Asymptotic Behavior of a Low-Temperature Non-Cascading 2-GREM Dynamics at Extreme Time Scales

LUIZ RENATO FONTES¹, SUSANA FRÓMETA²,† and LEONEL ZUAZNÁBAR¹,*

¹USP - Instituto de Matemática e Estatística.
R. do Matão, 1010 - Butantã, São Paulo - SP, CEP: 05508-090, Brasil.
E-mail: lrfontes@usp.br; lzuaznabar@ime.usp.br
²UFRGS - Instituto de Matemática e Estatística.
Campus do Vale, Av. Bento Gonçalves, Porto Alegre - RS, 9500. CEP: 91509-900, Brasil.
E-mail: susana.frometa@ufrgs.br

We derive the scaling limit for the Hierarchical Random Hopping dynamics for the non cascading 2-GREM at low temperatures and time scales where the dynamics is close to equilibrium. The fine tuning phenomenon plays a role (under certain choices of parameters of the model), yielding three dynamical regimes. In contrast to the cascading case, the pairs of first and second level energies have fluctuations which scale with the volume, and this leads to a family of time scales where we see the dynamics moving through only a part of the full low-temperature energy landscape.

Keywords: GREM; Random Hopping Dynamics; low temperature; fine tuning temperature; scaling limit; extreme time-scale; ergodic time-scale; K process; spin glasses

1. Introduction

This paper is part of an ongoing effort to understand and describe the long term, large volume behavior of mean field spin glass dynamics at low temperature. It is most directly related to [17], where the Hierarchical Random Hopping dynamics (HRHD) for the cascading 2-GREM was analysed. The present paper takes on the non-cascading case of the same dynamics.

We may trace the study of such questions/models in the mathematical literature to [1], [2], where the case of the REM was investigated. The main motivation in many mathematical papers, prompted by the analysis of the phenomenological trap models first appearing in the physics literature (see e.g. [8] and [26]), is to understand the aging phenomenon, which takes place away from equilibrium; see also [4], [3], [9], [23], [5], [23].

Looking at such dynamics when they are close but not quite in equilibrium, in the ergodic time scale has also been a matter interest, either in itself, as a way to understand the transition of the dynamics from an aging behavior to an equilibrium behavior, or yet as an approach to obtain aging results, by first taking a scaling limit at an ergodic time scale, then taking a second limit at a vanishing time scale; see [21], [15], [20], [6], [17].

As in [17], by an ergodic time scale we mean a time scale where, loosely speaking, and as pointed out above, the dynamics is close to but not quite in equilibrium. It is synonymous to the extreme time scale denomination in the present title, and also that of [17]. More precisely, the scaling limit of the dynamics (when both volume and time diverge) under these time scales should be a non trivial, ergodic process. And under any longer (in leading order) time scale, the dynamics should converge to a product of the (infinite volume) equilibrium distribution (over time) — in the present case, we should obtain...
such a longer time scale by multiplying the ergodic time scale by any factor which diverges in the scaling limit. In contrast, an aging time scale is a time scale under which the dynamics exhibits aging in the scaling limit, i.e., certain of its two-time correlation functions are functions of the quotient of the two times; this a phenomenon that takes place far from equilibrium, where such correlations would instead be functions of the time difference, and thus aging time scales should be shorter than ergodic time scales.

The description of scaling limits at ergodic time scales involve K-processes of one kind or another. This is the case for the cascading 2-GREM of [17], and likewise for the present non-cascading case, as will be shown below. K-processes come up also at ergodic time scales of other models, such as trap models on large tori; see [24], [25], [13]. We briefly discuss these processes at the end of this section.

In order to understand the HRHD at ergodic time scales, we need to have a good grip on the structure of the lowest (2-GREM) energies underlying the model, since the dynamics lives on the corresponding spin configurations at that time scale. We have information for the full energy profile from [12], but we need it broken down by hierarchy. Our first result, Theorem 1.2 in the next subsection, gives an asymptotic description of the lowest 2-GREM energies by hierarchy, detailing the result in [12] for the present non-cascading case. It says that in that case, while the full energies are at distances of order 1 apart from each other, as established in [12], their components by hierarchy are much farther apart, in a volume dependent way. This is in stark contrast with the cascading case, where both hierarchical components of the lowest total energies behave in this respect similarly as their total. This makes for a different behavior of the noncascading dynamics at ergodic time scales in two related aspects:

- as in the cascading case, a fine tuning temperature phenomenon takes place in the noncascading dynamics as well (in a certain region of the parameter space), but while in the cascading case the fine tuning range of temperatures is of order $N^{-1}$, in the noncascading case it is of order $N^{-1/2}$, where $N$ represents the volume;
- more strikingly, at and above fine tuning (low) temperatures, the ergodic time scale admits a range of values indexed by a real parameter $L$, which in a certain way selects which lowest energy configurations are visited by the dynamics at a given such time scale.

Differently from [17], we refrain in this paper to discuss aging behavior. One reason is concision, but perhaps the main reason is that we expect the behavior on those time scales to be no different than for the REM; there should be no surprises, in contrast to what we find at ergodic time scales.

We present a detailed discussion of the above points in the following subsections.

The HRHD may be seen as Markov jump process on the hypercube $\{-1, 1\}^N$ in a random environment. We start its description by the environment in the next subsection. A definition of the dynamics, with a discussion of its relevant characteristics, followed by our main (dynamical) results, occupy the remaining subsections of this introduction.

1.1. The environment.

The Generalized Random Energy Model (GREM) was introduced in [14] as a hierarchical spin glass in equilibrium. We consider here the case with two hierarchies (the 2-GREM), which we specify next.

Given a natural number $N$ and $p \in (0, 1)$ let us define $N_1 := \lfloor p N \rfloor$ and $N_2 := N - N_1$. Consider, in $V_N := \{-1, 1\}^N$, a vector $\sigma = \sigma_1 \sigma_2$, where $\sigma_1 \in V_{N_1} := \{-1, 1\}^{N_1}$ and $\sigma_2 \in V_{N_2} := \{-1, 1\}^{N_2}$. For a given $a \in (0, 1)$, and for each $N$, let us define the following Gaussian random variable

$$\Xi_{\sigma} := \sqrt{a} \Xi_{\sigma_1}^{(1)} + \sqrt{1 - a} \Xi_{\sigma_1 \sigma_2}^{(2)},$$

(1.1)
where $\Xi = \{ \Xi^{(1)}_\sigma, \Xi^{(2)}_\sigma : \sigma \in \mathcal{V}_N \}$ is a family of independent standard Gaussian random variables. This represents our random environment. The variables on the family $\Xi$ are correlated in the following way: \( \operatorname{cov}(\Xi^{(1)}_\sigma, \Xi^{(2)}_\tau) = 0 \) if $\sigma_1 \neq \tau_1$ and $\operatorname{cov}(\Xi^{(1)}_\sigma, \Xi^{(2)}_\tau) = a$ if $\sigma_1 = \tau_1$ and $\sigma_2 \neq \tau_2$.

The associated Gibbs measure at inverse temperature $\beta > 0$ is the (random) measure $G_{\beta,N}(\sigma) = e^{\beta \sqrt{N} \Xi^{(1)}_\sigma} / Z_{\beta,N}$ where $Z_{\beta,N}$ is a normalization. The Hamiltonian or energy of $\sigma$ would be, as usual, $H_N(\sigma) = -\sqrt{N} \Xi^{(1)}_\sigma$. We refer to the minima of $H_N(\cdot)$ as low energy or low-lying configurations. At low temperature, the Gibbs measure is essentially concentrated on the low-lying configurations, and at high temperature, the Gibbs measure is concentrated in a growing number of energy levels which get denser as $N \to \infty$.

There are two scenarios that may be distinguished: the cascading case and the non-cascading case. The cascading occurs when $a > p$, and the low-lying energies are achieved by adding up the low-lying energies of the two levels. The non-cascading case occurs when $a \leq p$ and it turns out that the correlations are too weak to have an impact on the extremes and the system “collapses” to a REM, see [12]. Also, in this case, the extremal configurations must differ in the first index, this phenomenon justifies the non-cascading denomination for our system.

As we mentioned before, we will consider the non-cascading 2-GREM evolving under the Random Hopping Dynamics at ergodic time scales. Before introducing the dynamics we will go into the behavior of the environment.

Our main results involve the point process that we define below.

\textbf{Definition 1.1.} Let us call $\mathcal{P} = \{ \xi_i, i \geq 1 \}$ the Poisson point process on $\mathbb{R}$ with intensity measure $Ke^{-x}dx$, with $K > 0$, such that $\xi_i > \xi_{i+1}$ for all $i \geq 1$. We note that this process has a finite maximum.

Concerning the environment, we need to establish a limit result for the first coordinate of the low-lying configurations. We mentioned above that in the non-cascading case the extremal values behaves like in the REM. Indeed, it was proved in Theorem 1.1 of [12] that, in the case $a \leq p$, the point process

\[
\mathcal{P}_N := \sum_{\sigma \in \mathcal{V}_N \times \mathcal{V}_N} \mathbb{1}_{u_N^{-1}(\xi_\sigma)} \to \mathcal{P} \text{ in distribution, as } N \to \infty, \tag{1.2}
\]

where $\mathcal{P}$ is the Poisson point process of Definition 1.1, with $K = 1$ when $a < p$ and $K = \frac{1}{2}$ when $a = p$, and the function $u_N$ is the scaling function for the maximum of $2N$ i.i.d. standard Gaussians, defined as

\[
u_N(x) = \frac{x}{\beta \sqrt{N} + \beta \sqrt{N} + \log N + \kappa} \tag{1.3}
\]

for $\kappa = \log 2 + \log 4\pi$ and $\beta = \sqrt{2 \log 2}$.

The constant $\beta$ above also coincides with the singularity point of the free energy function associated to the Gibbs measure. The low-temperature regime occurs when $\beta > \beta_s$. As we mentioned before, in this temperature region, the Gibbs measure becomes, in the large volume limit, fully concentrated on the set of low-lying configurations, corresponding to the (suitably exponentiated and normalized) points of $\mathcal{P}$ — see (1.5) below.

The convergence in (1.2), however, does not give us any information about the composition of the low-lying configurations when looking at the first and second level separately. The following result states the convergence of the first level low-lying configurations.
Theorem 1.2 (Environment behavior). Let us relabel our indices \( \sigma = (\sigma_1, \sigma_2) \) as \( \sigma(i) = (\sigma_1(i), \sigma_2(i)) \) in order to have \( \Xi_{\sigma(1)} > \Xi_{\sigma(2)} > \cdots > \Xi_{\sigma(2^N)} \). For \( a \leq p \) and \( P = \{\xi_i, i \geq 1\} \) the Poisson point process in Definition 1.1, we have that for every \( k \geq 1 \)

\[
\left( u_N^{-1}(\Xi_{\sigma(1)}), u_N^{-1}(\Xi_{\sigma(1)} - \sqrt{aN\beta_\sigma}); \ldots; u_N^{-1}(\Xi_{\sigma(k)}), u_N^{-1}(\Xi_{\sigma(1)} - \sqrt{aN\beta_\sigma}) \right)
\]

converges in distribution to \((\xi_1, W_1; \ldots; \xi_k, W_k)\), where, for \( a < p \), we have \( K = 1 \) and \( W_1, \ldots, W_k \) are independent Gaussian variables with mean zero and variance \( 1 - a \); which are also independent of the process \( P \), and for \( a = p \), we have \( K = \frac{1}{2} \) and the variables \( W_1, \ldots, W_k \) are now conditioned on being negative.

Remark 1.3. Theorem 1.2 can be written considering instead the second level of the low-lying configurations, obtaining an analogous result in which the centered Gaussian variables in the limit are also independent, but with variance \( a \).

Remark 1.4. It follows from Theorem 1.2 that in the limit \( \{\sigma_1(i) : i \geq 1\} \) are all distinct. In other words, for each \( \sigma_1(i) \), we have a single \( \sigma_2(i) \), and we thus say that \( \sigma_1(i) \) and \( \sigma_2(i) \) are matched. This is in contrast with the cascading phase, where to each \( \Xi_{\sigma_1(i)} \), there corresponds (in the limit, infinitely) many \( \Xi_{\sigma_2(i)} \)’s.

Remark 1.5. A similar issue as the one treated in Theorem 1.2 arises in the analysis of the maximum of two-speed branching Brownian motion, which is a model that is quite similar, and indeed related, to the 2-GREM; see [11]. The proof of Theorem 1.2 of the latter reference goes through an estimation of two components of maxima, similarly as in our proof of our Theorem 1.2, and a range of square root of the leading order of the contributions to the maxima, around such leading order, comes up as well; see (outline of) the proof of Theorem 1.2 of [11]. For the analysis of the HRHD of the 2-GREM we need however precise location of such contributions, beyond the order of magnitude of such range; this is provided by our Theorem 1.2 above, and, as far as we can see, not in [11], or elsewhere.

For later use, let us denote

\[
\gamma^N(\sigma_1\sigma_2) = e^{\frac{\beta}{\sigma_\star}u_N^{-1}(\Xi_{\sigma_1}\sigma_2)} \quad \text{and} \quad \gamma_i = e^{\frac{\beta}{\sigma_\star}\xi_i}.
\]

When \( \sigma_1\sigma_2 = \sigma(i) \), we denote \( \gamma^N_i = \gamma^N(\sigma_1\sigma_2) \).

As a consequence of (1.2) we have

\[
\lim_{N \to \infty} \left\{ \sum_{\sigma_2 \in \mathcal{V}_{2N}} \gamma^N(\sigma_1(i)\sigma_2), i \geq 1 \right\} = \Gamma := \{\gamma_i, i \geq 1\}, \quad \text{in distribution.}
\]

We note that \( \Gamma \) is a Poisson point process in \([0, \infty)\) with intensity \( \frac{\alpha}{2 + \pi} \) \( dx \), where \( \alpha = \frac{\beta_\star}{\beta} \); for \( \beta > \beta_\star \), the sum over the points of \( \Gamma \) is a.s. finite.

\(^1\)This is well defined a.s.
1.2. Dynamics.

We consider the Hierarchical Random Hopping dynamics (HRHD), which is a Markov jump process \{\sigma^N(t), t > 0\} that evolves in \mathcal{V}_N with transition rates given by

\[ w_N(\sigma, \sigma') = \frac{e^{-\beta\sqrt{N}X_i}}{N} \mathbbm{1}_{\sigma \preceq \sigma'} + \frac{e^{-\beta\sqrt{(1-a)N}X_i}}{N} \mathbbm{1}_{\sigma \preceq \sigma'} \]  

(1.6)

and \( w_N(\sigma, \sigma') = 0 \) else, where, for \( i = 1, 2 \), we say that \( \sigma \preceq \sigma' \) iff \( \sigma \sim \sigma' \) and \( \sigma_1 \sim \sigma'_1 \). Here, \( \sigma \sim \sigma' \) indicates that \( \sigma \) and \( \sigma' \) differ in exactly one coordinate. We say, in this context, that \( \sigma \) and \( \sigma' \) are nearest neighbors in \( \mathcal{V}_N \).

The low-temperature 2-GREM in the non-cascading case, as well as in the cascading case, exhibits, depending on the values of \( p \) and \( a \), a dynamical phase transition in the HRHD at ergodic time scales that may be described as a fine tuning phenomenon.

1.2.1. Fine tuning; heuristics

There are two competing factors governing the behavior of the HRHD at ergodic time scales. One factor is \( \#_2 \), the number of jumps until \( \sigma^N \) finds a second level low-lying configuration; this factor is of order \( 2N^2 \). The other factor is \( \#_1 \), the number of visits by \( \sigma^N \) to a first level low-lying configuration \( \sigma_1 \) before leaving it. This is a geometric random variable with mean \( 1 + \frac{N^2}{\beta} e^{\beta/2N^2} \), which, by Theorem 1.2, is of order \( e^{\beta\sqrt{\alpha N} + \beta\sqrt{\alpha N} w} \), where \( W \) is a centered Gaussian random variable with variance \( 1 - a \). This differs with the case \( a > p \), studied in [17], where the first level of a low-lying configuration is exponentially large in \( N \) (namely, of order \( e^{\beta\sqrt{\alpha N}} \)), but with a deterministic inverse square root correction, instead of the present correction, which is random and stretched exponential.

The Gaussian random variable \( W \) that appears in our case will play a crucial role in the behavior of the process \( \sigma^N \) at low temperature at or above fine tuning temperatures. The first effect is in the definition of the ergodic time scale in those regimes, which will incorporate a threshold \( L \) in the exponential root scale. This will effectively select low-lying configurations whose \( W \) values exceed the threshold, and those configurations are the only low-lying ones visited by the dynamics in that time scale. We thus get a family of ergodic time scales, indexed by a real parameter \( L \), the scaling limit of the dynamics under which yields an ergodic infinite volume dynamics supported on the selected low-lying configurations for each \( L \).

The analysis in the paragraph before last leads us to define a family of volume dependent critical parameters

\[ \beta_{FT} := \frac{(1 - p)}{2a} \beta_* - \frac{\theta}{\sqrt{N}}, \quad \theta \in \mathbb{R}, \text{ and } \tilde{\beta}_{FT} := \frac{(1 - p)}{2a} \tilde{\beta}_*, \]  

(1.7)

Let us assume henceforth that \( \tilde{\beta}_{FT} > \beta_* \); we will consider three different temperature regions: when \( \beta \) is larger than \( \tilde{\beta}_{FT} \), equal to \( \tilde{\beta}_{FT} \), and smaller than \( \tilde{\beta}_{FT} \).

(i) At relatively low temperature, when \( \beta > \tilde{\beta}_{FT} \), then \( \#_1 \gg \#_2 \), and while staying at a first level low-lying configuration, the process has time to reach the matching second level low-lying many times until it changes the first component. Here we have a similar situation as the one described in Theorem 2.7 of [17].

(ii) At fine tuning \( \beta = \beta_{FT} \), as pointed out above, the ergodic time scale has a threshold parameter \( L = L_0 \); then, we have \( \#_1 \gg \#_2 \) when \( W > L \), and \( \#_1 \ll \#_2 \) when \( W < L \). Hence for some low-lying configurations the process has time to reach the second level maximum (many times)
during each visit to the first level of such configurations, and there are other ones for which the
first component changes before the process hits the second level ground configurations. So, in
this case the HRHD behaves like for $\beta > \bar{\beta}_{FT}$, except that it does not visit part of the set of
low-lying configurations (those for which $W < L$), but rather spends virtually all of the time in
the complementary set of low-lying configurations.

(iii) At relatively high temperature, when $\beta_* < \beta < \bar{\beta}_{FT}$, we have that $\#_1 \ll \#_2$; then, after
leaving a low-lying configuration, $\sigma^N$ will visit many first level low-lying configurations without finding
the matching second level low-lying configuration. We again have the threshold-$L$ time scale,
and again low-lying configurations with $W < L$ are not visited. The difference here with respect to the
fine tuning regime is that each low-lying configuration that gets visited, is visited many times, in such a way that every time a configuration with a certain $W'$ is left, before it returns, the
ones with values of $W$ larger than $W'$ are visited many times; however, there is a compensating
factor in that the dynamics spends much more time in each visit to configurations with smaller
$W$'s, so that the total time spent on visits to low-lying configurations with $W$'s larger than $W'$
between visits to the low-lying configuration corresponding to $W'$ is comparable to that of
a single visit to the latter configuration. This provides a mixing mechanism for the visits
to the low-lying configurations which results in the infinite volume dynamics being in a sort of
partial equilibrium, or equilibrium in only a part of the configuration space (namely, those
configurations with $W > L$).

More details and precise statements are given below.

**Remark 1.6.** In the case where $\bar{\beta}_{FT} \leq \beta_*$, the fine tuning phenomenon takes place at high/critical
temperature, and since we are looking only at subcritical temperatures, we only have the behavior
described in (i), in that case.

### 1.3. Scaling limit results for the dynamics.

In this section we will state our limit results for the dynamics according to three different conditions
on the temperature parameter $\beta$: when $\beta$ is larger, equal or smaller than the parameter $\beta_{FT}$, defined in
(1.7). We will use the convergence stated in Theorem 1.2, which is weak in the original environment; however, via Skorohod’s Theorem, we are able to make it a strong convergence by going to another,
suitable probability space. From now on, we will assume that we are in the probability space where the
strong convergence takes place, and omit further reference to it.

For each $N$, we relabel the indices $\sigma = (\sigma_1, \sigma_2)$ as $\{\sigma(1), \ldots, \sigma(2^N)\}$ such that $\Xi_{\sigma(1)} > \Xi_{\sigma(2)} > \ldots > \Xi_{\sigma(2^N)}$ (such notation was already used in Theorem 1.2). Let us define the function $\phi^N : \mathcal{V}_N \rightarrow \mathbb{Z}_+$ as

$$\phi(\sigma) = \phi(\sigma_1 \sigma_2) := \min \{i = 1, \ldots, 2^N : \sigma_1(i) = \sigma_1\},$$

and the process $X^N$ as

$$X^N(t) := \phi(\sigma^N(t)), \text{ for all } t \geq 0. \quad (1.8)$$

Our scaling limit results will be stated for the process $X^N$ for the convenience of working with a
state space which naturally extends to the set of natural numbers, which will be the state space of the
limiting processes. Recall the assumption made right below (1.7). We will assume a uniform initial
distribution for $\sigma^N(t)$, for the convenience that the uniform distribution on $\mathcal{V}_N$, is invariant for the
random walk on the hypercube $\mathcal{V}_N$, $i = 1, 2$. It is a simple, if cumbersome matter to extend our results
to most other initial distributions.
Theorem 1.9
are interspersed with an increasing number of visits to other such configurations. The scale, but add up to it when summed over all visits, and successive visits to a given such configuration visit; furthermore, these single time scales are increasingly infinitesimal with respect to the overall time have different increasingly incomparable single time scales for different such configurations in each
converges in distribution as
\( N \rightarrow \infty \) from the increasingly fractal nature of the successive visits of the process to

Remark
The convergence above holds in probability.

Theorem 1.7
\( \frac{1}{N} \int_0^{CN_t} \mathbb{1}_{\{X^N(s)=\ell\}} ds = \pi^L_\ell. \)
The convergence above holds in probability.

Remark 1.8. Theorem 1.7 follows from an analysis based on the heuristics described in (iii) of Sub-subsection 1.2.1. It suggests that the single-time (and also the finite dimensional) distributions of \( \{X^N(cN_t), t \geq 0\} \) converge to (products of) \( \pi^L_\ell \); however, we did not find a way to show that; the main difficulty issues from the increasingly fractal nature of the successive visits of the process to the relevant low-lying configurations as \( N \rightarrow \infty \): as indicated in the above mentioned heuristics, we have different increasingly incomparable single time scales for different such configurations in each visit; furthermore, these single time scales are increasingly infinitesimal with respect to the overall time scale, but add up to it when summed over all visits, and successive visits to a given such configuration are interspersed with an increasing number of visits to other such configurations.

Theorem 1.9 (At fine tuning temperature). For \( \beta = \beta_{FT} \), with \( \theta = \frac{1-P_2}{2a^2} L \) in (1.7), we have that \( \{X^N(cN_t), t \geq 0\} \) converges in distribution as \( N \rightarrow \infty \) to a K-process on \( \mathbb{N}_L = \mathbb{N} \cup \{\infty\} \) with parameter set \( \Gamma_L \), starting from \( \infty \).

We have a longer ergodic time scale for the next result. Let
\[ \bar{c}_N = 2^{-N_2} e^{\beta N - \frac{\log N + \kappa}{2a}}. \]

Theorem 1.10 (Below fine tuning temperatures). For \( \beta > \beta_{FT} \), we have that \( \{X^N(\bar{c}_N t), t \geq 0\} \) converges in distribution as \( N \rightarrow \infty \) to a K-process on \( \mathbb{N} = \mathbb{N} \cup \{\infty\} \) with parameter set \( \Gamma \), starting from \( \infty \).

The convergence of Theorems 1.9 and 1.10 is on Skorohod space with the \( J_1 \) topology. Proofs of both these results are provided in the supplemental material [16].

Remark 1.11. Our choice of representation for the state of the dynamics at each time, involving the first level part of the spin configuration only, as prescribed in (1.8), allows for the convergence in the \( J_1 \)-Skorohod space. A reasonable alternative choice would be to include the second level part of the spin configuration, but then we would lose the \( J_1 \)-convergence in Theorems 1.9 and 1.10 (since in both cases we have increasingly many jumps in and out of the second level part of low-lying configurations, while the second level part rests). Since in the non cascading regime we have, for each low-lying configuration, a single second level part for each first level part, the choice of representation between both possibilities considered is more of a technical nature, pertaining rather to the mode of convergence than to the limiting dynamics.
Remark 1.12. As pointed out at the beginning of this subsection, Theorems 1.7, 1.9 and 1.10 are stated for a special version of the environment, given by Skorohod representation, which converges almost surely as $N \to \infty$. Convergence results for the original environment follow immediately for the integrated dynamics, that is, the distribution of the dynamics, before and after the limit, averaged over the respective environment.

Remark 1.13. A comparison to the corresponding results of [17] is in order. Theorem 1.10 is essentially the same as Theorem 2.7 of [17] if one considers only the first level motion in the latter theorem, which is in a sense the only relevant one (since, as will be argued in detail below, the dynamics spends virtually all the time visiting the low-lying configurations, and, as pointed out above, the second level part of the each low-lying configuration is a function of the first level part of that configuration). For Theorems 1.7 and 1.9, the different structure of the minima of the energies with respect to the cascading case of [17], revealed by Theorem 1.2, is felt, both in the (family of) scales in the ergodic time regime, as well as in the fact that in those time scales we only see part of the energy landscape (and not the full one, as in the corresponding results of [17]). Other than that (but this is of course a major point), Theorems 1.7 and 1.9 describe similar behavior of the non cascading dynamics, as Theorems 2.4 and 2.5 of [17] do for the cascading case, if one looks only at the first level motion of the latter dynamics; we notice however that, while for the cascading case, in the time scale of Theorem 2.4 of [17], the first level is in full equilibrium, this is not the case described by the present Theorem 1.7, which, as pointed out above, could be better characterized as partial equilibrium.

1.4. K-processes

We give a brief description of a process entering two of our main results. By a K-process in this paper we loosely mean a Markov process on an infinite subset $\mathcal{N}$ of $\mathbb{N}$ whose single visits to a given $x \in \mathcal{N}$ lasts an exponential time of mean $\gamma_x$, where $\gamma_x > 0$, $x \in \mathcal{N}$, are parameters of the process, satisfying moreover that $\sum_{x \in \mathcal{N}} \gamma_x < \infty$. The transition from $x$ is uniform in a certain sense, which can be made precise given the latter summability of the parameters. See [21], Section 3 (where $\mathcal{N} = \mathbb{N}$, and there is an extra parameter $c$, which for our purposes here should be taken as 0), for a detailed definition and properties of such a (version of this) process. In order to have regular (i.e., càdlàg) trajectories, we compactify $\mathcal{N}$, and the resulting process lives in $\mathcal{N} \cup \{\infty\}$. The set of times the process spends at $\infty$ is Cantor-like/perfect and with vanishing length. An interpretation of the sites of $\mathcal{N}$ in contrast to $\infty$ in the context of scaling limits of low-temperature spin dynamics is that $x \in \mathcal{N}$ represents singly a general low-lying (selected) configuration, while $\infty$ represents the remaining (higher energy) configurations lumped together.

Variants of this process come up in similar situations involving a range of dynamics which exhibit trapping, as pointed out above, under names such as weighted K-processes (as in [18]; see also [6, 24, 25]); (multi-dimensional) K-processes on a tree (as in [20, 19]), or spatial K-processes (as in [13]). The only version appearing in the present paper, however, is the one discussed in the previous paragraph.

1.5. Organization of the paper.

We devote the next section for the proof of Theorem 1.2, and the following three sections to Theorems 1.7, 1.9, and 1.10, respectively.
2. Proof of Theorem 1.2

We will prove first the case \( a < p \). The case \( a = p \) will follow the lines of the previous case, as we will see later.

Let us consider, for \( \delta > 0 \) and \( j \in \mathbb{Z} \), the intervals

\[
I_N^j := \left[ \sqrt{aN_\beta} + \frac{j}{N^{\frac{1}{2}+\delta}}, \sqrt{aN_\beta} + \frac{j+1}{N^{\frac{1}{2}+\delta}} \right].
\]  

(2.1)

For each \( j \), let us call \( C(I_N^j) := \# \{ \sigma_1 \in V_{N_1} : \Xi^{(1)}_{\sigma_1} \in I_N^j \} \) the cardinality of the set of index \( \sigma_1 \) in \( V_{N_1} \) such that \( \Xi^{(1)}_{\sigma_1} \) belongs to the interval \( I_N^j \). Let us consider

\[
M_N^j := \max_{\sigma \in V_{N_1}, \Xi^{(1)}_{\sigma} \in I_N^j} \Xi_{\sigma}.
\]  

(2.2)

Note that \( \max_{\sigma \in V_{N}} \Xi_{\sigma} = \max_{j \in \mathbb{Z}} M_N^j \). Let us consider

\[
Y_N^j := \sqrt{a} \left( \sqrt{aN_\beta} + \frac{j}{N^{\frac{1}{2}+\delta}} \right) + \sqrt{1-a} \left( \max_{\sigma \in V_{N_1}, \Xi^{(1)}_{\sigma} \in I_N^j} \Xi_{\sigma}^{(2)} \right).
\]  

(2.3)

Later, in Lemma 2.9, we will prove that \( M_N^j \) and \( Y_N^j \) are close. Still, in \( Y_N^j \) we are taking the maximum over a set of random size. In order to overcome this difficulty, we will replace \( C(I_N^j) \) by \( \mathbb{E}[C(I_N^j)] \) and, in Lemma 2.8, we establish a result that enables this replacement.

For each \( j \), let us define the variables

\[
W_N^j := \max_{k=1,\ldots,2N^2\mathbb{E}[C(I_N^j)]} \Xi_{j,k},
\]  

(2.4)

where \( \{ \Xi_N^{j,k} : j \geq 1, k \geq 1 \} \) is a family of i.i.d. standard Gaussian variables, and

\[
Z_N^j := \sqrt{a} \left( \sqrt{aN_\beta} + \frac{j}{N^{\frac{1}{2}+\delta}} \right) + \sqrt{1-a} W_N^j.
\]  

(2.5)

In order to prove Theorem 1.2, we will first prove some convergence results for the variables \( Z_N^j \), as an approximation to \( Y_N^j \). Later we prove that this approximation is good and Theorem 1.2 follows.

Lemma 2.1. Let \( B \) be a nonempty open subset of \( \mathbb{R} \) and \( \epsilon \in (0, \frac{1}{2}) \); then, we have that

\[
\lim_{N \to \infty} \mathbb{P} \left[ \max_{j, \frac{j}{N^{\frac{1}{2}+\delta}} \in B \cap [-N^\epsilon, N^\epsilon]} \mathbb{E}^{-1}(Z_N^j) \leq x \right] = e^{-\Phi_B e^{-x}},
\]

where \( \Phi_B = \int_B \frac{1}{\sqrt{2\pi(1-a)}} e^{-\frac{x^2}{2(1-a)}} dx \).

Remark 2.2. The constant \( \Phi_B \) in Lemma 2.1 hints at the appearance of the Gaussian random variables in the statement of Theorem 1.2.
Remark 2.3. It readily follows from Lemma 2.1 that its statement extends to other $B$’s, like closed subsets of $\mathbb{R}$.

Proof. For $x$ fixed, by independence, we have

$$
P \left[ \max_{j: \frac{j}{N^k + \delta} \in B \cap [-N^k, N^k]} Z_j \leq u_N(x) \right] = \prod_{j: \frac{j}{N^k + \delta} \in B \cap [-N^k, N^k]} P \left[ X \leq y_j^N \right] 2^{N^2 E[C(I^t_N)]},$$

where $y_j^N = y_j^N(x) = \frac{x}{\beta} + (1 - a) \beta N - \frac{\log N + \kappa}{2 \beta} - \frac{\sqrt{a j N}}{N^2}$ and $X$ is a standard Gaussian random variable.

Remark 2.4. Note that $\min_{j: \frac{j}{N^k + \delta} \in B \cap [-N^k, N^k]} y_j^N \to \infty$ as $N \to \infty$.

Applying logarithms we can see that it suffices to show that

$$
\lim_{N \to \infty} \sum_{j: \frac{j}{N^k + \delta} \in B \cap [-N^k, N^k]} 2^{N^2 E[C(I^t_N)]} \log \left( 1 - P \left[ X > y_j^N \right] \right) = -\Phi_B e^{-x}. \tag{2.6}
$$

By Taylor’s Theorem, we have that

$$
\log \left( 1 - P \left[ X > y_j^N \right] \right) = -P \left[ X > y_j^N \right]^2 \frac{1}{2(1 - \theta_j^N)} - \frac{P \left[ X > y_j^N \right]^3}{2}, \text{ where } 0 < \theta_j^N < P \left[ X > y_j^N \right] < 1. \tag{2.7}
$$

Then, since, by Remark 2.4, $\max_{j: \frac{j}{N^k + \delta} \in B \cap [-N^k, N^k]} P \left[ X > y_j^N \right]$ vanishes as $N \to \infty$, it suffices to prove

$$
\lim_{N \to \infty} \sum_{j: \frac{j}{N^k + \delta} \in B \cap [-N^k, N^k]} 2^{N^2 E[C(I^t_N)]} P \left[ X > y_j^N \right] = \Phi_B e^{-x}. \tag{2.8}
$$

A simple computation shows that

$$
2^{N^2 E[C(I^t_N)]} = 2^N P \left[ X \in I^t_N \right] = \frac{2^N}{\sqrt{2\pi}} \int_{I^t_N} e^{-x^2} dx = \frac{2^N \eta_j^N}{\sqrt{2\pi} N^k + \delta} \exp \left\{ - \frac{(\sqrt{a N^k} + \frac{j}{N^k + \delta})^2}{2} \right\}, \tag{2.9}
$$

where $\eta_j^N \to 1$ as $N \to \infty$ for all $|j| \leq N^k + \epsilon$ and $\epsilon < \frac{1}{2}$. Then, in order to prove (2.8), it suffices to prove

$$
\lim_{N \to \infty} \sum_{j: \frac{j}{N^k + \delta} \in B \cap [-N^k, N^k]} 2^N \exp \left\{ - \frac{(\sqrt{a N^k} + \frac{j}{N^k + \delta})^2}{2} \right\} P \left[ X > y_j^N \right] = \Phi_B e^{-x}. \tag{2.10}
$$
We recall the following standard tail estimates on the tail of the standard (unnormalized) Gaussian distribution
\[
\frac{x}{1 + x^2} e^{-\frac{x^2}{2}} \leq \int_x^\infty e^{-\frac{t^2}{2}} dt \leq \frac{1}{x} e^{-\frac{x^2}{2}}, \quad x > 0.
\] (2.11)

Then, again by Remark 2.4, and using (2.11), we have that in order to prove (2.10), it suffices to show that
\[
\lim_{N \to \infty} \sum_{j : \frac{1}{N^{1/3}+\delta} \in B \cap [-N^\epsilon, N^\epsilon]} \frac{2^N}{\sqrt{2\pi N^{1/3}}} \exp \left\{ -\frac{1}{2} \left( \sqrt{\frac{2}{N^{1/3}+\delta}} \right)^2 \frac{(y^j_N)^2}{2} \right\} = \Phi_B e^{-x}. \quad (2.12)
\]

A simple computation shows that the limit in (2.12) is equal to
\[
\lim_{N \to \infty} \frac{2(1-a)N}{\sqrt{2\pi N^{1/3}}} \sum_{j : \frac{1}{N^{1/3}+\delta} \in B \cap [-N^\epsilon, N^\epsilon]} \frac{1}{\sqrt{2\pi N^{1/3}}} \exp \left\{ -\frac{1}{2N^{1/3}+\delta} \right\} = e^{-x}. \quad (2.14)
\]

Then, in order to conclude the proof of (2.8), we only need to show that
\[
\lim_{N \to \infty} \sum_{j : \frac{1}{N^{1/3}+\delta} \in B \cap [-N^\epsilon, N^\epsilon]} \frac{1}{\sqrt{2\pi (1-a)N^{1/3}}} \exp \left\{ -\frac{1}{2(1-a)N^{1/3}} \right\} = \Phi_B. \quad (2.15)
\]

which holds because the left hand side in the limit above is a Riemann sum for the integral that defines \( \Phi_B \).

\[\square\]

**Corollary 2.5.** Let \( B \) be a Borel set of \( \mathbb{R} \) and \( \epsilon \in (0, \frac{1}{2}) \), then we have
\[
\lim_{N \to \infty} \sum_{j : \frac{1}{N^{1/3}+\delta} \in B \cap [-N^\epsilon, N^\epsilon]} P \left[ \frac{1}{u_N} (Z_N^j) > x \right] = \Phi_B e^{-x}. \]

The proof of Corollary 2.5 is provided in the supplemental material [16].

Recall from (2.5) the definition of \( Z_N^j \). For \( B \) as in Lemma 2.1, let us relabel the indices \( j \) belonging to the set \( \{ j : \frac{1}{N^{1/3}+\delta} \in B \cap [-N^\epsilon, N^\epsilon] \} \) as \( j_B(1), j_B(2), \cdots \) in such a way that \( Z_N^{j_B(1)} > Z_N^{j_B(2)} > \cdots \). We will omit \( B \) in the notation \( Z_N^{j_B(i)} \) when \( B = \mathbb{R} \). The next result describes the scaling limit distribution of the vector \( (Z_N^{j_B(1)}, Z_N^{j_B(2)}) \).
Lemma 2.6. Let $B$ be as in Lemma 2.1 and $\epsilon \in (0, \frac{1}{2})$. Then, for $x \geq y$, we have
\[
\lim_{N \to \infty} P \left[ u_N^{-1} (Z_N^{iB(1)}) \leq x, u_N^{-1} (Z_N^{iB(2)}) \leq y \right] = [1 + \Phi_B(e^{-y} - e^{-x})] e^{-\Phi_B e^{-y}}.
\]

Proof.
\[
P \left[ u_N^{-1} (Z_N^{iB(1)}) \leq x, u_N^{-1} (Z_N^{iB(2)}) \leq y \right]
= P \left[ Z_N^{iB(1)} \leq u_N(x), Z_N^{iB(2)} \leq u_N(y) \right]
= P \left[ Z_N^{iB(1)} \leq u_N(y) \right] + P \left[ u_N(y) < Z_N^{iB(1)} \leq u_N(x), Z_N^{iB(2)} \leq u_N(y) \right]
= P \left[ Z_N^{iB(1)} \leq u_N(y) \right] + \sum_{j: \frac{j}{N^{1/2}} + \epsilon \in B \cap [-N^\epsilon, N^\epsilon]} P \left[ u_N(y) < Z_N^{j} \leq u_N(x) \right] \left[ \max_{l \neq j} Z_N^{l} \leq u_N(y) \right].
\]

(2.16)

It readily follows from the arguments used in the proof of Lemma 2.1 that we have that
\[
\lim_{N \to \infty} \frac{P \left[ \max_{l \neq j} Z_N^{l} \leq u_N(y) \right]}{P \left[ \max_{l} Z_N^{l} \leq u_N(y) \right]} = 1 \text{ uniformly in } \left\{ j: \frac{j}{N^{1/2} + \epsilon} \in B \cap [-N^\epsilon, N^\epsilon] \right\}.
\]

Then the limit of the expression in (2.16) is equal to
\[
\lim_{N \to \infty} P \left[ Z_N^{iB(1)} \leq u_N(y) \right] \left[ 1 + \sum_{j: \frac{j}{N^{1/2}} + \epsilon \in B \cap [-N^\epsilon, N^\epsilon]} \left( P \left[ Z_N^{j} \leq u_N(x) \right] - P \left[ Z_N^{j} \leq u_N(y) \right] \right) \right].
\]

The proof is concluded using Lemma 2.1 and Corollary 2.5. \hfill \square

We next establish a convergence result for the variables $Z_N^j$, which constitutes a version of Theorem 1.2 when we replace $Y_N^j$ by $Z_N^j$.

Proposition 2.7. Let $P = \{\xi_i: i \geq 1\}$ be the Poisson point process in Definition 1.1, with $K = 1$. For $k \geq 1$, let $W_1, \ldots, W_k$ be independent standard Gaussian variables with variance $1 - a$, which are also independent of $P$. Then
\[
\left( \frac{j(1)}{N^{1/2} + \delta}, \ldots, \frac{j(k)}{N^{1/2} + \delta}, u_N^{-1} (Z_N^{j(1)}), \ldots, u_N^{-1} (Z_N^{j(k)}) \right)
\]
converges in distribution to $(W_1, \ldots, W_k; \xi_1, \ldots, \xi_k)$ as $N$ goes to infinity.

Proof. It is enough to show that, for $\infty > d_1 > c_1 \geq d_2 > c_2 \geq \cdots \geq d_k > c_k > -\infty$, and $B_1, \ldots, B_k$ disjoint intervals of $\mathbb{R}$, we have
\[
\lim_{N \to \infty} P \left[ \frac{j(1)}{N^{1/2} + \delta} \in B_1, \ldots, \frac{j(k)}{N^{1/2} + \delta} \in B_k; u_N^{-1} \left( Z_N^{j(1)} \right) \in (c_1, d_1), \ldots, u_N^{-1} \left( Z_N^{j(k)} \right) \in (c_k, d_k) \right]
= (\Pi_{i=1}^k \Phi_{B_i}) P \left[ \xi_1 \in (c_1, d_1), \ldots, \xi_k \in (c_k, d_k) \right]
\]

(2.17)
(see Theorem 2.2 from [7], which may be applied to \( \{R^k \setminus \{(x_1, \ldots, x_k) \in R^k : x_i = x_j \text{ for some } i \neq j\}\} \times S_k \), which has full measure under the limit law, where \( S_k = \{x \in R^k : x_1 > x_2 > \cdots > x_k\}\).

For short, let \( \tilde{B}_k = (B_1 \cup \cdots \cup B_k)^c \). Recalling the definition of \( j_B(i) \) above Lemma 2.6, the probability on the left hand side of (2.17) equals

\[
P \left[ \bigcap_{i=1}^k \left\{ u_N^{-1} \left( Z_N^{j_B_k(i)} \right) \in [c_i, d_i] \right\}, \bigcap_{i=1}^{k-1} \left\{ u_N^{-1} \left( Z_N^{j_B_k(i)} \right) > u_N^{-1} \left( Z_N^{j_B_k(2)} \right) \right\}, \right]
\]

\[
\left\{ u_N^{-1} \left( Z_N^{j_B_k(1)} \right) > u_N^{-1} \left( Z_N^{j_B_k(1)} \right) \right\}.
\]

(2.18)

Let \( f_{N,i} \) be the density of the random variable \( u_N^{-1} \left( Z_N^{j_B_k(1)} \right) \) and \( h_{N,i} \) be the joint density of the random vector \( \left( u_N^{-1} \left( Z_N^{j_B_k(1)} \right), u_N^{-1} \left( Z_N^{j_B_k(2)} \right) \right) \). The probability in (2.18) then equals

\[
\int_{c_k}^{d_k} \left[ u_N^{-1} \left( Z_N^{j_B_k(1)} \right) < x \right] \left( \Pi_{i=1}^{k-1} \int_{-\infty}^{x} \int_{c_i}^{d_i} \Phi_{N,i}(s,t) \, ds \, dt \right) f_{N,k}(x) \, dx.
\]

(2.19)

By Lemmas 2.1 and 2.6, we have, \( \lim_{N \to \infty} P \left[ u_N^{-1} \left( Z_N^{j_B_k(1)} \right) < x \right] = e^{-\left(1 - \sum_{i=1}^{k} \chi_{B_i} \right)} e^{-x} \) and

\[
\lim_{N \to \infty} \Pi_{i=1}^{k-1} \int_{-\infty}^{x} \int_{c_i}^{d_i} \Phi_{N,i}(s,t) \, ds \, dt = \Pi_{i=1}^{k-1} \left( \Phi_{B_i} e^{-\Phi_{B_i}} e^{-x} \right) \int_{c_i}^{d_i} e^{-s} \, ds
\]

It promptly follows that the limit in (2.19) equals

\[
\left( \Pi_{i=1}^{k} \Phi_{B_i} \right) \left( \Pi_{i=1}^{k-1} \int_{c_i}^{d_i} e^{-s} \, ds \right) \left( \int_{c_k}^{d_k} e^{-x} e^{-e^{-x}} \, dx \right),
\]

which is readily checked to equal the right hand side of (2.17). \( \Box \)

The next lemma states a deviation result for the number of indices \( \sigma_1 \) such that \( \Xi_{\sigma_1}^{(1)} \in I_N^j \). In Proposition 2.7 we stated the convergence of the variables \( Z_N^{j_k} \) and, with the next result, we will be able to prove the convergence of the variables \( Y_N^{j_k} \).

**Lemma 2.8.** For all \( 0 < \epsilon < \frac{1}{2} \) and \( A > 0 \) we have

\[
\lim_{N \to \infty} \sum_{|j| \leq N^{\frac{1}{2} + \epsilon + \delta}} P \left[ |C(I_N^j) - E[C(I_N^j)]| \geq AE[C(I_N^j)] \right] = 0.
\]

**Proof.** Recall (2.1). Note that

\[
C(I_N^j) = \sum_{\sigma_1 \in \mathcal{V}_{N_1}} \mathbb{1}_{\{\Xi_{\sigma_1}^{(1)} \in I_N^j\}}.
\]

Then, for any integer \( j \in \left[ -N^{\frac{1}{2} + \epsilon + \delta}, N^{\frac{1}{2} + \epsilon + \delta} \right] \), and all large \( N \), we get

\[
E[C(I_N^j)] = 2N^j \int_{I_N^j} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx
\]

(2.20)
Lemma 2.9. close variables are since $j$ which converges to zero as $N$ and hence $E$ and, then Recall from (2.2) and (2.3) the definition of $\sqrt{\alpha N \beta_s + \frac{j+1}{N^{\frac{1}{2}+\delta}}}$ such that $1 \in V$ such that $\sum_{\sigma_1 \in \mathcal{V}_{N_1}} \mathbb{I}_{\{\Xi^{(1)}_{\sigma_1} \in I_N^j\}} + \sum_{\sigma_1 \neq \overline{\sigma}_1} \mathbb{I}_{\{\Xi^{(1)}_{\sigma_1} \in I_N^j\}} \mathbb{I}_{\{\Xi^{(1)}_{\overline{\sigma}_1} \in I_N^j\}}'$ and, then $E[(C(I_N^j))^2] = E[C(I_N^j)] + \left(E[C(I_N^j)]\right)^2 - E[C(I_N^j)]P[\Xi^{(1)}_{\sigma_1} \in I_N^j]$. So we have

$$\frac{E[(C(I_N^j))^2] - \left(E[C(I_N^j)]\right)^2}{\left(E[C(I_N^j)]\right)^2} = \frac{1 - P[X_{\sigma_1} \in I_N^j]}{E[C(I_N^j)]} \leq \frac{1}{E[C(I_N^j)]},$$

and hence

$$\sum_{|j| \leq N^{\frac{1}{2}+\epsilon+\delta}} P\left[|C(I_N^j) - E[C(I_N^j)]| > A E[C(I_N^j)] \right]$$

$$\leq \sum_{|j| \leq N^{\frac{1}{2}+\epsilon+\delta}} \frac{E[(C(I_N^j))^2] - \left(E[C(I_N^j)]\right)^2}{A^2 \left(E[C(I_N^j)]\right)^2}$$

$$\leq \sum_{|j| \leq N^{\frac{1}{2}+\epsilon+\delta}} \frac{1}{A^2 E[C(I_N^j)]} \leq \frac{2N^{\frac{1}{2}+\epsilon+\delta}}{A^2 2(p-\alpha)\sqrt{N}} \exp\{o(N)\},$$

which converges to zero as $N$ goes to infinity, since $p > a$. $\Box$

Recall from (2.2) and (2.3) the definition of $M^j_N$ and $Y^j_N$, respectively. Next we prove that these two variables are close to each other.

**Lemma 2.9.** For all $N > 0$ and all $j \in \mathbb{Z}$, we have

$$0 \leq M^j_N - Y^j_N \leq \frac{\sqrt{\alpha}}{N^{\frac{1}{2}+\delta}}.$$

**Proof.** For $\sigma_1 \in \mathcal{V}_{N_1}$ such that $\Xi^{(1)}_{\sigma_1} \in I_N^j$, we have

$$\sqrt{\alpha} \Xi^{(1)}_{\sigma_1} + \sqrt{1 - a\Xi_{\sigma_1 \sigma_2}^{(2)}} \leq \sqrt{\alpha} \left(\sqrt{\alpha N \beta_s + \frac{j+1}{N^{\frac{1}{2}+\delta}}} \right) + \sqrt{1 - a\Xi_{\sigma_1 \sigma_2}^{(2)}} \leq Y^j_N + \frac{\sqrt{\alpha}}{N^{\frac{1}{2}+\delta}},$$
hence, taking maximum in \( \sigma_1 \in V_N \), such that \( \Xi^{(1)}_{\sigma_1} \in I^j_N \), we have \( M^j_N \leq Y^j_N + \frac{\sqrt{a}}{N^{\frac{1}{2} + \delta}} \). For the other inequality, let us call \( \tilde{\sigma}_1 \) and \( \tilde{\sigma}_2 \) the indices such that
\[
\Xi^{(2)}_{\tilde{\sigma}_1 \tilde{\sigma}_2} = \max_{\sigma_1 \in V_N, \Xi^{(1)}_{\sigma_1} \in I^j_N} \Xi^{(2)}_{\sigma_1 \sigma_2}.
\]
Then \( M^j_N \geq \sqrt{a} \Xi^{(1)}_{\tilde{\sigma}_1} + \sqrt{1 - a} \Xi^{(2)}_{\tilde{\sigma}_1 \tilde{\sigma}_2} \geq \sqrt{a} \left( \sqrt{aN\beta_s} + \frac{j}{N^{\frac{1}{2} + \delta}} \right) + \sqrt{1 - a} \Xi^{(2)}_{\tilde{\sigma}_1 \tilde{\sigma}_2} \). That is \( M^j_N \geq Y^j_N \), which completes the proof. \( \square \)

For our next result we will use a similar notation as the one used in the paragraph above Lemma 2.6. Let us consider \( j(1), j(2), \ldots \) in the set \( \{ j : |j| \leq N^{\frac{1}{2} + \delta + \epsilon} \} \) such that \( M^{j(1)}_N > M^{j(2)}_N > \cdots \). And let us call \( \hat{\sigma}(i) \) the configuration in \( V_N \) such that \( M^{(i)}_N = \Xi_{\hat{\sigma}(i)} \), for each positive integer \( i \).

**Proposition 2.10.** Let \( \mathcal{P} = \{ \xi_i : i \geq 1 \} \) be the Poisson point process in Definition 1.1 and consider \( p > a \). Let \( W_1, \ldots, W_k \) be independent centered Gaussian variables with variance \( 1 - a \), which also are independent of the process \( \mathcal{P} \). Then
\[
\left( u^{-1}_N \left( M^{j(1)}_N \right), \Xi^{(1)}_{\hat{\sigma}_1(1)} - \sqrt{aN\beta_s}, \ldots, u^{-1}_N \left( M^{j(k)}_N \right), \Xi^{(1)}_{\hat{\sigma}_1(k)} - \sqrt{aN\beta_s} \right)
\]
converges in distribution to \((\xi_1, W_1; \cdots; \xi_k, W_k)\) as \( N \) goes to infinity.

**Proof.** Since, for each \( i \), \( \Xi^{(1)}_{\hat{\sigma}(i)} \in I^j_N \), we have that \( 0 \leq \Xi^{(1)}_{\hat{\sigma}(i)} - \left[ \sqrt{aN\beta_s} + \frac{j(i)}{N^{\frac{1}{2} + \delta}} \right] \leq \frac{1}{N^{\frac{1}{2} + \delta}} \). Then, using Lemma 2.9, it is enough to prove that
\[
\left( u^{-1}_N \left( Y^{j(1)}_N \right), \frac{j(1)}{N^{\frac{1}{2} + \delta}}; \cdots; u^{-1}_N \left( Y^{j(k)}_N \right), \frac{j(k)}{N^{\frac{1}{2} + \delta}} \right) \to (\xi_1, W_1; \cdots; \xi_k, W_k), \text{ in distribution.}
\]

As part of the strategy, we compare the distribution of \( Y^j_N \) to that of a perturbation of \( Z^j_N \). For \( \eta > 0 \), let us define \( Z^{j(\eta)}_N \) as
\[
Z^{j(\eta)}_N = \sqrt{a} \left( \frac{j}{N^{\frac{1}{2} + \delta}} + \sqrt{aN\beta_c} \right) + \sqrt{1 - a} \max_{k=1,\ldots,\lfloor 2(1-p)N \rfloor} \Xi_{j,k} \quad (2.24)
\]
Obviously adapting the proof of Proposition 2.7, we find, under the same conditions of the latter result, that
\[
\lim_{N \to \infty} \mathbb{P} \left[ \bigcap_{i=1}^{k} \left\{ u^{-1}_N \left( Z^{j(i)\eta}_N \right) \in [c_i, d_i], \frac{j(i)}{N^{\frac{1}{2} + \delta}} \in B_i \right\} \right]
= \eta^k \left( \Pi_{i=1}^{k} \Phi_{B_i} \right) \left( \Pi_{i=1}^{k} \int_{c_i}^{d_i} e^{-s} ds \right) \left( \int_{c_k}^{d_k} e^{-x} e^{-y} e^{-x+y} dx \right), \quad (2.25)
\]
which may be readily checked to mean that
\[
\left( u^{-1}_N \left( Z^{j(1)\eta}_N \right), \frac{j(1)}{N^{\frac{1}{2} + \delta}}; \cdots; u^{-1}_N \left( Z^{j(k)\eta}_N \right), \frac{j(k)}{N^{\frac{1}{2} + \delta}} \right) \to (\xi^\eta_1, W_1; \cdots; \xi^\eta_k, W_k),
\]
where $\xi^\eta \equiv \xi + \log \eta$.

Then, for $0 < A < 1$, using Lemma 2.8 and (2.25), we get

$$\lim_{N \to \infty} \sup P \left[ \bigcap_{i=1}^k \left\{ u_N^{-1}(Y_N^{j(i)}) \geq x_i, \frac{j(i)}{N^{1/2} + \delta} \in B_i \right\} \right]$$

$$= \lim_{N \to \infty} \sup P \left[ \bigcap_{i=1}^k \left\{ u_N^{-1}(Y_N^{j(i)}) \geq x_i, \frac{j(i)}{N^{1/2} + \delta} \in B_i \right\} \right. \bigcap \left. \bigcap_{i=1}^k \left\{ |C(I_N^{j(i)}) - E[C(I_N^{j(i)})]| < AE[C(I_N^{j(i)})] \right\} \right]$$

$$\leq \lim_{N \to \infty} \sum_{j : |j| \leq N^{1/2 + \delta + \varepsilon}} 1 \left( P \left[ \bigcap_{i=1}^k \left\{ u_N^{-1}(Y_N^{j(i)}) \geq x_i, \frac{j(i)}{N^{1/2} + \delta} \in B_i \right\} \right] \right)$$

$$= (\Pi_{i=1}^k \Phi_{B_i}) \left( \bigcap_{i=1}^k \left\{ \xi_i \geq x_i - \log(1 + A) \right\} \right)$$

Similarly, we get a lower bound for the liminf of the probability on the left hand side of the above expression by exchanging $A$ with $-A$. Since $A$ is arbitrary, the result promptly follows. \hfill \Box

**Proposition 2.11.** For $0 < \epsilon < \frac{1}{2}$ and $\delta > 0$, the point process $\hat{P}_N$, defined as

$$\hat{P}_N := \sum_{j : |j| \leq N^{1/2 + \delta + \epsilon}} 1 \left( u_N^{-1}(M_N^j), \right)$$

converges in distribution to a Poisson point process $P$ with intensity measure $e^{-x} dx$.

The proof of Proposition 2.11 is provided in the supplemental material [16]

We need a final piece to complete the proof of Theorem 1.2. Up to this point, we have been assuming that our indices $j$ are in the interval $[-N^{1/2 + \delta + \epsilon}, N^{1/2 + \delta + \epsilon}]$. Now we would like to remove this restriction, and, for that, let us recall, from the paragraph above Proposition 2.10, that $\hat{\sigma}(i)$ is defined as the configuration that satisfies $M_N^{j(i)} = \Xi_{\hat{\sigma}(i)}$, where the indices $j(i) \in \left\{ j : |j| \leq N^{1/2 + \delta + \epsilon} \right\}$ are such that $M_N^{j(1)} > M_N^{j(2)} > \cdots$. Also, without any restriction in $j$, we say that $\sigma(i)$ is the configuration in which the $i^{th}$ maximum of $M_N^j$ is attained, that is, $M_N^{j(i)} = \Xi_{\sigma(i)}$. The next lemma shows that the mentioned assumption is removable.

**Lemma 2.12.** For any $k \geq 1$ we have

$$\lim_{N \to \infty} P \left[ \Xi_{\sigma(1)} = \Xi_{\hat{\sigma}(1)}, \ldots, \Xi_{\sigma(k)} = \Xi_{\hat{\sigma}(k)} \right] = 1.$$  

**Proof.** The proof follows readily from the fact that, on the one hand, $\hat{P}_N$ is (clearly) dominated by $P_N$, and on the other hand, from Proposition 2.11 and (1.2), both point processes have the same limit. \hfill \Box

**Proof of Theorem 1.2.** Note that the case $a < p$ follows readily from Proposition 2.10 and Lemma 2.12. It remains to prove the case $a = p$, for which all the results for $a < p$ extend with the corresponding changes in statements, and similar proofs. For that reason we will only point out some little modifications in the results above.

In the intervals $J_N^i$, defined in (2.1), let us only consider $j \leq -1$. Then, following the lines of the proof of the case $a < p$, we obtain a similar result as the one in Proposition 2.7, where, the independent and centered Gaussian variables $W_1, \cdots, W_k$ are now conditioned on being negative and $K = \frac{2}{3}$ in the
Poisson point process $\mathcal{P} = \{\xi_i : i \geq 1\}$ of Definition 1.1. Also, considering $-\frac{1}{2} < \delta < 0$ in the intervals $I^j_N$, the corresponding result obtained in Lemma 2.8 also holds.

As before, given (1.2), it will be enough to prove the following analogous result to Proposition 2.11:

$$\sum_{j:N} \mathbb{1}_{L_N^{-1}(I^j_N)} \text{ converges in distribution to a Poisson point process } \mathcal{P} \text{ with intensity measure } \frac{1}{2} e^{-x} dx,$$

which follows from a straightforward adaptation of the previous proof. \qed

3. Proof of Theorem 1.7

We recall that we are resorting to Skorohod’s Representation Theorem to have our environment family of random variables $\Xi$ realized in a space where the convergence stated in Theorem 1.2 (for the original space) holds almost surely; i.e., we have that the convergences in (1.2), (1.5) and

$$\lim_{N \to \infty} \frac{\Xi^{(1)}}{\sigma_1(k)} - \sqrt{a_N} \beta_n = W_k \text{, for all } k \geq 1,$$

(3.1)

hold almost surely; we further assume that we are in the event of that space where those convergences take place everywhere.

We start with some definitions. Consider the following sets: $I := \{i \geq 1 : W_i > L\}$ and $J := \{j \geq 1 : W_j < L\}$, and given $M \geq 1$, the subsets $I_M := \{i_1, \ldots, i_M\} \subset I$ and $J_M := \{j_1, \ldots, j_M\} \subset J$ characterized by:

(a) $W_{i_1} < W_{i_2} < \cdots < W_{i_M}$ and $W_{j_1} < W_{j_2} < \cdots W_{j_M}$,

(b) $\max I_M < \min I \setminus I_M$ and $\max J_M < \min J \setminus J_M$.

Let us denote by $X^N_M$ the restriction of the process $X^N$ to $I_M \cup J_M$.

Proposition 3.1. Given $M \geq 1$ such that $\ell \in I_M$, we have for all $t > 0$ that

$$\lim_{N \to \infty} \frac{1}{c_N t} \int_0^{c_N t} \mathbb{1}_{\{X^N_M(s) = \ell\}} ds = \frac{\gamma(\ell)}{\sum_{m=1}^M \gamma(i_m)},$$

in probability.

Proposition 3.2. Given $t > 0$, let $T^{N,\text{out}}_M(t)$ be the time spent by $X^N$ outside $I_M \cup J_M$ up to time $c_N t$. Then for any $\lambda > 0$ we have

$$\lim_{M \to \infty} \limsup_{N \to \infty} P\left[ \frac{1}{c_N t} T^{N,\text{out}}_M(t) > \lambda \right] = 0.$$
\[ \leq P \left[ \frac{1}{c_N t} \int_0^{c_N t} \mathbb{1}_{\{X^N(s) = \ell\}} ds - \frac{\gamma(\ell)}{\sum_{m=1}^{M} \gamma(i_m)} > \frac{\lambda}{2} \right]. \]

By (3.2), the probability in the right hand side above has the upper bound
\[ P \left[ \frac{1}{c_N t} \int_0^{c_N t} \mathbb{1}_{\{X^N(s) = \ell\}} ds - \frac{\gamma(\ell)}{\sum_{m=1}^{M} \gamma(i_m)} > \frac{\lambda}{4} \right] + P \left[ \frac{1}{c_N t} T_{N,\text{out}}(t) > \frac{\lambda}{4} \right]. \]

Taking \( \limsup \) as \( N \to \infty \) and then \( M \to \infty \), using Proposition 3.1 and 3.2 we obtain Theorem 1.7.

As will become clear in the proof, it is sufficient to show Proposition 3.1 and 3.2 for \( t = 1 \). We will prove Proposition 3.1 in the next section and the proof of Proposition 3.2 is provided in the supplemental material.

### 3.1. Proof of Proposition 3.1

Let us denote by \( \{J_{1,N}(j), j \geq 0\} \) and \( \{J_{2,N}(j), j \geq 0\} \) two independent, discrete time, simple, symmetric random walks evolving in \( \mathcal{V}_{N_1} \) and \( \mathcal{V}_{N_2} \), respectively. For any given probability measure \( \mu_i \) defined in \( \mathcal{V}_{N_i} \), we set
\[ P_{\mu_i}[J_{i,N}(0) = \sigma_i] = \mu_i(\sigma_i). \]

Let us denote by \( \pi_1 \) and \( \pi_2 \) the uniform distribution in \( \mathcal{V}_{N_1} \) and \( \mathcal{V}_{N_2} \), respectively. We will assume that \( J_{1,N} \) starts from \( \pi_1 \) and \( J_{2,N} \) starts from \( \pi_2 \). Let us consider a third probability space \( (\Omega, \mathcal{F}, \tilde{P}) \) in which a family of i.i.d. mean one exponential random variables \( \{T_j, j \geq 0\} \) is defined. We will describe the evolution of \( \sigma^N(t) \) through the product probability \( P = P_{\mu_1} \times P_{\mu_2} \times \tilde{P} \).

For the sake of simplicity we will denote
\[ P_{\mu_1} \times P_{\mu_2} = P_{\mu_1} \times P_{\mu_2} \times \tilde{P}. \]

In some steps of the proof, we will be dealing with situations in which \( J_{1,N} \) starts from certain configuration \( \sigma_1(i_m) \) and \( J_{2,N} \) starts from the uniform distribution \( \pi_2 \). In which case, we adopt the abbreviated notation
\[ P_{m} = P_{\delta_{\sigma_1(i_m)} \times \pi_2}, \quad (3.3) \]
where \( \delta_{\sigma_1} \) denotes the Dirac measure in \( \mathcal{V}_{N_1} \) concentrated in \( \sigma_1 \). Also, we will replace the notation \( P_{\delta_{\sigma_1}} \) by \( P_{\sigma_1} \).

Recalling the transition rate defined in (1.6), the first component of \( \sigma^N \) changes after a geometric number of jumps in the second component, and on each one of those jumps, the amount of time that the process spends has exponential distribution. So, when \( X^N \) arrives at some state \( i \), the time spent until it decides to jump is distributed as a sum of these exponential random variables:
\[ H^{N}(i) = \sum_{j=0}^{G^{N}(i)-1} \sum_{j=0}^{G^{N}(j)-1} N e^{\frac{\beta(1-a)N^{2}}{1-N^{2}}} T_{j}^{N,\text{out}} = \sum_{j=0}^{G^{N}(i)-1} \sum_{j=0}^{G^{N}(j)-1} N e^{\frac{\beta N^{2}}{\mu_{N}(i)+\beta N^{2}}} T_{j}, \quad (3.4) \]
Non-Cascading 2-GREM Dynamics at Extreme Time Scales

where \( G_N(i) = G^N(\sigma_1(i)) \) is a Geometric random variable with mean \( \mu_N(i) \), defined as

\[
\mu_N(i) = 1 + \frac{N_2}{N_1} \beta \sqrt{\alpha N} \xi^{(1)}_{\alpha_1(i)},
\]

(3.5)

Recall the conditions (a) and (b) above Proposition 3.1. In view of the independence and continuity of the distribution of the variables \( W_i, i \geq 1, \sigma_1(i_1), \ldots, \sigma_1(i_M), \sigma_1(j_1), \ldots, \sigma_1(i_M), \)

are well defined and distinct almost surely. Keeping this in mind, let us denote by \( H_N^{m,\ell} \) the time spent by \( X_N \) at site \( i_m \in I_M \) in its \( \ell \)th visit. Note that, for each \( m = 1, \ldots, M \), the variables \( \{ H_N^{m,\ell}; \ell \geq 1 \} \)
have the same distribution as \( H_N^{i_m} \), defined in (3.4); however, they are not independent, indeed they depend on the position of \( J^{2,N} \) in each arrival to \( i_m \), and these positions are not independent.

Given a simple, symmetric random walk \( J^{1,N} \) on \( V_{N_1} \), let us consider the sequence of times when \( J^{1,N} \) visits \( \sigma_1(i_1) \)

\[
\tau^k := \inf\{ n > \tau^{k-1} : J^{1,N}(n) = \sigma_1(i_1) \}, k \geq 1,
\]

(3.6)

with \( \tau^0 = -1 \), and for \( 1 \leq i \leq 2^{N_1} \),

\[
S^k_i = \sum_{n=0}^{\tau^k} \mathbb{1}\{ J^{1,N}(n) = \sigma_1(i) \},
\]

(3.7)

the number of visits of \( J^{1,N} \) to \( \sigma_1(i) \) up to \( \tau^k \).

Let us now consider the sequence of times spent by \( X^N \) on \( i_m \) between returns to \( i_1 \), that is, for \( k \geq 1 \), set

\[
F_N^{m,k} := \sum_{\ell=S^k_{i_m}+1}^{S^k_{i_m}} H_N^{m,\ell}.
\]

(3.8)

Similarly, let \( \hat{H}_N^{m,\ell} \) be the time spent by \( X^N \) at site \( j_m \in J_M \) in its \( \ell \)th visit, and

\[
Q_N^{m,k} := \sum_{\ell=S^k_{j_m}+1}^{S^k_{j_m}} \hat{H}_N^{m,\ell},
\]

(3.9)

is the time spent by \( X^N \) on \( j_m \) between returns to \( i_1 \).

Finally, let us consider the sequence of times that \( X^N \) spends in all sites \( i_m \) and \( j_m \), for \( m = 1, \ldots, M \), between returns to \( i_1 \):

\[
R_N^k := \sum_{m=1}^{M} [F_N^{m,k} + Q_N^{m,k}], k \geq 1.
\]

(3.10)

We note that for every \( m \), the random variables \( F_N^{m,k}, k \geq 2 \), are identically distributed among themselves, and identically distributed to \( F_N^{m,1} \) under \( P_1 \). The same holds for \( Q_N^{m,k}, k \geq 1 \), and \( R_N^k, k \geq 1 \).
For the next lemma, let us define $E_m$ the expectation with respect to the probability $P_m$ defined in (3.3).

**Lemma 3.3.** For $m = 1, \ldots, M$, we have

$$\lim_{N \to \infty} \frac{E_1 \left[ F_{m,1}^N \right]}{E_1 \left[ R_1^N \right]} = \frac{\gamma(i_m)}{\sum_{m=1}^{M} \left( \gamma(i_m) + \gamma(j_m) \right)}.$$

**Proof.** For $\ell \geq 1$, let us define

$$G_\ell = \sum_{n=0}^{\tau(m) - 1} G^N_1(j_1, N(n)), \quad g_\ell = G^N_2(\tau(m))(i_m),$$

where $\tau(m)$ is the number of jumps that $J_1, N$ executes up to its $\ell^{th}$ visit to $\sigma_1(i_m)$ and $\{ G^N_1(i) : n \geq 0, i = 1, \ldots, 2^{N_1} \}$ is a sequence of independent, Geometric random variables with mean $\mu_N(i)$, respectively, defined in (3.5). By (3.4), we can write

$$P_{m,\ell} = \frac{N}{N_1 \mu_N(i_m)} \sum_{j=1}^{g_\ell} e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), j_2, N(\sigma_1(j))} T_{G_\ell + j}}.$$

Conditioning on $J_1, N = \{ J_1, N(n) : n \geq 1 \}$ and $G = \{ G_\ell : \ell \geq 1 \}$, by (3.8), we have that

$$E_1 \left[ F_{m,1}^N \right] = \frac{N}{N_1 \mu_N(i_m)} E_1 \left[ \sum_{\ell=1}^{S_{i_m}^1} \sum_{j=1}^{g_\ell} e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), j_2, N(\sigma_1(j))} T_{G_\ell + j}} \right].$$

Recall that the random walks $J_1, N$ and $J_2, N$ are independent from each other and are also independent from $G$. Then

$$E_1 \left[ e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), j_2, N(\sigma_1(j))} T_{G_\ell + j}} \right] = \sum_{\sigma_2 \in \mathcal{V}_{N_2}} e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), \sigma_2, \sigma_2} \mathbb{P}_{\sigma_2} \left[ J_2, N(\sigma_1(j) + j) = \sigma_2 \right]} = \frac{1}{2^{N_2}} \sum_{\sigma_2 \in \mathcal{V}_{N_2}} e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), \sigma_2}}.$$

It is known from elementary theory of Markov chains (see e.g. (7.17) in [17])

$$E_1 \left[ S_{i_m}^1 \right] = 1.$$

Hence,

$$E_1 \left[ F_{m,1}^N \right] = \frac{N}{N_1 \mu_N(i_m)} \mathbb{E}[g_1] \frac{1}{2^{N_2}} \sum_{\sigma_2 \in \mathcal{V}_{N_2}} e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), \sigma_2}} = \frac{N}{N_1 \mu_N(i_m)} \frac{1}{2^{N_2}} \sum_{\sigma_2 \in \mathcal{V}_{N_2}} e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), \sigma_2}}. \tag{3.14}$$

\footnote{If $\tau_1(m) = 0$, then $G_1 = -1$.}
Analogously, we get
\[ E_1 \left[ Q_{m,1}^N \right] = \frac{N}{N_1} \sum_{\sigma_2 \in \mathbb{V}_{N_2}} e^{\beta \sqrt{N} \Xi_{\sigma_1 (j_m) \sigma_2}}. \] (3.15)

Hence
\[ E_1 \left[ R_1^N \right] = \frac{N}{N_1} \sum_{m=1}^{M} \sum_{\sigma_2 \in \mathbb{V}_{N_2}} e^{\beta \sqrt{N} \Xi_{\sigma_1 (i_m) \sigma_2}} + \sum_{\sigma_2 \in \mathbb{V}_{N_2}} e^{\beta \sqrt{N} \Xi_{\sigma_1 (j_m) \sigma_2}}. \] (3.16)

Finally, using the normalization \( u_{N}^{-1} \), and recalling the definition (1.4), we have
\[ \lim_{N \to \infty} E_1 \left[ F_{m,k}^N \right] = \frac{\sum_{\sigma_2 \in \mathbb{V}_{N_2}} \gamma^N (\sigma_1 (i_m) \sigma_2)}{\sum_{m=1}^{M} \sum_{\sigma_2 \in \mathbb{V}_{N_2}} [\gamma^N (\sigma_1 (i_m) \sigma_2) + \gamma^N (\sigma_1 (j_m) \sigma_2)]} = \frac{\gamma(i_m)}{\sum_{m=1}^{M} [\gamma(i_m) + \gamma(j_m)]}, \] (3.17)

where we have used (1.5) (in the strong form mentioned at the beginning of the section) in the past passage.

Let \( R_0^N \) be the time spent by \( X_M^N \) until its first time out of \( i_1 \). Notice that \( R_k^N \) is the time that \( X_M^N \) spends between the \((k-1)^{st}\) and \(k^{th}\) visit to \( i_1 \), \( k \geq 1 \).

**Lemma 3.4.** Given \( \delta > 0 \) let us define
\[ b_N := \left\lfloor \delta \frac{c_N}{E_1 \left[ R_1^N \right]} \right\rfloor. \] (3.18)

Then, for any fixed \( m = 1, \ldots, M \), we have
\[ \lim_{N \to \infty} \frac{1}{b_N} \sum_{k=1}^{b_N} F_{m,k}^N = \frac{\gamma(i_m)}{\sum_{m=1}^{M} [\gamma(i_m) + \gamma(j_m)]} \] (3.19)

and
\[ \lim_{N \to \infty} \frac{1}{b_N} \sum_{k=1}^{b_N} Q_{m,k}^N = 0. \] (3.20)

Both convergences above hold in probability.

**Proof.** We start with (3.19). Notice that the variables \( \{ F_{m,k}^N : k \geq 1 \} \) have the same distribution, but they are not independent, the dependence coming from the correlations along the trajectory of the random walk \( J_2^N \). In order to control these correlations, we proceed as follows.

We start with the following decomposition:
\[ F_{m,k}^N := F_{m,k}^{1,N} + F_{m,k}^{2,N}, \]
where \( F_{m,k}^{1,N} \) registers the first \( N^3 \) steps of the random walk \( J_{2,N}^{2} \) only, that is,

\[
F_{m,k}^{1,N} = \frac{N}{N_1 \mu_N(i_m)} \sum_{\ell = s_{i_m}+1}^{s_{i_m}^k} \sum_{j=1}^{(K_{i_m}^N - 1) \wedge d_{\ell}} e^{\beta \sqrt{N} \xi \gamma^N_{\sigma_1(i_m), \gamma^N_{\sigma_2}} T_{G_{\ell} + j}},
\]

where \( K_{i_m}^N = N^3 - \Upsilon_\ell \), and \( \Upsilon_0, \Upsilon_1, \ldots \) is a family of iid Bernoulli(\( \frac{1}{2} \)) random variables, independent of everything else.

We allow the \( N^3 \) steps to \( J_{2,N}^{2} \) in order to enable a coupling, after those many steps, to the uniform invariant measure; the \( \Upsilon \)'s comprise another enabler of such a coupling, as it helps break the periodicity of \( J_{2,N}^{2} \) — see paragraph below (3.24). The coupled process does not exhibit the above mentioned correlations.

We next show that \( F_{m,k}^{1,N} \) makes a negligible contribution to the expression whose limit is taken in (3.19). Let us write

\[
E_1 \left[ \frac{1}{b_N} \sum_{k=1}^{b_N} \frac{F_{m,k}^{1,N}}{E_1[R_{1}^{N}]} \right] = \frac{E_1[F_{m,1}^{1,N}]}{E_1[R_{1}^{N}]}.
\]

Reproducing the estimate of the expectation of \( F_{m,1}^{1,N} \) obtained in (3.14), we get

\[
E_1 \left[ F_{m,1}^{1,N} \right] \leq \frac{N}{N_1 2 N_2 \mu_N(i_m)} \sum_{\sigma_2 \in \mathcal{V}_{N_2}} e^{\sqrt{N} \xi \gamma^N_{\sigma_1(i_m), \sigma_2}} N^3.
\]

Therefore, using (3.16) and the normalization \( u_N^{-1} \), we obtain

\[
E_1 \left[ F_{m,1}^{1,N} \right] \leq \frac{N}{N_1 2 N_2 \mu_N(i_m)} \sum_{\sigma_2 \in \mathcal{V}_{N_2}} \gamma^N_{\sigma_1(i_m), \sigma_2} \mu_N(i_m) N^3.
\]

Using (1.5) and recalling, from (3.5), the definition of \( \mu_N(i_m) \), we have that, by (3.1), the last factor above goes to zero as \( N \to \infty \), which proves that

\[
\lim_{N \to \infty} \frac{1}{b_N} \sum_{k=1}^{b_N} \frac{F_{m,k}^{1,N}}{E_1[R_{1}^{N}]} = 0,
\]

in probability. It remains to show that

\[
\lim_{N \to \infty} \frac{1}{b_N} \sum_{k=1}^{b_N} \frac{F_{m,k}^{2,N}}{E_1[R_{1}^{N}]} = \frac{\gamma(i_m)}{\sum_{m=1}^{M} [\gamma(i_m) + \gamma(j_m)]},
\]

in probability.

Let us recall Lemma 3.1 of [10], which says that for a simple symmetric random walk \( \{ J_N^N(k) : k \geq 0 \} \) on \( \mathcal{V}_N \), where \( \pi \) is the uniform distribution on \( \mathcal{V}_N \), \( \theta_N := \frac{3 \ln 2}{2} N^2 \), \( \sigma, \bar{\sigma} \in \mathcal{V}_N \), and \( i \geq 1 \), we have

\[
\left| \sum_{\ell=0}^{1} \pi \bar{J}_N^N(\theta_N + i + \ell) = \bar{\sigma}, J_N^N(0) = \sigma \right| - 2 \pi(\sigma) \pi(\bar{\sigma}) \leq 2^{-3N+1}.
\]
Non-Cascading 2-GREM Dynamics at Extreme Time Scales

In the event that $g_\ell \geq N^3$, we couple $J^{2, N}(G_\ell + K_\ell^N)$ to $U_\ell^N$, where $U_0^N, U_1^N, \ldots$ are iid random variables which are uniformly distributed in $V_{N_2}$. (3.24) and $\beta < \beta_{FT}$ imply that the coupling may be set so that it holds for all $\ell \leq S_{im}^N$, for $b_N$ defined in (3.18), with probability vanishingly close to 1 as $N \to \infty$.

We may thus replace $J^{2, N}(G_\ell + K_\ell^N)$ by $U_\ell^N$ in $F_{m,k}^{2, N}$, thus obtaining an iid family $\tilde{F}_{m,k}^{2, N}$, $1 \leq k \leq b_N$, and it is enough to establish (3.23) with $\tilde{F}_{m,k}^{2, N}$ replacing $F_{m,k}^{2, N}$. The modification of $J^{2, N}$ and $J^N$ produced by these replacements will denoted by $\tilde{J}^{2, N}$ and $\tilde{J}^N$, respectively. We notice that $\tilde{F}_{m,k}^{2, N}$, $k \geq 1$, are iid.

Then (3.23) follows from Chebyshev’s inequality, once we use (3.17, 3.22), and show that

$$
\lim_{N \to \infty} \frac{1}{b_N} \frac{1}{E_1 \left[ (\tilde{F}_{m,1}^{2, N})^2 \right]} = 0.
$$

(3.25)

Since $\tilde{F}_{m,1}^{2, N}$ is stochastically bounded by $F_{m,1}^{N}$, it is enough to show that

$$
\lim_{N \to \infty} \frac{1}{b_N} \frac{1}{E_1 \left[ (F_{m,1}^{N})^2 \right]} = 0 \iff \lim_{N \to \infty} \frac{1}{c_N} \frac{1}{E_1 \left[ (F_{m,1}^{N})^2 \right]} = 0.
$$

Now, observe that

$$
(F_{m,1}^{N})^2 = \left( \sum_{\ell=1}^{S} H_{m,\ell}^N \right)^2 \leq S \sum_{\ell=1}^{S} \left( H_{m,\ell}^N \right)^2, \text{ where } S = S_{im}^1;
$$

so, conditioning on $J^{1, N}$, we get $E_1 \left[ (F_{m,1}^{N})^2 \right] \leq E_{\pi_2} \left[ (H_{m,1}^N)^2 \right] E_{\sigma_1(i_1)} [S^2]$. As a consequence of Corollary 1.5 in [22], to the effect that the steps of $J^{1, N}$ among the configurations $\sigma_1(i), i \in \mathcal{M}$, is approximately uniformly distributed, there exists a positive constant $D$ such that $E_{\sigma_1(i_1)} [S^2] \leq D$ for all $N$; it thus suffices to prove

$$
\lim_{N \to \infty} \frac{1}{c_N} \frac{E_{\pi_2} \left[ (H_{m,1}^N)^2 \right]}{E_1 \left[ R_1^N \right]} = 0.
$$

(3.26)

Let us write

$$
\left( \sum_{j=0}^{g_1-1} e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), j^2, N(j) T_j}} \right)^2 = \sum_{j=0}^{g_1-1} e^{2\beta \sqrt{N} \Xi_{\sigma_1(i_m), j^2, N(j) T_j}^2} + 2 \sum_{0 \leq j < \ell \leq g_1-1} e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), j^2, N(j)}} e^{\beta \sqrt{N} \Xi_{\sigma_1(i_m), j^2, N(\ell) T_j} T_\ell}.
$$

Observe that

$$
E_{\pi_2} \left[ \sum_{j=0}^{g_1-1} e^{2\beta \sqrt{N} \Xi_{\sigma_1(i_m), j^2, N(j) T_j}^2} \right] = \bar{E} \left[ T_2^2 \right] \frac{\mu_N(i_m)}{2N^2 \sqrt{N}} \sum_{\sigma_2 \in V_{N_2}} e^{2\beta \sqrt{N} \Xi_{\sigma_1(i_m), \sigma_2}}.
$$
Then

\begin{align*}
N^2 \quad & \frac{E_{\pi_2} \left[ \sum_{j=0}^{g_i-1} e^{2\beta\sqrt{aN}\xi_{1(i_m),j^2,N(j)}} T_j^2 \right]}{(N_1 \mu_N(i_m))^2} \quad c\, E_1 \left[ R^N \right] \\
&= \tilde{E} \left[ T^2 \right] \frac{N}{N_2} \frac{\sum_{\sigma_2 \in V_{N_2}} \left( \gamma_N(\sigma_1(i_m)\sigma_2) \right)^2}{\sum_{m=1}^{M} \sum_{\sigma_2 \in V_{N_2}} \left[ \gamma_N(\sigma_1(i_m)\sigma_2) + \gamma_N(\sigma_1(i_m)\sigma_2) \right]} \quad e^{-\beta\sqrt{aN} \left( \xi_{1(i_m)} - \beta_+ \sqrt{aN} - L \right)} \quad \frac{1 + N_1 e^{-\beta\sqrt{aN}\xi_{1(i_m)}}}{N_2}.
\end{align*}

By (1.5), (3.1) and the definition of \(i_m\), for which \(W_{i_m} > L\), we get that the expression above vanishes as \(N \to \infty\).

For the estimate of the remaining part of \(E_{\pi_2} \left[ (H_{m_1}^N)^2 \right]\), observe that

\begin{align*}
& E_{\pi_2} \left[ \sum_{0 \leq j < \ell \leq g_1-1} e^{\beta\sqrt{N}\xi_{1(i_m),j^2,N(j)}} e^{\beta\sqrt{N}\xi_{1(i_m),j^2,N(\ell)}} T_j T_\ell \right] \\
&= \sum_{\sigma_2, \sigma_2' \in V_{N_2}} e^{\beta\sqrt{N}\xi_{1(i_m)^2\sigma_2}} e^{\beta\sqrt{N}\xi_{1(i_m)^2\sigma_2'}} E_{\pi_2} \left[ \sum_{0 \leq j < \ell \leq g_1-1} \mathbb{I} \left( J_{j^2,N(j)} = \sigma_2 \right) \mathbb{I} \left( J_{j^2,N(\ell)} = \sigma_2' \right) \right].
\end{align*}

To estimate the second factor in the last line above, we split the summation inside the expectation considering separately values of \(j\) and \(\ell\) whose difference is smaller or bigger than \(N^3\), then using (3.5), (3.24) and some straightforward computations we obtain that there exists a constant \(C > 0\), independent of \(N\), such that, for all \(N\) large, holds

\[ E_{\pi_2} \left[ \sum_{0 \leq j < \ell \leq g_1-1} \mathbb{I} \left( J_{j^2,N(j)} = \sigma_2 \right) \mathbb{I} \left( J_{j^2,N(\ell)} = \sigma_2' \right) \right] \leq C \, N^3 \frac{e^{\beta\sqrt{aN}\xi_{1(i_m)}}}{2N_2}.
\]

Then

\begin{align*}
N^2 \quad & \frac{E_{\pi_2} \left[ \sum_{0 \leq j < \ell \leq g_1-1} e^{\beta\sqrt{N}\xi_{1(i_m),j^2,N(j)}} e^{\beta\sqrt{N}\xi_{1(i_m),j^2,N(\ell)}} T_j T_\ell \right]}{(N_1 \mu_N(i_m))^2} \quad c\, E_1 \left[ R^N \right] \\
&\leq \frac{C \, N}{N_2} \frac{\sum_{\sigma_2 \in V_{N_2}} \left( \gamma_N(\sigma_1(i_m)\sigma_2) \right)^2}{\sum_{m=1}^{M} \sum_{\sigma_2 \in V_{N_2}} \left[ \gamma_N(\sigma_1(i_m)\sigma_2) + \gamma_N(\sigma_1(i_m)\sigma_2) \right]} \frac{N^3}{1 + N_1 e^{-\beta\sqrt{aN}\xi_{1(i_m)}}},
\end{align*}

which goes to zero as \(N \to \infty\), and this establishes (3.25). The proof of (3.19) is complete.

Let us prove now (3.20). Notice that \(Q_{m,k}^N, k \geq 1\), have the same distribution. Now, consider the following indicator random variables:

\[ \tilde{B}^N_k := \begin{cases} 1; & \text{if } \sigma(j_m) \text{ is visited by } J^N \text{ between the } k^{th} \text{ and the } (k + 1)^{th} \text{ return of } J^{1,N} \text{ to } \sigma_1(i_1), \\ 0; & \text{otherwise}, \end{cases} \]

and let us write

\[ Q^N_{m,k} = Q^N_{m,k} \tilde{B}^N_k + Q^N_{m,k}(1 - \tilde{B}^N_k). \]
Proceeding as in the proof of Lemma 3.3, similarly as in (3.14), we find that
\[
E_1 \left[ Q_{m,1}^N (1 - \hat{B}_k^N) \right] \leq \frac{N}{N_1} \frac{1}{2N_2} \sum_{\sigma_2 \in \mathcal{V}_{N_2}, \sigma_2 \neq \sigma_2(j_m)} e^{\beta \sqrt{N} \Xi_{\sigma_1(j_m)\sigma_2}}.
\]
Then, using (3.16) and the normalization \( u_{N_1}^{-1} \), we obtain
\[
E_1 \left[ \frac{1}{b_N} \sum_{k=1}^{b_N} Q_{m,k}^N (1 - \hat{B}_k^N) \right] \leq \frac{\sum_{\sigma_2 \neq \sigma_2(j_m)} \gamma^N (\sigma_1(j_m)\sigma_2)}{\sum_{m=1}^M \sum_{\sigma_2 \in \mathcal{V}_{N_2}} \left[ \gamma^N (\sigma_1(i_m)\sigma_2) + \gamma^N (\sigma_1(j_m)\sigma_2) \right]}.
\]
By (1.5), the right hand side above goes to zero as \( N \to \infty \), which proves that
\[
\frac{1}{b_N} \sum_{k=1}^{b_N} Q_{m,k}^N (1 - \hat{B}_k^N) = 0, \quad \text{in probability. (3.27)}
\]
In order to conclude we will show that
\[
P_1 ( \bigcup_{k=1}^{b_N} \hat{B}_k^N ) \to 0 \text{ as } N \to \infty,
\]
which follows from
\[
b_N P_1 ( \hat{B}_k^N ) \to 0 \text{ as } N \to \infty. \quad (3.29)
\]
For \( m = 1, \ldots, M \) and \( \ell \geq 1 \), let \( \tau_\ell(m) \) and \( \hat{\tau}_\ell \) be as in the paragraph of (3.11), except that we replace \( i_m \) by \( j_m \). When \( J_{1,N}^1 \) is visiting state \( \sigma_1(j_m) \) for the \( \ell \text{th} \) time, \( \ell \geq 1 \), we consider \( \hat{\theta}_\ell \) defined as the number of steps that \( J_{2,N}^2 \) takes to reach \( \sigma_2(j_m) \), if ever during that visit. We then have that
\[
\hat{B}_1^N = \bigcup_{\ell=1}^{\hat{\ell}} \{ \hat{\tau}_\ell > \hat{\theta}_\ell \}, \quad \text{where } \hat{\ell} = S^1_{j_m}.
\]
From the elementary theory of Markov chains, it follows that \( E_1(S^1) = 1 \); see (7.17) in [17]. It is thus enough to show that
\[
b_N P_1 ( \hat{\tau}_1 > \hat{\theta}_1 ) \to 0 \text{ as } N \to \infty. \quad (3.31)
\]
Because of \( \hat{\tau}_1 \) is a Geometric random variable with mean \( 1 + \frac{N_2}{N_1} e^{\beta \sqrt{N} \Xi^{(1)}_{\sigma_1(j_m)}} \), the latter probability is readily seen to equal \( E_1[(1 - q)^{\hat{\theta}_1}] \), where \( q = \left( 1 + \frac{N_2}{N_1} e^{\beta \sqrt{N} \Xi^{(1)}_{\sigma_1(j_m)}} \right)^{-1} \). We may then use Kemperman’s formula to write
\[
E_1[(1 - q)^{\hat{\theta}_1}] = E_{\pi_2}[(1 - q)^{\hat{\theta}_1}] = \frac{1}{B_0(\lambda)} \frac{1}{2N_2} \sum_{i=0}^{N_2} \binom{N_2}{i} B_i(\lambda),
\]
where for \( i = 0, \ldots, N_2, \)
\[
B_i(\lambda) = \int_0^1 (1 - u)^i (1 + u)^{N_2 - i} u^{\lambda - 1} du = \sum_{j=0}^{N_2 - i} \binom{N - i}{j} \frac{\Gamma(i + 1) \Gamma(\lambda + j)}{\Gamma(\lambda + i + j + 1)},
\]
Indeed, using that

\[ \frac{b_N}{B_0(\lambda)} \int_0^1 u^{\lambda-1} du \sum_{i=0}^{N_2} \binom{N_2}{i} (1-u)^i(1+u)^{N_2-i} \]

Proof. Follows immediately from Lemma 3.4.

and we readily check that the third quotient is of order \(1\) (since the numerator is the right scale for the denominator, as follows from (1.2) and (1.3)),

The first quotient above is clearly of order \(1\), and the second quotient may be also checked to be of order \(1\) (since the numerator is the right scale for the denominator, as follows from (1.2) and (1.3)),

and that the sum in the denominator on the right hand side above is

\[ \sim \sum_{i=1}^{N_2} \binom{N_2}{i} \frac{1}{i} \sim \frac{2^{N_2+1}}{N_2}. \]

From the above and (3.18), we find that the left hand side of (3.31) equals

\[ \frac{N_1}{2N} \sum_{m=1}^{M} c_{V_{N_2}} e^{\beta(N-\log N+\alpha)} \sum_{\sigma_2 \in V_{N_2}} e^{\beta\sqrt{aN\Xi(1)(j_m)}} \sum_{\sigma_2 \in V_{N_2}} e^{\beta\sqrt{aN\Xi(1)(j_m)}} \]

\[ \times \frac{e^{-\beta aN\sqrt{aN\lambda}}}{2^{N_2} + N_1 e^{-\beta aN\sqrt{aN\lambda}(W_{j_m} + o_1)}}. \]

The first quotient above is clearly of order \(1\), and the second quotient may be also checked to be of order \(1\) (since the numerator is the right scale for the denominator, as follows from (1.2) and (1.3)),

and we readily check that the third quotient is of order \(e^{-\beta aN(L-W_{j_m} + o_1)}\) (here we have used that \(\beta < \beta_{FT}\)), and this vanishes as \(N \to \infty\) since by definition \(W_{j_m} < L\). The result follows.

\[ \square \]

**Corollary 3.5.** Given \(\delta > 0\), consider \(b_N := \left\lfloor \frac{\delta c_N}{E_1[R_1^N]} \right\rfloor\). Then for all \(m = 1, \ldots, M\) we have

\[ \lim_{N \to \infty} \frac{1}{b_N} \sum_{k=1}^{b_N} \frac{R_k^N}{E_1[R_1^N]} = \frac{\sum_{m=1}^{M} \gamma(i_m)}{\sum_{m=1}^{M} \gamma(i_m) + \gamma(j_m)}, \] in probability.

**Proof.** Follows immediately from Lemma 3.4. \(\square\)

**Lemma 3.6.** Let us define \(L^N(t) := \max \left\{ n \geq 0 : \sum_{k=0}^{n} R_k^N \leq t \right\}\), for \(t \geq 0\). Then

\[ \lim_{N \to \infty} \frac{E_1[R_1^N] L^N(c_N)}{c_N} = \frac{\sum_{m=1}^{M} \gamma(i_m) + \gamma(j_m)}{\sum_{m=1}^{M} \gamma(i_m)}, \] in probability.

**Proof.** Follows from Corollary 3.5. However, since the summation in the definition of \(L^N\) starts at \(k = 0\), we only need to prove that the first summand, when divided by \(c_N\), goes to zero in probability. Indeed, using that \(E_{\sigma_1} \left[ \sum_{k=0}^{t-1} \mathbb{1}_{\{j^*(N(k) = \sigma_1)\}} \right] \leq 2\) for any \(\sigma_1 \in V_n\) and \(t^1\) as defined in (3.6), (see
Lemma 7.4 in [17]), and reproducing the estimate (3.16), we get

$$E_{\pi_1 \times \pi_2}[R_0^N] \leq 2^{N_1} N_1 2^{N_2} \sum_{m=1}^M \left[ \sum_{\sigma_2 \in \mathcal{V}_{N_2}} e^{\beta_N N_2 \sigma_1(m) \sigma_2} + \sum_{\sigma_2 \in \mathcal{V}_{N_2}} e^{\beta_N N_2 \sigma_1(m) \sigma_2} \right].$$

Since $\beta < \beta_{FT}$ we obtain the desired convergence. 

**Lemma 3.7.** For $m = 1, \ldots, M$ and $L^N$ as defined in Lemma 3.6, we have

$$\lim_{N \to \infty} \frac{1}{L^N(c_N)} \sum_{k=0}^{L^N(c_N)} \frac{F^N_{m,k}}{E_1[F^N_{m,k}]} = 1, \text{ in probability.}$$

**Proof.** Follows readily from Lemmas 3.3 and 3.6. 

**Proof of Proposition 3.1.** Note that

$$\sum_{k=0}^{L^N(c_N)} F^N_{\ell,k} \leq \int_0^{c_N} \mathbb{1}_{\{X_M(s) = \ell\}} ds \leq \sum_{k=0}^{L^N(c_N)+1} F^N_{\ell,k}. \quad (3.37)$$

Writing

$$\frac{1}{c_N} \sum_{k=0}^{L^N(c_N)} F^N_{\ell,k} = \left( \frac{1}{L^N(c_N)} \sum_{k=0}^{L^N(c_N)} \frac{F^N_{\ell,k}}{E_1[F^N_{\ell,k}]} \right) \frac{E_1[R_N^1]}{E_1[R_N^1]}$$

and using Lemma 3.3, Lemma 3.6 and Lemma 3.7 we obtain the desired convergence. 

**Acknowledgements**

LRF was partially supported by CNPq grants 311257/2014-3 and 307884/2019-8; and FAPESP grant 2017/10555-0. SF was partially supported by FAPESP grant 2017/10555-0 during visits to USP to carry out this research. LZ was supported by a PNPD/CAPES grant 88882315481/2013-01 post doctoral fellowship. We thank an anonymous referee for many specific comments to an earlier version of this text that much helped improve the presentation.

**References**

[1] BÉN AROUS, G., BOVIER, A. and GAYRARD, V. (2003). Glauber dynamics of the random energy model. I. Metastable motion on the extreme states. *Comm. Math. Phys.* 235 379–425. MR1974509

[2] BÉN AROUS, G., BOVIER, A. and GAYRARD, V. (2003). Glauber dynamics of the random energy model. II. Aging below the critical temperature. *Comm. Math. Phys.* 236 1–54. MR1977880

[3] BÉN AROUS, G., BOVIER, A. and ČERNÝ, J. (2008). Universality of the REM for dynamics of mean-field spin glasses. *Comm. Math. Phys.* 282 663–695. MR2426140
[4] Ben Arous, G., Dembo, A. and Guionnet, A. (2001). Aging of spherical spin glasses. *Probab. Theory Related Fields* **120** 1–67. MR1856194

[5] Ben Arous, G. and Gün, O. (2012). Universality and extremal aging for dynamics of spin glasses on subexponential time scales. *Comm. Pure Appl. Math.* **65** 77–127. MR2846638

[6] Bezerra, S. C., Fontes, L. R. G., Gava, R. J., Gayrard, V. and Mathieu, P. (2012). Scaling limits and aging for asymmetric trap models on the complete graph and $K$ processes. *ALEA Lat. Am. J. Probab. Math. Stat.* **9** 303–321. MR3069367

[7] Billingsley, P. (2013). *Convergence of probability measures*. John Wiley & Sons.

[8] Bouchaud, J. P. and Dean, D. S. (1995). Aging on Parisi’s tree. *J. Phys. I France* **5** 265–286.

[9] Bovier, A. and Gayrard, V. (2013). Convergence of clock processes in random environments and aging in the $p$-spin SK model. *Ann. Probab.* **41** 817–847. MR3077527

[10] Bovier, A. and Kurkova, I. (2004). Derrida’s Generalised Random Energy models 1: models with finitely many hierarchies. *Markov Process. Related Fields* **10** 439–480.

[11] Cortines, A., Gold, J. and Louidor, O. (2018). Dynamical freezing in a spin glass system with logarithmic correlations. *Electron. J. Probab.* **23** Paper No. 59, 31. MR3814255

[12] Derrida, B. (1985). A generalization of the random energy model which includes correlations between energies. *Journal de Physique Lettres* **46** 401–407.

[13] Fontes, L. R. G. and Lima, P. H. S. (2009). Convergence of Symmetric Trap Models in the Hypercube. In *New Trends in Mathematical Physics* (V. Sidoravičius, ed.) 285-297. Springer, Dordrecht.

[14] Fontes, L. R. G., Gava, R. J. and Gayrard, V. (2019). Asymptotic behavior and aging of a low temperature cascading 2-GREM dynamics at extreme time scales. *Ann. Probab.* **47** 176–223. MR3814255

[15] Fontes, L. R. G. and Mattei, P. (2008). K-processes, scaling limit and aging for the trap model in the complete graph. *Ann. Probab.* **36** 1322–1358. MR2435851

[16] Gayrard, V. and Arous, G. B. (2008). Elementary potential theory on the hypercube. *Electronic Journal of Probability* **13** 1726–1807.

[17] Gayrard, V. and Gün, O. (2016). Aging in the GREM-like trap model. *Markov Process. Related Fields* **22** 165–202. MR3523983

[18] Jara, M., Landim, C. and Teixeira, A. (2011). Quenched scaling limits of trap models. *Ann. Probab.* **39** 176–223. MR2778800

[19] Jara, M., Landim, C. and Teixeira, A. (2014). Universality of trap models in the ergodic time scale. *Ann. Probab.* **42** 2497–2557. MR3265173

[20] Sasaki, M. and Nemoto, K. (2000). Analysis on Aging in the Generalized Random Energy Model. *J. Phys. Soc. Jpn.* **69** 3045–3050.