Another view on aging in the REM

Jiří Černý

ETH Zürich, Departement Mathematik, Rämistrasse 101, 8092 Zürich, Switzerland
jiri.cerny@math.ethz.ch

Summary. We give a new proof of aging for a version of a Glauber dynamics in the Random Energy Model. The proof uses ideas that were developed in [BBČ07] for studying the dynamics of a p-spin Sherrington-Kirkpatrick spin glass.

1 Introduction

Aging was proved recently for a simple Glauber-type dynamics of p-spin spin glasses, [BBČ07]. In order to overcome the difficulties stemming from the correlations of the Hamiltonian of these spin glasses, several new ideas were introduced. They allowed to show that the aging behaviour of the p-spin spin glass is essentially the same as the one of the Random Energy Model (REM), at least at some time scales.

In the present paper, we take a step back and apply these ideas to the REM. A new proof of aging in the REM, even if it be shorter than the older proofs, is, however, not the principal objective of this paper. Rather, it is written as a different presentation of ideas of [BBČ07], uncluttered from rather heavy computations that were necessary for the p-spin spin glass.

Let us start with a brief summary of the efforts that lead to [BBČ07]. Aging in spin glasses was for the first time observed experimentally in the beginning of the 1980's. In order to explain the observations, trap models were introduced by Bouchaud [Bou92, BD95] in the physics literature. Trap models are effective models which can be solved analytically using simple renewal arguments, and which nevertheless reproduce the characteristic power-law behaviour observed experimentally. While trap models are heuristically motivated to capture the behaviour of the dynamics of spin glass models, there is no clear theoretical, let alone mathematical, derivation of them from an underlying spin-glass dynamics.

The first steps to provide a such derivation were taken in [BBG03a, BBG03b] for a version of a Glauber dynamics in the REM. Technically very elaborate renewal arguments were used to prove aging and the relevance of the trap model ansatz in this case. The core of the argument is an analysis of visits of the dynamics to a finite set of extremes of the Hamiltonian. In order to allow the dynamics to discover these extremes, the time scales studied in these papers should be carefully fixed to be only slightly shorter than the equilibration scale. In particular, it means that the considered time scales should increase exponentially with the size of the system.

Another proof of aging in the REM, for a dynamics close to this of [BBG03a], was given in [BČ08]. It was inspired by methods developed for trap models on $\mathbb{Z}^d$ [BČM06]. The main idea of the approach of [BČ08] is to study the extremes of the Hamiltonian along the trajectory of the dynamics instead of concentrating on visits of dynamics to the extremes of the Hamiltonian. Apart from having some slight technical advantages, the new point of view gave more freedom in choosing the time scales. In [BČ08] time scales much shorter than the equilibration scale, but still increasing exponentially with the size of the system, were studied. The techniques can be however easily extended to apply also to the scales considered in [BBG03a].

The main technical tool of [BČ08] is the so-called *clock process*, which is, roughly speaking, a process that records the time needed for a given number of jumps of the dynamics (see (4) for the exact definition). It was argued that this process converges, after a proper rescaling, to a stable subordinator in many situations. From the convergence to a subordinator, aging can be deduced using the classical arc-sine law. Even if techniques and time scales are different, the obtained aging results are essentially the same as those predicted by the trap models.

Both above mentioned studies of the REM used substantially the fact that the Hamiltonian is particularly simple: it is a collection of i.i.d. random variables. This ceases to be true for 'more realistic' mean-field spin glasses, like Sherrington-Kirkpatrick model or p-spin spin glass. For statics of the spin glass the correlation between the energies impose that (at low temperature) the main contribution to the Gibbs measure does not come from a finite number of distant configurations as in the REM, but from a finite number of distant 'valleys', which, however, contain many configurations.

Interestingly, in [BBČ07] it was showed that the global behaviour of the clock process is not influenced by the correlations in the dynamics of p-spin spin glass with $p \geq 3$. The rescaled clock process hardly feels the correlations and large valleys and it converges to a stable subordinator, confirming the universality of this behaviour. However, this convergence can be proved only at the expense of restricting the range of time scales to the lower part of the range of [BČ08], remaining far from the equilibration time scale (see (14) for details).

Even if the asymptotic behaviour of the clock process remains unchanged, the non-i.i.d. character of the Hamiltonian disallows a direct application of methods of [BČ08]. New techniques used in [BBČ07] are based on ideas from extremal theory and exploit strongly the Gaussian character of the Hamiltonian. It should be remarked that this property of the Hamiltonian had been hardly used in the studies of the REM dynamics. E.g., the only Gaussian ingredient in [BČ08] is the standard asymptotic expression for the tail probability. The results of the [BČ08] can thus be extended to a much larger class of distributions of the Hamiltonian.

In this paper, we take full advantage of the Gaussian distribution of the Hamiltonian and we use the methods of [BBČ07] to provide new and relatively short proof of aging in the REM on time scales that are much shorter than the equilibration scale. Aside from this proof, we use the occasion to explicitly write out the modifications which are necessary to obtain a result on the longest possible time scales of [BBG03b].

To close the introduction, let us remark that the time scales of this paper (and also of [BBG03a, BČ08]) are 'much longer' than those used in the studies of the Langevin dynamics of soft-spin models [CK93, BG97, BDG06], where one considers the infinite volume limit at fixed time t, and then analyzes the ensuing dynamics as t tends to infinity. In this paper, the time depends exponentially on the size of the system and both tend to infinity together.

2 Model and results

We will study the same dynamics of the REM as in [BČ08]. It is defined as follows. Let $\mathcal{S}_N \equiv \{-1,1\}^N$ be a N-dimensional hypercube equipped with a distance

$$\mathrm{dist}(\sigma,\tau) = \frac{1}{2}\sum_{i=1}^{N} |\sigma_i - \tau_i|, \qquad \sigma,\tau \in \mathcal{S}_N. \tag{1}$$

The Hamiltonian of the REM is defined as $\sqrt{N}H_N$, where $H_N : \mathcal{S}_N \to \mathbb{R}$ is a centred i.i.d. Gaussian process on $\mathcal{S}_N$ with variance $\mathbb{E}[H_N(\sigma)^2] = 1$. We will use $\mathcal{H}$ to denote the σ-algebra generated by $\{H_N(\sigma), \sigma \in \mathcal{S}_N, N \in \mathbb{N}\}$. The corresponding Gibbs measure is given by

$$\mu_{\beta,N}(\sigma) \equiv Z_{\beta,N}^{-1} e^{\beta\sqrt{N}H_N(\sigma)}. \tag{2}$$

We consider a nearest-neighbour continuous-time Markov dynamics $\sigma_N = (\sigma_N(t), t \ge 0)$ on $\mathcal{S}_N$ which is given by its transition rates

$$w_N(\sigma,\tau) = \begin{cases} N^{-1}e^{-\beta\sqrt{N}H_N(\sigma)}, & \text{if } \mathrm{dist}(\sigma,\tau) = 1,\\ 0, & \text{otherwise.}\end{cases} \tag{3}$$

Obviously, σ_N is reversible with respect to the Gibbs measure $\mu_{\beta,N}$.

It is an important property, that this dynamics can be constructed as a time change of a simple random walk on $\mathcal{S}_N$: Let $(Y_N(k), k \in \mathbb{N})$ be the simple discrete-time random walk (SRW) on $\mathcal{S}_N$ started at some fixed point of $\mathcal{S}_N$, say at $\mathbf{1} = \{1, \dots, 1\}$. For $\beta > 0$ and $k \in \mathbb{N}$ we define the *clock-process* by

$$S_N(k) = \sum_{i=0}^{k-1} e_i \exp\left\{\beta\sqrt{N} H_N\big(Y_N(i)\big)\right\}, \tag{4}$$

where $(e_i, i \in \mathbb{N})$ is a sequence of mean-one i.i.d. exponential random variables. The process σ_N can be then written as

$$\sigma_N(t) \equiv Y_N(S_N^{-1}(t)). \tag{5}$$

We consider all random processes to be defined on an abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $\mathcal{Y}$ the σ-algebra generated by $\{Y_N(k), k \in \mathbb{N}, N \in \mathbb{N}\}$. The σ-algebra generated by $\{e_i, i \in \mathbb{N}\}$ will be denoted by $\mathcal{E}$. Note that the three σ-algebras $\mathcal{H}$, $\mathcal{Y}$, and $\mathcal{E}$ are all independent under $\mathbb{P}$.

For $\gamma > 0$ we define

$$\begin{aligned} r_N &= r_N(\gamma) = e^{N\gamma^2/2\beta^2}, \\ t_N &= t_N(\gamma) = N^{-\frac{1}{2\alpha}} e^{\gamma N}, \end{aligned} \tag{6}$$

and the rescaled clock process $\bar{S}_N^\gamma$ by

$$\bar{S}_N^\gamma(s) = t_N^{-1} S_N\big(\lfloor s r_N \rfloor\big), \qquad s \geq 0. \tag{7}$$

The function r_N is the *number-of-jumps* scale: We observe σ_N after making $O(r_N)$ jumps. The function t_N, the *time scale*, then gives the time that σ_N typically needs to make this number of steps. We view $\bar{S}_N^\gamma$ as an element of the space D of càdlàg functions from $[0, \infty)$ to $\mathbb{R}$ equipped with the standard Skorokhod J_1-topology.

Let $V_\alpha(t)$ be the α-stable subordinator with the Laplace transform given by

$$\mathbb{E}[e^{-\lambda V_\alpha(t)}] = \exp(-t\lambda^\alpha). \tag{8}$$

We will use $\beta_c = \sqrt{2\log 2}$ to denote the critical temperature of the REM.

The main result of this paper is the following theorem that provides the asymptotic behaviour of the clock.

Theorem 1. *For any fixed* γ *such that*

$$0 < \gamma < \min\left(\beta^2, \beta_c\beta\right), \tag{9}$$

under the conditional distribution $\mathbb{P}[\cdot|\mathcal{Y}]$*,* $\mathcal{Y}$*-a.s., the law of the stochastic process* $\bar{S}_N^\gamma$ *converges to the law of* α*-stable subordinator* $V_\alpha(\mathcal{K}\cdot)$*, where* $\alpha \equiv \gamma/\beta^2$*, and* $\mathcal{K}$ *is a constant which will be computed explicitly in Lemma 1.*

This theorem is very close to the results of [BČ08]. However, there is one important difference. In [BČ08], the convergence to the subordinator is proved under the law $\mathbb{P}[\cdot|\mathcal{H}]$, $\mathcal{H}$-a.s. The slightly non-physical conditioning on $\mathcal{Y}$ that appears in our theorem, and of course also in [BBČ07] whose methods we use, is the price to pay for having at hand Gaussian tools. As we have already remarked, the Gaussian character of the Hamiltonian was almost not exploited in the previous studies of the REM. Hence, the conditioning on $\mathcal{H}$ did not pose any problem. This conditioning, that is fixing the Gaussian disorder, is ruled out if we want to employ more advanced Gaussian techniques now.

It is an interesting open question, if it is possible to deduce the results of [BČ08] from Theorem 1 without using too much the properties of the REM. This could allow to prove the convergence under $\mathbb{P}[\cdot|\mathcal{H}]$ also for the p-spin spin glass.

The next theorem will be used to prove aging on the longest possible time scales, that means on time scales of [BBG03b].

Theorem 2. *Let $\beta > \beta_c$ and $\gamma = \beta\beta_c$. Then for all $n \geq 1$ finite, for all $0 \leq s_1 < \cdots < s_n$ and $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R})$*

$$\lim_{\xi\to 0}\lim_{N\to\infty} \mathbb{P}\Big[\bigcap_{i=1}^{n} \xi^{-1/\alpha}\bar{S}_N^{\gamma}(\xi s_i) \in A_i \Big| \mathcal{Y}\Big] = \mathbb{P}\Big[\bigcap_{i=1}^{n} V_\alpha(\mathcal{K}s_i) \in A_i\Big] \tag{10}$$

in probability.

As a consequence of Theorems 1 and 2 we get the following aging result

Theorem 3. *(a) Under the hypotheses of Theorem 1, for all $\theta > 1$, $\mathcal{Y}$-a.s.,*

$$\lim_{N\to\infty} \mathbb{P}[\sigma_N(t_N) = \sigma_N(\theta t_N)|\mathcal{Y}] = \mathsf{Asl}_\alpha(\theta), \tag{11}$$

where $\mathsf{Asl}_\alpha(\theta) = \frac{\sin(\alpha\pi)}{\pi}\int_0^{1/\theta} u^{\alpha-1}(1-u)^{-\alpha}\mathrm{d}u$,

(b) Under the hypotheses of Theorem 2, for all $\theta > 1$, in probability,

$$\lim_{\xi\to 0}\lim_{N\to\infty} \mathbb{P}[\sigma_N(\xi t_N) = \sigma_N(\xi\theta t_N)|\mathcal{Y}] = \mathsf{Asl}_\alpha(\theta). \tag{12}$$

Claim (b) of the last theorem relates to the aging result obtained in [BBG03b] just as claim (a) relates to [BČ08]: the role of σ-algebras $\mathcal{H}$ and $\mathcal{Y}$ is inverted.

Finally, let us compare our result with the result of [BBČ07] for the p-spin spin glass. The definition of the dynamics considered there is the same as in this paper. The only change is, of course, the Hamiltonian, which is given by a centred Gaussian process on $\mathcal{S}_N$ with covariance

$$\mathrm{Cov}(H_N(\sigma), H_N(\tau)) = \Big(\frac{1}{N}\sum_{i=1}^{N}\sigma_i\tau_i\Big)^p. \tag{13}$$

With two replacements, Theorem 1 holds in this case: First, condition (9) should be replaced by $p \geq 3$ and

$$\gamma < \min\left(\beta^2, \zeta(p)\beta\right), \tag{14}$$

where $\zeta(p)$ is an increasing function strictly smaller than β_c which converges to β_c as $p \to \infty$. Second, the space D should be equipped with a weaker topology.

Condition (14) implies that the longest time scales we are able to treat in the p-spin model are shorter than for the REM. We do not know what happens in the p-spin model on longer scales. As $p \to \infty$ (14) approaches (9), which is not surprising, since the REM can be considered as p-spin model with $p = \infty$.

Let us close this section by a rough description of the techniques which are used to prove Theorem 1. The behaviour of the rescaled clock process $\bar{S}_N^\gamma$ is determined by the energies of spin configurations that are visited during the first $O(r_N)$ steps. The energies of visited configurations form a Gaussian process $X_N^0(k) \equiv H_N(Y_N(k))$ which has random covariance $\mathrm{Cov}(X_N^0(k), X_N^0(j)) = \mathbf{1}\{Y_N(k) = Y_N(j)\}$. We are interested mainly in the visited configurations whose energy is very large, because these configurations contribute a lot to the clock process. We want thus to know how extremes of a Gaussian process X_N^0 with random correlation structure behave.

The standard method how to study such extremes is to replace the complicated Gaussian process with a simpler process, the behaviour of whose extremes can be determined more easily and whose correlation structure locally approximates well the correlation structure of the original process. A study of the behaviour of extremes of the original process then breaks into two parts. First, the behaviour of extremes of the simple process should be determined. Second, it should be proved that the approximation by the simple process is reasonable.

In the case of the REM, that is of Gaussian process X_N^0, the approximating process will be particularly simple. We define $(X^1(k), k \in \mathbb{N}, N \in \mathbb{N})$ as an i.i.d. sequence of standard Gaussian random variables. It is clearly a natural choice, since the simple random walk on the hypercube has a very small probability to return to an already visited configuration. More precisely, for any j fixed, $\mathbb{P}[\exists k : r_N > k > j, Y_N(k) = Y_N(j)] \sim 1/N$ (at least if $\gamma \le \beta\beta_c$).

In the first step of the proof we will thus analyse extremes of an i.i.d. sequence, which is, of course, quite simple. This is done in Section 3 using a method that allows control the clock process immediately. The second step of the proof, that is the verification if the approximation is justified, is done in Section 5. In Section 4 we collect several estimates on the simple random walk on the hypercube. Finally, all theorems are proved in Section 6.

3 Sum of i.i.d. exponentials

As explained in the last section, we will compare the clock process with the sum of i.i.d. random variables with the same distribution, that is with the sum

$$\tilde{S}_N^{\gamma}(s) = t_N^{-1} \sum_{i=0}^{\lfloor sr_N \rfloor} e^{\beta\sqrt{N} X_N^1(i)}, \tag{15}$$

Sums of this type were exhaustively studied in [BBM05]. The results of this paper imply directly that $\tilde{S}_N^{\gamma}$ converges to an γ/β^2 stable subordinator. For the sake of completeness we will provide here a simple proof of this claim. We start by an easy lemma.

Lemma 1. *For all β, γ satisfying hypotheses of Theorems 1 or 2 there exists $\mathcal{K} = \mathcal{K}_{\beta,\gamma} > 0$ such that, for r_N and t_N as in* (6) *and $u > 0$,*

$$\lim_{N\to\infty} r_N \Big(1 - \mathbb{E}\Big[\exp\Big\{ - \frac{u}{t_N} e_i e^{\beta\sqrt{N} X_N^1(i)} \Big\}\Big]\Big) = \mathcal{K} u^{\alpha}. \tag{16}$$

This lemma should be viewed as a 'large deviation' statement, since values of $X_N^1(i)$ that give the largest contribution to the Laplace transform in (16) differ significantly from the typical ones. From large deviation point of view, the proof below is simply a tilting of the Gaussian distribution. This tilting makes the most contributing values typical.

Proof. Recall that $\alpha = \gamma/\beta^2$. To save the notation we set $X_i = X_N^1(i)$ and define $g(x) = \ln(1+x)$. Performing the expectation over e_i we get

$$1 - \mathbb{E}\Big[\exp\Big\{ - \frac{ue_i}{t_N} e^{\beta\sqrt{N} X_i} \Big\}\Big] = \int_{-\infty}^{\infty} \frac{\mathrm{d}x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big(1 - \exp\Big\{ - g\Big(\frac{u}{t_N} e^{\beta\sqrt{N}x}\Big)\Big\}\Big). \tag{17}$$

We now tilt the measure. Setting $x = (\beta z + \log t_N - \log u)/(\beta\sqrt{N})$ we find that (17) equals

$$u^{\alpha} r_N^{-1} \int_{-\infty}^{\infty} \frac{\mathrm{d}z}{\sqrt{2\pi}} e^{-z^2/2N} \, \frac{e^{-\gamma z/\beta} e^{\beta z}}{1 + e^{\beta z}} e^{z f_N(\beta,\gamma,u)} \, (1 + o(1)). \tag{18}$$

where $f_N(\beta,\gamma,u) = (\beta \log N)/(2\gamma N) + \log u/(\beta N) \to 0$ as $N \to \infty$. The integrand converges to $e^{(1-\alpha)\beta z}(1 + e^{\beta z})^{-1}$ which decays exponentially as z tends both to ∞ and $-\infty$, since $\beta/\gamma^2 < 1$. An application of the dominated convergence theorem then yields the convergence of the integral to a positive constant $\mathcal{K}$ independent of u. □

By consequence, we get the analogue of Theorem 1 for the process X_N^1:

Proposition 1. *For all β, γ satisfying hypotheses of Theorems 1 or 2, the sequence of processes $\tilde{S}_N^{\gamma}(s)$ converges to the stable subordinator $V_\alpha(\mathcal{K}s)$ weakly in the Skorokhod J_1-topology.*

Proof. To check the convergence of finite-dimensional marginals from Lemma 1 is trivial. E.g.

$$\mathbb{E}[e^{-u\tilde{S}_N^{\gamma}(s)}] = \mathbb{E}\Big[\exp\Big\{ - \frac{u}{t_N} e_i e^{\beta\sqrt{N} X_i} \Big\}\Big]^{\lfloor sr_N \rfloor} \xrightarrow{N\to\infty} e^{-s\mathcal{K}u^{\alpha}}, \tag{19}$$

which is the Laplace transform of $V_\alpha(\mathcal{K}s)$. The proof of the tightness is an easy modification of the tightness proof for Theorem 1. We omit it therefore here. □

By simply changing the domains of integration in the last proof, the next lemma can be verified. We will need it later to check the tightness of $\bar{S}_N^\gamma$.

Lemma 2. *Let B_N^ε be such that*

$$t_N^{-1} e^{\beta\sqrt{N} B_N^\varepsilon} = \varepsilon \tag{20}$$

and let

$$f_N(\varepsilon) = r_N \Big(1 - \mathbb{E}\Big[\exp\Big\{ - t_N^{-1} e_i e^{\beta\sqrt{N} X_N^1(i)}\Big\}\mathbf{1}\{X_N^1(i) \le B_N^\varepsilon\}\Big]\Big). \tag{21}$$

Then $\lim_{\varepsilon\to 0} \limsup_{N\to\infty} f_N(\varepsilon) = 0$.

4 Random walk properties

To compare the clock processes $\tilde{S}_N^\gamma$ and $\bar{S}_N^\gamma$ we need to know to what extent the covariances of X^1 and X^0 differ. Since the non-zero covariances of X^0 rise from self-intersections of the simple random walk Y_N we should control their number.

Lemma 3. *Under the assumptions of Theorem 1 there exists $C = C(\beta,\gamma)$ such that $\mathcal{Y}$-a.s. for all but finitely many N,*

$$\sum_{i\neq j=1}^{sr_N} \mathbf{1}\{Y_N(i) = Y_N(j)\} \le CN^{-1} s r_N. \tag{22}$$

Proof. Note that the assumption (9) implies that $r_N \ll 2^N$. Let $p_k^N(x,y) = \mathbb{P}[Y_N(k) = y | Y_N(0) = x]$. To bound the sum (22) for i, j that are far from each other we use the fact that the random walk on the hypercube reaches the equilibrium very quickly. The next lemma can be proved by using the coupling argument of [Mat87]. Detailed proof is given in [BBČ07] and we will not repeat it here.

Lemma 4. *There exists K large enough such that for all $k \ge KN^2 \log N =: \mathfrak{K}(N)$ and $x, y \in \mathcal{S}_N$*

$$\left| \frac{p_k^N(x,y) + p_{k+1}^N(x,y)}{2} - 2^{-N} \right| \le 2^{-8N}. \tag{23}$$

Let $A_N = \{(i,j) : 0 \le i, j \le sr_N, |i-j| > \mathfrak{K}(N)\}$. From Lemma 4 it follows that

$$\mathbb{E}\Big[\sum_{(i,j)\in A_N} \mathbf{1}\{Y_N(i)=Y_N(j)\}\Big] \le C(sr_N)^2 2^{-N}. \tag{24}$$

Hence,

$$\mathbb{P}\Big[\sum_{(i,j)\in A_N} \mathbf{1}\{Y_N(i)=Y_N(j)\} \ge CN^{-1}sr_N\Big] \le CNsr_N 2^{-N}. \tag{25}$$

The Borel-Cantelli lemma then implies an a.s. bound for the sum over A_N.

To bound the contribution of pairs i,j with $|i-j|\le \mathfrak{K}(N)$ we need another lemma.

Lemma 5. *There exist $c_2 > c_1 > 0$ such that, for all N,*

$$\frac{c_1}{N} \le \mathbb{E}\Big[\sum_{i=1}^{\mathfrak{K}(N)} \mathbf{1}\{Y_N(i)=Y_N(0)\}\Big] \le \frac{c_2}{N}. \tag{26}$$

Proof. There are many ways how to prove this lemma. Let us sketch one of them. Let, for $A\subset \mathcal{S}_N$, $\tau_A=\min\{k\ge 1 : Y_N(k)\in A\}$, and let $B_4=\{z : d(z,\mathbf{1})=4\}$. Using the fact that $d(\mathbf{1},Y_N(k))$ is the Ehrenfest's-Urn Markov chain, one can check that $c'/N \le \mathbb{P}[\tau_{\mathbf{1}}<\tau_{B_4}|Y_N(0)=\mathbf{1}]\le c/N$. A similar argument gives also $\mathbb{P}[\tau_{\mathbf{1}}<\tau_{B_4}|Y_N(0)\in B_4]<c/N^{-4}$. Therefore to get from B_4 to 1 we need in average N^4 tries, but we have at most $\mathfrak{K}(N)=KN^2\log N$ of them. Hence, the probability of returning to $\mathbf{1}$ before $\mathfrak{K}(N)$ is smaller than c/N, which yields the claim of the lemma. □

Let $Z_i=\mathfrak{K}(N)^{-1}\sum_{j=i+1}^{i+\mathfrak{K}(N)}\mathbf{1}\{Y_N(j)=Y_N(i)\}$. Then $Z_i\in[0,1]$, and by the last lemma $c_1(N\mathfrak{K}(N))^{-1}\le \mathbb{E}[Z_i]\le c_2(N\mathfrak{K}(N))^{-1}$. Obviously,

$$\sum_{\substack{i,j=1\\|i-j|\le\mathfrak{K}(N)}}^{sr_N} \mathbf{1}\{Y_N(i)=Y_N(j)\} \le 2\mathfrak{K}(N)\sum_{k=1}^{\mathfrak{K}(N)}\sum_{j=1}^{m} Z_{j\mathfrak{K}(N)+i}, \tag{27}$$

where $m=\lceil sr_N/\mathfrak{K}(N)\rceil$. The inner sum in the last expression is an i.i.d. sum. Hoeffding's inequality [Hoe63] applied to the sequence $\{Z_i\}$ gives for any $u>0$,

$$\mathbb{P}\Big[\sum_{j=1}^{m} Z_{j\mathfrak{K}(N)+i} - m\mathbb{E}[Z_i]\ge um\Big]\le \exp\{-2mu^2\}. \tag{28}$$

Setting $u=\mathbb{E}[Z_i]$ and observing that the right-hand side of the last expression is summable even after a multiplication by $\mathfrak{K}(N)$, the Borel-Cantelli lemma and (27) imply that $\mathcal{Y}$-a.s., for all but finitely many N,

$$\sum_{\substack{i,j=1\\|i-j|\le\mathfrak{K}(N)}}^{sr_N} \mathbf{1}\{Y_N(i)=Y_N(j)\}\le CN^{-1}sr_N. \tag{29}$$

This completes the proof of Lemma 3. □

Lemma 6. *Let $s > 0$ and let for any $\omega \geq 1$*

$$I_\omega = I_\omega(N,s) = \{\sigma \in \mathcal{S}_N : \exists i_1 < \cdots < i_\omega \leq r_N s, Y_N(i_1) = \cdots = Y_N(i_\omega) = \sigma\} \tag{30}$$

be the set of configurations visited at least ω-times. Then,

(a) there exists $C > 0$ such that $\mathcal{Y}$-a.s., for all but finitely many N, for all $\omega \in \{2, \dots, N\}$

$$|I_\omega| \leq C^\omega (N^{1-\omega} r_N s \vee 1). \tag{31}$$

(b) $\mathcal{Y}$-a.s., for all but finitely many N,

$$|I_N| = 0. \tag{32}$$

Proof. The proof is very similar to the previous one: Lemma 4 can be used to show that number of elements of I_ω with $\max\{i_k - i_{k-1} : 2 \leq k \leq \omega\} \geq \mathfrak{K}(N)$ is much smaller than the right-hand side of (31). Defining $\tilde{Z}_i = \mathbf{1}\{Y_N(i) \in I_\omega, i_1 = i, i_\omega - i_1 \leq \omega \mathfrak{K}(N)\}$ and observing similarly as in Lemma 5 that $\mathbb{E}[\tilde{Z}_i] \sim cN^{1-\omega}$, the claim (a) can be proved by another application of Hoeffding's inequality.

Claim (b) follows from $\mathbb{P}[\exists i \leq r_N s : Y_N(i) \in I_N] \leq C^N s N^{-N+1} r_N$ and the Borel-Cantelli lemma. □

To prove Theorem 2, we need the following modification of Lemma 3.

Lemma 7. *Let $\gamma = \beta\beta_c$, that is $r_N = 2^N$. Then there exists a constant C such that for all $\xi < 1$*

$$\begin{aligned} &\mathbb{P}\Big[\sum_{i \neq j = 1}^{\xi r_N} \mathbf{1}\{Y_N(i) = Y_N(j)\} \geq \xi r_N\Big] \leq C\xi, \\ &\mathbb{P}[|I_\omega| \geq r_N \xi^{\omega/2}] \leq C\xi^{\omega/2}. \end{aligned} \tag{33}$$

Proof. Following the same reasoning as in the proof of Lemma 3 one can show that $\mathbb{E}\big[\sum_{i \neq j=1}^{\xi r_N} \mathbf{1}\{Y_N(i) = Y_N(j)\}\big] \leq C\xi^2 r_N$. The first claim then follows from the Markov inequality. The second claim can be obtained from $\mathbb{E}[|I_\omega|] \leq C(\xi r_N)^\omega 2^{-N(\omega-1)} \leq r_N \xi^\omega$. □

5 Comparison of two processes

We can now compare the clock process $\bar{S}_N^\gamma$ with the i.i.d. sum $\tilde{S}_N^\gamma$. We use Λ^0 and Λ^1 to denote the covariance matrices of X^0 and X^1.

$$\Lambda^0_{ij} = \mathbf{1}\{Y_N(i) = Y_N(j)\}, \qquad \Lambda^1_{ij} = \delta_{ij}. \tag{34}$$

For $h \in [0,1]$ we define the interpolating process $X_N^h(i) \equiv \sqrt{1-h}X_N^0(i) + \sqrt{h}X_N^1(i)$.

Let $\ell \in \mathbb{N}$, $0 = s_0 < \cdots < s_\ell = T$ and $u_1, \ldots, u_\ell > 0$ be fixed. For any Gaussian process X we define

$$\begin{aligned} F_N\big(X; \{s_i\}, \{u_i\}\big) &\equiv \mathbb{E}\Big[\exp\Big(-\sum_{k=1}^{\ell} \frac{u_k}{t_N} \sum_{i=s_{k-1}r_N}^{s_k r_N - 1} e_i e^{\beta\sqrt{N}X(i)}\Big)\Big| X\Big](X) \\ &= \exp\Big(-\sum_{k=1}^{\ell} \sum_{i=s_{k-1}r_N}^{s_k r_N - 1} g\Big(\frac{u_k}{t_N} e^{\beta\sqrt{N}X(i)}\Big)\Big). \end{aligned} \tag{35}$$

Note that $\mathbb{E}[F(X^0; \{s_i\}, \{u_i\})|\mathcal{Y}]$ is a joint Laplace transform of the distributions of the properly rescaled clock process at times s_i. The following proposition thus compares Laplace transforms of $\tilde{S}_N^\gamma$ and $\bar{S}_N^\gamma$.

Proposition 2. *(a) If the assumptions of Theorem 1 are satisfied, then for all sequences $\{s_i\}$ and $\{u_i\}$,*

$$\lim_{N\to\infty} \mathbb{E}\big[F_N\big(X_N^0; \{s_i\}, \{u_i\}\big)\big|\mathcal{Y}\big] - \mathbb{E}\big[F_N\big(X_N^1; \{s_i\}, \{u_i\}\big)\big] = 0, \qquad \mathcal{Y}\text{-a.s.} \tag{36}$$

(b) If the assumptions of Theorem 2 hold, then, in probability,

$$\lim_{\xi\to 0}\lim_{N\to\infty} \mathbb{E}\big[F_N\big(X_N^0; \{\xi s_i\}, \{\xi^{-\frac{1}{\alpha}} u_i\}\big)\big|\mathcal{Y}\big] - \mathbb{E}\big[F_N\big(X_N^1; \{\xi s_i\}, \{\xi^{-\frac{1}{\alpha}} u_i\}\big)\big] = 0. \tag{37}$$

Proof. We use the well-known interpolation formula for functionals of two Gaussian processes due (probably) to Slepian and Kahane (see e.g. [LT91])

$$\mathbb{E}[F_N(X_N^1) - F_N(X_N^0)|\mathcal{Y}] = \frac{1}{2}\int_0^1 \mathrm{d}h \sum_{\substack{i,j=1\\ i\neq j}}^{Tr_N} (\Lambda_{ij}^1 - \Lambda_{ij}^0)\mathbb{E}\Big[\frac{\partial^2 F_N(X_N^h)}{\partial X(i)\partial X(j)}\Big|\mathcal{Y}\Big]. \tag{38}$$

We will show that the integral in (38) converges to 0. To save on notation we assume that $\ell = 1$ and write $u = u_1$, $s = s_1$. Generalisation to larger ℓ is straightforward. The second derivative in (38) is 0 if at least one of i, j is larger than sr_N. For $i, j < sr_N$ the second derivative equals

$$\begin{aligned} &\frac{u^2\beta^2 N F_N(X_N^h)}{t_N^2} \prod_{\circ=i,j} e^{\beta\sqrt{N}X_N^h(\circ)} g'\Big(\frac{u}{t_N} e^{\beta\sqrt{N}X_N^h(\circ)}\Big) \\ &\le \frac{u^2\beta^2 N}{t_N^2} \prod_{\circ=i,j} e^{\beta\sqrt{N}X_N^h(\circ)} \exp\Big[-2g\Big(\frac{u}{t_N} e^{\beta\sqrt{N}X_N^h(\circ)}\Big)\Big], \end{aligned} \tag{39}$$

where we used the fact that $g'(x) = (\ln(1+x))' = (1+x)^{-1} = \exp(-g(x))$, and omitted in the summation of $F_N(X_N^h)$ all terms different from i and j. To estimate the expected value of this expression we need the following technical lemma.

Lemma 8. *Let $c \in [0,1)$ and let U_1, U_2 be two standard normal variables with the covariance $\mathbb{E}[U_1U_2] = c$. For $u > 0$ define $\Xi_N(c) = \Xi_N(c,\beta,\gamma,u)$ and $\bar{\Xi}_N(c) = \bar{\Xi}_N(c,\beta,\gamma,u)$ by*

$$\Xi_N(c) = \frac{u^2\beta^2 N}{t_N^2} \prod_{\circ=1,2} \mathbb{E}\Big[\exp\Big\{\beta\sqrt{N}U_\circ - 2g\big(ut_N^{-1}e^{\beta\sqrt{N}U_\circ}\big)\Big\}\Big] \tag{40}$$

and

$$\bar{\Xi}_N(c) = C(1-c)^{-1/2}(1\wedge u^2)N^{1/(1+c)}\exp\Big\{-\frac{\gamma^2 N}{\beta^2(1+c)}\Big\}, \tag{41}$$

where $C = C(\gamma,\beta)$ is a suitably chosen large constant. Then

$$\Xi_N(c) \le \bar{\Xi}_N(c). \tag{42}$$

Proof. Define $\kappa_\pm = \sqrt{2(1\pm c)}$. Since $c < 1$ both κ_+ and κ_- are positive. Let $\bar{U}_1$, $\bar{U}_2$ be two independent standard normal variables. Then U_1 and U_2 can be written as

$$U_1 = \frac{1}{2}(\kappa_+\bar{U}_1 + \kappa_-\bar{U}_2), \qquad U_2 = \frac{1}{2}(\kappa_+\bar{U}_1 - \kappa_-\bar{U}_2). \tag{43}$$

Hence, $U_1 + U_2 = \kappa_+\bar{U}_1$. Using $g(x)+g(y) = g(x+y+xy) \ge g(x+y)$ if $xy \ge 0$, and $e^x + e^{-x} \ge e^{|x|}$ we get

$$\sum_{\circ=1,2} g\big(ut_N^{-1}e^{\beta\sqrt{N}U_\circ}\big) \ge g\Big(ut_N^{-1}\exp\Big(\frac{\kappa_+\beta\sqrt{N}\bar{U}_1}{2} + \Big|\frac{\kappa_-\beta\sqrt{N}\bar{U}_2}{2}\Big|\Big)\Big). \tag{44}$$

Hence, $\Xi_N(c)$ is bounded from above by

$$\begin{aligned}&\frac{u^2\beta^2 N}{t_N^2}\int_{\mathbb{R}^2}\frac{\mathrm{d}y}{2\pi}\\ &\times\exp\Big\{-\frac{y_1^2+y_2^2}{2} + \beta\sqrt{N}\kappa_+y_1 - 2g\big(ut_N^{-1}e^{\kappa_+\beta\sqrt{N}y_1/2+\kappa_-\beta\sqrt{N}|y_2|/2}\big)\Big\}.\end{aligned} \tag{45}$$

Substituting $z_1 = y_1 - \beta\sqrt{N}\kappa_+$ and $z_2 = y_2$ we get

$$\begin{aligned}&\frac{u^2\beta^2 N}{t_N^2}e^{\beta^2\kappa_+^2N/2}\int_{\mathbb{R}^2}\frac{\mathrm{d}z}{2\pi}\exp\Big(-\frac{z_1^2+z_2^2}{2}\Big)\\ &\times\exp\Big(-2g\Big(u\exp\Big\{\sqrt{N}\Big[\Big(\frac{\beta^2\kappa_+^2}{2}-\bar{\gamma}_N\Big)\sqrt{N} + \frac{\beta\kappa_+}{2}z_1 + \frac{\beta\kappa_-}{2}|z_2|\Big]\Big\}\Big)\Big),\end{aligned} \tag{46}$$

where $\bar{\gamma}_N \equiv N^{-1}\log t_N = \gamma - \log N/(2\alpha N)$. Observe that $\exp(-2g(ue^{\sqrt{N}x}))$ converges to the indicator function $\mathbf{1}_{x<0}$, as $N \to \infty$. The role of x will be played by the square bracket in the expression (46). Since $\kappa_+ > \sqrt{2}$ and $\gamma/\beta^2 < 1$, this bracket is positive for 'typical' z_1, z_2 not far from 0. This means that the largest contribution to (46) comes again from non-typical values of z_1, z_2. We need another tilting:

$$z_1 = \frac{1}{\sqrt{N}}\Big[v_1 - \frac{\kappa_-}{\kappa_+}|v_2| - N\Big(\beta\kappa_+ - \frac{2\bar{\gamma}_N}{\beta\kappa_+}\Big)\Big], \quad z_2 = \frac{v_2}{\sqrt{N}}. \tag{47}$$

This substitution transforms the domain where the square bracket of (46) is negative into the half plane $v_1 < 0$: The expression inside of the braces in (46) equals $\beta\kappa_+ v_1/2$. Substituting (47) into $(z_1^2 + z_2^2)/2$ produces an additional prefactor $N^{-1/\alpha}N^{1/(1+c)}\exp\big(-\frac{(\beta^2\kappa_+^2-2\gamma)^2N}{2\beta^2\kappa_+^2}\big)$. Another prefactor N^{-1} comes from the Jacobian. The remaining terms can be bounded from above by

$$\begin{aligned} &C\int_{\mathbb{R}}\exp\Big\{-\frac{v_2^2}{2N} - \Big(\beta\kappa_- - \frac{2\bar{\gamma}_N\kappa_-}{\beta\kappa_+^2}\Big)|v_2|\Big\}\mathrm{d}v_2 \\ &\times\int_{\mathbb{R}}\exp\Big\{\Big(\beta\kappa_+ - \frac{2\bar{\gamma}_N}{\beta\kappa_+}\Big)v_1 - 2g(ue^{\beta\kappa_+ v_1/2})\Big\}\mathrm{d}v_1. \end{aligned} \tag{48}$$

Ignoring the quadratic term and using the facts that the parenthesis on the first line is always positive and $\gamma_N \to \gamma$ we can bound the first integral for N large by

$$C\Big(\beta\kappa_- - \frac{2\gamma\kappa_-}{\beta\kappa_+^2}\Big)^{-1} \le C\kappa_-^{-1} \le C(1-c)^{-1/2}. \tag{49}$$

To bound the second integral observe that the integrand behaves as $\exp\{-2v_1\bar{\gamma}_N/\beta\kappa_+\}$ as $v_1 \to \infty$, and as $\exp\{(\beta\kappa_+ - (2\bar{\gamma}_N/\beta\kappa_+))v_1\}$ as $v_1 \to -\infty$. Therefore, the second integral is bounded uniformly for all values of c. Moreover, as u increases, the second integral is $O(u^{-2})$.

Putting everything together we get

$$\begin{aligned} \Xi_N(c) &\le C(1-c)^{-\frac{1}{2}}\frac{u^2\beta^2N}{t_N^2}(1\wedge u^{-2})e^{\beta^2\kappa_+^2N/2}N^{\frac{1}{1+c}-\frac{1}{\alpha}-1}e^{-\frac{(\beta^2\kappa_+^2-2\gamma)^2N}{2\beta^2\kappa_+^2}} \\ &= C(\gamma,\beta)(1-c)^{-1/2}N^{1/(1+c)}(1\wedge u^2)\exp\Big\{-\frac{\gamma^2N}{\beta^2(1+c)}\Big\} = \bar{\Xi}_N(c). \end{aligned} \tag{50}$$

This finishes the proof of Lemma 8. $\square$

We can now finish the proof of Proposition 2. First, note that for all $K > 0$

$$\lim_{N\to\infty} e^{KN/2}\int_0^1 (1-h)^{-1/2}N^{1/(1+h)}e^{-KN/(1+h)}\mathrm{d}h = C_K > 0. \tag{51}$$

Lemmas 3 and 8 imply that the absolute value of (38) can be bounded from above, $\mathcal{Y}$-a.s. for N large enough, by

$$CN^{-1}r_N\int_0^1 (1-h)^{-1/2}N^{1/(1+h)}\exp\Big\{-\frac{\gamma^2N}{\beta^2(1+h)}\Big\}\mathrm{d}h \tag{52}$$

which converges to 0 as $N \to \infty$ using (51). This finishes the proof of Proposition 2(a).

In the case (b) we can use Lemmas 7 and 8 to show that, out of a set with probability smaller than $C\xi$, (38) is bounded from above by

$$C\xi r_N \int_0^1 (1-h)^{-1/2} N^{1/(1+h)} \exp\Big\{ -\frac{\gamma^2 N}{\beta^2(1+h)}\Big\} \mathrm{d}h, \tag{53}$$

which again converges to 0 after taking $N \to \infty$ and then $\xi \to 0$. $\square$

6 Proofs of the main results

Proof (of Theorem 1). Propositions 1 and 2(a) yield

$$\begin{aligned} \mathbb{E}\big[\exp\big\{-u\bar{S}_N^\gamma(s)\big\}\big|Y_N\big] &= \mathbb{E}\big[F_N(X_N^0;s,u)\big|Y_N\big] \\ &= \mathbb{E}\big[F_N(X_N^1;s,u)\big] + o(1) = \mathbb{E}[e^{-uV_\alpha(\mathcal{K}s)}] + o(1). \end{aligned} \tag{54}$$

This implies the convergence of fixed time distributions of $\bar{S}_N^\gamma$ to those of $V_\alpha(\mathcal{K}\cdot)$. Analogous computation gives the convergence of more-dimensional marginals.

We still need to check tightness in the J_1-topology of the sequence $\bar{S}_N^\gamma$. For increasing processes it amounts to check (see e.g. [EK86] Theorem 7.2 on page 128) that $\mathcal{Y}$-a.s.

$$\forall \eta > 0, T > 0 \quad \exists K \text{ such that } \mathbb{P}[\bar{S}_N(T) \ge K|\mathcal{Y}] \le \eta \quad \forall N. \tag{55}$$

and

$$\forall \eta > 0, T > 0 \quad \exists \delta \text{ such that } \mathbb{P}[w_\delta^T(\bar{S}_N) \ge \eta|\mathcal{Y}] \le \eta \quad \forall N, \tag{56}$$

where the modulus of continuity $w_\delta^T(f)$ is defined by

$$w_\delta^T(f) = \inf_{\{t_i\}} \max_i \sup \big\{|f(s)-f(t)| : s,t \in [t_{i-1},t_i)\big\}, \tag{57}$$

where $\{t_i\}$ ranges over all partitions of the form $0 = t_0 < t_1 < \dots < t_{n-1} < T \le t_n$ with $\min_{1\le i\le n}(t_i - t_{i-1}) > \delta$ and $n \ge 1$.

The condition (55) follows directly from the tightness of the fixed-time marginals, which is a consequence of the continuity of the limiting Laplace transform $e^{-T\mathcal{K}u^\alpha}$ at $u = 0$.

To prove (56) more work is necessary. First, we show that traps with energies 'much smaller' than $\gamma\sqrt{N}/\beta$ contribute hardly to the clock process. Recall the definition (20) of B_N^ε and define

$$\bar{S}_N^\gamma(s,\varepsilon) = t_N^{-1} \sum_{i=0}^{\lfloor sr_N \rfloor} e_i \exp\big\{\beta\sqrt{N}H_N(Y_N(i))\big\}\mathbf{1}\{H_N(Y_N(i)) \le B_N^\varepsilon\}. \tag{58}$$

Lemma 9. *For every T and $\eta > 0$ there exists ε such that $\mathcal{Y}$-a.s., for all but finitely many N*

$$\mathbb{P}[\bar{S}_N^\gamma(T,\varepsilon) \geq \eta|\mathcal{Y}] \leq \eta. \tag{59}$$

Proof. Using definition (30) of I_ω it is possible to rewrite $\bar{S}_N^\gamma(T,\varepsilon)$ as

$$t_N^{-1}\sum_{\omega=1}^{\infty}\sum_{\sigma\in I_\omega\setminus I_{\omega+1}}\sum_{i=1}^{\omega} e'_{\sigma,i} e^{\beta\sqrt{N}H_N(\sigma)}\mathbf{1}\{H_N(\sigma)\leq B_N^\varepsilon\} \equiv \sum_{\omega=1}^{\infty}\sum_{i=1}^{\omega} q_N^\varepsilon(\omega,i), \tag{60}$$

where $e'_{\sigma,i}$ are i.i.d. mean-one exponentials. By Lemma 6(b) we can restrict the sum to $\omega \leq N$. For these ω we have, using the notation of Lemma 2,

$$\mathbb{P}\Big[q_N^\varepsilon(\omega,i) \geq \frac{\eta}{\omega 2^\omega}\Big] \leq \frac{1-\mathbb{E}[e^{-q_N^\varepsilon(\omega)}]}{1-e^{-\eta/(2^\omega\omega)}} \leq \frac{1-(1-r_N^{-1}f_N(\varepsilon))^{|I_\omega|}}{1-e^{-\eta/(2^\omega\omega)}}. \tag{61}$$

For $\omega = 1$ this is bounded by $Cf_N(\varepsilon)$ which can be made smaller than $\eta/2$ by choosing ε small enough. Moreover, using Lemma 6, $\mathcal{Y}$-a.s,

$$\sum_{\omega=2}^{N}\sum_{i=1}^{\omega}\mathbb{P}\Big[q_N^\varepsilon(\omega,i) \geq \frac{\eta}{\omega 2^\omega}\Big] \leq \sum_{\omega=2}^{N} Cf_N(\varepsilon) r_N^{-1}(N^{-\omega+1}r_N \vee 1)\omega^2 2^\omega. \tag{62}$$

which is smaller than $\eta/2$ for N large enough. □

Lemma 10. *For any fixed $\varepsilon > 0$, $\mathcal{Y}$-a.s,*

$$\lim_{N\to\infty}\mathbb{P}[\max\{H_N(\sigma) : \sigma \in I_2\} \geq B_N^\varepsilon|\mathcal{Y}] = 0. \tag{63}$$

Proof. By Lemma 6, $\mathcal{Y}$-a.s., $|I_2| \leq TN^{-1}r_N$. Since the energies are i.i.d., the lemma follows by elementary arguments. □

We can now check (56). Fix $\eta > 0$ and $T > 0$. By Lemma 9 we can choose ε small such that $\mathbb{P}[\bar{S}_N^\gamma(T,\varepsilon) \geq \eta/2] \leq \eta/3$. Lemma 10 implies that $\mathbb{P}[\max\{H_N(\sigma) : \sigma \in I_2\} \geq B_N^\varepsilon|\mathcal{Y}] \leq \eta/3$. for all N large enough. Thus out of a set of probability smaller than $2\eta/3$, the contribution to the clock of the configurations with energies smaller than B_N^ε and of the configurations visited more than ones is bounded by $\eta/2$. Out of this set the modulus of continuity $w_\delta^T(\bar{S}_N^\gamma)$ can be larger than η only if

$$\exists i,j : |i-j| \leq \delta r_N, Y_N(i), Y_N(j) \notin I_2 \text{ and } \min_{k\in\{i,j\}} \frac{e_k}{t_N} e^{\beta\sqrt{N}H_N(Y_N(k))} \geq \frac{\eta}{2}. \tag{64}$$

However, the random variables $\{H_N(\sigma), \sigma \in I_1 \setminus I_2\}$ are i.i.d. An elementary calculation then shows that the probability of event in (64) can be made smaller than $\eta/3$ by choosing δ small. This finishes the proof of the tightness and thus of Theorem 1. □

Proof (of Theorem 2). The theorem follows from Propositions 1 and 2(b) by a reasoning analogous to (54). □.

Proof (of Theorem 3). The proof of (a) is standard. It is sufficient to observe that $\mathbb{P}[\sigma_N(t_N) = \sigma_N(\theta t_N)|\mathcal{Y}]$ is very well approximated by $\mathbb{P}[\{\bar{S}_N^\gamma(t) : t \ge 0\} \cap [1,\theta] = \emptyset|\mathcal{Y}]$. The last probability converges, $\mathcal{Y}$-a.s, as $N \to \infty$, to $\mathbb{P}[\{V_\alpha(t) : t \ge 0\} \cap [1,\theta] = \emptyset]$ by the weak convergence of $\bar{S}_N^\gamma$ in the J_1-topology.

To imply claim (b) Theorem 2 is not sufficient. We need in addition the following estimate.

Lemma 11. *Let $\gamma = \beta\beta_c$, that is $r_N = 2^N$, and let $\hat{S}_N^\xi(t) = \xi^{-1/\alpha}\bar{S}_N^\gamma(\xi t)$. Then, for any $T > 0$ and $\eta > 0$ there exists $\delta > 0$ such that*

$$\limsup_{\xi \to 0} \limsup_{N \to \infty} \mathbb{P}[\mathbb{P}[w_\delta^T(\hat{S}_N^\xi) \ge \eta|\mathcal{Y}] \le \eta] = 0. \tag{65}$$

Proof. We first set (see (58)), $\hat{S}_N^\xi(t,\varepsilon) = \xi^{-1/\alpha}\bar{S}_N^\gamma(\xi t, \xi^{1/\alpha}\varepsilon)$. We claim that, similarly as in Lemma 9, if ε is small enough, then

$$\mathbb{P}[\hat{S}_N^\xi(T,\varepsilon) \ge \eta|\mathcal{Y}] \le \eta \tag{66}$$

holds with probability converging to 1 as $N \to \infty$ and $\xi \to 0$. The proof of this claim is analogous to the proof of Lemma 9; Lemma 7 is used instead of Lemmas 3 and 6.

Moreover, similarly to Lemma 10, $\mathbb{P}[\max\{H_N(\sigma) : \sigma \in I_2\} \ge B_N^{\xi^{1/\alpha}\varepsilon}|\mathcal{Y}]$ converges to 0 in probability as $N \to \infty$ and $\xi \to 0$, by Lemma 7 again. The proof then follows the same line as the proof of the tightness in Theorem 1. $\square$

We now finish the proof of Theorem 3(b). As before, $\mathbb{P}[\sigma_N(\xi t_N) = \sigma_N(\xi\theta t_N)|\mathcal{Y}]$ is well approximated by

$$\mathbb{P}[\{\hat{S}_N(t) : t \ge 0\} \cap [1,\theta] = \emptyset|\mathcal{Y}]. \tag{67}$$

The last probability can be bounded from above by

$$\mathbb{P}\Big[\bigcup_{i=0}^{K} \big\{\hat{S}_N(i\delta) \le 1 \cap \hat{S}_N((i+1)\delta) \ge \theta\big\} \cup \{\hat{S}_N(K\delta) \le 1\}\Big|\mathcal{Y}\Big]. \tag{68}$$

The last quantity converges in probability, by Theorem 2, to

$$\mathbb{P}\Big[\bigcup_{i=0}^{K} \big\{V_\alpha(\mathcal{K}i\delta) \le 1 \cap V_\alpha(\mathcal{K}(i+1)\delta) \ge \theta\big\} \cup \{\hat{S}_N(\mathcal{K}K\delta) \le 1\}\Big], \tag{69}$$

which can be made arbitrarily close to $\mathsf{Asl}_\alpha(\theta)$ by choosing δ small and K large. A lower bound on (67) can be obtained by considering the event

$$\bigcup_{i=0}^{\lfloor T\delta^{-1}\rfloor} \big\{\hat{S}_N(i\delta) \le 1-\eta \cap \hat{S}_N((i+1)\delta) \ge \theta+\eta\big\} \cap \{w_\delta^T(\hat{S}_N) \le \eta\}. \tag{70}$$

By Lemma 11, with probability converging to 1 as $N \to \infty$ and $\xi \to 0$, the conditional probability of the event on the right of the last expression is very close to 1, and the conditional probability of of the union over i converges to a number that can be made arbitrarily close to $\mathsf{Asl}_\alpha(\theta)$ again. $\square$

References

[BBČ07] G. Ben Arous, A. Bovier, and J. Černý. Universality of the REM for dynamics of mean-field spin glasses. submitted, 2007.

[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(3):379–425, 2003.

[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(1):1–54, 2003.

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

[BČ08] G. Ben Arous and J. Černý. The arcsine law as a universal aging scheme for trap models. *Comm. Pure Appl. Math.*, 61(3):289–329, 2008.

[BČM06] G. Ben Arous, J. Černý, and T. Mountford. Aging in two-dimensional Bouchaud's model. *Probab. Theory Related Fields*, 134(1):1–43, 2006.

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

[BDG06] Gérard Ben Arous, Amir Dembo, and Alice Guionnet. Cugliandolo-Kurchan equations for dynamics of spin-glasses. *Probab. Theory Related Fields*, 136(4):619–660, 2006.

[BG97] G. Ben Arous and A. Guionnet. Symmetric Langevin spin glass dynamics. *Ann. Probab.*, 25(3):1367–1422, 1997.

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

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

[EK86] S. N. Ethier and T. G. Kurtz. *Markov processes*. John Wiley & Sons Inc., New York, 1986.

[Hoe63] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[LT91] M. Ledoux and M. Talagrand. *Probability in Banach spaces*, volume 23. Springer-Verlag, 1991.

[Mat87] P. Matthews. Mixing rates for a random walk on the cube. *SIAM J. Algebraic Discrete Methods*, 8(4):746–752, 1987.