The Annals of Probability
2014, Vol. 42, No. 6, 2497–2557
DOI: 10.1214/13-AOP886

UNIVERSALITY OF TRAP MODELS IN THE ERGODIC TIME SCALE

BY M. JARA, C. LANDIM AND A. TEIXEIRA

IMPA, IMPA and Université de Rouen, and IMPA

Consider a sequence of possibly random graphs $G_N = (V_N, E_N)$, $N \ge 1$, whose vertices's have i.i.d. weights $\{W_x^N : x \in V_N\}$ with a distribution belonging to the basin of attraction of an α-stable law, $0 < \alpha < 1$. Let X_t^N, $t \ge 0$, be a continuous time simple random walk on G_N which waits a *mean* W_x^N exponential time at each vertex x. Under considerably general hypotheses, we prove that in the ergodic time scale this trap model converges in an appropriate topology to a K-process. We apply this result to a class of graphs which includes the hypercube, the d-dimensional torus, $d \ge 2$, random d-regular graphs and the largest component of super-critical Erdős–Rényi random graphs.

1. Introduction. Trap models were introduced to investigate aging, a non-equilibrium phenomenon of considerable physical interest [4, 8, 9, 13, 29]. These trap models are defined as follows: consider an unoriented graph $G = (V, E)$ with finite degrees and a sequence of i.i.d. strictly positive random variables $\{W_z : z \in V\}$ indexed by the vertices. Let $\{X_t : t \ge 0\}$ be a continuous-time random walk on V which waits a mean W_z exponential time at site z, at the end of which it jumps to one of its neighbors with uniform probability.

The expected time spent by the random walk on a vertex z is proportional to the value of W_z. It is thus natural to regard the environment W as a landscape of valleys or traps with depth given by the values of the random variables $\{W_z : z \in V\}$. As the random walk evolves, it explores the random landscape, finding deeper and deeper traps, and aging appears as a consequence of the longer and longer times the process remains at the same vertex.

Assume that the distribution of W_x belongs to the domain of attraction of an α-stable law, $0 < \alpha < 1$. The variables $\{W_x : x \in V\}$ take now large values in certain sites, forcing the random walk to stay still for a long time when it reaches one of them, causing a macroscopic subdiffusive behavior.

In dimension 1, Fontes, Isopi and Newman [18] proved under these hypotheses that for almost all environments, the random walk converges, in the time scale $t^{1+(1/\alpha)}$, to a singular diffusion with a random discrete speed measure. In dimension $d \ge 2$, Ben Arous and Černý [6] proved that for almost all environments the

Received August 2012; revised August 2013.
MSC2010 subject classifications. 60J27, 60K37, 82C41.
Key words and phrases. Trap models, scaling limit, metastability.

Bouchaud trap model converges in a proper time scale, $t^{2/\alpha}$ in dimension $d \geq 3$ and a scale logarithmic smaller than $t^{2/\alpha}$ in dimension 2, to the fractional-kinetic process, a self-similar, non-Markovian, continuous process, obtained as the time change of a Brownian motion by the inverse of an independent α-stable subordinator. In fact, they proved, under quite general conditions on the environment, that the clock process converges to an α-stable subordinator, for a large range of time scales [7]. In these time scales, the random walk does not visit the deepest traps, but exhibit an aging behavior. During the exploration of the random scenery, the process discovers deeper and deeper traps which slow down its evolution, the mechanism responsible for the aging phenomenon. We refer to [5, 10] for recent reviews.

The investigation of trap models on graphs in the time scale in which the deepest traps are visited started with Fontes and Mathieu [20]. The authors proved that the random walk among random traps in the complete graph converges to the K-process, a continuous-time, Markov dynamics on $\bar{\mathbb{N}}$, the one point compactification of $\mathbb{N}$, which hits any finite subset A of $\mathbb{N}$ with uniform distribution. This latter result was extended by Fontes and Lima [19] to the hypercube and by us [24] to the d-dimensional torus, $d \geq 2$.

In the present paper, we exhibit simple conditions that imply the convergence to the K-process in the scaling limit. Our conditions are general enough to include the hypercube and the torus, as well as random d-regular graphs and the largest component of the super-critical Erdős–Rényi random graphs. These are good examples to keep in mind throughout the text.

Let $\{G_N : N \geq 1\}$, $G_N = (V_N, E_N)$, be a sequence of possibly random, finite, connected graphs defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where V_N represents the set of vertices and E_N the set of unoriented edges. Assume that the number of vertices, $|V_N|$, converges to $+\infty$ in $\mathbb{P}$-probability.

Assume that on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we are given an i.i.d. collection of random variables $\{W_j^N : j \geq 1\}$, $N \geq 1$, independent of the random graph G_N and whose common distribution belongs to the basin of attraction of an α-stable law, $0 < \alpha < 1$. Hence, for all $N \geq 1$ and $j \geq 1$,

$$\mathbb{P}[W_1^N > t] = \frac{L(t)}{t^\alpha}, \qquad t > 0, \tag{1.1}$$

where L is a slowly varying function at infinity.

For each $N \geq 1$, reenumerate in decreasing order the weights $W_1^N, \dots, W_{|V_N|}^N$: $\hat{W}_j^N = W_{\sigma(j)}^N$, $1 \leq j \leq |V_N|$ for some permutation σ of the set $\{1, \dots, |V_N|\}$ and $\hat{W}_j^N \geq \hat{W}_{j+1}^N$ for $1 \leq j < |V_N|$. Let $(x_1^N, \dots, x_{|V_N|}^N)$ be a random enumeration of the vertices of G_N and define $W_{x_j^N}^N = \hat{W}_j^N$, $1 \leq j \leq |V_N|$, turning $G_N = (V_N, E_N, W^N)$ into a finite, connected, vertex-weighted graph.

Consider for each $N \geq 1$, a continuous-time random walk $\{X_t^N : t \geq 0\}$ on V_N, which waits a mean W_x^N exponential time at site x, after which it jumps to one of its neighbors with uniform probability. The generator $\mathcal{L}_N$ of this walk is given by

$$(\mathcal{L}_N f)(x) = \frac{1}{\deg(x)} \frac{1}{W_x^N} \sum_{y \sim x} [f(y) - f(x)] \tag{1.2}$$

for every $f : V_N \to \mathbb{R}$, where $y \sim x$ means that $\{x, y\}$ belongs to the set of edges E_N and where $\deg(x)$ stands for the degree of x: $\deg(x) = \#\{y \in V_N : y \sim x\}$.

Heuristics. The main results of this article assert that, under fairly general conditions on the graph sequence G_N, the random walk X_t^N converges in the ergodic time scale to a K-process. Let us now give an informal description of the above statement.

Given the graph sequence G_N and the associated weights W_x^N, suppose that:

(a) A small number of sites supports most of the stationary measure of the process X_t^N [see (B.0)],

and that we are able to find a sequence ℓ_N satisfying the following conditions:

(b) the ball $B(x, \ell_N)$ around a typical point x has a volume much smaller than $|V_N|$; see (B.1),

(c) starting outside of the above ball, the random walk "mixes" before hitting its center x; see (B.2) and

(d) the graphs G_N are transitive [or satisfy the much weaker hypothesis (B.3)].

Under the above conditions, we are able to show that

$$X_t^N \text{ converges to a } K\text{-process,} \tag{1.3}$$

introduced in [20, 31], after proper scaling, see Theorems 2.1 and 2.2.

Still on a heuristic level, let us give a brief explanation of why the above conditions should imply the stated convergence. Let M_N be a sequence of integers converging to $+\infty$ slowly enough for the balls $B(x_j^N, \ell_N)$, $1 \le j \le M_N$, to be disjoint. We call the vertices $\{x_1^N, \dots, x_{M_N}^N\}$ the *deep traps* and the remaining vertices $\{x_{M_N+1}^N, \dots, x_{|V_N|}^N\}$ the *shallow traps*. The idea is to decompose the trajectory of the random walk in excursions between the successive visits to the balls $B(x_j^N, \ell_N)$.

Denote by $v_{\ell_N}(x_j^N)$ the escape probability from x_j^N. This is the probability that the random walk X_t^N starting from x_j^N attains the boundary of the ball $B(x_j^N, \ell_N)$ before returning to x_j^N. The random walk X_t^N starting from x_j^N visits x_j^N on average $v_{\ell_N}(x_j^N)^{-1}$ times before it escapes. After escaping, it mixes and then it reaches a new deep trap with a distribution determined by the topology of the graph. This distribution does not depend on the last deep trap visited because the process has mixed before reaching the next trap. In an excursion between two deep traps, the random walk visits only shallow traps, which should not influence the asymptotic behavior.

Hence, if the escape probabilities and the degrees of the random graph have a reasonable asymptotic behavior [see (B.3)], we expect the random walk X_t^N to evolve as a Markov process on $\{1, \dots, M_N\}$ which waits at site j a mean $W^N_{x^N_j} v_{\ell_N}(x_j^N)^{-1}$ exponential time, at the end of which it jumps to a point in $\{1, \dots, M_N\}$ whose distribution does not depend on j. This latter process can be easily shown to converge to the K-process, proving the main result of this article.

There are several interesting examples of random graphs which are not considered in this article, either because assumptions (B.0)–(B.3) fail or because they have not been proved yet. We leave as open problems the asymptotic behavior of a random walk among random traps on uniform trees on N vertices, on the critical component of an Erdős–Rényi graph, on Sierpinski carpets, on the giant component of the percolation cluster on a torus or on the invasion percolation cluster.

The article is organized as follows. In the next section, we give a precise statement of our main results. In the following two sections, we present some preliminary results on hitting probabilities and holding times of a random walk among random traps. In Section 5, we present the topology in which the convergence to the K-process takes place and in Section 6 we construct a coupling between the random walk and a Markov process on the set $\{1, \dots, M\}$. This latter process can be seen as the trace of the K-process on the set $\{1, \dots, M\}$ and the coupling as the main step of the proof. In Section 7, we show that this latter process converges to the K-process. Putting together the assertions of Sections 5, 6, 7, we derive in Section 8 a result which provides sufficient conditions for the convergence to the K-process of a sequence of random walks among random traps on deterministic graphs. We adapt this result in Section 9 to random pseudo-transitive graphs and in Section 10 to graphs with asymptotically random conductances. We show in Section 11 that this latter class includes the largest component of a super-critical Erdős–Rényi graphs.

2. Notation and results.

Recall the notation introduced in the previous section up to the Section Heuristics. Denote by ν_N the unique stationary distribution of the process $\{X_t^N : t \ge 0\}$. An elementary computation shows that ν_N is in fact reversible and given by

$$\nu_N(x) = \frac{\deg(x) W_x^N}{Z_N}, \qquad x \in V_N, \tag{2.1}$$

where Z_N is the normalizing constant $Z_N = \sum_{y \in V_N} \deg(y) W_y^N$.

For a fixed graph G_N and a fixed environment $\mathbf{W} = \{W_z^N : z \in V_N\}$, denote by $\mathbf{P}_x^N = \mathbf{P}_x^{G_N, \mathbf{W}}$, $x \in V_N$, the probability on the path space $D(\mathbb{R}_+, V_N)$ induced by the Markov process $\{X_t^N : t \ge 0\}$ starting from x. Expectation with respect to $\mathbf{P}_x^N$ is represented by $\mathbf{E}_x^N$. We denote sometimes X_t^N by $X^N(t)$ to avoid small characters.

Let $\{\mathbb{X}_n^N : n \ge 0\}$ be the lazy embedded discrete-time chain in X_t^N, that is, the discrete-time Markov chain which jumps from x to y with probability

$(1/2)\deg(x)^{-1}$ if $y \sim x$ and which jumps from x to x with probability $(1/2)$. Denote by π_N the unique stationary, in fact reversible, distribution of the skeleton chain, given by

$$\pi_N(x) = \frac{\deg(x)}{\sum_{y \in V_N} \deg(y)}. \tag{2.2}$$

For a subset B of V_N, we denote by H_B the hitting time of B and by H_B^+ the return time to B:

$$H_B = \inf\{t \geq 0 : X_t^N \in B\},$$
$$H_B^+ = \inf\{t \geq 0 : X_t^N \in B \text{ and } \exists s < t \text{ s.t. } X_s^N \notin B\}.$$

When B is a singleton $\{x\}$, we denote H_B, H_B^+ by H_x, H_x^+, respectively. We also write $\mathbb{H}_B$ (resp., $\mathbb{H}_B^+$) for the hitting time of a set B (resp., return time to B) for the discrete chain $\mathbb{X}_n^N$.

K-processes. To describe the asymptotic behavior of the random walk X_t^N, consider two sequences of positive real numbers $\mathbf{u} = \{u_k : k \in \mathbb{N}\}$ and $\mathbf{Z} = \{Z_k : k \in \mathbb{N}\}$ such that

$$\sum_{k \in \mathbb{N}} Z_k u_k < \infty, \qquad \sum_{k \in \mathbb{N}} u_k = \infty. \tag{2.3}$$

Consider the set $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ of nonnegative integers with an extra point denoted by ∞. We endow this set with the metric induced by the isometry $\phi : \bar{\mathbb{N}} \to \mathbb{R}$, which sends $n \in \bar{\mathbb{N}}$ to $1/n$ and ∞ to 0. This makes the set $\bar{\mathbb{N}}$ into a compact metric space.

In Section 7, based on [20], we construct a Markov process on $\bar{\mathbb{N}}$, called the K-process with parameter (Z_k, u_k) which can be informally described as follows. Being at $k \in \mathbb{N}$, the process waits a mean Z_k exponential time, at the end of which it jumps to ∞. Immediately after jumping to ∞, the process returns to $\mathbb{N}$. The hitting time of any finite subset A of $\mathbb{N}$ is almost surely finite. Moreover, for each fixed $n \geq 1$, the probability that the process hits the set $\{1, \dots, n\}$ at the state k is equal to $u_k / \sum_{1 \leq j \leq n} u_j$. In particular, the trace of the K-process on the set $\{1, \dots, n\}$ is the Markov process which waits at k a mean Z_k exponential time at the end of which it jumps to j with probability $u_j / \sum_{1 \leq i \leq n} u_i$.

Topology. Between two successive sojourns in deep traps, the random walk X_t^N visits in a short time interval several shallow traps. If we want to prove the convergence of the process X_t^N to a process which visits only the deep traps, we need to consider a topology which disregard short excursions. With this in mind, we introduce the following topology.

Fix $T > 0$. For any function $f : [0, T] \to \mathbb{R}$ and any point $t \in [0, T]$, we say that f is locally constant at t if f is constant in a neighborhood of t. Let

$$\mathcal{C}(f) = \{t \in [0, T];\ f \text{ is locally constant in } t\}, \tag{2.4}$$

and $\mathcal{D}(f) = \mathcal{C}(f)^c$. Notice that the set $\mathcal{D}(f)$ is always closed. Let Λ denote the Lebesgue measure in $[0, T]$ and denote by $\mathfrak{M}_0$ the space of functions which are locally constant a.e., that is,

$$\mathfrak{M}_0 := \{f : [0, T] \to \mathbb{R}; \Lambda(\mathcal{D}(f)) = 0\}. \tag{2.5}$$

We say that two locally constant functions f and $g \in \mathfrak{M}_0$ are equivalent if $f(t) = g(t)$ for any $t \notin \mathcal{D}(f) \cup \mathcal{D}(g)$. Note that if f and g are equivalent then $f = g$ almost everywhere.

Let make the space $\mathfrak{M}_0$ into a metric space by introducing the distance

$$d_T(f, g) = \inf_{A \in \mathcal{B}} \{\|f - g\|_{\infty, A^c} + \Lambda(A)\}, \tag{2.6}$$

where $\mathcal{B} = \mathcal{B}([0, T])$ is the set of Borel subsets of $[0, T]$, and $\|f - g\|_{\infty, A^c}$ stands for the supremum norm of $f - g$ restricted to A^c. Intuitively speaking, the distance between f and g is small if they are close to each other, except for a set of small measure.

We prove in Section 5 that d_T is well defined and that it introduces a metric in $\mathfrak{M}_0$ which generates the topology of convergence in measure with respect to the Lebesgue measure in $[0, T]$. With this metric, $\mathfrak{M}_0$ is separable but not complete.

Main result. Let $\mathbb{V} = \mathbb{V}_N = |V_N|$ and let $\Psi_N : V_N \to \{1, \dots, \mathbb{V}_N\}$ be the random function defined by $\Psi_N(x_j^N) = j$. The first main result of this article relies on three assumptions. We first require the sequence of invariant measures ν_N to be almost surely tight. Assume that for any increasing sequence J_N, with $\lim_N J_N = \infty$,

$$\lim_{N \to \infty} \mathbb{E}[\nu_N(\{x_1^N, \dots, x_{\min\{J_N, \mathbb{V}_N\}}^N\}^c)] = 0. \tag{B.0}$$

Denote by $B(x, \ell)$ the ball of radius ℓ centered at $x \in V_N$ with respect to the graph distance $d = d_N$ in G_N. Fix a sequence $\{\ell_N : N \geq 1\}$ of positive numbers, representing the radius of balls we place around each deep trap. Let $\mathfrak{x}$ be a vertex chosen uniformly among the vertices of V_N. We assume that

$$\lim_{N \to \infty} \mathbb{E}\left[\frac{|B(\mathfrak{x}, 2\ell_N)|}{\mathbb{V}_N}\right] = 0. \tag{B.1}$$

It follows from this condition that the number of vertices $\mathbb{V}_N$ of the graph G_N diverges in probability:

$$\lim_{N \to \infty} \mathbb{P}[\mathbb{V}_N \geq K] = 1$$

for every $K \geq 1$.

Let $\|\mu - \nu\|_{\mathrm{TV}}$ be the total variation distance between two probability measures μ, ν defined on V_N, and let $t_{\mathrm{mix}} = t_{\mathrm{mix}}^N$ be the mixing time of the discrete chain $\{\mathbb{X}_n^N : n \geq 0\}$; see equation (4.33) in [26].

We assume that the typical point $\mathfrak{x}$ is not hit before the mixing time if one starts the random walk at distance at least ℓ_N from $\mathfrak{x}$. More precisely, we suppose that there exists an increasing sequence L_N, $\lim_{N\to\infty} L_N = \infty$, such that

$$\lim_{N\to\infty} \mathbb{E}\Big[\sup_{y\notin B(\mathfrak{x},\ell_N)} \mathbf{P}_y[\mathbb{H}_{\mathfrak{x}} \le L_N t_{\text{mix}}]\Big] = 0. \tag{B.2}$$

We finally introduce the notion of pseudo-transitive graphs, which includes the classical definition of transitive graphs but also encompasses other important examples such as random regular graphs, discussed in Proposition 9.3.

Consider a sequence of possibly random graphs $G_N = (V_N, E_N)$. We say that two subsets A, B of V_N with distinguished vertices $\mathfrak{x} \in A$, $\mathfrak{y} \in B$, are isomorphic, $(\mathfrak{x}, A) \equiv (\mathfrak{y}, B)$, if there exists a bijection $\varphi : A \to B$ with the property that $\varphi(\mathfrak{x}) = \mathfrak{y}$ and that for any $a, b \in A$, $\{a, b\}$ is an edge of G_N if and only if $\{\varphi(a), \varphi(b)\}$ is an edge of G_N.

Let $\mathfrak{x}, \mathfrak{y} \in V_N$ be two vertices chosen independently and uniformly in V_N. We say that G_N is pseudo-transitive for the sequence ℓ_N, if

$$\lim_{N\to\infty} \mathbb{P}\big[\big(\mathfrak{x}, B(\mathfrak{x}, \ell_N)\big) \not\equiv \big(\mathfrak{y}, B(\mathfrak{y}, \ell_N)\big)\big] = 0. \tag{2.7}$$

Clearly, any sequence of transitive graphs is pseudo-transitive for any given sequence ℓ_N.

For $x \in V_N$, let $v_\ell(x) = v^N_{\ell_N}(x)$ be the probability of escape from x:

$$v_\ell(x) = \mathbf{P}^N_x\big[\mathbb{H}_{R(x,\ell)} < \mathbb{H}^+_x\big],$$

where $R(x, \ell) = B(x, \ell)^c$. Let $\{c_k : k \ge 1\}$ be the sequence defined by

$$c_k^{-1} = \inf\{t \ge 0 : \mathbb{P}\big[W^N_1 > t\big] \le k^{-1}\}. \tag{2.8}$$

The constant c_N^{-1} represents the typical size of $\max_{1\le k\le N} W^N_k$, so that $c_{\mathbb{V}} W^N_{x_j}$ for fixed j is of order one.

THEOREM 2.1. *Fix a sequence of pseudo-transitive graphs G_N with respect to a sequence ℓ_N. Suppose that* (B.0)–(B.2) *hold and that $\Psi_N(X^N_0)$ converges in probability to some $k \in \mathbb{N}$. Then, letting $\beta_N^{-1} = c_{\mathbb{V}} v^N_{\ell_N}(x^N_1)$, we have that*

$$\big(c_{\mathbb{V}}\mathbf{W}^N, \Psi_N\big(X^N_{t\beta_N}\big)\big) \text{ converges weakly to } (\mathbf{w}, K_t),$$

where the sequence $\mathbf{w} = (w_1, w_2, \dots)$ is defined in (8.4) *and where for each fixed $\mathbf{w}$, K_t is a K-process with parameter* $(\mathbf{w}, 1)$ *starting from k. In the convergence, we adopted $L^1(\mathbb{N})$ topology in the first coordinate and d_T-topology in the second.*

It is not difficult to show from the definition of the random sequence $\mathbf{w} = (w_1, w_2, \dots)$ that w_1 has a Fréchet distribution. In Section 9, we apply Theorem 2.1 to the hypercube, the d-dimensional torus, $d \ge 2$, and to a sequence of random d-regular graphs, $d \ge 3$.

The second main result of the article concerns graphs in which assumption (2.7) of isometry of neighborhoods is replaced by an asymptotic independence and a second moment bound.

Assume that there exists a coupling $\mathcal{Q}_N$ between the random graph $\{G_N : N \geq 1\}$ and a sequence of i.i.d. random vectors $\{(D_k, E_k) : k \geq 1\}$ (independent of N) such that for every $K \geq 1$ and $\delta > 0$,

$$\lim_{N\to\infty} \mathcal{Q}_N\Big[\max_{1\leq j\leq K} |v_\ell(\mathfrak{x}_j)^{-1} - E_j^{-1}| > \delta\Big] = 0,$$

$$\lim_{N\to\infty} \mathcal{Q}_N\left[\bigcup_{j=1}^{K} \{\deg(\mathfrak{x}_j) \neq D_j\}\right] = 0, \tag{B.3}$$

$$\mathcal{Q}_N[D_1 \geq 1, 0 < E_1 \leq 1] = 1, \qquad E_{\mathcal{Q}_N}\big[(D_1/E_1)^2\big] < \infty$$

for one and, therefore, all $N \geq 1$, where $\ell = \ell_N$ is the radius of the balls placed around each trap and introduced right above (B.1), and $\mathfrak{x}_1, \dots, \mathfrak{x}_K$ is a collection of distinct vertices chosen uniformly in V_N. We can now state our second main result, which can be seen as a generalization of Theorem 2.1.

THEOREM 2.2. *Fix a sequence of random graphs G_N. Suppose that* (B.0)–(B.3) *hold and that $\Psi_N(X_0^N)$ converges in probability to some $k \in \mathbb{N}$. Then, defining $\beta_N = c_{\mathbb{V}}^{-1}$, we have that*

$$\big(c_{\mathbb{V}}\mathbf{W}^N, \Psi_N\big(X_{t\beta_N}^N\big)\big) \text{ converges weakly to } (\mathbf{w}, K_t),$$

where the sequence $\mathbf{w} = (w_1, w_2, \dots)$ is defined in (8.4) *and where for each fixed $\mathbf{w}$, K_t is a K-process starting from k with parameter $(\mathbf{Z}, \mathbf{u})$, where $Z_k = w_k/E_k$ and $u_k = D_k E_k$. In the convergence, we adopted $L^1(\mathbb{N})$ topology in the first coordinate and d_T-topology in the second.*

In Section 11, we apply this result to the largest component of a super-critical Erdős–Rényi random graph. We expect this statement to be applicable in a wider context, such as random graphs with random degree sequences, or percolation clusters on certain graphs.

3. Hitting probabilities. We prove in this section general estimates on the hitting distribution of a random walk on a finite graph. These estimates will be useful in the description of the trace of our trap model on the deepest traps. Since N will be kept fixed throughout the section, we omit N from the notation almost everywhere.

Recall that we denote by $d = d_N$ the graph distance on V_N: $d(x, y) = m$ if there exists a sequence $x = z_0, z_1, \dots, z_m = y$ such that $z_{i+1} \sim z_i$ for $0 \leq i \leq m-1$, and if there do not exist shorter sequences with this property. For $x \in V_N$ and a subset C of V_N, denote by $d(x, C)$ the distance from x to C: $d(x, C) = \min_{y\in C} d(x, y)$.

For $\ell \geq 1$, denote by $B(C, \ell)$ the vertices at distance at most ℓ from C: $B(C, \ell) = \{x \in V_N : d(x, C) \leq \ell\}$ and let $R(C, \ell) = B(C, \ell)^c$ as before. When the set C is a singleton $\{x\}$, we write $B(x, \ell)$, $R(x, \ell)$ for $B(\{x\}, \ell)$, $R(\{x\}, \ell)$, respectively.

Fix $M \geq 1$, a subset $A = \{x_1, \dots, x_M\}$ of V_N and $\ell \geq 1$. Recall from Section 2 that we denote by $v_\ell(x)$, $x \in A$, the escape probability from x, and let $p(x, A)$ be the probability of reaching the set A at x, when starting at equilibrium:

$$(3.1) \qquad v_\ell(x) = \mathbf{P}_x\big[\mathbb{H}_{R(x,\ell)} < \mathbb{H}_x^+\big], \qquad p(x, A) = \mathbf{P}_{\pi_N}\big[\mathbb{X}^N(\mathbb{H}_A) = x\big],$$

where π_N is the stationary state of the discrete-time chain $\mathbb{X}_n^N$, introduced in (2.2).

LEMMA 3.1. *Fix a subset $A = \{x_1, \dots, x_M\}$ of V. For any $z \notin A$ and for any $L \geq 1$,*

$$\sum_{j=1}^{M} \big|\mathbf{P}_z[\mathbb{X}_{\mathbb{H}_A} = x_j] - p(x_j, A)\big| \leq 2\big(2^{-L} + \mathbf{P}_z[\mathbb{H}_A < L t_{\mathrm{mix}}]\big).$$

Moreover, if there exists $\ell \geq 1$ such that $d(x_a, x_b) > 2\ell + 1$ for $a \neq b$, then for all $L \geq 1$ and for all $1 \leq i \leq M$,

$$\sum_{j \neq i} \big|\mathbf{P}_{x_i}[\mathbb{X}_{\mathbb{H}_A} = x_j] - v_\ell(x_i) p(x_j, A)\big| \leq 2 v_\ell(x_i) \max_{z \in R(A,\ell)} \big\{2^{-L} + \mathbf{P}_z[\mathbb{H}_A < L t_{\mathrm{mix}}]\big\}.$$

PROOF. Fix a subset $A = \{x_1, \dots, x_M\}$ of V and $z \notin A$. By definition of the mixing time t_{mix} and by the definition of the total variation distance,

$$\begin{aligned}
&\sum_{j=1}^{M} \big|\mathbf{E}_z\big[\mathbf{P}_{\mathbb{X}(L t_{\mathrm{mix}})}[\mathbb{X}_{\mathbb{H}_A} = x_j]\big] - \mathbf{P}_\pi[\mathbb{X}_{\mathbb{H}_A} = x_j]\big| \\
&\quad = \sum_{j=1}^{M} \bigg| \sum_{w \in V} \{\mathbf{P}_z[\mathbb{X}(L t_{\mathrm{mix}}) = w] - \pi(w)\} \mathbf{P}_w[\mathbb{X}_{\mathbb{H}_A} = x_j] \bigg| \\
&\quad \leq 2 \big\|\mathbf{P}_z[\mathbb{X}_{L t_{\mathrm{mix}}} = \cdot] - \pi(\cdot)\big\|_{\mathrm{TV}} \leq 2 \cdot 2^{-L}.
\end{aligned}$$

To prove the first claim of the lemma, apply the Markov property to get that

$$\mathbf{P}_z[\mathbb{X}_{\mathbb{H}_A} = x_j] \leq \mathbf{E}_z\big[\mathbf{P}_{\mathbb{X}(L t_{\mathrm{mix}})}[\mathbb{X}_{\mathbb{H}_A} = x_j]\big] + \mathbf{P}_z[\mathbb{X}_{\mathbb{H}_A} = x_j, \mathbb{H}_A \leq L t_{\mathrm{mix}}]$$

and that

$$\begin{aligned}
&\mathbf{P}_z[\mathbb{X}_{\mathbb{H}_A} = x_j] \\
&\quad \geq \mathbf{P}_z[\mathbb{X}_{\mathbb{H}_A} = x_j, \mathbb{H}_A > L t_{\mathrm{mix}}] \\
&\quad = \mathbf{E}_z\big[\mathbf{P}_{\mathbb{X}(L t_{\mathrm{mix}})}[\mathbb{X}_{\mathbb{H}_A} = x_j]\big] - \mathbf{E}_z\big[\mathbf{P}_{\mathbb{X}(L t_{\mathrm{mix}})}[\mathbb{X}_{\mathbb{H}_A} = x_j], \mathbb{H}_A \leq L t_{\mathrm{mix}}\big].
\end{aligned}$$

The triangular inequality together with the previous two bounds and the estimate presented in the beginning of the proof show that

$$\sum_{j=1}^{M} \big|\mathbf{P}_z[\mathbb{X}_{\mathbb{H}_A} = x_j] - \mathbf{P}_\pi[X_{\mathbb{H}_A} = x_j]\big| \le 2\big(2^{-L} + \mathbf{P}_z[\mathbb{H}_A < Lt_{\mathrm{mix}}]\big).$$

This proves the first claim of the lemma.

We turn now to the proof of the second claim of the lemma. Since $d(x_i, A \setminus \{x_i\}) > \ell$ and $i \neq j$, the expression inside the absolute value on the left-hand side of the inequality can be written as

$$\mathbf{P}_{x_i}\big[\mathbb{X}(\mathbb{H}_A) = x_j | \mathbb{H}_{R(x_i,\ell)} < \mathbb{H}^+_{x_i}\big] v_\ell(x_i) - v_\ell(x_i) p(x_j, A).$$

The absolute value is thus bounded by

$$\sum_{z\in V} \big|\mathbf{P}_z\big[\mathbb{X}(\mathbb{H}_A) = x_j\big] - p(x_j, A)\big| \mathbf{P}_{x_i}\big[\mathbb{H}_{R(x_i,\ell)} < \mathbb{H}^+_{x_i}, \mathbb{X}(\mathbb{H}_{R(x_i,\ell)}) = z\big].$$

Since $d(x_a, x_b) > 2\ell + 1$, $a \neq b$, the set of vertices z at distance $\ell + 1$ from x_i is disjoint from A. Hence, by the first part of the proof, the sum over $j \neq i$ of this expression is bounded above by

$$2v_\ell(x_i) \max_{z\in R(A,\ell)} \big\{2^{-L} + \mathbf{P}_z[\mathbb{H}_A < Lt_{\mathrm{mix}}]\big\}$$

for every $L \ge 1$. This proves the lemma. □

Denote by $\mathcal{D}(f)$ the Dirichlet form of a function $f : V \to \mathbb{R}$:

$$\mathcal{D}(f) = \frac{1}{2} \sum_{x\in V} \sum_{y\sim x} \frac{\nu(x)}{\deg(x) W_x} \big(f(x) - f(y)\big)^2.$$

For disjoint subsets A and B of V, denote by $\mathrm{cap}(A, B)$ the capacity between A and B:

$$\mathrm{cap}(A, B) = \inf_f \mathcal{D}(f),$$

where the infimum is carried over all functions $f : V \to \mathbb{R}$ such that $f(x) = 1$ for $x \in A$, $f(y) = 0$, $y \in B$. Let $g : V \to [0, 1]$ be given by

$$g_{A,B}(x) = \mathbf{P}_x[H_A \le H_B].$$

It is a known fact that

$$\mathrm{cap}(A, B) = \mathcal{D}(g_{A,B}) = \sum_{y\in A} \nu(y) W_y^{-1} \mathbf{P}_y\big[H_B < H_A^+\big]. \tag{3.2}$$

Note that we may replace in the above identity H_B, H_A^+ by $\mathbb{H}_B$, $\mathbb{H}_A^+$, respectively.

Take a set $A \subset V$ composed of M points which are far apart and let x be a point in A. In the next lemma, we are going to estimate the probability $p(x, A) =$

$\mathbf{P}_\pi[\mathbb{X}_{\mathbb{H}_A} = x]$. This probability will be roughly proportional to $\deg(x)v_\ell(x)$. Let us first introduce a normalizing constant. For $\ell \ge 1$ and a finite subset A of V, let

$$\Gamma_\ell(A) = \sum_{x \in A} \deg(x) v_\ell(x).$$

LEMMA 3.2. *Fix a subset $A = \{x_1, \dots, x_M\}$ of V such that $d(x_a, x_b) > 2\ell + 1$, $a \ne b$, for some $\ell \ge 1$. Then*

$$\max_{1 \le i \le M} \left| p(x_i, A) - \frac{\deg(x_i) v_\ell(x_i)}{\Gamma_\ell(A)} \right| \le 2 \max_{z \in R(A,\ell)} \{2^{-L} + \mathbf{P}_z[\mathbb{H}_A \le L t_{\mathrm{mix}}]\}.$$

PROOF. Fix $1 \le i \le M$ and let $A_i = A \setminus \{x_i\}$. Since $\mathcal{D}(g_{\{x_i\},A_i}) = \mathcal{D}(1 - g_{\{x_i\},A_i})$, by (3.2)

$$\deg(x_i)\mathbf{P}_{x_i}[\mathbb{H}_{A_i} < \mathbb{H}^+_{x_i}] = \sum_{j \ne i} \deg(x_j) \mathbf{P}_{x_j}[\mathbb{H}_{x_i} < \mathbb{H}^+_{A_i}]. \tag{3.3}$$

On the other hand, since $d(x_i, A_i) > \ell$,

$$\begin{aligned}\mathbf{P}_{x_i}[\mathbb{H}_{A_i} < \mathbb{H}^+_{x_i}] &= \mathbf{E}_{x_i}[\mathbf{1}\{\mathbb{H}_{R(x_i,\ell)} < \mathbb{H}^+_{x_i}\}\mathbf{P}_{\mathbb{X}(\mathbb{H}_{R(x_i,\ell)})}[\mathbb{H}_{A_i} < \mathbb{H}_{x_i}]] \\ &= \mathbf{E}_{x_i}[\mathbf{1}\{\mathbb{H}_{R(x_i,\ell)} < \mathbb{H}^+_{x_i}\}(1 - \mathbf{P}_{\mathbb{X}(\mathbb{H}_{R(x_i,\ell)})}[\mathbb{X}_{\mathbb{H}_A} = x_i])].\end{aligned}$$

Therefore,

$$\begin{aligned}&\mathbf{P}_{x_i}[\mathbb{H}_{A_i} < \mathbb{H}^+_{x_i}] - v_\ell(x_i)[1 - p(x_i, A)] \\ &\quad = \mathbf{E}_{x_i}[\mathbf{1}\{\mathbb{H}_{R(x_i,\ell)} < \mathbb{H}^+_{x_i}\}\{p(x_i, A) - \mathbf{P}_{\mathbb{X}(\mathbb{H}_{R(x_i,\ell)})}[\mathbb{X}_{\mathbb{H}_A} = x_i]\}].\end{aligned}$$

Since $d(x_a, x_b) > 2\ell + 1$, we may replace in the previous expression $\mathbb{X}(\mathbb{H}_{R(x_i,\ell)})$ by $\mathbb{X}(\mathbb{H}_{R(A,\ell)})$. By the first assertion of Lemma 3.1, the absolute value of the difference inside braces is less than or equal to $2\max_{z \in R(A,\ell)}\{2^{-L} + \mathbf{P}_z[\mathbb{H}_A \le L t_{\mathrm{mix}}]\}$ for every $L \ge 1$. Hence,

$$\begin{aligned}&|\mathbf{P}_{x_i}[\mathbb{H}_{A_i} < \mathbb{H}^+_{x_i}] - v_\ell(x_i)[1 - p(x_i, A)]| \\ &\qquad \le 2 v_\ell(x_i) \max_{z \in R(A,\ell)} \{2^{-L} + \mathbf{P}_z[\mathbb{H}_A \le L t_{\mathrm{mix}}]\}\end{aligned} \tag{3.4}$$

for every $L \ge 1$.

Similarly, from (3.3) one obtains that

$$\begin{aligned}&\deg(x_i)\mathbf{P}_{x_i}[\mathbb{H}_{A_i} < \mathbb{H}^+_{x_i}] \\ &\quad = \sum_{j \ne i} \deg(x_j) \mathbf{E}_{x_j}[\mathbf{1}\{\mathbb{H}_{R(x_j,\ell)} < \mathbb{H}^+_A\}\mathbf{P}_{\mathbb{X}(R(x_j,\ell))}[\mathbb{X}_{\mathbb{H}_A} = x_i]].\end{aligned}$$

It follows from this identity and the previous argument that

$$\begin{aligned}&\Big|\deg(x_i)\mathbf{P}_{x_i}\big[\mathbb{H}_{A_i} < \mathbb{H}_{x_i}^+\big] - \sum_{j\neq i}\deg(x_j)v_\ell(x_j)p(x_i,A)\Big|\\ &\quad\le 2\sum_{j\neq i}\deg(x_j)v_\ell(x_j)\max_{z\in R(A,\ell)}\big\{2^{-L}+\mathbf{P}_z[\mathbb{H}_A\le Lt_{\mathrm{mix}}]\big\}\end{aligned}$$

for all $L \ge 1$.

The two previous estimates yield the bound

$$\begin{aligned}&\Big|\deg(x_i)v_\ell(x_i)\big[1-p(x_i,A)\big] - \sum_{j\neq i}\deg(x_j)v_\ell(x_j)p(x_i,A)\Big|\\ &\quad\le 2\sum_{j=1}^{M}\deg(x_j)v_\ell(x_j)\max_{z\in R(A,\ell)}\big\{2^{-L}+\mathbf{P}_z[\mathbb{H}_A\le Lt_{\mathrm{mix}}]\big\}.\end{aligned}$$

To conclude the proof of the lemma, it remains to divide both sides of the inequality by $\Gamma_\ell(A)$. □

4. Holding times of the trace process.

We present in this section a general result on Markov chains computing the time spent by this chain on a subset of the state space. This will be useful later in proving that the time spent by the walk on the shallow traps can be disregarded.

Consider an *irreducible* continuous-time Markov process $\{X_t : t \ge 0\}$ on a *finite* state space V. Denote by $\{W_x : x \in V\}$ the mean of the exponential waiting times, by ν the unique stationary probability measure, and by $\{\tau_j : j \ge 0\}$ the sequence of jump times.

Denote by $\mathbf{P}_x$, $x \in V$, the probability measure on the path space $D(\mathbb{R}_+, V)$ induced by the Markov process X_t starting from x. Expectation with respect to $\mathbf{P}_x$ is represented by $\mathbf{E}_x$. For a probability measure μ on V, let $\mathbf{P}_\mu = \sum_{x\in V}\mu(x)\mathbf{P}_x$.

Fix a set $A \subset V$ and let U be a stopping time such that for all $x \in A$,

$$\mathbf{P}_x[\tau_1 \le U] = 1, \qquad \mathbf{P}_x[H_{A\setminus\{x\}} \ge U] = 1, \qquad \mathbf{E}_x[U] < \infty.$$

$U = H_{R(A,\ell)}$ is the example to keep in mind, where ℓ is chosen so that $d(x, y) > 2\ell + 1$ for all $x \ne y \in A$. Let $S_A = U + H_A \circ \theta_U$ be the hitting time of the set A after time U. Denote by $v(x)$ the probability that starting from x the stopping time U occurs before the process returns to x: $v(x) = \mathbf{P}_x[U < H_x^+]$, which should be understood as an escape probability.

Let D_k, $k \ge 0$, be the time of the kth return to A after escaping: $D_0 = 0$, $D_1 = S_A$, $D_{k+1} = D_k + S_A \circ \theta_{D_k}$, $k \ge 1$. Clearly, if X_0 belongs to A, $\{X_{D_k} : k \ge 0\}$ is a discrete time Markov chain on A. On the other hand, by assumption $\mathbf{E}_x[D_1] = \mathbf{E}_x[U + H_A \circ \theta_U]$ is finite.

LEMMA 4.1. *The Markov chain* $\{X_{D_k} : k \ge 0\}$ *is irreducible. Moreover, for every* $f : V \to \mathbb{R}$,

$$\lim_{k\to\infty} \frac{1}{k} \int_0^{D_k} f(X_t)\, dt = \sum_{z\in A} \rho(z) \mathbf{E}_z \biggl[\int_0^{D_1} f(X_t)\, dt \biggr]$$

$\mathbf{P}_\nu$*-almost surely, where* ρ *is the unique stationary state of the discrete time chain* $\{X_{D_k} : k \ge 0\}$.

PROOF. We first prove the irreducibility of the chain $\{X_{D_k} : k \ge 0\}$. Fix x, $y \in A$ and consider a self-avoiding path $x_0 = x, \dots, x_n = y$ such that the discrete-time Markov chain associated to the Markov process X_t jumps from x_i to x_{i+1}, $0 \le i < n$, with positive probability. Such path exists by the irreducibility of X_t. Let x_j be the first state in the sequence $x_1, \dots, x_n$ which belongs to A. Since $\mathbf{P}_x[H_{A\setminus\{x\}} \ge U] = 1$,

$$\begin{aligned} \mathbf{P}_x[X_{D_1} = x_j] &\ge \mathbf{P}_x[X_{D_1} = x_j, Z_1 = x_1, \dots, Z_j = x_j] \\ &= \mathbf{P}_x[X_{U+H_A\circ\theta_U} = x_j, U \le H_{A\setminus\{x\}}, Z_1 = x_1, \dots, Z_j = x_j], \end{aligned}$$

where $\{Z_n : n \ge 0\}$ is the discrete-time jump chain associated to the process $\{X_t : t \ge 0\}$. Since $U \ge \tau_1$, on the event $\{Z_1 = x_1, \dots, Z_j = x_j\} \cap \{U \le H_{A\setminus\{x\}}\}$, $U + H_A \circ \theta_U = \tau_j$. The previous probability is thus equal to

$$\mathbf{P}_x[X_{\tau_j} = x_j, Z_1 = x_1, \dots, Z_j = x_j] = \mathbf{P}_x[Z_1 = x_1, \dots, Z_j = x_j] > 0.$$

Repeating this argument for the subsequent states in the sequence $x_1, \dots, x_n$ which belong to A, we prove that the chain X_{D_k} is irreducible.

Fix a function $f : V \to \mathbb{R}$. Clearly,

$$\frac{1}{k} \int_0^{D_k} f(X_t)\, dt = \frac{1}{k} \sum_{x\in A} \sum_{j=0}^{k-1} \int_{D_j}^{D_{j+1}} f(X_t)\, dt\, \mathbf{1}\{X_{D_j} = x\}.$$

For $x \in A$, let $K_1^x = \min\{j \ge 0 : X_{D_j} = x\}$, $K_{n+1}^x = \min\{j > K_n^x : X_{D_j} = x\}$, $n \ge 1$, and let $L_k^x = \#\{j < k : X_{D_j} = x\}$. With this notation, we can rewrite the previous sum as

$$\frac{1}{k} \sum_{x\in A} \sum_{n=1}^{L_k^x} \int_{D_{K_n^x}}^{D_{K_n^x+1}} f(X_t)\, dt = \sum_{x\in A} \frac{L_k^x}{k} \frac{1}{L_k^x} \sum_{n=1}^{L_k^x} \int_{D_{K_n^x}}^{D_{K_n^x+1}} f(X_t)\, dt.$$

By the irreducibility of the chain X_{D_k}, for each $x \in A$, L_k^x / k converges a.s. as $k \uparrow \infty$ to $\rho(x)$. Moreover, for each x, the variables $\int_{[D_{K_n^x}, D_{K_n^x+1})} f(X_t)\, dt$, $n \ge 1$, are independent and identically distributed. Hence, since $L_k^x \uparrow \infty$, by the law of large numbers, $\mathbf{P}_\nu$-almost surely,

$$\lim_{k\to\infty} \frac{1}{L_k^x} \sum_{n=1}^{L_k^x} \int_{D_{K_n^x}}^{D_{K_n^x+1}} f(X_t)\, dt = \mathbf{E}_x \biggl[\int_0^{D_1} f(X_t)\, dt \biggr].$$

The lemma follows from the two previous convergences. □

PROPOSITION 4.2. *The unique stationary state ρ of the discrete-time Markov chain $\{X_{D_k} : k \geq 0\}$ satisfies*

$$\rho(x) = \nu(x)v(x)W_x^{-1}\mathbf{E}_\rho[D_1] = \frac{\nu(x)v(x)W_x^{-1}}{\sum_y \nu(y)v(y)W_y^{-1}}. \tag{4.1}$$

Moreover, for every $g : V \to \mathbb{R}$,

$$\sum_{x\in A} v(x)\nu(x)W_x^{-1}\mathbf{E}_x\left[\int_0^{D_1} g(X_t)\,dt\right] = \sum_{x\in V} g(x)\nu(x). \tag{4.2}$$

PROOF. Applying Lemma 4.1 to $f = 1$, we obtain that $\mathbf{P}_\nu$-almost surely

$$\lim_{k\to\infty} \frac{D_k}{k} = \lim_{k\to\infty} \frac{1}{k}\int_0^{D_k} dt = \mathbf{E}_\rho[D_1]. \tag{4.3}$$

By Lemma 4.1 with $f(y) = \mathbf{1}\{y = x\}$, we get that $\mathbf{P}_\nu$-almost surely

$$\lim_{k\to\infty} \frac{1}{k}\int_0^{D_k} \mathbf{1}\{X_t = x\}\,dt = \rho(x)\mathbf{E}_x\left[\int_0^{D_1} \mathbf{1}\{X_t = x\}\,dt\right]$$

because starting from $y \neq x$, the process does not visit x before time D_1. In particular, all terms on the right-hand side in the statement of Lemma 4.1, but the one $z = x$, vanish. On the other hand, dividing and multiplying the expression on the left-hand side of the previous equation by D_k, we obtain by the ergodic theorem and by (4.3) that

$$\mathbf{E}_\rho[D_1]\nu(x) = \rho(x)\mathbf{E}_x\left[\int_0^{D_1} \mathbf{1}\{X_t = x\}\,dt\right]. \tag{4.4}$$

The time spent at x before D_1 is the time spent at x before U which is a geometric sum of independent exponential times. The success probability of the geometric is $v(x)$ and the mean of the exponential distributions is W_x. Hence, the right-hand side of the previous formula is equal to $\rho(x)W_x/v(x)$. This proves the first identity in (4.1). To derive the second identity, note that $\mathbf{E}_\rho[D_1]$ does not depend on x, and it is therefore only a normalizing constant to make ρ into a probability distribution.

By the ergodic theorem, for every $g : V \to \mathbb{R}$,

$$\lim_{k\to\infty} \frac{1}{D_k}\int_0^{D_k} g(X_t)\,dt = \sum_{x\in V} g(x)\nu(x).$$

To conclude the proof of the proposition, it remains to show that the left-hand side of this expression is equal to the left-hand side of (4.2). To this end, we will use the previous lemma.

For a function $g : V \to \mathbb{R}$, by Lemma 4.1 for $f = g$ and (4.3), we get

$$\lim_{k\to\infty} \frac{1}{D_k}\int_0^{D_k} g(X_t)\,dt = \lim_{k\to\infty} \frac{k}{D_k}\frac{1}{k}\int_0^{D_k} g(X_t)\,dt = \frac{1}{\mathbf{E}_\rho[D_1]}\mathbf{E}_\rho\left[\int_0^{D_1} g(X_t)\,dt\right].$$

To conclude the proof of the proposition, it suffices to use (4.1). □

COROLLARY 4.3. *We have that*

$$\mathbf{E}_\rho[D_1] = \frac{E_\rho[W_x/v(x)]}{1-\nu(V\setminus A)}.$$

Furthermore, for any function $g : V \to \mathbb{R}$,

$$\mathbf{E}_\rho\left[\int_0^{D_1} g(X_t)\,dt\right] = E_\nu[g]\mathbf{E}_\rho[D_1].$$

PROOF. We can write

$$\mathbf{E}_\rho[D_1] = \mathbf{E}_\rho\left[\int_0^{D_1} dt\right] = \mathbf{E}_\rho\left[\int_0^{D_1} \mathbf{1}\{X_t\in A\}\,dt\right] + \mathbf{E}_\rho\left[\int_0^{D_1} \mathbf{1}\{X_t\notin A\}\,dt\right].$$

By the same reasoning as below (4.4), we conclude that the first expectation in the sum above equals $E_\rho[W_x/v(x)]$. To evaluate the second expectation, we use Proposition 4.2 with $g = \mathbf{1}\{V\setminus A\}$ to conclude that

$$\mathbf{E}_\rho\left[\int_0^{D_1} \mathbf{1}\{X_t\notin A\}\,dt\right] = \mathbf{E}_\rho[D_1]\nu(V\setminus A).$$

Putting together the above equations, we conclude the proof of the first assertion of the corollary.

The second claim follows from the first identity in (4.1) and from (4.2). □

5. On the topology of convergence in measure. Fix $T > 0$ and let us denote by $\mathfrak{M}$ the space of measurable functions $f : [0, T] \to \mathbb{R}$. We consider the interval $[0, T]$ equipped with the Lebesgue measure, which will be denoted by Λ. As usual, we say two functions $f, g \in \mathfrak{M}$ are equal if they differ on a set of zero Lebesgue measure on $[0, T]$. Let $\mathcal{B}([0, T])$ denote the set of Borel subsets of $[0, T]$.

We introduce the following distance in $\mathfrak{M}$:

$$d_T(f,g) = \inf_{A\in\mathcal{B}([0,T])}\{\|f-g\|_{\infty,A^c} + \Lambda(A)\}, \tag{5.1}$$

where $\|f-g\|_{\infty,A^c}$ stands for the supremum of $f-g$ on the set A^c.

LEMMA 5.1. *The distance d_T metrizes the topology of convergence in measure in $\mathfrak{M}$. Moreover, the space $\mathfrak{M}$ is complete and separable under this distance.*

PROOF. Let us recall the definition of the *Ky Fan distance* in $\mathfrak{M}$ as

$$d_{\mathrm{KF}}(f,g) = \inf\{\varepsilon > 0; \Lambda(|f-g| > \varepsilon) \le \varepsilon\}.$$

It is well-known that the Ky Fan distance metrizes the topology of convergence in measure [14], and that the space $\mathfrak{M}$ is complete and separable under this metric. Therefore, it is enough to show that the distances d_T and d_{KF} are equivalent. First, we notice that we can assume that the sets A in the definition of d_T are of the form $\{|f-g|>\varepsilon\}$. In fact, if a set A is not of this form, let us write $\varepsilon = \|f - g\|_{\infty,A^c}$. We can take out the points of A such that $|f-g| \le \varepsilon$ without changing the supremum, and this lowers the value of $\Lambda(A)$. This procedure transforms the set A into $\{|f-g|>\varepsilon\}$. Therefore,

$$d_T(f,g) = \inf_{\varepsilon>0}\bigl\{\varepsilon + \Lambda\bigl(|f-g|>\varepsilon\bigr)\bigr\}, \tag{5.2}$$

which looks very close to the Ky Fan distance. Let us prove the aforementioned equivalence starting from (5.2). In one hand, if $d_{\mathrm{KF}}(f,g)=\varepsilon$ then there exists a sequence $\delta_n \downarrow 0$ such that

$$\Lambda\bigl(|f-g|>\varepsilon+\delta_n\bigr) \le \varepsilon + \delta_n.$$

Therefore,

$$d_T(f,g) \le \varepsilon + \delta_n + \Lambda\bigl(|f-g|>\varepsilon+\delta_n\bigr) \le 2(\varepsilon+\delta_n),$$

which shows that $d_T(f,g) \le 2 d_{\mathrm{KF}}(f,g)$. On the other hand, if $d_T(f,g) = a$ then there exist sequences $\delta_n \downarrow 0$ and $\varepsilon_n > 0$ such that

$$a + \delta_n = \varepsilon_n + \Lambda\bigl(|f-g|>\varepsilon_n\bigr).$$

In particular, $\varepsilon_n \le a + \delta_n$. Therefore,

$$\Lambda\bigl(|f-g|>a+\delta_n\bigr) \le \Lambda\bigl(|f-g|>\varepsilon_n\bigr) = a + \delta_n - \varepsilon_n \le a + \delta_n,$$

from where we conclude that $d_{\mathrm{KF}}(f,g) \le d_T(f,g)$. $\square$

Now we define the set of *locally constant functions* as a subset of the space $\mathfrak{M}$. Let $B(t,\delta)$ be the ball of radius δ centered at t. For any function $f:[0,T]\to\mathbb{R}$ and any point $t\in[0,T]$, we say that f is *locally constant* at t if there exists $\delta>0$ such that f is (Λ-almost surely) constant in $B(t,\delta)$. Define the set

$$\mathcal{C}(f) = \bigl\{t\in[0,T];\ f \text{ is locally constant in } t\bigr\},$$

and notice that $\mathcal{C}(f)$ is open. Let $\mathcal{D}(f)$ be the closed set $\mathcal{D}(f)=\mathcal{C}(f)^c$. Let $\mathfrak{M}_0$ be the set

$$\mathfrak{M}_0 := \bigl\{f\in\mathfrak{M};\ \Lambda\bigl(\mathcal{D}(f)\bigr)=0\bigr\}.$$

We call $\mathfrak{M}_0$ the set of *locally constant functions*. Let $f\in\mathfrak{M}_0$. Notice that the value of f in $\mathcal{D}(f)$ is not relevant, since $\Lambda(\mathcal{D}(f))=0$, and that the space of locally constant functions $\mathfrak{M}_0$ is not closed. In fact, the closure of $\mathfrak{M}_0$ is the whole space $\mathfrak{M}$.

Let $f\in\mathfrak{M}_0$. From the point of view of the topological properties of $\mathfrak{M}$, the values of f on $\mathcal{D}(f)$ are not relevant. However, since f is locally constant, it has a modification which is continuous Λ-a.e. Therefore, it makes sense to fix a representative of f. A simple way to do this is the following. We say that $x\in\mathcal{C}_0(f)$ if there exists $\delta>0$ such that $f(y)=f(x)$ Λ-a.e. in $B(x,\delta)$. We will write

$\mathcal{D}_0(f) = \mathcal{C}_0(f)^c$. Notice that $\Lambda(\mathcal{C}(f) \setminus \mathcal{C}_0(f)) = 0$. Now let $\tilde{f} : [0, T] \to \mathbb{R}$ be given by

$$\tilde{f}(t) = \frac{1}{2}\Big\{\liminf_{\substack{s\to t\\ s\in\mathcal{C}_0(f)}} f(s) + \limsup_{\substack{s\to t\\ s\in\mathcal{C}_0(f)}} f(s)\Big\}. \tag{5.3}$$

When $\liminf_{s\to t, s\in\mathcal{C}_0(f)} f(s) = -\infty$ and $\limsup_{s\to t, s\in\mathcal{C}_0(f)} f(s) = +\infty$, we set $\tilde{f}(t) = 0$. Clearly, $\tilde{f} = f$ on $\mathcal{C}_0(f)$ so that $\mathcal{D}_0(\tilde{f}) \subset \mathcal{D}_0(f)$, where inclusion may be strict.

LEMMA 5.2. *Fix $f, g \in \mathfrak{M}_0$. We have that*

$$\limsup_{\substack{s\to t\\ s\in\mathcal{C}_0(f)}} f(s) = \limsup_{\substack{s\to t\\ s\in\mathcal{C}_0(g)}} g(s)$$

whenever $f = g$ Λ-a.e., with a similar identity if we replace lim sup *by* lim inf. *In particular, $\tilde{f} = \tilde{g}$ if $f = g$ Λ-a.e. and equation* (5.3) *distinguishes a unique representative for each equivalence class of $\mathfrak{M}_0$.*

PROOF. Consider two functions f, g such that $f = g$ Λ-a.e. It is enough to show that

$$\limsup_{\substack{s\to t\\ s\in\mathcal{C}_0(f)}} f(s) \le \limsup_{\substack{s\to t\\ s\in\mathcal{C}_0(g)}} g(s) \quad \text{and} \quad \liminf_{\substack{s\to t\\ s\in\mathcal{C}_0(f)}} f(s) \ge \liminf_{\substack{s\to t\\ s\in\mathcal{C}_0(g)}} g(s).$$

We prove the first inequality, the derivation of the second one being similar. There exists a sequence $\{s_j : j \ge 1\}$ such that $s_j \in \mathcal{C}_0(f)$, $\lim_j s_j = t$,

$$\limsup_{\substack{s\to t\\ s\in\mathcal{C}_0(f)}} f(s) = \lim_{j\to\infty} f(s_j).$$

Since s_j belongs to $\mathcal{C}_0(f)$, f is Λ-a.e. constant in an interval $(s_j - \varepsilon, s_j + \varepsilon)$ and, therefore, in the interval $I_j = (s_j - \varepsilon, s_j + \varepsilon) \cap (s_j - (1/j), s_j + (1/j))$. Of course, $I_j \subset \mathcal{C}(f)$. As $\mathcal{D}_0(g)$ has Lebesgue measure 0, $\mathcal{C}_0(g) \cap I_j \neq \varnothing$. Take an element s'_j of this latter set. Since I_j is contained in $\mathcal{C}(f)$, s'_j belongs to $\mathcal{C}_0(f) \cap \mathcal{C}_0(g)$ so that $g(s'_j) = f(s'_j)$. Moreover, since f is Λ-a.e. constant in I_j and s_j, s'_j belong to I_j, $f(s_j) = f(s'_j)$. On the other hand, $\lim_j s'_j = t$ because s_j converges to t and $|s'_j - s_j| < (1/j)$. Hence,

$$\lim_{j\to\infty} f(s_j) = \lim_{j\to\infty} g(s'_j) \le \limsup_{\substack{s\to t\\ s\in\mathcal{C}_0(g)}} g(s),$$

which proves the lemma. $\square$

From now on when considering a function in $\mathfrak{M}_0$, we always refer to the representative defined by (5.3). For example, if we say that f is continuous at x, we actually mean that $\tilde{f}$ is continuous at x.

Let us introduce the following modulus of continuity in $\mathfrak{M}$. For a measurable function $f:[0,T]\to\mathbb{R}$ and $\delta>0$, let

$$\omega_\delta(f)=\Lambda\big(B\big(\mathcal{D}(f),\delta\big)\big).$$

The modulus of continuity $\omega_\delta(f)$ converges to 0 as $\delta\to 0$ if and only if f belongs to $\mathfrak{M}_0$. We extend this definition to the space $\mathfrak{M}$. Notice that $\mathcal{D}(\tilde{f})\subseteq\mathcal{D}(f)$. Therefore, the modulus of continuity of $\tilde{f}$ goes to 0 at least as fast as the modulus of continuity of f. Following the convention made above, when we write $\omega_\delta(f)$ we really mean $\omega_\delta(\tilde{f})$:

$$\omega_\delta(f)=\Lambda\big(B\big(\mathcal{D}(\tilde{f}),\delta\big)\big).$$

With this convention, Lemma 5.2 ensures that the modulus of continuity is well defined, that is, $\omega_\delta(f)=\omega_\delta(g)$ if $f=g$ Λ-a.e. The main motivation for the Introduction of the modulus of continuity $\omega_\delta(f)$ will be a comparison criterion between the topology in $\mathfrak{M}_0$ induced by d_T and the one induced by Skorohod's M_2 topology. We postpone the discussion of this criterion to Lemma 5.4, and we present here another motivation which we consider to be of independent interest.

PROPOSITION 5.3. *A subset $\mathcal{F}\subseteq\mathfrak{M}_0$ is sequentially precompact with respect to d_T if*

$$\sup_{f\in\mathcal{F}}\|f\|_\infty<\infty \quad\text{and}\quad \lim_{\delta\to 0}\sup_{f\in\mathcal{F}}\omega_\delta(f)=0. \tag{5.4}$$

PROOF. For $f\in\mathcal{F}$, define $\ell_f^\delta(t)=\operatorname{dist}(t,B(\mathcal{D}(\tilde{f}),\delta)^c)$. Since ℓ_f^δ is 1-Lipschitz for any $f\in\mathfrak{M}_0$ and any $\delta>0$, the family $\{\ell_f^\delta, f\in\mathcal{F}\}$ is equicontinuous. Fix a sequence f_n in $\mathcal{F}$ and a sequence $\{\delta_m:m\geq 1\}$ of positive numbers such that $\lim_m\delta_m=0$. Since $\sup_{f\in\mathcal{F}}\|f\|_\infty<\infty$, by a standard Cantor diagonal argument, we can extract a subsequence, still denoted by f_n, for which, as $n\uparrow\infty$, $\ell_{f_n}^{\delta_m}$ converges uniformly to some function ℓ^{δ_m} for every m, and for which $\tilde{f}_n(t)$ converges to some limit $F(t)$ for any rational t in $[0,T]$.

Let $\varepsilon_m=\limsup_{n\to\infty}\omega_{\delta_m}(f_n)$. By (5.4), $\lim_m\varepsilon_m=0$. Since $\ell_{f_n}^{\delta_m}$ converges uniformly to ℓ^{δ_m} and since $\{\ell_{f_n}^{\delta_m}\neq 0\}=B(\mathcal{D}(\tilde{f}_n),\delta_m)$,

$$\Lambda\big(\ell^{\delta_m}\neq 0\big)\leq\limsup_{n\to\infty}\Lambda\big(\ell_{f_n}^{\delta_m}\neq 0\big)=\limsup_{n\to\infty}\omega_{\delta_m}(f_n)=\varepsilon_m. \tag{5.5}$$

We claim that for every $t\in[0,T]$ such that $\ell^{\delta_m}(t)=0$ for some $m\geq 1$, there exist a neighborhood $N(t)$ of t and an integer $n_0\geq 1$ for which F is constant on

$N(t) \cap \mathbb{Q}$ and $\tilde{f}_n(t)$ is constant on $N(t)$ for $n \ge n_0$. We postpone the proof of this claim.

As $\lim_m \varepsilon_m = 0$, by (5.5) $\lim_m \Lambda(\ell^{\delta_m} \neq 0) = 0$. There exists therefore a subsequence $\{m(j) : j \ge 1\}$ such that $\sum_j \Lambda(\ell^{\delta_{m(j)}} \neq 0) < \infty$. Let $A = \bigcap_{k \ge 1} \bigcup_{j \ge k} \{\ell^{\delta_{m(j)}} \neq 0\}$ so that $\Lambda(A) = 0$. If t belongs to the set A^c, which has full measure, $\ell^{\delta_{m(j)}}(t) = 0$ for some j. By the conclusions of the previous paragraph, there exist a neighborhood $N(t)$ of t and an integer $n_0 \ge 1$ for which F is constant on $N(t) \cap \mathbb{Q}$ and $\tilde{f}_n(t)$ is constant on $N(t)$ for $n \ge n_0$.

In view of the previous result, we may define a function $\hat{F} : [0, T] \to \mathbb{R}$ which vanishes on the set A, and which on each element t of the set A^c is locally constant with value given by the value of F on a rational point close to t. In particular, $A^c \subset \mathcal{C}(\hat{F})$ which ensures that $\hat{F}$ belongs to $\mathfrak{M}_0$. Moreover, it follows from the convergence of $\tilde{f}_n$ to F on the rationals that $\tilde{f}_n(t)$ converges to $\hat{F}(t)$. Since set A has Lebesgue measure 0, f_n converges almost surely to $\hat{F}$. Therefore, by the Egoroff theorem, f_n converges to $\hat{F}$ with respect to the metric d_T.

To conclude the proof of the proposition, it remains to verify the assertion assumed in the beginning of the argument. Fix $t \in [0, T]$ and suppose that $\ell^{\delta_m}(t) = 0$ for some $m \ge 1$. In this case, since $\ell^{\delta_m}_{\tilde{f}_n}(t)$ converges to $\ell^{\delta_m}(t) = 0$, $\lim_n \operatorname{dist}(t, B(\mathcal{D}(\tilde{f}_n), \delta_m)^c) = 0$. Take a point t_n in the compact set $B(\mathcal{D}(\tilde{f}_n), \delta_m)^c$ realizing this distance to conclude that there exists a sequence t_n converging to t for which $\ell^{\delta_m}_{\tilde{f}_n}(t_n) = 0$. As $\ell^{\delta_m}_{\tilde{f}_n}(t_n) = 0$, $\tilde{f}_n$ is constant in the interval $(t_n - \delta_m, t_n + \delta_m)$. Therefore, functions $\tilde{f}_n$ are constant in a neighborhood $N(t)$ of t for n large enough. Since $\tilde{f}_n$ converges on the rationals to F, we conclude, as claimed, that F is constant in $N(t) \cap \mathbb{Q}$. □

Another topology which can be defined in the space $\mathfrak{M}_0$ corresponds to the projection of the Skorohod's M_2 topology, which is generated by the Hausdorff distance between the graphs of the functions. For $f, g \in \mathfrak{M}_0$, define the distance $d_T^{(2)}(f, g)$ by

$$d_T^{(2)}(f, g) := d_H(\Gamma_{\tilde{f}}, \Gamma_{\tilde{g}}),$$

where $\tilde{f}$, $\tilde{g}$ are the representatives of f, g defined in (5.3),

$$\Gamma_{\tilde{f}} = \bigcup_{t \in [0,T]} \{t\} \times \Big[\liminf_{s \to t} \tilde{f}(s), \limsup_{s \to t} \tilde{f}(s)\Big],$$

and d_H is the Hausdorff distance.

Recall the definition of the modulus of continuity $\omega_\delta(f)$ and note that $\omega_\delta(f) \ge 2\delta$ unless f is constant. Denote by $B(f; r)$, $B^{(2)}(f; r)$ the ball of center f and radius r with respect to the metric d_T, $d_T^{(2)}$, respectively.

LEMMA 5.4. *For any $f \in \mathfrak{M}_0$ and any $\delta > 0$,*

$$B^{(2)}(f;\delta) \subseteq B\big(f;\delta + \omega_{2\delta}(f)\big).$$

PROOF. Fix $f \in \mathfrak{M}_0$, $\delta > 0$ and $g \in B^{(2)}(f;\delta)$. By definition of d_T,

$$\begin{aligned} d_T(g,f) = d_T(\tilde{g},\tilde{f}) &\le \|\tilde{f} - \tilde{g}\|_{\infty, B(\mathcal{D}(\tilde{f}),2\delta)^c} + \Lambda\big(B\big(\mathcal{D}(\tilde{f}),2\delta\big)\big) \\ &= \|\tilde{f} - \tilde{g}\|_{\infty, B(\mathcal{D}(\tilde{f}),2\delta)^c} + \omega_{2\delta}(\tilde{f}). \end{aligned}$$

In order to evaluate the first term above, fix $t \notin B(\mathcal{D}(\tilde{f}),2\delta)$ so that $\tilde{f}$ is constant in $B(t,2\delta)$. In particular, $\Gamma_{\tilde{f}} \subset \Sigma = [0,t-2\delta]\times\mathbb{R} \cup [0,T]\times\{\tilde{f}(t)\} \cup [t+2\delta,T]\times\mathbb{R}$. Since $d_T^{(2)}(\tilde{g},\tilde{f}) = d_T^{(2)}(g,f) \le \delta$, by definition of the Hausdorff distance,

$$\delta \ge \operatorname{dist}\big(\big(t,\tilde{g}(t)\big),\Gamma_{\tilde{f}}\big) \ge \operatorname{dist}\big(\big(t,\tilde{g}(t)\big),\Sigma\big) = 2\delta \wedge \big|\tilde{f}(t) - \tilde{g}(t)\big|.$$

This implies that $|\tilde{f}(t) - \tilde{g}(t)| \le \delta$ for every $t \notin B(\mathcal{D}(\tilde{f}),2\delta)$, which finishes the proof of the lemma. □

Consider a sequence $\{Y_n : 1 \le n \le \infty\}$ of real-valued stochastic processes defined on some probability space $(\Omega,\mathcal{F},P)$. Assume that the trajectories of each Y_n, $1 \le n \le \infty$, belong to $\mathfrak{M}_0$ P-almost surely. This is the case, for instance, of continuous-time Markov chains taking values on a countable subset of $\mathbb{R}$.

THEOREM 5.5. *Fix $T > 0$. If $d_T^{(2)}(Y_n,Y_\infty)$ converges to 0 in probability as $n \uparrow \infty$, then $d_T(Y_n,Y_\infty)$ converges to 0 in probability as $n \uparrow \infty$.*

PROOF. It is enough to show that for each $\varepsilon > 0$, $\lim_{n\to\infty} P[d_T(Y_n,Y_\infty) > 2\varepsilon] = 0$. Fix $\delta < \varepsilon$ so that the previous probability is bounded by $P[d_T(Y_n,Y_\infty) > \varepsilon + \delta]$. This latter probability is in turn less than or equal to

$$P\big[d_T(Y_n,Y_\infty) > \varepsilon + \delta, \omega_{2\delta}(\tilde{Y}_\infty) \le \varepsilon\big] + P\big[\omega_{2\delta}(\tilde{Y}_\infty) > \varepsilon\big].$$

Since Y_∞ has trajectories in $\mathfrak{M}_0$ P-almost surely, the second term vanishes as $\delta \downarrow 0$. The first one is bounded by $P[d_T(Y_n,Y_\infty) > \delta + \omega_{2\delta}(\tilde{Y}_\infty)]$ which by the previous lemma is less than or equal to $P[d_T^{(2)}(Y_n,Y_\infty) > \delta]$. By assumption, this term vanishes as $n \uparrow \infty$. □

Assume that in the probability space $(\Omega,\mathcal{F},P)$ introduced before the statement of the previous theorem is also defined a sequence $\{X_n : 1 \le n < \infty\}$ of real-valued stochastic processes whose trajectories belong to $\mathfrak{M}_0$ P-almost surely.

COROLLARY 5.6. *Fix $T > 0$. If both $d_T(X_n,Y_n)$ and $d_T^{(2)}(Y_n,Y_\infty)$ converge to zero in probability as $n \uparrow \infty$, then $d_T(X_n,Y_\infty)$ also converges to zero in probability as $n \uparrow \infty$.*

REMARK 5.7. We would like to justify the Introduction of the topology of convergence in measure. In particular, we explain why we did not choose one of the Skorohod topologies which are canonically used to define convergence of càdlàg processes. For this, let us present some shortcomings of the Skorohod topologies in this context.

In [20], the authors introduce a compactification of $\bar{\mathbb{N}} = \{0, 1, \dots\} \cup \{\infty\}$, induced by the isometry $\phi : \bar{\mathbb{N}} \to \mathbb{R}$ which sends n to $1/n$ and ∞ to zero. The Skorohod's J_1 topology induced by this metric in $D(\mathbb{R}_+, \bar{\mathbb{N}})$ is used in [20] when developing a criterion for convergence toward the K-process. However, this choice is not convenient in the current context, as we explain below.

Consider a sequence of graphs in which the escape probabilities v_ℓ do not converge to one (e.g., the torus case in Proposition 9.2, or the Erdős–Rényi in Theorem 11.10). In such examples, the random walk will perform small excursions around a deep trap x before escaping from the ball $B(x, \ell)$. Due to the acceleration factor β_N, these excursions will last shorter and shorter times as we increase N and should be neglected in the scaling limit. However, this is not the case for any of the Skorohod topologies. For example, the sequence of functions $f_N(t) = 1_{\{1 \le t < 1+1/N\}}$ does not converge in any of the Skorohod topologies to $f(t) \equiv 0$.

There is a simple solution for the above problem, based on the fact that the excursions around x before escaping from $B(x, \ell)$ vanish in the supremum norm for the Euclidean metric of the torus. There is however a different shortcoming in this case. Consider for instance the discrete torus $\mathbb{T}_N^d$ embedded in the continuous torus $\mathbb{T}^d$. As we said above, this naturally introduces a metric on $\mathbb{T}_N^d$ for which the small excursions around a deep trap x do not pose any problems in the Skorohod's J_1 topology since they stay close to x in the supremum norm. In this case, the problem arises when an excursion exits the neighborhood $B(x, \ell_N)$. In this situation, the random walk typically performs a very short and "dense" excursion around the torus before finding the next deep trap to settle. Again, this phenomenon prevents convergence in any of the associated Skorohod topologies. Actually, not even the limiting process belongs to the Skorohod topology of $\mathbb{T}^d$ as its trajectories are not right continuous.

The topology of convergence in measure deals with these two obstructions, as it ignores what happens in vanishing time intervals. Due to its variational character, it turns out that our metric d_T is extremely well suited for computations, when compared with the equivalent Ky Fan metric d_{KF}.

6. Main result. We prove in this section that under certain assumptions the continuous time Markov process X_t^N, introduced in Section 2, is close, in an appropriate time scale and with respect to the topology introduced in Section 5, to a simple random walk Y_t^N which only visits the set A_N of the deepest traps and which has identically distributed jump probabilities: $p_N(x, y) = \rho_N(y)$, $x, y \in A_N$. For such result we need, roughly speaking, the set of deepest traps A_N

- to support most of the stationary measure ν,
- to consist of well-separated points,
- to be unlikely to be hit in a short time,
- to have comparable escape probabilities from each of its points.

The main result presented below holds in a more general context than the one described in Section 2. We suppose throughout this section that $\{G_N : N \geq 1\}$ is a sequence of finite, connected, vertex-weighted graphs, where $\{W_x^N : x \in V_N\}$ represents the positive weights. The vertices of V_N are enumerated in decreasing order of weights, $V_N = \{x_1^N, \dots, x_{|V_N|}^N\}$, $W_{x_j}^N \geq W_{x_{j+1}}^N$, $1 \leq j \leq |V_N| - 1$.

Denote by X_t^N the Markov process on V_N with generator given by (1.2). We do not assume that the depths W_x^N are chosen according to (1.1), but we impose some conditions presented below in (A.0)–(A.3).

We write in this section $J_N \uparrow \infty$ to represent a nondecreasing sequence of natural numbers $\{J_N : N \geq 1\}$ such that $\lim_{N\to\infty} J_N = \infty$. To keep notation simple, we sometimes omit the dependence on N of states, measures and sets.

Recall that $\nu = \nu_N$, defined in (2.1), is the stationary measure of the random walk X_t^N. Assume that $\nu(B_N^c)$ vanishes asymptotically for any sequence of subsets $B_N = \{x_1^N, \dots, x_{J_N}^N\} \subset V_N$ such that $J_N \uparrow \infty$:

$$\lim_{N\to\infty} \nu_N(B_N^c) = 0. \tag{A.0}$$

We now fix sequences $M_N \uparrow \infty$ and $\ell_N \uparrow \infty$ ($M_N \leq |V_N|$). The sequence M_N represents the number of deep traps selected, and ℓ_N a lower bound on the minimal distance among these deepest traps. We formulate three assumptions on these sequences. Let $A_N = \{x_1^N, \dots, x_{M_N}^N\}$ be the set of the deepest traps. We first require the deepest traps to be well separated:

$$d(x_i^N, x_j^N) > 2\ell_N + 1, \qquad 1 \leq i \neq j \leq M_N \tag{A.1}$$

for all N large enough. This condition, which is analogous to condition (B.1), ensures that any path $\{x_i^N = z_0, z_1, \dots, z_m = x_j^N\}$ from x_i^N to x_j^N has a state z_k which belongs to $R(A_N, \ell_N)$.

The second assumption is somehow related to (B.3) and requires, as explained below, the different escape probabilities v_x, $x \in A_N$, to have similar order of magnitude. For a subset B of V_N, let ν_B be the measure ν conditioned on B:

$$\nu_B(x) = \frac{W_x^N \deg(x)}{\sum_{y\in B} W_y^N \deg(y)}, \qquad x \in B.$$

Expectation with respect to ν_B is denoted by E_{ν_B}.

We suppose that there exists a sequence $\{\beta_N : N \geq 1\}$ such that for any sequence of subsets $B_N = \{x_1^N, \dots, x_{J_N}^N\} \subset A_N$ such that $|B_N| = J_N \uparrow \infty$

$$\limsup_{N\to\infty} E_{\nu_B}\left[\frac{W_x^N}{\beta_N v_\ell(x)}\right] < \infty, \qquad \limsup_{N\to\infty} \frac{1}{|B_N|} E_{\nu_B}\left[\frac{\beta_N v_\ell(x)}{W_x^N}\right] < \infty. \tag{A.2}$$

This hypothesis postulates essentially a law of large numbers for $\deg(x_j)v_\ell(x_j)$ and a bound for the sum of $(W^N_{x_j})^2 \deg(x_j)/v_\ell(x_j)$.

In analogy with (B.2), we will also assume that the hitting time of A_N is much larger than the mixing time of the discrete-time random walk on G_N. For $L \ge 1$ let

$$\kappa_N = \kappa(L, M_N, \ell_N) = \max_{x\in A_N} \max_{z \notin B(x,\ell_N)} \mathbf{P}^N_z[\mathbb{H}_x < L t^N_{\text{mix}}]. \tag{6.1}$$

Assume that for some sequence $L_N \uparrow \infty$,

$$\lim_{N\to\infty} M_N^3 \kappa_N = 0, \qquad \lim_{N\to\infty} M_N^2 2^{-L_N} = 0. \tag{A.3}$$

REMARK 6.1. Consider three sequences $M_N \uparrow \infty$, $\ell_N \uparrow \infty$ and $L_N \uparrow \infty$ satisfying (A.0)–(A.2) and such that

$$\lim_{N\to\infty} \kappa(L_N, M_N, \ell_N) = 0. \tag{6.2}$$

Then, there exists a sequence $M'_N \uparrow \infty$, $M'_N \le M_N$, for which the three sequence M'_N, ℓ_N, L_N satisfy (A.0)–(A.3).

Indeed, it follows from (6.2) and the fact that $L_N \uparrow \infty$ that there exists a sequence $K_N \uparrow \infty$ such that $\lim_{N\to\infty} K_N^2 2^{-L_N} = 0$, $\lim_{N\to\infty} K_N^3 \kappa(L_N, M_N, \ell_N) = 0$. Define a new sequence M'_N by $M'_N = \min\{M_N, K_N\}$ and define A'_N accordingly. Since $A'_N \subset A_N$ and $\kappa'_N \le \kappa_N$, (A.0)–(A.3) hold for the sequences M'_N, ℓ_N, L_N.

Hence, in applications, if one is able to prove (6.2), one can redefine the sequence M_N to obtain (A.3) which is the condition assumed in the main result of this section. Moreover, if a sequence M_N satisfies conditions (A.1), (A.2), (6.2), then any sequence $M'_N \uparrow \infty$ which increases to infinity with N at a slower pace than M_N, $M'_N \le M_N$, also satisfies these three conditions. The same observation holds for the sequence L_N. Hence, in the applications, both sequences shall increase very slowly to infinity, in a way that (A.3) is fulfilled, and all the problem rests on the identification of a convenient space scale ℓ_N, large for the process to mix before returning to a state, as required in condition (6.2), but not too large, to permit a good description of a ball of radius ℓ_N and a good estimate of the escape probability $v_\ell(x)$.

Let ρ_N be the probability measure on the set A_N given by

$$\rho_N(x_j) = \frac{\deg(x_j)v_\ell(x_j)}{\sum_{1\le i\le M_N} \deg(x_i)v_\ell(x_i)}, \tag{6.3}$$

where $v_\ell(x_j) = v^N_{\ell_N}(x_j)$ is the escape probability introduced in (3.1). By (2.1), ρ_N can also be written as

$$\rho_N(x_j) = \frac{\nu(x_j)v_\ell(x_j)W^{-1}_{x_j}}{\sum_{1\le i\le M_N} \nu(x_i)v_\ell(x_i)W^{-1}_{x_i}}, \tag{6.4}$$

which corresponds to (4.1) with $U = H_{R(A_N,\ell_N)}$.

For each $N \geq 1$, consider the continuous-time Markov process $\{Y_t^N : t \geq 0\}$ on A_N defined as follows. While at $x \in A_N$, the process waits a mean $W_x^N / v_\ell(x)$ exponential time at the end of which it jumps to $y \in A_N$ with probability $\rho_N(y)$. Note that the jump distribution is independent of the current state and that the process may jump to its current state since we did not impose y to be different from x. Moreover, the probability measure $\nu^N(x)/\nu^N(A_N)$ is the (reversible) stationary state of the Markov chain $\{Y_t^N : t \geq 0\}$.

We are now in a position to state the main result of this paper, from which we will deduce Theorems 2.1 and 2.2.

THEOREM 6.2. *Suppose that conditions* (A.0)–(A.3) *are in force. Then, for every* $N \geq 1$, *there exists a coupling* Q_N *between the stationary, continuous-time Markov chain* $\{Y_{\beta_N t}^N : t \geq 0\}$ *described above, and the Markov chain* $\{X_{\beta_N t}^N : t \geq 0\}$ *such that* $Q_N[X_0^N = Y_0^N = y] = \rho(y)$, $y \in A_N$, *and*

$$\lim_{N\to\infty} Q_N\big[d_T\big(X_{\beta_N \cdot}^N, Y_{\beta_N \cdot}^N\big) > \delta\big] = 0$$

for every $T \geq 0$ *and* $\delta > 0$, *where* d_T *stands for the distance introduced in* (5.1).

Theorem 6.2 follows from Lemmas 6.3, 6.4 and Proposition 6.5 below. Theorem 6.2 asserts that the process $X_{\beta_N t}^N$ is close to the process $Y_{\beta_N t}^N$ which jumps at rate $\beta_N v_\ell(x)/W_x^N$. If this latter expression is not of order one, the asymptotic behavior of $Y_{\beta_N t}^N$ will not be meaningful and our approximation of $X_{\beta_N t}^N$ by $Y_{\beta_N t}^N$ devoid of interest. Hence, in the applications we expect

$$\beta_N \approx \frac{W_{x_j}^N}{v_\ell(x_j)}.$$

LEMMA 6.3. *Assume that hypotheses* (A.0)–(A.3) *are in force. Then there exists a subset* $B_N = \{x_1^N, \dots, x_{J_N}^N\} \subset A_N$ *such that*

$$\lim_{N\to\infty} M_N\big(2^{-L_N} + M_N\kappa_N\big)\frac{\beta_N}{E_\rho[W_x^N/v_\ell(x)]} = 0, \tag{6.5}$$

$$\lim_{N\to\infty} \nu\big(B_N^c\big) = 0, \tag{6.6}$$

$$\limsup_{N\to\infty} \frac{\beta_N}{E_\rho[W_x^N/v_\ell(x)]}\nu\big(A_N^c\big)\rho(B_N) = 0. \tag{6.7}$$

PROOF. We start proving (6.5). By definition of the probability measure ρ_N this expression is equal to

$$M_N^2\big(2^{-L_N} + M_N\kappa_N\big)\frac{1}{M_N}E_{\nu_A}\left[\frac{\beta_N v_\ell(x)}{W_x^N}\right] \qquad \text{recall (6.4).}$$

This term vanishes as $N \uparrow \infty$ in view of (A.3) and (A.2) with $B_N = A_N$.

By (A.0), $\nu(A_N^c)$ vanishes as $N \uparrow \infty$. There exists, therefore, a sequence $K_N \uparrow \infty$ such that $\lim_{N\to\infty} K_N \nu(A_N^c) = 0$. Let $B_N = \{x_1^N, \dots, x_{J_N}^N\}$, where $J_N = \min\{M_N, K_N\}$ so that $|B_N|\nu(A_N^c) \to 0$. The second assertion of the lemma follows from assumption (A.0) because $J_N \uparrow \infty$. Moreover, as

$$\beta_N E_\rho\left[\frac{W_x^N}{v_\ell(x)}\right]^{-1} \rho(B_N) \le E_{\nu_B}\left[\frac{\beta_N v_\ell(x)}{W_x^N}\right],$$

by (A.2) and by definition of the set B_N, we have that

$$\limsup_{N\to\infty} \frac{\beta_N}{E_\rho[W_x^N/v_\ell(x)]} \nu(A_N^c)\rho(B_N) \le C_0 \limsup_{N\to\infty} |B_N|\nu(A_N^c) = 0 \tag{6.8}$$

for some finite constant. This concludes the proof of the lemma. □

LEMMA 6.4. *Assume that conditions* (A.2), (A.3), (6.5)–(6.7) *are in force. Then there exists a sequence* $\{K_N : N \ge 1\}$ *such that*

$$\lim_{N\to\infty} K_N M_N 2^{-L_N} = 0, \qquad \lim_{N\to\infty} K_N M_N^2 \kappa_N = 0, \tag{6.9}$$

$$\lim_{N\to\infty} \frac{K_N \nu(V_N \setminus A_N)}{\beta_N \nu(A_N)} E_\rho\left[\frac{W_x^N}{v_\ell(x)}\right] = 0, \tag{6.10}$$

$$\lim_{N\to\infty} \frac{K_N}{\beta_N} E_\rho\left[\frac{W_x^N}{v_\ell(x)} \mathbf{1}\{x \notin B_N\}\right] = 0, \tag{6.11}$$

$$\lim_{N\to\infty} \frac{K_N^2 \nu(V_N \setminus A_N)}{\beta_N \nu(A_N)} E_\rho\left[\frac{W_x^N}{v_\ell(x)}\right]\rho(B_N) = 0, \tag{6.12}$$

$$\lim_{N\to\infty} K_N \nu(V_N \setminus A_N)\rho(B_N) = 0, \tag{6.13}$$

$$\lim_{N\to\infty} \frac{K_N}{\beta_N} E_\rho\left[\frac{W_x^N}{v_\ell(x)}\right] = \infty. \tag{6.14}$$

PROOF. In view of (A.3), there exists a sequence $\psi_N \uparrow \infty$ such that $\psi_N M_N^2 2^{-L_N}$, $\psi_N M_N^3 \kappa_N$ vanish as $N \uparrow \infty$. We may choose this sequence ψ_N so that the limits in (6.5) and (6.6) still hold when multiplied by ψ_N, as well as the one in (6.7) when multiplied by ψ_N^2. Given this sequence ψ_N, let

$$K_N = \frac{\psi_N \beta_N}{E_\rho[W_x^N/v_\ell(x)]} = \psi_N \beta_N E_{\nu_A}\left[\frac{v_\ell(x)}{W_x^N}\right].$$

Conditions (6.9) follow the definition of ψ_N and from (6.5), while condition (6.10) follows from (6.6) since $B_N \subseteq A_N$. To verify (6.11), it is enough to remember that $\rho(x) W_x^N v_\ell(x)^{-1} = \nu(x)$ and to recall (6.6). Condition (6.12) follows from assumptions (6.7), (6.6) and the definition of K_N. Condition (6.13) follows from (6.7) and the definition of K_N. Finally, condition (6.14) requires ψ_N to diverge. □

PROPOSITION 6.5. *Suppose that conditions* (A.1), (A.2), (6.9)–(6.14) *are in force. Then, for every $N \geq 1$, there exists a coupling Q_N between the stationary, continuous-time Markov chain $\{Y^N_{\beta_N t} : t \geq 0\}$ on A_N with mean $W^N_x/\beta_N v_\ell(x)$ exponential waiting times and uniform jump probabilities $p_N(x,y) = \rho_N(y)$, $x, y \in A_N$, and the Markov chain $\{X^N_{\beta_N t} : t \geq 0\}$ such that $Q_N[X^N_0 = Y^N_0 = y] = \rho(y)$, $y \in A_N$, and*

$$\lim_{N\to\infty} Q_N\big[d_T\big(X^N_{\beta_N \cdot}, Y^N_{\beta_N \cdot}\big) > \delta\big] = 0$$

for every $T \geq 0$ and $\delta > 0$, where d_T stands for the distance introduced in (5.1).

PROOF. Recall the definition of the sequence of stopping times $\{D_k : k \geq 0\}$ introduced in Section 4 with $U = H_{R(A_N,\ell_N)}$. Since by (A.1) $R(A_N, \ell_N) \neq \varnothing$ and since the state space is finite and irreducible, $\mathbf{E}_x[U] < \infty$ for all $x \in A$. It also follows from assumption (A.1) that $\mathbf{P}_x[H_{A\setminus\{x\}} \geq U] = 1$ for all $x \in A$. Therefore, by Lemma 4.1 and Proposition 4.2, the discrete-time Markov chain $X^N_{D_k}$ is irreducible and its unique stationary state is the measure ρ defined in (6.3).

We start the construction of the measure Q_N by coupling the discrete skeleton of the chain Y^N_t with the chain $X^N_{D_k}$, and by coupling the waiting times of the chain Y^N_t with the times spent by X^N_t at each site of A_N. It follows from Lemma 3.2, which presents an estimate of the distance between the measure ρ and the measure $p(\cdot, A)$, from Lemma 3.1 and from the strong Markov property at time $H_{R(A,\ell)}$ that

$$\sup_{y\in A} \big\|\mathbf{P}_y\big[X^N_{D_1} = \cdot\big] - \rho(\cdot)\big\|_{\mathrm{TV}} \leq (M_N + 1)\big(2^{-L_N} + M_N\kappa_N\big) =: a_N. \tag{6.15}$$

Let $\sigma_0 = 0$ and denote by $\{\sigma_i : i \geq 1\}$ the jump times of the chain Y^N_t, including among these jumps the ones to the same site. We couple the initial state X^N_0 and Y^N_0 so that $Q_N[X^N_0 = Y^N_0] = 1$, $Q_N[X^N_0 = x] = \rho(x)$, $x \in A$. As $Y^N_{\sigma_1}$ is distributed according to ρ, by (6.15) we can couple $X^N_{D_1}$ and $Y^N_{\sigma_1}$ in a way that they coincide with probability at least $1 - a_N$. Moreover, conditioned on $X^N_{D_i} = x$, the number of visits of X^N_t to the point x between times D_i and D_{i+1} is a geometric random variable with success probability $v_\ell(x)$, so that

$$\int_{[D_i, D_{i+1})} \mathbf{1}\{X^N_t = x\}\, dt$$

is an exponential random variable with expectation $W_x/v_\ell(x)$. This is also the distribution of the time that Y^N_t spends in x. Proceeding by induction and using the strong Markov property at times D_i (for X_t^N) and σ_i (for Y_t^N), we obtain a coupling Q_N between X_t^N and Y_t^N such that

$$Q_N\left[\begin{array}{c} X_{D_i}^N = Y_{\sigma_i}^N, \int_{D_i}^{D_{i+1}} \mathbf{1}\{X_t^N = X_{D_i}^N\}\,dt = \sigma_{i+1} - \sigma_i \\ \text{for every } 0 \le i \le K_N \end{array}\right] \ge 1 - K_N a_N,$$

where K_N is the sequence introduced in Lemma 6.4. Denote the event appearing in the previous formula by $\mathcal{G}$. By (6.9),

$$\lim_{N\to\infty} Q_N[\mathcal{G}^c] = 0. \tag{6.16}$$

We claim that the coupling Q_N defined above satisfies the statement of the theorem. To estimate the distance between the processes X_t^N and Y_t^N, we introduce a third process $\bar{X}_t^N$ close to X_t^N in the distance d_T. Following [2], consider the process $\bar{X}_t^N$ defined by

$$\bar{X}_t^N = X^N\big(\sup\{s \le t : X_s^N \in A_N\}\big). \tag{6.17}$$

The (non-Markovian) process $\bar{X}_t^N$ indicates the last site in A_N visited by X_s^N before time t. We adopt for $\bar{X}_t^N$ the same convention agreed for the process Y_t^N and consider that the process $\bar{X}_t^N$ jumped from $y \in A_N$ to y at time t' if the process X_t^N being at y at time $s < t'$, reached $R(A_N, \ell)$ and then returned to y at time t' before hitting another site $z \in A_N \setminus \{y\}$. With this convention, the jump times of the process $\bar{X}_t^N$ are exactly the stopping times $\{D_i : i \ge 1\}$.

We assert that for every $T > 0$ and $\delta > 0$,

$$\lim_{N\to\infty} \mathbf{P}_\rho\big[d_T\big(\bar{X}_{\beta_N\cdot}^N, X_{\beta_N\cdot}^N\big) > \delta\big] = 0. \tag{6.18}$$

Fix $T > 0$ and $\delta > 0$. By definition of the process $\bar{X}^N$,

$$d_T\big(\bar{X}_{\beta_N\cdot}^N, X_{\beta_N\cdot}^N\big) \le \frac{1}{\beta_N}\int_0^{\beta_N T} \mathbf{1}\{X_t^N \notin A_N\}\,dt. \tag{6.19}$$

Therefore,

$$\mathbf{P}_\rho\big[d_T\big(\bar{X}_{\beta_N\cdot}^N, X_{\beta_N\cdot}^N\big) > \delta\big] \le \frac{1}{\beta_N\delta}\mathbf{E}_\rho\left[\int_0^{D_{K_N}} \mathbf{1}\{X_t^N \notin A_N\}\,dt\right] + \mathbf{P}_\rho[D_{K_N} \le \beta_N T].$$

Let us define

$$\Delta_N := \int_0^{D_{K_N}} \mathbf{1}\{X_t^N \notin A_N\}\,dt = D_{K_N} - \sigma_{K_N}.$$

This quantity will appear a couple of times in the computations below. By (6.22), $\mathbf{P}_\rho[D_{K_N} \le \beta_N T]$ vanishes as $N \uparrow \infty$ because $\sigma_{K_N} \le D_{K_N}$. On the other hand, by definition of the process $\bar{X}_t^N$ and by stationarity,

$$\frac{1}{\beta_N\delta}\mathbf{E}_\rho[\Delta_N] = \frac{K_N}{\beta_N\delta}\mathbf{E}_\rho\left[\int_0^{D_1} \mathbf{1}\{X_t^N \notin A_N\}\,dt\right].$$

By Corollary 4.3, the previous expression equals

$$\frac{K_N}{\beta_N \delta} \nu_N(V_N \setminus A_N) \mathbf{E}_\rho[D_1] = \frac{K_N \nu_N(V_N \setminus A_N)}{\beta_N \nu_N(A_N)\delta} E_\rho\left[\frac{W_x^N}{v_\ell(x)}\right]. \tag{6.20}$$

By (6.10), this expression vanishes as $N \uparrow \infty$. This proves (6.18).

Now we turn into the estimation of the distance between $\bar{X}_t^N$ and Y_t^N. On the event $\mathcal{G}$, the first K_N jumps of the processes $\bar{X}_t^N$ and Y_t^N are the same, and the process Y_t^N is always "ahead of" $\bar{X}_t^N$ in the sense that $\bar{X}_t^N$ spends more time at each site than Y_t^N. We need to show that the delay between $\bar{X}_t^N$ and Y_t^N is small. Let $B_N = \{x_1^N, \dots, x_{M'_N}^N\} \subseteq A_N$ be the set introduced in Lemma 6.3 and which satisfies conditions (6.11) and (6.12), and let $\mathfrak{N}_N$ be the number of times the process Y^N visits B_N before σ_{K_N}:

$$\mathfrak{N}_N := \#\{j < K_N : Y_{\sigma_j}^N \in B_N\}.$$

Denote by $\mathcal{G}_1$ the event $\mathcal{G} \cap \{\sigma_{K_N} \ge \beta_N T\}$. Since we have that $d_T(\bar{X}_{\beta_N \cdot}^N, Y_{\beta_N \cdot}^N) \le \beta_N^{-1} \int_0^{\beta_N T} \mathbf{1}\{\bar{X}_t^N \neq Y_t^N\}\, dt$, on the set $\mathcal{G}_1$, $d_T(\bar{X}_{\beta_N \cdot}^N, Y_{\beta_N \cdot}^N) \le \beta_N^{-1} \int_0^{\sigma_{K_N}} \mathbf{1}\{\bar{X}_t^N \neq Y_t^N\}\, dt$. Therefore, on the set $\mathcal{G}_1$,

$$\begin{aligned} & d_T(\bar{X}_{\beta_N \cdot}^N, Y_{\beta_N \cdot}^N) \\ & \quad \le \frac{1}{\beta_N} \sum_{j=1}^{K_N} \int_{\sigma_{j-1}}^{\sigma_j} \mathbf{1}\{Y_t^N \neq \bar{X}_t^N\}\, dt \\ & \quad \le \frac{1}{\beta_N} \int_0^{\sigma_{K_N}} \mathbf{1}\{Y_t^N \notin B_N\}\, dt + \frac{1}{\beta_N} \sum_{j=1}^{K_N} \int_{\sigma_{j-1}}^{\sigma_j} \mathbf{1}\{Y_t^N \in B_N, Y_t^N \neq \bar{X}_t^N\}\, dt. \end{aligned}$$

We claim that each integral in the second term of the previous sum is bounded by Δ_N. Indeed, the total delay of the process $\bar{X}_t^N$ with respect to the process Y_t^N in the interval $[0, \sigma_{K_N}]$ is $D_{K_N} - \sigma_{K_N} = \Delta_N$. On the other hand, either the length of time interval $[\sigma_{j-1}, \sigma_j]$ is bounded by Δ_N, in which case the claim is trivial, or the length is greater than Δ_N. In this latter situation, since the total delay between Y and $\bar{X}$ in the interval $[0, \sigma_{K_N}]$ is Δ_N, $D_{j-1} - \sigma_{j-1} \le \Delta_N$ for $1 \le j \le K_N$. Hence, in the interval $[\sigma_{j-1} + \Delta_N, \sigma_j)$ we have that $\bar{X}_t = Y_t$. This proves our assertion. In conclusion, if one recalls the definition of $\mathfrak{N}_N$, on the set $\mathcal{G}_1$,

$$d_T(\bar{X}_{\beta_N \cdot}^N, Y_{\beta_N \cdot}^N) \le \frac{1}{\beta_N} \int_0^{\sigma_{K_N}} \mathbf{1}\{Y_t^N \notin B_N\}\, dt + \frac{1}{\beta_N} \Delta_N \mathfrak{N}_N.$$

In conclusion,

$$\begin{aligned} & Q_N[d_T(\bar{X}_{\beta_N \cdot}^N, Y_{\beta_N \cdot}^N) > \delta] \\ & \quad \le Q_N[\mathcal{G}_1^c] + \frac{2}{\beta_N \delta} Q_N\left[\int_0^{\sigma_{K_N}} \mathbf{1}\{Y_t^N \notin B_N\}\, dt\right] \\ & \qquad + Q_N[\Delta_N \mathfrak{N}_N > (1/2)\delta\beta_N]. \end{aligned} \tag{6.21}$$

The first term vanishes as $N \uparrow \infty$ by (6.16) and (6.22). By Chebyshev and Cauchy–Schwarz inequalities, $P[ZW > \delta] = P[\sqrt{ZW} > \sqrt{\delta}] \le (\delta^{-1} E[Z]E[W])^{1/2}$ for any pair of nonnegative random variables Z, W. Therefore, the sum of the second and third terms is bounded by

$$\frac{2K_N}{\beta_N \delta} E_\rho\left[\frac{W_x^N}{v_\ell(x)} \mathbf{1}\{x \notin B_N\}\right] + \sqrt{\frac{2Q_N[\Delta_N] Q_N[\mathfrak{N}_N]}{\delta \beta_N}}.$$

Since $Q_N[\mathfrak{N}_N] = K_N \rho(B_N)$, by (6.20) this expression is less than or equal to

$$\frac{2K_N}{\beta_N \delta} E_\rho\left[\frac{W_x^N}{v_\ell(x)} \mathbf{1}\{x \notin B_N\}\right] + \sqrt{\frac{2K_N^2}{\beta_N \delta} \frac{\nu(V^N \setminus A_N)}{\nu(A_N)} E_\rho\left[\frac{W_x^N}{v_\ell(x)}\right] \rho(B_N)}.$$

By assumptions (6.11) and (6.12), this expression vanishes as $N \uparrow \infty$.

To conclude the proof of the theorem, it remains to show that

$$\lim_{N \to \infty} Q_N[\sigma_{K_N} \le \beta_N T] = 0. \tag{6.22}$$

For any random variable Z and any $T \ge 0$ such that $E[Z] \ge 2T$, by Chebyshev inequality we have that

$$P[Z < T] \le \frac{4\,\mathrm{Var}(Z)}{E[Z]^2}.$$

Note that

$$Q_N[\sigma_{K_N}] = K_N E_\rho\left[\frac{W_x^N}{v_\ell(x)}\right], \qquad \mathrm{Var}_{Q_N}(\sigma_{K_N}) \le 2K_N E_\rho\left[\left(\frac{W_x^N}{v_\ell(x)}\right)^2\right],$$

and that, by assumption (6.14), $K_N E_\rho[W_x^N / v_\ell(x)] \ge 2\beta_N T$ for N sufficiently large. By the previous elementary inequality,

$$\begin{aligned} Q_N[\sigma_{K_N} \le \beta_N T] &\le \frac{8 E_\rho[(W_x^N / v_\ell(x))^2]}{K_N E_\rho[W_x^N / v_\ell(x)]^2} \\ &\le \frac{8\beta_N}{K_N E_\rho[W_x^N / v_\ell(x)]} \frac{E_\rho[(W_x^N / v_\ell(x))^2]}{\beta_N E_\rho[W_x^N / v_\ell(x)]}. \end{aligned} \tag{6.23}$$

By assumption (6.14), the first term of this expression vanishes as $N \uparrow \infty$. The second one is equal to

$$E_{\nu_A}\left[\frac{W_x^N}{\beta_N v_\ell(x)}\right].$$

By (A.2) this expression is bounded uniformly in N. This concludes the proof of (6.22) and the one of Proposition 6.5. $\square$

Instead of starting from the stationary measure ρ_N, we may also start from any state x_i^N.

COROLLARY 6.6. *Assume that*

$$\liminf_{N} \nu_N(x_i^N) > 0 \qquad \textit{for every } i \ge 1. \tag{6.24}$$

Under the assumptions of Proposition 6.5, for every $i \ge 1$, $N \ge 1$, there exists a coupling $Q^\star_N$ between the stationary, continuous-time Markov chains $\{Y^N_{\beta_N t} : t \ge 0\}$ and $\{X^N_{\beta_N t} : t \ge 0\}$ such that $Q^\star_N[X^N_0 = Y^N_0 = x^N_i] = 1$, and

$$\lim_{N\to\infty} Q^\star_N[d_T(X^N_{\beta_N \cdot}, Y^N_{\beta_N \cdot}) > \delta] = 0$$

for every $T \ge 0$ and $\delta > 0$.

PROOF. The coupling is constructed as in Proposition 6.5, with the condition $Q^\star_N[X^N_0 = Y^N_0 = x^N_i] = 1$ replacing the analogous condition there. Consider the sequence K_N introduced in Lemma 6.4 and recall the definition of the set $\mathcal{G}$ introduced just before (6.16).

Since ν is the stationary state of the process X^N, for every $\delta > 0$,

$$\begin{aligned}
\mathbf{P}_{x_i^N}\bigg[\int_0^{T\beta_N} \mathbf{1}\{X^N_t \notin A_N\}\,dt > \delta\beta_N\bigg]
&\le \frac{1}{\beta_N \delta \nu(x_i^N)} \mathbf{E}_\nu\bigg[\int_0^{T\beta_N} \mathbf{1}\{X^N_t \notin A_N\}\,dt\bigg] \\
&\le \frac{T\nu(A_N^c)}{\delta\nu(x_i^N)}.
\end{aligned}$$

Hence, in view of (6.19), $\mathbf{P}_{x_i^N}[d_T(\bar{X}^N_{\beta_N \cdot}, X^N_{\beta_N \cdot}) > \delta]$ vanishes by (6.24) and assumption (A.0).

Let B_N be the set introduced in Lemma 6.3. Since ν_A is the stationary measure for the process Y_t, for the same reasons,

$$\mathbf{E}_{x_i^N}\bigg[\frac{1}{\beta_N}\int_0^{T\beta_N} \mathbf{1}\{Y^N_t \notin B_N\}\,dt\bigg] \le \frac{T\nu(B_N^c)}{\nu(x_i^N)}.$$

The distribution of the jump times $\{\sigma_j - \sigma_1 : j \ge 1\}$ of the process Y constructed in this corollary is the same as the distribution of the jump times $\{\sigma_j : j \ge 0\}$ of Proposition 6.5. In particular, by (6.23),

$$Q^\star_N[\sigma_{K_N+1} \le \beta_N T] \le Q_N[\sigma_{K_N} \le \beta_N T] \le \frac{8}{K_N E_\rho[W^N_x / v_\ell(x)]} E_{\nu_A}\bigg[\frac{W^N_x}{v_\ell(x)}\bigg].$$

Let

$$\Delta^\star_N = \int_0^{\beta_N T} \mathbf{1}\{X^N_t \notin A_N\}\,dt.$$

As in the proof of Proposition 6.5, on the set $\mathcal{G}_1^\star = \mathcal{G} \cap \{\sigma_{K_N+1} \geq \beta_N T\}$,

$$\begin{aligned}\int_0^{T\beta_N} \mathbf{1}\{\bar{X}_t^N \neq Y_t^N, Y_t^N \in B_N\}\,dt &= \sum_{j=0}^{K_N} \int_{\sigma_j \wedge \beta_N T}^{\sigma_{j+1}\wedge\beta_N T} \mathbf{1}\{\bar{X}_t^N \neq Y_t^N, Y_t^N \in B_N\}\,dt \\ &\leq \sum_{j=0}^{K_N} \mathbf{1}\{Y_{\sigma_j}^N \in B_N\}\Delta_N^\star = (1+\mathfrak{N}_N^\star)\Delta_N^\star,\end{aligned}$$

where $\mathfrak{N}_N^\star := \#\{1 \leq j < K_N : Y_{\sigma_j}^N \in B_N\}$ has the same distribution as $\mathfrak{N}_N$. Therefore, on the set $\mathcal{G}_1^\star$, $\beta_N d_T(\bar{X}_{\beta_N \cdot}, Y_{\beta_N \cdot}) \leq \int_0^{T\beta_N} \mathbf{1}\{Y_t^N \notin B_N\}\,dt + (1+\mathfrak{N}_N^\star)\Delta_N^\star$. In view of the argument below (6.21) and the previous estimates,

$$Q_N^\star[d_T(\bar{X}_{\beta_N\cdot}, Y_{\beta_N\cdot}) > \delta] \leq a_N K_N + \frac{2T\nu(B_N^c)}{\nu(x_i^N)} + \sqrt{\frac{2T\nu(A_N^c)}{\delta\nu(x_i^N)}[1+K_N\rho(B_N)]},$$

where a_N is given by (6.15). By (6.24), (6.13) and as in Proposition 6.5, this expression vanishes as $N \uparrow \infty$. □

The following remark will be important when proving Theorem 2.2.

REMARK 6.7. Assumption (A.0) has only been used in Lemma 6.3 to prove the existence of a sequence of subsets B_N satisfying (6.6), (6.7). In particular, Theorem 6.2 remains in force if hypothesis (A.0) is replaced by the existence of a sequence $I_N \leq M_N$, $I_N \uparrow \infty$, for which $B_N = \{x_1^N, \ldots, x_{I_N}^N\}$ satisfies (6.6) and such that

$$\lim_{N\to\infty} |B_N|\nu_N(A_N^c) = 0 \qquad \text{see (6.8).} \tag{6.25}$$

7. *K*-processes. We introduce in this section K-processes, a class of strong Markov processes on $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ with one fictitious state. We refer to [20] for historical remarks and to [31] for a detailed presentation and the proofs omitted here. The main result of this section presents sufficient conditions for the convergence of a sequence of finite-state Markov processes to a K-process.

Throughout this section, we fix two sequences of positive real numbers $\{u_k : k \in \mathbb{N}\}$ and $\{Z_k : k \in \mathbb{N}\}$. The first sequence represents the “entrance measure” and the second one the “hopping times” of the K-process. The only assumption we make over these sequences is that

$$\sum_{k\in\mathbb{N}} Z_k u_k < \infty. \tag{7.1}$$

However, the process will be more interesting in the case

$$\sum_{k\in\mathbb{N}} u_k = \infty. \tag{7.2}$$

If this sum is finite, the K-process associated to the sequences u_k and Z_k corresponds to a Markov process on $\mathbb{N}$ with no fictitious state.

Consider the set $\bar{\mathbb{N}}$ of nonnegative integers with an extra point denoted by ∞. We endow this set with the metric induced by the isometry $\phi : \bar{\mathbb{N}} \to \mathbb{R}$ which sends $n \in \bar{\mathbb{N}}$ to $1/n$ and ∞ to 0. This makes the set $\bar{\mathbb{N}}$ into a compact metric space. We use the notation $\mathrm{dist}(x, y) = |\phi(y) - \phi(x)|$ for this metric.

For each $k \in \mathbb{N}$, define independent Poisson process $\{N_t^k : t \geq 0\}$ with jump rate given by u_k. Denote by σ_i^k, $i \geq 1$, the time of the ith jump performed by the process N_t^k. Independently from the Poisson processes, let $\{T_0, T_i^k; k \in \mathbb{N}, i \geq 1\}$ be a collection of mean one independent exponential random variables.

Let $Z_\infty = 0$ and for $y \in \bar{\mathbb{N}}$ consider the process

$$\Gamma^y(t) = Z_y T_0 + \sum_{k \in \mathbb{N}} Z_k \sum_{i=1}^{N_t^k} T_i^k.$$

Define the K-process with parameter (Z_k, u_k), starting from y as follows:

$$X^y(t) = \begin{cases} y, & \text{if } 0 \leq t < Z_y T_0, \\ k, & \text{if } \Gamma^y(\sigma_i^k -) \leq t < \Gamma^y(\sigma_i^k) \text{ for some } i \geq 1, \text{ and} \\ \infty, & \text{otherwise.} \end{cases} \tag{7.3}$$

Note that $X^y(0) = y$ almost surely if $y \in \mathbb{N}$, and even in the case $y = \infty$ if (7.2) holds. We summarize in the next result the main properties of the process X_t^y. Its proof can be found in [31] or adapted from [20] where the case in which $u_k = 1$ for all $k \geq 1$ is examined. Recall that we denote by H_A the hitting time of a set A and that $Z_\infty = 0$.

THEOREM 7.1. *For any $y \in \bar{\mathbb{N}}$, the process $\{X^y(t) : t \geq 0\}$ is a strong Markov process on $\bar{\mathbb{N}}$ with right-continuous paths with left limits. Being at $k \in \mathbb{N}$, the process waits a mean Z_k exponential time at the end of which it jumps to ∞. For any finite subset A of $\mathbb{N}$, H_A is a.s. finite and*

$$\mathbf{P}[X^y(H_A) = j] = \frac{u_j}{\sum_{i \in A} u_i}, \qquad j \in A.$$

We investigate in this section the convergence of a sequence of Markov processes in finite state spaces toward the process $X^y(t)$. Let $\{M_N : N \geq 1\}$ be a sequence of integers such that $M_N \uparrow \infty$, and consider the sequences of positive real numbers

$$u_k^N, \qquad Z_k^N, \qquad 1 \leq k \leq M_N, N \geq 1. \tag{7.4}$$

In analogy with (7.3), we define processes $X_N^y(t)$ with "entrance measure" given by u_k^N and "hopping times" given by Z_k^N. For $N \geq 1$, let T_0^N, $T_i^{N,k}$, $N_t^{N,k}$ and

$\sigma_i^{N,k}$, $1 \le k \le M_N$, $i \ge 1$, be defined as above and write

$$\Gamma_N^y(t) = Z_y^N T_0^N + \sum_{k=1}^{M_N} Z_k^N \sum_{i=1}^{N_t^{N,k}} T_i^{N,k} \qquad \text{for } 1 \le y \le M_N$$

and

$$(7.5) \qquad X_N^y(t) = \begin{cases} y, & \text{if } 0 \le t < Z_y^N T_0^N, \\ k, & \text{if } \Gamma_N^y(\sigma_i^{N,k}-) \le t < \Gamma_N^y(\sigma_i^{N,k}) \text{ for some } i \ge 1. \end{cases}$$

One can easily see that the process X_N^y is a continuous-time càdlàg, Markov chain over $\{1, \dots, M_N\}$. The order in which the points $\{1, \dots, M_N\}$ are visited by X_N^y, after the starting position, is given by the order of the times $\sigma_i^{N,k}$. From this fact, we can conclude that the law of X_N^y is characterized by the following properties:

- The state space is $\{1, \dots, M_N\}$ and the process starts from y almost surely,
- The process X_N^y remains at any site k an exponential time with mean Z_k^N, after which it jumps to a site j with probability $u_j^N / \sum_{1 \le i \le M_N} u_i^N$.

REMARK 7.2. Note that the dynamics of the process X_N^y does not change if one replaces the vector $\{u_k^N : 1 \le k \le M_N\}$ by the vector $\{\gamma_N u_k^N : 1 \le k \le M_N\}$ for some $\gamma_N > 0$. In particular, when applying the theorem below we may multiply the sequence u_k^N by a constant γ_N to ensure the convergence of $\gamma_N u_k^N$ to u_k.

The main result of this section is stated below. Recall from [17], (5.2) the definition of the Skorohod's J_1 topology.

THEOREM 7.3. *Assume that for every $k \in \mathbb{N}$*

$$(7.6) \qquad \lim_{N \to \infty} (Z_k^N, u_k^N) = (Z_k, u_k)$$

and that

$$\lim_{m \to \infty} \limsup_{N \to \infty} \sum_{k=m}^{M_N} Z_k^N u_k^N = 0.$$

Then, for any given $y \in \mathbb{N}$, X_N^y converges weakly, as $N \uparrow \infty$, toward X^y in the Skorohod's J_1 topology.

PROOF. The proof is a modification of the one of Lemma 3.11 in [20]. We first couple the Poisson point processes used to define Γ_N^y and Γ^y. In some probability space $(\Omega, \mathcal{A}, \mathbf{Q})$, we construct a collection $\{N^k : k \in \mathbb{N}\}$ of Poisson point processes in $\mathbb{R}_+ \times \mathbb{R}_+$ with respect to the Lebesgue measure. Let $N^k(u, t)$ be the number of points falling in the rectangle $[0,t]\times[0,u]$. For fixed $k\in\mathbb{N}$ and $u\ge 0$, $N^k(u,\cdot)$ is distributed as a Poisson counting process with rate u. Define Γ^y and Γ^y_N as before, but using these coupled arrival processes, with corresponding intensities u_k and u^N_k. Moreover, we also use the same jump clocks $\{T^k_i : k\in\mathbb{N}, i\ge 1\}$ in their constructions.

Fix an integer $m\in\mathbb{N}$ and denote by $\{S^m_i : i\ge 1\}$ the arrival times of the process $N^1(u_1,\cdot)+N^2(u_2,\cdot)+\cdots+N^m(u_m,\cdot)$, with $S^m_0=0$. Fix $T\ge 0$ and let

$$L^m_T=\inf\{i\ge 1;\ \Gamma^y_N(S^m_i)\ge T \text{ for every } N\ge 1\}.$$

Since (Z^N_1,u^N_1) converges to (Z_1,u_1) and since $\Gamma^y_N(s)\ge\sum_{1\le i\le N^1(u^N_1,s)} Z^N_1 T^1_i$, by the law of large numbers the above infimum is finite.

Since the sequence $\{u_k : k\in\mathbb{N}\}$ is not summable, there exists a random integer m' large enough so that almost surely

$$\sum_{k=m+1}^{m'} Z_k \sum_{j=N^k(u_k,S^m_i)}^{N^k(u_k,S^m_{i+1}-)} T^k_j>0,\qquad i=0,\dots,L^m_T, \tag{7.7}$$

where $f(s-)$ stands for the left limit at s of a càdlàg function f.

Since u^N_k converges to u_k, almost surely there exists $N(m)$ such that

$$N^k(u^N_k,t)=N^k(u_k,t) \tag{7.8}$$

for all $1\le k\le m$, $0\le t\le S^m_{L^m_T}$ and all $N\ge N(m)$. By possibly increasing $N(m)$, we can also assume that

$$\inf_{N\ge N(m)}\sum_{k=m+1}^{m'} Z^N_k \sum_{j=N^k(u^N_k,S^m_j)}^{N^k(u^N_k,S^m_{j+1}-)} T^k_j>0,\qquad i=0,\dots,L^m_T. \tag{7.9}$$

It follows from (7.8) that the arrival times S^m_i are the same for the process X^y and X^y_N. Furthermore, by (7.7), (7.9), on each interval (S^m_i,S^m_{i+1}) there is at least one arrival of a Poisson process $N^k(u_k,\cdot)$ for some $k>m$ and one arrival for a Poisson process $N^k(u^N_k,\cdot)$ for some $k>m$. In particular, in the time interval $[\Gamma^y(S^m_i),\Gamma^y(S^m_{i+1}-))$ (resp., $[\Gamma^y_N(S^m_i),\Gamma^y_N(S^m_{i+1}-))$), $0\le i<L^m_T$, the process X^y (resp., X^y_N) performs an excursion in the set $\{1,\dots,m\}^c$, while on each time interval $[\Gamma^y(S_i-),\Gamma^y(S_i))$ (resp., $[\Gamma^y_N(S_i-),\Gamma^y_N(S_i))$), $1\le i\le L^m_T$, the processes X^y and X^y_N sit on the same site of $\{1,\dots,m\}$.

For $N\ge N(m)$, define the time changes $\lambda^m_N : [0,\Gamma^y_N(S^m_{L^m_T})]\to\mathbb{R}_+$ by

$$\lambda^m_N(t)=\frac{Z_y}{Z^N_y}t\qquad \text{for } 0\le t<Z^N_y T_0.$$

For $0 \le i \le L_T^m - 1$, let

$$\lambda_N^m(t) = \Gamma^y(S_i^m) + \frac{\Gamma^y(S_{i+1}^m-) - \Gamma^y(S_i^m)}{\Gamma_N^y(S_{i+1}^m-) - \Gamma_N^y(S_i^m)}[t - \Gamma_N^y(S_i^m)]$$

if $\Gamma_N^y(S_i^m) \le t \le \Gamma_N^y(S_{i+1}^m-)$ and let

$$\lambda_N^m(t) = \Gamma^y(S_{i+1}^m-) + \frac{\Gamma^y(S_{i+1}^m) - \Gamma^y(S_{i+1}^m-)}{\Gamma_N^y(S_{i+1}^m) - \Gamma_N^y(S_{i+1}^m-)}[t - \Gamma_N^y(S_{i+1}^m-)]$$

if $\Gamma_N^y(S_{i+1}^m-) \le t \le \Gamma_N^y(S_{i+1}^m)$.

In view of our previous discussion,

$$\text{dist}\big(X^y(\lambda_N^m(t)), X_N^y(t)\big) \le (1+m)^{-1} \qquad \text{for every } t \le T. \tag{7.10}$$

Indeed, whenever $X^y(\lambda_N^m(t))$ differs from $X_N^y(t)$, they are both above m, and the diameter of the set $\{m+1, m+2, \dots\}$ under $\text{dist}(\cdot,\cdot)$ is given by $(m+1)^{-1}$.

We claim that λ_N^m is close to the identity: for any $\delta > 0$,

$$\lim_m \limsup_N \mathbf{Q}\Big[\sup_{0\le t\le T} |\lambda_N^m(t) - t| > \delta\Big] = 0. \tag{7.11}$$

To prove this claim, fix $m \ge 1$ and note that

$$\sup_{0\le t\le T} |\lambda_N^m(t) - t| \le \max_{0\le i\le L_T^m} \{|\Gamma^y(S_i^m) - \Gamma_N^y(S_i^m)| \vee |\Gamma^y(S_i^m-) - \Gamma_N^y(S_i^m-)|\}.$$

By construction, the right-hand side is bounded above by

$$\begin{aligned} &|Z_y^N - Z_y|T_0 + \sum_{k=1}^m |Z_k^N - Z_k| \sum_{j=1}^{N^k(u_k^N, S_{L_T^m}^m)} T_j^k \\ &\quad + \sum_{k=m+1}^{\infty} Z_k \sum_{j=1}^{N^k(u_k, S_{L_T^m}^m)} T_j^k + \sum_{k=m+1}^{M_N} Z_k^N \sum_{j=1}^{N^k(u_k^N, S_{L_T^m}^m)} T_j^k. \end{aligned} \tag{7.12}$$

For each fixed m, the first two terms vanish almost surely as N goes to infinity. To estimate the other two terms, note that $L_T^m \ge L_T^{m+1}$, that $S_{L_T^{m+1}}^{m+1} \le S_{L_T^m}^m$ and that N^k, $\{T_j^k : j \ge 1\}$ are independent of $S_{L_T^m}^m$ for $k > m$. In particular, for $k > m$ and $u > 0$,

$$E_{\mathbf{Q}}\Bigg[\sum_{j=1}^{N^k(u, S_{L_T^m}^m)} T_j^k\Bigg] = u E_{\mathbf{Q}}[S_{L_T^m}^m] \le u E_{\mathbf{Q}}[S_{L_T^1}^1].$$

Last expectation is bounded because $S_{L_T^1}^1$ is defined through a Poisson process. Therefore, as $Z_k u_k$ is summable in k, the third term in (7.12), which does not depend on N, has finite expectation and converges to zero almost surely and in $L^1(\mathbf{Q})$ as m tends to infinity. Similarly,

$$E_{\mathbf{Q}}\left[\sum_{k=m+1}^{M_N} Z_k^N \sum_{j=1}^{N^k(u_k^N, S_{L_T^m}^m)} T_j^k\right] \le \sum_{k\ge m+1} Z_k^N u_k^N E_{\mathbf{Q}}[S_{L_T^1}^1].$$

By assumption, this expression vanishes as $N \uparrow \infty$ and then $m \uparrow \infty$. This proves that (7.11) holds in fact in $L^1(\mathbf{Q})$.

As a consequence of (7.11), one can extract a sequence m_N growing slowly enough such that

$$\sup_{0\le t\le T} \left|\lambda_N^{m_N} - t\right| \text{ converges to zero in probability as } N \uparrow \infty.$$

This, together with (7.10) provides the two conditions of Proposition 5.3(c) in [17]. Hence, X_N^y converges in probability to X^y in the Skorohod's J_1 topology as N tends to infinity. □

8. Scaling limit of trap models. In this section, we join the results of the last three sections to establish the asymptotic behavior of random walks on vertex-weighted graphs.

Throughout this section, we restrict our attention to weights given by an i.i.d. sequence of random variables in the basin of attraction of an α-stable distribution, as in (1.1). Let us first collect some consequences of this choice of random variables. In particular, we obtain the convergence of the environment to a limiting distribution.

Recall that $\alpha \in (0, 1)$ is the parameter of the stable distribution. Let λ be the measure on $\mathbb{R} \times (0, \infty)$ given by $\lambda = \alpha w^{-(1+\alpha)}\,dx\,dw$. Denote by $\{(z_i, \hat{w}_i) \in \mathbb{R} \times (0, \infty) : i \ge 1\}$ the marks of a Poisson point process of intensity λ independent of the sequence of graphs $\{G_N : N \ge 1\}$ and defined on a probability space $(\Omega', \mathcal{F}', P)$. Define the random measure ζ on $\mathbb{R}$ by

$$\zeta = \sum_{i\ge 1} \hat{w}_i \delta_{z_i}, \tag{8.1}$$

and let $\zeta_t = \zeta((0, t])$, $t \ge 0$, be the ζ-measure of the interval $(0, t]$. Let $F : [0, \infty) \to [0, \infty)$ be defined by

$$P[\zeta_1 > F(t)] = \mathbb{P}[W_x^N > t], \qquad t \ge 0.$$

The function F is nondecreasing and right continuous. Denote its right-continuous generalized inverse by F^{-1} and let

$$\hat{\tau}_i^N = F^{-1}(\mathbb{V}^{1/\alpha}[\zeta_{i/\mathbb{V}} - \zeta_{(i-1)/\mathbb{V}}]), \qquad 1 \le i \le \mathbb{V}. \tag{8.2}$$

Denote by τ_i^N, $1 \le i \le \mathbb{V}$, the sequence $\hat{\tau}_i^N$ in decreasing order: $\hat{\tau}_i^N = \tau_{\sigma(i)}^N$ for some permutation σ of $\{1, \dots, \mathbb{V}\}$ and $\tau_i^N \ge \tau_{i+1}^N$.

By [18], Proposition 3.1, $\{\hat{\tau}_i^N : 1 \le i \le \mathbb{V}\}$ has the same distribution as $\{W_x^N : x \in V_N\}$. Therefore, $(\tau_1^N, \dots, \tau_{\mathbb{V}}^N)$ has the same distribution as $(W_{x_1^N}^N, \dots, W_{x_{\mathbb{V}}^N}^N)$. Moreover, since $\mathbb{V}_N = |V_N| \to \infty$ $\mathbb{P}$-almost surely, the same result implies that $(\mathbb{P} \times P)$-almost surely,

$$\lim_{N\to\infty} \sum_{j\ge 1} |c_{\mathbb{V}} \tau_j^N - w_j| = 0, \tag{8.3}$$

where $\mathbf{W} = \{w_i : i \ge 1\}$ represents the weights in decreasing order of the measure ζ restricted to [0, 1]:

$$\begin{aligned} w_1 &= \max\{\hat{w}_i : z_i \in [0,1]\}, \\ w_{j+1} &= \max\{\hat{w}_i : z_i \in [0,1], \hat{w}_i \notin \{w_1, \dots, w_j\}\}, \qquad j \ge 1, \end{aligned} \tag{8.4}$$

and $\{c_k : k \ge 1\}$ is the sequence defined by (2.8).

Recall the definition of the function Ψ_N introduced just before the statement of Theorem 2.1.

THEOREM 8.1. *Let $G_N = (V_N, E_N)$ be a sequence of finite vertex-weighted graphs fulfilling assumptions* (A.0)–(A.2) *for some sequences M_N, ℓ_N. Assume, furthermore, that there exist sequences $L_N \uparrow \infty$, $\{\beta_N : N \ge 1\}$ and $\{\gamma_N : N \ge 1\}$ such that*

$$\lim_{N\to\infty} \kappa(L_N, M_N, \ell_N) = 0, \tag{8.5}$$

and such that

$$\lim_{N\to\infty} \left(\frac{W_{x_j}^N}{\beta_N v_\ell(x_j)}, \gamma_N v_\ell(x_j) \deg(x_j) \right) = (Z_j, u_j) \qquad \text{for all } j \ge 1, \tag{8.6}$$

$$\lim_{m\to\infty} \limsup_{N\to\infty} \sum_{j=m}^{M_N} \frac{W_{x_j}^N}{\beta_N v_\ell(x_j)} \gamma_N v_\ell(x_j) \deg(x_j) = 0. \tag{8.7}$$

Suppose, finally, that $\Psi_N(X_0^N)$ converges weakly to $k \in \mathbb{N}$. Then, for every $T > 0$, the Markov chain $\{\Psi_N(X_{\beta_N t}^N) : 0 \le t \le T\}$ converges to the K-process with parameters (Z_j, u_j) starting from k, in the topology introduced in Section 5.

PROOF. Repeating the arguments presented below (6.2), we obtain a new sequence M_N' for which (A.3) holds, as well as (8.6) with M_N' instead of M_N. Denote this new sequence by M_N. Under assumptions (A.0)–(A.3), Theorem 6.2 furnishes a coupling between the random walk $X_{\beta_N t}^N$ and a Markov process $Y_{\beta_N t}^N$ on $\{1, \dots, M_N\}$ whose d_T-distance converges to 0 in probability. In view of Remark 7.2 and by Theorem 7.3, under conditions (8.6), the Markov process $Y_{\beta_N t}^N$

converges to the K-process with parameters (Z_j, u_j) in the Skorohod's J_1 topology. By Skorohod's representation theorem, there exists a probability space in which this convergence take place almost surely. It remains to apply Corollary 5.6. □

In view of Remark 6.7, we may replace condition (A.0) by assumptions (6.6) and (6.25).

THEOREM 8.2. *Let $G_N = (V_N, E_N)$ be a sequence of finite vertex-weighted graphs fulfilling assumptions* (A.1)–(A.3) *for some sequences M_N, ℓ_N, L_N. Assume that there exists a sequence of subsets $B_N = \{x_1^N, \dots, x_{I_N}^N\}$, $I_N \le M_N$, $I_N \uparrow \infty$, satisfying* (6.6), (6.25). *Suppose, furthermore, that condition* (8.6) *is in force and that $\Psi_N(X_0^N)$ converges weakly to $k \in \mathbb{N}$. Then, for every $T > 0$, the Markov chain $\{\Psi_N(X_{\beta_N t}^N) : 0 \le t \le T\}$ converges to the K-process with parameters (Z_j, u_j) starting from k, in the topology introduced in Section* 5.

PROOF. By Remark 6.7, there exists a coupling between the random walk $X_{\beta_N t}^N$ and a Markov process $Y_{\beta_N t}^N$ on $\{1, \dots, M_N\}$ whose d_T-distance converges to 0 in probability. By Theorem 7.3, under conditions (8.6), the Markov process $Y_{\beta_N t}^N$ converges to the K-process with parameters (Z_j, u_j) in the Skorohod's J_1 topology. By Skorohod representation theorem, there exists a probability space in which this convergence take place almost surely. It remains to apply Corollary 5.6. □

9. Pseudo-transitive graphs. We prove in this section Theorem 2.1, inspired by Theorem 8.2, and we apply this result to some pseudo-transitive graphs. The assumptions (A.1)–(A.3), (6.6), (6.25), (8.6) simplify in this context because the degree and the escape probability from the deep traps do not depend on the specific vertex.

PROOF OF THEOREM 2.1. Fix an increasing sequence ℓ_N and a sequence of pseudo-transitive graphs G_N with respect to the sequence ℓ_N. We first derive some consequences of assumptions (B.0)–(B.2) and (2.7).

It follows from these hypotheses that there exists an increasing sequence $M_N \uparrow \infty$ such that

$$\lim_{N\to\infty} M_N^2 \mathbb{E}\Big[\frac{|B(\mathfrak{x}, 2\ell_N)|}{\mathbb{V}_N}\Big] = 0,$$

$$\lim_{N\to\infty} M_N \mathbb{P}\big[\big(\mathfrak{x}, B(\mathfrak{x}, \ell_N)\big) \not\equiv \big(\mathfrak{y}, B(\mathfrak{y}, \ell_N)\big)\big] = 0,$$

$$\lim_{N\to\infty} M_N^5 \mathbb{E}\Big[\sup_{y \notin B(\mathfrak{x},\ell_N)} \mathbf{P}_y[\mathbb{H}_{\mathfrak{x}} \le L_N t_{\mathrm{mix}}]\Big] = 0, \qquad \lim_{N\to\infty} M_N^3 2^{-L_N} = 0.$$

Let Σ_N^j, $1 \le j \le 3$ be the events

$$\Sigma_N^1 = \bigcap_{1 \le i \ne j \le M_N} \{B(x_i^N, \ell_N) \cap B(x_j^N, \ell_N) = \varnothing\},$$

$$\Sigma_N^2 = \bigcap_{j=1}^{M_N} \{(x_1^N, B(x_1^N, \ell_N)) \equiv (x_j^N, B(x_j^N, \ell_N))\},$$

$$\Sigma_N^3 = \Big\{ M_N^3 \max_{1 \le j \le M_N} \sup_{y \notin B(x_j^N, \ell_N)} \mathbf{P}_y[\mathbb{H}_{x_j^N} \le L_N t_{\mathrm{mix}}] \le M_N^{-1} \Big\}.$$

In the places where the vertices of the graph appear, as in the definition of the set Σ_N^1, the sequence M_N obtained above has to be replaced by $\min\{M_N, \mathbb{V}_N\}$, where $\mathbb{V}_N$ stands for the number of vertices of the random graph G_N. It is easy to see that all three events have probability asymptotically equal to one. We prove this assertion for Σ_N^1 and leave to the reader the proof for the other two. By definition, $\mathbb{P}[(\Sigma_N^1)^c]$ is bounded above by

$$\sum_{1 \le i \ne j \le M_N} \mathbb{P}[B(x_i^N, \ell_N) \cap B(x_j^N, \ell_N) \ne \varnothing] \le M_N^2 \mathbb{P}[B(x_1^N, \ell_N) \cap B(x_2^N, \ell_N) \ne \varnothing]$$

because $x_1^N, \dots, x_{\mathbb{V}}^N$ is uniformly distributed. By this same reason, conditioning on x_1^N, we obtain that the right-hand side is equal to

$$M_N^2 \mathbb{E}\Big[\frac{|B(x_1^N, 2\ell_N)| - 1}{\mathbb{V}_N - 1}\Big],$$

which vanishes as $N \uparrow \infty$ in view of the definition of the sequence M_N.

Let $A_N = \{x_1^N, \dots, x_{M_N}^N\}$. By hypothesis (B.0), $\nu_N(A_N^c)$ converges to 0 in $\mathbb{P}$-probability. In particular, there exists a deterministic sequence $I_N \uparrow \infty$, $I_N \le M_N$, such that $I_N \nu_N(A_N^c)$ converges to 0 in $\mathbb{P}$-probability. Let $B_N = \{x_1^N, \dots, x_{I_N}^N\}$. Since $I_N \uparrow \infty$, by hypothesis (B.0), $\nu_N(B_N^c)$ converges to 0 in $\mathbb{P}$-probability. Therefore, there exists a sequence $\varepsilon_N \downarrow 0$ for which

$$\lim_{N \to \infty} \mathbb{P}[\nu_N(B_N^c) + I_N \nu_N(A_N^c) \ge \varepsilon_N] = 0.$$

Let $\Sigma_N^4 = \{\nu_N(B_N^c) + I_N \nu_N(A_N^c) < \varepsilon_N\}$.

We turn now into the proof of the theorem which relies on Theorem 8.2. Recall the definition of the random weights $\hat{\tau}_j^N$, $1 \le j \le \mathbb{V}_N$, introduced at the beginning of Section 8. Since $\{\hat{\tau}_j^N : 1 \le j \le \mathbb{V}_N\}$ has the same distribution as $\{W_j^N : 1 \le j \le N\}$, we may replace the latter random weights by the former and assume that the random walk X_t^N evolves among random traps with depth τ_j^N instead of $W_{x_j}^N$.

To show that the pair $(c_{\mathbb{V}} \tau^N, \Psi_N(X_{t\beta_N}^N))$ converges weakly to $(\mathbf{w}, K_t)$, it is enough to show that any subsequence $\{N_j : j \ge 1\}$ possesses a sub-subsequence $\mathfrak{n}$ such that $(c_{\mathfrak{n}} \tau^{\mathfrak{n}}, \Psi_{\mathfrak{n}}(X_{t\beta_{\mathfrak{n}}}^{\mathfrak{n}}))$ converges to $(\mathbf{w}, K_t)$. Fix, therefore, a subsequence N_j.

By (8.3), the ordered sequence $(c_{N_j}\tau_1^{N_j},\dots,c_{N_j}\tau_{\mathbb{V}}^{N_j})$ converges almost surely in $L^1(\mathbb{N})$ to $\mathbf{w}=(w_1,w_2,\dots)$. This proves the weak convergence of the first coordinate. Let $\Sigma_{N_j}=\bigcap_{1\le k\le 4}\Sigma_{N_j}^k$. There exists a sub-subsequence, denoted by $\mathfrak{n}$, for which

$$\mathbb{P}\Big[\bigcup_{\mathfrak{n}_0\ge 1}\bigcap_{\mathfrak{n}\ge\mathfrak{n}_0}\Sigma_{\mathfrak{n}}\Big]=1.$$

We affirm that all assumptions of Theorem 8.2 hold on the set $\bigcup_{\mathfrak{n}_0\ge 1}\bigcap_{\mathfrak{n}\ge\mathfrak{n}_0}\Sigma_{\mathfrak{n}}$. Indeed, recall that $\beta_{\mathfrak{n}}^{-1}=c_{\mathfrak{n}}v_{\ell_{\mathfrak{n}}}^{\mathfrak{n}}(x_1^{\mathfrak{n}})$. Condition (A.1) follows from the definition of the set $\Sigma_{\mathfrak{n}}^1$. On the set $\Sigma_{\mathfrak{n}}^2$, the escape probabilities $v_\ell(x_j^{\mathfrak{n}})$ and the degrees $\deg(x_j^{\mathfrak{n}})$ are all the same for $1\le j\le M_{\mathfrak{n}}$. In particular, by definition of the sequence $\beta_{\mathfrak{n}}$, condition (A.2) becomes

$$\limsup_{\mathfrak{n}\to\infty}\frac{\sum_{j=1}^{J_{\mathfrak{n}}}c_{\mathfrak{n}}(\tau_j^{\mathfrak{n}})^2}{\sum_{j=1}^{J_{\mathfrak{n}}}\tau_j^{\mathfrak{n}}}<\infty,\qquad \limsup_{\mathfrak{n}\to\infty}\frac{1}{\sum_{j=1}^{J_{\mathfrak{n}}}c_{\mathfrak{n}}\tau_j^{\mathfrak{n}}}<\infty$$

for all sequences $J_{\mathfrak{n}}$ such that $J_{\mathfrak{n}}\le M_{\mathfrak{n}}$, $J_{\mathfrak{n}}\uparrow\infty$. Since the sequence $\tau_j^{\mathfrak{n}}$ is decreasing in j, the first ratio is bounded by $c_{\mathfrak{n}}\tau_1^{\mathfrak{n}}$, and these bounds are a consequence of (8.3). Condition (A.3) follows from the definition of the sequence M_N and from the definition of the set $\Sigma_{\mathfrak{n}}^3$. Conditions (6.6), (6.25) follow from the definition of the set $\Sigma_{\mathfrak{n}}^4$. Finally, on the set $\Sigma_{\mathfrak{n}}^2$, $v_\ell(x_j^{\mathfrak{n}})\deg(x_j^{\mathfrak{n}})$, $1\le j\le M_{\mathfrak{n}}$, is constant and the hypotheses (8.6) with $\gamma_{\mathfrak{n}}=[v_\ell(x_1^{\mathfrak{n}})\deg(x_1^{\mathfrak{n}})]^{-1}$ and $(Z_j,u_j)=(w_j,1)$ follow from (8.3). This proves the affirmation.

We may now apply Theorem 8.2 to conclude that the Markov chain $\Psi_{\mathfrak{n}}(X_{\beta_{\mathfrak{n}}t}^{\mathfrak{n}})$ converges to the K-process with parameters $(w_j,1)$ starting from k, in the topology introduced in Section 5. This concludes the proof of Theorem 2.1. □

We conclude this section with some examples of graphs satisfying the assumptions of Theorem 2.1.

9.1. *Hypercube.* We prove in this subsection the convergence of the trap model on the n-dimensional hypercube toward the K-process associated to constant entrance measure. This result has been established in [19] under the stronger Skorohod's J_1 topology with a different approach. Here, we give a proof as an application of Theorem 2.1.

Let $N=2^n$, $n\ge 1$, and let G_N be the n-dimensional hypercube $\{0,1\}^n$ with edges connecting any two points that differ by only one coordinate. By estimate (6.15) in [26], $t_{\mathrm{mix}}^N\ll n^2$.

PROPOSITION 9.1. *The assumptions of Theorem 2.1 are in force for the hypercube G_N with $\ell_N=\log_2(N)/10=n/10$.*

PROOF. Since the graph is transitive, condition (2.7) is satisfied and (B.0) follows from (8.3). To estimate the ratio in (B.1) note that $|B(0, 2\ell_N)|/\mathbb{V}_N$ is equal to the probability that the sum of n Bernoulli(1/2) independent random variables is less than or equal to $2\ell_N = n/5$. By the law of large numbers, this probability vanishes as $n \uparrow \infty$.

To show that (B.2) is in force, we could compare the distance $d(0, X_t)$ with an Ehrenfest's urn; see [26], Section 2.3, and proceed with a calculation based on a birth and death chain. For simplicity, we give instead a reference implying the result. By Lemmas 3.6(i) and 3.2(i) of [11], with $m(N) = N^2$ and $a = 1$, there exists a finite constant C_0 independent of n such that

$$\sup_{y \notin B(0,\ell_N)} \mathbf{P}_y[\mathbb{H}_0 \le n^2] \le C_0\left(n^2/N + \binom{n}{n/10}^{-1} n^{1/2}\log(n)\right)$$
$$\le C_0\big(n^2/N + (10)^{-n/10} n^{1/2}\log(n)\big),$$

which vanishes as $n \uparrow \infty$, proving (B.2). □

To complete the description of the asymptotic behavior of the trap model on the hypercube, it remains to determine the time scale β_N. By a computation based on a birth-and-death chain, the escape probability converges to 1 as $N \uparrow \infty$ and, therefore, $\lim_N \beta_N c_N = 1$.

9.2. *Discrete torus for $d \ge 2$.* In this subsection, the graph G_N stands for the d-dimensional discrete torus $\mathbb{T}_N^d = (\mathbb{Z}/N\mathbb{Z})^d$, $d \ge 2$, endowed with nearest neighbors edges. By [26], Theorem 5.5,

$$t^N_{\mathrm{mix}} \le C_0 N^2 \tag{9.1}$$

for some $C_0 = C_0(d)$. This constant may change from line to line, but will only depend on d.

We proved in [24] that in this context the trap model converges to the K-process. The next proposition shows that this result follows from Theorem 2.1.

PROPOSITION 9.2. *The assumptions of Theorem 2.1 are in force for the d-dimensional torus G_N with*

$$\ell_N = \begin{cases} N^{1/2}, & d \ge 3, \\ \dfrac{N}{\log^{1/4} N}, & d = 2, \end{cases} \qquad L_N = \begin{cases} \log^2 N, & d \ge 3, \\ \log^{1/4} N, & d = 2. \end{cases}$$

PROOF. Since the graph is transitive, condition (2.7) is satisfied and (B.0) follows from (8.3). On the other hand, assumption (B.1) is clearly in force by definition of ℓ_N. It remains to check hypothesis (B.2). Recall the definition of the sequence L_N. The case $d \ge 3$ follows directly from Lemma 3.1 of [33], and we focus on the case $d = 2$. Fix $x \in \mathbb{T}_N^d$ and $z \notin B(x, \ell_N)$. If Π stands for the canonical projection from $\mathbb{Z}^2$ to $\mathbb{T}_N^2$ and $\mathcal{P}_z$ for the probability corresponding to the symmetric nearest neighbor discrete time random walk on $\mathbb{Z}^2$,

$$\mathbf{P}_z[\mathbb{H}_x < L_N t_{\mathrm{mix}}^N] = \mathcal{P}_z[\mathbb{H}_{\Pi^{-1}(x)} < L_N t_{\mathrm{mix}}^N].$$

We may bound the previous probability by

$$\mathcal{P}_z[\mathbb{H}_{B(z, N\log^{1/4} N)^c} < L_N t_{\mathrm{mix}}^N] + \sum_i \mathcal{P}_z[\mathbb{H}_{x_i} < L_N t_{\mathrm{mix}}^N], \tag{9.2}$$

where the sum is performed over all sites x_i in the preimage of x which belong to the ball $B(z, N\log^{1/4} N)$.

The first term can be bounded using the estimate (9.1) for the mixing time and an exponential Doob inequality since each component of the random walk is a martingale. This argument shows that the first term is bounded by $4\exp\{-a\log^{1/4} N\}$ for some $a > 0$. Since there are no more than $C_0\sqrt{\log N}$ terms in the sum, the second expression in the previous decomposition is bounded above

$$C_0\sqrt{\log N}\mathcal{P}_0[\mathbb{H}_x < L_N t_{\mathrm{mix}}^N],$$

where x is a site at distance ℓ_N from the origin. Decomposing this probability according to whether the random walk reached the boundary of the ball with radius $N\log^{1/4} N$ before time $C_0 N^2 \log^{1/4} N$ or not, and recalling the argument employed to bound the first term in (9.2), we conclude that the previous expression is bounded by

$$C_0\sqrt{\log N} e^{-a\log^{1/4} N} + C_0\sqrt{\log N}\mathcal{P}_0[\mathbb{H}_x < \mathbb{H}_{B(0, N\log^{1/4} N)^c}]$$

for some finite constant C_0 and some positive a. By [25], Proposition 1.6.7, and the reversibility of the random walk, the second term is less than or equal to

$$C_0\sqrt{\log N}\left(1 - \frac{\log \ell_N}{\log(N\log^{1/4} N)} + \frac{C_0}{\log^2 N}\right) \le C_0 \log^{-1/4} N,$$

which proves condition (B.2). □

To complete the description of the asymptotic behavior of the trap model on the discrete torus $\mathbb{T}_N^d$, it remains to determine the time scale β_N. Let v_d, $d \ge 3$, be the escape probability of a simple random walk on $\mathbb{Z}^d$, and let

$$\beta'_N = \begin{cases} c_{|\mathbb{T}_N^d|}^{-1} (2/\pi)\log(N), & d = 2, \\ c_{|\mathbb{T}_N^d|}^{-1} v_d^{-1}, & d \ge 3. \end{cases}$$

In view of the definition of β_N and of [25], Theorem 1.6.6, $\lim_{N\to\infty} \beta_N/\beta'_N = 1$.

9.3. *Random d-regular graphs*. In this subsection, we consider a sequence of graphs G_N with N vertices satisfying the following three assumptions:

(G1) G_N is d-regular for some $d \geq 3$;
(G2) There is a constant $\alpha > 0$ such that for any vertex x of V_N, the ball $B(x, \alpha \log N)$ contains at most one cycle;
(G3) The spectral gap λ_N of the continuous time random walk on G_N is bounded below by some positive constant: $\lambda_N \geq \gamma > 0$ for all $N \geq 1$.

It follows from [12], Remark 1.4, that these three hypotheses hold, with probability approaching 1 as $N \uparrow \infty$, for a sequence of random d-regular graphs on N vertices. They are also satisfied by the so-called Lubotzky–Phillips–Sarnak graphs [28].

By [32] page 328, under conditions (G1) and (G3), the mixing time t^N_{mix} is bounded above by $C_0 \log N$ for some finite constant C_0.

PROPOSITION 9.3. *Let $\{G_N : N \geq 1\}$ be a sequence of random graphs defined on some probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *satisfying the assumptions* (G1)–(G3) *with a $\mathbb{P}$-probability converging to* 1 *as $N \uparrow \infty$. Then the conditions of Theorem* 2.1 *are fulfilled with $L_N = \log N$ and $\ell_N = \alpha' \log N$ for some $\alpha' < \min\{\alpha, [2\log(d-1)]^{-1}\}$, where α is the constant appearing in condition* (G2).

PROOF. Condition (B.0) follows from assumption (G1) and (8.3). The rest of the proof is based on estimates obtained in [12].

By [12], Lemma 6.1, with $\Delta = \ell_N$, the probability that a ball $B(x, \ell_N)$ is not a tree is bounded by $(d-1)^{-(\alpha-\alpha')\log N}$. Let Σ_N be the event

$$\Sigma_N = \{B(x^N_1, \ell_N) \text{ and } B(x^N_2, \ell_N) \text{ are disjoint trees}\}. \tag{9.3}$$

We claim that $\mathbb{P}[\Sigma_N]$ converges to 1 as $N \uparrow \infty$. Indeed, if $\tilde{\Sigma}_N$ stands for the event that $B(x^N_1, \ell_N)$, $B(x^N_2, \ell_N)$ are trees, in view of the estimate of the previous paragraph, $\mathbb{P}[\tilde{\Sigma}^c_N]$ is bounded by $2(d-1)^{-(\alpha-\alpha')\log N}$ which vanishes as $N \uparrow \infty$. On the other hand, since $|B(x_1, r)| \leq 4(d-1)^r$ for any ball in a d-regular graph and since x^N_1, x^N_2 are uniformly distributed,

$$\mathbb{P}[B(x_1, \ell_N) \cap B(x_2, \ell_N) \neq \varnothing] \leq 4\frac{(d-1)^{2\ell_N}}{N}.$$

As $\alpha' < [2\log(d-1)]^{-1}$, this expression vanishes as $N \uparrow \infty$. This proves the claim and assumption (2.7), which clearly follows from the claim. Condition (B.1) is also in force because $|B(x_1, 2\ell_N)| \leq 4(d-1)^{2\ell_N}$.

It remains to examine the escape probability appearing in condition (B.2). It follows from the bound for the mixing time presented just before the statement of the proposition and from our choice of the sequence L_N that

$$\mathbf{P}_z[\mathbb{H}_x < L_N t^N_{\mathrm{mix}}] \leq \mathbf{P}_z[\mathbb{H}_x < C_0(\log N)^2].$$

By [12], Lemma 3.3, with $r = 0$ and $s = \alpha' \log N$, the previous expression for $z \notin B(x, \ell_N)$, is bounded by $C_0 N^{-a}$ for some finite constant C_0 and some positive $a > 0$. This concludes the proof of the proposition. □

We conclude this section computing the scaling factor β_N in the context of graphs satisfying assumptions (G1)–(G3). On the event (9.3), which has asymptotic probability equal to one, $B(x_1, \ell_N)$ is a d-regular tree so that

$$v_{\ell_N}(x_1) = \frac{d-2}{d-1}\left(\frac{1}{1-(d-1)^{-\ell_N}}\right).$$

In particular, $\lim_{N\to\infty} \beta_N c_N = (d-1)/(d-2)$.

10. Graphs with asymptotically random conductances. We prove in this section Theorem 2.2. The proof follows the one of Theorem 2.1. However, the absence of regularity of the graph requires some extra effort in establishing (A.2).

Recall the coupling $\mathcal{Q}_N$ defined in (B.3) between the random graph G_N and the sequence of i.i.d. random vectors $\{(D_j, E_j) : j \geq 1\}$. We extend this coupling $\mathcal{Q}_N$ to a coupling $\mathcal{Q}$ between all random graphs G_N and the sequence of i.i.d. random vectors $\{(D_j, E_j) : j \geq 1\}$ using $\mathcal{Q}_N$ as the conditional probability:

$$\mathcal{Q}\big[G_N = G \big| \{(D_j, E_j) : j \geq 1\}\big] = \mathcal{Q}_N\big[G_N = G \big| \{(D_j, E_j) : j \geq 1\}\big],$$

with the further condition that the graphs G_N, $N \geq 1$, are conditionally independent, given $\{(D_j, E_j) : j \geq 1\}$. Include in the probability space just defined the random measure ζ introduced in (8.1) which is associated to the marks of a Poisson point process independent from the variables (D_j, E_j) and from the random graphs G_N. The probability measure on this new space is still denoted by $\mathcal{Q}$.

Recall the definition of the random weights $\hat{\tau}_j^N$, $1 \leq j \leq \mathbb{V}_N$, introduced in Section 2. Since $\{\hat{\tau}_j^N : 1 \leq j \leq \mathbb{V}_N\}$ has the same distribution as $\{W_j^N : 1 \leq j \leq |V_N|\}$, we may replace the latter random weights by the former and assume that the random walk X_t^N evolves among random traps with depth τ_j^N instead of $W_{x_j}^N$.

Since w_j is a.s. summable, since by (B.3) D_1/E_1 has finite $\mathcal{Q}$-expectation and since the sequences $\{w_j\}$ and $\{(D_j, E_j)\}$ are independent,

$$\sum_{j\geq 1} w_j \frac{D_j}{E_j} \quad \text{is } \mathcal{Q}\text{-almost surely finite.} \tag{10.1}$$

By the strong law of large numbers, almost surely

$$\frac{1}{n}\sum_{j=1}^{n} D_j/E_j \leq C_1 \tag{10.2}$$

for all large enough n, where $C_1 = 2E_{\mathcal{Q}}[D_1/E_1]$.

By hypotheses (B.1)–(B.3), there exists an increasing sequence $M_N \uparrow \infty$ such that

$$\lim_{N\to\infty} M_N^2 E_{\mathbb{Q}}\Big[\frac{|B(\mathfrak{x}, 2\ell_N)|}{\mathbb{V}_N}\Big] = 0, \qquad \lim_{N\to\infty} M_N^3 2^{-L_N} = 0,$$

$$\lim_{N\to\infty} \mathbb{Q}\Big[\max_{1\le j\le M_N} \big|[v_\ell(\mathfrak{x}_j)]^{-1} - E_j^{-1}\big| > M_N^{-2}\Big] = 0,$$

$$\lim_{N\to\infty} \mathbb{Q}\Big[\bigcup_{j=1}^{M_N}\{\deg(\mathfrak{x}_j) \neq D_j\}\Big] = 0,$$

$$\lim_{N\to\infty} M_N^5 E_{\mathbb{Q}}\Big[\sup_{y\notin B(\mathfrak{x},\ell_N)} \mathbf{P}_y[\mathbb{H}_{\mathfrak{x}} \le L_N t_{\mathrm{mix}}]\Big] = 0.$$

As before, in the places where the vertices of the graph appear, as in the definition of the set Σ_N^1, the sequence M_N obtained above has to be replaced by $\min\{M_N, \mathbb{V}_N\}$, where $\mathbb{V}_N$ stands for the number of vertices of the random graph G_N.

Let Σ_N^j, $1 \le j \le 4$ be the events

$$\Sigma_N^1 = \bigcap_{1\le i\neq j\le M_N} \{B(x_i^N, \ell_N) \cap B(x_j^N, \ell_N) = \varnothing\},$$

$$\Sigma_N^2 = \Big\{\max_{1\le j\le M_N} \big|[v_\ell(x_j^N)]^{-1} - E_j^{-1}\big| \le M_N^{-2}\Big\},$$

$$\Sigma_N^3 = \bigcap_{j=1}^{M_N} \{\deg(x_j^N) = D_j\},$$

$$\Sigma_N^4 = \Big\{M_N^3 \max_{1\le j\le M_N} \sup_{y\notin B(x_j^N,\ell_N)} \mathbf{P}_y[\mathbb{H}_{x_j^N} \le L_N t_{\mathrm{mix}}] \le M_N^{-1}\Big\}.$$

Similarly to what was done in the proof of Theorem 2.1, we can show that these events have probability asymptotically equal to one.

By (8.3), we may replace the sequence M_N by a possibly random increasing sequence $M_N' \le \min\{M_N, \mathbb{V}_N\}$, $M_N' \uparrow \infty$ $\mathbb{Q}$-a.s., still denoted by M_N, for which all the previous estimates hold and such that for all $N \ge 1$,

$$\sum_{j\ge 1} |c_{\mathbb{V}}\tau_j^N - w_j| \le M_N^{-2}. \tag{10.3}$$

By hypothesis (B.0), even though the sequence M_N is random, the expectation $\mathbb{E}[\nu_N(\{x_1^N, \dots, x_{M_N}^N\}^c)]$ vanishes as $N \uparrow \infty$. Let $A_N = \{x_1^N, \dots, x_{M_N}^N\}$. As in the proof of Theorem 2.1, presented in the previous sections, using again hypothesis (B.0) we construct a set $B_N = \{x_1^N, \dots, x_{I_N}^N\}$, $|B_N| = I_N$, and a sequence $\varepsilon_N \downarrow 0$ for which

$$\lim_{N\to\infty} \mathbb{Q}\big[\nu_N(B_N^c) + I_N \nu_N(A_N^c) \ge \varepsilon_N\big] = 0.$$

Let $\Sigma_N^5 = \{\nu_N(B_N^c) + I_N \nu_N(A_N^c) \le \varepsilon_N\}$.

To show that the pair $(c_{\mathbb{V}} \tau^N, \Psi_N(X_{t\beta_N}^N))$ converges weakly to $(\mathbf{w}, K_t)$, it is enough to show that any subsequence $\{N_j : j \ge 1\}$ possesses a sub-subsequence $\mathfrak{n}$ such that $(c_{\mathfrak{n}} \tau^{\mathfrak{n}}, \Psi_{\mathfrak{n}}(X_{t\beta_{\mathfrak{n}}}^{\mathfrak{n}}))$ converges to $(\mathbf{w}, K_t)$. Fix, therefore, a subsequence N_j. By (8.3), the ordered sequence $(c_{N_j} \tau_1^{N_j}, \dots, c_{N_j} \tau_{\mathbb{V}}^{N_j})$ converges almost surely in $L^1(\mathbb{N})$ to $\mathbf{w} = (w_1, w_2, \dots)$. This proves the weak convergence of the first coordinate. Let $\Sigma_{N_j} = \bigcap_{1 \le k \le 5} \Sigma_{N_j}^k$. There exists a sub-subsequence, denoted by $\mathfrak{n}$, for which

$$\mathbb{Q}\Big[\bigcup_{\mathfrak{n}_0 \ge 1} \bigcap_{\mathfrak{n} \ge \mathfrak{n}_0} \Sigma_{\mathfrak{n}}\Big] = 1.$$

We affirm that all assumptions of Theorem 8.2 hold on the event $\bigcup_{\mathfrak{n}_0 \ge 1} \bigcap_{\mathfrak{n} \ge \mathfrak{n}_0} \Sigma_{\mathfrak{n}}$ intersected with the ones in (10.1), (10.2) and (10.3). Indeed, condition (A.1) follows from the definition of the set $\Sigma_{\mathfrak{n}}^1$. Similarly to the proof of Theorem 2.1, condition (A.3) follows from the definitions of the sequence $M_{\mathfrak{n}}$ and the set $\Sigma_{\mathfrak{n}}^4$. Conditions (6.6), (6.25) follow from the definition of the set $\Sigma_{\mathfrak{n}}^5$.

We turn to condition (A.2). Recall that $\beta_{\mathfrak{n}} = c_{\mathfrak{n}}^{-1}$. Fix a sequence $J_{\mathfrak{n}} \uparrow \infty$ such that $J_{\mathfrak{n}} \le M_{\mathfrak{n}}$, and let $B_{\mathfrak{n}} = \{x_1^{\mathfrak{n}}, \dots, x_{J_{\mathfrak{n}}}^{\mathfrak{n}}\}$. Since we replaced the weights $W_{x_j^{\mathfrak{n}}}^{\mathfrak{n}}$ by $\tau_j^{\mathfrak{n}}$, the first expectation appearing in this hypothesis can be rewritten as

$$\frac{\sum_{1 \le j \le J_{\mathfrak{n}}} [c_{\mathfrak{n}} \tau_j^{\mathfrak{n}}]^2 (\deg(x_j^{\mathfrak{n}}) / v_\ell(x_j^{\mathfrak{n}}))}{\sum_{1 \le j \le J_{\mathfrak{n}}} c_{\mathfrak{n}} \tau_j^{\mathfrak{n}} \deg(x_j^{\mathfrak{n}})}. \tag{10.4}$$

By definition of the set $\Sigma_{\mathfrak{n}}^3$ we may replace $\deg(x_j^{\mathfrak{n}})$ by D_j. Since $\tau_j^{\mathfrak{n}}$ is decreasing, by definition of the set Σ_N^2 the numerator is bounded by

$$c_{\mathfrak{n}} \tau_1^{\mathfrak{n}} \sum_{j=1}^{J_{\mathfrak{n}}} c_{\mathfrak{n}} \tau_j^{\mathfrak{n}} \frac{D_j}{E_j} + \frac{c_{\mathfrak{n}} \tau_1^{\mathfrak{n}}}{M_{\mathfrak{n}}^2} \sum_{j=1}^{J_{\mathfrak{n}}} c_{\mathfrak{n}} \tau_j^{\mathfrak{n}} D_j.$$

The second term divided by the denominator in (10.4) is less than or equal to $c_{\mathfrak{n}} \tau_1^{\mathfrak{n}} M_{\mathfrak{n}}^{-2}$ which goes to 0 as $\mathfrak{n} \to \infty$ in view of (10.3). Also, by (10.3), the first term is bounded by

$$c_{\mathfrak{n}} \tau_1^{\mathfrak{n}} \sum_{j=1}^{J_{\mathfrak{n}}} w_j \frac{D_j}{E_j} + c_{\mathfrak{n}} \tau_1^{\mathfrak{n}} \frac{1}{M_{\mathfrak{n}}} \max_{1 \le j \le J_{\mathfrak{n}}} \frac{D_j}{E_j}.$$

Since the denominator in (10.4) is bounded below by $c_{\mathfrak{n}} \tau_1^{\mathfrak{n}} D_1 \ge c_{\mathfrak{n}} \tau_1^{\mathfrak{n}}$, the first condition in (A.2) follows from (10.1), (10.2).

The second condition of assumption (A.2) can be written as

$$\frac{1}{J_{\mathfrak{n}}} \frac{\sum_{1 \le j \le J_{\mathfrak{n}}} v_\ell(x_j^{\mathfrak{n}}) \deg(x_j^{\mathfrak{n}})}{\sum_{1 \le j \le J_{\mathfrak{n}}} c_{\mathfrak{n}} \tau_j^{\mathfrak{n}} \deg(x_j^{\mathfrak{n}})}.$$

By definition of the set $\Sigma_{\mathfrak{n}}^3$ we may replace $\deg(x_j^{\mathfrak{n}})$ by D_j. The sum in the denominator is bounded below by $c_{\mathfrak{n}}\tau_1^{\mathfrak{n}}D_1 \ge c_{\mathfrak{n}}\tau_1^{\mathfrak{n}}$, which is uniformly bounded. Since the escape probability is bounded by one and since by (B.3) E_j is bounded by one, the numerator is less than or equal to $\sum_{1\le j\le J_{\mathfrak{n}}}(D_j/E_j)$, whose average by (10.2) is bounded.

It remains to establish (8.6) with $\gamma_N = 1$, $Z_j = w_j/E_j$ and $u_j = E_j D_j$. The convergence of the first term follows from (10.3), the definition of $\Sigma_{\mathfrak{n}}^2$ and $\Sigma_{\mathfrak{n}}^3$ and the fact that the variables E_j are bounded by one. The second part of (8.6) amounts to estimate

$$\sum_{j=m}^{M_{\mathfrak{n}}} c_{\mathfrak{n}}\tau_j^{\mathfrak{n}} \deg(x_j^{\mathfrak{n}}) = \sum_{j=m}^{M_{\mathfrak{n}}} c_{\mathfrak{n}}\tau_j^{\mathfrak{n}} D_j \le \sum_{j=m}^{M_{\mathfrak{n}}} w_j(D_j/E_j) + \frac{1}{M_{\mathfrak{n}}^2}\max_{1\le j\le M_{\mathfrak{n}}}(D_j/E_j),$$

where the identity follows from the definition of $\Sigma_{\mathfrak{n}}^3$ and the inequality from (10.3) and the boundedness of E_j. The first term on the right-hand side vanishes in view of (10.1) and the second one by (10.2). This concludes the proof of the theorem.

11. Supercritical Erdős–Rényi random graphs. We show in this section that supercritical Erdős–Rény random graphs satisfy the assumptions of Theorem 2.2. Let $\mathscr{V}_N$ be the set of vertices $\mathscr{V}_N = \{1,\dots,N\}$. For $\lambda > 1$ fixed, let $\{\xi_{x,y} : x, y \in \mathscr{V}_N\}$ be i.i.d. Bernoulli(λ/N) random variables constructed in a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The Erdős–Rényi random graph is defined as $\mathscr{G}_N = (\mathscr{V}_N, \mathscr{E}_N)$, where $\mathscr{E}_N$ is the random set of edges given by $\{\{x,y\}; \xi_{x,y} = 1\}$. Throughout this section, c_j, C_j, $j \ge 0$, represent positive constants depending on λ and sometimes on further parameters, the first ones being tipically small and the last ones large. The next result can be found in [16], Theorem 2.3.2.

THEOREM 11.1. *There is a constant c_0 such that with $\mathbb{P}$-probability converging to one as N tends to infinity, there is a unique component $\mathcal{C}_{\max}$ in $(\mathscr{V}_N, \mathscr{E}_N)$ with $|\mathcal{C}_{\max}| > c_0 \log N$. Moreover, there exists $0 < \mathfrak{v}_\lambda < 1$ such that*

$$\lim_{N\to\infty} \mathbb{P}\left[\left|\frac{|\mathcal{C}_{\max}|}{N} - \mathfrak{v}_\lambda\right| > \varepsilon\right] = 0$$

for all $\varepsilon > 0$.

We will be interested in analyzing the trap model in $\mathcal{C}_{\max}$, providing another interesting example for which our theory can be applied. For the sake of simplicity, we shall assume that the common distribution of the traps $\{W_j^N : j \ge 1\}$ is α-stable. More precisely, recall the definition of the variables $\hat{\tau}_i^N$, $1 \le i \le \mathbb{V}$, introduced in (8.2) with $\mathbb{V} = N$ and $F(t) = t$. We assume in this section that $W_i^N = \hat{\tau}_i^N$, $1 \le i \le N$.

Let $V_N = \mathcal{C}_{\max}$ be the random set of vertices and let $E_N = \{\{x,y\} \subset \mathcal{C}_{\max} : \{x,y\} \in \mathscr{E}_N\}$ be the random set of edges of the random graph G_N. In contrast with the previous examples presented in Section 9, the number of vertices of the random graph G_N is also random. The weights are distributed as follows. Given V_N, reenumerate the weights W_j^N, $1 \le j \le |V_N|$, in decreasing order and denote by $\hat{W}_j^N$ the new sequence, so that $\hat{W}_j^N \ge \hat{W}_{j+1}^N$, $1 \le j < |V_N|$, $\hat{W}_{\sigma(j)}^N = W_j^N$ for some permutation σ of V_N. Randomly enumerate the vertices of V_N, obtaining a vector $(x_1^N, \dots, x_{|V_N|}^N)$, and set $W_{x_j^N}^N = \hat{W}_j^N$. Given this random vertex-weighted graph, we examine the continuous-time random walk X_t^N on G_N with generator given by (1.2).

Note that to define the random weights $W_j^N = \hat{\tau}_j^N$ we divided the interval [0, 1] in N subintervals instead of dividing it in $|V_N|$ intervals. In particular, in contrast with the examples of Section 9, $N^{-1/\alpha} W_{x_1^N}^N$ does not converge to a Fréchet distribution, but so does $\mathfrak{v}_\lambda^{-1/\alpha} N^{-1/\alpha} W_{x_1^N}^N$, where $\mathfrak{v}_\lambda$ is given by Theorem 11.1.

In the rest of this section, we prove that the assumptions of Theorem 2.2 are fulfilled. By Theorem 11.1, the number of vertices converges in probability to $+\infty$. To establish (B.0), fix a sequence $J_N \uparrow \infty$ and denote by $\mathscr{W}_1^N, \dots, \mathscr{W}_N^N$ the sequence $W_1^N, \dots, W_N^N$ enumerated in decreasing order. Note that $\mathscr{W}_j^N \ge \hat{W}_j^N$, $1 \le j \le |V_N|$. By (8.3) and (2.8), for every $\varepsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}\Big[\sum_{j\ge 1} |N^{-1/\alpha} \mathscr{W}_j^N - w_j| \ge \varepsilon\Big] = 0.$$

Since $\sum_{j\ge J_N} w_j$ vanishes almost surely as $N \uparrow \infty$, if Σ_N^0 stands for the event $\sum_{j \ge J_N} N^{-1/\alpha} \mathscr{W}_j^N \le 1$,

$$\lim_{N\to\infty} \mathbb{P}[\Sigma_N^0] = 1.$$

Denote by Σ_N^1 the event $\{|V_N - \mathfrak{v}_\lambda N| \le \varepsilon N\}$ for some $0 < \varepsilon < \min\{\mathfrak{v}_\lambda, 1 - \mathfrak{v}_\lambda\}$. By Theorem 11.1, $\mathbb{P}[\Sigma_N^1] \to 1$. In conclusion, to prove (B.0) we need to show that

$$\lim_{N\to\infty} \mathbb{E}\big[\nu_N(\{x_1, \dots, x_{\min\{J_N, |V_N|\}}\}^c) \mathbf{1}\{\Sigma_N^0 \cap \Sigma_N^1\}\big] = 0.$$

By definition of ν_N, and since all vertices in V_N have degree at least equal to one,

$$\nu_N(\{x_1, \dots, x_{\min\{J_N, |V_N|\}}\}^c) \le \frac{\sum_{j=J_N+1}^{|V_N|} W_{x_j}^N \deg(x_j)}{W_{x_1}^N}.$$

Since $\mathscr{W}_j^N \ge \hat{W}_j^N$, $1 \le j \le |V_N|$,

$$\sum_{j=J_N+1}^{|V_N|} W_{x_j}^N \deg(x_j) \le \sum_{j=J_N+1}^{|V_N|} \mathscr{W}_j^N \deg(x_j) \le \sum_{j=J_N+1}^{N} \mathscr{W}_j^N \deg(x_j),$$

if $x_{|V_N|+1}, \dots, x_N$ represents a random enumeration of the vertices of $\mathscr{V}_N$ which do not belong to the largest component. On the set Σ_N^1, $W_{x_1}^N \ge \max_{1\le k\le c_\lambda N} W_k^N$, where $c_\lambda = \mathfrak{v}_\lambda - \varepsilon$. This latter variable as well as the variables $\mathscr{W}_j^N$ depend only on the Poisson point process defined at the beginning of Section 8. Hence, if we denote by $\mathfrak{W}$ the σ-algebra generated by this process and let $\Sigma_N^{0,1} = \Sigma_N^0 \cap \Sigma_N^1$, we obtain that

$$\begin{aligned}
&\mathbb{E}\left[\frac{\sum_{j=J_N+1}^N \mathscr{W}_j^N \deg(x_j)}{W_{x_1}^N}\mathbf{1}\{\Sigma_N^{0,1}\}\right] \\
&\quad\le \mathbb{E}\left[\frac{\sum_{j=J_N+1}^N \mathscr{W}_j^N \deg(x_j)}{\max_{1\le k\le c_\lambda N} W_k^N}\mathbf{1}\{\Sigma_N^{0,1}\}\right] \\
&\quad\le \mathbb{E}\left[\frac{\mathbf{1}\{\Sigma_N^0\}}{\max_{1\le k\le c_\lambda N} W_k^N}\sum_{j=J_N+1}^N \mathscr{W}_j^N \mathbb{E}[\deg(x_j)\mathbf{1}\{\Sigma_N^1\}|\mathfrak{W}]\right].
\end{aligned}$$

We first estimate the conditional expectation and then the remaining expression. Since the law of the graph $\mathscr{G}_N$ is independent of the σ-algebra $\mathfrak{W}$, the previous conditional expectation is equal to $\mathbb{E}[\deg(x_j)\mathbf{1}\{\Sigma_N^1\}]$. By construction if $j \le |V_N|$, $\deg(x_j)$ has the same distribution as $\deg(x_k)$ for $1 \le k \le |V_N|$, with a similar fact if $j > |V_N|$. Therefore, for a fixed j, the previous expectation is bounded by

$$\sum_{\ell\le j-1} \mathbb{E}\left[\mathbf{1}\{|V_N| = \ell\}\frac{1}{N-\ell}\sum_{y\notin V_N}\deg(y)\right] + \sum_{\ell\ge j}\mathbb{E}\left[\mathbf{1}\{|V_N| = \ell\}\frac{1}{\ell}\sum_{y\in V_N}\deg(y)\right],$$

where the sum is carried over all ℓ such that $|\ell - \mathfrak{v}_\lambda N| \le \varepsilon N$. Estimating the denominators by the worst case, we get that the sum is less than or equal to

$$\frac{1}{\min\{\mathfrak{v}_\lambda - \varepsilon, 1-\varepsilon-\mathfrak{v}_\lambda\}}\mathbb{E}\left[\frac{1}{N}\sum_{y=1}^N \deg(y)\right].$$

This expectation is equal to λ.

It remains to estimate the expectation involving the weights. On the set Σ_N^0, $\sum_{J_N+1\le j\le N} \mathscr{W}_j^N \le N^{1/\alpha}$. On the other hand, using the notation introduced in (8.1), $\max_{1\le k\le c_\lambda N} N^{-1/\alpha} W_k^N \ge w(\lambda)$, where $w(\lambda) = \max_i \hat{w}_i$, and where the maximum is carried over all indices i such that $z_i \le c_\lambda$. Hence,

$$\mathbb{E}\left[\frac{\mathbf{1}\{\Sigma_N^0\}}{\max_{1\le k\le c_\lambda N} W_k^N}\sum_{j=J_N+1}^N \mathscr{W}_j^N\right] \le \mathbb{E}\left[\frac{1}{w(\lambda)}\right].$$

Since $w' = w(\lambda)/c_\lambda^{1/\alpha}$ has a Fréchet distribution, $P(w' \le t) = \exp\{-1/t^\alpha\}$, this expectation is finite, which proves condition (B.0).

The results of this section should still hold if we require the variables W_j^N to belong to the domain of attraction of an α-stable law and to satisfy the bound

$$\limsup_{N\to\infty}\mathbb{E}\left[\left(c_N \sup_{1\le i\le N} W_i^N\right)^{-1}\right] < +\infty,$$

where c_N has been introduced in (2.8).

To understand the asymptotic law of the escape probabilities, we need to introduce a related branching process. Let $\mathcal{T}$ be the random tree obtained by the Galton–Watson process with offspring distribution Poisson(λ) and denote its law by $\mathcal{P}$. Since λ is assumed to be greater than one, the event that $\mathcal{T}$ is infinite has positive $\mathcal{P}$-probability, [16], Theorem 2.1.4. We denote by $\varnothing$ the root of $\mathcal{T}$.

We first show that the neighborhood of a random point in the Erdős–Rényi graph looks like the neighborhood of $\varnothing$ in $\mathcal{T}$. This is made precise as follows. We write (x, G) for a graph with a marked vertex x. We say that (x, G) is isometric to (x', G') if there exists an isometry between G and G', sending x to x'. As an abuse of notation, we consider $A \subseteq \mathscr{V}_N$ both as a set of vertices and as the corresponding induced subgraph of $\mathscr{G}_N$.

PROPOSITION 11.2. *Let $0 < \gamma < (3 \log \lambda)^{-1}$. There exist constants C_1 and $N_0 = N_0(\lambda, \gamma)$ such that given a random point $z \in \mathscr{V}_N$, we can find a coupling Q_N between the random graph $\mathscr{G}_N$ under $\mathbb{P}$ and the Galton–Watson tree $\mathcal{T}$ under $\mathcal{P}$ such that for all $N \geq N_0$,*

$$Q_N\big[\big(z, B(z, \gamma \log N)\big) \textit{ is isometric to } \big(\varnothing, B(\varnothing, \gamma \log N)\big)\big] \geq 1 - C_1 N^{3\gamma \log \lambda - 1}.$$

PROOF. We follow an argument similar to the one in [16], Section 2.2. Assume, without loss of generality, that $z = 1$ and define an exploration of the cluster $\mathcal{C}_1$ containing 1 in the following way. Let $S_0 = \{2, 3, \dots, N\}$, $I_0 = \{1\}$ and $R_0 = \varnothing$. These sets represent respectively the "susceptible," the "infected" and the "removed" sites. Define a discrete time evolution by

$$\begin{aligned} R_{t+1} &= R_t \cup I_t, \\ I_{t+1} &= \{y \in S_t; \xi_{x,y} = 1 \text{ for some } x \in I_t\}, \\ S_{t+1} &= S_t \setminus I_{t+1}. \end{aligned}$$

Note that the cluster $\mathcal{C}_1$ is given by $\bigcup_{t=1}^{\infty} I_t$ and that $B(1, r) = \bigcup_{t=1}^{r} I_t$.

In order to couple the above exploration process with a Galton–Watson branching process, we introduce a new set of independent Bernoulli(λ/N) random variables $\zeta_{x,y}^t$, $t \geq 1$, $x \geq 1$, $1 \leq y \leq N$. Let $Z_0 = 1$ and

$$Z_{t+1} = \sum_{\substack{x \in I_t \\ y \in S_t}} \xi_{x,y} + \sum_{\substack{x \in I_t \\ y \in \mathscr{V}_N \setminus S_t}} \zeta_{x,y}^t + \sum_{x=N+1}^{N+Z_t-|I_t|} \sum_{y=1}^{N} \zeta_{x,y}^t. \tag{11.1}$$

The first term in the above sum can be written as $|I_{t+1}| + C_{t+1}$, where C_{t+1} represents the number of "collisions" occurring in the exploration process, that is, individuals in I_{t+1} connected to more than one individual in I_t. The second term stands for the "immigrants" introduced to compensate the fact that $|S_t| < N$, and the third term for children of individuals that are not in I_t.

It is easy to check that the process $\{Z_t : t \geq 0\}$ is a branching process with offspring distribution Binomial$(N, \lambda/N)$. Let $\mathcal{T}'$ be the random tree associated with Z_t. More precisely, if x is the ith individual in the tth generation of $\mathcal{T}'$, the number of offsprings of x will be given by

$$\begin{cases} \sum_{y \in S_t} \xi_{x,y} + \sum_{y \in \mathscr{V}_N \setminus S_t} \zeta_{x,y}^t, & \text{if } i \leq |I_t|, \\ \sum_{y=1}^{N} \zeta_{x,y}^t, & \text{otherwise.} \end{cases}$$

It is immediate to check that Z_t is the size of the tth generation of $\mathcal{T}'$ and that $Z_t \geq |I_t|$.

On the event $Z_s = |I_s|$, $1 \leq s \leq t$, there were no collisions and no immigrants. Therefore, in this event the subgraph $(1, B(1,t))$ of $\mathscr{G}_N$ is isometric to the subgraph $(\varnothing, B(\varnothing, t))$ of $\mathcal{T}'$. Hence, by [16], Theorem 2.2.2, with $t = \gamma \log N$, there exist a constant $C_1 < \infty$ and a coupling Q' between $\mathscr{G}_N$ and $\mathcal{T}'$ such that with probability at least $1 - C_1 N^{2\gamma \log \lambda - 1}$, $(1, B(1,t))$ is isometric to $(\varnothing, B(\varnothing, t))$.

Claim A: Let $0 < \gamma < (3 \log \lambda)^{-1}$. There exist n_0 and a coupling Q'' between the tree $\mathcal{T}'$ with Binomial$(N, \lambda/N)$ offsprings and the tree $\mathcal{T}$ with Poisson(λ) offsprings, such that, with probability at least $1 - C_1 N^{3\gamma \log \lambda - 1}$, $(\varnothing, B(\varnothing, \gamma \log N))$ (in $\mathcal{T}'$) is isometric to $(\varnothing, B(\varnothing, \gamma \log N))$ (in $\mathcal{T}$) for $N \geq n_0$.

It is well known that a Poisson(λ) random variable Y can be coupled with a Binomial$(N, \lambda/N)$ random variable Y', in a way that

$$P[Y = Y'] \geq 1 - 2\lambda^2 N^{-1}, \tag{11.2}$$

see, for instance, [15], Chapter 2.6, or [27], Theorem 1, for a bound on the total variation distance and [26], Chapter 4, for a connection between total variation distance and coupling. On the other hand, by [1], Theorem 4, there exist $\theta = \theta(\lambda) > 0$ and C_3 such that for and any t, $A \geq 0$,

$$\mathcal{P}[Z_t \geq A\lambda^t] = \mathcal{P}[e^{\theta(Z_t/\lambda^t)} \geq e^{\theta A}] \leq e^{-\theta A} \mathcal{E}[e^{\theta(Z_t/\lambda^t)}] \leq C_3 e^{-\theta A}.$$

This bound permits to estimate the volume of the subgraph $B(\varnothing, \gamma \log N)$ of $\mathcal{T}$. Fix $\gamma \in (0,1)$. Since $|B(\varnothing, \gamma \log N)| = \sum_{0 \leq t \leq \gamma \log N} Z_t$, we have that

$$\begin{aligned} &\mathcal{P}[|B(\varnothing, \gamma \log N)| \geq N^{3\gamma \log \lambda}] \\ &\quad \leq \sum_{t=0}^{\gamma \log N} \mathcal{P}[Z_t \geq N^{2\gamma \log \lambda}] \leq \sum_{t=0}^{\gamma \log N} \mathcal{P}[Z_t \geq N^{\gamma \log \lambda} \lambda^t] \end{aligned}$$

for all N greater than some constant $N_0 = N_0(\lambda, \gamma)$. Therefore, applying the previous estimate, we conclude that for every $0 < \gamma < 1$, there exist $C_3 < \infty$ and $N_0(\lambda, \gamma) < \infty$ such that

$$\mathcal{P}[|B(\varnothing, \gamma \log N)| \geq N^{3\gamma \log \lambda}] \leq C_3 \exp\{-\theta N^{\gamma \log \lambda}\} \tag{11.3}$$

for all $N \geq N_0$.

Claim A follows from (11.2) and (11.3), which concludes the proof of Proposition 11.2. □

In the proof of the previous lemma, we also obtained a bound on the size of a ball $B(z, \gamma \log N)$ around a typical point z.

COROLLARY 11.3. *For any $0 < \gamma < (3 \log \lambda)^{-1}$, there exist a finite constant C_2 and an integer N_0, depending only on λ and γ, such that for any random point $z \in \{1, \dots, N\}$,*

$$\mathbb{P}\big[\big|B(z, \gamma \log N)\big| \geq N^{3\gamma \log \lambda}\big] \leq C_2 N^{3\gamma \log \lambda - 1}$$

for all $N \geq N_0$.

As required in (B.3), we extend the local isometry obtained in Proposition 11.2 to various balls in the random graph $\mathscr{G}_N$.

COROLLARY 11.4. *Fix positive numbers b and γ such that $0 < 2b + 6\gamma \log \lambda < 1$. There exist constants C_0, N_0, depending only on λ and γ, and a coupling $Q' = Q'_N$ between the random graph $\mathscr{G}_N$ and N^b independent Galton–Watson trees $\mathcal{T}_i$, $1 \leq i \leq N^b$, such that for all $N \geq N_0$,*

$$Q'\big[\mathscr{B}^c\big] \leq C_0 N^{2b + 6\gamma \log \lambda - 1},$$

where $\mathscr{B}$ is the event "The balls $(z_i, B(z_i, \gamma \log N))$, $1 \leq i \leq N^b$, are disjoint and isometric to $(\varnothing_i, B(\varnothing_i, \gamma \log N))$," and $z_1, \dots, z_{N^b}$ are sites randomly chosen in $\{1, \dots, N\}$.

PROOF. Choose randomly N^b sites on $\{1, \dots, N\}$, denoted by $z_1, \dots, z_{N^b}$. By Proposition 11.2, for N large, there is a coupling Q' between independent Erdős–Rényi random graphs $\mathscr{G}_N^i$, $1 \leq i \leq N^b$, and independent Galton–Watson trees $\mathcal{T}_i$ in a way that with probability at least $1 - C_1 N^b N^{3\gamma \log \lambda - 1}$ each ball $(z_i, B(z_i, \gamma \log N))$ in $\mathscr{G}_N^i$ is isomorphic to $(\varnothing_i, B(\varnothing_i, \gamma \log N))$ in $\mathcal{T}_i$.

We construct an Erdős–Rényi-distributed graph $\mathscr{G}_N$ which is partially determined by the above $\mathscr{G}_N^i$'s. We first explore the ball $B(z_1, \gamma \log N)$ in $\mathscr{G}_N^1$. Every edge $\{x, y\}$ revealed during this exploration is open in $\mathscr{G}_N$ if and only if it is open in $\mathscr{G}_N^1$. Then we proceed by exploring $B(z_2, \gamma \log N)$ in $\mathscr{G}_N$ observing only that we do not reassign values to edges in $\mathscr{G}_N$ that were already established in the previous step. After proceeding with this exploration for $i = 1, \dots, N^b$, we assign the remaining edges of $\mathscr{G}_N$ independently.

It is clear from the above exploration procedure that the graph $\mathscr{G}_N$ is distributed as an Erdős–Rényi random graph. Moreover, on the event $\mathscr{A}$ defined as "the balls $B(z_i, \gamma \log N)$, $i = 1, \dots, N^b$, are pairwise disjoint in $\{1, \dots, N\}$," we have that

$(z_i, B(z_i, \gamma \log N))$ in $\mathscr{G}_N$ is isomorphic to the corresponding pair in $\mathscr{G}_N^i$. Consequently, they will be isomorphic to $(\varnothing_i, B(\varnothing_i, \gamma \log N))$ in $\mathcal{T}_i$. Therefore, to conclude the proof of the corollary, it remains to estimate $Q'[\mathscr{A}^c]$.

Since all the vertices are indistinguishable, $Q'[\mathscr{A}^c]$ is bounded by

$$N^{2b} Q'[B(z_1, \gamma \log N) \cap B(z_2, \gamma \log N) \neq \varnothing] = N^{2b} Q'[z_1 \in B(z_2, 2\gamma \log N)].$$

Since z_2 is independent of z_1, this latter probability is bounded by

$$Q'[|B(z_2, 2\gamma \log N)| \geq N^{6\gamma \log \lambda}] + \frac{1}{N} N^{6\gamma \log \lambda}.$$

By Corollary 11.3, for N large, the first term is bounded above by $C_2 N^{6\gamma \log \lambda - 1}$ for some finite constant C_2. Hence,

$$Q'[\mathscr{A}^c] \leq C_2 N^{2b + 6\gamma \log \lambda - 1},$$

which proves the corollary. $\square$

It is a well-known fact that

$$\text{conditioned on being infinite, } \mathcal{T} \text{ is } \mathcal{P}\text{-a.s. transient;} \tag{11.4}$$

see Theorem 3.5 and Corollary 5.10 in [30]. We denote by $v_\varnothing$ the probability that a simple random walk starting at $\varnothing$ never returns to this site, the so called escape probability. As we will show, the distribution of $v_\varnothing$ under $\mathcal{P}$ is close to that of the probability that a random walk on the giant component $\mathcal{C}_{\max}$ of the random graph $\mathscr{G}_N$ escapes from a certain neighborhood of a random vertex.

Since the isometry obtained in Corollary 11.4 is local, we need a tool to show that looking at a neighborhood of $\varnothing \in \mathcal{T}$ we can obtain precise estimates on the escape probability $v_\varnothing$. The next result plays a central role in this respect. Denote by Δ_l, $l \geq 0$, the points of the lth generation of a tree: $\Delta_l = B(\varnothing, l) \setminus B(\varnothing, l-1)$.

For a fixed tree $\mathcal{T}$, we denote by $\mathbf{P}_y = \mathbf{P}_y^\tau$, $y \in \mathcal{T}$ the probability induced by the discrete-time simple random walk on $\mathcal{T}$ starting from y.

PROPOSITION 11.5. *There exist constants c_1, c_2, depending only on λ, such that, for every $l \geq 1$,*

$$\mathcal{P}\Big[\sup_{y \in \Delta_l} \mathbf{P}_y[H_\varnothing < \infty] \geq \exp\{-c_1 l\}\Big] \leq \exp\{-c_2 l\}.$$

PROOF. Throughout the proof of this lemma, given a rooted tree $\mathcal{T}$ and a vertex $y \in \mathcal{T}$, we denote by $\mathcal{T}_y$ the subtree formed by the root y together with the descendants of y in $\mathcal{T}$.

The idea is to show that in the path between y and $\varnothing$ there are many tunnels from which the random walk can escape to infinity. In order to properly define these tunnels, we need to introduce some extra notation. For an arbitrary tree $\mathcal{T}$

rooted at $\varnothing$, we define the tree $\mathcal{T}^{\text{tail}}$, obtained by adding a vertex $\varnothing'$ which is connected to $\varnothing$ by an edge. This extra element should be regarded as the ancestor of $\varnothing$. In the proof, we use the notation $\mathbf{P}_x^{\mathcal{T}}$ to specify on which tree the random walk is defined.

For a given $\delta > 0$, we say that a tree $\mathcal{T}$ with root $\varnothing$ satisfies the property $\mathcal{Q}^\delta$ if

$$\mathbf{P}_{\varnothing}^{\mathcal{T}^{\text{tail}}}[H_{\varnothing'} = \infty] \geq \delta.$$

In other words, the property $\mathcal{Q}^\delta$ is saying that a random walk on $\mathcal{T}^{\text{tail}}$ has probability at least δ of never hitting the ancestor $\varnothing'$ of the root $\varnothing$.

It is clear from (11.4) that for every $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon, \lambda) > 0$ such that

$$\mathcal{P}\big[\mathcal{T} \text{ does not satisfy } \mathcal{Q}^\delta\big] \leq q + \varepsilon, \tag{11.5}$$

where q is the extinction probability: $q = \mathcal{P}[\mathcal{T} \text{ is finite}]$.

If y is in the lth generation of $\mathcal{T}$, we write $\varnothing = y_0, y_1, \dots, y_l = y$ to denote the unique simple path connecting $\varnothing$ to y. Moreover, we denote by $\Gamma(y)$ the number of elements y_k, $0 \leq k < l$, having at least one descendant $y_k' \neq y_{k+1}$ such that $\mathcal{T}_{y_k'}$ satisfies $\mathcal{Q}^\delta$.

We can now use (11.5) together with [23], Lemma 1, to conclude that there exist constants c_3 and c_4 such that

$$\mathcal{P}\big[\exists y \in \Delta_l \text{ such that } \Gamma(y) < c_3 l\big] \leq \exp\{-c_4 l\}.$$

To conclude the proof of the lemma, it remains to show that there exists $c_1 > 0$ for which the event "$\exists y \in \Delta_l$ such that $\mathbf{P}_y[H_\varnothing < \infty] \geq \exp\{-c_1 l\}$" is contained in the event "$\exists y \in \Delta_l$ such that $\Gamma(y) < c_3 l$."

Assume that all points z in generation l of $\mathcal{T}$ are such that $\Gamma(z) \geq c_3 l$ and fix a point $y \in \Delta_l$. Recall the definition of $y_0, \dots, y_l$ given above and consider a subsequence k_j, $1 \leq j \leq c_3 l$, for which y_{k_j} has a descendant $y_{k_j}' \neq y_{k_j+1}$ such that $\mathcal{T}_{y_{k_j}'}$ satisfies $\mathcal{Q}^\delta$. These points are the entrance to the tunnels $\mathcal{T}_{y_{k_j}'}$ that we have referred to in the beginning of the proof.

Let $\mathcal{T}_-$ be the subtree of $\mathcal{T}$ with all the descendants of y_{k_j} removed, $1 \leq j \leq c_3 l$, with the exception of y_{k_j+1} and y_{k_j}'. An argument based on flows or capacities shows that $\mathbf{P}_y^{\mathcal{T}}[H_\varnothing < \infty] \leq \mathbf{P}_y^{\mathcal{T}_-}[H_\varnothing < \infty] \leq \mathbf{P}_{y_{k_m}}^{\mathcal{T}_-}[H_\varnothing < \infty]$ where $m = c_3 l$. By the strong Markov property,

$$\mathbf{P}_{y_{k_m}}^{\mathcal{T}_-}[H_\varnothing < \infty] \leq \mathbf{P}_{y_{k_m}}^{\mathcal{T}_-}[H_{y_{k_{m-1}}} < \infty]\mathbf{P}_{y_{k_{m-1}}}^{\mathcal{T}_-}[H_\varnothing < \infty].$$

Since $\mathcal{T}_{y_{k_j}'}$ satisfies $\mathcal{Q}^\delta$ and since we removed all descendants of y_{k_j} with the exception of y_{k_j}' and y_{k_j+1}, $\mathbf{P}_{y_{k_m}}^{\mathcal{T}_-}[H_{y_{k_{m-1}}} = \infty] \geq (1/3)\mathbf{P}_{y_{k_m}'}^{\mathcal{T}_-}[H_{y_{k_m}} = \infty] \geq \delta/3$. Hence, the previous expression is bounded by

$$\big[1 - (\delta/3)\big]\mathbf{P}_{y_{k_{m-1}}}^{\mathcal{T}_-}[H_\varnothing < \infty].$$

Iterating this argument $m-1$ times, we finally get that $\mathbf{P}_y^{\mathcal{T}}[H_\varnothing < \infty]$ is bounded by $[1-(\delta/3)]^{c_3 l-1}$, which concludes the proof of the lemma. $\square$

Proposition 11.5 permits to approximate the inverse of the escape probability $v_\varnothing$ by a local quantity. Fix a infinite tree $\mathcal{T}$ and $m \geq 1$. Let $v_\varnothing^{(m)}$ be the probability to escape from $B(\varnothing, m)$, $v_\varnothing^{(m)} = \mathbf{P}_\varnothing[H_\varnothing^+ > H_{B(\varnothing,m)^c}]$. Recall from [26], Chapter 9, the notion of flow and energy of a flow. Since $|\mathcal{T}| = \infty$, we can define a trivial unit flow from $\varnothing$ to $B(\varnothing, m)^c$ which has energy equal to m. Hence, by Proposition 9.5 and Theorem 9.10 of [26],

$$v_\varnothing^{(m)} \geq (d_\varnothing m)^{-1}, \tag{11.6}$$

where $d_\varnothing$ is the degree of the root.

COROLLARY 11.6. *There exist positive constants c_1 and c_2, depending only on λ, such that*

$$\mathcal{P}\big[|\Delta_l| < \exp\{c_1 l\} | \Delta_l \neq \emptyset\big] \leq \exp\{-c_2 l\}$$

for every $l \geq 1$.

PROOF. For a tree with at least l generations, let $\mathcal{G}_l$ be the graph obtained by identifying all points in Δ_l, naming this vertex z_l. All other sites are left untouched, and the number of vertices of this new graph is $|B(\varnothing, l)| - |\Delta_l| + 1$. Since the stationary measure of a simple random walk is proportional to the degree of the vertex,

$$|\Delta_l|/d_\varnothing = \pi(z_l)/\pi(\varnothing),$$

where π stands for the stationary measure of a simple random walk on $\mathcal{G}_l$. The ratio in the right-hand side of the above equation can be estimated using the escape probabilities from these two points. Let $\mathbf{P}_x^{\mathcal{G}}$, $x \in \mathcal{G}_l$ stand for the probability on the path space induced by a discrete-time random walk on $\mathcal{G}_l$ starting from x. Recall that the resistance between $\varnothing$ and z_l is the same as the resistance between z_l and $\varnothing$, so that

$$\frac{\pi(z_l)}{\pi(\varnothing)} = \frac{\mathbf{P}_\varnothing^{\mathcal{G}}[H_{z_l} < H_\varnothing^+]}{\mathbf{P}_{z_l}^{\mathcal{G}}[H_\varnothing < H_{z_l}^+]}.$$

We may couple the random walk on $\mathcal{G}_l$ with a random walk on the tree in such a way that $\mathbf{P}_\varnothing^{\mathcal{G}}[H_{z_l} < H_\varnothing^+] = \mathbf{P}_\varnothing[H_{\Delta_l} < H_\varnothing^+]$ and that $\mathbf{P}_{z_l}^{\mathcal{G}}[H_\varnothing < H_{z_l}^+] \leq \max_{y\in\Delta_l} \mathbf{P}_y[H_\varnothing < H_{\Delta_l}^+]$. By (11.6), $\mathbf{P}_\varnothing[H_{\Delta_l} < H_\varnothing^+] \geq (d_\varnothing l)^{-1}$. Putting together all previous estimates, we get that on the set $\Delta_l \neq \emptyset$,

$$|\Delta_l|^{-1} \leq l \max_{y\in\Delta_l} \mathbf{P}_y\big[H_\varnothing < H_{\Delta_l}^+\big] \leq l \max_{y\in\Delta_l} \mathbf{P}_y[H_\varnothing < \infty]. \tag{11.7}$$

Since there is a positive probability that a super-critical tree survives, the probability appearing in the statement of the lemma is bounded by $C_0\mathcal{P}[|\Delta_l| < \exp\{c_1 l\}, \Delta_l \neq \emptyset]$. By (11.7), this probability is bounded by $C_0\mathcal{P}[l \max_{y\in\Delta_l} \mathbf{P}_y[H_\varnothing < \infty] \geq \exp\{-c_1 l\}]$, which is bounded by $\exp\{-c_2 l\}$ by Proposition 11.5. □

COROLLARY 11.7. *For any $0 < \gamma < 1$, there exist positive constants c_0 and $N_0 \geq 1$, depending only on γ and λ, such that for all $N \geq N_0$,*

$$\mathcal{P}\left[\left|\frac{1}{v_\varnothing} - \frac{1}{v'_\varnothing}\right| \geq d_\varnothing N^{-c_0} \Big| |\mathcal{T}| = \infty\right] \leq N^{-c_0},$$

where $d_\varnothing$ represents the degree of $\varnothing$ and $v'_\varnothing = \mathbf{P}_\varnothing[H^+_\varnothing > H_{B(\varnothing,\gamma\log N)^c}]$.

PROOF. Fix $0 < \gamma < 1$ and an infinite tree $\mathcal{T}$. To keep notation simple, let $B = B(\varnothing, \gamma \log N)$ and let ∂B be the set of points in B^c which have a neighbor in B. By the strong Markov property,

$$\mathbf{P}_\varnothing[H^+_\varnothing > H_{B^c}] \geq \mathbf{P}_\varnothing[H^+_\varnothing = \infty] \geq \mathbf{P}_\varnothing[H^+_\varnothing > H_{B^c}] \inf_{x\in\partial B} \mathbf{P}_x[H^+_\varnothing = \infty].$$

Inverting these terms, we obtain

$$0 \leq \frac{1}{v'_\varnothing} - \frac{1}{v_\varnothing} \leq \frac{1}{v'_\varnothing}\left(\frac{1}{\inf_{x\in\partial B}\mathbf{P}_x[H_\varnothing = \infty]} - 1\right).$$

By Proposition 11.5 with $l = \gamma \log N$, there exists constants $c_1, c_2 > 0$, depending on λ, such that on a set with probability at least $1 - N^{-\gamma c_2}$ the previous infimum is bounded below $1 - N^{-\gamma c_1}$. Since $(1-x)^{-1} \leq 1 + 2x$ for $x \in (0, 1/2)$, there exists $N_0 = N_0(\gamma, \lambda)$ such that for $N \geq N_0$,

$$\left|\frac{1}{v'_\varnothing} - \frac{1}{v_\varnothing}\right| \leq \frac{2}{N^{\gamma c_1}}\frac{1}{v'_\varnothing}.$$

Estimate (11.6) permits to conclude the proof of the corollary, changing the values of the exponents if necessary. □

COROLLARY 11.8. *Let $\mathcal{T}$ be a Galton–Watson tree with* Poisson(λ) *offsprings, $\lambda > 1$. Then there exist finite constants c_0, C_0 and $s_0 < \infty$, depending only on λ, such that*

$$\mathcal{P}[(v_\varnothing)^{-1} \geq s \,|\, |\mathcal{T}| = \infty] \leq C_0 \exp\{-c_0\sqrt{s}\}$$

for all $s \geq s_0$.

PROOF. Since $\mathcal{T}$ is supercritical, the probability appearing in the statement of the lemma is bounded by $C_3\mathcal{P}[(v_\varnothing)^{-1} \geq s, |\mathcal{T}| = \infty]$ for some finite constant C_3 depending only on λ. Fix an integer $n \geq 1$. By the strong Markov property, $v_\varnothing$

is bounded below by $\mathbf{P}_\varnothing[H_{B^c} \le H_\varnothing^+] \inf_{y\in B^c} \mathbf{P}_y[H_\varnothing = \infty]$, where $B = B(\varnothing, n)$. Therefore, $\mathcal{P}[(v_\varnothing)^{-1} \ge s, |\mathcal{T}| = \infty]$ is less than or equal to

$$\mathcal{P}\big[\mathbf{P}_\varnothing\big[H_{B^c} \le H_\varnothing^+\big]^{-1} \ge s/2, |\mathcal{T}| = \infty\big] + \mathcal{P}\Big[\inf_{y\in B^c} P_y[H_\varnothing = \infty] \le 1/2\Big].$$

By (11.6), $\mathbf{P}_\varnothing[H_{B^c} \le H_\varnothing^+] \ge (d_\varnothing n)^{-1}$. The previous expression is thus bounded by

$$\mathcal{P}[d_\varnothing n \ge s/2] + \mathcal{P}\Big[\sup_{y\in B^c} P_y[H_\varnothing < \infty] \ge 1/2\Big].$$

Set $n = \sqrt{s}$, recall that $d_\varnothing$ has a Poisson(λ) distribution. Apply an exponential Chebyshev inequality to estimate the first term. By Proposition 11.5 with $l = \sqrt{s}$, the second term is bounded by $\exp\{-c_2\sqrt{s}\}$ provided s is large enough. □

The following corollary allows us to bound the quantity ε_N appearing in (6.5) and (6.7). This corresponds to the probability of entering the neighborhood of a deep trap before L_N times the mixing time.

COROLLARY 11.9. *Fix an arbitrary vertex $y \in \{1, \dots, N\}$ and $0 < \gamma < (3\log\lambda)^{-1}$. Then there exists positive constants c_0 and $N_0 \ge 1$, depending only on γ and λ, such that for all $N \ge N_0$,*

$$\mathbb{P}\Big[\sup_{z\in B(y,\gamma\log N)^c} \mathbf{P}_z\big[H_y \le \log^4 N\big] > N^{-c_0}\Big] \le N^{-c_0}.$$

PROOF. Denote by $\partial_i A$ the internal boundary of a set A: $\partial_i A = \{x \in A : d(x, A^c) = 1\}$. Fix $0 < \gamma < (3\log\lambda)^{-1}$. By Propositions 11.2 and 11.5, there exist positive constants c_1, c_2 and C_1, depending only on λ, such that

$$\mathbb{P}\Big[\sup_{z\in\partial_i B} \mathbf{P}_z[H_y \le H_{B^c}] > N^{-\gamma c_1}\Big]$$
$$\le C_1 N^{-c} + \mathcal{P}\Big[\sup_{z\in\partial_i B} \mathbf{P}_z[H_\varnothing \le H_{B^c}] > N^{-\gamma c_1}\Big] \le C_1 N^{-c} + N^{-\gamma c_2},$$

where $c = 3\gamma\log\lambda - 1$ and $B = B(y, \gamma\log N)$.

Assume that $\sup_{z\in\partial_i B} \mathbf{P}_z[H_y \le H_{B^c}] \le N^{-\gamma c_1}$. We claim that in this case

$$\sup_{z\in B^c} \mathbf{P}_z\big[H_y \le \log^4 N\big] \le N^{-\gamma c_1} + \sup_{z\in B^c} \mathbf{P}_z\big[H_y \le \log^4 N - 1\big]. \tag{11.8}$$

Iterating this estimate $\log^4 N$ times, we conclude the proof of the corollary. It is enough, therefore, to prove (11.8). By the strong Markov property, $\mathbf{P}_z[H_y \le \log^4 N]$ is bounded by $\sup_{w\in\partial_i B} \mathbf{P}_w[H_y \le \log^4 N]$. If $\{H_y < H_{B^c}\}$, by the initial assumption we may bound the probability by $N^{-\gamma c_1}$. This gives the first term on the right-hand side of (11.8). On the other hand, on the set $\{H_y > H_{B^c}\}$,

$H_y = H_{B^c} + H_y \circ \theta_{H_{B^c}}$ and $H_y \circ \theta_{H_{B^c}} \le \log^4 N - 1$. Hence, by the strong Markov property, for every $w \in \partial_i B$,

$$\mathbf{P}_w\big[H_y \le \log^4 N, H_{B^c} < H_y\big] \le \mathbf{P}_w[H_{B^c} < H_y] \sup_{z \in B^c} \mathbf{P}_z\big[H_y \le \log^4 N - 1\big],$$

which proves (11.8) and the corollary. $\square$

We conclude this section deriving the scaling limit of the random walk X_t^N on the giant component of the supercritical Erdős–Rényi random graph.

THEOREM 11.10. *Consider the trap model X_t^N on the largest component $\mathcal{C}_{\max}$ of the Erdős–Rényi random graph with traps W_x^N, $x \in \mathcal{C}_{\max}$, as described in the beginning of this section. Assume that $\Psi_N(X_0^N)$ converges in probability to some $k \in \mathbb{N}$. Let $\beta_N = (\mathfrak{v}_\lambda N)^{1/\alpha}$. Then*

$$\big(\beta_N^{-1}\mathbf{W}^N, \Psi_N\big(X_{t\beta_N}^N\big)\big) \qquad \text{converges weakly to } (\mathbf{w}, K_t),$$

where $\mathbf{w}$ is the sequence defined in (8.4) and where, for each fixed $\mathbf{w}$, K_t is a K-process starting from k with parameter $(\mathbf{Z}, \mathbf{u})$, where $Z_k = w_k/E_k$ and $u_k = D_k E_k$. Here, (D_k, E_k), $k \ge 1$ is an i.i.d. sequence, distributed as $(d_\varnothing, v_\varnothing)$ under $\mathcal{P}[\cdot||\mathcal{T}| = \infty]$. The above convergence refers to the L^1-topology in the first coordinate and d_T-topology in the second.

PROOF. We need to establish conditions (B.0)–(B.3) for the above sequence of graphs and to apply Theorem 2.2. Condition (B.0) has been proven in the beginning of this section. The main difficulty in checking the remaining hypotheses comes from the fact that we are dealing with the giant component $\mathcal{C}_{\max}$, which has a random size, instead of the whole set $\{1, \dots, N\}$ as in the above lemmas and propositions.

In order to prove (B.1), let $\ell_N = (\gamma/2)\log N$ with γ satisfying the conditions of Corollary 11.4. Since the term inside the expectation in (B.1) is bounded by one, the expectation in (B.1) is less than or equal to

$$\mathbb{E}\bigg[\frac{1}{|\mathcal{C}_{\max}|}\sum_{x \in \mathcal{C}_{\max}} \frac{|B(x, 2\ell_N)|}{|\mathcal{C}_{\max}|}\bigg]$$
$$\le \mathbb{P}\big[|\mathcal{C}_{\max}| < (\mathfrak{v}_\lambda/2)N\big] + \frac{4}{(\mathfrak{v}_\lambda N)^2}\mathbb{E}\Bigg[\sum_{x=1}^{N}\big|B(x, 2\ell_N)\big|\Bigg].$$

By Theorem 11.1, the first term vanishes as $N \uparrow \infty$, while by Proposition 11.2 and Corollary 11.3 the second term vanishes. This proves that condition (B.1) is fulfilled.

By [3, 21, 22], with high probability the mixing time of a random walk on $\mathcal{C}_{\max}$ is less than or equal to $C_0 \log^2 N$ for some finite constant C_0. Choosing

$L_N = C_0^{-1} \log^2 N$, the hypothesis (B.2) becomes a direct consequence of Corollary 11.9. It is indeed enough to condition the event appearing in the statement of Corollary 11.9 on the set that y belongs to $\mathcal{C}_{\max}$ and to recall from Theorem 11.1 that the giant component has a positive density with probability converging to 1.

It remains to check (B.3). Let Q'_N be the coupling between the random graph $\mathscr{G}_N$ and N^b independent Galton–Watson trees $\mathcal{T}_i$ constructed in Corollary 11.4. We assume that these trees are the first N^b trees of an infinite i.i.d. sequence of Galton–Watson trees.

Fix $K \geq 1$ and let $\mathfrak{x}_1, \mathfrak{x}_2, \dots, \mathfrak{x}_K$ be the first K points z_i which belongs to $\mathcal{C}_{\max}$: $\mathfrak{x}_1 = z_j$ if $z_j \in \mathcal{C}_{\max}$ and $z_i \notin \mathcal{C}_{\max}$ for $1 \leq i < j$, and so on. It is clear that $\mathfrak{x}_1, \dots, \mathfrak{x}_K$ is uniformly distributed among all possible choices and that the probability of not finding K points in $\mathcal{C}_{\max}$ among N^b points uniformly distributed in $\mathscr{V}_N$ converges to 0.

Let $\mathfrak{y}_j$, $1 \leq j \leq K$, be the first K indices of trees $\mathcal{T}_i$ which are infinite and let (D_j, E_j) be the degree and the escape probabilities $(d_\varnothing, v_\varnothing)$ in $\mathcal{T}_{\mathfrak{y}_j}$. Note that the vectors (D_j, E_j) are independent and identically distributed and that $Q'_N[(D_1, E_1) \in A] = \mathcal{P}[(d_\varnothing, v_\varnothing) \in A \,|\, |\mathcal{T}| = \infty]$. In particular, by Corollary 11.8 and Schwarz inequality the last two conditions in (B.3) are fulfilled.

Let $\mathbb{A}_N$ be the event "The graphs $(\mathfrak{x}_i, B(\mathfrak{x}_i, \gamma \log N))$, $1 \leq i \leq K$, are isometric to the graphs $(\mathfrak{y}_i, B(\mathfrak{y}_i, \gamma \log N))$, $1 \leq i \leq K$." In view of Corollary 11.7, on the set $\mathbb{A}_N$, the first two condition in (B.3) are fulfilled. To conclude the proof of condition (B.3), it remains to show that

$$\lim_{N \to \infty} \mathbb{P}[\mathbb{A}_N^c] = 0. \tag{11.9}$$

We define six sets $\Sigma_{N,j}$, $0 \leq j \leq 5$, such that $\bigcap_{0 \leq j \leq 5} \Sigma_{N,j} \subset \mathbb{A}_N$ and then prove that each of this set has asymptotic full measure. Recall that b and γ satisfy the assumptions of Corollary 11.4 and let $\Sigma_{N,0} = \mathscr{B}$ be the set introduced in that corollary. Since $\{(\mathfrak{x}_i)_{i=1}^K = (\mathfrak{y}_i)_{i=1}^K\} \cap \mathscr{B} \subset \mathbb{A}_N$, it is enough to find conditions which guarantee that $\mathfrak{x}_i = \mathfrak{y}_i$, $1 \leq i \leq K$.

Let $\Sigma_{N,1} = \{\mathrm{diam}(\mathcal{C}_{\max}) \geq \gamma \log N\}$, let $\Sigma_{N,2} = \{|\{z_1, \dots, z_{\log N}\} \cap \mathcal{C}_{\max}| \geq K\}$ and let $\Sigma_{N,3}$ be the event "Every three $\mathcal{T}_i$, $1 \leq i \leq \log N$, with diameter greater or equal to $\gamma \log N$ survives." On $\Sigma_{N,0} \cap \Sigma_{N,1} \cap \Sigma_{N,2} \cap \Sigma_{N,3}$, the graphs $(\mathfrak{x}_i, B(\mathfrak{x}_i, \gamma \log N))$, $1 \leq i \leq K$, are coupled to infinite trees.

It remains to guarantee that there is no infinite tree coupled with a graph $(z_i, B(z_i, \gamma \log N))$ whose root z_i does not belong to $\mathcal{C}_{\max}$. Let $\Sigma_{N,4}$ be the event "Every tree $\mathcal{T}_i$, $1 \leq i \leq \log N$, with diameter greater of equal than $\gamma \log N$ has at least N^δ elements among the first $\gamma \log N$ generations," and let $\Sigma_{N,5}$ be the event "Every connected subset of $\mathscr{V}_N$ with more than N^δ elements is contained in $\mathcal{C}_{\max}$." On $\Sigma_{N,0} \cap \Sigma_{N,4} \cap \Sigma_{N,5}$, all infinite trees $\mathcal{T}_i$, $1 \leq i \leq \log N$, are coupled with graphs whose root belongs to $\mathcal{C}_{\max}$.

Putting together the previous assertions, we get that $\bigcap_{0 \leq j \leq 5} \Sigma_{N,j} \subset \mathbb{A}_N$, as claimed. We next show that each event introduced above has asymptotic full probability. By Corollary 11.4, $\mathbb{P}[\Sigma_{N,0}^c]$ vanishes, by Theorem 11.1 and by Corollary 11.3 $\mathbb{P}[\Sigma_{N,1}^c]$, and by Theorem 11.1, $\mathbb{P}[\Sigma_{N,2}^c]$ vanishes. By Corollary 11.6,

$\mathbb{P}[\Sigma_{N,4}^c]$ vanishes for some $\delta > 0$, and by Theorem 11.1 $\mathbb{P}[\Sigma_{N,5}^c]$ vanishes. Finally, by Corollary 11.6, there exists $\delta = \delta(\gamma, \lambda) > 0$ with the following property. A tree which has diameter $\gamma \log N$ has at least N^δ elements at generation $\gamma \log N$ with probability converging to 1. Since from each element of the generation $\gamma \log N$ descends an independent super-critical tree which has positive probability to survive, $\mathbb{P}[\Sigma_{N,3}^c]$ vanishes. □

REFERENCES

[1] ATHREYA, K. B. (1994). Large deviation rates for branching processes. I. Single type case. *Ann. Appl. Probab.* **4** 779–790. MR1284985

[2] BELTRÁN, J. and LANDIM, C. (2010). Tunneling and metastability of continuous time Markov chains. *J. Stat. Phys.* **140** 1065–1114. MR2684500

[3] BENJAMINI, I., KOZMA, G. and WORMALD, N. (2006). The mixing time of the giant component of a random graph. Available at arXiv:math/0610459v1.

[4] BEN AROUS, G. and ČERNÝ, J. (2005). Bouchaud's model exhibits two different aging regimes in dimension one. *Ann. Appl. Probab.* **15** 1161–1192. MR2134101

[5] BEN AROUS, G. and ČERNÝ, J. (2006). Dynamics of trap models. In *Mathematical Statistical Physics* 331–394. Elsevier, Amsterdam. MR2581889

[6] BEN AROUS, G. and ČERNÝ, J. (2007). Scaling limit for trap models on $\mathbb{Z}^d$. *Ann. Probab.* **35** 2356–2384. MR2353391

[7] BEN AROUS, G. and ČERNÝ, J. (2008). The arcsine law as a universal aging scheme for trap models. *Comm. Pure Appl. Math.* **61** 289–329. MR2376843

[8] BEN AROUS, G., ČERNÝ, J. and MOUNTFORD, T. (2006). Aging in two-dimensional Bouchaud's model. *Probab. Theory Related Fields* **134** 1–43. MR2221784

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

[10] ČERNÝ, J. (2006). The behaviour of aging functions in one-dimensional Bouchaud's trap model. *Comm. Math. Phys.* **261** 195–224. MR2193209

[11] ČERNÝ, J. and GAYRARD, V. (2008). Hitting time of large subsets of the hypercube. *Random Structures Algorithms* **33** 252–267. MR2436848

[12] ČERNÝ, J., TEIXEIRA, A. and WINDISCH, D. (2011). Giant vacant component left by a random walk in a random d-regular graph. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 929–968. MR2884219

[13] CUGLIANDOLO, L. F. and KURCHAN, J. (1993). Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model. *Phys. Rev. Lett.* **71** 173–176.

[14] DUDLEY, R. M. (2004). *Real Analysis and Probability*. Cambridge Univ. Press, Cambridge.

[15] DURRETT, R. (2010). *Probability: Theory and Examples*, 4th ed. Cambridge Univ. Press, Cambridge. MR2722836

[16] DURRETT, R. (2010). *Random Graph Dynamics*. Cambridge Univ. Press, Cambridge. MR2656427

[17] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York. MR0838085

[18] FONTES, L. R. G., ISOPI, M. and NEWMAN, C. M. (2002). Random walks with strongly inhomogeneous rates and singular diffusions: Convergence, localization and aging in one dimension. *Ann. Probab.* **30** 579–604. MR1905852

[19] 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.

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

[21] FOUNTOULAKIS, N. and REED, B. A. (2007). Faster mixing and small bottlenecks. *Probab. Theory Related Fields* **137** 475–486. MR2278465

[22] FOUNTOULAKIS, N. and REED, B. A. (2008). The evolution of the mixing rate of a simple random walk on the giant component of a random graph. *Random Structures Algorithms* **33** 68–86. MR2428978

[23] GRIMMETT, G. and KESTEN, H. (1984). Random electrical networks on complete graphs II: Proofs. Available at arXiv:math/0107068.

[24] JARA, M., LANDIM, C. and TEIXEIRA, A. (2011). Quenched scaling limits of trap models. *Ann. Probab.* **39** 176–223. MR2778800

[25] LAWLER, G. F. (1991). *Intersections of Random Walks*. Birkhäuser, Boston, MA. MR1117680

[26] LEVIN, D. A., PERES, Y. and WILMER, E. L. (2009). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. MR2466937

[27] LE CAM, L. (1960). An approximation theorem for the Poisson binomial distribution. *Pacific J. Math.* **10** 1181–1197. MR0142174

[28] LUBOTZKY, A., PHILLIPS, R. and SARNAK, P. (1988). Ramanujan graphs. *Combinatorica* **8** 261–277. MR0963118

[29] LUNDGREN, L., SVEDLINDH, P., NORDBLAD, P. and BECKMAN, O. (1983). Dynamics of the relaxation-time spectrum in a CuMn spin-glass. *Phys. Rev. Lett.* **51** 911–914.

[30] LYONS, R. and PERES, Y. (2011). *Probability on Trees and Networks*. Cambridge Univ. Press, Cambridge.

[31] PEIXOTO, G. R. C. (2011). Um estudo sobre o processo K não homogêneo. Master thesis, Univ. São Paulo.

[32] SALOFF-COSTE, L. (1997). Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1996). *Lecture Notes in Math.* **1665** 301–413. Springer, Berlin. MR1490046

[33] TEIXEIRA, A. and WINDISCH, D. (2011). On the fragmentation of a torus by random walk. *Comm. Pure Appl. Math.* **64** 1599–1646. MR2838338

M. JARA
A. TEIXEIRA
IMPA
ESTRADA DONA CASTORINA 110
CEP 22460 RIO DE JANEIRO
BRASIL
E-MAIL: mjara@impa.br
augusto@impa.br

C. LANDIM
IMPA
ESTRADA DONA CASTORINA 110
CEP 22460 RIO DE JANEIRO
BRASIL
AND
UNIVERSITÉ DE ROUEN
CNRS UMR 6085
AVENUE DE L'UNIVERSITÉ, BP.12
TECHNOPÔLE DU MADRILLET
F76801 SAINT-ÉTIENNE-DU-ROUVRAY
FRANCE
E-MAIL: landim@impa.br