# CONCENTRATION OF THE CLOCK PROCESS NORMALISATION FOR THE METROPOLIS DYNAMICS OF THE REM

JIŘÍ ČERNÝ

Abstract. In [ČW17], it was shown that the clock process associated with the Metropolis dynamics of the Random Energy Model converges to an $\alpha$-stable process, after being scaled by a random, Hamiltonian dependent, normalisation. We prove here that this random normalisation can be replaced by a deterministic one.

## 1. Introduction

Recently, in [ČW17], it was shown that the out-of-equilibrium Metropolis dynamics of the Random Energy Model (REM) in a broad range of time scales falls into the universality class of Bouchaud's trap model [Bou92], at least at the level of the scaling limit of the so-called clock process. Later, in [Gay18], this result was extended to a usual aging statement, in terms of two-time observables, using different set of techniques. This concluded, to a certain extent, the long line of studies of aging in the REM, started in [BBG03a, BBG03b] (we refer to [ČW17, Gay18] for in-depth bibliographies).

The scaling limit results of [ČW17] and [Gay18] have one slightly infuriating (at least for the author of this paper) feature: the scaling functions used to normalise the clock process depend on the Hamiltonian of the REM and are therefore random (cf. Theorem 1.1 in [ČW17], and Proposition 1.5 with the subsequent remarks in [Gay18]).

There are several heuristic arguments why to believe that this apparent necessity to choose random scaling functions is actually just a shortcoming of the techniques used in [ČW17, Gay18]. Some of these arguments will be given later in this paper, others appear in Remark 4 under Theorem 1.1 of [ČW17]. In this remark, we conjectured that the scaling function may be chosen deterministic. The main aim of this paper is to prove this conjecture.

## 2. Setting and result

We work in the setting of [ČW17] which we recall now. We consider the standard REM whose state space is the $N$-dimensional hypercube $\mathbb{H}_N = \{-1, 1\}^N$, and whose Hamiltonian is a collection $(E_x)_{x\in\mathbb{H}_N}$ of i.i.d. standard Gaussian random

*Date*: February 14, 2019.

variables defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The equilibrium distribution of this model at the inverse temperature $\beta > 0$ is given by the (non-normalized) Gibbs measure $\tau_x = \mathrm{e}^{\beta\sqrt{N}E_x}$.

The Metropolis dynamics of the REM is a continuous-time Markov chain $X = (X_t)_{t\geq 0}$ on $\mathbb{H}_N$ with transition rates

$$r_{xy} = \mathrm{e}^{-\beta\sqrt{N}(E_x-E_y)_+}\mathbf{1}_{\{x\sim y\}} = \left(1 \wedge \frac{\tau_y}{\tau_x}\right)\mathbf{1}_{\{x\sim y\}}, \qquad x, y \in \mathbb{H}_N. \tag{2.1}$$

Here, $x \sim y$ means that $x$ and $y$ are neighbours on $\mathbb{H}_N$, that is they differ in exactly one coordinate. In order to understand the behaviour of $X$, [ČW17] introduces its 'accelerated' version $Y = (Y_t)_{t\geq 0}$ which is a continuous-time Markov chain with transition rates

$$q_{xy} = \frac{\tau_x \wedge \tau_y}{1 \wedge \tau_x}\mathbf{1}_{\{x\sim y\}}, \qquad x, y \in \mathbb{H}_N. \tag{2.2}$$

It can easily be checked that $Y$ is reversible, with the equilibrium distribution

$$\nu_x = \frac{1 \wedge \tau_x}{Z_N}, \qquad x \in \mathbb{H}_N, \tag{2.3}$$

where $Z_N = \sum_{x\in\mathbb{H}_N}(1\wedge\tau_x)$ is a $\tau$-dependent normalisation constant. Finally, since $r_{xy} = (1 \vee \tau_x)^{-1}q_{xy}$, $X$ can be written as a time change of $Y$,

$$X(t) = Y(S^{-1}(t)), \tag{2.4}$$

where $S^{-1}$ is the generalised right-continuous inverse of the *clock process* $S$,

$$S(t) = \int_0^t (1 \vee \tau_{Y_s})\,\mathrm{d}s. \tag{2.5}$$

Given the environment $\tau = (\tau_x)_{x\in\mathbb{H}_N}$, we use $P^\tau_\nu$ and $P^\tau_x$ to denote the laws of the process $Y$ started from its stationary distribution $\nu$ or from $x \in \mathbb{H}_N$, respectively, and write $E^\tau_\nu$, $E^\tau_x$ for the corresponding expectations. $D([0, T], \mathbb{R})$ stands for the space of $\mathbb{R}$-valued càdlàg functions on $[0, T]$.

The following theorem is the main result of [ČW17]:

**Theorem 2.1** ([ČW17],Theorem 1.1)**.** *Let $\alpha \in (0, 1)$ and $\beta > 0$ be such that*

$$\frac{1}{2} < \frac{\alpha^2\beta^2}{2\ln 2} < 1, \tag{2.6}$$

*and define*

$$g_N = \mathrm{e}^{\alpha\beta^2 N}\left(\alpha\beta\sqrt{2\pi N}\right)^{-\frac{1}{\alpha}}. \tag{2.7}$$

*Then there are random variables $R_N$ which depend on the Hamiltonian $(E_x)_{x\in\mathbb{H}_N}$ only, such that for every $T > 0$ the rescaled clock processes*

$$S_N(t) = g_N^{-1}S(tR_N), \qquad t \in [0, T], \tag{2.8}$$

*converge in $\mathbb{P}$-probability as $N \to \infty$, in $P_\nu^\tau$-distribution on the space $D([0,T],\mathbb{R})$ equipped with the Skorokhod $M_1$-topology, to an $\alpha$-stable subordinator $V_\alpha$. The random variables $R_N$ satisfy*

$$\lim_{N\to\infty} \frac{\ln R_N}{N} = \frac{\alpha^2\beta^2}{2}, \quad \mathbb{P}\text{-a.s.} \tag{2.9}$$

In fact, [ČW17] not only states the existence of the random normalisation scale $R_N$, but provides an explicit formula for it, (cf. (2.10) there): For $\alpha \in (0,1)$, $\beta > 0$ as in Theorem 2.1, fix $\gamma'$ such that

$$\frac{1}{2} < \gamma' < \gamma := \frac{\alpha^2\beta^2}{2\ln 2} < 1, \tag{2.10}$$

and define the set of *deep traps*

$$\mathcal{D}_N = \{x \in \mathbb{H}_N : \ \tau_x \geq g'_N\}, \tag{2.11}$$

where the scale $g'_N$ is chosen so that

$$\mathbb{P}[x \in \mathcal{D}_N] = 2^{-\gamma' N}(1 + o(1)). \tag{2.12}$$

Let $H_x = \inf\{t \geq 0 : Y_t = x\}$ be the hitting time of $x \in \mathbb{H}_N$ by $Y$, and let $\ell_t(x) = \int_0^t \mathbf{1}_{\{Y_s = x\}} \mathrm{d}s$ be the local time of $Y$ at time $t \geq 0$ and position $x \in \mathbb{H}_N$. Finally, let $T_{\text{mix}}$ be a certain randomized stopping time at which $Y$ is "well mixed". Its exact definition is slightly complicated (cf. [ČW17, Proposition 3.3]), but it is irrelevant here. Then $R_N$ is defined by

$$R_N = 2^{N(\gamma-\gamma')}\Bigg(\sum_{x\in\mathcal{D}_N} \frac{E_x^\tau[\ell_{T_{\text{mix}}}(x)^\alpha]}{E_\nu^\tau[H_x]}\Bigg)^{-1}. \tag{2.13}$$

As mentioned in the introduction, the fact that the normalisation scale $R_N$ in (2.8) is random is rather displeasing, even if (2.9) proves that at least its exponential growth is deterministic. We now improve (2.9) and show the behaviour conjectured in [ČW17].

**Theorem 2.2.** *For every $\alpha, \beta$ as in (2.6), there exists a sequence $h_N$ independent of the choice of $\gamma'$ in (2.10), satisfying $\lim_{N\to\infty} N^{-1}\ln h_N = 0$, such that*

$$\lim_{N\to\infty} h_N^{-1} \mathrm{e}^{-\alpha^2\beta^2 N/2} R_N = 1, \qquad \mathbb{P}\text{-a.s.} \tag{2.14}$$

*In particular, the main claim of Theorem 2.1 holds true also when the definition (2.8) of the rescaled clock process $S_N$ is replaced by $S_N(t) = g_N^{-1} S(h_N \mathrm{e}^{\alpha^2\beta^2 N/2} t)$.*

*Remark* 2.3. (a) While we decided to work in the setting of [ČW17], we are confident that similar techniques can be applied in order to show that the random normalisation $b_n$ defined in (1.41)–(1.43) of [Gay18] has a deterministic asymptotic behaviour, as well.

(b) The proof of Theorem 2.2 does not use the assumption $\gamma = \frac{\alpha^2\beta^2}{2\ln 2} > \frac{1}{2}$ from (2.6). This assumption was taken in [ČW17] to make certain arguments simpler.

As shown in [Gay18], (a variant of) Theorem 2.1 holds for every $\gamma, \alpha \in (0,1)$. Hence, our arguments should provide a concentration of the random normalisation in the whole aging regime.

(c) While the main result of this paper is very model specific, the technique that we develop here is rather general and can, e.g., be used to show that quenched expected hitting time of "sparse" random sets by certain processes in random environment concentrates around its annealed average. Obtaining such a technique was another motivation for writing this paper.

We close this section by a heuristic explanation why it should be expected that the quantity $R_N$ exhibits a law of large numbers (2.14), as we promised in the introduction. Remark first that the points of $\mathcal{D}_N$ are typically well separated when $\gamma' > 1/2$, in fact their typical minimal distance is of order $N$, cf. [ČW17, Lemma 2.1]. Assume now that it is possible to put around every point $x \in \mathcal{D}_N$ (or at least around most of them) a set $A_x \ni x$, so that $A_x$ is not connected to $A_{x'}$ for all $x \neq x' \in \mathcal{D}_N$ (ideally, $A_x$ would be a ball $B(x, \rho_N)$ around $x$ with a radius $1 \ll \rho_N \ll N$), and have the property that "when started out of $A_x$, the process $Y$ mixes well before hitting $x$", that is, slightly more formally,

$$P^\tau_y[T_{\mathrm{mix}} \le H_x] \ge 1 - o(1), \quad \text{for all } y \notin A_x. \tag{2.15}$$

If such sets exist, then, viewing the hypercube as an electrical network with conductances $c_{xy} = Z_N^{-1}(\tau_x \wedge \tau_y)\mathbf{1}_{\{x \sim y\}}$ (cf. [ČW17, (2.4)]), it is relatively standard to relate the fraction in (2.13) to the effective conductance $\mathcal{C}(x, A_x^c)$ from $x$ to the complement of $A_x$. Indeed, if (2.15) holds, then, under $P^\tau_x$, $\ell_{T_{\mathrm{mix}}}(x)$ can be approximated by $\ell_{T_{A_x}}(x)$, where $T_{A_x}$ denotes the exit time from $A_x$. It is a known fact that, under $P^\tau_x$, $\ell_{T_{A_x}}(x)$ has exponential distribution whose mean can easily be calculated and equals $Z_N^{-1}\mathcal{C}(x, A_x^c)^{-1}$ for every $x \in \mathcal{D}_N$, see the proof of Corollary 4.3 in [ČW17]. Hence, $E^\tau_x[\ell_{T_{\mathrm{mix}}}(x)^\alpha]$ is approximately equal to $c_{N,\alpha}\mathcal{C}(x, A_x^c)^{-\alpha}$. On the other hand, using e.g. arguments as in [ČTW11, Proposition 3.2], if (2.15) holds, then $E^\tau_\nu[H_x]$ can be approximated by $c'_N\mathcal{C}(x, A_x^c)^{-1}$. Hence, assuming (2.15), the sum in the definition (2.13) of $R_N$ approximately equals

$$c_{N,\alpha} \sum_{x \in \mathcal{D}_N} \mathcal{C}(x, A_x^c)^{1-\alpha}. \tag{2.16}$$

Recalling that $A_x$ are mutually disconnected, and thus the effective conductances $\mathcal{C}(x, A_x^c)$, $x \in \mathcal{D}_N$, independent (or even i.i.d. depending on the construction of $A_x$), (2.14) then should follow by invoking a suitable law of large numbers for triangular arrays.

The problem with the reasoning above is that it seems very difficult to find the sets $A_x$ such that (2.15) holds, due to some "singular" behaviour of $Y$. Therefore, in this paper we resort to a second moment computation and estimate the variance of $R_N^{-1}$ using the classical Efron-Stein inequality.

A key ingredient in the application of this inequality is the observation of [Gay18] (cf. Proposition 3.8 there) that relations like (2.15), which are hard to prove uniformly for all $x \in \mathcal{D}_N$ and $y \in A_x^c$, typically hold on average, cf. Lemma 4.3 below.

Finally, let us introduce an additional notation. For any $A \subset \mathbb{H}_N$, we write $H_A = \inf\{t \geq 0 : Y_t \in A\}$ for its hitting time by $Y$. We use $\lambda_Y$ to denote the spectral gap of $Y$. Since $\lambda_Y$ depends on the random environment $\tau$, we write $\lambda_Y^\tau$ when we want to point out this dependence. The same holds true for $\mathcal{D}_N = \mathcal{D}_N^\tau$, $Z_N = Z_N^\tau$, etc. We use $c, C, \dots$ to denote generic finite positive constants whose value might change from line to line; they may depend on $\alpha, \beta$ but not on $N$. For a function $f : \mathbb{N} \to (0, \infty)$ and $a \in \mathbb{R}$ we often write $f(N) \leq 2^{aN(1+o(1))}$ to abbreviate $\limsup_{N\to\infty} N^{-1} \ln f_N \leq a \ln 2$. If $f$ depends on additional parameters, this is meant to be uniform in them.

## 3. Preliminaries

This section contains several preparatory steps which will later allow to construct another random scaling function $F_N$ providing a very good approximation of $R_N$, and whose variance will be easier to estimate.

We start by replacing the slightly unpleasant randomized stopping time $T_{\mathrm{mix}}$ appearing in the definition (2.13) of $R_N$ by a deterministic time horizon $\mu_N$,

$$\mu_N = N^2 \mathrm{e}^{\beta\sqrt{N}}. \tag{3.1}$$

**Lemma 3.1.** *$\mathbb{P}$-a.s. for all $N$ large enough, for all $x \in \mathbb{H}_N$,*

$$E_x^\tau\big[\big|\ell^\alpha_{T_{\mathrm{mix}}}(x) - \ell^\alpha_{\mu_N}(x)\big|\big] \leq 2^{-N(1+o(1))} E_x^\tau[\ell^\alpha_{\mu_N}(x)] + \mathrm{e}^{-N^2}. \tag{3.2}$$

*Proof.* We decompose the expectation appearing in the lemma as

$$\begin{aligned} &E_x^\tau\big[\big|\ell^\alpha_{T_{\mathrm{mix}}}(x) - \ell^\alpha_{\mu_N}(x)\big|\big] \\ &= E_x^\tau\big[\big(\ell^\alpha_{T_{\mathrm{mix}}}(x) - \ell^\alpha_{\mu_N}(x)\big)\mathbf{1}_{\{T_{\mathrm{mix}} \geq \mu_N\}}\big] + E_x^\tau\big[\big(\ell^\alpha_{\mu_N}(x) - \ell^\alpha_{T_{\mathrm{mix}}}(x)\big)\mathbf{1}_{\{T_{\mathrm{mix}} < \mu_N\}}\big]. \end{aligned} \tag{3.3}$$

By the construction of the mixing time $T_{\mathrm{mix}}$ in [ČW17, Proposition 3.3], $\mathbb{P}$-a.s. for all $N$ large enough, under $P_x^\tau$, $T_{\mathrm{mix}}/m_N$ has the geometrical distribution with parameter $1 - \mathrm{e}^{-1}$, where $m_N = N^{c(\beta)}$ with $c(\beta) > 0$. Using the Cauchy-Schwarz inequality and the fact that $\ell_t(x) \leq t$, the first summand on the right-hand side of (3.3) satisfies

$$\begin{aligned} &E_x^\tau\big[\big(\ell^\alpha_{T_{\mathrm{mix}}}(x) - \ell^\alpha_{\mu_N}(x)\big)\mathbf{1}_{\{T_{\mathrm{mix}} \geq \mu_N\}}\big] \\ &\quad\leq E_x^\tau\big[\big(\ell^\alpha_{T_{\mathrm{mix}}}(x) - \ell^\alpha_{\mu_N}(x)\big)^2\big]^{1/2} P_x^\tau[T_{\mathrm{mix}} \geq \mu_N]^{1/2} \\ &\quad\leq c(m_N^\alpha + \mu_N^\alpha)\mathrm{e}^{-\mu_N/2m_N} \leq \mathrm{e}^{-N^2}, \end{aligned} \tag{3.4}$$

for all $N$ large enough, since $\mu_N \gg 2N^2 m_N$.

For the second summand in (3.3), we observe that on $T_{\mathrm{mix}} < \mu_N$, since $\alpha < 1$, $\ell^\alpha_{\mu_N}(x) - \ell^\alpha_{T_{\mathrm{mix}}}(x) \leq (\ell_{\mu_N}(x) - \ell_{T_{\mathrm{mix}}}(x))^\alpha$. In addition, by [ČW17, Proposition 3.3] again, $Y_{T_{\mathrm{mix}}}$ is $\nu$-distributed and independent of $T_{\mathrm{mix}}$. Therefore, using twice the strong Markov property, once with $T_{\mathrm{mix}}$ and once with $H_x$,

$$\begin{aligned} E_x^\tau\big[\big(\ell^\alpha_{\mu_N}(x) - \ell^\alpha_{T_{\mathrm{mix}}}(x)\big)\mathbf{1}_{\{T_{\mathrm{mix}}<\mu_N\}}\big] &\le E_x^\tau\big[\big(\ell_{\mu_N}(x) - \ell_{T_{\mathrm{mix}}}(x)\big)^\alpha\mathbf{1}_{\{T_{\mathrm{mix}}<\mu_N\}}\big] \\ &\le E_\nu^\tau[\ell^\alpha_{\mu_N}] \\ &\le P_\nu^\tau[H_x \le \mu_N] E_x^\tau[\ell_{\mu_N}(x)^\alpha]. \end{aligned} \tag{3.5}$$

By [AB92, Theorem 1], under $P_\nu^\tau$, the hitting time $H_x$ is approximately exponentially distributed in the sense that

$$\Big|P_\nu^\tau[H_x > t] - \mathrm{e}^{-\frac{t}{E_\nu^\tau[H_x]}}\Big| \le \frac{1}{\lambda_Y E_\nu^\tau[H_x]}. \tag{3.6}$$

It follows that the right-hand side of (3.5) is bounded by

$$\big(1 - \mathrm{e}^{-\frac{\mu_N}{E_\nu^\tau[H_x]}} + (\lambda_Y E_\nu^\tau[H_x])^{-1}\big) E_x^\tau[\ell_{\mu_N}(x)^\alpha]. \tag{3.7}$$

Finally, by [ČW17, Propositions 3.1 and 4.1], $\lambda_Y \ge N^{-c}$, and $E_\nu^\tau[H_x] = 2^{N(1+o(1))}$, $\mathbb{P}$-a.s. for $N$ large enough, for all $x \in \mathbb{H}_N$. Hence, the second summand in (3.3) is bounded by $2^{-N(1-o(1))}E_x^\tau[\ell^\alpha_{\mu_N}(x)]$, which completes the proof. □

The second goal of this section is to estimate the probability of certain bad random environments $\tau$ for which a control of $R_N$ is very difficult. To this end, we fix $\eta > 0$ small and call $\tau$ good if the following conditions are satisfied:

(i) The normalisation factor $Z_N^\tau$ of (2.3) satisfies $2^{N-2} \le Z_N^\tau \le 2^N$.
(ii) The set $\mathcal{D}_N^\tau$ of deep traps satisfies $|\mathcal{D}_N^\tau| \in (1-\eta, 1+\eta)2^{N(1-\gamma')}$.
(iii) The size of the largest connected component of the set $\{x \in \mathbb{H}_N : \tau_x \ge \mathrm{e}^{\beta N^{3/4}}\}$ is smaller than $N$.
(iv) The hitting times from equilibrium are well behaving: $E_\nu^\tau[H_x] \ge 2^{N-N^\eta}$ for all $x \in \mathbb{H}_N$.
(v) The spectral gap $\lambda_Y^\tau$ is not too small, $\lambda_Y^\tau \ge \exp\{-\beta\sqrt{N}\}$.

We write $\mathcal{G}$ for the set of good $\tau$'s. In the next lemma we show that bad environments have extremely small probability.

**Lemma 3.2.** *There exist small constants $\eta, \varepsilon > 0$ such that $\mathbb{P}[\tau \notin \mathcal{G}] \le 2^{-(1+\varepsilon)N}$ for all $N$ large enough.*

*Proof.* We estimate the probabilities of the complements of the events in conditions (i)–(v) one by one:

For (i), recall that $Z_N = \sum_{x\in\mathbb{H}_N}(1 \wedge \tau_x)$, and thus $Z_N \le 2^N$. On the other hand, since $\tau_x$ are i.i.d., $Z_N$ stochastically dominates a binomial random variable with parameters $(2^N, 1/2)$. Hence, $P[Z_N \le 2^{N-2}] \le \exp(-c2^N)$ for some $c > 0$, by a standard large deviation estimate.

Similar argument apply for (ii). Since $\tau_x$'s are independent and (2.12) holds, the random variable $|\mathcal{D}_N|$ has binomial distribution with parameters $(2^N, 2^{-\gamma' N(1+o(1))})$.

The standard estimates on large deviations of binomial distribution then lead to $\mathbb{P}[|\mathcal{D}_N| \notin (1-\eta, 1+\eta)2^{(1-\gamma')N}] \le \exp(-2^{(1-\gamma'-\varepsilon)N})$.

For (iii), observe that by standard Gaussian tail estimates $\mathbb{P}[\tau_x \ge \mathrm{e}^{\beta N^{3/4}}] = \mathbb{P}[E_x \ge N^{1/4}] \le \mathrm{e}^{-\sqrt{N}/2}$. Therefore, for any $y \in \mathbb{H}_N$, the usual percolation arguments imply that the size of the connected component $\mathcal{C}_y$ of the level set $\{x : \tau_x \ge \mathrm{e}^{\beta\sqrt{N}}\}$ containing $y$ is stochastically dominated by the total progeny $\mathcal{T}$ of a Galton-Watson process with binomial$(N, \mathrm{e}^{-\sqrt{N}/2})$ offspring distribution. By, e.g., [vdH17, Theorem 3.13], for every $k \ge 1$, $\mathbb{P}[\mathcal{T} = k] = k^{-1}\mathbb{P}\big[\sum_{i=1}^k \xi_i = k-1\big]$, where $\xi_i$ are i.i.d. binomial random variables with parameters $(N, \mathrm{e}^{-\sqrt{N}/2})$. Therefore, by the exponential Markov inequality,

$$\begin{aligned} \mathbb{P}[|\mathcal{C}_y| \ge N] &\le \sum_{k=N}^{\infty} \mathbb{P}\Big[\sum_{i=1}^{k} \xi_i \ge k-1\Big] \\ &\le \sum_{k=N}^{\infty} \mathrm{e}^{-\lambda(k-1)}(1+\mathrm{e}^{-\sqrt{N}/2}(\mathrm{e}^{\lambda}-1))^{Nk} \le \mathrm{e}^{-cN^{3/2}}, \end{aligned} \tag{3.8}$$

for some $c > 0$, where the last inequality follows after taking $\lambda = \sqrt{N}/4$, after an easy computation. Summing over all $y \in \mathbb{H}_N$ then completes the proof for the condition (iii).

The probabilities of (iv) and (v) are slightly more difficult to estimate. We therefore rely on the computations of [ČW17]. For (iv), it was proved in [ČW17, Proposition 4.1], that $\mathbb{P}$-a.s. for all $N$ large enough $E_\nu^\tau H_x \ge 2^{N-N^\eta}$ for $\eta$ sufficiently small. Inspecting the proof of this proposition, reveals that $E_\nu^\tau H_x \ge 2^{N-N^\eta}$ if $\tau$ satisfies a certain property introduced in Lemma 4.2 of [ČW17]. From the proof of this lemma then follows that this property is not satisfied with probability smaller than $\exp(-N^{1+\varepsilon})$, see the last formula of the proof of Lemma 4.2 on page 271 in [ČW17], which is sufficient to deal with the case (iv).

For (v), Proposition 3.1 of [ČW17] provides a lower bound $\lambda_Y \ge N^{-c(\beta)}$, $\mathbb{P}$-a.s. for all $N$ large. Inspecting the proof of this proposition however reveals that the estimate on the probability of the complementary event is too large, namely $\mathrm{e}^{-c\sqrt{N}\ln N}$, which is not sufficient for our purposes. To show that

$$\mathbb{P}[\lambda_Y < \exp\{-\beta\sqrt{N}\}] \le 2^{-(1+\varepsilon)N}, \tag{3.9}$$

we thus need to rerun the proof of Proposition 3.1 of [ČW17] with different parameters. The required modifications are luckily rather self-contained, so we only describe them here: In Lemma 3.2 of [ČW17] and its proof, all occurrences of $N^{-\beta C_0}$ should be replaced by $\mathrm{e}^{-\beta\sqrt{N}}$, in particular a point $x \in \mathbb{H}_N$ should be called good if $\tau_x \ge \mathrm{e}^{-\beta\sqrt{N}}$, that is $E_x \ge -1$. It follows that $\mathbb{P}[x \text{ is good}] \ge 4/5$, and therefore the inequality (3.2) of [ČW17] becomes

$$\begin{aligned}\mathbb{P}[\exists x \in \mathbb{H}_N : x \text{ has fewer than } C_0\sqrt{N}/2 \text{ good neighbours}]\\ \le 2^N \mathbb{P}[\mathrm{Bin}(N, 4/5) \le C_0\sqrt{N}/2] \le 2^{-(1+2\varepsilon)N},\end{aligned} \tag{3.10}$$

for $\varepsilon > 0$ sufficiently small, where the last inequality again follows by a large deviation argument. With this change, the remaining parts of the proof require only straightforward modifications and yield the estimate (3.9). □

The last lemma of this section explains the importance of condition (iii) of the definition of $\mathcal{G}$. Its proof is inspired by [Gay18, Proposition 3.8], but it is simpler since we require a weaker statement.We write

$$\mathcal{R}_{\mu_N} = \{Y_t : t \le \mu_N\} \tag{3.11}$$

for the range of $Y$ up to time $\mu_N$.

**Lemma 3.3.** *If the condition (iii) of $\tau \in \mathcal{G}$ is satisfied, then for every $x \in \mathbb{H}_N$*

$$E_x[|\mathcal{R}_{\mu_N}|] \le N^2 \mu_N \mathrm{e}^{\beta N^{3/4}} \le 2^{No(1)}. \tag{3.12}$$

*Proof.* Recall (2.2) and observe that $q_{xy} \ge \mathrm{e}^{\beta N^{3/4}}$ iff both $\tau_x$ and $\tau_y$ is larger than $\mathrm{e}^{\beta N^{3/4}}$. On the other hand, since $\tau \in \mathcal{G}$ and thus the size of the largest connected component of $\{x : \tau_x \ge \mathrm{e}^{\beta N^{3/4}}\}$ is at most $N$, $|\mathcal{R}_{\mu_N}| \le N J_{\mu_N}$, where $J_{\mu_N}$ is the number of jumps of $Y$ before $\mu_N$ along edges with rate smaller than $\mathrm{e}^{\beta N^{3/4}}$,

$$J_{\mu_N} = \left|\{t \le \mu_N : Y_{t-} = x \ne Y_t = y \text{ such that } q_{xy} < \mathrm{e}^{\beta N^{3/4}}\}\right|. \tag{3.13}$$

Since any $x \in \mathcal{H}_N$ is incident to $N$ edges, the maximal instantaneous rate at which a new point is added to $J_{\mu_N}$ is $N\mathrm{e}^{\beta N^{3/4}}$, and thus $J_{\mu_N}$ is stochastically dominated by a Poisson random variable with mean $\mu_N N \mathrm{e}^{\beta N^{3/4}}$, in particular $E^\tau_x[|\mathcal{R}_{\mu_N}|] \le N E^\tau_x[J_{\mu_N}] \le N^2 \mu_N \mathrm{e}^{\beta N^{3/4}}$. The last inequality of the lemma follows from the definition (3.1) of $\mu_N$. □

## 4. Proof of Theorem 2.2

We have now all ingredients needed to show our main result. To this end, we introduce two convenient abbreviations

$$\mathcal{L}^\tau_x = E^\tau_x[\ell_{\mu_N}(x)^\alpha], \tag{4.1}$$

$$\mathcal{H}^\tau_x = E^\tau_\nu[H_x]. \tag{4.2}$$

and define random variables

$$F_N = F^\tau_N = \mathbf{1}_{\{\tau \in \mathcal{G}\}} \sum_{x \in \mathcal{D}_N} \frac{\mathcal{L}^\tau_x}{\mathcal{H}^\tau_x}. \tag{4.3}$$

In view of the definition (2.13) of $R_N$ and Lemmas 3.1, 3.2, $\mathbb{P}$-a.s.,

$$\lim_{N\to\infty} R_N F_N 2^{(\gamma'-\gamma)N} = 1. \tag{4.4}$$

Hence, to show Theorem 2.2, we should prove that, for $h_N$ as in the theorem,

$$\lim_{N\to\infty} h_N 2^{\gamma' N} F_N = 1, \qquad \mathbb{P}\text{-a.s,} \tag{4.5}$$

that is that $F_N$ concentrates around its expectation.

To this end, observe first that conditions (ii), (iv) of the definition of $\mathcal{G}$ and the fact that $\mathcal{L}_x^\tau \le \mu_N$ imply that, uniformly for all $\tau$,

$$F_N^\tau \le 2^{-\gamma' N(1-o(1))}. \tag{4.6}$$

On the other hand, due to (2.9) and (4.4), $\mathbb{P}$-a.s. for all $N$ large,

$$F_N^\tau \ge 2^{-\gamma' N(1+o(1))}, \tag{4.7}$$

and thus $\mathbb{E}F_N = 2^{-\gamma' N(1+o(1))}$. To prove the concentration we should thus show

$$\operatorname{Var} F_N(\tau) \le 2^{-(2\gamma'+\varepsilon)N}, \quad \text{for some } \varepsilon > 0. \tag{4.8}$$

Statement (2.14) of Theorem 2.2 then follows from (4.5)–(4.8) by a Borel-Cantelli argument. The independence of $h_N$ of $\gamma'$ is a consequence of Theorem 2.1: Since the limit of the rescaled clock process $S_N$ of (2.8) does not depend on the choice of $\gamma'$ in the definition (2.13) of $R_N$, and (2.14) allows to replace $R_N$ by $h_N e^{\alpha^2\beta^2/2}$, it must be possible to choose $h_N$ independent of $\gamma'$.

The rest of this paper proves (4.8). As we announced in the introduction, its proof uses the classical Efron-Stein inequality (see [ES81] for the original reference and [BLM13, Theorem 3.1] for the version of this inequality that we use). Let $(E'_x : x \in \mathbb{H}_N)$ be i.i.d. standard normal random variables defined on $(\Omega, \mathcal{A}, \mathbb{P})$ which are independent of the original energies $(E_x : x \in \mathbb{H}_N)$. Set $\tau'_x = \exp\{\beta\sqrt{E'_x}\}$, and for every $z \in \mathbb{H}_N$ define a new random environment $\tau^z$ by

$$\tau_x^z = \begin{cases} \tau'_x, & \text{if } x = z, \\ \tau_x, & \text{otherwise.} \end{cases} \tag{4.9}$$

Then, by Efron-Stein inequality,

$$\operatorname{Var} F_N \le \sum_{z\in\mathbb{H}_N} \mathbb{E}\big[\big(F_N(\tau) - F_N(\tau^z)\big)^2\big]. \tag{4.10}$$

We start with few preparatory claims. For $z \in \mathbb{H}_N$, let

$$B_z = B(z,1) = \{y \in \mathbb{H}_N : \operatorname{dist}(y,z) \le 1\}, \tag{4.11}$$

and, for $x, z \in \mathbb{H}_N$, let $H_x^z$ be the first time when $Y$ hits $x$ after hitting $B_z$,

$$H_x^z = \inf\{t \ge 0 : Y_t = x \text{ and there is } s < t \text{ such that } Y_s \in B_z\}. \tag{4.12}$$

Finally, let

$$\mathcal{P}(x,z,\tau) = P_x^\tau[H_x^z \le \mu_N] + P_x^{\tau^z}[H_x^z \le \mu_N]. \tag{4.13}$$

be the probability that $Y$ makes a round from $x$ to $z$ and back before time $\mu_N$, either in environment $\tau$ or $\tau^z$.

The next two lemmas bound the differences $\mathcal{H}_x^\tau - \mathcal{H}_x^{\tau^z}$ and $\mathcal{L}_x^\tau - \mathcal{L}_x^{\tau^z}$ in terms of $\mathcal{P}(x,z,\tau)$:

**Lemma 4.1.** *For every $\tau \in \mathcal{G}$ and $z \in \mathbb{H}_N$ such that $\tau^z \in \mathcal{G}$ as well,*

$$\big|\mathcal{L}_x^\tau - \mathcal{L}_x^{\tau^z}\big| \le \mu_N^\alpha \mathcal{P}(x,z,\tau). \tag{4.14}$$

*Proof.* Observe that, by (2.2), the transition rates $q_{yy'}$ of the process $Y$ depend on $\tau_z$ only if $y \in B_z$. Hence, the measures $P_x^\tau$ and $P_x^{\tau^z}$ agree on the "stopped" $\sigma$-algebra $\sigma(Y_s : s \le H_{B_z})$. In particular,

$$\begin{aligned} E_x^\tau\big[\ell_{\mu_N}^\alpha(x)\mathbf{1}_{\{H_{B_z}>\mu_N\}}\big] &= E_x^{\tau^z}\big[\ell_{\mu_N}^\alpha(x)\mathbf{1}_{\{H_{B_z}>\mu_N\}}\big],\\ E_x^\tau\big[\ell_{H_{B_z}}^\alpha(x)\mathbf{1}_{\{H_{B_z}\le\mu_N\}}\big] &= E_x^{\tau^z}\big[\ell_{H_{B_z}}^\alpha(x)\mathbf{1}_{\{H_{B_z}\le\mu_N\}}\big]. \end{aligned} \tag{4.15}$$

Therefore, using $\ell_{\mu_N}^\alpha(x) \le \ell_{H_{B_z}}^\alpha(x) + (\ell_{\mu_N}(x) - \ell_{H_{B_z}}(x))^\alpha$ on $H_{B_z} \le \mu_N$, since $\alpha < 1$,

$$\begin{aligned} \mathcal{L}_x^\tau - \mathcal{L}_x^{\tau^z} &= E_x^\tau\big[\ell_{\mu_N}^\alpha(x)\mathbf{1}_{\{H_{B_z}\le\mu_N\}}\big] - E_x^{\tau^z}\big[\ell_{\mu_N}^\alpha(x)\mathbf{1}_{\{H_{B_z}\le\mu_N\}}\big]\\ &\le E_x^\tau\big[\big(\ell_{H_{B_z}}^\alpha(x) + (\ell_{\mu_N}(x) - \ell_{H_{B_z}}(x))^\alpha\big)\mathbf{1}_{\{H_{B_z}\le\mu_N\}}\big]\\ &\quad - E_x^{\tau^z}\big[\ell_{H_{B_z}}^\alpha(x)\mathbf{1}_{\{H_{B_z}\le\mu_N\}}\big]\\ &= E_x^\tau\big[(\ell_{\mu_N}(x) - \ell_{H_{B_z}}(x))^\alpha\mathbf{1}_{\{H_{B_z}\le\mu_N\}}\big]\\ &\le \mu_N^\alpha P_x^\tau[H_x^z \le \mu_N], \end{aligned} \tag{4.16}$$

where, in the last inequality, we applied the strong Markov property at the time $H_{B_z}$ and used $\ell_{\mu_N}(x) \le \mu_N$. Repeating the same argument with the role of $\tau$ and $\tau^z$ reversed, the claim of the lemma follows by recalling the definition (4.13) of $\mathcal{P}(x,z,\tau)$. ☐

**Lemma 4.2.** *Let $\tau \in \mathcal{G}$, $x \in \mathcal{D}_N^\tau$, and let $z \in \mathbb{H}_N \setminus \{x\}$ be such that $\tau^z \in \mathcal{G}$ as well. Then*

$$\big|\mathcal{H}_x^\tau - \mathcal{H}_x^{\tau^z}\big| \le \mu_N\big(1 + Z_N\mathcal{P}(x,z,\tau)\big). \tag{4.17}$$

*Proof.* By, e.g., [AF02, Lemma 2.12] (cf. also Section 2.2.3 in the same reference),

$$\mathcal{H}_x^\tau = E_\nu^\tau[H_x] = \frac{1}{\nu_x}\int_0^\infty \big(P_x^\tau[Y_s = x] - \nu_x\big)\,\mathrm{d}s. \tag{4.18}$$

In addition, by the usual spectral decomposition for Markov chains,

$$P_x^\tau[Y_s = y] = \nu_x^{-1}\sum_{i=0}^{2^N-1}(\mathbf{1}_x,\psi_i)\mathrm{e}^{-\lambda_i s}(\mathbf{1}_y,\psi_i), \tag{4.19}$$

where $\lambda_i$ and $\psi_i$, $i = 0,\dots,2^N-1$, are the eigenvalues and the corresponding orthonormal eigenfunctions of the operator $(Qf)(x) = \sum_{y\in\mathbb{H}_N} q_{xy}(f(y) - f(x))$ acting on $L^2(\nu)$, such that $\psi_0 = \mathbf{1}$, and $0 = \lambda_0 < \lambda_Y = \lambda_1 \le \lambda_i$ for any $i \ge 2$, and $(\,\cdot\,,\,\cdot\,)$ denotes the scalar product on $L^2(\nu)$. Taking $x = y$, this implies

$$|P_x^\tau[Y_s = x] - \nu_x| \le \mathrm{e}^{-\lambda_Y s}. \tag{4.20}$$

Hence, the part of (4.18) corresponding to integral over $s \geq \mu_N$ can be bounded by $\nu_x^{-1} \mathrm{e}^{-\lambda_Y \mu_N}$. If $\tau \in \mathcal{G}$, then $\lambda_Y \mu_N \geq N^2$, cf. (3.1) and condition (v) of $\tau \in \mathcal{G}$. Therefore, this part of the integral is smaller than $\mathrm{e}^{-N^2/2}$, which is significantly smaller than the right-hand side of (4.17). The same holds true for $\tau^z$ in place of $\tau$. Therefore, ignoring those negligible errors,

$$\begin{aligned} \big|\mathcal{H}_x^\tau - \mathcal{H}_x^{\tau^z}\big| &\leq \Big| \int_0^{\mu_N} \Big( \frac{P_x^\tau[Y_s = x]}{\nu_x^\tau} - \frac{P_x^{\tau^z}[Y_s = x]}{\nu_x^{\tau^z}} \Big) \,\mathrm{d}s \Big| \\ &\leq \Big| \int_0^{\mu_N} \frac{1}{\nu_x^\tau} \big( P_x^\tau[Y_s = x] - P_x^{\tau^z}[Y_s = x] \big) \,\mathrm{d}s \Big| \\ &\quad + \Big| \int_0^{\mu_N} P_x^{\tau^z}[Y_s = x] \Big( \frac{1}{\nu_x^\tau} - \frac{1}{\nu_x^{\tau^z}} \Big) \,\mathrm{d}s \Big|. \end{aligned} \tag{4.21}$$

By similar arguments as in the proof of the previous lemma, for every $s \leq \mu_N$,

$$\big| P_x^\tau[Y_s = x] - P_x^{\tau^z}[Y_s = x] \big| \leq \mathcal{P}(x, z, \tau). \tag{4.22}$$

In addition, by (2.3), for every $x \in \mathcal{D}_N^\tau$, $\nu_x^\tau = (Z_N^\tau)^{-1}$. Since $z \neq x$, and thus $x \in \mathcal{D}_N^{\tau^z}$ as well,

$$\Big| \frac{1}{\nu_x^\tau} - \frac{1}{\nu_x^{\tau^z}} \Big| = |Z_N^\tau - Z_N^{\tau^z}| \leq 1, \tag{4.23}$$

by the definition of $Z_N$ (see (2.3)). Inserting these estimates into (4.21) completes the proof of the lemma. □

The last two lemmas together imply that for any $\tau \in \mathcal{G}$, $x \in \mathcal{D}_N^\tau$ and $z \neq x$ such that $\tau^z \in \mathcal{G}$,

$$\Big| \frac{\mathcal{L}_x^\tau}{\mathcal{H}_x^\tau} - \frac{\mathcal{L}_x^{\tau^z}}{\mathcal{H}_x^{\tau^z}} \Big| \leq \frac{\mathcal{L}_x^\tau |\mathcal{H}_x^\tau - \mathcal{H}_x^{\tau^z}| + \mathcal{H}_x^{\tau^z} |\mathcal{L}_x^\tau - \mathcal{L}_x^{\tau^z}|}{\mathcal{H}_x^\tau \mathcal{H}_x^{\tau^z}}. \tag{4.24}$$

Hence, using also conditions (ii), (iv) of $\tau \in \mathcal{G}$ and $\mathcal{L}_x^\tau \leq \mu_N^\alpha$, the Lemmas 4.1 and 4.2 imply that, for $z \neq x$ such that $\tau, \tau^z \in \mathcal{G}$ and $x \in \mathcal{D}_N^\tau$,

$$\begin{aligned} &\Big| \frac{\mathcal{L}_x^\tau}{\mathcal{H}_x^\tau} - \frac{\mathcal{L}_x^{\tau^z}}{\mathcal{H}_x^{\tau^z}} \Big| \\ &\quad \leq \mu_N^{1+\alpha} 2^{-2N(1+o(1))} \big( 1 + 2^N \mathcal{P}(x, z, \tau) \big) + 2^{-N(1+o(1))} \mu_N^\alpha \mathcal{P}(x, z, \tau) \\ &\quad \leq 2^{-N(1+o(1))} \big( \mathcal{P}(x, z, \tau) + 2^{-N(1+o(1))} \big) =: \mathcal{E}(x, z, \tau). \end{aligned} \tag{4.25}$$

We will need the following estimates on $\mathcal{P}(x, z, \tau)$ and $\mathcal{E}(x, z, \tau)$.

**Lemma 4.3.** *Uniformly in $x \in \mathbb{H}_N$ and $\tau \in \mathcal{G}$,*

$$\sum_{z \in \mathbb{H}_N} \mathcal{P}(x, z, \tau) \leq 2^{N o(1)}, \tag{4.26}$$

*and therefore*

$$\sum_{z\in\mathbb{H}_N}\mathcal{E}(x,z,\tau)\le 2^{-N(1+o(1))}, \tag{4.27}$$

$$\sum_{z\in\mathbb{H}_N}\mathcal{E}(x,z,\tau)^2\le 2^{-2N(1+o(1))}. \tag{4.28}$$

*Proof.* Recall (3.11) and (4.13) and observe that

$$\begin{aligned}\sum_{z\in\mathbb{H}_N}\mathcal{P}(x,z,\tau)&\le\sum_{z\in\mathbb{H}_N}P_x^\tau[H_{B_z}\le\mu_N]+\sum_{z\in\mathbb{H}_N}P_x^{\tau^z}[H_{B_z}\le\mu_N]\\&\le(N+1)\big(E_x^\tau[|\mathcal{R}(\mu_N)|]+E_x^{\tau^z}[|\mathcal{R}(\mu_N)|]\big),\end{aligned} \tag{4.29}$$

where we used the fact that $|B_z|=(N+1)$. Therefore, the first claim follows from Lemma 3.3

The remaining claims are easy consequences of the first one and the fact that $\mathcal{P}(x,\tau,z)\le 2$, and so $\mathcal{P}(x,z,\tau)^2\le 2\mathcal{P}(x,z,\tau)$. □

We can now finally come back to the computation of the variance of $F_N$. By (4.10), recalling the definition (4.3) of $F_N$,

$$\begin{aligned}\mathrm{Var}F_N&=\sum_{z\in\mathbb{H}_N}\mathbb{E}\Big[\Big(\sum_{x\in\mathcal{D}_N^\tau}\frac{\mathcal{L}_x^\tau}{\mathcal{H}_x^\tau}\mathbf{1}_{\{\tau\in\mathcal{G}\}}-\sum_{x\in\mathcal{D}_N^{\tau^z}}\frac{\mathcal{L}_x^{\tau^z}}{\mathcal{H}_x^{\tau^z}}\mathbf{1}_{\{\tau^z\in\mathcal{G}\}}\Big)^2\Big]\\&=\sum_{z\in\mathbb{H}_N}\mathbb{E}\Big[\sum_{x,y\in\mathbb{H}_N}\Big(\frac{\mathcal{L}_x^\tau}{\mathcal{H}_x^\tau}\mathbf{1}_{\{x\in\mathcal{D}_N^\tau\}}-\frac{\mathcal{L}_x^{\tau^z}}{\mathcal{H}_x^{\tau^z}}\mathbf{1}_{\{x\in\mathcal{D}_N^{\tau^z}\}}\Big)\\&\qquad\times\Big(\frac{\mathcal{L}_y^\tau}{\mathcal{H}_y^\tau}\mathbf{1}_{\{y\in\mathcal{D}_N^\tau\}}-\frac{\mathcal{L}_y^{\tau^z}}{\mathcal{H}_y^{\tau^z}}\mathbf{1}_{\{y\in\mathcal{D}_N^{\tau^z}\}}\Big)\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big]\\&\quad+\sum_{z\in\mathbb{H}_N}\mathbb{E}\Big[\Big(\sum_{x\in\mathcal{D}_N^\tau}\frac{\mathcal{L}_x^\tau}{\mathcal{H}_x^\tau}\Big)^2\mathbf{1}_{\{\tau\in\mathcal{G},\tau^z\notin\mathcal{G}\}}+\Big(\sum_{x\in\mathcal{D}_N^{\tau^z}}\frac{\mathcal{L}_x^{\tau^z}}{\mathcal{H}_x^{\tau^z}}\Big)^2\mathbf{1}_{\{\tau\notin\mathcal{G},\tau^z\in\mathcal{G}\}}\Big].\end{aligned} \tag{4.30}$$

We first estimate the second sum on the right-hand side of (4.30). On $\tau\in\mathcal{G}$, we have $\mathcal{H}_x^\tau\ge 2^{N(1+o(1))}$, $|\mathcal{D}_N|\le c2^{(1-\gamma')N}$, $\mathcal{L}_x^\tau\le\mu_N^\alpha$, and similarly for $\tau^z$. Hence, this sum is bounded by

$$\begin{aligned}&\Big(2^{(1-\gamma')N}\frac{\mu_N^\alpha}{2^{N(1+o(1))}}\Big)^2\sum_{z\in\mathbb{H}_N}\big(\mathbb{P}[\tau\in\mathcal{G},\tau^z\notin\mathcal{G}]+\mathbb{P}[\tau\notin\mathcal{G},\tau^z\in\mathcal{G}]\big)\\&=2^{-2\gamma'N(1+o(1))}2^N2\,\mathbb{P}[\tau\notin\mathcal{G}]\\&\le 2^{-(2\gamma'+\varepsilon)N},\end{aligned} \tag{4.31}$$

where in the last inequality we used Lemma 3.2.

We split the first sum on the right-hand side of (4.30) into four parts, according to possible mutual equalities of $x$, $y$ and $z$:

(1) In the case $x = y = z$, observing that $z \in \mathcal{D}_N^{\tau^z}$ iff $\tau'_z \geq g'_N$ (cf. (2.11) and (4.9))

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{z\in\mathbb{H}_N}\Big(\frac{\mathcal{L}_z^\tau}{\mathcal{H}_z^\tau}\mathbf{1}_{\{z\in\mathcal{D}_N^\tau\}}-\frac{\mathcal{L}_z^{\tau^z}}{\mathcal{H}_z^{\tau^z}}\mathbf{1}_{\{z\in\mathcal{D}_N^{\tau^z}\}}\Big)^2\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big]\\
&\leq 2\mathbb{E}\Big[\sum_{z\in\mathcal{D}_N^\tau}\Big(\frac{\mathcal{L}_x^\tau}{\mathcal{H}_x^\tau}\Big)^2\mathbf{1}_{\{\tau\in\mathcal{G}\}}\Big]+2\mathbb{E}\Big[\sum_{z\in\mathbb{H}_N}\Big(\frac{\mathcal{L}_x^{\tau^z}}{\mathcal{H}_x^{\tau^z}}\Big)^2\mathbf{1}_{\{\tau'_z\geq g'_N,\tau^z\in\mathcal{G}\}}\Big]\\
&\leq c\mu_N^{2\alpha}2^{(1-\gamma')N}2^{-2N(1+o(1))}\leq 2^{-(2\gamma'+\varepsilon)N},
\end{aligned}\tag{4.32}$$

where we used conditions (ii) and (iv) of $\tau\in\mathcal{G}$, relation (2.12), $\mathcal{L}_x^\tau\leq\mu_N^\alpha$, and the fact that $\gamma'<1$.

(2) In the case when $x=y$ and $x\neq z$, $x\in\mathcal{D}_N^\tau$ iff $x\in\mathcal{D}_N^{\tau^z}$. Therefore, by (4.25) and Lemma 4.3, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{z\in\mathbb{H}_N}\sum_{x\in\mathcal{D}_N^\tau\setminus\{z\}}\Big(\frac{\mathcal{L}_x^\tau}{\mathcal{H}_x^\tau}-\frac{\mathcal{L}_x^{\tau^z}}{\mathcal{H}_x^{\tau^z}}\Big)^2\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big]\\
&\qquad\leq\mathbb{E}\Big[\sum_{x\in\mathcal{D}_N^\tau}\sum_{z\in\mathbb{H}_N\setminus\{x\}}\mathcal{E}(x,z,\tau)^2\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big],\\
&\qquad\leq 2^{N(1-\gamma')}2^{-2N(1+o(1))}\leq 2^{-(2\gamma'+\varepsilon)N},
\end{aligned}\tag{4.33}$$

where we used condition (ii) of $\tau\in\mathcal{G}$ and $\gamma'<1$, again.

(3) In the case when $y\neq x$ and exactly one of $x,y$ equals $z$, say for simplicity that it is $y$, we have again $x\in\mathcal{D}_N^\tau$ iff $x\in\mathcal{D}_N^{\tau^z}$. Therefore, the contribution of this case to the right-hand side of (4.30) is, by (4.25), at most

$$\begin{aligned}
&\sum_{z\in\mathbb{H}_N}\mathbb{E}\Big[\sum_{x\in\mathcal{D}_N^\tau}\mathcal{E}(x,z,\tau)\Big|\frac{\mathcal{L}_z^\tau}{\mathcal{H}_z^\tau}\mathbf{1}_{\{y\in\mathcal{D}_N^\tau\}}-\frac{\mathcal{L}_z^{\tau^z}}{\mathcal{H}_z^{\tau^z}}\mathbf{1}_{\{y\in\mathcal{D}_N^{\tau^z}\}}\Big|\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big]\\
&\qquad\leq\mathbb{E}\Big[\frac{\mu_N^\alpha}{2^{N(1+o(1))}}\sum_{x\in\mathcal{D}_N^\tau}\sum_{z\in\mathbb{H}_N}\mathcal{E}(x,z,\tau)\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big]\\
&\qquad\leq 2^{(1-\gamma')N}\mu_N 2^{-2N(1+o(1))}\leq 2^{-(2\gamma'+\varepsilon)N},
\end{aligned}\tag{4.34}$$

where we used conditions (ii), (iv) of $\tau\in\mathcal{G}$, $\gamma'<1$, and Lemma 4.3 again.

(4) We now treat the case when $x,y,z$ are different points. This the most difficult case and the only place where the fact that $\mathcal{P}(x,z,\tau)$ is defined using $H_x^z$ and not $H_{B_z}$ is used. In this case, by (4.25), we obtain the upper bound

$$\mathbb{E}\Big[\sum_{z\in\mathbb{H}_N}\sum_{x\neq y\in\mathcal{D}_N^\tau\setminus\{z\}}\mathcal{E}(x,z,\tau)\mathcal{E}(y,z,\tau)\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big].\tag{4.35}$$

Recalling first the definition of $\mathcal{E}(x,z,\tau)$ in (4.25), the summands not containing either $\mathcal{P}(x,y,\tau)$ or $\mathcal{P}(y,z,\tau)$ can be bounded by

$$2\mathbb{E}\Big[\sum_{x,y\in\mathcal{D}_N^\tau}\sum_{z\in\mathbb{H}_N}2^{-3N(1+o(1))}\big(\mathcal{P}(x,z,\tau)+2^{-N(1+o(1))}\big)\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big]\le 2^{-(2\gamma'+\varepsilon)N}, \tag{4.36}$$

similarly as in the previous steps. For the remaining summand

$$2^{-2N(1+o(1))}\mathbb{E}\Big[\sum_{z\in\mathbb{H}_N}\sum_{x,y\in\mathcal{D}_N^\tau\setminus\{z\}}\mathcal{P}(x,z,\tau)\mathcal{P}(y,z,\tau)\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big] \tag{4.37}$$

we recall the definition of $\mathcal{P}(x,z,\tau)$, and observe that

$$P_x^\tau[H_x^z\le\mu_N]\le\sum_{a\in B_z}P_x^\tau[H_a=H_{B_z}\le\mu_N]P_a^\tau[H_x\le\mu_N]. \tag{4.38}$$

Therefore,

$$\mathcal{P}(x,z,\tau)\mathcal{P}(y,z,\tau)\le\sum_{\tau',\tau''\in\{\tau,\tau^z\}}\sum_{a,b\in B_z}P_x^{\tau'}[H_a\le\mu_N]P_b^{\tau''}[H_y\le\mu_N], \tag{4.39}$$

and thus

$$\begin{aligned}&\mathbb{E}\Big[\sum_{z\in\mathbb{H}_N}\sum_{x,y\in\mathcal{D}_N^\tau\setminus\{z\}}\mathcal{P}(x,z,\tau)\mathcal{P}(y,z,\tau)\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big]\\&\le\mathbb{E}\Big[\sum_{\tau',\tau''\in\{\tau,\tau^z\}}\sum_{x\in\mathcal{D}_N^\tau}\sum_{z\in\mathbb{H}_N}\sum_{a,b\in B_z}P_x^{\tau'}[H_a\le\mu_N]\sum_{y\in\mathbb{H}_N}P_b^{\tau''}[H_y\le\mu_N]\mathbf{1}_{\{\tau,\tau^z\in\mathcal{G}\}}\Big]\end{aligned} \tag{4.40}$$

By Lemma 4.3, the summation over $y$ can be bounded by $2^{No(1)}$ uniformly in $b$. Due to the symmetry of $\mathbb{H}_n$, since $|B_z|=N+1$, the summations over $a$, and $z$ can be written as $(N+1)\sum_{z\in\mathbb{H}_N}$. For the summation over $z$ one can then apply the Lemma 4.3 again. The summation over $x$ contributes the factor $|\mathcal{D}_N^\tau|\le c2^{(1-\gamma')N}$. Therefore, the last expression is bounded by $2^{(1-\gamma'+o(1))N}$. Inserting this back into (4.37), implies that the contribution (4.35) of the forth case is bounded by $2^{-(1+\gamma'+o(1))N}\le 2^{-(2\gamma'+\varepsilon)N}$.

Putting (4.31) and the estimates from the cases (1)–(4) back into (4.30) implies that $\operatorname{Var}F_N\le 2^{-(2\gamma'+\varepsilon)N}$. This completes the proof of (4.8) and thus of Theorem 2.2.

## References

[AB92] David J. Aldous and Mark Brown, *Inequalities for rare events in time-reversible Markov chains. I*, Stochastic inequalities (Seattle, WA, 1991), IMS Lecture Notes Monogr. Ser., vol. 22, Inst. Math. Statist., Hayward, CA, 1992, pp. 1–16. MR 1228050

[AF02] David Aldous and James Allen Fill, *Reversible markov chains and random walks on graphs*, 2002, Unfinished monograph, recompiled 2014, available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[BBG03a] Gérard Ben Arous, Anton Bovier, and Véronique Gayrard, *Glauber dynamics of the random energy model. I. Metastable motion on the extreme states*, Comm. Math. Phys. **235** (2003), no. 3, 379–425. MR 1974509

[BBG03b] Gérard Ben Arous, Anton Bovier, and Véronique Gayrard, *Glauber dynamics of the random energy model. II. Aging below the critical temperature*, Comm. Math. Phys. **236** (2003), no. 1, 1–54. MR 1977880

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013, A nonasymptotic theory of independence, With a foreword by Michel Ledoux. MR 3185193

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

[ČTW11] Jiří Černý, Augusto Teixeira, and David Windisch, *Giant vacant component left by a random walk in a random d-regular graph*, Ann. Inst. Henri Poincaré Probab. Stat. **47** (2011), no. 4, 929–968. MR 2884219

[ČW17] Jiří Černý and Tobias Wassmer, *Aging of the Metropolis dynamics on the Random Energy Model*, Probab. Theory Related Fields **167** (2017), no. 1, 253–303.

[ES81] B. Efron and C. Stein, *The jackknife estimate of variance*, Ann. Statist. **9** (1981), no. 3, 586–596. MR 615434

[Gay18] Véronique Gayrard, *Aging in metropolis dynamics of the rem: a proof*, Probability Theory and Related Fields (2018), Advance online publication. doi:10.1007/s00440-018-0873-6.

[vdH17] Remco van der Hofstad, *Random graphs and complex networks. Vol. 1*, Cambridge Series in Statistical and Probabilistic Mathematics, [43], Cambridge University Press, Cambridge, 2017. MR 3617364

Jiří Černý, Department of Mathematics and Computer Science, University of Basel, 4051 Basel, Switzerland

*Email address*: jiri.cerny@unibas.ch