

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

Preprint

ISSN 2198-5855

**Aging in the GREM-like trap model**

Véronique Gayrard<sup>1</sup>, Onur Gün<sup>2</sup>

submitted: January 19, 2015

<sup>1</sup> I2M, Aix-Marseille Université

39 rue F. Joliot Curie

13453 Marseille cedex 13

France

E-Mail: [veronique.gayrard@math.cnrs.fr](mailto:veronique.gayrard@math.cnrs.fr)

[veronique@gayrard.net](mailto:veronique@gayrard.net)

<sup>2</sup> Weierstrass Institute

Mohrenstr. 39

10117 Berlin

Germany

E-Mail: [onur.guen@wias-berlin.de](mailto:onur.guen@wias-berlin.de)

No. 2066

Berlin 2015



---

2010 *Mathematics Subject Classification*. Primary 82C44, 60K37, 60J27; secondary 60G55.

*Key words and phrases*. Random walk, random environment, trap models, aging, spin glasses, aging.

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

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

## Abstract

The GREM-like trap model is a continuous time Markov jump process on the leaves of a finite volume  $L$ -level tree whose transition rates depend on a *trapping landscape* built on the vertices of the whole tree. We prove that the natural two-time correlation function of the dynamics ages in the infinite volume limit and identify the limiting function. Moreover, we take the limit  $L \rightarrow \infty$  of the two-time correlation function of the infinite volume  $L$ -level tree. The aging behavior of the dynamics is characterized by a collection of *clock processes*, one for each level of the tree. We show that for any  $L$ , the joint law of the clock processes converges. Furthermore, any such limit can be expressed through Neveu's continuous state branching process. Hence, the latter contains all the information needed to describe aging in the GREM-like trap model both for finite and infinite levels.

## 1 Introduction

Trap models are main theoretical tools to quantify the out-of-equilibrium dynamics of spin glasses, and more specifically their aging behavior (see [BCKM97] for a review). Introduced in this context by J.P. Bouchaud [Bou92] to model the dynamics of mean field spin glasses such as the REM, GREM, and  $p$ -spin SK models, trap models are simple Markov jump processes that describe dynamics on microscopic spin space in terms of thermally activated barrier crossing between the valleys (or *traps*) of a random landscape on reduced state space devised so as to retain some of the key features of the free energy landscape of the underlying spin system. Activated aging occurs if, on time scales that diverge with the size of the system, the process observed through suitably chosen time-time correlation functions becomes slower and slower as time elapses.

The first rigorous connection between the microscopic dynamics of a spin system and a trap model was established in [BABG03a, BABG03b] where it is proved that a particular Glauber dynamics of the REM has the same aging behavior as Bouchaud's symmetric trap model on the complete graph or "REM-like trap model" [BD95]. This result was followed up by a series of results yielding a detailed understanding of the aging behavior of the REM for a wide range of time scales and temperatures, and various dynamics [BAČ08, Gay10b, Gay14]. Spin glasses with non-trivial correlations, namely the  $p$ -spin SK models, could also be dealt with albeit in a restricted domain of the time scales and temperature parameters, where it was proved that aging is just as in the REM [BAČ08, BAG12, BG13, BGŠ13]. This reflects the fact that the dynamics is insensitive to the correlation structure of the random environment. Although we do expect that, on longer time scales, the aging dynamics of the  $p$ -spin SK models belongs to (a) different universality class(es), there is yet no rigorous result supporting this idea.

At a heuristic level, possible effects of strongly correlated random environments on activated aging were first modeled by Bouchaud and Dean [BD95] using a trap model whose random traps are organized according to a hierarchical tree structure inspired by Parisi's ultrametric construction [MPS<sup>+</sup>84]. More recently, Sasaki and Nemoto [SN00, SN01] introduced a trap model on a tree with a view to modeling the aging dynamics of the GREM. From a mathematical perspective this GREM-like trap model seems more promising. Indeed a detailed (rigorous) analysis of the statics of the GREM (as well as that of a more general class of continuous random energy models, or CREMs) is available

(see [BK07] and the references therein), making it plausible to expect that the GREM-like trap model correctly predicts the behavior of the aging dynamics of the GREM itself, at least in some domain of the temperature and time scale parameters.

With this in mind, our aim in this paper is two-fold. Firstly, we want to identify the aging behavior of the GREM-like trap model as the branch size of the tree,  $n$ , diverges, and also as the number of levels,  $L$ , diverges after the limit  $n \rightarrow \infty$  is taken. Secondly, we want to emphasize that Neveu's continuous state branching process (hereafter abbreviated CSBP) naturally describes aging in the GREM-like trap model, namely, the aging behavior of the dynamics is encoded in a collection of *clock processes* and all possible limits of these clock processes are extracted from Neveu's CSBP.

## 1.1 Sasaki and Nemoto GREM-like trap model [SN00]

We start by specifying the underlying tree structure. For  $n \in \mathbb{N}$ , we write  $[n] = \{1, \dots, n\}$ . Let  $L \in \mathbb{N}$  be fixed and set  $V|_k = [n]^k$  for  $k = 1, \dots, L$ . We define the rooted  $L$ -level perfect  $n$ -ary tree by

$$\mathbb{T}_L = \bigcup_{k=0}^L V|_k, \quad (1.1)$$

where  $V|_0 = \{\emptyset\}$  is the root. We use the notation  $\mu|_k = \mu_1 \mu_2 \cdots \mu_k$  for a generic element of  $V|_k$ . We sometimes simply use  $\mu$  for  $\mu|_L \in V|_L$ . By convention the root belongs to the 0-th generation of the tree and  $\mu|_k$  to the  $k$ -th generation. For  $0 \leq k_1 < k_2 \leq L$ , we say that  $\mu|_{k_1} \in V|_{k_1}$  is an ancestor of  $\mu'|_{k_2} \in V|_{k_2}$  if  $\mu|_{k_1} = \mu'|_{k_1}$ . The set  $V|_L$  consists of the leaves of the tree, that is, the vertices that do not have any offspring. Whenever convenient we add the root to the notation by writing  $\mu|_k = \mu_0 \mu_1 \cdots \mu_k$ , where  $\mu_0 \equiv \emptyset$ . Note that the trees  $\mathbb{T}_L$  are parametrized by  $n \in \mathbb{N}$ . However, for notational convenience we do not keep  $n$  in the notation.

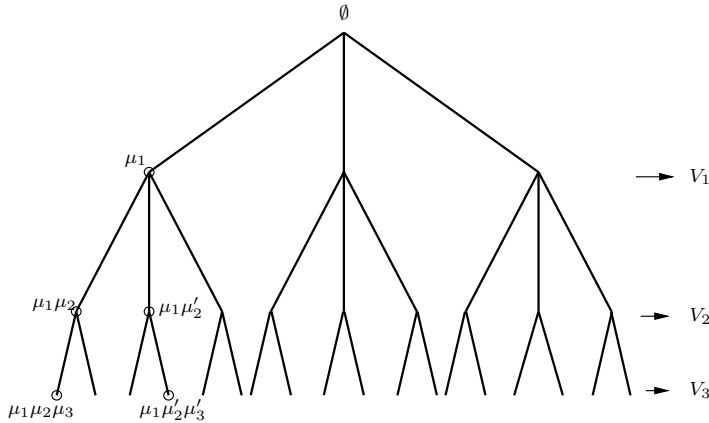


Figure 1: A representation of  $\mathbb{T}_L$  with  $L = n = 3$ .

Given two vertices  $\mu|_k, \mu'|_k \in V|_k$ , we denote by

$$\langle \mu|_k, \mu'|_k \rangle = \max\{l : \mu|_l = \mu'|_l\}, \quad \mu, \mu' \in V|_k, \quad (1.2)$$

the generation of their last common ancestor (hereafter abbreviated g.l.c.a.). The trapping landscape, or random environment, is a collection of independent random variables on the vertices of the tree  $\mathbb{T}_L$ ,

$$\{\lambda(\mu|_k) : \mu \in V|_L, k = 1, \dots, L\}, \quad (1.3)$$

where  $\lambda(\emptyset) \equiv 0$  and, for  $k = 1, \dots, L$ ,

$$\mathbb{P}(\lambda^{-1}(\mu|_k) \geq u) = u^{-\alpha_{k,L}}, \quad u \geq 1. \quad (1.4)$$

Here, the  $\alpha_{k,L}$ 's are real numbers satisfying

$$0 < \alpha_{1,L} < \alpha_{2,L} < \dots < \alpha_{L,L} < 1, \quad (1.5)$$

which implies that the  $\lambda^{-1}$ 's are heavy tailed. We assume that the two-parameter sequences (in  $n$  and  $L$ ) of random environments are independent and defined on a common probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . We denote by  $\mathbb{E}$  the expectation under  $\mathbb{P}$ .

For  $k = 1, \dots, L$ , let  $Y|_k$  be the discrete time Markov chain on  $V|_k$  with transition probabilities

$$W_k(\mu|_k, \mu'|_k) = \sum_{l=0}^{\langle \mu|_k, \mu'|_k \rangle \wedge (k-1)} \frac{(1 - \lambda(\mu|_l)) \prod_{l'=l+1}^{k-1} \lambda(\mu|_{l'})}{n^{k-l}}, \quad (1.6)$$

with the convention that  $\prod_{j=k}^{k-1} \lambda(\mu|_j) = 1$ . The GREM-like trap model, denoted by  $(X_L(t) : t \geq 0)$ , is a continuous time Markov jump process on the set of leaves of the tree,  $V|_L$ , whose transition rates are given by

$$w_L(\mu, \mu') = \lambda(\mu) W_L(\mu, \mu'). \quad (1.7)$$

Thus, we see that after waiting at a leaf  $\mu$  an exponential time with mean value  $\lambda^{-1}(\mu)$ , the process jumps to  $\mu'$  with probability  $W_L(\mu, \mu')$ . We also see that  $X_L$  is a reversible process whose unique invariant measure assigns to  $\mu \in V|_L$  the mass  $\prod_{k=1}^L \lambda^{-1}(\mu|_k)$ .

A comment on the form of the transition probabilities  $W_L(\mu, \mu')$  is now in order. Note that  $\lambda(\mu|_l) \in [0, 1]$ ,  $1 \leq l \leq L$ , so that the product  $(1 - \lambda(\mu|_l)) \prod_{l'=l+1}^{L-1} \lambda(\mu|_{l'})$  appearing in the summation in (1.6) is a probability. In physicists' term, this is the probability that, along the transition from  $\mu$  to  $\mu'$ , the system is activated from  $\mu$  to  $\mu|_{l+1}$  but not from  $\mu|_{l+1}$  to  $\mu|_l$ , meaning that the process jumps out of the traps attached to the vertices  $\mu|_{l'}$ ,  $l+1 \leq l' \leq L-1$ , but stays stranded in the trap attached to  $\mu|_l$ . The process then chooses its next state uniformly at random among the  $n^{L-l}$  leaves descending from  $\mu|_l$ .

Throughout this paper the initial distribution of  $X_L$  is taken to be the uniform distribution on  $V|_L$ . We embed the distributions of the chains  $X_L(t)$  for each  $n \in \mathbb{N}$  and  $L \in \mathbb{N}$ , on a common probability space whose distribution and expectation we denote by  $\mathcal{P}$  and  $\mathcal{E}$ , respectively. We suppress any references to the trapping landscape in the notation.

The two-time correlation function that we use to quantify aging in the GREM-like trap model is defined as follows. For  $k = 1, \dots, L$ , we first set

$$\Pi_k(t, s) = \mathcal{P}(\langle X_L(t), X_L(t+u) \rangle \geq k, \forall u \in [0, s]), \quad t, s > 0. \quad (1.8)$$

We next choose a non-decreasing smooth function  $q : [0, 1] \rightarrow [0, 1]$  with  $q(0) = 0$  and  $q(1) = 1$ , and then use it to define the two-time correlation function

$$C_L(t, s) = \sum_{k=1}^L \left[ q\left(\frac{k}{L}\right) - q\left(\frac{k-1}{L}\right) \right] \Pi_k(t, s). \quad (1.9)$$

Let us point out the key connections between the GREM-like trap model and the GREM (or the CREM, its generalization to continuous hierarchies). For this we rely on the paper [BK07] by Bovier and

Kurkova, that reviews the state of the art on the statics of these models. Using (1.2) one naturally defines a distance,  $1 - L^{-1}\langle \mu, \mu' \rangle$ ,  $\mu, \mu' \in V|_L$ , on the set of leaves of the tree, which is nothing but the ultrametric distance (see (1.1) in [BK07]) that endows the space of spin configurations of the GREM. Observe moreover that the function  $q$  in (1.9) is the analogue of the function  $A$  that enters the definition of the covariance structure of the GREM (see (1.2) in [BK07]). Thus  $C_L(t, s)$  in (1.8) naturally plays the role of the spin-spin correlation function between two configurations of some microscopic GREM dynamics observed at times  $t$  and  $t+s$ . Let us now turn to the trapping landscape (1.3). Its key features are modeled on those of the point process of extremes of the GREM's Boltzmann weights at temperature  $T$ . Under certain conditions on the parameters of the model, this process is known to converge, as the system's size diverges, to Ruelle's (non-normalized) Poisson cascades, a hierarchical point process constructed from a collection of Poisson point processes of intensity  $c_k x^{-(1+T/T_k)} dx$ , one for each level,  $k$ , of the hierarchy; here  $T_k$  is an associated critical temperature, and all processes that may result from this limiting procedure must be such that  $T_1 > \dots > T_L$  (see Theorem 2.3, Theorem 3.2, and definition 3.1 in [BK07]). Hence, at low enough temperature,  $T/T_1 < \dots < T/T_L < 1$ . Setting  $\alpha_{k,L} = T/T_k$  in (1.5), the non-normalized invariant measure  $\prod_{k=1}^L \lambda^{-1}(\mu|_k)$ ,  $\mu \in V|_L$ , of the process  $X_L$  in (1.7) models Ruelle's (non-normalized) cascade, while each  $\lambda^{-1}(\mu|_k)$  models the depth of a trap at  $\mu|_k$ . Based on heuristic ideas derived from metastability, the GREM dynamics is then replaced in an ad hoc manner by the process  $X_L$ . We refer to Section 1.2 of [BABG03b] for a more precise explanation of the correspondence between microscopic dynamics and trap models in the 1-level (REM-like) trap model. Note finally that in the GREM, the size of the hypercubes of different levels can vary and this correspond to varying branch sizes of the tree in the GREM-like trap model. Our results are valid for a range of trees with varying branch sizes, however, in order to keep the notations simple we opted to work with regular trees.

*Remark 1.1.* Observe that the two-time correlation function  $C_L$  can be expressed in terms of the distribution of the overlap observed by the dynamics. More precisely,

$$C_L(t, s) = \mathcal{E} \left( q(T_L(t, s)) \right), \quad (1.10)$$

where

$$T_L(t, s) = \sup \left\{ l : \frac{1}{L} \langle X_L(t), X_L(t+u) \rangle \geq l, \forall u \in [0, s] \right\}. \quad (1.11)$$

Indeed, in our proofs we identify the limiting distribution of (1.11), and thus, our results are valid for any choice of  $q$ .

Following the by now well-established strategy to analyze the aging properties of disordered systems, for each level  $k$ , we introduce the so-called  $k$ -th level clock process,  $S_{k,L}$ , that is the partial sum process defined through

$$X_L(t)|_k = Y(S_{k,L}^{\leftarrow}(t))|_k, \quad t \geq 0. \quad (1.12)$$

In contrast with earlier works, in order to fully describe the behavior of the two-time correlation function we need to control a whole collection of clock processes, one for each level of the tree.

The following description of  $S_{k,L}$  is going to be useful. We say that a jump of  $X_L$  is from a level  $k \in \{1, \dots, L\}$ , if the last activated vertex is on the  $(k-1)$ -th level of the tree, and that a jump is beyond and including the  $k$ th level if it is a jump from a level in  $\{1, \dots, k\}$ . Then,  $S_{k,L}(i)$  is the time it takes for the dynamics  $X_L$  to make  $i$  jumps beyond and including the  $k$ -th level of the tree.

## 1.2 Convergence of the two-time correlation function.

To study aging one has to choose a time scale  $c_n$  on which the dynamics is observed. In this paper we work with  $c_n$  of the form

$$c_n = n^\rho, \quad \rho > 0. \quad (1.13)$$

We then say that the system is aging on the  $k$ -th level, for  $\theta > 0$ , if  $\Pi_k(c_n, \theta c_n)$  has a non-trivial limit that depends only on  $\theta$ , and that it is non-aging on the  $k$ -th level if  $\Pi_k(c_n, \theta c_n)$  converges to 0. For  $k = 1, \dots, L$ , we set

$$d_{k,L} = \frac{1}{\alpha_{L,L}} + \frac{1}{\alpha_{L-1,L}} + \dots + \frac{1}{\alpha_{k,L}} - (L - k). \quad (1.14)$$

Note that by (1.5)

$$d_{1,L} > d_{2,L} > \dots > d_{L,L} > 1. \quad (1.15)$$

For a given exponent  $\rho$ , we define  $l^*(\rho) \in \{1, \dots, L\}$  by

$$l^*(\rho) = \sup\{k : \rho < d_{k,L}\}. \quad (1.16)$$

For  $\alpha \in (0, 1)$  let  $Asl_\alpha$  be the classical arcsine law distribution function

$$Asl_\alpha(u) = \frac{\sin \alpha \pi}{\pi} \int_0^u x^{-\alpha} (1-x)^{\alpha-1} dx, \quad u \in [0, 1]. \quad (1.17)$$

**Theorem 1.1.** *For any fixed  $L \in \mathbb{N}$  and  $\rho \in (0, d_{1,L}) \setminus \{d_{L,L}, \dots, d_{2,L}\}$  set  $l^* = l^*(\rho)$  and  $\bar{\alpha}_k = \prod_{i=k}^{l^*} \alpha_{i,L}$ . There exists a subset  $\Omega_L \subseteq \Omega$  with  $\mathbb{P}(\Omega_L) = 1$  such that for any environment in  $\Omega_L$ , for any  $\theta > 0$ ,*

$$\lim_{n \rightarrow \infty} C_L(c_n, \theta c_n) = \sum_{k=1}^{l^*} \left[ q \left( \frac{k}{L} \right) - q \left( \frac{k-1}{L} \right) \right] Asl_{\bar{\alpha}_k} \left( \frac{1}{1+\theta} \right). \quad (1.18)$$

*Remark 1.2.* Indeed, we are going to prove that, under the assumptions of Theorem 1.1, for  $k = 1, \dots, l^*$ ,  $\Pi_k(c_n, \theta c_n) \rightarrow Asl_{\bar{\alpha}_k} \left( \frac{1}{1+\theta} \right)$ , and for  $k = l^*(\rho) + 1, \dots, L$ ,  $\Pi_k(c_n, \theta c_n) \rightarrow 0$ .

The collection in (1.14) can be seen as critical time scale exponents. More precisely, as a consequence of Theorem 1.1, if  $\rho < d_{k,L}$ , the dynamics is aging on the  $k$ -th level, and if  $\rho > d_{k,L}$ , it is non-aging on the  $k$ -th level. Then, the levels  $1, \dots, l^*(\rho)$  are aging whereas the levels below  $l^*(\rho)$  are non-aging, and the inequalities in (1.15) reflect that as the time scale of observation gets longer the aging behavior disappears from bottom to the top of the tree.

Next, we take the  $L \rightarrow \infty$  limit of the limiting two-time correlation function in (1.18). In order to do this, we need to define a family of collections of  $\alpha$  parameters satisfying (1.5). Let  $R : [0, 1] \rightarrow [0, \infty)$  be a strictly increasing, strictly concave, smooth function with  $R(0) = 0$ . For  $L \in \mathbb{N}$ , set

$$\alpha_{k,L} = \exp \left\{ - \left( R(k/L) - R((k-1)/L) \right) \right\}, \quad k = 1, \dots, L. \quad (1.19)$$

As a result, the collection  $\{\alpha_{1,L}, \dots, \alpha_{L,L}\}$  satisfies (1.5) for all  $L \in \mathbb{N}$ . Note that although for a given collection  $\{\alpha_{1,L}, \dots, \alpha_{L,L}\}$  as in (1.5), one can find a function  $R$  satisfying (1.19), it is not necessarily possible to do that for a given family of such collections  $\{\alpha_{1,L}, \dots, \alpha_{L,L} : L \in \mathbb{N}\}$ . Also observe that, in the light of Theorem 1.1, the choice of (1.19) is natural for an infinite levels limit because the infinite product of  $\alpha$  parameters is non-trivial.

For a given exponent  $\rho > 0$ , we set

$$r^*(\rho) = \sup\{s \geq 0 : R(1) - R(s) + 1 > \rho\}. \quad (1.20)$$

We will see in Section 3 (see (3.23) and the paragraph before it) that  $r^*(\rho)$  is the limit of the ratio  $l^*(\rho)/L$  as  $L$  diverges. As a result, there is aging for  $\rho$  such that  $r^*(\rho) > 0$ . Since  $R(0) = 0$ , the latter is equivalent to  $\rho < R(1) + 1$ .

**Corollary 1.2.** *Let  $\rho \in (0, R(1) + 1) \setminus \{d_{L,L}, \dots, d_{2,L}\}$  for all  $L$  large enough. Set  $r^* = r^*(\rho)$  and let  $\alpha(x) = \exp\{- (R(r^*) - R(x))\}$  for  $x \in [0, r^*]$ . There exists a subset  $\Omega' \subseteq \Omega$  with  $\mathbb{P}(\Omega') = 1$  such that for any environment in  $\Omega'$ , for any  $\theta > 0$ ,*

$$\lim_{L \rightarrow \infty} \lim_{n \rightarrow \infty} C_L(c_n, \theta c_n) = \int_0^{r^*} q'(x) A s l_{\alpha(x)} \left( \frac{1}{1 + \theta} \right) dx. \quad (1.21)$$

*Remark 1.3.* The result in Corollary 1.20 can be extended to functions  $q$  that are right continuous and left limits. However, since the formula in 1.21 is very transparent and neat, we use only smooth functions.

Let us discuss the connection of our results with earlier literature and comment about some directions for future research. Observe that our results do not cover the time scales  $c_n \sim n^{d_{k,L}}$ . Moreover, varying branch sizes can yield time scales that are critical for several levels simultaneously. For these levels the inequalities in (1.15) become equalities. The cases where the branch sizes are fine tuned in such a way that all the critical time scale exponents are the same were studied in [FGG14]. There, the authors constructed a  $K$  process in an infinite  $L$ -level tree and proved that for any critical time scale, the scaling limit of the GREM-like trap model exists and is given by such an  $L$ -level  $K$  process. In [FP14], the authors constructed, for a special choice of parameters, an infinite levels, infinite volume  $K$  process and showed that it is the limit  $L \rightarrow \infty$  of the  $L$ -level  $K$  process obtained in [FGG14]. An investigation of intermediate cases, where there are times scales that are critical for some levels and aging for others, would require a non-trivial combination of analysis of this paper and that of [FGG14].

### 1.3 Convergence of the clock processes.

In this section we state our results on the convergence of the clock processes. We express the limiting clock processes using Neveu's CSBP which is a time-homogeneous Markov process  $(W(r) : r \geq 0)$  whose semigroup is characterized by

$$E [e^{-\kappa W(r)} | W(0) = t] = \exp(-t\kappa e^{-r}), \quad \kappa > 0, t \geq 0. \quad (1.22)$$

We write  $W(r, t)$  for  $W$  starting from  $t \geq 0$ , that is,  $W(0, t) = t$ . Using Kolmogorov's extension theorem one can construct a process  $(W(r, t) : r, t \geq 0)$  such that  $W(\cdot, 0) \equiv 0$  and,  $W(\cdot, t + s) - W(\cdot, t)$  is independent of  $(W(\cdot, c) : 0 \leq c \leq t)$  and has the same law as  $W(\cdot, s)$ . Hence, for any fix  $r \geq 0$ , the right continuous version of  $W(r, \cdot)$  has independent, stationary increments. From its Laplace transform we see that it is an  $e^{-r}$ -stable subordinator with Laplace exponent  $\kappa e^{-r}$ . Moreover, by (1.22),  $W(r + p, \cdot)$  has the same distribution as the Bochner subordination of  $W(r, \cdot)$  with the directing process  $W'(p, \cdot)$ , where  $W'(p, \cdot)$  is an independent copy of  $W(p, \cdot)$ , that is,

$$W(r + p, \cdot) \stackrel{d}{=} W(r, W'(p, \cdot)). \quad (1.23)$$

The above description is taken from [BLG00] where it is pointed out that, in general, (1.23) allows to connect CSBPs and Bochner's subordination, a connection developed further in the following:

**Proposition 1.3** (Proposition 1 in [BLG00] applied to Neveu's CSBP). *On some probability space there exists a process  $(Z_{p,r}(t) : 0 \leq p \leq r, \text{ and } t \geq 0)$  such that:*

(i) *For every  $0 \leq p \leq r$ ,  $Z_{p,r} = (Z_{p,r}(t) : t \geq 0)$  is an  $e^{-(r-p)}$ -stable subordinator with Laplace exponent  $\kappa e^{-(r-p)}$ .*



(ii) For any integer  $m \geq 2$  and  $0 \leq r_1 \leq \dots \leq r_m$  the subordinators  $Z_{r_1, r_2}, Z_{r_2, r_3}, \dots, Z_{r_{m-1}, r_m}$  are independent and

$$Z_{r_1, r_m}(t) = Z_{r_{m-1}, r_m} \circ \dots \circ Z_{r_1, r_2}(t), \quad \forall t \geq 0 \text{ a.s.} \quad (1.24)$$

Finally,  $\{Z_{0,r}(t) : r \geq 0, t \geq 0\}$  and  $\{W(r, t) : r, t \geq 0\}$  have the same finite dimensional distributions.

To each level of the tree we assign a pair of sequences,  $a_n(k)$  and  $c_n(k)$ , defined by

$$a_n(k) = \begin{cases} n^{1+\rho-d_{k,L}} & k = L, \dots, l^* + 1, \\ n^{\bar{\alpha}_k(1+\rho-d_{l^*,L})} & k = l^*, \dots, 1, \end{cases} \quad (1.25)$$

and

$$c_n(k) = a_n(k+1), \quad k = 1, \dots, L-1, \quad c_n(L) = c_n. \quad (1.26)$$

Note that as  $n \rightarrow \infty$ ,  $a_n(k) \ll n$  and  $c_n(k) = a_n(k)^{1/\alpha_{k,L}}$  for all  $k \leq l^*$  (i.e. for all the aging levels), whereas  $a_n(k) \gg n$  and  $c_n(k) = a_n(k)n^{1/\alpha_{k,L}-1}$  for all  $k \geq l^* + 1$  (i.e. for all the non-aging levels). Moreover, under the assumptions of Theorem 1.1, in both cases the decay or growth of  $a_n(k)/n$  is at least polynomial.

For  $k = 1, \dots, L$ , we define the rescaled clock processes

$$S_{k,L}^{(n)}(t) = \frac{S_{k,L}(\lfloor ta_n(k) \rfloor)}{c_n}, \quad t \geq 0, \quad (1.27)$$

where we set  $S_{k,L}(0) = 0$ . Hence,  $S_{k,L}^{(n)} \in D([0, \infty))$  where  $D([0, \infty))$  denotes the space of càdlàg functions on  $[0, \infty)$ . The following is our main result on the convergence of clock processes.

**Theorem 1.4.** For any  $L \in \mathbb{N}$  there exists a subset  $\tilde{\Omega}_L \subseteq \Omega$  with  $\mathbb{P}(\tilde{\Omega}_L) = 1$  such that for some positive constants  $b_{1,L}, \dots, b_{l^*,L}$  setting

$$\tilde{Z}_{k,l^*}(\cdot) = Z_{R((k-1)/L), R(l^*/L)}(b_{k,L} \cdot), \quad (1.28)$$

for any environment in  $\tilde{\Omega}_L$ , as  $n \rightarrow \infty$

$$\left( S_{k,L}^{(n)} : k = 1, \dots, l^* \right) \Longrightarrow \left( \tilde{Z}_{k,l^*} : k = 1, \dots, l^* \right) \quad (1.29)$$

weakly on the space  $D^{l^*}([0, \infty))$  equipped with the product Skorohod  $J_1$  topology.

*Remark 1.4.* Note that by (1.19) and the definition of Neveu's CSBP, for any  $i = 1, \dots, L$  and  $b > 0$ ,  $Z_{R((i-1)/L), R(i/L)}(b \cdot)$  is a stable subordinator with index  $\alpha_{i,L}$ . Therefore, by (1.24) the distribution of the right hand side of (1.29) is given by compositions of stable subordinators.

*Remark 1.5.* Together with the previous remark, Theorem 1.4 implies that  $S_{k,L}^{(n)}$  converges weakly to an  $\bar{\alpha}_k$ -stable subordinator, where  $\bar{\alpha}_k$  is given as in Theorem 1.1. If one is only interested in marginal distributions of the clock processes (which is enough to obtain our results on two-time correlation functions) a shorter proof is available through a technique based on Durrett and Resnick [DR78], which has been recently proved to be very useful in the context of dynamics in disordered systems, see [BG13, BGŠ13, Gay10a, Gay10b, Gay12]. However, in this paper we choose to prove the stronger result of the joint convergence of clock processes in order to make the connection to Neveu's CSBP more transparent.

In [BLG00], Bertoin and Le Gall gave a representation of the genealogical structure of CSBPs using Bochner's subordination. We say that an individual  $t$  at generation  $r$  has an ancestor  $c$  at generation  $p \in [0, r]$  if  $c$  is a jumping time of  $Z_{p,r}$  and

$$Z_{p,r}(c-) < t < Z_{p,r}(c). \quad (1.30)$$

Since the Lévy measure of  $Z_{r,p}$  has no atoms, the set of individuals at generation  $d$  who do not have an ancestor at generation  $r < d$  has a.s. Lebesgue measure 0. In view of this, for individuals  $t_1$  and  $t_2$  at generation  $r$ , we let

$$T_r(t_1, t_2) = \sup\{p \geq 0 : t_1 \text{ and } t_2 \text{ have a common ancestor at generation } p\},$$

and set  $T_r(t_1, t_2) = -\infty$  if  $t_1$  and  $t_2$  do not have a common ancestor. Using  $T_r$  we can express the limiting two-time functions in Theorem 1.1 and Corollary 1.2 as follows. Note that, by definition, we have

$$P(T_r(t_1, t_2) \geq p) = P(\{Z_{p,r}(t) : t \geq 0\} \cap [t_1, t_2] = \emptyset), \quad (1.31)$$

and since  $Z_{p,r}$  is a stable subordinator with index  $e^{-(r-p)}$ , the right hand side of (1.31) is nothing but  $Asl_\alpha(b/d)$ , where  $\alpha = e^{-(r-p)}$ . Thus, we can rewrite the  $Asl$  terms in Theorem 1.1 and Corollary 1.2 as, respectively,

$$Asl_{\bar{\alpha}_k}\left(\frac{1}{1+\theta}\right) = P\left(T_{R(k-1)/L}(1, 1+\theta) \geq R((k-1)/L)\right), \quad (1.32)$$

and

$$Asl_{\alpha(x)}\left(\frac{1}{1+\theta}\right) = P\left(T_{R(r^*)}(1, 1+\theta) \geq R(x)\right). \quad (1.33)$$

*Remark 1.6.* Neveu's CSBP was first used in the study of the statics of the GREM and CREM [Nev92]. Namely, the limiting geometric structure of the Gibbs measure of these models can be expressed in terms of the genealogy of Neveu's CSBP (see Section 5 of [BK07]). However, in the context of the dynamics, it appears in a different way, describing the limits of the clock processes.

The rest of this paper is organized as follows. In Section 2 we describe clock processes through a certain cascade of point processes associated to the dynamics, prove that they converge weakly to a cascade of Poisson point processes, and finally, using little more than the continuous mapping theorem, we establish Theorem 1.4. In Section 3 we prove Theorem 1.1 and Corollary 1.2.

## 2 Convergence of the clock processes.

### 2.1 Description of the clock processes through a cascade of point processes.

We first give definitions and introduce notation for general cascade processes.

For a complete, separable metric space  $A$ , we designate by  $M(A)$  the space of point measures on  $A$ , and by  $\varepsilon_x$  the Dirac measure at  $x \in A$ , i.e.  $\varepsilon_x(F) = 1$  if  $x \in F$  and  $\varepsilon_x(F) = 0$  if  $x \notin F$ . We set

$$H = (0, \infty) \times (0, \infty), \quad \text{and} \quad M_{1,l} = M(H^1) \otimes \cdots \otimes M(H^l), \quad l \in \mathbb{N}. \quad (2.1)$$

All the point measures we use are indexed by  $\mathbb{N}^k$  and we use the notation  $j|_k = j_1 j_2 \cdots j_k$  for a member of  $\mathbb{N}^k$ .

The set of  $l$ -level cascade point measures,  $\mathcal{M}_{1,l}$ , is the subset of  $M_{1,l}$  where for each  $m = (m_1, \dots, m_l) \in \mathcal{M}_{1,l}$  there corresponds a collection of points in  $H$  of the form

$$\{(t_{j|_k}, x_{j|_k}) : j|_l \in \mathbb{N}^l, k = 1, \dots, l\} \quad (2.2)$$

such that for each  $k = 1, \dots, l$

$$m_k = \sum_{j|_k \in \mathbb{N}^k} \varepsilon_{(t_{j_1}, x_{j_1}, \dots, t_{j|_k}, x_{j|_k})}. \quad (2.3)$$

We refer to the collection in (2.2) as the marks of  $m$ . Throughout this paper we assume that all the point measures are simple. Let  $\widetilde{\mathcal{M}}_{1,l}$  be the subset of  $\mathcal{M}_{1,l}$  such that for each  $m \in \widetilde{\mathcal{M}}_{1,l}$ ,  $k = 1, \dots, l$ ,  $j|_{k-1} \in \mathbb{N}^{k-1}$  and  $t > 0$ ,

$$m_k \left( \{(t_{j_1}, x_{j_1}, \dots, t_{j|_{k-1}}, x_{j|_{k-1}})\} \times ((0, t] \times (0, \infty)) \right) < \infty. \quad (2.4)$$

Then, for  $m \in \widetilde{\mathcal{M}}_{1,l}$ , there exists a unique labeling of the marks so that for any  $j|_{k-1} \in \mathbb{N}^{k-1}$

$$t_{j|_{k-1}1} < t_{j|_{k-1}2} < \dots. \quad (2.5)$$

From now on we only use this labeling for the marks of  $m \in \widetilde{\mathcal{M}}_{1,l}$ .

We define  $T_l : \widetilde{\mathcal{M}}_{1,l} \rightarrow D([0, \infty))$  by

$$T_l(m)(t) = \sum_{t_{j_1} \leq t} \sum_{t_{j_1 j_2} \leq x_{j_1}} \dots \sum_{t_{j_l} \leq x_{j_{l-1}}} x_{j_l}. \quad (2.6)$$

Next, we introduce a map  $\bar{T}_l : \widetilde{\mathcal{M}}_{1,l} \rightarrow \widetilde{\mathcal{M}}_{1,l-1}$ . For  $m \in \widetilde{\mathcal{M}}_{1,l}$  let

$$Z(i) = m_2 \left( \{(t_i, x_i)\} \times ((0, x_i] \times (0, \infty)) \right), \quad i \in \mathbb{N}, \quad (2.7)$$

where we set  $Z(0) = 0$ ,

$$g(i) = \max\{r : Z(0) + \dots + Z(r) < i\}, \quad h(i) = Z(0) + \dots + Z(g(i)), \quad i \in \mathbb{N}, \quad (2.8)$$

and

$$s(i) = x_1 + \dots + x_i, \quad i \in \mathbb{N}, \quad s(0) = 0. \quad (2.9)$$

Then, we define  $\bar{T}_l(m)$  as a point in  $\mathcal{M}_{1,l}$  whose marks are given by

$$\{(\bar{t}_{j|_k}, \bar{x}_{j|_k}) : j|_{l-1} \in \mathbb{N}^{l-1}, k = 1, \dots, l-1\}, \quad (2.10)$$

where, for  $j_1 \in \mathbb{N}$ ,

$$\bar{t}_{j_1} = s(g(j_1)) + t_{(g(j_1)+1)(j_1-h(j_1))}, \quad \bar{x}_{j_1} = x_{(g(j_1)+1)(j_1-h(j_1))}, \quad (2.11)$$

and for  $k > 1$  and  $j|_k \in \mathbb{N}^k$ ,

$$\bar{t}_{j|_k} = t_{(g(j_1)+1)(j_1-h(j_1))j_2 \dots j_k}, \quad \bar{x}_{j|_k} = x_{(g(j_1)+1)(j_1-h(j_1))j_2 \dots j_k}. \quad (2.12)$$

It is clear that  $\bar{T}_l(m) \in \widetilde{\mathcal{M}}_{1,l}$  and the marks in (2.10) is already ordered in jump times, that is, (2.5) is satisfied.

For  $k \leq l$ , we define  $T_{k,l} : \widetilde{\mathcal{M}}_{1,l} \rightarrow D([0, \infty))$  by

$$T_{k,l} = T_{l-k+1}(\bar{T}_{l-k+2} \circ \dots \circ \bar{T}_{l-1} \circ \bar{T}_l), \quad (2.13)$$

where it is understood that  $T_{1,l} = T_l$ .

For  $l_1 \leq l_2$  and  $m = (m_1, \dots, m_{l_2}) \in \widetilde{\mathcal{M}}_{1,l_2}$ , let  $m|_{l_1} = (m_1, \dots, m_{l_1}) \in \widetilde{\mathcal{M}}_{1,l_1}$ . We use the following property later: for  $k < l_1 \leq l_2$  and  $m \in \widetilde{\mathcal{M}}_{1,l_2}$ ,

$$T_{k,l_2}(m) = T_{l_1,l_2}(m) \circ T_{k,l_1-1}(m|_{l_1-1}), \quad (2.14)$$

where the composition in the above display is on the space  $D([0, \infty))$ .

We now describe a cascade of point processes associated to the dynamics  $X_L$  whose image under the functionals  $T_{k,L}$  yields the clock processes. A cascade of simple random walks with  $L$  levels on  $\mathbb{T}_L$  is a collection of random variables

$$J = \left\{ J_k(j_k; j|_{k-1}) : j|_L \in \mathbb{N}^L, k = 1, \dots, L \right\} \quad (2.15)$$

that is characterized as follows:

(i) For each  $k = 1, \dots, L$  and  $j|_{k-1}$  fixed,  $\{J_k(j_k; j|_{k-1}) : j_k \in \mathbb{N}\}$  is a collection of i.i.d. random variables distributed uniformly on  $[n]$ .

(ii) The families  $\{J_k(j_k; j|_{k-1}) : j_k \in \mathbb{N}\}$  for  $k = 1, \dots, L$  and  $j|_{k-1}$  are independent.

For  $k = 1, \dots, L$ , the jump chain  $J(j|_k)$  on  $V|_k$  is defined by

$$J(j|_k) = J_1(j_1)J_2(j_2; j|_1) \cdots J_k(j_k; j|_{k-1}). \quad (2.16)$$

In (i) above we make use of the fact that a simple random walk on the complete graph  $[n]$  starting from a uniform distribution is the same as a sequence of i.i.d. uniform distributions on  $[n]$ . For a fixed realization of the random environment, consider the collection of random variables

$$\{(t_{j|_k}, \xi_{j|_k}) : j|_L \in \mathbb{N}^L, k = 1, \dots, L\} \quad (2.17)$$

given as follows: for a given realization of  $J$ , it is an independent collection with  $\xi_{j|_k} \stackrel{d}{=} G(\lambda(J(j|_k)))$  for  $k = 1, \dots, L-1$ , and  $\xi_{j|_L} \stackrel{d}{=} \lambda^{-1}(J(j|_L))e$ . Here,  $G(p)$  denotes a geometric random variable with success probability  $p$  and  $e$  is a mean one exponential random variable. We also set  $t_{j|_k} = j_k$ . Using the collection in (2.17) we define  $\zeta^L = (\zeta_1^L, \dots, \zeta_L^L)$  by

$$\zeta_k^L = \sum_{j_1, \dots, j_k \in \mathbb{N}} \varepsilon_{(t_{j_1}, \xi_{j_1}, \dots, t_{j_k}, \xi_{j_k})}. \quad (2.18)$$

Clearly  $\zeta^L \in \widetilde{\mathcal{M}}_{1,L}$  and the marks of it are already ordered in jump times. Finally, the clock process  $S_{k,L}$  is given by

$$S_{k,L} = T_{k,L}(\zeta^L), \quad k = 1, \dots, L. \quad (2.19)$$

## 2.2 Convergence of the cascade of point processes of the dynamics

We have seen in the previous section how the collection in (2.17) is used to describe all the clock processes. Since we are interested in proving the joint convergence of clock processes  $S_{1,L}, \dots, S_{l^*,L}$  rescaled as in (1.27), where  $l^* = l^*(\rho)$ , we consider the following collection

$$\left\{ (t_{j|_k}^{(n)}, \xi_{j|_k}^{(n)}) : j|_{l^*} \in \mathbb{N}^{l^*}, k = 1, \dots, l^* \right\} \quad (2.20)$$

obtained from (2.17) by setting  $t_{j|_k}^{(n)} = j_k/a_n(k)$ ,

$$\xi_{j|_k}^{(n)} = \frac{\xi_{j|_k}}{c_n(k)}, \quad k = 1, \dots, l^* - 1, \quad \text{and} \quad \xi_{j|_{l^*}}^{(n)} = \frac{\Lambda(j|_{l^*})}{c_n}, \quad (2.21)$$

where

$$\Lambda(j|_{l^*}) = \sum_{j_{l^*+1}=1}^{\xi_{j|_{l^*}}} \cdots \sum_{j_L=1}^{\xi_{j|_{L-1}}} \xi_{j|_L}. \quad (2.22)$$

Here, the terms corresponding to the aging level  $l^*$  collects the whole waiting times until the dynamics jump over the vertex  $J(j|_{l^*})$ . Defining  $\zeta^{(n),l^*} = (\zeta_1^{(n),l^*}, \dots, \zeta_{l^*}^{(n),l^*})$  by

$$\zeta_k^{(n),l^*} = \sum_{j|_k \in \mathbb{N}^k} \varepsilon_{(t_{j_1}^{(n)}, \xi_{j_1}^{(n)}, \dots, t_{j_k}^{(n)}, \xi_{j_k}^{(n)})}, \quad (2.23)$$

we get

$$S_{k,L}^{(n)} = T_{k,l^*}(\zeta^{(n),l^*}), \quad k = 1, \dots, l^*. \quad (2.24)$$

Now we describe a cascade of Poisson point processes (PPP). For  $l \in \mathbb{N}$ , for constants  $0 < \beta_1 < \dots < \beta_l < 1$  and  $D_1, \dots, D_l > 0$ , let

$$\left\{ (t_{j|_k}, \eta_{j|_k}) : j|_k \in \mathbb{N}^k, k = 1, \dots, l \right\} \quad (2.25)$$

be a collection of random variables whose distribution is characterized by the following properties:

(i) For each  $k = 1, \dots, l$  and  $j|_{k-1}$  fixed, the distribution of  $\{(t_{j|_k}, \eta_{j|_{k-1}j_k}) : j_k \in \mathbb{N}\}$  is that of marks of a PPP on  $H$  with mean measure  $dt \times D_k \beta_k x^{-1-\beta_k} dx$ .

(ii) The families  $\{(t_{j|_{k-1}j_k}, \eta_{j|_{k-1}j_k}) : j_k \in \mathbb{N}\}$  for  $k = 1, \dots, l$  and  $j|_{k-1}$  are independent.

We define  $\chi^l = (\chi_1^l, \dots, \chi_l^l)$  by

$$\chi_k^l = \sum_{j|_k \in \mathbb{N}^k} \varepsilon_{(t_{j_1}, \eta_{j_1}, \dots, t_{j_k}, \eta_{j_k})}, \quad (2.26)$$

and refer to  $\chi^l$  as Ruelle's Poisson Cascade (RPC) with parameters  $\beta_1, \dots, \beta_l$  and constants  $D_1, \dots, D_l$ .

Our first goal is to prove that cascade of point measures  $\zeta^{(n),l^*}$  converges weakly to an RPC. Next, we prove that the versions of clock process functionals where the very small jumps are ignored are continuous. Finally, using the continuous mapping theorem and controlling the very small jumps we establish the convergence of clock processes. In this regard, we extend the classical results on convergence of sum of random variables with heavy tails (see e.g. Theorem 3.7.2 in [Dur10] and Proposition 3.4. in [Res86]).

Next, we state our weak convergence result. Recall that  $\zeta^{(n),l^*}$  and  $\chi^{l^*}$  are random elements of product of space of point measures  $M_{1,l^*} = M(H) \otimes \cdots \otimes M(H^{l^*})$ , and we use the weak convergence induced by the product vague topology, denoted by  $\Rightarrow$ .

**Proposition 2.1.** *There exist positive constants  $D_1, \dots, D_{l^*}$  and a subset  $\Omega'_L \subseteq \Omega$  with  $\mathbb{P}(\Omega'_L) = 1$  such that for any environment in  $\Omega'_L$ , as  $n \rightarrow \infty$*

$$\zeta^{(n),l^*} \Rightarrow \chi^{l^*}, \quad (2.27)$$

where  $\chi^{l^*}$  is a RPC with parameters  $\alpha_{1,L}, \dots, \alpha_{l^*,L}$  and constants  $D_1, \dots, D_{l^*}$ .

We first recall some basic facts about Laplace functionals of point processes since we use them to prove Proposition 2.1. For the basic concepts about point processes we mainly follow the book [Res08] and refer readers to the same source for further details.

Let  $m \in M(A)$  and  $f$  be a non-negative Borel measurable function on  $A$ . Define

$$m(f) = \int_A f(x) m(dx). \quad (2.28)$$

Let  $C_K^+(A)$  be the set of non-negative continuous functions on  $A$  with compact support. Then,  $m^{(n)} \in M(A)$  converges vaguely to  $m \in M(A)$  if

$$m^{(n)}(f) \rightarrow m(f), \quad \forall f \in C_K^+(A). \quad (2.29)$$

The vague topology on  $M(A)$  induced by the vague convergence is metrizable as a complete separable metric space. The weak convergence in  $M(A)$  is with respect to the vague topology on  $M(A)$ . Let  $N$  be a point process on  $A$  and let  $P$  and  $E$  denote its distribution and expectation, respectively. The Laplace functional of  $N$  is a map which takes Borel measurable non-negative functions into  $[0, \infty)$ , defined by

$$\Phi_N(f) = E[\exp(-N(f))] = \int_{M(A)} \exp(-m(f)) P(dm). \quad (2.30)$$

The Laplace functionals give a useful criteria for the weak convergence of point processes:  $N^{(n)} \Rightarrow N$  if and only if  $\Phi_{N^{(n)}}(f) \rightarrow \Phi_N(f)$  for all  $f \in C_K^+(A)$ .

In Proposition 2.1 we are concerned with sequences of vectors of point processes so now we describe the Laplace functionals for such vectors. Let  $l \in \mathbb{N}$  and  $N = (N_1, \dots, N_l)$  be a random variable on  $M_{1,l}$  and again  $P$  and  $E$  denote its distribution and expectation, respectively. The Laplace functional of  $N$  is a map which takes a vector of  $l$  non-negative Borel measurable functions,  $f_1, \dots, f_l$  on  $H, \dots, H^l$ , respectively, into  $[0, \infty)$ , defined by

$$\begin{aligned} \Phi_N(f_1, \dots, f_l) &= E[\exp(-f_1(N_1) - \dots - f_l(N_l))] \\ &= \int_{M_{1,l}} \exp(-m_1(f_1) - \dots - m_l(f_l)) P(dm_1 \dots dm_l). \end{aligned} \quad (2.31)$$

As already stated in Proposition 2.1, the weak convergence we use is the one induced by the product vague topology on  $M_{1,l}$ . It is trivial to extend the characterization of the weak convergence of point processes by the Laplace functionals to the product space to get

$$N^{(n)} = (N_1^{(n)}, \dots, N_l^{(n)}) \Rightarrow N = (N_1, \dots, N_l), \quad (2.32)$$

if and only if for all  $f_1 \in C_K^+(H), \dots, f_l \in C_K^+(H^l)$ ,

$$\Phi_{N^{(n)}}(f_1, \dots, f_l) \rightarrow \Phi_N(f_1, \dots, f_l). \quad (2.33)$$

Thus, our goal is to show that  $\Phi_{\zeta^{(n),l^*}} \rightarrow \Phi_{\chi^{l^*}}$ .

For  $k = 1, \dots, l^* - 1$  and  $\mu|_k \in V|_k$ , let  $\mathcal{P}_{\mu|_k}^{(n)}$  be the probability distribution

$$\begin{aligned} \mathcal{P}_{\mu|_k}^{(n)}((u, \infty)) &= \mathcal{P}\left(G(\lambda(J(j|_k))) \geq uc_n(k) \mid J(j|_k) = \mu|_k\right), \quad u > 0, \\ &= \mathcal{P}\left(G(\lambda(\mu|_k)) \geq uc_n(k)\right), \end{aligned} \quad (2.34)$$

and for  $\mu|_{l^*} \in V|_{l^*}$ , let  $\mathcal{P}_{\mu|_{l^*}}^{(n)}$  be the probability distribution

$$\mathcal{P}_{\mu|_{l^*}}^{(n)}((u, \infty)) = \mathcal{P}\left(\Lambda(j|_{l^*}) \geq uc_n \mid J(j|_{l^*}) = \mu|_{l^*}\right), \quad u > 0, \quad (2.35)$$

where  $\Lambda$  is given by (2.22). We set

$$\bar{\nu}_k^{(n)}(u; \mu|_{k-1}) = \frac{a_n(k)}{n} \sum_{\mu|_k=1}^n \mathcal{P}_{\mu|_k}^{(n)}((u, \infty)), \quad u > 0. \quad (2.36)$$

For simplicity, in the rest of this section we write  $\alpha_1, \dots, \alpha_L$  for  $\alpha_{1,L}, \dots, \alpha_{L,L}$ . The following lemma is the basic step in proving Proposition 2.1.

**Lemma 2.2.** For each  $k = 1, \dots, l^*$ , there exist a subset  $\Omega_{k,L} \subseteq \Omega$  with  $\mathbb{P}(\Omega_{k,L}) = 1$  such that the following holds:

(i) for  $k = 1, \dots, l^* - 1$ , for any environment in  $\Omega_{k,L}$ , for any  $u > 0$ , uniformly in  $\mu|_{k-1}$ , as  $n \rightarrow \infty$

$$\bar{\nu}_k^{(n)}(u; \mu|_{k-1}) \rightarrow D_k u^{-\alpha_k}, \quad (2.37)$$

where  $D_k = \Gamma(1 + \alpha_k)$ ,

(ii) for  $k = l^*$ , for any environment in  $\Omega_{l^*,L}$ , for any  $u > 0$ , uniformly in  $\mu|_{l^*-1}$ , as  $n \rightarrow \infty$

$$\bar{\nu}_{l^*}^{(n)}(u; \mu|_{l^*-1}) \rightarrow D_{l^*} u^{-\alpha_{l^*}}, \quad (2.38)$$

where  $D_{l^*}$  is positive a deterministic constant.

We first finish the proof of Proposition 2.1 using Lemma 2.2.

*Proof of Proposition 2.1.* Choose the constants  $D_1, \dots, D_{l^*}$  as in Lemma 2.2 and write the measures

$$\nu_k(dx_k) = D_k \alpha_k x_k^{-1-\alpha_k}, \quad x_k \geq 0. \quad (2.39)$$

We calculate the Laplace functional of  $\chi^{l^*}$ . Let  $\phi_k : C_K^+(H^k) \rightarrow C_K^+(H^{k-1})$  be

$$\phi_k f_k(t_1, x_1, \dots, t_{k-1}, x_{k-1}) = \int \int (1 - e^{-f_k(t_1, x_1, \dots, t_{k-1}, x_{k-1}, t_k, x_k)}) dt_k \nu_k(dx_k).$$

Note that  $\chi_1^{l^*}$  is a PPP on  $H$  with mean measure  $dt \times D_1 \alpha_1 x_1^{-1-\alpha_1} dx_1$ . Hence,

$$\Phi_{\chi_1^{l^*}}(f_1) = \exp(-\phi_1(f_1)). \quad (2.40)$$

For  $k = 2, \dots, l^*$ , using the correlation structure of  $\chi^{l^*}$  we get

$$\begin{aligned} & E \left[ \exp\left(-\sum_{j_1 \in \mathbb{N}} f_1(t_{j_1}, \eta_{j_1}) - \dots - \sum_{j|k \in \mathbb{N}^k} f_k(t_{j_1}, \eta_{j_1}, \dots, t_{j|k}, \eta_{j|k})\right) \middle| \chi_1^{l^*}, \dots, \chi_{k-1}^{l^*} \right] \\ &= \exp\left(-\sum_{j_1 \in \mathbb{N}} f_1(t_{j_1}, \eta_{j_1}) - \dots - \sum_{j|_{k-1} \in \mathbb{N}^{k-1}} f_{k-1}(t_{j_1}, \eta_{j_1}, \dots, t_{j|_{k-1}}, \eta_{j|_{k-1}})\right) \times \\ & \quad \prod_{j|_{k-1} \in \mathbb{N}^{k-1}} E \left[ \exp\left(-\sum_{j_k \in \mathbb{N}} f_k(t_{j_1}, \eta_{j_1}, \dots, t_{j|_{k-1}}, \eta_{j|_{k-1}}, t_{j|_k}, \eta_{j|_k})\right) \middle| \chi_{k-1}^{l^*} \right] \\ &= \exp\left(-\sum_{j_1 \in \mathbb{N}} f_1(t_{j_1}, \eta_{j_1}) - \dots - \sum_{j|_{k-1} \in \mathbb{N}^{k-1}} (f_{k-1} + \phi_k f_k)(t_{j_1}, \eta_{j_1}, \dots, t_{j|_{k-1}}, \eta_{j|_{k-1}})\right). \end{aligned}$$

In the last step above we used part (i) of the description of RPC and the Laplace transform of a PPP. Taking the expectations of first and last terms in the above display gives

$$\Phi_{(\chi_1^{l^*}, \dots, \chi_k^{l^*})}(f_1, \dots, f_k) = \Phi_{(\chi_1^{l^*}, \dots, \chi_{k-1}^{l^*})}(f_1, \dots, f_{k-2}, f_{k-1} + \phi_k f_k). \quad (2.41)$$

Thus, the Laplace functional  $\Phi_{\chi^{l^*}}$  is given by (2.41) recursively from  $k = l^*$  to  $k = 2$  and by (2.40).

We proceed the proof by induction. Recall that  $\xi_{j_1}^{(n)} \stackrel{d}{=} G(\lambda(J(j_1)))/c_n(1)$ . Hence, since  $\{J(j_1) : j_1 \in \mathbb{N}\}$  is i.i.d. so is  $\{\xi_{j_1}^{(n)} : j_1 \in \mathbb{N}\}$ , and

$$a_n(1) \mathcal{P}(\xi_{j_1}^{(n)} \geq u) = \frac{a_n(1)}{n} \sum_{\mu_1=1}^n \mathcal{P}_{\mu_1}^{(n)}((u, \infty)) =: \bar{\nu}_1^{(n)}(u), \quad u > 0. \quad (2.42)$$

Therefore, by Lemma 2.2 and Proposition 3.21 in [Res08], for any environment in  $\Omega_{1,L}$ ,

$$\zeta_1^{(n),l^*} \Rightarrow \chi_1^{l^*}. \quad (2.43)$$

Let  $k \in \{2, \dots, l^*\}$  and assume that  $\exists \Omega'_{k-1,L} \subseteq \Omega$  with  $\mathbb{P}(\Omega'_{k-1,L}) = 1$  such that for any environment in  $\Omega'_{k-1,L}$ ,

$$(\zeta_1^{(n),l^*}, \dots, \zeta_{k-1}^{(n),l^*}) \Rightarrow (\chi_1^{l^*}, \dots, \chi_{k-1}^{l^*}). \quad (2.44)$$

Let  $\mathcal{G}_{k-1}^{(n)}$  and  $\mathcal{F}_{k-1}^{(n)}$  be the  $\sigma$ -algebras

$$\mathcal{G}_{k-1}^{(n)} = \sigma\left(\zeta_1^{(n),l^*}, \dots, \zeta_{k-1}^{(n),l^*}\right), \quad \mathcal{F}_{k-1}^{(n)} = \sigma\left(J(j|i) : j|_{k-1} \in \mathbb{N}^{k-1}, i = 1, \dots, k-1\right). \quad (2.45)$$

Observe that then,

$$\mathcal{E} \left[ \exp\left(-\zeta_1^{(n),l^*}(f_1) - \dots - \zeta_k^{(n),l^*}(f_k)\right) \middle| \mathcal{G}_{k-1}^{(n)} \right] \quad (2.46)$$

$$= \mathcal{E} \left[ \mathcal{E} \left[ \exp\left(-\zeta_1^{(n),l^*}(f_1) - \dots - \zeta_k^{(n),l^*}(f_k)\right) \middle| \mathcal{G}_{k-1}^{(n)}, \mathcal{F}_{k-1}^{(n)} \right] \middle| \mathcal{G}_{k-1}^{(n)} \right] \quad (2.47)$$

$$= \exp\left(-\zeta_1^{(n),l^*}(f_1) - \dots - \zeta_{k-1}^{(n),l^*}(f_{k-1}) - g^{(n)}(\zeta_1^{(n),l^*}, \dots, \zeta_{k-1}^{(n),l^*})\right), \quad (2.48)$$

where for  $m = (m_1, \dots, m_{k-1}) \in \mathcal{M}_{1,k-1}$  with marks  $\{(t_{j|i}, x_{j|i}) : j|_{k-1} \in \mathbb{N}^{k-1}, i = 1, \dots, k-1\}$ ,  $g^{(n)}(m)$  is given by

$$- \sum_{j|_k \in \mathbb{N}^k} \log \mathcal{E} \left[ 1 - \int \left(1 - e^{-f_k(t_{j_1, x_{j_1}}, \dots, t_{j_{k-1}, x_{j_{k-1}}}, j_k/a_n(k), x_k)}\right) \frac{\nu_k^{(n)}(dx_k; J(j|_{k-1}))}{a_n(k)} \right]. \quad (2.49)$$

By Lemma 2.2, for any environment in  $\Omega_{k,L}$  where  $\Omega_{k,L}$  is as given in Lemma 2.2, for any  $m \in \mathcal{M}_{1,k-1}$ ,

$$g^{(n)}(m) \rightarrow \sum_{j|_{k-1} \in \mathbb{N}^{k-1}} \phi_k f_k(t_{j_1, x_{j_1}}, \dots, t_{j_{k-1}, x_{j_{k-1}}}) = m_{k-1}(\phi_k f_k). \quad (2.50)$$

By Proposition 3.16 in [Res08], we have for any  $A \subseteq \mathcal{M}_{1,k-1}$  relatively compact and  $F \subseteq H^{k-1}$  also relatively compact,

$$\sup_{m \in A} \sum_{j|_{k-1} \in \mathbb{N}^{k-1}} \mathbb{1} \{(t_{j_1, x_{j_1}}, \dots, t_{j_{k-1}, x_{j_{k-1}}}) \in F\} < \infty. \quad (2.51)$$

This, together with Lemma 2.2, yield

$$\sup_{m \in A} \|g^{(n)}(m) - m_{k-1}(\phi_k f_k)\| \rightarrow 0. \quad (2.52)$$

Hence,  $g^{(n)}(m) \rightarrow m_{k-1}(\phi_k f_k)$  uniformly on compact sets. Also, it is obvious that  $\|\exp(-g^{(n)}(m))\| \leq 1$  for all  $m \in \mathcal{M}_{1,k-1}$ . Hence, by taking the expectation of (2.48) and using the induction step (2.44), we conclude that for any environment in  $\Omega'_{k,L} = \Omega'_{k-1,L} \cap \Omega_{k,L}$ , as  $n \rightarrow \infty$

$$\begin{aligned} & \mathcal{E} \left[ \exp\left(-\zeta_1^{(n),l^*}(f_1) - \dots - \zeta_k^{(n),l^*}(f_k)\right) \right] \\ & \rightarrow \Phi_{(\chi_1^{l^*}, \dots, \chi_{k-1}^{l^*})}(f_1, \dots, f_{k-2}, f_{k-1} + \phi_k f_k). \end{aligned} \quad (2.53)$$

This and (2.43) finish the proof of Proposition 2.1 with  $\Omega'_L = \Omega'_{l^*,L}$ .  $\square$



*Proof of Lemma 2.2 part (i).* We first prove that for any  $u > 0$ ,  $\exists \Omega_{k,L}^u \subset \Omega$  with  $\mathbb{P}(\Omega_{k,L}^u) = 1$  such that for every environment in  $\Omega_{k,L}^u$ , uniformly in  $\mu|_{k-1}$ , as  $n \rightarrow \infty$   $\bar{\nu}_k^{(n)}(u; \mu|_{k-1}) \rightarrow \Gamma(1 + \alpha_k)u^{-\alpha_k}$ . We first calculate the expectation of  $\bar{\nu}_k^{(n)}(u; \mu|_{k-1})$  over the environment. Note that under  $\mathbb{P}$ ,  $\left\{ \mathcal{P}_{\mu|_k}^{(n)}(u, \infty) : \mu|_k \in V|_k \right\}$  is an i.i.d. collection. Hence, using (1.4) and (2.34) we have

$$\begin{aligned} \mathbb{E}[\bar{\nu}_k^{(n)}(u; \mu|_{k-1})] &= a_n(k) \mathbb{E} \left[ (1 - \lambda(\mu|_k))^{\lfloor uc_n(k) \rfloor} \right] \\ &= a_n(k) \int_0^1 (1-x)^{\lfloor uc_n(k) \rfloor} x^{\alpha_k-1} dx \\ &= a_n(k) \int_0^\infty C_k^{(n)}(z) \alpha_k z^{-1+\alpha_k} u^{-\alpha_k} (c_n(k))^{-\alpha_k} dz \end{aligned}$$

where we performed the change of variables  $x \rightarrow uc_n(k)/z$  and set

$$C_k^{(n)}(z) = \left( 1 - \frac{z}{uc_n(k)} \right)^{\lfloor uc_n(k) \rfloor} \mathbb{1} \{z \leq uc_n(k)\}. \quad (2.54)$$

Since  $k < l^*$ , by the scaling relations in (1.25), we have  $a_n(k) = (c_n(k))^{\alpha_k}$ . Using the bounds  $uc_n(k) - 1 < \lfloor uc_n(k) \rfloor \leq uc_n(k)$  and the fact that  $c_n(k)$  is diverging, we have for all  $z \geq 0$ ,  $C_k^{(n)}(z) \rightarrow e^{-z}$  and  $C_k^{(n)}(z) \leq e^{-z}$ . Hence, by the dominated convergence theorem, we can conclude that

$$\mathbb{E}[\bar{\nu}_k^{(n)}(u; \mu|_{k-1})] \longrightarrow u^{-\alpha_k} \alpha_k \Gamma(\alpha_k) = u^{-\alpha_k} \Gamma(1 + \alpha_k). \quad (2.55)$$

To control the fluctuations of  $\bar{\nu}_k^{(n)}(u; \mu|_{k-1})$  we write, for  $\theta_n \geq 0$ ,

$$\begin{aligned} \mathbb{P} \left( \exists \mu|_{k-1} \in V|_{k-1} : \left| \bar{\nu}_k^{(n)}(u; \mu|_{k-1}) - \mathbb{E}[\bar{\nu}_k^{(n)}(u; \mu|_{k-1})] \right| \geq \theta_n \right) \\ \leq n^{k-1} \mathbb{P} \left( \left| \bar{\nu}_k^{(n)}(u; \mu|_{k-1}) - \mathbb{E}[\bar{\nu}_k^{(n)}(u; \mu|_{k-1})] \right| \geq \theta_n \right), \end{aligned} \quad (2.56)$$

and bound the above probabilities using Bennett's bound (see [Ben62]), which states that if  $(X(\mu) : \mu = 1, \dots, n)$  is a family of centered i.i.d. random variables that satisfies  $\max_{\mu=1, \dots, n} |X(\mu)| \leq a$ , then for any  $b^2 \geq \sum_{\mu=1}^n \mathbb{E}X^2(\mu)$  and  $t < b^2/(2a)$ ,

$$\mathbb{P} \left( \left| \sum_{\mu=1}^n X(\mu) \right| > t \right) \leq \exp \left( -\frac{t^2}{4b^2} \right). \quad (2.57)$$

For fixed  $\mu|_{k-1}$ , let  $\{X(\mu_k) : \mu_k = 1, \dots, n\}$  be the collection of i.i.d. random variables given by

$$X(\mu_k) = \mathcal{P} \left( G(\lambda(\mu|_k)) \geq uc_n(k) \right) - \mathbb{E} \left[ \mathcal{P} \left( G(\lambda(\mu|_k)) \geq uc_n(k) \right) \right]. \quad (2.58)$$

Then,  $\max |X(\mu_k)| \leq 2$  and  $\sum_{\mu_k} \mathbb{E}[X^2(\mu_k)] \leq Cn/a_n(k)$  for  $n$  large enough, for some positive constant  $C$ . Hence, we can choose  $b^2 = Cn/a_n(k)$  and get that (2.56) is bounded above by

$$n^{k-1} \exp \left( -C' \frac{n\theta_n^2}{a_n(k)} \right), \quad (2.59)$$

provided that  $\theta_n < C/2$ . Since  $k < l^*$ , the ratio  $n/a_n(k)$  diverges at least polynomially fast and thus, we can choose  $\theta_n$  such that  $\lim_{n \rightarrow \infty} \theta_n = 0$  and  $n\theta_n^2/a_n(k)$  diverges at least polynomially fast. Hence,

$$\mathbb{P} \left( \exists \mu|_{k-1} \in V|_{k-1} : \left| \bar{\nu}_k^{(n)}(u; \mu|_{k-1}) - \mathbb{E}[\bar{\nu}_k^{(n)}(u; \mu|_{k-1})] \right| \geq \theta_n \right) \leq \exp(-n^c) \quad (2.60)$$

for some  $c > 0$ . Borel-Cantelli Lemma and (2.55) prove that  $\mathbb{P}(\Omega_{k,L}^u) = 1$ , where for any  $u > 0$ ,  $\Omega_{k,L}^u$  is the set of environments where  $\bar{\nu}_k^{(n)}(u; \mu|_{k-1}) \rightarrow \bar{\nu}_k(u)$ , uniformly in  $\mu|_{k-1} \in V|_{k-1}$ . Let  $\mathbb{Q}^+$  be the set of positive rational numbers. The fact that  $\bar{\nu}^{(n)}(u, \mu|_{k-1})$  is monotone,  $\bar{\nu}_k^{(n)}(u)$  is continuous and the countable intersection of probability one events is also a probability one event finish the proof with  $\Omega_{k,L} = \bigcap_{u \in \mathbb{Q}^+} \Omega_{k,L}^u$ .  $\square$

*Proof of Lemma 2.2 part (ii).* For  $m = 1, \dots, L$  and  $\mu|_{m-1} \in V|_{m-1}$ , we define random variables  $\Theta_m(\cdot; \mu|_{m-1})$  recursively, from  $m = L$  to  $m = 1$ , as follows. Let

$$\Theta_L(r; \mu|_{L-1}) = \sum_{j=1}^r \lambda^{-1}(\mu|_{L-1} J(j)) e_j, \quad r \in \mathbb{N}, \quad (2.61)$$

where the independent collections  $\{J(j) : j \in \mathbb{N}\}$  and  $\{e_j : j \in \mathbb{N}\}$  are i.i.d. with uniform distribution on  $[n]$  and the distribution of a mean one exponential random variable, respectively. Also, let

$$\Theta_m(r; \mu|_{m-1}) = \sum_{j=1}^r \Theta_{m+1}^{(j)}(G(\lambda(\mu|_{m-1} J(j))); \mu|_{m-1} J(j)) \quad (2.62)$$

where  $\{J(j) : j \in \mathbb{N}\}$  is i.i.d. with uniform distribution on  $[n]$ , and given this collection,  $\{\Theta_{m+1}^{(j)}(\cdot; \mu|_{m-1} J(j)) : j \in \mathbb{N}\}$  and  $\{G(\lambda(\mu|_{m-1} J(j))) : j \in \mathbb{N}\}$  are independent collections of independent random variables with each  $\Theta_{m+1}^{(j)}(\cdot; \mu|_{m-1} J(j))$  having the distribution of  $\Theta_{m+1}(\cdot; \mu|_{m-1} J(j))$ , and  $G(\lambda(\mu|_{m-1} J(j)))$  having the distribution of a geometric random variable with probability of success  $\lambda(\mu|_{m-1} J(j))$ . In words,  $\Theta_m(\cdot; \mu|_{m-1})$  has the distribution of the first level clock process of the GREM-like trap model, reduced to the subtree attached to  $\mu|_{m-1}$ . Note that the distribution of  $\Theta_m(\cdot; \mu|_{m-1})$  is i.i.d. in  $\mu|_{m-1}$ . Finally, for  $m = l^* + 1, \dots, L$  and  $\mu|_{m-1} \in V|_{m-1}$ , we define

$$\Theta_m^{(n)}(r; \mu|_{m-1}) = \frac{\Theta_m(\lfloor r a_n(m) \rfloor; \mu|_{m-1})}{c_n}, \quad r > 0. \quad (2.63)$$

**Lemma 2.3.** For any  $m = l^* + 1, \dots, L$ ,  $r > 0$  and  $u > 0$ ,

$$\mathbb{E} \left[ \mathcal{P}(\Theta_m^{(n)}(r; \mu|_{m-1}) \geq u) \right] \longrightarrow P(r Z_m \geq u), \quad (2.64)$$

where  $Z_m$  is a positive random variable whose Laplace transform is given by

$$E[e^{-\kappa Z_m}] = e^{-d_m \kappa^{\alpha_m}}, \quad \kappa > 0 \quad (2.65)$$

for some  $d_m > 0$ .

*Proof of Lemma 2.3.* Note that it is enough to show that for any  $r > 0$  and  $\kappa > 0$

$$\mathbb{E} \left[ \mathcal{E} \left[ \exp(-\kappa \Theta_m^{(n)}(r; \mu|_{m-1})) \right] \right] \longrightarrow e^{-d_m (\kappa r)^{\alpha_m}}. \quad (2.66)$$

We use the following fact several times in the proof: Let  $X_n, X$  be random variables on  $[0, 1]$  and  $a_n$  be a diverging scale, then

$$E[X_n^{r a_n}] \xrightarrow{n \rightarrow \infty} E[e^{-r X}] \quad \forall r > 0 \Leftrightarrow E[e^{-r a_n (1 - X_n)}] \xrightarrow{n \rightarrow \infty} E[e^{-r X}] \quad \forall r > 0. \quad (2.67)$$

We proceed the proof by induction. We first prove the case where  $m = L$ . By (2.61) and (2.63),

$$\mathcal{E} \left[ \exp \left( -\kappa \Theta_L^{(n)}(r; \mu|_{L-1}) \right) \right] = \left\{ \sum_{\mu_L=1}^n \frac{1}{n} \frac{1}{1 + \kappa \lambda^{-1}(\mu|_L)/c_n} \right\}^{\lfloor ra_n(L) \rfloor}. \quad (2.68)$$

By (1.25), we have  $a_n(L) \gg n$  and  $c_n = c_n(L) = a_n(L)n^{1/\alpha_L-1}$ . Thus,

$$\begin{aligned} & n \mathbb{E} \left[ 1 - \exp \left( -\frac{a_n(L)}{n} \frac{r \kappa \lambda^{-1}(\mu|_L)/c_n(L)}{1 + \kappa \lambda^{-1}(\mu|_L)/c_n(L)} \right) \right] \\ &= \int_{n^{-1/\alpha_L}}^{\infty} \left\{ 1 - \exp \left( -\frac{r \kappa y}{1 + \kappa y n/a_n(L)} \right) \right\} \alpha_L y^{-1-\alpha_L} dy \\ &\xrightarrow{n \rightarrow \infty} \int_0^{\infty} \left( 1 - \exp(-r \kappa y) \right) \alpha_L y^{-1-\alpha_L} dy = \Gamma(1 - \alpha_L)(r \kappa)^{\alpha_L}. \end{aligned}$$

In taking the above limit we used  $\mathbb{P}(\lambda^{-1}(\mu|_L) \geq u) = u^{-\alpha_L}$ , the bound  $1 - e^{-cy} \leq 1 \wedge cy$  and the dominated convergence theorem. By the independence structure of the landscape,

$$\begin{aligned} & \mathbb{E} \left[ \exp \left( -\sum_{\mu_L=1}^n \frac{a_n(L)}{n} \frac{r \kappa \lambda^{-1}(\mu|_L)/c_n(L)}{1 + \kappa \lambda^{-1}(\mu|_L)/c_n(L)} \right) \right] \\ &= \left( 1 - \mathbb{E} \left[ 1 - \exp \left( -\frac{a_n(L)}{n} \frac{r \kappa \lambda^{-1}(\mu|_L)/c_n(L)}{1 + \kappa \lambda^{-1}(\mu|_L)/c_n(L)} \right) \right] \right)^n \\ &\xrightarrow{n \rightarrow \infty} \exp \left( -\Gamma(1 - \alpha_L)(r \kappa)^{\alpha_L} \right). \end{aligned} \quad (2.69)$$

Finally, using (2.67) with  $X_n = \sum_{\mu_L=1}^n \frac{1}{n} \frac{1}{1 + \kappa \lambda^{-1}(\mu|_L)/c_n(L)}$  and  $a_n = a_n(L)$  finish the proof of (2.66) for  $m = L$ , where  $d_L = \Gamma(1 - \alpha_L)$ .

Now assume that (2.66) is true for  $m + 1$ . By (2.62), we have

$$\mathcal{E} \left[ \exp \left( -\kappa \Theta_m(1; \mu|_{m-1}) \right) \right] = \sum_{\mu_m=1}^n \frac{1}{n} \mathcal{E} \left[ \exp \left( -\kappa \Theta_{m+1}(G(\lambda(\mu|_m)); \mu|_m) \right) \right]. \quad (2.70)$$

Using the geometric distribution and (2.62) we get

$$\mathcal{E} \left[ \exp \left( -\kappa \Theta_{m+1}(G(\lambda(\mu|_m)); \mu|_m) \right) \right] = \frac{1}{1 + \lambda^{-1}(\mu|_m) \varphi_{m+1}(\kappa; \mu|_m)},$$

where

$$\varphi_{m+1}(\kappa; \mu|_m) = \frac{1 - \mathcal{E}[e^{-\kappa \Theta_{m+1}(1; \mu|_m)}]}{\mathcal{E}[e^{-\kappa \Theta_{m+1}(1; \mu|_m)}]}. \quad (2.71)$$

Hence,

$$\mathcal{E} \left[ \exp \left( -\kappa \Theta_m^{(n)}(r; \mu|_{m-1}) \right) \right] = \left\{ \sum_{\mu_m} \frac{1}{n} \frac{1}{1 + \lambda^{-1}(\mu|_m) \varphi_{m+1}(\kappa/c_n; \mu|_m)} \right\}^{\lfloor ra_n(m) \rfloor}.$$

By the induction step, for any  $\kappa, r > 0$ ,

$$\mathbb{E} \left[ \mathcal{E} \left[ e^{-\frac{\kappa}{c_n} \Theta_{m+1}(1; \mu|_m)} \right]^{\lfloor ra_n(m+1) \rfloor} \right] \longrightarrow E[e^{-\kappa r Z_{m+1}}], \quad (2.72)$$

and consequently,

$$\mathbb{E} \left[ \exp(-ra_n(m+1)\varphi_{m+1}(\kappa/c_n; \mu|_m)) \right] \longrightarrow E[e^{-r\kappa Z_{m+1}}] = e^{-d_{m+1}(r\kappa)^{\alpha_{m+1}}}. \quad (2.73)$$

We calculate

$$\begin{aligned} & n\mathbb{E} \left[ 1 - \exp \left( -\frac{a_n(m)}{n} \frac{r\lambda^{-1}(\mu|_m)\varphi_{m+1}(\kappa/c_n; \mu|_m)}{1 + \lambda^{-1}(\mu|_m)\varphi_{m+1}(\kappa/c_n; \mu|_m)} \right) \right] \\ &= \mathbb{E} \left[ n \int_1^\infty \left( 1 - \exp \left( -\frac{a_n(m)}{n} \frac{rx\varphi_{m+1}(\kappa/c_n; \mu|_m)}{1 + x\varphi_{m+1}(\kappa/c_n; \mu|_m)} \right) \right) \alpha_m x^{-1-\alpha_m} dx \right]. \end{aligned}$$

Using the change of variables  $x = yn^{1/\alpha_m}$  and  $a_n(m+1) = c_n(m) = a_n(m)n^{1/\alpha_m-1}$  we get that the above display is equal to

$$\int_{n^{-1/\alpha_m}}^\infty \mathbb{E} \left[ 1 - \exp \left( -\frac{ry\varphi_{m+1}(\kappa/c_n; \mu|_m)a_n(m+1)}{1 + y\varphi_{m+1}(\kappa/c_n; \mu|_m)a_n(m+1)n/a_n(m)} \right) \right] \alpha_{m+1}y^{-1-\alpha_{m+1}} dy. \quad (2.74)$$

Since  $\alpha_m < \alpha_{m+1}$  and  $n \ll a_n(m)$ , by (2.73), we can conclude that the sequence in (2.74) converges to

$$\int_0^\infty \left( 1 - \exp(-d_{m+1}(ry\kappa)^{\alpha_{m+1}}) \right) \alpha_m y^{-1-\alpha_m} dy = d_m(r\kappa)^{\alpha_m} \quad (2.75)$$

where  $d_m = d_{m+1}^{\alpha_m/\alpha_{m+1}} \Gamma(1 - \frac{\alpha_m}{\alpha_{m+1}})$ . A calculation as in (2.69) and (2.67) finish the proof.  $\square$

As before, we have  $\mathbb{E}[\bar{V}_{l^*}^{(n)}(u; \mu|_{l^*-1})] = a_n(l^*)\mathbb{E}[\mathcal{P}_{\mu|_{l^*}}^{(n)}((u, \infty))]$ . Note that the distribution of  $\Lambda(j|_{l^*})$ , when  $J(j|_{l^*}) = \mu|_{l^*}$ , is that of  $\Theta_{l^*+1}(G(\lambda(\mu|_{l^*})); \mu|_{l^*})$  where  $\Theta_{l^*+1}(\cdot; \mu|_{l^*})$  and  $G(\lambda(\mu|_{l^*}))$  are independent. Hence, by (2.35),

$$\mathcal{P}_{\mu|_{l^*}}^{(n)}((u, \infty)) = \sum_{i=1}^\infty \lambda(\mu|_{l^*})(1 - \lambda(\mu|_{l^*}))^{i-1} \mathcal{P}(\Theta_{l^*+1}(i; \mu|_{l^*}) \geq uc_n). \quad (2.76)$$

Thus,

$$\mathbb{E}[\bar{V}_{l^*}^{(n)}(u; \mu|_{l^*-1})] = \sum_{i=1}^\infty g_{n,u} \left( \frac{i}{c_n(l^*)} \right) h_n \left( \frac{i}{c_n(l^*)} \right) \frac{1}{c_n(l^*)} \quad (2.77)$$

where

$$g_{n,u}(r) = \mathbb{E} \left[ \mathcal{P}(\Theta_{l^*+1}^{(n)}(r; \mu|_{l^*}) \geq u) \right], \quad (2.78)$$

and

$$h_n(r) = a_n(l^*)c_n(l^*)\mathbb{E} \left[ \lambda(\mu|_{l^*})(1 - \lambda(\mu|_{l^*}))^{\lfloor rc_n(l^*) \rfloor - 1} \right]. \quad (2.79)$$

By Proposition 2.3 we get

$$g_{n,u}(r) \longrightarrow g_u(r) := P(rZ_{l^*+1} \geq u), \quad (2.80)$$

and a simple calculation yields

$$h_n(r) \longrightarrow \alpha_{l^*} r^{-1-\alpha_{l^*}} \Gamma(1 + \alpha_{l^*}). \quad (2.81)$$

Using (2.65) and a Tauberian theorem (see e.g. Corollary 8.1.7 in [BGT87]) we can conclude that there exists a  $C > 0$  such that for all  $r$  small enough,

$$g_u(r) \leq Cr^{\alpha_{l^*+1}} u^{-\alpha_{l^*+1}}. \quad (2.82)$$

Via (2.77) and (2.80)-(2.82),

$$\mathbb{E}[\bar{\nu}_{l^*}^{(n)}(u; \mu|_{l^*-1})] \longrightarrow \Gamma(1 + \alpha_{l^*}) \int_0^\infty g_u(r) \alpha_{l^*} r^{-1-\alpha_{l^*}} dr. \quad (2.83)$$

We use Bennett's bound once again to finish the proof. For fixed  $\mu|_{l^*-1}$ , consider the collection  $\{X_{\mu_{l^*}} : \mu_{l^*} \in \mathcal{M}_{l^*}\}$  where

$$X_{\mu_{l^*}} = \mathcal{P}_{\mu_{l^*}}^{(n)}((u, \infty)) - \mathbb{E}[\mathcal{P}_{\mu_{l^*}}^{(n)}((u, \infty))]. \quad (2.84)$$

Clearly,  $\max_{\mu_{l^*}} |X_{\mu_{l^*}}| \leq 2$ , and using Jensen's inequality we get

$$\sum_{\mu_{l^*}} \mathbb{E}[X_{\mu_{l^*}}^2] \leq C \frac{n}{a_n(l^*)}. \quad (2.85)$$

Hence, since  $a_n(l^*) \ll n$ , we can proceed exactly as in the proof of Lemma 2.2 part (i) to conclude that there exists  $\Omega_{l^*,L} \subset \Omega$  with  $\mathbb{P}(\Omega_{l^*,L}) = 1$  such that for any environment in  $\Omega_{l^*,L}$ , for any  $u > 0$ , uniformly in  $\mu|_{l^*-1}$

$$\bar{\nu}_{l^*}^{(n)}(u; \mu|_{l^*-1}) \longrightarrow \Gamma(1 + \alpha_{l^*}) \int_0^\infty g_u(r) \alpha_{l^*} r^{-1-\alpha_{l^*}} dr. \quad (2.86)$$

We have

$$\int_0^\infty g_u(r) \alpha_{l^*} r^{-1-\alpha_{l^*}} dr = \int_0^\infty P(Z_{l^*+1} \geq u/r) \alpha_{l^*} r^{-1-\alpha_{l^*}} dr \quad (2.87)$$

$$= u^{-\alpha_{l^*}} \int_0^\infty P(Z_{l^*+1} \geq s) \alpha_{l^*} s^{-1+\alpha_{l^*}} ds. \quad (2.88)$$

Since  $\alpha_{l^*} < \alpha_{l^*+1}$ , by the Laplace transform, (2.65), of  $Z_{l^*+1}$ ,

$$D_{l^*} = \Gamma(1 + \alpha_{l^*}) \int_0^\infty P(Z_{l^*+1} \geq s) \alpha_{l^*} s^{\alpha_{l^*}-1} ds < \infty. \quad (2.89)$$

Thus, we are finished with the proof of Lemma 2.2 part (ii).  $\square$

### 2.3 Continuity of functionals on the space cascade of point measures

In this subsection we prove that functionals used to define the clock processes are, after certain truncations, continuous.

For  $\gamma > 0$ , let  $\mathcal{M}_{1,l}^{(\gamma)}$  be the subset of  $\mathcal{M}_{1,l}$  such that for  $m \in \mathcal{M}_{1,l}^{(\gamma)}$ , whose marks are given by

$$\{(t_{j|k}, x_{j|k}) : j|l \in \mathbb{N}^l, k = 1, \dots, l\}, \quad (2.90)$$

it holds true that

$$m_1((0, \infty) \times \{\gamma, \gamma^{-1}\}) = 0, \quad (2.91)$$

and for  $k = 1, \dots, l-1$  and  $j|_k \in \mathbb{N}^k$ ,

$$m_{k+1}\left(\{(t_{j_1}, x_{j_1}, \dots, t_{j|k}, x_{j|k})\} \times \partial((0, x_{j|k}] \times (\gamma, \gamma^{-1}))\right) = 0. \quad (2.92)$$

Here,  $\partial$  denotes the boundary of a set. For  $m \in \mathcal{M}_{1,l}$ , let  $m^{(\gamma)} = (m_1^{(\gamma)}, \dots, m_l^{(\gamma)})$  be

$$m_k^{(\gamma)} = \sum_{j|_k \in \mathbb{N}^k} \varepsilon_{(t_{j_1}, x_{j_1}, \dots, t_{j_k}, x_{j_k})} \prod_{i=1}^k \mathbb{1}\{x_{j_i} \in (\gamma, \gamma^{-1})\}. \quad (2.93)$$

Note that then,  $m^{(\gamma)} \in \widetilde{\mathcal{M}}_{1,l}$ . We introduce the maps  $T_l^{(\gamma)} : \mathcal{M}_{1,l} \rightarrow D([0, \infty))$  and  $\overline{T}_l^{(\gamma)} : \mathcal{M}_{1,l} \rightarrow \widetilde{\mathcal{M}}_{1,l-1}$  by

$$T_l^{(\gamma)}(m) = T_l(m^{(\gamma)}), \quad \text{and} \quad \overline{T}_l^{(\gamma)}(m) = \overline{T}_l(m^{(\gamma)}). \quad (2.94)$$

**Lemma 2.4.** *For any  $\gamma > 0$ , the maps  $T_l^{(\gamma)}$  and  $\overline{T}_l^{(\gamma)}$  are continuous on  $\mathcal{M}_{1,l}^{(\gamma)}$ .*

We use the following proposition which is also a generalization from the usual point processes.

**Proposition 2.5.** *Let  $A_1, \dots, A_l \subset H$  be compact sets with*

$$m_1(\partial A_1) = \dots = m_l(\partial(A_1 \times \dots \times A_l)) = 0.$$

*Let  $m^{(n)}, m \in \mathcal{M}_{1,l}$  be such that  $m^{(n)} \xrightarrow{n \rightarrow \infty} m$ . Then, for any  $k = 1, \dots, l$ , after relabeling,*

$$m_k(\cdot \cap (A_1 \times \dots \times A_k)) = \sum_{j_1=1}^{q_1} \sum_{j_2=1}^{q_2(j_1)} \dots \sum_{j_k=1}^{q_k(j|_{k-1})} \varepsilon_{(t_{j_1}, x_{j_1}, \dots, t_{j_k}, x_{j_k})}$$

and

$$m_k^{(n)}(\cdot \cap (A_1 \times \dots \times A_k)) = \sum_{j_1=1}^{q_1} \sum_{j_2=1}^{q_2(j_1)} \dots \sum_{j_k=1}^{q_k(j|_{k-1})} \varepsilon_{(t_{j_1}^{(n)}, x_{j_1}^{(n)}, \dots, t_{j_k}^{(n)}, x_{j_k}^{(n)})}$$

*for all  $n \geq n(A_1, \dots, A_l)$ . Moreover, for any  $r = 1, \dots, k$ ;  $j_1 = 1, \dots, q_1$ ;  $j_2 = 1, \dots, q_2(j_1)$ ;  $\dots$ ;  $j_r = 1, \dots, q_r(j|_{r-1})$ , as  $n \rightarrow \infty$*

$$\| (t_{j_r}^{(n)}, x_{j_r}^{(n)}) - (t_{j_r}, x_{j_r}) \| \rightarrow 0.$$

*Proof of Proposition 2.5.* When  $k = 1$  we have the usual point processes case and it is covered by Proposition 3.13 in [Res08]. We now consider  $k = 2$ . Via, once again, Proposition 3.13 in [Res08], there exists a  $q_1 \in \mathbb{N}$  such that, after relabeling, for any  $n \geq n(A_1)$

$$m_1(\cdot \cap A_1) = \sum_{j_1=1}^{q_1} \varepsilon_{(t_{j_1}, x_{j_1})}, \quad m_1^{(n)}(\cdot \cap A_1) = \sum_{j_1=1}^{q_1} \varepsilon_{(t_{j_1}^{(n)}, x_{j_1}^{(n)})}, \quad (2.95)$$

and

$$\| (t_{j_1}^{(n)}, x_{j_1}^{(n)}) - (t_{j_1}, x_{j_1}) \| \rightarrow 0, \quad \forall j_1 = 1, \dots, q_1. \quad (2.96)$$

Let  $B_r(z)$  and  $\overline{B}_r(z)$  denote the open and closed ball, respectively, around  $z \in H^k$  with radius  $r$ . We choose  $\epsilon > 0$  small enough so that  $\overline{B}_\epsilon(t_1, x_1), \dots, \overline{B}_\epsilon(t_{q_1}, x_{q_1})$  are disjoint and for any  $j_1 = 1, \dots, q_1$ , we have  $m_1(\overline{B}_\epsilon((t_{j_1}, x_{j_1}))) = 1$  and  $\overline{B}_\epsilon(t_{j_1}, x_{j_1}) \subseteq A_1^o$ , where  $A_1^o$  denotes the interior of  $A_1$  (recall that we assume that all the point measures are simple). Hence, for any  $j_1 = 1, \dots, q_1$  we can find a  $q_2(j_1)$  such that, after relabeling,

$$m_2(\cdot \cap \overline{B}_\epsilon(t_{j_1}, x_{j_1}) \times A_2) = \sum_{j_2=1}^{q_2(j_1)} \varepsilon_{(t_{j_1}, x_{j_1}, t_{j_1 j_2}, x_{j_1 j_2})}. \quad (2.97)$$

We choose  $n$  large enough so that for any  $j_1 = 1, \dots, q_1$ , we have  $\| (t_{j_1}^{(n)}, x_{j_1}^{(n)}) - (t_{j_1}, x_{j_1}) \| < \epsilon/2$ . Thus, by (2.95), (2.97) and Proposition 3.13 in [Res08], after relabeling, for  $n \geq n(\bar{B}_\epsilon(t_{j_1}, x_{j_1}) \times A_2)$ ,

$$m_2^{(n)}(\cdot \cap \bar{B}_\epsilon(t_{j_1}, x_{j_1}) \times A_2) = \sum_{j_2=1}^{q_2(j_1)} \varepsilon_{(t_{j_1}^{(n)}, x_{j_1}^{(n)}, t_{j_1 j_2}^{(n)}, x_{j_1 j_2}^{(n)})} \quad (2.98)$$

and

$$\| (t_{j_1 j_2}^{(n)}, x_{j_1 j_2}^{(n)}) - (t_{j_1 j_2}, x_{j_1 j_2}) \| \rightarrow 0, \quad \forall j_2 = 1, \dots, q_2(j_1). \quad (2.99)$$

Also, by (2.95)-(2.99), for  $n \geq \max_{j_1=1, \dots, q_1} n(\bar{B}_\epsilon(t_{j_1}, x_{j_1}) \times A_2) = n(A_1, A_2)$ ,

$$m_2(\cdot \cap A_1 \times A_2) = \sum_{j_1=1}^{q_1} \sum_{j_2=1}^{q_2(j_1)} \varepsilon_{(t_{j_1}, x_{j_1}, t_{j_1 j_2}, x_{j_1 j_2})} \quad (2.100)$$

and

$$m_2^{(n)}(\cdot \cap A_1 \times A_2) = \sum_{j_1=1}^{q_1} \sum_{j_2=1}^{q_2(j_1)} \varepsilon_{(t_{j_1}^{(n)}, x_{j_1}^{(n)}, t_{j_1 j_2}^{(n)}, x_{j_1 j_2}^{(n)})}. \quad (2.101)$$

This proves the case  $k = 2$ . Iterating the exact same procedure finishes the proof.  $\square$

*Proof of Lemma 2.4.* Let  $m \in \mathcal{M}_{1,l}^{(\gamma)}$  and  $m^{(n)} \in \mathcal{M}_{1,l}$  with  $m^{(n)} \rightarrow m$  as  $n \rightarrow \infty$ . We first prove that  $T_l^{(\gamma)}(m^{(n)}) \rightarrow T_l^{(\gamma)}(m)$  on  $D([0, \infty))$  equipped with the Skorohod  $J_1$  topology. It is enough to show that  $T_l^{(\gamma)}(m^{(n)}) \rightarrow T_l^{(\gamma)}(m)$  on  $D([0, t'])$  for any continuity point  $t'$  of  $T_l^{(\gamma)}(m)$ , which is equivalent to  $m_1(\{t'\} \times (\gamma, \gamma^{-1})) = 0$ . Since  $m \in \mathcal{M}_{1,l}^{(\gamma)}$ , for any such  $t'$  there exists a  $q_1$  such that, after relabeling,

$$m_1(\cdot \cap (0, t'] \times (\gamma, \gamma^{-1})) = \sum_{j_1=1}^{q_1} \varepsilon_{(t_{j_1}, x_{j_1})} \quad (2.102)$$

Recursively, from  $k = 2$  to  $k = l$ , for any  $j_1 = 1, \dots, q_1; j_2 = 1, \dots, q_2(j_1); \dots; j_{k-1} = 1, \dots, q_{k-1}(j_{k-2})$ , there exists a  $q_k(j_{k-1})$  such that

$$\begin{aligned} & m_k(\cdot \cap (t_{j_1}, x_{j_1}, \dots, t_{j_{k-1}}, x_{j_{k-1}}) \times ((0, x_{j_{k-1}}] \times (\gamma, \gamma^{-1}))) \\ &= \sum_{j_k=1}^{q_k(j_{k-1})} \mathbb{1} \{ (t_{j_1}, x_{j_1}, \dots, t_{j_k}, x_{j_k}) \in \cdot \}. \end{aligned} \quad (2.103)$$

By (2.92), we can find a  $\delta > 0$  small enough so that, for any  $k = 1, \dots, l-1$  and  $j_1 = 1, \dots, q_1; j_2 = 1, \dots, q_2(j_1); \dots; j_k = 1, \dots, q_k(j_{k-1})$ ,

$$m_{k+1} \left( \{ (t_{j_1}, x_{j_1}, \dots, t_{j_k}, x_{j_k}) \} \times ([x_{j_k} - \delta, x_{j_k} + \delta] \times (\gamma, \gamma^{-1})) \right) = 0. \quad (2.104)$$

Since  $m_1$  is simple, using (2.91) we can choose an  $\epsilon < \delta/2$  small enough so that  $\bar{B}_\epsilon(t_1, x_1), \dots, \bar{B}_\epsilon(t_{q_1}, x_{q_1})$  are disjoint subsets of  $(0, t') \times (\gamma, \gamma^{-1})$ . Similarly, using (2.92) for  $\epsilon > 0$  small enough we have, for  $k = 1, \dots, l-1$  and for any  $j_1 = 1, \dots, q_1; j_2 = 1, \dots, q_2(j_1); \dots; j_k = 1, \dots, q_k(j_{k-1})$ ,

$$\{ \bar{B}_\epsilon(t_{j_{k+1}}, x_{j_{k+1}}) : j_{k+1} = 1, \dots, q_{k+1}(j_k) \} \quad (2.105)$$

is a disjoint collection of subsets in  $(0, x_{j_k} - \delta) \times (\gamma, \gamma^{-1})$ .

Via  $m_1(\{(t')\} \times (\gamma, \gamma^{-1})) = 0$  and (2.91), we can use Proposition 2.5 to get

$$m_1^{(n)}(\cdot \cap (0, t'] \times (\gamma, \gamma^{-1})) = \sum_{j_1=1}^{q_1} \varepsilon_{(t_{j_1}^{(n)}, x_{j_1}^{(n)})} \quad (2.106)$$

and

$$(t_{j_1}^{(n)}, x_{j_1}^{(n)}) \in \bar{B}_\epsilon(t_{j_1}, x_{j_1}), \quad \forall j_1 = 1, \dots, q_1. \quad (2.107)$$

Similarly, using (2.92), (2.104) and that (2.105) is a disjoint collection of subsets in  $(0, x_{j|k} - \delta) \times (\gamma, \gamma^{-1})$ , we can employ Proposition 2.5 to conclude the following: for any  $k = 1, \dots, l-1$  and for any  $j_1 = 1, \dots, q_1; j_2 = 1, \dots, q_2(j_1); \dots; j_k = 1, \dots, q_k(j|_{k-1})$ ,

$$\begin{aligned} & m_{k+1}^{(n)}\left(\cdot \cap \bar{B}_\epsilon(t_{j_1}, x_{j_1}) \times \dots \times \bar{B}_\epsilon(t_{j|k}, x_{j|k}) \times (0, x_{j|k}] \times (\gamma, \gamma^{-1})\right) \\ &= m_{k+1}^{(n)}\left(\cdot \cap \bar{B}_\epsilon(t_{j_1}, x_{j_1}) \times \dots \times \bar{B}_\epsilon(t_{j|k}, x_{j|k}) \times (0, x_{j|k}^{(n)}] \times (\gamma, \gamma^{-1})\right) \\ &= \sum_{j_{k+1}=1}^{q_{k+1}(j|k)} \varepsilon_{(t_{j_1}^{(n)}, x_{j_1}^{(n)}, \dots, t_{j|k+1}^{(n)}, x_{j|k+1}^{(n)})}, \end{aligned} \quad (2.108)$$

and

$$(t_{j|k+1}^{(n)}, x_{j|k+1}^{(n)}) \in \bar{B}_\epsilon(t_{j|k+1}, x_{j|k+1}), \quad \forall j_{k+1} = 1, \dots, q_{k+1}(j|k). \quad (2.109)$$

Let  $f, g \in D([0, t'])$ . Designating by  $d(f, g)$  the  $J_1$  distance on  $D([0, t'])$ , recall that

$$d(f, g) = \inf_{\lambda \in \Lambda} \left\{ \sup_{t \in [0, t']} \|\lambda(t) - t\| \vee \sup_{t \in [0, t']} \|f(\lambda(t)) - g(t)\| \right\}, \quad (2.110)$$

where the  $\Lambda$  is the set of strictly increasing, continuous mappings of  $[0, t']$  onto itself and  $\vee$  stands for maximum.

Let  $\lambda^{(n)} : [0, t'] \rightarrow [0, t']$  be the piecewise linear map that maps  $t_{j_1}$  to  $t_{j_1}^{(n)}$  for all  $j_1 = 1, \dots, q_1$  with  $\lambda^{(n)}(0) = 0$  and  $\lambda^{(n)}(t') = t'$ . Then by (2.106) and (2.107), we have for all  $n$  large enough,

$$\sup_{t \in [0, t']} \|\lambda^{(n)}(t) - t\| \leq \epsilon q_1. \quad (2.111)$$

Moreover, by (2.108) and (2.109), for all  $n$  large enough,

$$\sup_{t \in [0, t']} \|T_l^{(\gamma)}(m^{(n)})(\lambda^{(n)}(t)) - T_l^{(\gamma)}(m)\| \leq \epsilon \sum_{j_1=1}^{q_1} \dots \sum_{j_{l-1}=1}^{q_l(j|_{l-1})} 1. \quad (2.112)$$

Hence, by (2.110), we are finished by the proof of continuity of  $T_l^{(\gamma)}$ .

Next, we prove that  $\bar{T}_l^{(\gamma)}(m^{(n)}) \rightarrow \bar{T}_l^{(\gamma)}(m)$ . We use the notation  $\bar{T}_l^{(\gamma)}(m) = \bar{m}$  with  $\bar{m} = (\bar{m}_1, \dots, \bar{m}_{l-1})$  and similarly,  $\bar{T}_l(m^{(n)}) = \bar{m}^{(n)}$  with  $\bar{m}^{(n)} = (\bar{m}_1^{(n)}, \dots, \bar{m}_{l-1}^{(n)})$ . We first prove that  $\bar{m}_1^{(n)} \rightarrow \bar{m}_1$ . It is enough to show that for any  $0 < s_1 < s_2$  and  $F \subset (0, \infty)$  relatively compact with  $\bar{m}_1(\partial([s_1, s_2] \times F)) = 0$ ,  $\bar{m}_1^{(n)}([s_1, s_2] \times F) \rightarrow \bar{m}_1([s_1, s_2] \times F)$ . Let  $\{(t_{j_1}^{(\gamma)}, x_{j_1}^{(\gamma)}) : j_1 \in \mathbb{N}\}$  denote the marks of  $m_1^{(\gamma)}$  labeled so that  $t_1^{(\gamma)} < t_2^{(\gamma)} < \dots$  and let  $p = \inf\{m : x_1^{(\gamma)} + \dots + x_m^{(\gamma)} > t\} + 1$ . Let  $t_p < t' < t_{p+1}$ . Hence, after relabeling,

$$m_1((0, t'] \times (\gamma, \gamma^{-1})) = \sum_{j_1=1}^p \varepsilon_{(t_{j_1}, x_{j_1})} \quad (2.113)$$



and  $m_1(\partial((0, t'] \times (\gamma, \gamma^{-1}))) = 0$ . We choose  $\delta > 0$  small enough as in (2.104). As before, we choose  $\epsilon < \delta/2$  small enough so that  $\bar{B}_\epsilon(t_1^{(\gamma)}, x_1^{(\gamma)}), \dots, \bar{B}_\epsilon(t_p^{(\gamma)}, x_p^{(\gamma)}) \subset (0, t'] \times (\gamma, \gamma^{-1})$  is a disjoint collection. Using Proposition 2.5 we have for any  $\epsilon > 0$ , after relabeling, for  $n$  large enough,

$$m_1^{(n)}(\cdot \cap (0, t'] \times (\gamma, \gamma^{-1})) = \sum_{j_1=1}^p \varepsilon_{(t_{j_1}^{(n)}, x_{j_1}^{(n)})} \quad (2.114)$$

and

$$(t_{j_1}^{(n)}, x_{j_1}^{(n)}) \in \bar{B}_\epsilon(t_{j_1}, x_{j_1}), \quad \forall j_1 = 1, \dots, p. \quad (2.115)$$

Hence, using Proposition (2.5) we get for all  $j_1 = 1, \dots, p$ ,

$$m_2(\bar{B}_\epsilon(t_{j_1}, x_{j_1}) \times ((0, x_{j_1}^{(\gamma)}] \times (\gamma, \gamma^{-1}))) = m_2^{(n)}(\bar{B}_\epsilon(t_{j_1}^{(n)}, x_{j_1}^{(n)}) \times ((0, x_{j_1}^{(n)}] \times (\gamma, \gamma^{-1}))). \quad (2.116)$$

Finally, choosing  $\epsilon$  small enough so that  $2p\epsilon$  is smaller than the distance of the marks of  $\bar{m}_1$  in  $[s_1, s_2] \times F$  to the boundary of  $[s_1, s_2] \times F$  proves that  $\bar{m}_1^{(n)}([s_1, s_2] \times F) \rightarrow \bar{m}_1([s_1, s_2] \times F)$ . The convergence of  $\bar{m}_k^{(n)}$  to  $\bar{m}_k$  is trivial using the same exact proof and definitions.  $\square$

## 2.4 Proof of the convergence of the clock processes

*Proof of Theorem 1.4.* Let  $\chi^{l^*}$  be a RPC with parameters  $\alpha_1, \dots, \alpha_{l^*}$  and constants  $D_1, \dots, D_{l^*}$ , where the latter collection is given by Lemma 2.2. Note that a.s.  $\chi^{l^*} \in \mathcal{M}_{1, l^*}^{(\gamma)}$ . As before, we define  $\chi^{(\gamma), l^*} = (\chi_1^{(\gamma), l^*}, \dots, \chi_{l^*}^{(\gamma), l^*})$  by

$$\chi_k^{(\gamma), l^*} = \sum_{j|k \in \mathbb{N}^k} \varepsilon_{(t_{j_1}, x_{j_1}, \dots, t_{j_k}, x_{j_k})} \prod_{i=1}^k \mathbb{1}\{\chi_{j|i} \in (\gamma, \gamma^{-1})\}. \quad (2.117)$$

For  $k = 1, \dots, l^*$ , we define the maps  $T_{k, l^*}^{(\gamma)} : \mathcal{M}_{1, l^*} \rightarrow D([0, \infty))$  by

$$T_{k, l^*}^{(\gamma)} = T_{l^*-k+1}^{(\gamma)}(\bar{T}_{l^*-k+2}^{(\gamma)} \circ \dots \circ \bar{T}_{l^*}^{(\gamma)}). \quad (2.118)$$

We also set  $T_{k, l^*}^{(\gamma)}$  as the identity map for any  $k > l^*$ .

Let  $(\Upsilon_k : k = 1, \dots, l^*)$  be an independent collection of PPPs on  $H$  where for each  $k$ , the mean measure of  $\Upsilon_k$  is  $dt \times D_k \alpha_k x^{-1-\alpha_k} dx$ . We denote the marks of  $\Upsilon_k$  by  $\Upsilon_k = \sum_{i \in \mathbb{N}} \varepsilon_{(t_i^{(k)}, \eta_i^{(k)})}$ . For  $\gamma > 0$ , set  $\Upsilon_k^{(\gamma)} = \sum_{i \in \mathbb{N}} \varepsilon_{(t_i^{(k)}, \eta_i^{(k)})} \mathbb{1}\{\eta_i^{(k)} \in (\gamma, \gamma^{-1})\}$ . Due to correlation structure of  $\chi^{l^*}$ , we have  $\bar{T}_2^{(\gamma)} \circ \dots \circ \bar{T}_{l^*}^{(\gamma)}(\chi^{l^*})$  is independent of  $(\chi_1^{l^*}, \dots, \chi_{l^*-1}^{l^*})$  and has the same distribution as  $\Upsilon_{l^*}^{(\gamma)}$ . Using this, the fact that  $(\chi_1^{l^*}, \dots, \chi_{l^*-1}^{l^*}) \stackrel{d}{=} \chi^{l^*-1}$  where  $\chi^{l^*-1}$  has the distribution of a PRC with parameters  $\alpha_1, \dots, \alpha_{l^*-1}$  and constants  $D_1, \dots, D_{l^*-1}$ , and the identity (2.14) we get

$$\begin{aligned} & (T_{k, l^*}^{(\gamma)}(\chi^{l^*}) : k = 1, \dots, l^*) \\ &= (T_{l^*, l^*}^{(\gamma)}(\chi^{l^*}) \circ T_{k, l^*-1}^{(\gamma)}(\chi_1^{l^*}, \dots, \chi_{l^*-1}^{l^*}) : k = 1, \dots, l^*) \\ &= (T_1^{(\gamma)}(\bar{T}_2^{(\gamma)} \circ \dots \circ \bar{T}_{l^*}^{(\gamma)})(\chi^{l^*}) \circ T_{k, l^*-1}^{(\gamma)}(\chi_1^{l^*-1}, \dots, \chi_{l^*-1}^{l^*-1}) : k = 1, \dots, l^*) \\ &\stackrel{d}{=} (T_1^{(\gamma)}(\Upsilon_{l^*}^{(\gamma)}) \circ T_{k, l^*-1}^{(\gamma)}(\chi^{l^*-1}) : k = 1, \dots, l^*), \end{aligned} \quad (2.119)$$

where, in the last display above  $\Upsilon_{l^*}^{(\gamma)}$  and  $\chi^{l^*-1}$  are independent. Proceeding inductively, we reach

$$\left(T_{k,l^*}^{(\gamma)}(\chi^{l^*}) : k = 1, \dots, l^*\right) = \left(T_1^{(\gamma)}(\Upsilon_{l^*}^{(\gamma)}) \circ \dots \circ T_1^{(\gamma)}(\Upsilon_k^{(\gamma)}) : k = 1, \dots, l^*\right). \quad (2.120)$$

Let  $V_{\alpha_k}(t) = \sum_{t_i^{(k)} \leq t} \eta_i^{(k)}$  and  $V_{\alpha_k}^{(\gamma)}(t) = \sum_{t_i^{(k)} \leq t} \eta_i^{(k)} \mathbb{1}\{\eta_i^{(k)} \in (\gamma, \gamma^{-1})\}$ . Note that,  $V_{\alpha_k}$  and  $V_{\alpha_k}^{(\gamma)}$  are Lévy subordinators with corresponding Lévy measures  $D_k \alpha_k x^{-1-\alpha_k} dx$  and  $\mathbb{1}\{x \in (\gamma, \gamma^{-1})\} D_k \alpha_k x^{-1-\alpha_k} dx$ , respectively. We set

$$\tilde{V}_{k,l^*} = V_{\alpha_{l^*}} \circ V_{\alpha_{l^*-1}} \circ \dots \circ V_{\alpha_k}, \quad \text{and} \quad \tilde{V}_{k,l^*}^{(\gamma)} = V_{\alpha_{l^*}}^{(\gamma)} \circ V_{\alpha_{l^*-1}}^{(\gamma)} \circ \dots \circ V_{\alpha_k}^{(\gamma)}.$$

Hence, by (2.120), we get

$$\left(T_{k,l^*}^{(\gamma)}(\zeta^{l^*}) : k = 1, \dots, l^*\right) \stackrel{d}{=} \left(\tilde{V}_{k,l^*}^{(\gamma)} : k = 1, \dots, l^*\right). \quad (2.121)$$

It is trivial that as  $\gamma \rightarrow 0$

$$\left(\tilde{V}_{k,l^*}^{(\gamma)} : k = 1, \dots, l^*\right) \implies \left(\tilde{V}_{k,l^*} : k = 1, \dots, l^*\right) \quad (2.122)$$

weakly on the space  $D^{l^*}([0, \infty))$  equipped with the product Skorohod  $J_1$  topology.

By Lemma 2.4, for each  $k$ ,  $T_{k,l^*}^{(\gamma)}$  is continuous on  $\mathcal{M}_{1,l^*}^{(\gamma)}$ . By Proposition 2.1  $\zeta^{(n),l^*} \Rightarrow \chi^{l^*}$  for every environment in  $\Omega'_L$ , where  $\mathbb{P}(\Omega'_L) = 1$ . Finally, since a.s.  $\chi^{l^*} \in \mathcal{M}_{1,l^*}^{(\gamma)}$  for any  $\gamma > 0$  we can employ the continuous mapping theorem (see Theorem 2.7 in [Bil99]) to conclude that for any  $\gamma > 0$

$$\left(T_{k,l^*}^{(\gamma)}(\zeta^{(n),l^*}) : k = 1, \dots, l^*\right) \implies \left(\tilde{V}_{k,l^*}^{(\gamma)} : k = 1, \dots, l^*\right). \quad (2.123)$$

Now we want to prove that,  $\mathbb{P}$ -a.s. for any  $k = 1, \dots, l^*$  and  $\epsilon > 0$

$$\lim_{\gamma \rightarrow 0} \lim_{n \rightarrow \infty} \mathcal{P}\left(d(S_{k,L}^{(n)}, T_{k,l^*}^{(\gamma)}(\zeta^{(n),l^*})) \geq \epsilon\right) = 0 \quad (2.124)$$

where  $d$  denotes the  $J_1$  distance on  $D([0, \infty))$ . It is enough to check that  $\mathbb{P}$ -a.s. for any  $k = 1, \dots, l^*$  and  $t > 0$ ,

$$\lim_{\gamma \rightarrow 0} \lim_{n \rightarrow \infty} \mathcal{P}(A_{k,l^*}^{(n),(\gamma)}(t) \geq \epsilon) = 0 \quad (2.125)$$

where for  $k_1 = 1, \dots, l^*$  and  $k_2 = k_1, \dots, l^*$ ,

$$A_{k_1,k_2}^{(n),\gamma}(t) = \sum_{t_{j_1}^{(n)} \leq t} \sum_{t_{j_2} \leq \xi_{j_1}^{(n)}} \dots \sum_{t_{j_{k_2}} \leq \xi_{j_{k_2-1}}^{(n)}} \xi_{j_{k_2}}^{(n)} \mathbb{1}\left\{\xi_{j_{k_1}}^{(n)} \notin (\gamma, \gamma^{-1})\right\}.$$

Then, using Lemma 2.2 and the conditioning argument in the proof of Proposition 2.1 we can conclude that  $\exists \bar{\Omega}_L$  with  $\mathbb{P}(\bar{\Omega}_L) = 1$  such that for any environment in  $\bar{\Omega}_L$ , for any  $k = 1, \dots, l^*$  and  $t > 0$ ,

$$A_{k,l^*}^{(n),(\gamma)}(t) \Rightarrow V_{\alpha_{l^*}} \circ \dots \circ V_{\alpha_{k+1}} \circ V_{\alpha_k,(\gamma)} \circ V_{\alpha_{k-1}} \circ \dots \circ V_{\alpha_1}(t),$$

where  $V_{\alpha_1}, \dots, V_{\alpha_{k-1}}, V_{\alpha_{k+1}}, \dots, V_{\alpha_{l^*}}$  are as before and  $V_{\alpha_k,(\gamma)}(t) = \sum_{t_i \leq t} \eta_i^{(k)} \mathbb{1}\{\eta_i^{(k)} \notin (\gamma, \gamma^{-1})\}$ . Thus,

$$A_k^{(n),\gamma}(t) \Rightarrow V_{\alpha_{k+1}\alpha_{k+2}\dots\alpha_{l^*}} \circ V_{\alpha_k,(\gamma)} \circ V_{\alpha_1\dots\alpha_{k-1}}(t), \quad (2.126)$$

where  $V_{\alpha_{k+1}\alpha_{k+2}\dots\alpha_{l^*}}$  and  $V_{\alpha_1\dots\alpha_{k-1}}$  are two independent subordinators that are  $\alpha_{k+1}\alpha_{k+2}\dots\alpha_{l^*}$  and  $\alpha_1\dots\alpha_{k-1}$  stable, respectively, and independent from  $V_{\alpha_k,(\gamma)}$ . For any  $0 < \alpha, \alpha' < 1$ , let  $V_{\alpha,(\gamma)}$  and

$V_{\alpha'}$  be independent subordinators with Lévy measures  $\nu_{(\gamma)}(dx) = \mathbf{1}\{x \notin (\gamma, \gamma^{-1})\} D\alpha x^{-1-\alpha} dx$  and  $\nu'(dx) = D'\alpha' x^{-1-\alpha'} dx$ , respectively. For any  $T > 0$  we have  $E[V_{\alpha,(\gamma)}(T)] \leq CT\gamma^{1-\alpha}$  for some positive constant  $C$ . For any  $\epsilon' > 0$  given, if we choose  $T$  large enough so that  $P(V_{\alpha'}(t) > T) \leq \epsilon'$ , by Chebyshev inequality we get

$$P(V_{\alpha,(\gamma)}(V_{\alpha'})(t) \geq \epsilon) \leq \frac{CT\gamma^{1-\alpha}}{\epsilon} + \epsilon'. \quad (2.127)$$

Moreover, by the stability of  $\alpha$ -stable subordinators we have  $P(V_{\alpha'}(V_{\alpha,(\gamma)})(t) \geq \epsilon) = P(V_{\alpha'}(1)(V_{\alpha,(\gamma)}(t))^{1/\alpha} \geq \epsilon)$ . Hence, if we choose now  $M$  large enough so that  $P(V_{\alpha'}(1) \geq M) \leq \epsilon'$  we get

$$P(V_{\alpha'}(V_{\alpha,(\gamma)})(t) \geq \epsilon) \leq \frac{M^\alpha t \gamma^{1-\delta}}{\epsilon^\alpha} + \epsilon'. \quad (2.128)$$

Hence, combining (2.127) and (2.128) with (2.126) we get (2.125). Thus, using (2.122) and setting  $\tilde{\Omega}_L = \Omega'_L \cap \tilde{\Omega}_L$ , we have  $\mathbb{P}(\tilde{\Omega}_L) = 1$  and for any environment in  $\tilde{\Omega}_L$ , as  $n \rightarrow \infty$

$$\left( S_{k,L}^{(n)} : k = 1, \dots, l^* \right) \Longrightarrow \left( \tilde{V}_{k,l^*} : k = 1, \dots, l^* \right). \quad (2.129)$$

Recall (1.19) and that  $\alpha_{k,L} = \alpha_k$ . Observe that,

$$Z_{R(k-1/L), R(k/L)}(c_k \cdot) \stackrel{d}{=} V_{\alpha_k}(\cdot), \quad c_k = D_k \Gamma(1 - \alpha_k). \quad (2.130)$$

The independence of  $Z_{R(k-1/L), R(k/L)}$  in  $k$  and the stability property finish the proof of Theorem 1.4 with the constants  $b_{l^*,L} = c_{l^*}$  and  $b_{k,L} = c_k c_{k+1}^{\alpha_k} c_{k+2}^{\alpha_k \alpha_{k+1}} \dots c_{l^*}^{\alpha_k \dots \alpha_{l^*} - 1}$ , for  $k = 1, \dots, l^* - 1$ .  $\square$

### 3 Convergence of the two-time correlation function.

*Proof of Theorem 1.1.* Observe that  $\{S_{k,L}(i) : i \in \mathbb{N}\} \cap [t, t+s] = \emptyset$  implies that  $\langle X(t), X(t+u) \rangle \geq k$ ,  $\forall u \in [0, s]$ . On the other hand, if  $\{S_{k,L}(i) : i \in \mathbb{N}\} \cap [t, t+s] \neq \emptyset$  there is at least one jump beyond and including the  $k$ -th level during the time interval  $[t, t+s]$ . Thus, the only way  $\langle X(t), X(t+u) \rangle \geq k$ ,  $\forall u \in [0, s]$  can happen is that at each such jump the discrete Markov chain  $Y|_k$  jumps back to the same vertex it jumped from, and since  $Y|_k$  chain chooses the last coordinate  $\mu_k$  uniform at random on  $[n]$ , we arrive at

$$\Pi_k(t, s) = P\left(\{S_{k,L}(i) : i \in \mathbb{N}\} \cap [t, t+s] = \emptyset\right) + O(1/n). \quad (3.1)$$

By Theorem 1.4, for  $k = 1, \dots, l^*$ , for any environment in  $\tilde{\Omega}_L$ , as  $n \rightarrow \infty$

$$S_{k,L}^{(n)}(\cdot) \Longrightarrow \tilde{Z}_{k,l^*}(\cdot) \quad (3.2)$$

on  $D([0, \infty))$  equipped with  $J_1$  topology, where  $\tilde{Z}_{k,l^*}$  is as in Theorem 1.4. Since  $\tilde{Z}_{k,l^*}$  is an  $\alpha$ -stable subordinator with  $\alpha = \alpha_{k,L} \dots \alpha_{l^*,L}$ , and thus, its Lévy measure has no atoms. Hence, using the continuous mapping theorem and (3.1), we can conclude that for any environment in  $\tilde{\Omega}_L$ , as  $n \rightarrow \infty$

$$\Pi_k^{(n)}(c_n, \theta c_n) \longrightarrow P(\{\tilde{Z}_{k-1,l^*}(r) : r \geq 0\} \cap [1, 1+\theta] = \emptyset) = Asl_{\alpha_{k,L} \dots \alpha_{l^*,L}} \left( \frac{1}{1+\theta} \right). \quad (3.3)$$

We now prove that  $\mathbb{P}$ -a.s.  $\Pi_{l^*+1}(c_n, \theta c_n) \rightarrow 0$  as  $n \rightarrow \infty$ . This convergence implies that  $\Pi_k(c_n, \theta c_n) \rightarrow 0$  for  $k \geq l^* + 1$ , since  $\Pi_{k+1}(c_n, \theta c_n) \leq \Pi_k(c_n, \theta c_n)$ , and finishes the proof. Let  $S_{l^*, l^*} = T_{l^*, l^*}(m) = T_{l^*, l^*}(m|_{l^*})$ . Recalling (2.13), we have

$$S_{l^*+1, L} \circ S_{l^*, l^*} = S_{l^*, L}. \quad (3.4)$$

Observe that,  $S_{l^*, l^*}$  corresponds to the  $l^*$ -th clock process of the GREM-like trap model on  $l^*$ -levels tree where the environment is kept the same on those levels, the only difference is that it is in discrete time where the exponential waiting times are replaced by geometric random variables. Also, note that  $c_n(l^*)$  is a time scale of observation for this GREM-like trap model where all the levels are aging. Therefore, with the same exact proof of Theorem 1.4 we can see that  $\exists \Omega_L''$  with  $\mathbb{P}(\Omega_L'') = 1$  such that for any environment in  $\Omega_L''$ , as  $n \rightarrow \infty$

$$\frac{S_{l^*, l^*}(\cdot a_n(l^*))}{c_n(l^*)} \Longrightarrow V_{\alpha_{l^*, L}}(\cdot) \quad (3.5)$$

weakly on  $D([0, \infty))$  equipped with Skorohod  $J_1$  topology, where  $V_{\alpha_{l^*, L}}$  is an  $\alpha_{l^*, L}$ -stable subordinator. Using Theorem 1.4, for any  $\epsilon > 0$  we can choose  $t$  large enough so that for all  $n$  large,

$$\mathcal{P}\left(\frac{S_{l^*, L}(ta_n(l^*))}{c_n} > 1 + 2\theta\right) \geq 1 - \epsilon/4. \quad (3.6)$$

For  $t$  chosen as above, using (3.5) we can choose  $T > 0$  large enough so that for all  $n$  large,

$$\mathcal{P}\left(\frac{S_{l^*, l^*}(ta_n(l^*))}{c_n(l^*)} \leq T\right) \geq 1 - \epsilon/4. \quad (3.7)$$

Hence, by (3.5) we have for all  $n$  large (recall that  $c_n(l^*) = a_n(l^* + 1)$ ),

$$\mathcal{P}\left(\frac{S_{l^*+1, L}(Ta_n(l^* + 1))}{c_n} \geq 1 + 2\theta\right) \geq 1 - \epsilon/2. \quad (3.8)$$

Recall the collection in (2.17) and how we have constructed it from the chain  $J$ . For  $r = l^* + 1, \dots, L$ , let

$$S_{1, L, (\gamma)}^r(t) = \sum_{j_1=1}^t \sum_{j_2=1}^{\xi_{j_1}} \cdots \sum_{j_{l^*+1}=1}^{\xi_{j_{l^*}}} \Lambda_{(\gamma)}^r(j|_{l^*+1}) \quad (3.9)$$

where

$$\Lambda_{(\gamma)}^r(j|_{l^*+1}) = \sum_{j_{l^*+2}=1}^{\xi_{j_{l^*+1}}} \cdots \sum_{j_L=1}^{\xi_{j_{L-1}}} \xi_{j_L} \mathbb{1}\{\lambda^{-1}(J(j|_r)) \leq \gamma n^{1/\alpha_r}\}. \quad (3.10)$$

Here,  $S_{1, L, (\gamma)}^r$  is the 1st level clock process, where, restricted to jumps where the walk is at a vertex  $\mu$  with  $\lambda^{-1}(\mu|_r) \leq \gamma n^{1/\alpha_r}$ . A modification of the proof of Lemma 2.3 yields that  $\mathbb{P}$ -a.s. for any  $t > 0$ , as  $n \rightarrow \infty$

$$S_{1, L, (\gamma)}^r(ta_n(1))/c_n \Rightarrow S_{(\gamma)}^r(t) \quad (3.11)$$

where  $S_{(\gamma)}^r(t)$  is a positive random variable with the Laplace transform

$$E[e^{-\kappa S_{(\gamma)}^r(t)}] = e^{-\phi_{(\gamma)}(t, \kappa)}, \quad \kappa > 0, \quad (3.12)$$

and for any  $\kappa > 0$  and  $t > 0$ ,  $\phi_{(\gamma)}(t, \kappa) \rightarrow 0$  as  $\gamma \rightarrow 0$ . Hence,  $\mathbb{P}$ -a.s. for any  $r = l^* + 1, \dots, L$ ,  $t > 0$  and  $\epsilon > 0$

$$\lim_{\gamma \rightarrow 0} \limsup_{n \rightarrow \infty} \mathcal{P}(S_{1, L, (\gamma)}^r(ta_n(1)) \geq \epsilon c_n) = 0. \quad (3.13)$$

For  $T' > 0$ , define

$$A_n = \{\exists j_1 = 1, \dots, T' a_n(1), \exists j_2 = 1, \dots, T' a_n(2), \dots, \exists j_{l^*+1} = 1, \dots, T' a_n(l^* + 1) \text{ s.t.} \\ \forall k = 1, \dots, l^* \xi^{(n)}(j|_k) \geq \gamma \text{ and } \Lambda^{(\gamma)}(j|_{l^*+1}) \geq \theta c_n/2\} \quad (3.14)$$

where

$$\Lambda^{(\gamma)}(j|_{l^*+1}) = \sum_{j_{l^*+2}=1}^{\eta_{j|_{l^*+1}}} \cdots \sum_{j_L=1}^{\eta_{j|_{L-1}}} \eta_{j|_L} \prod_{r=l^*+1}^L \mathbb{1}\{\lambda^{-1}(J(j|_r)) > \gamma n^{1/\alpha_r}\}. \quad (3.15)$$

By (2.125), (3.8) and (3.13) we have  $\mathbb{P}$ -a.s. for  $T'$  large enough, for all  $n$  large and  $\gamma$  small  $\Pi_{l^*+1}^{(n)}(tc_n, sc_n) \leq \mathcal{P}(A_n) + \epsilon/2$ . Recall that  $c_n = n^{1/\alpha_L+c}$ , for some  $c > 0$ . We define

$$W^{(n)} = \left\{ \mu|_{l^*} \in V|_{l^*} : \text{if } \mu_{l^*+1}, \dots, \mu_L \text{ s.t. } \forall r = l^* + 1, \dots, L - 1 \\ \lambda^{-1}(\mu|_{l^*} \mu_{l^*+1} \cdots \mu_r) \geq \gamma n^{1/\alpha_r} \text{ then } \max_{\mu_L \in [n]} \lambda^{-1}(\mu|_L) \leq n^{1/\alpha_L+c/2} \right\}, \quad (3.16)$$

and

$$B_n = \{\exists j_1 = 1, \dots, T' a_n(1), \dots, \exists j_{l^*} = 1, \dots, T' a_n(l^*) \text{ s.t.} \\ \forall k = 1, \dots, l^* \xi^{(n)}(j|_k) \geq \gamma \text{ and } J(j|_{l^*}) \notin W^{(n)}\} \quad (3.17)$$

Since the maximum of  $n^d$  i.i.d. mean one exponential random variables is of order  $\log n$ , it follows that  $\mathbb{P}$ -a.s. for all  $n$  large enough  $\mathcal{P}(A_n) \leq \mathcal{P}(B_n) + \epsilon/2$ . Note that

$$\mathbb{P}(\mu|_{l^*} \notin W^{(n)}) \leq C n^{-\alpha_L c/2}, \quad (3.18)$$

for some  $C > 0$ . Since under  $\mathbb{P}$ ,  $\mathbb{1}\{\mu|_{l^*} \notin W^{(n)}\}$  is independent of the random environment on the top  $l^*$  levels we have

$$\mathbb{E}[\mathcal{P}(B_n)] = \left( \prod_{k=1}^{l^*} a_n(k) \mathbb{E}[\mathcal{P}(G(\lambda(\mu|_k) \geq \gamma c_n(k)))] \right) \mathbb{P}(\mu|_{l^*} \notin W^{(n)}). \quad (3.19)$$

By Lemma 2.2 and (3.18) the above quantity converges to 0 as  $n \rightarrow \infty$ . Since  $a_n(1) \ll n$ , we can control the fluctuations as in the proof of Lemma 2.2 to conclude that  $\mathbb{P}$ -a.s. as  $n \rightarrow \infty$ ,  $\mathcal{P}(B_n) \rightarrow 0$ . Hence, denoting by  $\bar{\Omega}_L$  the subset of  $\Omega$  such that for any environment in  $\bar{\Omega}_L$ , as  $n \rightarrow \infty$ ,  $\mathcal{P}(B_n) \rightarrow 0$ , we are finished with the proof of Theorem 1.1 where  $\Omega_L = \tilde{\Omega}_L \cap \bar{\Omega}_L$ .  $\square$

*Proof of Corollary 1.2.* Since  $R$  is uniformly continuous, for any  $\epsilon > 0$  given, we have for all  $L$  large enough

$$0 \leq R(k/L) - R((k-1)/L) \leq \epsilon \quad \forall k = 1, \dots, L. \quad (3.20)$$

Recall the definition of  $\alpha_{k,L}$

$$\alpha_{k,L} = \exp \left\{ - \left( R(k/L) - R(k/L - 1/L) \right) \right\}. \quad (3.21)$$

Using (3.20) for all  $k = 1, \dots, L$ ,

$$0 \leq e^{-\epsilon} (R(k/L) - R((k-1)/L)) \leq 1 - \alpha_{k,L} \leq R(k/L) - R((k-1)/L) \leq \epsilon.$$

Hence, using the definition of  $d_{k,L}$  (see (1.14)), for all  $k = 1, \dots, L$ , we have

$$R(1) - R((k-1)/L) + 1 \leq d_{k,L} \leq e^\epsilon (R(1) - R((k-1)/L)) + 1. \quad (3.22)$$

As a consequence,

$$\lim_{L \rightarrow \infty} \frac{l^*(\rho)}{L} = r^*(\rho) = \sup\{s \geq 0 : R(1) - R(s) + 1 > \rho\}. \quad (3.23)$$

Since  $\rho(0, d_{1,L}) \setminus \{d_{L,L}, \dots, d_{2,L}\}$  for all  $L$  large enough, using Theorem 1.1, we can conclude that for any  $t, s > 0$ , for any environment in  $\Omega' = \cap_L \Omega_L$ , for any  $L$  large enough, (1.18) is satisfied. Via (3.23) and the uniform continuity of  $R$  we get

$$\bar{\alpha}_k = \exp\left(\left(R(l^*(\rho)/L) - R((k-1)/L)\right)\right) = \exp\left(\left(-R(r^*(\rho)) - R(k/L)\right) + c(\epsilon)\right) \quad (3.24)$$

where  $c(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ . Finally, since  $q$  is smooth, using (3.23) one last time, we see that the left hand side of (1.18) is the Riemann sum approximation of the integral on the left hand side of (1.21). Thus, we are finished with the proof.  $\square$

## References

- [BABČ08] G. Ben Arous, A. Bovier, and J. Černý. Universality of the REM for dynamics of mean-field spin glasses. *Comm. Math. Phys.*, 282(3):663–695, 2008.
- [BABG03a] 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.
- [BABG03b] 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.
- [BAČ08] 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.
- [BAG12] G. Ben Arous and O. Gün. Universality and extremal aging for dynamics of spin glasses on subexponential time scales. *Comm. Pure Appl. Math.*, 65(1):77–127, 2012.
- [BCKM97] J.-P. Bouchaud, L.F. Cugliandolo, J. Kurchan, and M. Mezard. Out of equilibrium dynamics in spin-glasses and other glassy systems. *arXiv:9702070*, 1997.
- [BD95] J.-P. Bouchaud and D.S. Dean. Aging on parisi’s tree. *J. Phys. I France*, 5(3):265–286, 1995.
- [Ben62] G. Bennett. Probability inequalities for the sum of independent random variables. *Journal of the American Statistical Association*, 57(297):33–45, 1962.
- [BG13] A. Bovier and V. Gayrard. Convergence of clock processes in random environments and ageing in the  $p$ -spin SK model. *Ann. Probab.*, 41(2):817–847, 2013.
- [BGŠ13] A. Bovier, V. Gayrard, and A. Švejšda. Convergence to extremal processes in random environments and extremal ageing in SK models. *Probab. Theory Related Fields*, 157(1-2):251–283, 2013.
- [BGT87] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1987.

- [Bil99] P Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999.
- [BK07] A. Bovier and I. Kurkova. Much ado about Derrida's GREM. In *Spin glasses*, volume 1900 of *Lecture Notes in Math.*, pages 81–115. Springer, Berlin, 2007.
- [BLG00] J. Bertoin and J.-F. Le Gall. The Bolthausen-Sznitman coalescent and the genealogy of continuous-state branching processes. *Probab. Theory Related Fields*, 117(2):249–266, 2000.
- [Bou92] J.-P. Bouchaud. Weak ergodicity breaking and aging in disordered systems. *J. Phys. I France*, 2(9):1705–1713, 1992.
- [DR78] R. Durrett and S.I. Resnick. Functional limit theorems for dependent variables. *Ann. Probab.*, 6(5):829–846, 1978.
- [Dur10] R. Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.
- [FGG14] L. R. G. Fontes, R. J. Gava, and V. Gayrard. The  $K$ -process on a tree as a scaling limit of the GREM-like trap model. *Ann. Appl. Probab.*, 24(2):857–897, 2014.
- [FP14] L. R. Fontes and G. R. C. Peixoto. GREM-like K processes on trees with infinite depth. [arXiv:1412.4291](https://arxiv.org/abs/1412.4291), December 2014.
- [Gay10a] V. Gayrard. Aging in reversible dynamics of disordered systems. I. Emergence of the arcsine law in Bouchaud's asymmetric trap model on the complete graph. [arXiv:1008.3855](https://arxiv.org/abs/1008.3855), August 2010.
- [Gay10b] V. Gayrard. Aging in reversible dynamics of disordered systems. II. Emergence of the arcsine law in the random hopping time dynamics of the REM. [ArXiv:1008.3849](https://arxiv.org/abs/1008.3849), August 2010.
- [Gay12] V. Gayrard. Convergence of clock process in random environments and aging in Bouchaud's asymmetric trap model on the complete graph. *Electron. J. Probab.*, 17:no. 58, 33, 2012.
- [Gay14] V. Gayrard. Convergence of clock processes and aging in Metropolis dynamics of a truncated REM. [ArXiv:1402.0388](https://arxiv.org/abs/1402.0388), February 2014.
- [MPS<sup>+</sup>84] M. Mézard, G. Parisi, N. Surlas, G. Toulouse, and M. Virasoro. Nature of the spin-glass phase. *Phys. Rev. Lett.*, 52:1156–1159, 1984.
- [Nev92] J. Neveu. A continuous state branching process in relation with the grem model of the spin glass theory. *Rapport interne 267, Ecole Polytechnique Paris*, 1992.
- [Res86] S.I. Resnick. Point processes, regular variation and weak convergence. *Adv. in Appl. Probab.*, 18(1):66–138, 1986.
- [Res08] S.I. Resnick. *Extreme values, regular variation and point processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original.

- [SN00] M. Sasaki and K. Nemoto. Analysis on aging in the generalized random energy model. *Journal of the Physical Society of Japan*, 69(9):3045–3050, 2000.
- [SN01] M. Sasaki and K. Nemoto. Numerical study of aging in the generalized random energy model. *J. Phys. Soc. Jpn*, 70(4):1099–1104, 2001.