers can be easily proved in our context using the standard methods, because the variables $\bar{r}_i^\varepsilon$ are bounded and thus their moments of arbitrary large degree are finite. The uniform convergence on compact intervals is easy to prove using the fact that $\bar{S}^\varepsilon$ is increasing and the identity function is continuous. $\qquad\square$

3.3 Aging results

The aging results for the one-dimensional BTM follow essentially from Theorem 3.4. To control the two-time functions R and R^q it is only necessary to extend its validity to the joint distribution of $(X^\varepsilon(1), X^\varepsilon(1+\theta))$ at two fixed times, which is not difficult. This extension then yields the following aging result.

Theorem 3.11 (Aging in the one-dimensional BTM)**.** *For any $\alpha \in (0,1)$, $\theta > 0$ and $a \in [0,1]$ there exist aging functions $R_1(\theta)$, $R^q(\theta)$ such that*

$$\lim_{t_w \to \infty} R(t_w, t_w + \theta t_w) = \lim_{t_w \to \infty} \mathbb{EP}[X((1+\theta)t_w) = X(t_w)|\boldsymbol{\tau}] = R_1(\theta),$$

$$\lim_{t_w \to \infty} R^q(t_w, t_w + \theta t_w) = \lim_{t_w \to \infty} \mathbb{E} \sum_{i \in \mathbb{Z}} [\mathbb{P}(X((1+\theta)t_w) = i|\boldsymbol{\tau}, X(t_w))]^2 = R^q(\theta). \tag{3.31}$$

Moreover, $R_1(\theta)$ and $R^q(\theta)$ can be expressed using the analogous quantities defined using the singular diffusion Z:

$$R_1(\theta) = \mathbb{EP}[Z(1+\theta) = Z(1)|\rho],$$
$$R^q(\theta) = \mathbb{E} \sum_{x \in \mathbb{R}} [\mathbb{P}(Z(1+\theta) = x|\rho, Z(1))]^2. \tag{3.32}$$

Remark. 1. This result is contained in [FIN02] for $a = 0$ and in [BČ05] for $a > 0$.

2. Let us emphasise that the functions $R_1(\theta)$, $R^q(\theta)$ do not depend on the parameter a, since the diffusion $Z(t)$ and the measure ρ do not depend on it. This is the result of the compensation of shorter visits of deep traps by the attraction to them.

3. It should be also underlined that only averaged functions are considered in the theorem. For a fixed realisation of $\boldsymbol{\tau}$ there is no limit of $R(t_w, t_w + \theta t_w; \boldsymbol{\tau})$. From the proof of Theorem 3.11, it is however not difficult to derive the following weaker result (see Theorem 1.3 in [Čer06]).

Theorem 3.12 (Quenched aging on $\mathbb{Z}$, in distribution)**.** *As $t_w \to \infty$, the distribution of $R(t_w, t_w + \theta t_w; \boldsymbol{\tau})$ converges weakly to the distribution of $\mathbb{P}[Z(1 + \theta) = Z(1)|\rho]$.*

3.4 Subaging results

In the case of the two-time function Π much shorter times t should be considered, $t \ll t_w$, i.e. subaging takes place:

Theorem 3.13. *For any $\alpha \in (0, 1)$, $\theta > 0$ and $a \in [0, 1]$ there exist an aging function $\Pi_{1,a}(\theta)$ such that*

$$\lim_{t \to \infty} \Pi(t, t + f_a(t, \theta)) = \lim_{t \to \infty} \mathbb{EP}\big[X(t') = X(t) \forall t' \in [t, t + f_a(t, \theta)] | \boldsymbol{\tau}\big] = \Pi_{1,a}(\theta), \quad (3.33)$$

where the function f_a is given by

$$f_a(t, \theta) = \theta t^{\gamma(1-a)} L(t)^{1-a}. \quad (3.34)$$

Here we use γ to denote the subaging exponent, $\gamma = (1 + \alpha)^{-1}$, *and $L(t)$ is a slowly varying function that is determined only by the distribution of τ_0. The function $\Pi_{1,a}(\theta)$ can again be written using the singular diffusion Z,*

$$\Pi_{1,a}(\theta) = \int_0^\infty g_a^2(\theta u^{a-1}) \mathrm{d}F(u), \quad (3.35)$$

where $F(u) = \mathbb{EP}[\rho(\{Z(1)\}) \leq u | \rho]$, and where $g_a(\lambda)$ is the Laplace transform of the random variable $\nu_a \tau_0^a$,

$$g_a(\lambda) = \mathbb{E}(e^{-\lambda \nu_a \tau_0^a}). \quad (3.36)$$

If $a = 0$, (3.35) can be written as

$$\Pi_{1,0}(\theta) = \int_0^\infty e^{-\theta/u} \mathrm{d}F(u). \quad (3.37)$$

Remark. 1. As can be seen, in this case the function $\Pi_{1,a}(\theta)$ depends on a. This is not surprising since the compensation by attraction has no influence here and the jump rates clearly depend on a.

2. Of course, an analogous result to Theorem 3.12 holds for quenched subaging in distribution.

In the RHT case, $a = 0$, the proof of Theorem 3.13 is straightforward. It follows from the convergence of the distribution of $\tau^\varepsilon_{X^\varepsilon(1)}$ to the distribution of $\rho(\{Z(1)\})$ as stated in Theorem 3.4. This then implies the convergence of the distribution of $\tau_{X(t_w)}/(t_w^\gamma L(t_w))$.

The proof in the case $a > 0$ is more complicated. It essentially involves a control of the distribution of the depth of those traps that are nearest neighbours of the traps where $X(t_w)$ is with a large probability. It turns out that this distribution converges to the the distribution of τ_0, i.e. there is nothing special on the neighbours of deep traps.

The behaviour of the two-point functions $\Pi(t_w, t+t_w)$ and $R(t_w, t+t_w)$ is not difficult to understand and guess. We give here a heuristic explanation for these results, first in the case $a = 0$. After the first n jumps the process typically visits $O(n^{1/2})$ sites. The deepest trap that it finds during n jumps has therefore a depth of order $O(n^{1/2\alpha})$, which is the order of the maximum of $n^{1/2}$ heavy-tailed random variables τ_x as can be verified from Assumption 2.3. This trap is typically visited $O(n^{1/2})$ times. Since the depths are in the domain of attraction of an α-stable law with $\alpha < 1$, the time needed for n jumps is essentially determined by the time spent in the deepest trap. This time is therefore $O(n^{(1+\alpha)/2\alpha})$. Inverting this expression we get that the process visits typically $O(t^{\alpha\gamma})$ sites before time t. The deepest traps it finds during this time have a depth of order t^γ. Moreover, the process is located in one of these deep traps at time t. From this, one sees that the main contribution to quantity $R(t_w, t_w + t)$ comes from the trajectories of X that, between times t_w and $t_w + t$, leave the original site $X(t_w)$ a number of times of order $t_w^{\alpha\gamma}$, and then return to it. Each visit of the original site lasts an amount of time of order t_w^γ, which is the time scale on which Π ages.

For $a > 0$, such heuristics is not directly accessible. However, the Theorem 3.4 yields that $\tau_{X(t_w)}$ is of the same order, $O(t_w^\gamma)$, as in the RHT case. Each visit of $X(t_w)$ lasts a shorter time, $O(t_w^{\gamma(1-a)})$, as follows from the definition of the process and the fact that the depths of the neighbours of $X(t_w)$ are $O(1)$. On the other hand, the process makes more excursions from $X(t_w)$, their number being $O(t_w^{\gamma(\alpha+a)})$.

3.5 Behaviour of the aging functions on different time scales

Having found two interesting time scales $t = O(t_w)$ and $t = O(t_w^\gamma)$ in the model[1] one may ask if there are other interesting time scales for the functions Π and R. This question was raised by Bouchaud and Bertin in [BB03]. The negative answer was given in [Čer06]:

Theorem 3.14 (Behaviour of Π on different scales). *(a) Short time scales. Let $f(t)$ be an increasing function satisfying $t^\kappa \geq f(t) \geq t^\mu$ for all t large and for some $\gamma > \kappa \geq \mu > 0$. Then*

$$\lim_{t \to \infty} \left(\frac{f(t)}{t^\gamma} \right)^{\alpha-1} \big(1 - \Pi(t, t + f(t)) \big) = K_1, \qquad (3.38)$$

with $0 < K_1 < \infty$.

---

[1]We suppose here that $a = 0$ and Assumption 2.3 holds, that is $L(t) \to 1$.

(b) Long time scales. *Let $g(t)$ be such that $t^\gamma = o(g(t))$. Then*

$$\lim_{t\to\infty} \left(\frac{g(t)}{t^\gamma}\right)^\alpha \Pi\big(t, t + g(t)\big) = K_2, \tag{3.39}$$

with $0 < K_2 < \infty$.

(c) Behaviour of $\Pi_{1,0}(\theta)$. *The function $\Pi(\theta)$ satisfies*

$$\lim_{\theta\to 0} \theta^{\alpha-1}(1 - \Pi_{1,0}(\theta)) = K_1, \tag{3.40}$$

$$\lim_{\theta\to\infty} \theta^\alpha \Pi_{1,0}(\theta) = K_2. \tag{3.41}$$

Remark. We emphasise that the constant K_1 occur both in (3.38) and (3.40). That means that the behaviour of $\Pi_{1,0}(\theta)$ at $\theta \sim 0$ gives also the behaviour of $\Pi(t, t + f(t))$ for $f(t) \ll t^\gamma$. An analogous remark applies for time scales $f(t) \gg t^\gamma$. Both constants K_1 and K_2 can again be expressed using the F.I.N. diffusion. The formulae, and also similar (slightly weaker) results for the two-time function R can be found in [Čer06].

We give again a heuristic description of this result. As we know, at time t the process X is typically in a trap of depth $O(t^\gamma)$, it needs a time of the same order to jump out. In Theorem 3.14(a) we look at $1 - \Pi(t, t + f(t))$ with $f(t) \ll t^\gamma$, that is at the probability that a jump occurs in a time much shorter than t^γ. There are essentially two possible extreme strategies which lead to such an event:

1. $\tau_{X(t)}$ has the typical order t^γ but the jump occurs in an exceptionally short time.

2. $X(t)$ is in an atypically shallow trap and stays there a typical time.

In [Čer06] it is proved that the second strategy dominates. Therefore, one has to study the probability of being in a very shallow trap or, equivalently, to describe the tail of $\mathbb{P}[\tau_{X(t)}/t^\gamma \leq u]$ for $u \sim 0$. To control this tail the proof use the fact that although the BTM never reaches equilibrium in a finite time in infinite volume, it is nearby equilibrium if one observes only traps that are much shallower than the typical depth t^γ on intervals that are small with respect to the typical size of $X(t)$. This puts on a rigorous basis, at least in dimension one, the concept of *local equilibrium* that was introduced in the physics literature by [RMB00]. The concept does not give the right predictions for the values of the limiting functions $R_1(\theta)$ and $\Pi_{1,0}(\theta)$ but it is useful to describe their asymptotic behaviour. A very similar heuristics applies also for time scales $f(t) \gg t^\gamma$.

3.6 References

In the RHT case, $a = 0$, this model has first time been studied by Fontes, Isopi and Newman in [FIN99]. It was used there as a tool to control the behaviour of the voter model with random rates. It was proved there that the process is sub-diffusive, i.e. $\mathbb{E}[X(t)/\sqrt{t}] \to 0$ as $t \to \infty$, and that the dynamics localises in the sense that (for a.e. $\boldsymbol{\tau}$)

$$\sup_{x\in\mathbb{Z}} \mathbb{P}[X(t) = x|\boldsymbol{\tau}] \nrightarrow 0 \qquad \text{as } t \to \infty. \tag{3.42}$$

That means that there will be always a site (dependent on time t) where the process X can be found with a non-negligible probability. This localisation occurs only in one-dimensional BTM and is at the heart of the majority of the differences between $\mathbb{Z}$ and other graphs.

The scaling limit has been established for the case $a = 0$ in [FIN02] and by [Čer03, BČ05] for $a \neq 0$. Following [BB03], the results of Section 3.5 have been given by [Čer06].

For aging in another interesting one-dimensional dynamics, i.e. Sinai's Random Walk, see [DGZ01, BF06]. Let us mention two interesting open questions on this one-dimensional trap model:

(a) What is the behaviour of the edge of the spectrum for the generator of the dynamics. This might be close to, but easier than the same question solved for Sinai's Random Walk by [BF06].

(b) What is the influence of a drift in the BTM? Monthus [Mon04] gives a very interesting picture based on renormalisation arguments.

4 The trap model in dimension larger than one

After resolving the BTM on $\mathbb{Z}$, the next natural step is to study the Bouchaud model on the d-dimensional lattice, $\mathsf{BTM}(\mathbb{Z}^d, \boldsymbol{\tau}, 0)$, $d > 1$. Observe that we set $a = 0$, that means that only the RHT dynamics is considered. In this section we always assume that Assumption 2.3 holds.

4.1 The fractional-kinetics process

As for the one-dimensional model, we first identify a scaling limit of the BTM on $\mathbb{Z}^d$. The result of this section is contained in the forthcoming paper [BČ06b]. We will, from now on, use frequently the theory of Lévy processes and subordinators. A very short summary of this theory can be found in Appendix A.

Let us first define the process that appears as the scaling limit.

Definition 4.1 (Fractional kinetics)**.** Let $B_d(t)$ be the standard d-dimensional Brownian motion started at 0 and let V be the α-stable subordinator given by its Laplace transform $\mathbb{E}[e^{-\lambda V(t)}] = e^{-t\lambda^\alpha}$. Let $T(s) = \inf\{t : V(t) > s\}$ be the inverse of $V(t)$. We define the *fractional-kinetics process* Ψ_d by

$$\Psi_d(s) = B_d(T(s)). \tag{4.1}$$

We list here without proofs several properties of the process Ψ_d.

Proposition 4.2. *1. The stable subordinator V is strictly increasing, therefore its inverse T and also the process Ψ_d are continuous.*

2. The name of the process is due to the following fact. Let $p(t, \cdot)$ be the probability density of $\Psi_d(t)$. Then p is a solution of the fractional kinetic equation,

$$\frac{\partial^\alpha}{\partial t^\alpha} p(t, x) = \frac{1}{2} \Delta p(t, x) + \delta(0) \frac{t^{-\alpha}}{\Gamma(1 - \alpha)}. \tag{4.2}$$

Here, the fractional derivative $\partial^\alpha p(t,x)/\partial t^\alpha$ is the inverse Laplace transform of $s^\alpha \tilde{p}(s,x)$, where $\tilde{p}(s,x) = \int_0^\infty e^{-st} p(t,x)\,\mathrm{d}t$ is the usual Laplace transform. The equation (4.2) should be understood in the weak sense, i.e. it holds after the integration against smooth test functions.

    3. The process Ψ_d is not Markov, as can be seen easily from the previous point.

    4. The fixed-time distribution of Ψ_d is the Mittag-Leffler distribution,

$$\mathbb{E}\big(e^{i\xi\cdot\Psi_d(t)}\big) = E_\alpha(-|\xi|^2 t^\alpha), \tag{4.3}$$

where $E_\alpha(z) = \sum_{m=0}^\infty z^m/\Gamma(1+m\alpha)$.

    5. The process Ψ_d is self-similar:

$$\Psi_d(t) \stackrel{\text{law}}{=} \lambda^{-\alpha/2}\Psi(\lambda t). \tag{4.4}$$

The process Ψ_d is well known in the physics literature (see [Zas02] for a broad survey and earlier references). It is the scaling limit of a very classical object, a Continuous Time Random Walk (CTRW) introduced by [MW65]. More precisely consider a simple random walk Y on $\mathbb{Z}^d$ and a sequence of positive i.i.d. random variables $\{s_i : i \in \mathbb{N}\}$ with the distribution in the domain of attraction of an α-stable law. Define the CTRW $U(t)$ by

$$U(t) = Y(k) \qquad \text{if } t \in \Big[\sum_{i=1}^{k-1} s_i, \sum_{i=1}^k s_i\Big). \tag{4.5}$$

It is proved in [MS04] that there is a constant C (depending only on the distribution of s_i) such that

$$Cn^{-\alpha/2}U(tn) \xrightarrow{n\to\infty} \Psi_d(t). \tag{4.6}$$

4.2   Scaling limit

Observe that the Bouchaud model (for $a = 0$) can be expressed as a time change of the simple random walk. The time-change process is crucial for us:

Definition 4.3. Let $S(0) = 0$ and let $S(k)$, $k \in \mathbb{N}$, be the time of the kth jump of X. For $s \in \mathbb{R}$ we define $S(s) = S(\lfloor s \rfloor)$. We call $S(s)$ the *clock process*. Obviously, $X(t) = Y(k)$ for all $S(k) \le t < S(k+1)$.

The following result shows that the limit of the d-dimensional Bouchaud model and its clock process on $\mathbb{Z}^d$ ($d \ge 2$) is trivial, in the sense that it is identical with the scaling limit of the much simpler ("completely annealed") dynamics of the CTRW.

Theorem 4.4 (Scaling limit of BTM on $\mathbb{Z}^d$). *Let*

$$f(n) = \begin{cases} n^{\alpha/2}(\log n)^{(1-\alpha)/2}, & \text{if } d = 2, \\ n^{\alpha/2}, & \text{if } d \ge 3. \end{cases} \tag{4.7}$$

Then for all $d \geq 2$ and for a.e. $\boldsymbol{\tau}$,

$$\frac{C_d(\alpha)X(nt)}{f(n)} \xrightarrow{n\to\infty} \Psi_d(t) \quad and \quad \frac{S\big(C_d(\alpha)^{-2}f(n)^2 s\big)}{n} \xrightarrow{n\to\infty} V(s). \tag{4.8}$$

weakly in the Skorokhod topology on $D([0,T],\mathbb{R}^d)$ (the space of cadlag functions from $[0,T]$ to $\mathbb{R}^d$). If $G_d(0)$ denotes Green's function of the d-dimensional random walk at 0, then

$$C_d(\alpha) = \begin{cases} \sqrt{2\pi^{1-\alpha}\Gamma(1-\alpha)\Gamma(1+\alpha)}, & if \ d = 2, \\ \sqrt{dG_d(0)^{\alpha}\Gamma(1-\alpha)\Gamma(1+\alpha)}, & if \ d \geq 3. \end{cases} \tag{4.9}$$

The main ideas of the proof of this theorem will be explained in Section 4.4. At this place, let us only compare the fractional-kinetics process Ψ_d with the F.I.N. diffusion Z. Both these processes are defined as a time change of the Brownian motion $B_d(t)$. The clock processes however differ considerably. For $d = 1$, the clock equals $\phi(t) = \int \ell(t,x)\rho(\mathrm{d}x)$, where ρ is the random speed measure obtained as the scaling limit of the environment. Moreover, since ℓ is the local time of the Brownian motion B_1, the processes B_1 and ϕ are *dependent*. For $d \geq 2$, the Brownian motion B_d and the clock process, i.e. the stable subordinator V, are *independent*. The asymptotic independence of the clock process S and the location Y of the BTM is a very remarkable feature distinguishing $d \geq 2$ and $d = 1$. It explains the "triviality" of the scaling limit in dimension $d \geq 2$, but is, by no means, trivial matter to prove. We will come back to an intuitive explanation of the independence in Section 5.3. Note also that nothing like a scaling limit of the random environment appears in the definition of Ψ_d, moreover, the convergence holds $\boldsymbol{\tau}$-a.s. The absence of the scaling limit of the environment in the definition of Ψ_d transforms into the non-Markovianity of Ψ_d. Note however that it is considerably easier to control the behaviour of Ψ_d than of Z even if Ψ_d is not Markov: many quantities related to Ψ_d can be computed explicitly, as can be seen from Proposition 4.2.

4.3 Aging results

The following two theorems describe the aging behaviour of the two-time functions R and Π.

Theorem 4.5 (Quenched aging on $\mathbb{Z}^d$)**.** *For all $\alpha \in (0,1)$ and $d \geq 2$ there exists a deterministic function $R(\theta)$ independent of d (but dependent on α) such that for $\mathbb{P}$-a.e. realisation of the random environment $\boldsymbol{\tau}$*

$$\lim_{t_w \to \infty} R(t_w, t_w + \theta t_w; \boldsymbol{\tau}) = R(\theta). \tag{4.10}$$

The function $R(\cdot)$ can be written explicitly: Let $\mathsf{Asl}_\alpha(u)$ be the distribution function of the generalised arcsine law with parameter α,

$$\mathsf{Asl}_\alpha(u) := \frac{\sin \alpha \pi}{\pi} \int_0^u u^{\alpha-1}(1-u)^{-\alpha}\,\mathrm{d}u. \tag{4.11}$$

Then $R(\theta) = \mathsf{Asl}_\alpha(1/1+\theta)$.

Theorem 4.6 (Quenched (sub-)aging on $\mathbb{Z}^d$)**.** *For all $\alpha \in (0,1)$ and $d \geq 2$ there exists a deterministic function $\Pi_d(\theta)$ such that for $\mathbb{P}$-a.e. realisation of the random environment $\boldsymbol{\tau}$*

$$\lim_{t_w \to \infty} \Pi\big(t_w, t_w + \theta f(t_w); \boldsymbol{\tau}\big) = \Pi_d(\theta), \tag{4.12}$$

with

$$f(t_w) = \begin{cases} \frac{t_w}{\log t_w}, & \text{if } d = 2, \\ t_w, & \text{if } d \geq 3. \end{cases} \tag{4.13}$$

The function Π_d does depend on d. Moreover, for all $\theta > 0$ it satisfies

$$\lim_{d \to \infty} \Pi_d(\theta) = R(\theta). \tag{4.14}$$

Remark. 1. Both theorems are proved in [BČM06] for $d = 2$ and in [Čer03] for $d \geq 3$. The proofs are relatively technical and exceed the scope of these notes. The main ideas however do not use specific properties of the integer lattice $\mathbb{Z}^d$ and can be generalised to different graphs. These ideas will be explained in Section 5. The other important part of the proof, that is the coarse-graining of the trajectory of the process, is explained in the next subsection.

2. The function Π_d can also be explicitly calculated but the formula is tedious [BČM06, Čer03].

3. We will see later that the function R is closely related to the arcsine law for Lévy processes. As we will also see, the same function appears as the limit in the case of the BTM on a large complete graph. Therefore, the mean-field dimension of the BTM for the two-time function R is $d = 2$.

4. Both presented results are *quenched*, i.e. they hold for a.e. $\boldsymbol{\tau}$. This should be compared with the results for the $d = 1$ case (Theorems 3.11, 3.12 and 3.13) where only averaged aging holds. The reason for this difference is explained in Section 5.3 below. It is, of course, trivial to get averaged results from Theorems 4.5 and 4.6: the dominated convergence theorem yields

$$\begin{aligned} &\lim_{t_w \to \infty} R(t_w, t_w + \theta t_w) = R(\theta), \\ &\lim_{t_w \to \infty} \Pi\big(t_w, t_w + \theta f(t_w)\big) = \Pi(\theta). \end{aligned} \tag{4.15}$$

5. The fact that in $d \geq 3$ the time scale $f(t_w) = t_w$ is the same for both functions R and Π is a consequence of the transience of the simple random walk. The traps are visited only a finite number of times, so that different time scales cannot appear for R and Π.

4.4 The coarse-graining procedure

We would like to describe here the coarse-graining procedure which was introduced in [BČM06], and which is the main tool in proving Theorems 4.4–4.6. Even this short sketch of the procedure might be considered technical and can be skipped on a first

reading. The reader can also decide to return here after being acquainted with the general ideas of Sections 5.2 and 5.3. We will deal here only with the convergence of the clock process to an α-stable subordinator. For the sake of concreteness we set $d = 2$, however the same arguments apply also for $d > 2$. The discussion in this section is valid *for a.e. realisation of the random environment $\boldsymbol{\tau}$.*

The random environment is heavy tailed. Therefore, as for $d = 1$, the behaviour of the clock process $S(k)$ is determined by the time spent in the deepest traps that the process visits during the first k steps. It is thus necessary to find the depth scale of these traps and then study how these traps contribute to the clock process.

The coarse-graining procedure of [BČM06] studies the process Y (resp. X) only before the exit from a large disk $\mathbb{D}(n)$, $n \in \mathbb{N}$, of area $m2^n n^{1-\alpha}$ centred at the origin. We denote by $R(n)$ the radius of this disk, $R_n = \sqrt{\pi^{-1} 2^n n^{1-\alpha}}$. The random walk Y makes $O(R_n^2)$ steps in $\mathbb{D}(n)$ and visits $O(R_n^2 / \log R_n^2)$ different traps, as is well known. Therefore, the deepest traps that it visits have a depth of order $O((R_n^2 / \log R_n^2)^{1/\alpha}) = O(2^{n/\alpha} n^{-1})$. We therefore approximate the clock process by the time that X spends in the set

$$T_\varepsilon^M = \{x \in \mathbb{D}(n) : \varepsilon 2^{n/\alpha} n^{-1} \leq \tau_x < M 2^{n/\alpha} n^{-1}\}, \tag{4.16}$$

where ε is a small and M a large constant. We call the traps in this set the *deep traps*. It must be proved that this approximation is correct. We do not want, however, to deal with this problem here (see [BČM06]).

After giving the proper depth scale, we can study how the deep traps are visited. This is where the coarse-graining is crucial. We cut the trajectory of the process Y into short pieces. Every such piece of the trajectory ends when Y exits for the first time the disk of area $2^n n^\gamma$ around its initial point. At this moment a new piece starts. Clearly, we should take $\gamma < 1 - \alpha$. Formally, we set $j_0^n = 0$, and then we define recursively

$$j_i^n = \min\left\{k > j_{i-1}^n : \operatorname{dist}(Y(k), Y(j_{i-1}^n)) \geq \sqrt{\pi^{-1} 2^n n^\gamma}\right\}. \tag{4.17}$$

We use x_i^n to denote the starting points of the pieces of the trajectory, $x_i^n = Y(j_i^n)$. It can be seen easily that the number of pieces of the trajectory before the exit from $\mathbb{D}(n)$ is of order $O(n^{1-\alpha-\gamma})$.

We look at the time that the walk spends in the deep traps during one piece of trajectory: we define the *score* of the piece i by

$$s_i^n = \sum_{k=j_i^n}^{j_{i+1}^n - 1} e_k \tau_{Y(k)} \mathbb{1}\{Y(k) \in T_\varepsilon^M\}. \tag{4.18}$$

To study the scores it is convenient to introduce another family of random variables s_x^n indexed by $x \in \mathbb{D}(n)$. We set the distribution of s_x^n to be the same as the distribution of s_i^n conditioned on the fact that the ith piece of the trajectory starts at x, i.e. conditioned on $x_i^n = x$. The main technical piece of the proof is to show that the law of s_x^n does not depend on x (with a small exceptional set):

24

Lemma 4.7. *Fix $\kappa > 0$ large enough and define*

$$\mathcal{E}(n) = \{x \in \mathbb{D}(n) : \operatorname{dist}(x, T_\varepsilon^M) \geq \sqrt{\pi^{-1} 2^n n^{-\kappa}}\}. \tag{4.19}$$

Then for $\mathbb{P}$-a.e. random environment $\boldsymbol{\tau}$, uniformly for $x \in \mathcal{E}(n)$

$$\lim_{n \to \infty} \frac{1 - \mathbb{E}[\exp(-\frac{\lambda s_x^n}{2^{n/\alpha}}) | s_x < \infty, \boldsymbol{\tau}]}{n^{\alpha + \gamma - 1}} = F(\lambda). \tag{4.20}$$

Here

$$F(\lambda) = F(\lambda; \varepsilon, M, \alpha) = \mathcal{K}\left(p_\varepsilon^M - \int_\varepsilon^M \frac{\alpha}{1 + \mathcal{K}'\lambda z} \cdot \frac{1}{z^{\alpha + 1}} \, \mathrm{d}z\right) \tag{4.21}$$

with $\mathcal{K}' = \pi^{-1}\log 2$ and $\mathcal{K} = (\log 2)^{-1}$.

This lemma is a consequence of the following four facts, whose proofs are based on classical sharp estimates on the Green's function for the simple random walk on $\mathbb{Z}^d$ and on certain "homogeneity" properties of the random environment. These proofs can be found in [BČM06].

1. s_x^n is equal to 0 with probability $1 - \mathcal{K}p_\varepsilon^M n^{\alpha + \gamma - 1}(1 + o(1))$, where $p_\varepsilon^M = \varepsilon^{-\alpha} - M^{-\alpha}$. That means that typically no deep trap is visited in a piece of the trajectory. Further, it implies that only a finite number of pieces has a non-zero score before the exit from $\mathbb{D}(n)$.

2. With probability $\mathcal{K}p_\varepsilon^M n^{\alpha + \gamma - 1}(1 + o(1))$ the random walk visits (many times) only one deep trap during one piece, call it y. The probability that two or more deep traps are visited during one piece is $O(n^{2(\alpha + \gamma - 1)})$, therefore, with overwhelming probability, this event does not occur before the exit from $\mathbb{D}(n)$.

3. In the case when one deep trap y is visited, the distribution of its normalised depth $2^{-n/\alpha} n \tau_y$ converges to the distribution on $[\varepsilon, M]$ with the density $p(u)$ proportional to $u^{-\alpha - 1}$.

4. The number of visits to y is geometrically distributed with mean $\mathcal{K}'n(1 + o(1))$. Therefore, conditionally on hitting y the score has an exponential distribution with mean $\mathcal{K}'n\tau_y$ which is of order $O(2^{n/\alpha})$.

Using Lemma 4.7, it is not difficult to check that the scores s_i^n are asymptotically i.i.d. Actually, to transfer the uniform convergence of distributions of s_x^n into the asymptotic i.i.d. property of s_i^n it is sufficient to check that with an overwhelming probability all pieces of trajectory before the exit of $\mathbb{D}(n)$ do start in $\mathcal{E}(n)$.

The asymptotic i.i.d. property then yields the convergence of the normalised sum of scores, $2^{-n/\alpha} \sum_{i=0}^{tn^{1-\alpha-\gamma}} s_i^n$, to a Lévy process. The Lévy measure of this process can be computed from Lemma 4.7. Using the knowledge of the Lévy measure, one can then prove that for any $T > 0$ it is possible to choose m large enough such that as $\varepsilon \to 0$ and $M \to \infty$ the distribution of this Lévy process on $[0, T]$ approaches the distribution of an α-stable subordinator on $[0, T]$. The convergence of the clock process to the same subordinator then follows since the sum of scores is a good approximation of the clock process.

4.5 References

The trap model in dimension $d \geq 2$ is studied in [MB96, RMB01]. In the case $a = 0$, a mathematical proof of aging has been given in [Čer03, BČM06]. This proof is based on the coarse-graining of the trajectories of the BTM sketched in the last section. In [BČ06a] we establish the fractional-kinetics scaling limit based on the arguments of [BČM06]. The case where $a \neq 0$ is discussed in [RMB01], but is still an open problem.

Let us mention that the clock process introduced in Definition 4.3 is close to the problem of Random Walk in Random Scenery (RWRS), except for the extra randomisation due to exponential waiting times. However, here the tails of the scenery distribution are heavier than in the recent works on RWRS.

In dimension $d = 1$, RWRS with heavy tails have been studied by Kesten and Spitzer [KS79]. The F.I.N. diffusion could indeed be seen as a Brownian motion time-changed by a (dependent) Kesten-Spitzer clock.

5 The arcsine law as a universal aging scheme

In this section we explain a general strategy that can be used to prove aging of the functions Π and R in the BTM on many different graphs (including $\mathbb{Z}^d$ for $d \geq 2$, tori in $\mathbb{Z}^d$, large complete graphs, and high-dimensional hypercubes). When this strategy can be used, the behaviour of $R(t_w, t_w + \theta t_w; \boldsymbol{\tau})$ for large t_w can be expressed using the distribution function of the generalised arcsine distribution with parameter α, $\mathsf{Asl}_\alpha(\cdot)$.

More formally, we will consider in this section a sequence of the Bouchaud trap models $\{\mathsf{BTM}(G_n, \boldsymbol{\tau}_n, 0) : n \in \mathbb{N}\}$. We will prove that for properly chosen time scales $t_w(n)$

$$\lim_{n \to \infty} R_n(t_w(n), (1 + \theta)t_w(n); \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1 + \theta), \tag{5.1}$$

where

$$R_n(t_w, t_w + t; \boldsymbol{\tau}) := \mathbb{P}[X_n(t_w) = X_n(t_w + t)|\boldsymbol{\tau}]. \tag{5.2}$$

The possible time scales $t_w(n)$ will depend on the graphs G_n and the laws of the depths $\boldsymbol{\tau}_n$. We will always try to prove aging results on the widest possible range of time scales.

We will consider only the case of RHT dynamics, i.e. $a = 0$. So that, the embedded discrete-time process Y_n is the simple random walk on G_n, and the continuous time Markov chain X_n is a time change of Y. Our strategy relies on an approximation of the clock process (see Definition 4.3), more precisely of its rescaling, by an α-stable subordinator. We will then show that the event $X_n(t_w) = X_n((1+\theta)t_w)$ is approximated by the event that the subordinator jumps over the interval $[1, 1+\theta]$. The classical arcsine law for Lévy processes (see Proposition A.4 in the Appendix) will then imply the aging result (5.2).

5.1 Aging on large complete graphs

As a warm-up, and in order to explain our strategy, we give here a complete proof of aging on a large complete graph in this sub-section. The advantage of this graph is that

the embedded simple random walk is particularly simple. On the other hand, almost all effects of the trapping landscape are already present. The method of the proof that we use here is probably not the simplest one. Its main ideas can however be adapted to more complex graphs; the structure of the proof stays the same, but a relatively fine control of the simple random walk on these graphs is then required.

Note that the complete graph is the graph that was proposed in the original paper of Bouchaud [Bou92]. The aging on this graph is proved using renewal arguments in [BD95], see also the introduction to [BBG03b]. Another proof of aging and much more can be found in [BF05] where eigenvalues and eigenvectors of the generator of X are very carefully analysed.

Let G_n be a complete graph with n vertices, $G_n = (\mathcal{V}_n, \mathcal{E}_n)$, where $\mathcal{V}_n = \{1, \ldots, n\}$ and $\mathcal{E}_n = \langle x, y \rangle, x, y \in \mathcal{V}_n$. Note that we include loops $\langle x, x \rangle$ into the graph, so that jumps from x to x are possible. This makes the embedded random walk of the RHT dynamics extremely simple: the positions $Y_n(i)$, $i \in \mathbb{N}$, are i.i.d. uniform random variables on $\mathcal{V}_n$. We also suppose that the starting position of the process, $X_n(0) = Y_n(0)$, is uniformly distributed on $\mathcal{V}_n$. Finally, we assume that the trapping landscape is given by an i.i.d. sequence $\{\tau_x : x \in \mathbb{N}\}$ independent of n such that Assumption 2.3 holds.

As the process X_n is a time change of the i.i.d. sequence $Y_n(i)$, we need to study the clock process $S_n(t)$. Observe that $S_n(k)$ for $k \in \mathbb{N}$ is the time of kth jump of X_n,

$$S_n(k) = \sum_{i=0}^{k-1} \tau_{Y_n(i)} e_i, \tag{5.3}$$

where $\{e_i\}$ is an i.i.d. sequence of exponential random variables with mean one.

Since we have included the loops into the graph, we redefine the two-time function Π slightly:

$$\Pi_n(t_w, t_w + t; \boldsymbol{\tau}) := \mathbb{P}\big[\{S_n(k) : k \in \mathbb{N}\} \cap [t_w, t_w + t] = \emptyset \big| \boldsymbol{\tau}\big]. \tag{5.4}$$

Note that this definition differs from the original one only if a jump from $X(t_w)$ to $X(t_w)$ occurs. The probability of this event is $1/n$ and is thus negligible for large n. Similarly, for the complete graph the function R_n differs very little from the function Π_n, a correction is again of order $1/n$.

We prove aging on time scales smaller than $n^{1/\alpha}$. The scale $n^{1/\alpha}$ appears because the deepest trap in $\mathcal{V}_n$ has a depth of this order as can be easily verified from Assumption 2.3. Therefore, at time scales shorter than $n^{1/\alpha}$ the process has not enough time to reach the equilibrium, and aging can be observed:

Theorem 5.1 (Quenched aging on the complete graph). *Let $0 < \kappa < 1/\alpha$ and let $t_w = t_w(n) = n^\kappa$. Then for a.e. random environment $\boldsymbol{\tau}$*

$$\lim_{n \to \infty} \Pi_n(t_w(n), (1 + \theta)t_w(n); \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1 + \theta). \tag{5.5}$$

Moreover, the clock process converges to the stable subordinator V with the Lévy measure $\alpha\Gamma(1 + \alpha)u^{-\alpha-1}\mathrm{d}u$,

$$\frac{S_n(sn^\kappa)}{n^{\kappa/\alpha}} \xrightarrow{n \to \infty} V(s). \tag{5.6}$$

Remark. 1. A similar result holds on the shortest possible time scale $\kappa = 0$: if $t_w(n) = tn^0 = t$, then

$$\lim_{t \to \infty} \lim_{n \to \infty} \Pi_n(t_w, (1 + \theta)t_w; \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1 + \theta). \tag{5.7}$$

This is actually proved in [BD95] and [BF05].

2. For the longest possible time scales $t_w(n) = tn^{1/\alpha}$ a double limiting procedure is also necessary. Essentially the same arguments that are used to prove Theorem 5.1 yield that $\Pi_n(t_w(n), (1 + \theta)t_w(n); \boldsymbol{\tau})$ converges to $\mathsf{Asl}_\alpha(1/1 + \theta)$ in probability, that is for any $\varepsilon > 0$

$$\lim_{t \to 0} \lim_{n \to \infty} \mathbb{P}\left[\left|\Pi(tn^{1/\alpha}, (1 + \theta)tn^{1/\alpha}; \boldsymbol{\tau}) - \mathsf{Asl}_\alpha(1/1 + \theta)\right| \geq \varepsilon\right] = 0. \tag{5.8}$$

3. No scaling limit for X exists since there is no such thing as a scaling limit for the complete graph.

To prove Theorem 5.1 we consider a rescaling of the time change process,

$$\mathcal{S}_n(t) = \frac{1}{n^\kappa} S_n(k) \qquad \text{for } t \in [u_n(k), u_n(k + 1)), \tag{5.9}$$

where $u_n(k) = n^{-\kappa\alpha} \sum_{i=1}^{k} e_i'$, and $\{e_i', i \in \mathbb{N}\}$ is another sequence of mean one i.i.d. exponential random variables. We introduce this sequence in order to randomise the times of jumps of $\mathcal{S}_n$. Thanks to this randomisation $\mathcal{S}_n$ is the Lévy process. The standard rescaling of S_n, $n^{-\kappa} S_n(tn^{\kappa\alpha})$ does not have this property. $\mathcal{S}_n$ is compound Poisson process. The intensity of its jumps is $n^{\alpha\kappa}$. Every jump has the same distribution as $n^{-\kappa}\tau_x e_i$, where x is uniformly distributed in $\mathcal{V}_n$. Therefore, the Lévy measure μ_n of $\mathcal{S}_n$ is

$$\mu_n(\mathrm{d}u) = n^{\alpha\kappa} \frac{1}{n} \sum_{x \in \mathcal{V}_n} \frac{n^\kappa e^{-un^\kappa/\tau_x}}{\tau_x} \, \mathrm{d}u = n^{\alpha\kappa+\kappa-1} \sum_{x \in \mathcal{V}_n} \frac{e^{-un^\kappa/\tau_x}}{\tau_x} \, \mathrm{d}u. \tag{5.10}$$

It follows from definitions of Π_n and $\mathcal{S}_n$ that

$$\Pi_n(n^\kappa, (1 + \theta)n^\kappa; \boldsymbol{\tau}) = \mathbb{P}\left[\{\mathcal{S}_n(t) : t \in \mathbb{R}\} \cap [1, 1 + \theta] = \emptyset\right], \tag{5.11}$$

that is the probability that $\mathcal{S}_n$ jumps over the interval $[1, 1 + \theta]$.

The idea of the proof is the following. We write $\mathcal{S}_n$ as a sum of three independent Lévy processes, $\mathcal{S}_{n,M} = \mathcal{S}_{n,M}^\infty$, $\mathcal{S}_{n,\varepsilon}^M$ and $\mathcal{S}_n^\varepsilon = \mathcal{S}_{n,0}^\varepsilon$ with the Lévy measures $\mu_{n,M} = \mu_{n,M}^\infty$, $\mu_{n,\varepsilon}^M$ and $\mu_n^\varepsilon = \mu_{n,0}^\varepsilon$, where

$$\mu_{n,a}^b(\mathrm{d}u) = n^{\alpha\kappa+\kappa-1} \sum_{x \in \mathcal{V}_n} \mathbb{1}\{x \in T_a^b(n)\} \frac{1}{\tau_x} e^{-un^\kappa/\tau_x} \, \mathrm{d}u, \tag{5.12}$$

and

$$T_a^b(n) = \left\{x \in \mathcal{V}_n : \frac{\tau_x}{n^\kappa} \in [a, b)\right\} \tag{5.13}$$

That is we divide traps into three categories: (a) *the very deep traps, $x \in T_M(n)$,* i.e. $\tau_x \geq Mn^\kappa$, (b) *the deep traps, $x \in T_\varepsilon^M(n)$,* i.e. $\tau_x/n^\kappa \in [\varepsilon, M)$, (c) *the shallow traps,*

$x \in T^\varepsilon$, i.e. $\tau_x < \varepsilon n^\kappa$. We consider the contributions of these different categories to the clock process separately.

We show that the contribution of the deep traps, i.e. the process $\mathcal{S}_{n,\varepsilon}^M$, is well approximated by an α-stable subordinator, at least if ε is small and M large enough. The proof of this fact relies on the weak convergence of the Lévy measures $\mu_{n,\varepsilon}^M$. We prove this convergence only on the interval $[0, T]$ where T is chosen such that $\mathcal{S}_{n,\varepsilon}^M$ is larger than $(1 + \theta)$ with a large probability.

Further we prove that the very deep and shallow traps can be almost neglected. More exactly, we prove that $\mathcal{S}_{n,M}(T) = 0$ with a large probability, that is $\mathcal{S}_{n,M}$ does not jump before T, and that $\mathcal{S}_n^\varepsilon(T)$ can be made small by choosing ε small enough. These fact together will imply that $\mathcal{S}_n$ is well approximated on $[0, T]$ by an α-stable subordinator, and the claim of the theorem follows from (5.11) and the arcsine law for stable subordinators, more precisely from Corollary A.5.

5.1.1 Deep traps

The following proposition describes the contribution of the deep traps to the time change, that is the process $\mathcal{S}_{n,\varepsilon}^M$.

Proposition 5.2. *The Lévy measures $\mu_{n,\varepsilon}^M$ of $\mathcal{S}_{n,\varepsilon}^M$ converge weakly as $n \to \infty$ to the measure μ_ε^M given by*

$$\mu_\varepsilon^M(\mathrm{d}u) = \int_\varepsilon^M \frac{\alpha}{z^{\alpha+2}} \, e^{-u/z} \, \mathrm{d}z \, \mathrm{d}u. \tag{5.14}$$

This proposition has an important corollary that states that $\mathcal{S}_{n,\varepsilon}^M$ can be well approximated for large n by an α-stable subordinator. To this end, let $\mathcal{Z}_\varepsilon^M$ be a subordinator with the Lévy measure

$$\tilde{\mu}_\varepsilon^M(\mathrm{d}u) = \int_0^\varepsilon \frac{\alpha}{z^{\alpha+2}} \, e^{-u/z} \, \mathrm{d}z \, \mathrm{d}u + \int_M^\infty \frac{\alpha}{z^{\alpha+2}} \, e^{-u/z} \, \mathrm{d}z \, \mathrm{d}u \tag{5.15}$$

independent of all already introduced random variables.

Corollary 5.3. *The process $\mathcal{S}_{n,\varepsilon}^M + \mathcal{Z}_\varepsilon^M$ converge weakly in the Skorokhod topology on $D([0, T], \mathbb{R})$ to an α-stable subordinator. Moreover, for any T and $\delta > 0$ it is possible to choose ε small and M large enough such that*

$$\mathbb{P}[\mathcal{Z}_\varepsilon^M(T) > \delta] < \delta. \tag{5.16}$$

Proof of Corollary 5.3. The first claim is consequence of the convergence of Lévy measures (Proposition 5.2) together with Lemma A.2. The fact that the limit is a stable subordinator follows from

$$\tilde{\mu}_\varepsilon^M(\mathrm{d}u) + \mu_\varepsilon^M(\mathrm{d}u) = \alpha\Gamma(1 + \alpha)u^{-\alpha-1} \, \mathrm{d}u. \tag{5.17}$$

Finally, to prove (5.16) observe that for all $\lambda > 0$

$$\mathbb{E}[e^{-\lambda \mathcal{Z}_\varepsilon^M(T)}] = e^{-T \int_0^\infty (1 - e^{-\lambda x}) \tilde{\mu}_\varepsilon^M(\mathrm{d}x)} \xrightarrow{\varepsilon \to 0, M \to \infty} 1, \tag{5.18}$$

that is the law of $\mathcal{Z}_\varepsilon^M(T)$ converges weakly to the Dirac mass at 0. $\qquad \square$

To prove Proposition 5.2 we first show that for a.e. $\boldsymbol{\tau}$ there is the right number of deep traps with depths approximately un^κ in $\mathcal{V}_n$, $u \in [\varepsilon, M]$. This follows from a "law-of-large-number-type" argument using the fact that we have a large number of traps with such depth. This property fails to be true if $\kappa \geq 1/\alpha$ (which explains why we cannot get a.s. result for $t_w = cn^{1/\alpha}$ and why we need the double limit procedure).

Lemma 5.4. *Let $0 < \kappa < 1/\alpha$. Then there exists a function $h(n)$, $h(n) \to 0$ as $n \to \infty$, such that*

$$\lim_{n\to\infty} \sup_{u\in[\varepsilon,M]\cap\{\mathbb{Z}/h(n)\}} \left| \frac{|T_u^{u+h(n)}(n)|}{h(n)n^{1-\alpha\kappa}} - \alpha u^{-\alpha-1} \right| = 0, \qquad \boldsymbol{\tau}\text{-a.s.} \tag{5.19}$$

Proof. Let $g(u)$ be defined by $\mathbb{P}[\tau_x \geq u] = u^{-\alpha}(1+g(u))$. It follows from Assumption 2.3 that $\lim_{u\to\infty} g(u) = 0$. Take $h(n)$ such that $h(n) \to 0$, $h(n) \geq (\log n)^{-1}$, $h(n) \gg \sup_{u\in[\varepsilon,M]} g(un^\kappa)$. Then, as $n \to \infty$,

$$\mathbb{P}\big[\tau_x \in [un^\kappa, (u+h(n))n^\kappa]\big] = \alpha u^{-1-\alpha} h(n) n^{-\alpha\kappa}(1+o(1)). \tag{5.20}$$

Indeed, by definition of g, the left-hand side is equal to

$$u^{-\alpha}n^{-\kappa\alpha}\left[1 - \left(1 + \frac{h(n)}{u}\right)^{-\alpha}\right] + u^{-\alpha}n^{-\kappa\alpha}\left[g(un^\kappa) - g((u+h(n))n^\kappa)\left(1 + \frac{h(n)}{u}\right)^{-\alpha}\right]. \tag{5.21}$$

The last expression is equal to the right-hand side of (5.20) as follows from the definitions of g and h.

Fix $u \in [\varepsilon, M]$ and $\varepsilon > 0$. Using twice the exponential Chebyshev inequality together with (5.20) we get

$$\mathbb{P}\left[\frac{|T_u^{u+h(n)}(n)|}{h(n)n^{1-\alpha\kappa}} \notin (1-\varepsilon, 1+\varepsilon)\alpha u^{-\alpha-1}\right] \leq C \exp\{-c(\varepsilon)u^{-\alpha-1}n^{1-\alpha\kappa}h(n)\}. \tag{5.22}$$

Summing over $u \in [\varepsilon, M] \cap \{\mathbb{Z}/h(n)\}$, the claim of the lemma then follows using the Borel-Cantelli Lemma and the fact that $h(n) \geq (\log n)^{-1}$. $\qquad\square$

Proof of Proposition 5.2. By definition of $\mu_{n,\varepsilon}^M$,

$$\mu_{n,\varepsilon}^M(\mathrm{d}u) = n^{\alpha\kappa+\kappa-1} \sum_{x\in T_\varepsilon^M(n)} \frac{1}{\tau_x} e^{-un^\kappa/\tau_x}\,\mathrm{d}u$$

$$= n^{\alpha\kappa+\kappa-1} \sum_{i=0}^{(M-\varepsilon)/h(n)} \sum_{x\in T_{ih(n)}^{(i+1)h(n)}} \frac{e^{-u/ih(n)}}{n^\kappa ih(n)}\,\mathrm{d}u(1+o(1)) \tag{5.23}$$

$$= n^{\alpha\kappa-1} \sum_{i=0}^{(M-\varepsilon)/h(n)} |T_{ih(n)}^{(i+1)h(n)}(n)|\frac{1}{ih(n)} e^{-u/ih(n)}\,\mathrm{d}u(1+o(1)).$$

By the previous lemma, for large n, $\boldsymbol{\tau}$-a.s

$$\mu_{n,\varepsilon}^M(\mathrm{d}u) = \sum_{i=0}^{(M-\varepsilon)/h(n)} \alpha(ih(n))^{-\alpha-1}\frac{e^{-u/ih(n)}}{ih(n)} h(n)\,\mathrm{d}u(1+o(1)), \tag{5.24}$$

30

which is the Riemann sum of the integral in (5.14) with the mesh size $h(n)$. Since $h(n) \to 0$ as $n \to \infty$, the proof is finished. $\qquad\square$

5.1.2 Shallow traps

We prove that the contribution of the shallow traps at any final instant T, $\mathcal{S}_n^\varepsilon(T)$, is small if ε is chosen small enough.

Lemma 5.5. *There exists a large constant K such that for all $T > 0$ and n large*

$$\mathbb{E}[\mathcal{S}_n^\varepsilon(T)|\boldsymbol{\tau}] < KT\varepsilon^{1-\alpha}, \qquad \boldsymbol{\tau}\text{-a.s.} \tag{5.25}$$

Proof. Let ξ be such that $T \in [u_n(\xi), u_n(\xi+1))$ (see (5.9)). It is easy to see that ξ has Poisson distribution, $\mathbb{E}[\xi] = Tn^{\kappa\alpha}$. By definition of $\mathcal{S}_n^\varepsilon$,

$$\mathcal{S}_n^\varepsilon(T) = \frac{1}{n^\kappa} \sum_{i=0}^\xi e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) \in T^\varepsilon(n)\}. \tag{5.26}$$

To bound its expected value we divide $T^\varepsilon(n)$ into slices $T_{\varepsilon2^{-i}}^{\varepsilon2^{-i+1}}$, $i \in \mathbb{N}$. We define

$$A_n(i) := \Big\{\boldsymbol{\tau} : \mathbb{E}\Big[\frac{1}{n^\kappa}\sum_{i=0}^\xi e_i \tau_{Y_n(i)} \mathbb{1}\big\{Y_n(i) \in T_{\varepsilon2^{-i}}^{\varepsilon2^{-i+1}}(n)\big\}\Big|\boldsymbol{\tau}\Big] < K'\varepsilon^{1-\alpha}2^{i(\alpha-1)}\Big\}. \tag{5.27}$$

We show that there is K' such that a.s.

$$A_n := \bigcap_{i=0}^\infty A_n(i) \quad \text{occurs for all but finitely many } n. \tag{5.28}$$

Lemma 5.5 is then a direct consequence of this claim.

To show (5.28) we first estimate the probability that a fixed site, say 1, is in $T_{\varepsilon2^{-i}}^{\varepsilon2^{-i+1}}$,

$$p_{n,i} := \mathbb{P}[1 \in T_{\varepsilon2^{-i}}^{\varepsilon2^{-i+1}}] \leq C\varepsilon^{-\alpha}2^{i\alpha}n^{-\kappa\alpha} \tag{5.29}$$

as follows from Assumption 2.3. For any $x \in \mathcal{V}_n$,

$$\mathbb{E}\Big[\sum_{i=0}^\xi e_i \mathbb{1}\{Y_n(i) = x\}\Big] = Tn^{\kappa\alpha-1}, \tag{5.30}$$

therefore

$$\mathbb{P}\Big[\mathbb{E}\Big[\frac{1}{n^\kappa}\sum_{i=0}^\xi e_i \tau_{Y_n(i)} \mathbb{1}\big\{Y_n(i) \in T_{\varepsilon2^{-i}}^{\varepsilon2^{-i+1}}(n)\big\}\Big|\boldsymbol{\tau}\Big] \geq K'\varepsilon^{1-\alpha}2^{i(\alpha-1)}\Big]$$

$$\leq \mathbb{P}\Big[\sum_{x\in\mathcal{V}_n} Tn^{\kappa\alpha-1}\varepsilon2^{-i+1}\mathbb{1}\big\{x \in T_{\varepsilon2^{-i}}^{\varepsilon2^{-i+1}}(n)\big\} \geq K'\varepsilon^{1-\alpha}2^{i(\alpha-1)}\Big]. \tag{5.31}$$

By Chebyshev inequality this is bounded from above by

$$\exp\{-\lambda n^{1-\kappa\alpha} c K' \varepsilon^{-\alpha} 2^{i\alpha}\} \prod_{x\in\mathcal{V}_n} (p_{n,i}(e^\lambda - 1) + 1). \tag{5.32}$$

Using $1 + x \le e^x$ and (5.29), it is easy to see that for K' large enough this is smaller than $\exp\{-cn^{1-\kappa\alpha} 2^{i\alpha} \varepsilon^{-\alpha}\}$. Therefore, summing over $i \in \mathbb{N}$,

$$\mathbb{P}[A_n^c] \le \exp\{-c' n^{1-\kappa\alpha} \varepsilon^{-\alpha}\} \tag{5.33}$$

An application of the Borel-Cantelli lemma finishes the proof of (5.28) and therefore of Lemma 5.5. $\qquad\square$

5.1.3 Very deep traps

We show that, with a large probability, none of the very deep traps is visited during the first $Tn^{\kappa\alpha}$ steps:

Lemma 5.6. *If $0 < \kappa < 1/\alpha$, then for all n large enough, $\boldsymbol{\tau}$-a.s.*

$$\mathbb{P}[\mathcal{S}_{n,M}(T) \neq 0 | \boldsymbol{\tau}] \le cTM^{-\alpha}. \tag{5.34}$$

Proof. Define $H = \inf\{k \in \mathbb{N}_0 : Y_n(k) \in T_M(n)\}$. By definition of $\mathcal{S}_{n,M}$, the claim of the lemma is equivalent to $\mathbb{P}[H \le \xi | \boldsymbol{\tau}] \le cTM^{-\alpha}$, with the same ξ as in the previous proof. Since

$$\mathbb{P}[H \le \xi | \boldsymbol{\tau}] \le \mathbb{P}[H < 2Tn^{\kappa\alpha} | \boldsymbol{\tau}] + \mathbb{P}[\xi > 2Tn^{\kappa\alpha}], \tag{5.35}$$

and the second term on the right-hand side decreases exponentially in n, it is sufficient to bound the first term. Using the fact that $\boldsymbol{\tau}$-a.s. $|T_M(n)| \le CM^{-\alpha} n^{1-\kappa\alpha}$ (which can be proved similarly as Lemma 5.4), we have

$$\mathbb{P}[H < 2Tn^{\kappa\alpha} | \boldsymbol{\tau}] = \mathbb{P}\Big[\bigcup_{k=0}^{2Tn^{\kappa\alpha}} \{Y_n(k) \in T^M(n)\}\Big|\boldsymbol{\tau}\Big] \le 2Tn^{\kappa\alpha} \frac{|T^M(n)|}{n} \le cM^{-\alpha}T. \tag{5.36}$$

This finishes the proof. $\qquad\square$

5.1.4 Proof of Theorem 5.1

We can now finish the proof of aging on a large complete graph. All claims here are valid $\boldsymbol{\tau}$-a.s.

First, for $\delta > 0$ we fix $T > 0$ such that (for all n large enough)

$$\mathbb{P}[\mathcal{S}_{n,\varepsilon}^M(T) + \mathcal{Z}_\varepsilon^M(T) \le 2 + \theta | \boldsymbol{\tau}] < \frac{\delta}{4}, \tag{5.37}$$

which is possible due to Corollary 5.3. Further, we use this corollary, Lemmas 5.5 and 5.6 to fix ε and M such that

$$\mathbb{P}[\mathcal{S}_{n,M}(T) \neq 0 | \boldsymbol{\tau}] \le \frac{\delta}{4}, \qquad \mathbb{P}\Big[\mathcal{S}_n^\varepsilon \ge \frac{\delta}{2}\Big|\boldsymbol{\tau}\Big] \le \frac{\delta}{4} \quad \text{and} \quad \mathbb{P}\Big[\mathcal{Z}_\varepsilon^M \ge \frac{\delta}{2}\Big] \le \frac{\delta}{4}. \tag{5.38}$$

Let $B(n)$ be the intersection of all events from the two previous displays. It follows that $\mathbb{P}[B(n)^c] \leq \delta$. Let $E(n)$ be the event whose probability we are trying to estimate (see (5.11)):

$$E(n) := \big\{\{\mathcal{S}_n(t) : t \in \mathbb{R}\} \cap [1, 1+\theta] = \emptyset\big\}. \tag{5.39}$$

On $B(n)$ we can approximate $\mathcal{S}_n(t)$ by $\mathcal{S}_{n,\varepsilon}^M(t) + \mathcal{Z}_\varepsilon^M(t)$:

$$B \implies \sup_{t \in [0,T]} |\mathcal{S}_n(t) - \mathcal{S}_{n,\varepsilon}^M(t) + \mathcal{Z}_\varepsilon^M(t)| < \delta. \tag{5.40}$$

Further, let $\mathcal{R}_{n,\varepsilon}^M = \{\mathcal{S}_{n,\varepsilon}^M(t) + \mathcal{Z}_\varepsilon^M(t) : t \in \mathbb{R}\}$. We define

$$\begin{aligned}
G_1(n) &:= \big\{\mathcal{R}_{n,\varepsilon}^M \cap \big((1-\delta, 1+\varepsilon) \cup (1+\theta-\delta, 1+\theta+\delta)\big) \neq \emptyset\big\}, \\
G_2(n) &:= G_1(n)^c \cap \big\{\mathcal{R}_{n,\varepsilon}^M \cap (1, 1+\theta) \neq \emptyset\big\}, \\
G_3(n) &:= G_1(n)^c \cap \big\{\mathcal{R}_{n,\varepsilon}^M \cap (1, 1+\theta) = \emptyset\big\} = G_1(n)^c \cap G_2(n)^c.
\end{aligned} \tag{5.41}$$

These three events form a partition of the probability space. The reason for this partition is the following. If $G_1(n) \cap B(n)$ happens, then the process $\mathcal{S}_{n,\varepsilon}^M + \mathcal{Z}_\varepsilon^M$ intersects δ-neighbourhoods of 1 or $1+\theta$, and it is therefore not possible to decide if $E(n)$ is true. On the other hand,

$$B(n) \cap G_3(n) \implies E(n) \qquad \text{and} \qquad B(n) \cap G_2(n) \implies E(n)^c \tag{5.42}$$

as can be seen from (5.40). Therefore

$$\mathbb{P}[B(n) \cap G_3(n)|\boldsymbol{\tau}] \leq \mathbb{P}[E(n)|\boldsymbol{\tau}] \leq \mathbb{P}[B(n)^c|\boldsymbol{\tau}] + \mathbb{P}[G_1(n)|\boldsymbol{\tau}] + \mathbb{P}[G_3(n)|\boldsymbol{\tau}]. \tag{5.43}$$

Thanks to Corollary 5.3 it is possible to estimate the probabilities of $G_1(n)$ and $G_3(n)$. Indeed, $\mathcal{S}_{n,\varepsilon}^M + \mathcal{Z}_\varepsilon^M$ converge weakly in the Skorokhod topology to an α-stable subordinator. The probability that the subordinator hits any of the boundary points $1 \pm \delta$, $1 + \theta \pm \delta$ is zero. Therefore, it follows from the weak convergence and Corollaries 5.3 and A.5 that

$$\lim_{n \to \infty} \mathbb{P}[G_3(n)|\boldsymbol{\tau}] = \mathsf{Asl}_\alpha\left(\frac{1-\delta}{1+\theta+\delta}\right). \tag{5.44}$$

Similarly, by (A.9),

$$\mathbb{P}[G_1(n)|\boldsymbol{\tau}] \leq 1 - \mathsf{Asl}_\alpha\left(\frac{1-\delta}{1+\delta}\right) + 1 - \mathsf{Asl}_\alpha\left(\frac{1+\theta-\delta}{1+\theta+\delta}\right) \leq C\delta^{1-\alpha}. \tag{5.45}$$

Since δ is arbitrary, the claim of the theorem follows from (5.43)–(5.45) and from the continuity of $\mathsf{Asl}_\alpha(\cdot)$. This proves aging on the complete graph for $\kappa < 1/\alpha$.

5.2   The α-stable subordinator as a universal clock

We will now give a general set of conditions which ensures that the result and the proof that we have given in the simple case of the complete graph apply. Consider an arbitrary sequence of graphs G_n and the Bouchaud trap models on them, $\mathsf{BTM}(G_n, \boldsymbol{\tau}_n, 0)$. We set

$\nu = 1$, therefore at every vertex x the Markov chain X waits an exponentially distributed time with mean τ_x/d_x (d_x is the degree of x) and then it jumps to one of the neighbouring vertices with an equal probability. We suppose that Assumption 2.3 holds, in particular that means that τ_x are i.i.d.

We want to prove aging for this model at a time scale $t_w = t_w(n)$ using the same strategy as for the complete graph. That means, depending on the sequence G_n and the time scale t_w:

- to divide traps into three groups: shallow, deep, and very deep,

- to prove that the shallow traps can be ignored since the time spent there is negligible,

- to prove that the very deep traps can be ignored because they are not visited in the proper time scale,

- to show that the contribution of the deep traps to the time change can be approximated by a stable subordinator, which will show the convergence of the clock process,

- and finally, to deduce aging for the two-time function R from this convergence.

The idea is that, as in the case of the complete graph, the time change $S_n(j)$ should be dominated by a relatively small number of large contributions coming from the deep traps.

On the other hand we want to stay as general as possible: we do not want to use any particular properties of the graph. Therefore, we formulate six conditions on the sequence G_n and the simple random walk $Y_n(\cdot)$ on it. If these conditions are verified, the proof of aging can be finished in the spirit of Section 5.1.4. Proving these conditions should be dependent on the graphs G_n.

To formulate the conditions it is necessary to choose several objects that depend on the particular sequence G_n and on the observation time scale $t_w(n)$.

First, it is necessary to fix a (random) time ξ_n up to which we observe Y_n. Second, a scale $g(n)$ for deep traps should be chosen according to G_n and $t_w(n)$. This scale defines the set of the deep traps by

$$T_\varepsilon^M(n) := \{x \in \mathcal{V}_n : \varepsilon g(n) \leq \tau_x < M g(n)\}. \tag{5.46}$$

A possible generalisation of the definition (5.46) is described in the remark after the Theorem 5.7.

For the complete graph we used $g(n) = n^\kappa$ and $\xi_n \sim n^{\kappa\alpha}$ (see the proof of Lemma 5.5).

As before, we use $T_M(n) = \{x : \tau_x \geq M g(n)\}$ to denote the set of *very deep traps*. Similarly, we write $T^\varepsilon(n) = \{x : \tau_x < \varepsilon g(n)\}$ for the set of *shallow traps*.

It should be possible to almost ignore these two sets. This is ensured by the following two conditions. Compare them with Lemmas 5.5 and 5.6.

Condition 1. There is a function $h(\varepsilon)$ satisfying $\lim_{\varepsilon \to 0} h(\varepsilon) = 0$, such that for a.e. realisation of $\boldsymbol{\tau}$ and for all n large enough

$$\mathbb{E}\Big[\sum_{i=0}^{\xi_n} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) \in T^\varepsilon(n)\} \Big| \boldsymbol{\tau} \Big] \le h(\varepsilon) t_w(n). \tag{5.47}$$

In words, the expected time spent in the shallow traps before ξ_n is small with respect to $t_w(n)$.

Let $H_n(A)$, $A \subset \mathcal{V}_n$, denotes the hitting time of A by the simple random walk Y_n,

$$H_n(A) := \inf\{i \ge 0 : Y_n(i) \in A\}. \tag{5.48}$$

Condition 2. Given ξ_n, for any $\delta > 0$ there exists M large enough such that for a.e. realisation of $\boldsymbol{\tau}$ and for all n large

$$\mathbb{P}\big[H_n(T_M(n)) \le \xi_n \big| \boldsymbol{\tau} \big] \le \delta. \tag{5.49}$$

We should now ensure that the contribution of the deep traps to the time change can be approximated by an α-stable subordinator. The following facts were crucial for the proof in the complete graph case: asymptotically, as $n \to \infty$,

- every time a deep trap is visited, its depth is independent of depths of the previously visited deep traps.

- the probability that a deep trap with depth $ug(n) = un^\kappa$ is visited is proportional to $u^{-\alpha}$.

- for fixed ε and M only a finite number of deep traps was visited before the time-horizon ξ.

However, these facts could not be true for a general graph. Here, after leaving a deep trap, the process Y typically hits this trap with larger probability than any other deep trap. On recurrent graphs, such as $\mathbb{Z}^2$, it even visits this trap a number of times that diverges with n.

To overcome this problem we "group" the successive visits of one deep trap. The time spend during these visits there will then be considered as one contribution to the clock process. We define $r_n(j)$ as the sequence of times when a new deep trap is visited, $r_n(0) = 0$, and

$$r_n(i) = \min \big\{ j > r_n(i-1) : Y_n(j) \in T_\varepsilon^M(n) \setminus \{Y_n(r_n(i-1))\} \big\}. \tag{5.50}$$

We use ζ_n to denote the largest j such that $r_n(j) \le \xi_n$,

$$\zeta_n := \max\{j : r_n(j) \le \xi_n\}. \tag{5.51}$$

We define the process $U_n(j)$ that records the trajectory of Y_n (and thus of X_n) restricted to the deep traps,

$$U_n(j) := Y_n(r_n(j)), \qquad j \in \mathbb{N}_0. \tag{5.52}$$

35

Finally, define the score $s_n(j)$ be the time that X_n spends at site $U_n(j)$ between steps $r_n(j)$ and $r_n(j+1)$,

$$s_n(j) := \sum_{i=r_n(j)}^{r_n(j+1)} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) = U_n(j)\}, \qquad j < \zeta_n. \tag{5.53}$$

Denoting by $G_A^n(x, y)$ the Green's function of the simple random walk Y_n killed on the first visit to a set $A \subset \mathcal{V}_n$, it is easy to observe that $s_n(j)$ has the exponential distribution with mean

$$d_x^{-1} \tau_{U_n(j)} G_{T_\varepsilon^M \setminus \{U_n(j)\}}^n (U_n(j), U_n(j)). \tag{5.54}$$

Since Conditions 1 and 2 ensure that the visits of deep traps determine the behaviour of the time change $S_n(j)$, the sum $\sum_{i=1}^{j-1} s_n(i)$ can be considered as a good approximation of $S_n(r_n(j))$. The next condition guarantees that the scores $s_n(i)$ have a good asymptotic behaviour, i.e. are independent and have the right tail.

Condition 3. Let $(s_\infty(i) : i \in \mathbb{N})$ be an i.i.d. sequence given by $s_\infty(i) := \hat{e}_i \sigma_\varepsilon^M(i)$, where $\sigma_\varepsilon^M(i)$ is a sequence of i.i.d. random variables taking values between ε and M with common distribution function

$$\mathbb{P}[\sigma_\varepsilon^M(i) \leq u] = \frac{\varepsilon^{-\alpha} - u^{-\alpha}}{\varepsilon^{-\alpha} - M^{-\alpha}} =: \frac{\varepsilon^{-\alpha} - u^{-\alpha}}{p_\varepsilon^M}, \qquad u \in [\varepsilon, M]. \tag{5.55}$$

and $\hat{e}_i$ is an i.i.d. sequence of exponential, mean-one random variables independent of σ_ε^M. Then there exists a constant $\mathcal{K} > 0$ such that for all ε, M and for a.e. $\boldsymbol{\tau}$, the sequence $(s_n(j)/t_w(n), j \in \mathbb{N})$ converges as $n \to \infty$ in law to the sequence $(\mathcal{K}s_\infty(j), j \in \mathbb{N})$. (For notational convenience we define $s_n(j) = s_\infty(j)$ for all $j \geq \zeta_n$.)

The last three conditions will ensure that the approximation by the α-stable subordinator is relevant for the aging.

First, we need that $S_n(r_n(\zeta_n))$ is larger than $(1+\theta)t_w(n)$. Since $r_n(\zeta_n) \geq \sum_{i=1}^{\zeta_n-1} s_n(i)$, and $s_n(i)$ are easier to control than $S_n(r_n(j))$ we require

Condition 4. For a.e. $\boldsymbol{\tau}$ and for any fixed $\theta > 0$, $\delta > 0$ it is possible to choose ξ_n such that for all ε small and M large enough, and for ζ_n defined in (5.51)

$$\mathbb{P}\Big[\sum_{i=1}^{\zeta_n-1} s_n(i) \geq (1+\theta)t_w(n) \Big| \boldsymbol{\tau} \Big] \geq 1 - \delta. \tag{5.56}$$

Second, to prove aging for the two-point function R we need to show that for any time t' between $S_n(r_n(j))$ and $S_n(r_n(j+1))$ the probability that $X_n(t') = U_n(j)$ is large. For a formal statement of this claim we need some definitions. Let t'_n be a deterministic time sequence satisfying $t_w(n)/2 \leq t'_n \leq (1+\theta)t_w(n)$, and let $\delta > 0$. We define $j_n \in \mathbb{N}$ by

$$S_n(r_n(j_n)) \leq t'_n \leq S_n(r_n(j_n+1)) - \delta t_w(n), \tag{5.57}$$

and $j_n = \infty$ if (5.57) is not satisfied for any integer. Let $A_n(\delta)$ be an event defined by

$$A_n(\delta) = \{0 < j_n < \zeta_n\}. \tag{5.58}$$

We require

Condition 5. For any δ it is possible to choose ε small and M large enough such that for a.e. $\boldsymbol{\tau}$ and all n large enough

$$\mathbb{P}[X_n(t_n') = U_n(j_n)|A_n(\delta), \boldsymbol{\tau}] \geq 1 - \delta. \tag{5.59}$$

The last condition that we need excludes repetitions in the sequence U_n.

Condition 6. For any fixed ε and M and a.e. $\boldsymbol{\tau}$

$$\lim_{n\to\infty} \mathbb{P}[\exists 0 < i, j \leq \zeta_n \text{ such that } i \neq j \text{ and } U_n(i) = U_n(j)|\boldsymbol{\tau}] = 0. \tag{5.60}$$

We have formulated the six conditions that are inspired by the complete graph proof. It should be then not surprising that they imply the same result as for the complete graph:

Theorem 5.7 (Aging on general graphs). *Assume that Conditions 1–6 holds. Then for a.e. realisation of the random environment $\boldsymbol{\tau}$*

$$\lim_{n\to\infty} R_n(t_w(n), (1+\theta)t_w(n); \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1+\theta). \tag{5.61}$$

This theorem can be proved in a very similar way as Theorem 5.1. The complete proof can be found in [BČ06a].

Remark. 1. The set $T_\varepsilon^M(n)$ as defined in (5.46) is sometimes too large and not all conditions that we formulated can be verified easily. Typically this happens when two points of the top are too close to each other with non-negligible probability. In such case it is useful to define a set $B(n) \subset T_\varepsilon^M(n)$ of bad traps which cause the difficulties. If it is then possible to verify Conditions 1–6 for the set $T'^M_\varepsilon(n) := T_\varepsilon^M(n) \setminus B(n)$ and, moreover, if $B(n)$ satisfies for a.e. $\boldsymbol{\tau}$ a similar condition as the set of the very deep traps T_M,

$$\limsup_{n\to\infty} \mathbb{P}[H_n(B(n)) \leq \xi_n|\boldsymbol{\tau}] = 0, \tag{5.62}$$

then the conclusions of Theorem 5.7 hold without change.

2. The last pair of conditions is, in principal, necessary only for a "post-processing". If they are not verified, it is possible to prove aging for a *top-dependent* correlation function

$$R_n'(t_w, t_w + t; \boldsymbol{\tau}) = \mathbb{P}\big[\exists j : S_n(r_n(j)) \leq t_w < t_w + t \leq S_n(r_n(j+1))\big|\boldsymbol{\tau}\big], \tag{5.63}$$

which gives the probability that at most one site in the top is visited by X_n during the observed time interval. A two-point function similar to R_n' was considered in [BBG03a].

Theorem 5.8. *If only Conditions 1–4 hold, then $\boldsymbol{\tau}$-a.s.*

$$\lim_{n\to\infty} R_n'(t_w(n), (1+\theta)t_w(n); \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1+\theta). \tag{5.64}$$

5.3 Potential-theoretic characterisation

In the previous paragraph we have stated six conditions that allow to prove aging on an arbitrary sequence of graphs. It is however not clear if these conditions can be verified for any concrete model. In fact, they are satisfied for the graphs: $G_n = G = \mathbb{Z}^d$ with $d > 1$; for G_n a large torus in d dimensions, $G_n = \mathbb{Z}^d/n\mathbb{Z}^d$; for the n-dimensional hypercube, $G_n = \{-1,1\}^n$ among others (included of course the complete graph!). These examples will be developed in the next section, but we wont be able to give proofs, which can be pretty difficult (see [BČM06, BČ06a]). Rather, we want to give here a few hints with potential-theoretic flavor on how to verify our conditions.

We want mainly discuss the crucial Condition 3: the convergence of the scores $s_n(i)$ to the i.i.d. sequence s_∞. In the discussion we suppose that G_n are finite and sufficiently regular. Observe first that the set of deep traps T_ε^M is a *random cloud* on $\mathcal{V}_n$ (i.e. set of points chosen independently from $\mathcal{V}_n$, we assume that τ_x are i.i.d.). The intensity ρ_n of this cloud depends on n; under Assumption 2.3 $\rho_n \sim p_\varepsilon^M g(n)^{-\alpha} = (\varepsilon^{-\alpha} - M^{-\alpha})g(n)^{-\alpha}$. We, obviously, need that $\rho_n \to 0$, so that the random cloud is sufficiently sparse, but also $\rho_n|\mathcal{V}_n| \to \infty$, so that the mean size of the random cloud diverges (otherwise fluctuations of the random depths are important and no a.s. convergence holds).

Going back to sequence $s_n(i)$, we have already remarked that conditionally on $U_n(j)$, the jth visited deep trap, the score $s_n(j)$ is exponentially distributed with mean

$$\nu_n^{-1} d_{U_n(j)}^{-1} \tau_{U_n(j)} G_{T_\varepsilon^M(n)\setminus\{U_n(j)\}}^n (U_n(j), U_n(j)), \qquad (5.65)$$

where $G_A^n(\cdot, \cdot)$ is the Green's function of the simple random walk Y_n that is killed on the first hit of the set $A \subset \mathcal{V}_n$.

If the graphs G_n are "sufficiently regular", then all vertices have similar degree, that means, e.g., that there is a scale $d(n)$ such that $d_x/d(n)$ is uniformly bounded from 0 and ∞. Then by setting $\nu_n = d(n)^{-1}$ we can, at least theoretically, ignore the first two terms in (5.65). This problem does not appear in all examples we consider, there always $d_x = \text{const}(n)$. Hence, it remains to control $\tau_{U_n(j)}$ and the Green's function in (5.65).

For the Green's function, one typically prove that there is a scale $f(n) \sim t_w(n)/g(n)$ such that $\boldsymbol{\tau}$-a.s.

$$f(n)^{-1} G_{T_\varepsilon^M(n)\setminus\{x\}}^n (x, x) \xrightarrow{n\to\infty} \text{const.} \qquad (5.66)$$

uniformly for all $x \in T_\varepsilon^M(n)$. This result is again reasonable if the graph G_n is sufficiently regular and finite, and the cloud $T_\varepsilon^M(n)$ is very diluted.

To control the distribution of $\tau_{U_n(j)}$ consider first an arbitrary random cloud $A_n \subset \mathcal{V}_n$ with intensity $c\rho_n$. This random cloud will represent the set of the deep traps or its subsets. Recall that $H_n(A)$ denotes the hitting time of A by Y_n. Let $\mathbb{P}_x$ be the law of Y_n started at x. Suppose that it is possible to show for all $u > 0$ and some scale $r(n)$ independent of c

$$\sup_{x \in A_n} \left| \mathbb{P}_x\left[\frac{H_n(A_n \setminus \{x\})}{r(n)} \geq u \right] - \exp(-cu) \right| \xrightarrow{n\to\infty} 0, \qquad \boldsymbol{\tau} - \text{a.s.}, \qquad (5.67)$$

so that the distribution of the normalised hitting time converges uniformly to the exponential distribution with mean c^{-1}. This is again a reasonable property for very diluted

clouds. If (5.67) holds for all c then the lack-of-memory property of the exponential distribution allows to prove the following claim: Let A_n be a random cloud with intensity $(a + b)\rho_n$ and let $B_n \subset A_n$ be its sub-cloud with intensity $b\rho_n$. Then uniformly for all $x \in A_n$

$$\mathbb{P}_x[H(B_n \setminus \{x\}) < H(A_n \setminus \{x\})] \xrightarrow{n \to \infty} \frac{b}{a + b}. \qquad (5.68)$$

This claim yields, e.g. under Assumption 2.3, that for $u \in (\varepsilon, M)$

$$\mathbb{P}_{U_n(j-1)}[\tau_{U_n(j)} \leq u] \xrightarrow{n \to \infty} \frac{p_\varepsilon^u}{p_\varepsilon^M}, \qquad (5.69)$$

which is exactly what gives the "right tail" and the asymptotic independence of $s_n(i)$. The Condition 3 is then consequence of the claims of the last three paragraphs.

If the graphs G_n are infinite one typically cannot prove uniformity in (5.66) and (5.67). One can however restrict only to those deep traps that are reachable in ξ_n steps from the starting position. On this set the uniform control is usually possible.

There is also a heuristic reason for Condition 3. As the random cloud $T_\varepsilon^M(n)$ becomes more diluted, the hitting measure of this cloud charges more and more points because the random walk can "pass more easily around". If this happens, the law-of-large-numbers-type arguments hold not only for the whole sets $T_\varepsilon^u(n)$ but also for these sets *as sampled by the hitting measure*. That is why Condition 3 holds.

Remark. This is *not true* on $\mathbb{Z}$! The simple random walk on $\mathbb{Z}$ cannot "pass around": it hits necessarily one of two points of the random cloud that are "neighbours" of its starting position. This explains the special properties of the BTM on $\mathbb{Z}$ and also the necessity of averaging in Theorems 3.11 and 3.13.

The Conditions 2 and 4 are also easy consequences of claim (5.67). One need first fix $\xi_n = mr(n)$ with m large enough to accumulate large enough (but finite) number K of "independent" scores $s_n(i)$ in order to have with a large probability $\sum_{i=1}^K s_n(i) > 1 + \theta$. This satisfies Condition 4. Then one fix M large enough, such that the intensity $c(M)\rho_n$ of $T_M(n)$ satisfies $c(M) > \delta^{-1}m$. Then Condition 2 holds.

The preceding discussion can be summarised as follows:

Claim. *If* (5.66) *and* (5.67) *can be checked for* $T_\varepsilon^M(n)$ *on* G_n, *then, under Assumption 2.3, Conditions 2–4 can be verified.*

The Condition 1 does not follow directly from (5.67) since Poisson clouds with larger intensity than ρ_n should be considered. The "slicing strategy" as presented in the proof for the complete graph however usually works. The remaining Conditions 5 and 6 are not substantial and we do not discuss them here.

Remark also that Condition 1 together with (5.66) and (5.67) allows to prove the approximation of the clock process by an α-stable subordinator. This approximation is not a consequence of Conditions 1–6 only.

There are at least two methods of proof for facts (5.66) and (5.67). The first is the coarse-graining procedure of [BČM06] that is explained in Section 4.4. The advantage of

this procedure is that it should work on many different graphs. It is however relatively technical and many special cases should be treated apart. One can also use the formula given by Matthews [Mat88]. This method is used in [BČ06a] to prove aging in the REM and on the torus in $\mathbb{Z}^2$. It however applies only if Y_n "completely forgets" the position of $U_n(i)$ before hitting $U_n(i+1)$, i.e. that the hitting measure of the random cloud is essentially uniform. Therefore this method does not apply e.g. on $\mathbb{Z}^d$.

6 Applications of the arcsine law

We now describe two examples where the approach of Section 5 can be used to prove aging.

6.1 Aging in the REM

The Random Energy Model is the simplest mean-field model for spin-glasses and its static behaviour is well understood. The studies of dynamics are much more sparse. The first proof of aging in the REM was given in [BBG03a, BBG03b], based on renewal theory. The general approach of Section 5 gives another, shorter proof. This approach allows to prove aging on a broader range of time scales, but on the other hand do not include quite exactly the results of [BBG03a, BBG03b]. We will compare both results later. Before doing it, let us define the model and give some motivation why and in which ranges of times and temperatures aging occurs.

The Random Energy model is a mean-field model of a spin-glass. It consists of n spins that can take values -1 or 1, that is configurations of the REM are elements of $\mathcal{V}_n = \{-1, 1\}^n$. The energies $\{E_x, x \in \mathcal{V}_n\}$ of the configurations are i.i.d. random variables. The standard choice of the marginal distribution of E_x is centred normal distribution with variance n. We will, however, deviate from the standard choice to simplify the computations and we will assume that $-E_x$ are i.i.d. positive random variables with the common distribution given by

$$\mathbb{P}[-E_x/\sqrt{n} \geq u] = e^{-u^2/2}, \qquad u \geq 0. \tag{6.1}$$

We then define

$$\tau_x = \exp(-\beta E_x). \tag{6.2}$$

The distribution (6.1) has almost the same tail behaviour as the normal distribution. As we already know, it is the tail behaviour of τ_x (and thus of E_x) that is responsible for aging. Therefore, the use of "faked normal distribution" (6.1) is not substantial for our discussion. Remark also that a similar trick, i.e. to take E_x to be minimum of 0 and the normal variable, was used in [BBG03b].

For the dynamics of the REM we require that only one spin can be flipped at a given moment. This corresponds to

$$\mathcal{E}_n = \{\langle x, y \rangle \in \mathcal{V}_n^2 : \sum_{i=1}^{n} |x_i - y_i| = 2\}, \tag{6.3}$$

where $(x_1, \ldots, x_n)$ are the values of individual spins. We use G_n to denote the n-dimensional hypercube $(\mathcal{V}_n, \mathcal{E}_n)$. There are many choices for the dynamics of REM, that has the Gibbs measure $\boldsymbol{\tau}$ as a reversible measure. We will naturally consider the trap model dynamics $\mathsf{BTM}(G_n, \boldsymbol{\tau}, 0)$, which is one of the simplest choices. We fix $\nu_n = 1/n$ in definition (2.2), so that τ_x is the mean waiting time at x. We always suppose that

$$Y_n(0) = X_n(0) = \mathbf{1} = (1, \ldots, 1). \tag{6.4}$$

6.1.1 Short time scales

Theorem 6.1. *Let the parameters $\alpha \in (0, 1)$ and $\beta > 0$ be such that*

$$3/4 < \alpha^2 \beta^2 / 2 \log 2 < 1. \tag{6.5}$$

Define

$$t_w(n) := \exp(\alpha \beta^2 n). \tag{6.6}$$

Then, for a.e. $\boldsymbol{\tau}$,

$$\lim_{n \to \infty} R_n(t_w(n), (1 + \theta)t_w(n)) = \mathsf{Asl}_\alpha(1/1 + \theta). \tag{6.7}$$

Let us first explain the appearance of scale $t_w(n)$ together with one problem that is specially related to REM. We have seen in Section 5 that aging occurs only if τ_x are sufficiently heavy-tailed. This certainly fails to be true for the REM: an easy calculation gives $\mathbb{P}[\tau_x \geq u] = u^{-\log u/2\beta^2 n}$, which decreases faster than any polynomial. It is therefore clear that, if the system is given enough time to explore a large part of the configuration space and thus to discover the absence of heavy tails, then no aging occurs, at least not in our picture. On the other hand, at shorter time scales the system does not feel the non-existence of heavy tails as can be seen from the following estimate. Let $\alpha > 0$, then

$$
\begin{aligned}
e^{\alpha^2 \beta^2 n/2} \mathbb{P}\big[\tau_x \geq u e^{\alpha\beta^2 n}\big] &= e^{\alpha^2 \beta^2 n/2} \mathbb{P}\Big[E_x \geq \frac{\log u + \alpha\beta^2 n}{\beta\sqrt{n}}\Big] \\
&= \exp\Big\{ -\frac{\log^2 u}{2\beta^2 n} - \alpha \log u \Big\} \xrightarrow{n \to \infty} u^{-\alpha}.
\end{aligned}
\tag{6.8}
$$

and therefore

$$\mathbb{P}\Big[\frac{\tau_x}{e^{\alpha\beta^2 n}} \geq u\Big] = e^{-\alpha^2 \beta^2 n/2} \cdot u^{-\alpha}(1 + o(1)) \qquad (n \to \infty). \tag{6.9}$$

In view of the fact that the simple random walk on the hypercube almost never backtracks, it seems reasonable to presume that if the process had time to make only approximately $e^{\alpha^2 \beta^2 n/2}$ steps, then it has no time to discover the absence of heavy tails and aging could be observed. The above theorem shows this presumption to be true.

Let us remark that there is much stronger relation between "random exponentials" τ_x and heavy-tailed random variables. Let $(F_i, i \in \mathbb{N})$ be an i.i.d. sequence of centred

normal random variables with variance one. It was proved in [BBM05] that for some properly chosen $Z(n)$ and $N(n)$ the normalised sum

$$\frac{1}{Z(n)} \sum_{i=1}^{N(n)} e^{-\beta\sqrt{n}F_i} \tag{6.10}$$

converges as $n \to \infty$ in law to an α-stable distribution with α depending on β and $N(n)$. We show that the same is true for the properly normalised clock process $S(n)$, which is a properly normalised sum of correlated random variables, more precisely of i.i.d. random variables sampled by a random walk.

In view of (6.9) it is easy to fix objects for which Conditions 1–6 should be verified: we define

$$t_w(n) := \exp(\alpha\beta^2 n), \tag{6.11}$$

$$\xi_n := m \exp(\alpha^2\beta^2 n/2), \tag{6.12}$$

$$T_\varepsilon^M(n, \alpha) := \{x \in \mathcal{V}_n : \tau_x \in (\varepsilon, M)e^{\alpha\beta^2 n}\}. \tag{6.13}$$

The Theorem 6.1 is then the consequence of the following proposition and Theorem 5.7.

Proposition 6.2. *Let α and β be as in Theorem 6.1. Then for any θ it is possible to choose m large enough such that Conditions 1–6 hold for $\mathbb{P}$-a.e. $\boldsymbol{\tau}$.*

We believe that the range of the validity (6.5) of the Theorem 6.1 is not the broadest possible. The upper bound 1 is correct. If $\alpha^2\beta^2/2\log 2 > 1$, then $\xi_n \gg 2^n$. That means that the state space $\mathcal{V}_n$ becomes too small and the process can feel its finiteness. On the other hand, the lower-bound $3/4$ is purely technical and can probably be improved.

Observe also that the condition (6.5) can be rewritten as

$$\alpha^{-1}\beta_c\sqrt{3/4} < \beta < \alpha^{-1}\beta_c, \tag{6.14}$$

where $\beta_c = \sqrt{2\log 2}$ is the critical temperature of the usual REM. This, in particular, means that aging can be observed in REM also *above the critical temperature, $\beta < \beta_c$.*

The proof of this proposition in [BČ06a] follows the strategy outlined in Section 5 and uses the results of Matthews [Mat88] for the fine control of the simple random walk on the hypercube. In particular, (5.67) is a consequence of the following potential-theoretic result which might be of independent interest.

Proposition 6.3. *(i) Let for all $n \geq 1$ sets $A_n \subset \mathcal{V}_n$ be such that $|A_n| = \rho_n 2^n$ with "densities" ρ_n satisfying $\lim_{n\to\infty} \rho_n 2^{\gamma n} = \rho \in (0, \infty)$ for some $\gamma \in (1/2, 1)$. Let further the sets A_n satisfy the* minimal distance condition

$$\min\{d(x, y) : x, y \in A_n\} \geq (\omega(\gamma) + \varepsilon)n \tag{6.15}$$

for some small constant $\varepsilon > 0$ and for the unique solution $\omega(\gamma)$ of

$$\omega\log\omega + (1 - \omega)\log(1 - \omega) + \log 2 = (2\gamma - 1)\log 2, \qquad \omega \in (0, 1/2). \tag{6.16}$$

Then for all $s \geq 0$

$$\lim_{n \to \infty} \max_{x \in A_n} \left| \mathbb{E}_x \left[\exp\left(-\frac{s}{2^{\gamma n}} H_n(A_n \setminus \{x\}) \right) \right] - \frac{\rho}{s + \rho} \right| = 0. \tag{6.17}$$

That means that the hitting time $H_n(A_n \setminus \{x\})/2^{\gamma n}$ is asymptotically exponentially distributed with mean $1/\rho$.

(ii) If A_n are random clouds with intensity ρ_n such that $\lim_{n \to \infty} \rho_n 2^{\gamma n} = \rho \in (0, \infty)$ and $\gamma \in (3/4, 1)$, then the assumptions of (i) are a.s. satisfied.

This result also explains the appearance of the lower bound $3/4$ in the range of the validity of Theorem 6.1: the set of deep traps satisfies the assumptions of Proposition 6.3(ii) only if $\alpha^2 \beta^2 / 2 \log 2 > 3/4$.

6.1.2 Long time scales

The result of [BBG03a, BBG03a] deals with the longest possible time scales where aging appears in REM. The continuous-time Markov process X is replaced by a discrete-time process X', which at every step has the possibility not to move. The number of tries before leaving x has geometrical distribution with mean τ_x. As $n \to \infty$, this dynamics differs very little from the usual trap model dynamics. The random environment is given by

$$\tau_x = \exp(\beta \sqrt{n} \max(E_x, 0)), \tag{6.18}$$

where E_x are i.i.d. centred normal random variables with variance one.

To define the relevant time and depth scales we set

$$u_n(E) = \beta_c \sqrt{n} + E/\beta_c \sqrt{n} - \log(4\pi n \log 2)/2\beta_c \sqrt{n}. \tag{6.19}$$

We define the set of deep traps (*the top*) by

$$T_n(E) = \{x \in \mathcal{V}_n : E_x \geq u_n(E)\}. \tag{6.20}$$

The function $u_n(E)$ is chosen in such way that, as $n \to \infty$ and E is kept fixed the distribution of $|T_n(E)|$ converges to the Poisson distribution with mean that depends only on E. The mean diverges if $E \to -\infty$ afterwards. The initial position of the discrete-time Markov chain X' is chosen to be uniformly distributed in $T_n(E)$.

A different correlation function considered in [BBG03a]: If $x_n(k)$ denotes the last trap visited by X' before step k then

$$\Pi'_n(k, k + l, E; \boldsymbol{\tau}) = \mathbb{P}\big[\{X'(i) : i \in \{k+1, \ldots, k+l\} \cap (T_n(E) \setminus \{x_n(k)\}) = \emptyset \big| \boldsymbol{\tau} \big]. \tag{6.21}$$

It is essentially the same as the function R' (see (5.63))

This is the main aging result of [BBG03b]

Theorem 6.4. *For any $\beta > \beta_c = \sqrt{2 \log 2}$ and any $\varepsilon > 0$*

$$\lim_{t \to \infty} \lim_{E \to -\infty} \lim_{n \to \infty} \mathbb{P}\left[\left|\frac{\Pi'_n(c_n t, (1+\theta)c_n t, E; \boldsymbol{\tau})}{\mathsf{Asl}_{\beta_c/\beta}(1/1+\theta)} - 1\right| > \varepsilon\right] = 0, \tag{6.22}$$

where $c_n \sim e^{\beta\sqrt{n}u_n(E)}$.

In fact one can see that the dynamics of the REM when observed only on the top $T_n(E)$ can be approximated very well when $n \to \infty$ and $E \to -\infty$ by a BTM on the complete graph with $M = |T_n(E)| \sim e^{-E}$ vertices (see [BBG02, BBG03a]).

Let us now compare the results of Theorems 6.1 and 6.4. First, different correlation functions R and Π' are considered. This difference is not substantial, we believe that it is possible to eliminate the top dependence (i.e. to convert something like R' to something like R) of (6.22) by some post-processing in the direction of Condition 5.

The a.s. convergence in Theorem 6.1 is stronger than the convergence in probability in Theorem 6.4. It is a consequence of the fact that much larger time scales are considered and the set of the deep traps $T_n(E)$ is finite for fixed E. Therefore, we cannot use law-of-large-numbers-type arguments for the number of deep traps with depth in a fixed interval. We have seen this effects already in the case of the complete graph (see (5.8)).

The main difference between the two theorems is in the considered top sizes and time scales. In Theorem 6.4 the size of the top is kept bounded as $n \to \infty$. This allows to apply "lumping techniques" to describe the properties of the projection of a simple random walk on the hypercube to the top, that is to prove that that an equivalent of the process U_n (see (5.52)) converges to the simple random walk on the complete graph with the vertex set $T_n(E)$. In Theorem 6.1 (or more precisely, in Proposition 6.2) the size of the top $T_\varepsilon^M(n)$ increases exponentially with n. This makes the application of the lumping not possible and techniques based on Matthews results, i.e. Proposition 6.3, should be used.

The time scale $c_n \sim e^{\beta\sqrt{n}u_n(E)} \sim e^{\beta\beta_c n + \beta E/\beta_c}$ of Theorem 6.4 corresponds to the case $\alpha\beta/\beta_c = 1$ and is much larger than the scale $t_w(n) \sim e^{\alpha\beta^2 n} = e^{\frac{\alpha\beta}{\beta_c}\beta\beta_c n}$. These scales approach if $\alpha\beta/\beta_c$ tends to 1, which is the upper limit of the validity of Theorem 6.1. It would be possible to improve this theorem by setting $t_w(n) = e^{\beta\beta_c n}f(n)$ with some $f(n) \to 0$ as $n \to \infty$ sufficiently fast, but even then $t_w(n) \ll c_n$. Exactly at $\alpha\beta/\beta_c = 1$ Theorem 6.1 does not hold. As we have already remarked, in this case it is necessary to use the double-limit procedure and the convergence in probability.

6.1.3   Open questions and conjectures

The Theorems 6.1 and 6.4 give rigorous proofs of aging in the REM. They are however only partly satisfactory. It would be nice to replace the RHT dynamics by a more physical dynamics, like e.g. Glauber, or, at least, to explore the $a \neq 0$ case. We believe that the long-time behaviour of the model should not change dramatically, however we do not know any proof of it. The problem is that the Markov chain X becomes a time change of a random walk in random environment on the hypercube. Moreover, the clock process and the random walk are dependent.

Another natural direction of research is to extend the results for the RHT dynamics on the REM to other mean-field spin-glasses, like the SK model or the p-spin SK model. In these models the energies of the spin configurations E_x are no longer independent. We strongly believe that the approach of Section 5 can be applied, at least for large p and for well chosen time scales. These scales should be short enough not to feel the extreme values of the E_x's which rule the (model-dependent) statics, but long enough for the convergence to a stable subordinator to take place for the clock-process. The difficulty is to verify Conditions 1–6 if the E_x's are not i.i.d. The assumption that the E_x's are independent is used twice in the proof for the REM. First, we use it to verify Condition 1, that is to prove that the time spent in the shallow traps is small. We believe that this condition stays valid also for dependent spin-glass models. The second use of the independence is more substantial. It is used to describe the geometrical structure of the set of the deep traps. More exactly it is used to bound from below the minimal distance between deep traps and to show that the number of traps in $T_u^{u+\delta}(n)$ is proportional to $u^{-\alpha}$. It is an open question if these properties remain valid for dependent spin-glasses.

6.2 Aging on large tori

Another graph where the approach of Section 5 can be used to prove aging is a torus in $\mathbb{Z}^d$. For convenience we will consider only $d = 2$ here, although similar results are expected to hold for $d \geq 3$.

Let $G_n = (\mathcal{V}_n, \mathcal{E}_n)$ be the two-dimensional torus of size 2^n with nearest-neighbours edges, i.e. $\mathcal{V}_n = \mathbb{Z}^2/2^n\mathbb{Z}^2$, and edge $\langle x, y \rangle$ is in $\mathcal{E}_n$ iff

$$\sum_{i=1}^{2} |x_i - y_i| \bmod 2^n = 1. \tag{6.23}$$

We use $d(x, y)$ to denote the graph distance of $x, y \in \mathcal{V}_n$. Let further $\boldsymbol{\tau} = \{\tau_x^n\}$, $x \in \mathcal{V}_n$, $n \in \mathbb{N}$, be a collection of positive i.i.d. random variables satisfying Assumption 2.3. We consider the Bouchaud trap model, $\mathsf{BTM}(G_n, \boldsymbol{\tau}, 0)$ with $\nu = 1/4$.

Theorem 6.5 (Aging on the torus). *Let $t_w(n) = 2^{2n/\alpha}n^{1-(\gamma/\alpha)}$ with $\gamma \in (0, 1/6)$. Then for $\mathbb{P}$-a.e. realisation of the random environment $\boldsymbol{\tau}$*

$$\lim_{n \to \infty} R(t_w(n), (1+\theta)t_w(n); \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1+\theta). \tag{6.24}$$

The theorem follows from the following proposition whose proof can be found in [BČ06a].

Proposition 6.6. *For any θ there exist m large enough such that Conditions 1–6 hold for*

$$\begin{aligned} t_w(n) &= 2^{2n/\alpha}n^{1-\gamma/\alpha}, & \xi_n &= m2^{2n}n^{1-\gamma}, \\ T_\varepsilon^M(n) &= \{x \in \mathcal{V}_n : \tau_x \in (\varepsilon, M)2^{2n/\alpha}n^{-\gamma/\alpha}\}, \end{aligned} \tag{6.25}$$

The main motivation for Theorem 6.5 was to extend the range of aging scales on $\mathbb{Z}^2$ and mainly to really explore the extreme values of the random landscape. Namely, the BTM on the whole lattice $\mathbb{Z}^2$ does not find the deepest traps that are close to its starting position. In the first 2^{2n} steps, it gets to the distance 2^n and visits $O(2^{2n}/\log(2^{2n})) = O(2^{2n}/n)$ sites. Therefore, the deepest visited trap has a depth of order $2^{2n/\alpha}/n^{1/\alpha}$, which is much smaller that the depth of the deepest trap in the disk with radius 2^n, that is $2^{2n/\alpha}$. Eventually, the process visits also this deepest trap, however it will be too late. This trap will no longer be relevant for the time change since much deeper traps will have already be visited. The deepest trap is relevant only if the random walk stays in the neighbourhood of its starting point a long enough time. One way to force it to stay is to change $\mathbb{Z}^2$ to the torus. By changing the size of the torus relatively to the number of considered steps, i.e. by changing γ, different depth scales become relevant for aging.

The range of possible values of $\gamma \in (0, 1/6)$ has, as in the REM case, a natural bound and an artificial one. It is natural that $\gamma < 0$ cannot be considered, since if the simple walk makes more than $2^{2n}(\log 2^n)^{1+\varepsilon}$ steps, $\varepsilon > 0$, inside the torus of size 2^n, its occupation probabilities are very close to the uniform measure on the torus, that is the process is almost in equilibrium. The other bound, $\gamma = 1/6$, comes from the techniques that we use. We do not believe it to be meaningful since we expect the theorem to hold for all $\gamma > 0$. Actually, the result for $\gamma > 1$ follows easily from the Theorem 4.5 for the whole lattice. In this case the size of the torus is much larger than ξ_n^2. So that, the process has no time to discover the difference between the torus and $\mathbb{Z}^2$. We also know that Theorem 6.5 holds also in the window $[1/6, 1]$ since it can be proved by the same methods as for $\mathbb{Z}^2$, [BČM06]. Nevertheless the complete proof in this window has never been written.

The $\gamma = 0$ case corresponds to the longest possible time scales. We expect that a similar result as Theorem 6.4 is valid.

The proof of Proposition 6.6 uses again Matthews' results: the following equivalent of Proposition 6.3 can be proved.

Proposition 6.7. *(i) Let $A_n \subset \mathcal{V}_n$ be such that $|A_n| = \rho_n 2^{2n}$ with the density ρ_n satisfying $\lim_{n\to\infty} 2^{2n} n^{-\gamma} \rho_n = \rho$ for some $\gamma \in (0, 1/6)$ and $\rho \in (0, \infty)$. Let further A_n satisfy the minimal distance condition*

$$\min\{d(x, y) : x, y \in A_n\} \geq 2^n n^{-\kappa}, \tag{6.26}$$

for some $\kappa > 0$. Then, for $\mathcal{K} = (2\log 2)^{-1}$,

$$\lim_{n\to\infty} \max_{x\in A_n} \left| \mathbb{E}_x\left[\exp\left(-\frac{s}{2^{2n} n^{1-\gamma}} H(A_n \setminus \{x\}) \right) \right] - \frac{\mathcal{K}\rho}{s + \mathcal{K}\rho} \right| = 0. \tag{6.27}$$

(ii) If A_n are, in addition, random clouds with the densities given above, then the minimal distance condition is a.s. satisfied for all n large.

Appendix. Subordinators

We use frequently the theory of increasing Lévy processes in these notes. We summarise in this appendix the facts that are important for us. For a complete treatment of this theory the reader is referred to the beautiful book by Bertoin [Ber96].

Definition A.1. We say that V is a Lévy process if for every $s, t \geq 0$, the increment $V(t+s) - V(t)$ is independent of the process $(V(u), 0 \leq u \leq t)$ and has the same law as $V(s)$. That means in particular $V(0) = 0$.

We work only with the class of increasing Lévy processes, so called *subordinators*. There is a classical one-to-one correspondence between subordinators and the set of pairs $(\mathtt{d}, \mu)$, where $\mathtt{d} \geq 0$ and μ is a measure on $(0, \infty)$, satisfying

$$\int_0^\infty (1 \wedge x)\mu(\mathrm{d}x) < \infty. \tag{A.1}$$

The law of a subordinator is uniquely determined by the Laplace transform of $V(t)$,

$$\mathbb{E}[e^{-\lambda V(t)}] = e^{-t\Phi(\lambda)}, \tag{A.2}$$

where the *Laplace exponent*

$$\Phi(\lambda) = \mathtt{d} + \int_0^\infty (1 - e^{-\lambda x})\mu(\mathrm{d}x). \tag{A.3}$$

The constant $\mathtt{d}$ corresponds to the deterministic constant drift. All processes appearing in these notes have no drift, therefore we suppose always $\mathtt{d} \equiv 0$. The measure μ is called the *Lévy measure* of the subordinator V.

There are two important families of subordinators. The first consists of the stable subordinators. A subordinator is stable with index $\alpha \in (0, 1)$ if for some $c > 0$ its Laplace exponent satisfies

$$\Phi(\lambda) = c\lambda^\alpha = \frac{c\alpha}{\Gamma(1-\alpha)} \int_0^\infty (1 - e^{-\lambda x})x^{-1-\alpha}\,\mathrm{d}x. \tag{A.4}$$

Here Γ is the usual Gamma-function.

The second important family of subordinators are the compound Poisson processes. They correspond to finite Lévy measures, $\mu((0, \infty)) < \infty$. In this case V can be constructed from a Poisson point process J on $(0, \infty)$ with constant intensity $Z_\mu := \mu((0, \infty))$ and a family of i.i.d. random variables s_i with marginal $Z_\mu^{-1}\mu$ as follows. Let $J = \{x_i, i \in \mathbb{N}\}$, $x_1 < x_2 < \ldots$, and $x_0 = 0$. Then V is the process with $V(0) = 0$ that is constant on all intervals (x_i, x_{i+1}), and at x_i it jumps by s_i, i.e. $V(x_i) - V(x_i-) = s_i$.

We need to deduce convergence of subordinators from the convergence of Lévy measures.

Lemma A.2. *Let V_n be subordinators with Lévy measures μ_n. Suppose that the sequence μ_n converges weakly to some measure μ satisfying* (A.1). *Then V_n converge to V weakly in the Skorokhod topology on $D = D([0, T), \mathbb{R})$ for all final instants $T > 0$.*

Sketch of the proof. To check the convergence on the space of cadlag path D endowed with Skorokhod topology, it is necessary check two facts: (a) the convergence of finite-dimensional distributions, and (b) tightness. To check (a) it is sufficient to look at distributions at one fixed time, since V_n have independent, stationary increments. From the weak convergence of μ_n it follows that for all $\lambda > 0$,

$$\mathbb{E}[e^{-\lambda V_n(t)}] = e^{-t \int (1-e^{-\lambda x})\mu_n(\mathrm{d}x)} \xrightarrow{n \to \infty} e^{-t \int (1-e^{-\lambda x})\mu(\mathrm{d}x)} = \mathbb{E}[e^{-\lambda V(t)}], \qquad \text{(A.5)}$$

which implies the weak convergence of $V_n(t)$. Since V_n are increasing, to check the tightness it is sufficient to check the tightness of $V_n(T)$, which is equivalent to

$$\lim_{\lambda \to 0} \lim_{n \to \infty} \mathbb{E}[e^{-\lambda V_n(T)}] = 1. \qquad \text{(A.6)}$$

This is easy to verify using the weak convergence of μ_n and the validity of (A.1) for μ. $\square$

Definition A.3 (The generalised arcsine distributions)**.** For any $\alpha \in (0,1)$, the *generalised arcsine distribution with parameter α* is the distribution on $[0,1]$ with density

$$\frac{\sin \alpha \pi}{\pi} u^{\alpha-1}(1-u)^{-\alpha}. \qquad \text{(A.7)}$$

We use Asl_α to denote its distribution function,

$$\mathsf{Asl}_\alpha(u) := \int_0^u \frac{\sin \alpha \pi}{\pi} u^{\alpha-1}(1-u)^{-\alpha}\,\mathrm{d}u, \qquad u \in [0,1]. \qquad \text{(A.8)}$$

Note that $\mathsf{Asl}_\alpha(z) = \pi^{-1}\sin(\alpha\pi)B(z; \alpha, 1-\alpha)$ where $B(z; a, b)$ is the incomplete Beta function. It is easy to see that

$$\mathsf{Asl}_\alpha(z) = 1 - O((1-z)^{1-\alpha}), \qquad \text{as } z \to 1. \qquad \text{(A.9)}$$

The following fact is crucial for us.

Proposition A.4 (The arcsine law)**.** *Let V be an α-stable subordinator and let $T(x) = \inf\{t : V(t) > x\}$. Then the random variable $V(T(x)-)/x$ has the generalised arcsine distribution with parameter α.*

This proposition has an important corollary.

Corollary A.5. *The probability that the α-stable subordinator V jumps over interval $[a, b]$ (i.e. there is no $t \in \mathbb{R}$ such that $V(t) \in [a, b]$) is equal to*

$$\mathbb{P}[V(T(b)-) < a] = \mathsf{Asl}_\alpha(a/b). \qquad \text{(A.10)}$$

Sketch of the proof of Proposition A.4. Consider the potential measure U of the subordinator V,

$$U(A) = \mathbb{E}\Big[\int_0^\infty \mathbb{1}\{V(t) \in A\}\,\mathrm{d}t\Big]. \qquad \text{(A.11)}$$

Its Laplace transform is given by

$$\int_0^\infty e^{-\lambda x} U(\mathrm{d}x) = \mathbb{E}\Big[\int_0^\infty e^{-\lambda V(t)}\,\mathrm{d}t\Big] = \frac{1}{\Phi(\lambda)}. \tag{A.12}$$

Define further $\bar{\mu}(x) = \mu((x,\infty))$. Then

$$\int_0^\infty e^{-\lambda x}\bar{\mu}(x)\,\mathrm{d}x = \Phi(\lambda)/\lambda. \tag{A.13}$$

Fix $x > 0$. For every $0 \le y \le x < z$, we can write

$$\mathbb{P}[V(T(x)-) \in \mathrm{d}y, V(T(x)) \in \mathrm{d}z] = U(\mathrm{d}y)\mu(\mathrm{d}z - y). \tag{A.14}$$

(For a proof of this intuitively obvious claim see p. 76 of [Ber96].) Define now $A_t(x) = x^{-1}V(T(tx)-)$ and consider its "double" Laplace transform

$$\tilde{A}(q,\lambda) = \int_0^\infty e^{-qt}\mathbb{E}[\exp(-\lambda A_t(x))]\,\mathrm{d}t. \tag{A.15}$$

This Laplace transform can be explicitly calculated. Indeed using (A.12) and (A.13) we obtain

$$\begin{aligned}
\tilde{A}(q,\lambda) &= \int_0^\infty e^{-qt}\int_0^{tx} e^{-\lambda y/x}\bar{\mu}(tx - y)U(\mathrm{d}y)\,\mathrm{d}t \\
&= \int_0^\infty \int_0^\infty e^{-\lambda y/x}e^{-q(s+y)/x}\bar{\mu}(s)x^{-1}\,U(\mathrm{d}y)\,\mathrm{d}s \\
&= \frac{1}{x}\cdot\frac{1}{\Phi((\lambda + q)/x)}\cdot\frac{\Phi(q/x)}{q/x} = \frac{\Phi(q/x)}{q\Phi((q+\lambda)/x)}.
\end{aligned} \tag{A.16}$$

Using that V is α-stable, i.e. $\Phi(x) = cx^\alpha$, we get

$$\tilde{A}(q,\lambda) = \frac{q^{\alpha-1}}{(q+\lambda)^\alpha} = \int_0^\infty \int_0^t e^{-qt}e^{-\lambda s}\frac{s^{\alpha-1}(t-s)^{-\alpha}}{\Gamma(\alpha)\Gamma(1-\alpha)}\,\mathrm{d}s\,\mathrm{d}t. \tag{A.17}$$

Observing that $\pi^{-1}\sin(\alpha\pi) = (\Gamma(\alpha)\Gamma(1-\alpha))^{-1}$ yields the claim of the proposition. $\square$

References

[BBM05] G. Ben Arous, L. V. Bogachev, and S. A. Molchanov, *Limit theorems for sums of random exponentials*, Probab. Theory Related Fields **132** (2005), 579–612.

[BBG02] G. Ben Arous, A. Bovier, and V. Gayrard, *Aging in the random energy model*, Physical Review Letters **88** (2002), no. 8, 087201.

[BDG01] G. Ben Arous, A. Dembo, and A. Guionnet, *Aging of spherical spin glasses*, Probab. Theory Related Fields **120** (2001), no. 1, 1–67.

[BDG06]   G. Ben Arous, A. Dembo, and A. Guionnet, *Cugliandolo-Kurchan equations for dynamics of spin-glasses*, to appear in Probab. Theory Relelated Fields (2006).

[BG97]   G. Ben Arous and A. Guionnet, *Symmetric Langevin spin glass dynamics*, Ann. Probab. **25** (1997), no. 3, 1367–1422.

[Ben02]   G. Ben Arous, *Aging and spin-glass dynamics*, Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002) (Beijing), Higher Ed. Press, 2002, pp. 3–14.

[BBG03a]   G. Ben Arous, A. Bovier, and V. Gayrard, *Glauber dynamics of the random energy model. I. Metastable motion on the extreme states*, Comm. Math. Phys. **235** (2003), no. 3, 379–425.

[BBG03b]   G. Ben Arous, A. Bovier, and V. Gayrard, *Glauber dynamics of the random energy model. II. Aging below the critical temperature*, Comm. Math. Phys. **236** (2003), no. 1, 1–54.

[BČ05]   G. Ben Arous and J. Černý, *Bouchaud's model exhibits two aging regimes in dimension one*, Ann. Appl. Probab. **15** (2005), no. 2, 1161–1192.

[BČ06a]   G. Ben Arous and J. Černý, *The arcsine law as a universal aging scheme for trap models*, preprint, 2006.

[BČ06b]   G. Ben Arous and J. Černý, *The scaling limit of d-dimensional Bouchaud's model*, in preparation, 2006.

[BČM06]   G. Ben Arous, J. Černý, and T. Mountford, *Aging for Bouchaud's model in dimension two*, Probab. Theory Related Fields **134** (2006), no. 1, 1–43.

[BB03]   E. Bertin and J.-P. Bouchaud, *Sub-diffusion and localisation in the one dimensional trap model*, Physical Review E **67** (2003), 026128.

[Ber96]   J. Bertoin, *Lévy processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996.

[Bou92]   J.-P. Bouchaud, *Weak ergodicity breaking and aging in disordered systems*, J. Phys. I (France) **2** (1992), 1705–1713.

[BCKM98]   J.-P. Bouchaud, L. Cugliandolo, J. Kurchan, and M. Mézard, *Out of equilibrium dynamics in spin-glasses and other glassy systems*, Spin glasses and random fields (A. P. Young, ed.), World Scientific, Singapore, 1998.

[BD95]   J.-P. Bouchaud and D. S. Dean, *Aging on Parisi's tree*, J. Phys I(France) **5** (1995), 265.

[BG90]   J.-P. Bouchaud and A. Georges, *Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications*, Phys. Rep. **195** (1990), no. 4–5, 127–293.

[BF06]   A. Bovier and A. Faggionato, *Spectral analysis of Sinai's walk for small eigenvalues*, WIAS preprint 1056, 2006.

[BF05]   A. Bovier and A. Faggionato, *Spectral characterization of aging: the REM-like trap model*, Ann. Appl. Probab. **15** (2005), no. 3, 1997–2037.

[Čer03]   J. Černý, *On two properties of strongly disordered systems, aging and critical path analysis*, Ph.D. thesis, EPF Lausanne, 2003.

[Čer06]   J. Černý, *The behaviour of aging functions in one-dimensional Bouchaud's trap model*, Communications in Mathematical Physics **261** (2006), no. 1, 195–224.

[CHS93]   H. Crisanti, H. Horner, and H.-J. Sommers, *The spherical p-spin interaction spin glass model: the dynamics*, Zeitschrift für Physik B **92** (1993), 257–271.

[CK93]   L. F. Cugliandolo and J. Kurchan, *Analytical solution of the off-equilibrium dynamics of a long-range spin-glass*, Phys. Rev. Lett. **71** (1993), 173.

[DGZ01]   A. Dembo, A. Guionnet, and O. Zeitouni, *Aging properties of Sinai's random walk in random environment*, arXiv, math.PR/0105215, 2001.

[Der81]   B. Derrida, *Random Energy Model: An exactly solvable model of disordered systems*, Phys. Rev. B **24** (1981), no. 5, 2613–2626.

[DM76]   H. Dym and H. P. McKean, *Gaussian processes, function theory, and the inverse spectral problem*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976, Probability and Mathematical Statistics, Vol. 31.

[FIN99]   L. R. G. Fontes, M. Isopi, and C. M. Newman, *Chaotic time dependence in a disordered spin system*, Probab. Theory Related Fields **115** (1999), no. 3, 417–443.

[FIN02]   L. R. G. Fontes, M. Isopi, and C. M. Newman, *Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension*, Ann. Probab. **30** (2002), no. 2, 579–604.

[Gui97]   A. Guionnet, *Averaged and quenched propagation of chaos for spin glass dynamics*, Probab. Theory Related Fields **109** (1997), no. 2, 183–215.

[IM65]   K. Itô and H. P. McKean, Jr., *Diffusion processes and their sample paths*, Die Grundlehren der Mathematischen Wissenschaften, Band 125, Academic Press Inc., Publishers, New York, 1965.

[KS79]   H. Kesten and F. Spitzer, *A limit theorem related to a new class of self-similar processes*, Z. Wahrsch. Verw. Gebiete **50** (1979), no. 1, 5–25.

[KW82]   S. Kotani and S. Watanabe, *Kreĭn's spectral theory of strings and generalized diffusion processes*, Functional analysis in Markov processes (Katata/Kyoto, 1981), Lecture Notes in Math., vol. 923, Springer, Berlin, 1982, pp. 235–259.

[Mac85]   J. Machta, *Random walks on site disordered lattices*, J. Phys. A: Math. Gen. **18** (1985), no. 9, L531–L538.

[Mat88]   P. Matthews, *Covering problems for Markov chains*, Ann. Probab. **16** (1988), no. 3, 1215–1228.

[MS04]   M. M. Meerschaert and H.-P. Scheffler, *Limit theorems for continuous-time random walks with infinite mean waiting times*, J. Appl. Probab. **41** (2004), no. 3, 623–638.

[MB97]   R. Mélin and P. Butaud, *Glauber dynamics and ageing*, J. de Physique I **7** (1997), 691–710.

[Mon04]   C. Monthus, *Nonlinear response of the trap model in the aging regime: Exact results in the strong-disorder limit*, Phys. Rev. E **69** (2004), 026103.

[MB96]   C. Monthus and J.-P. Bouchaud, *Models of traps and glass phenomenology*, J. Phys. A **29** (1996), 3847–3869.

[MW65]   E. W. Montroll and G. H. Weiss, *Random walks on lattices. II*, Journal of Mathematical Physics **6** (1965), no. 2, 167–181.

[RMB00]   B. Rinn, P. Maass, and J.-P. Bouchaud, *Multiple scaling regimes in simple aging models*, Phys. Rev. Lett **84** (2000), 5403–5406.

[RMB01]   B. Rinn, P. Maass, and J.-P. Bouchaud, *Hopping in the glass configuration space: Subaging and generalized scaling laws*, Phys. Rev. B **64** (2001), 104417.

[Sto63]   C. Stone, *Limit theorems for random walks, birth and death processes, and diffusion processes*, Illinois J. Math. **7** (1963), 638–660.

[Tal03]   M. Talagrand, *Spin glasses: a challenge for mathematicians*, Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics, vol. 46, Springer-Verlag, Berlin, 2003, Cavity and mean field models.

[Zas02]   G. Zaslavsky, *Chaos, fractional kinetics, and anomalous transport*, Physics Reports **371** (2002), 461–580.