Random Walks on Directed Covers of Graphs

Lorenz A. Gilch*, Sebastian Müller†

Institut für mathematische Strukturtheorie
Graz University of Technology
Steyrergasse 30/III
A-8010 Graz
Austria

July 28, 2009

Abstract

Directed covers of finite graphs are also known as periodic trees or trees with finitely many cone types. We expand the existing theory of directed covers of finite graphs to those of infinite graphs. While the lower growth rate still equals the branching number, upper and lower growth rates do not longer coincide in general. Furthermore, the behaviour of random walks on directed covers of infinite graphs is more subtle. We provide a classification in terms of recurrence and transience and point out that the critical random walk may be recurrent or transient. Our proof is based on the observation that recurrence of the random walk is equivalent to the almost sure extinction of an appropriate branching process. Two examples in random environment are provided: homesick random walk on infinite percolation clusters and random walk in random environment on directed covers. Furthermore, we calculate, under reasonable assumptions, the rate of escape with respect to suitable length functions and prove the existence of the asymptotic entropy providing an explicit formula which is also a new result for directed covers of finite graphs. In particular, the asymptotic entropy of random walks on directed covers of finite graphs is positive if and only if the random walk is transient.

KEYWORDS: trees, random walk, recurrence, transience, upper Collatz-Wielandt number, branching process, rate of escape, asymptotic entropy

AMS 2000 MATHEMATICS SUBJECT CLASSIFICATION: 05C05, 60J10 , 60F05, 60J85

1 Introduction

Suppose we are given a connected, directed graph G with vertex set V, edge set E, and root i_0. We construct a labelled tree $\mathcal{T}$ from G. We start with the root that is labelled with i_0.

*email: gilch@TUGraz.at; supported by German Research Foundation (DFG) grant GI 746/1-1

†email: dr.sebastian.mueller@gmail.com; supported by German Research Foundation (DFG) grant MU 2868/1-1

Recursively, if x is a vertex in the tree with label $i \in V$, then x has $d(i,j)$ successors with label j if and only if there are $d(i,j)$ edges from i to j in G. The tree $\mathcal{T}$ is called the *directed cover* of G. This model generalizes previously investigated ones, namely random walks on directed covers of finite graphs. These trees are also known as *periodic trees*, compare with Lyons [17], or *trees with finitely many cone types*, compare with Nagnibeda and Woess [20].

If G is finite the directed cover has the property that the growth rate exists and equals the branching number of $\mathcal{T}$, compare with Lyons [18]. Lyons [17] investigated the behaviour of homesick random walks on those trees. Takacs [23] computed a formula for the rate of escape of homesick random walks. Nagnibeda and Woess [20] studied more general random walks on these trees and gave a criterion for transience, null-recurrence, and positive recurrence. This criterion depends on the largest eigenvalue of a positive matrix arising from the transition probabilities. They also computed, among other things, a formula for the rate of escape of random walks on directed covers of finite graphs. It is worth mentioning that the model of random strings discussed in Gairat et al. [12] offers a different point of view of this model.

Most of the arguments in the finite case are based on the existence of the Perron-Frobenius eigenvalue of appropriate non-negative matrices. In the infinite setting, the matrices become non-negative operators and the existence of a largest eigenvalue can no longer be guaranteed. If there exists a largest eigenvalue with positive left and right eigenvectors the study is analogous to the finite case, but in general the behaviour becomes more subtle. In particular, the lower and upper growth rates of the cover, that are defined by $\liminf_{n\to\infty} |\mathcal{T}_n|^{1/n}$ and $\limsup_{n\to\infty} |\mathcal{T}_n|^{1/n}$, where $\mathcal{T}_n$ is the number of vertices in $\mathcal{T}$ at height n, are no longer equal, compare with Example 3.1. However, the lower growth rate and branching number coincide and equal the upper Collatz-Wielandt number of the adjacency matrix of G, see Theorem 3.1.

The first main result, Theorem 3.3, is the classification of random walks on directed covers according to their transience and recurrence behaviour. This result is given in terms of a bounded operator M that describes the relation between *forward* and *backward* probabilities of the random walk. We show that if the upper Collatz-Wielandt number, $\lambda^+(M)$, of this operator is smaller than 1 the process is recurrent and if it is greater than 1 it is transient. In the critical case, $\lambda^+(M) = 1$, the random walk may be recurrent or transient, compare with Subsection 3.3 and Example 3.2. This is in contrast with the finite setting, where the critical random walk is recurrent; e.g. see Nagnibeda and Woess [20]. The idea of our proof is based on the observation that recurrence of the random walk is equivalent to a.s. global extinction of an appropriate infinite-type Galton-Watson process, compare with Theorem 3.2. Therefore, our approach gives an alternative proof, that uses standard results for multi-type Galton-Watson processes, for random walks on directed covers of finite graphs. In Subsection 3.3 we study several examples where a complete classification is given.

Another result is that the rate of escape w.r.t. different appropriate length functions exists under reasonable assumptions. For this purpose, we assume the spectral radius to be strictly smaller than 1 and positive recurrence of a new Markov chain on G, which describes the end of the tree to which the random walk on $\mathcal{T}$ converges. As in Nagnibeda and Woess [20], the existence of the rate of escape can not be shown by straight-forward arguments using Kingman's subadditive ergodic theorem. Thus, we use modified exit times from [20] to prove existence of the rate of escape with respect to different length functions, see Theorem 3.8. This enables us to prove another main result of the paper, namely the existence of the asymptotic entropy $h = \lim_{n\to\infty} \mathbb{E}\bigl[-\frac{1}{n}\log \pi_n(X_n)\bigr]$ under reasonable assumptions, where $(X_n)_{n\in\mathbb{N}_0}$ is a random walk on a directed cover and π_n is its distribution at time n. This result, whose proof envolves generating functions techniques, is also new for the case when G is finite. While it is well-known that entropy (introduced by Avez [1]) exists for random walks on groups existence for random walks on other structures is not known a priori. For more details about entropy of random walks on groups see Kaimanovich and Vershik [16] and Derriennic [8]. In particular, we show that the asymptotic entropy equals the rate of escape with respect to a distance function in terms of Green functions. This also implies that the asymptotic entropy equals the rate of escape with respect to the Green metric (introduced by Blachère and Brofferio [5]), which is given by $d_G(x,y) = -\log F(x,y)$, where $F(x,y)$ is the probability of ever hitting y when starting at x. The technique of our proof was motivated by Benjamini and Peres [2], where it is shown that for random walks on finitely generated groups the entropy equals the rate of escape w.r.t. the Green metric. Blachère, Haïssinsky and Mathieu [6] generalized this result to random walks on countable groups. We also want to mention the work of Björklund [4], who gave an interpretation of the Green metric in terms of Hilbert metrics. Our result also includes an explicit formula for the entropy and shows that the entropy can be computed along almost every sample path. Furthermore, we get convergence in L_1 of $-\frac{1}{n}\log \pi_n(X_n)$ to h, see Theorem 3.11.

The paper is organized as follows. In Section 2 we give the basic notations and definitions concerning graphs and trees. The main results, Theorem 3.3, Theorem 3.8, and Theorems 3.9, 3.11, together with examples and discussions are presented in Section 3. All proofs are given in Section 4.

2 Preliminaries

Most of the time we follow the notation of [18], where the reader can find the basic definitions and results concerning graphs and random walks on trees. For general information on non-negative matrices we refer to [22] and on Banach lattices and positive operators to [21].

2.1 Notations and Definitions

Let $G = (V, E)$ be a directed graph with countable vertex set V, edge set E and root o. For ease of presentation, we also identify a graph with its vertex set, i.e., $x \in G$ means $x \in V$. The *adjacency matrix* $A = \bigl(a(i,j)\bigr)_{i,j\in G}$ of G is defined by $a(i,j) = 1$ if there is an edge from i to j and 0 otherwise. A graph is called *locally finite* if the row and column sums of A are finite. A graph is (strongly) *connected* if there is a directed path from every vertex to any other vertex, and it has *bounded geometry* if the vertex degrees are uniformly bounded. A *tree* $\mathcal{T}$ is an undirected graph in which every two vertices are connected by exactly one shortest path. A *rooted tree* is a tree with a distinguished vertex, the root o. This endows the tree with a natural orientation: towards or away from the root. Furthermore, denote by $|x|$ the natural distance from a vertex x of $\mathcal{T}$ to o, i.e., the length of the shortest path (we call such a shortest path also a *geodesic*) from o to x. Let $\mathcal{T}_n$ be the set of vertices at distance n from the root o. Every vertex $x \in \mathcal{T}_n$ has a unique geodesic $\langle o = x_0, x_1, \dots, x_n = x\rangle$ coming from the origin o. The vertex x_{n-1} is called the *ancestor* x^- of x and x is called the *direct descendent* or *successor* of x^-. In general, a vertex y is a *descendent* of x if x lies on the unique geodesic from o to y. A *ray* $\xi = \langle o = x_0, x_1, \dots\rangle$ is an infinite path from o to infinity that doesn't backtrack, i.e., $x_i \neq x_j$ for all $i \neq j$. We call the set of all rays of $\mathcal{T}$ the *(end) boundary* of $\mathcal{T}$, denoted by $\partial\mathcal{T}$. There is a natural metric on $\partial\mathcal{T}$: if two rays ξ and η have exactly n edges in common their distance is defined as $d(\xi,\eta) := e^{-n}$. A *flow* θ is a non-negative function on the vertices of $\mathcal{T}$ such that $\theta(x) = \sum_{y\in\mathcal{T}, y^-=x}\theta(y)$. A flow θ is called a *unit flow* if $\theta(o) = 1$. The *lower* and *upper growth rates* of a tree $\mathcal{T}$ are defined as

$$\underline{gr}(\mathcal{T}) := \liminf_{n\to\infty} |\mathcal{T}_n|^{1/n} \quad \text{and} \quad \overline{gr}(\mathcal{T}) := \limsup_{n\to\infty} |\mathcal{T}_n|^{1/n},$$

where $|\mathcal{T}_n|$ is the cardinality of the set $\mathcal{T}_n$. If these numbers are equal we speak of the *growth rate* $gr(\mathcal{T}) = \lim_{n\to\infty}|\mathcal{T}_n|^{1/n}$ of $\mathcal{T}$. Another method to measure the growth of a tree is the *branching number* $br(\mathcal{T})$. We recall two possible definitions. The first uses the concept of flows while the second uses the Hausdorff dimension $\dim\partial\mathcal{T}$ of the boundary $\partial\mathcal{T}$ of the tree $\mathcal{T}$:

$$\begin{aligned} br(\mathcal{T}) &:= \sup\left\{\lambda > 0 \,\middle|\, \exists \text{ flow } \theta\ \forall x\in\mathcal{T}:\ 0 \le \theta(x) \le \lambda^{-|x|}\right\} \\ &:= \exp\dim\partial\mathcal{T}. \end{aligned}$$

Let us remark that there is the following general connection between the lower growth rate and the branching number

$$br(\mathcal{T}) \le \underline{gr}(\mathcal{T}). \tag{1}$$

2.2 Non-negative Infinite Matrices

Let $M := \big(m(x,y)\big)_{x,y\in G}$ be an infinite matrix with non-negative entries. For $n \in \mathbb{N}$, let $M^n = \big(m^{(n)}(x,y)\big)_{x,y\in G}$ be the n-th matrix power of M and set $M^0 := I$, the identity matrix over $\mathbb{N}$. A non-negative matrix M is called irreducible if for all $x, y \in G$ there exists some $k \in \mathbb{N}$ such that $m^{(k)}(x,y) > 0$. We will assume throughout the paper that there exist constants $c, C > 0$ such that

$$0 < c < \sum_{y\in G} m(x,y) < C < \infty \qquad \text{for all } x \in G. \tag{2}$$

Due to the upper bound the matrix M can be interpreted as a bounded linear operator on ℓ_p, $p \in [1,\infty]$. For each $f \in \ell_p$, let $Mf(x) = \sum_{y\in G} m(x,y)f(y)$. The spectrum of M is $\sigma(M,\ell_p) := \big\{\lambda \in \mathbb{C} \mid \lambda I - T \text{ is not a bijection of } \ell_p\big\}$. It is compact and non-void. Denote by

$$r_p(M) := \sup\big\{|\lambda| \,\big|\, \lambda \in \sigma(M,\ell_p)\big\} = \lim_{n\to\infty}\sqrt[n]{\|M^n\|_p}$$

the ℓ_p-spectral radius of M. Observe that in general the spectral radius is not an eigenvalue, and hence no direct generalization of the Perron-Frobenius eigenvalue from the finite dimensional case exists. A priori it is not clear if some ℓ_p-spectral radius may serve as an analogon for the Perron-Frobenius eigenvalue in the finite dimensional setting. Indeed, the following characteristic turns out to be appropriate:

$$\lambda^+(M) := \sup\bigl\{\lambda > 0 \,\big|\, \exists\, 0 < f \in \ell_\infty :\ Mf \geq \lambda f\bigr\}.$$

This number is also known as the *upper Collatz-Wielandt number*, compare with [11]. The number $\lambda^+(M)$ is well-defined, since under the general assumption (2) we have

$$c \leq \inf_{x\in G} \sum_{y\in G} m(x,y) \leq \lambda^+(M) \leq \sup_{x\in G} \sum_{y\in G} m(x,y) \leq C.$$

The lower bound for $\lambda^+(M)$ is obtained by investigating $M\mathbf{1}$, where $\mathbf{1}$ is the vector with all entries equal to 1. The upper bound is obvious, since with $f \in \ell_\infty$ such that $\|f\|_\infty = 1$ we obtain $Mf(x) \leq \sum_{y\in G} m(x,y)$. Furthermore, it is easy to see that the upper Collatz-Wielandt number is less than or equal to the ℓ_∞-spectral radius of M, that is,

$$\lambda^+(M) \leq r_\infty(M).$$

If M is homogeneous in the sense of quasi-transitiveness, we have that $\lambda^+(M) = r_\infty(M)$, compare with Example 3.3.2. We also refer to [24] where several other relations between the ℓ_p-spectral radii and $\lambda^+(M)$ for symmetric matrices are discussed.

Remark 2.1. The sup in the definition of $\lambda^+(M)$ may or may not be attained. Consider the adjacency matrix $A_\mathbb{Z}$ of the graph $G = \mathbb{Z}$. By the latter we mean the graph G with $V = \mathbb{Z}$ and $E = \bigl\{(x,y) \in \mathbb{Z}^2 \mid |x-y| = 1\bigr\}$. Clearly, $\lambda^+(A_\mathbb{Z}) = 2$ and the sup is attained with the vector $\mathbf{1}$. On the other hand, consider $G = \mathbb{N}$ with its adjacency matrix $A_\mathbb{N}$. We still have that $\lambda^+(A_\mathbb{N}) = 2$. This can be seen using a recurrence argument or observing that $r_2(A_\mathbb{N}) = 2$ and $r_2(A_\mathbb{N}) \leq r_\infty(A_\mathbb{N}) \leq 2$. Assume that $\lambda^+(A_\mathbb{N})$ is attained using the function f. Since $f(1) \geq 2f(0)$ and $f(n+1) \geq 2f(n) - f(n-1)$ for all $n \geq 1$ we see that f is unbounded and hence obtain a contradiction.

Remark 2.2. Another characteristic that might serve as a generalization of the Perron-Frobenius eigenvalue to the infinite setting is $\rho(M) := \limsup_{n\to\infty} \bigl(m^{(n)}(x,y)\bigr)^{1/n}$. The latter is independent of the specific choice of x, y and equals the ℓ_2-spectral radius if M is symmetric, e.g. compare with [25, Chapter II, Section 10]. It can be seen analogously to von Below [24, Corollary 4.6] that $\rho(M) \leq \lambda^+(M)$. Furthermore, $\rho(M)$ can be given in terms of convergence parameters of Green functions, rate functions of large deviation principles, and super-harmonic functions. Moreover, $\rho(M)$ depends on local properties, since

$$\rho(M) = \sup_{F\subset G, |F|<\infty} \rho(M_F), \tag{3}$$

where $M_F = \bigl(m_F(x,y)\bigr)_{x,y\in F}$ is the matrix induced by F, i.e., $m_F(x,y) = m(x,y)$ for $x, y \in F$ and 0 otherwise. This fact indicates that $\rho(M)$ is not the *good* characteristic for recurrence and transience, since these properties describe a global behaviour.

Remark 2.3. Let A be the adjacency matrix of a graph G, which is symmetric, that is, there is an edge from $i \in G$ to $j \in G$ if and only if there is an edge from j to i. While in the finite case we always have $r_2(A) = r_\infty(A)$, in general we only have $r_2(A) \leq r_\infty(A)$ for infinite A. Let $G = \mathcal{T}_d$ be the regular tree with degree $d \geq 3$. Clearly, we have $r_\infty = d$ and it is very well-known that $r_2(A) = 2\sqrt{d-1}$. It is worth noting that if $r_2(A) < r_\infty(A)$ we can not approximate $r_\infty(A)$ using Perron-Frobenius eigenvalues of finite subgraphs, compare with Equation (3).

3 Results

3.1 Trees as Directed Covers of Infinite Graphs of Bounded Geometry

Suppose we are given a connected and directed graph $G = (V, E)$ of bounded geometry without multiple edges. In some cases, e.g. Remark 3.3, we can drop the assumption of bounded geometry. Let $i_0 \in V$ be a distinguished vertex of G. The *directed cover* $\mathcal{T}$ of G (rooted in i_0) is defined recursively as a rooted tree $\mathcal{T}$, whose vertices are labelled by the vertex set V. The root o of $\mathcal{T}$ has label i_0; recursively, if $x \in \mathcal{T}$ is labelled with i, then x has one direct descendent with label j if and only if there is an edge from i to j in G. And vice versa we define the *label function* $\tau : \mathcal{T} \to V$ to be the function that associates to each vertex in $\mathcal{T}$ its label in V. In order to distinguish between the two graphs G and V we use in general the variables i, j for vertices in G (and labels in T) and x, y for the vertices of $\mathcal{T}$. We refer to [17] and [20] for more details and references for finite graphs G. In the finite case these trees are known as *periodic trees* ([17], [23]) or *trees with finitely many cone types* ([20]).

We remark that we may assume w.l.o.g. that the graph has no multiple edges. Indeed, suppose that G has multiple edges. Then one can replace G by a new graph G', which arises from G by introducing new vertices and edges and deleting those multiple edges. That is, if there are $m > 2$ edges from i to j in G, then we add to G the new vertices $j_1, \ldots, j_{m-1}$ and add an edge from i to each of these new vertices. Each of these new vertices is again connected with every vertex which has an edge coming from j. Repeating this operation gives a new graph G', which has now no multiple edges.

Note that there is a one-to-one correspondence between all finite paths $\langle i_0, i_1, i_2, \ldots i_n \rangle$ in G starting at i_0 and all vertices in $\mathcal{T}$. Each vertex $x_n \in \mathcal{T}$ at height n corresponds to a unique geodesic $\langle o, x_1, x_2, \ldots, x_n \rangle$ coming from the origin. This geodesic is uniquely determined by its labels, i.e., the path $\langle i_0, \tau(x_1), \tau(x_2), \ldots, \tau(x_n) \rangle$ in G.

The *cone* $\mathcal{T}^x := \{y \in \mathcal{T} \mid y \text{ is a descendent of } x\}$ is the subtree rooted at x and spanned by all vertices y such that x lies on the geodesic from o to y. We say that $\mathcal{T}^x$ has *cone type* $\tau(x)$. Observe that if $x, y \in \mathcal{T}$ with $\tau(x) = \tau(y)$ then the trees $\mathcal{T}^x$ and $\mathcal{T}^y$ are isomorphic as rooted trees.

For general trees we have $br(\mathcal{T}) \leq \underline{gr}(\mathcal{T})$. While for directed covers of finite graphs $br(\mathcal{T}) = gr(\mathcal{T})$ the growth rate does not exist in general, compare with Example 3.1. Nevertheless we have the following result for directed covers of infinite graphs.

Theorem 3.1. *Let A be the adjacency matrix of G and $\mathcal{T}$ be the directed cover of G. Then:*

1. $\lambda^+(A) = br(\mathcal{T}) = \underline{gr}(\mathcal{T})$.
2. $\overline{gr}(\mathcal{T}) \le r_\infty(A)$.

Example 3.1. We give an example where the growth rate does not exist. The desired graph G will be constructed inductively. To this end, let $(k_n)_{n\in\mathbb{N}}$ be the sequence defined by $k_1 := 2$ and $k_{n+1} := 3\sum_{j=1}^{n} k_j$. For n odd, let G_n be the *circle* with k_n vertices $\{1, 2, \ldots, k_n\}$ and directed edges $(i, i+1)$ for $1 \le i < k_n$ and $(k_n, 1)$. We call 1 the starting point and k_n the end point of the circle. For n even, G_n is constructed as follows. Let $T_2^{k_n-1}$ be a rooted binary tree of height $k_n - 1$ with directed edges all leading away from the root. The leaves at level $k_n - 1$ are enumerated by $\{1, 2, \ldots, 2^{k_n-1}\}$. Now we add direct paths of length k_{n+1} from leaf i to leaf $i+1$ for $1 \le i < 2^{k_n-1}$ and one directed path from leaf 2^{k_n-1} to o_n of length k_{n+1}. Here, o_n is an additional vertex with a directed edge from o_n to the root of $T_2^{k_n-1}$. Observe that all graphs G_n are connected. The graph G is now defined inductively. We start with G_1. We add a copy of G_2 in *gluing* (vertices are identified) the end point of G_1 with o_2 of G_2. At each leave of G_2 we glue a copy of G_3 and continue inductively. Observe that each vertex has outdegree at most 2 and the vertices (except the starting and end point) in copies of the circle G_n (n odd) have only outdegree 1 as well as the vertices on the directed paths leading away from the leaves of G_n (n even). Let $\beta_n = \sum_{i=1}^{n} k_i$. Therefore, for n even we find that

$$|T_{\beta_n}| \ge 1^{\frac{\beta_n}{4}} 2^{\frac{3\beta_n}{4}}.$$

Due to the construction of G every path emanating from the origin of length β_n for n odd has at most $\beta_n/2$ vertices of degree 2. This yields for n odd

$$|T_{\beta_n}| \le 2^{\frac{\beta_n}{2}}.$$

Eventually,

$$\underline{gr}(\mathcal{T}) \le 2^{\frac{1}{2}} < 2^{\frac{3}{4}} \le \overline{gr}(\mathcal{T}). \tag{4}$$

3.2 Recurrence and Transience of Random Walks on Directed Covers

We consider the model of nearest neighbour random walks on $\mathcal{T}$ according to [20]. Suppose we are given transition probabilities $p_G(i,j)$ on G, where $(i,j) \in E$. We hereby assume that $p_G(i,j) > 0$ if and only if there is an edge from i to j in G. Furthermore, suppose we are given *backward* probabilities $p(-i) \in (0,1)$ for each $i \in G$. For ease of presentation and technical reasons, we add to $\mathcal{T}$ a loop at the origin o. Then the random walk on the tree $\mathcal{T}$ is defined through the following transition probabilities $p(x,y)$, where $x, y \in \mathcal{T}$:

$$p(o,y) := \begin{cases} \big(1 - p(-i_0)\big) p_G\big(i_0, \tau(y)\big), & \text{if } y \ne o, \\ p(-i_0), & \text{if } y = o, \end{cases} \tag{5}$$

and for $x \neq o$ with $\tau(x) = i$,

$$p(x,y) := \begin{cases} \big(1 - p(-i)\big) p_G\big(i, \tau(y)\big), & \text{if } x = y^-, \\ p(-i), & \text{if } y = x^-. \end{cases} \tag{6}$$

We will also write $p(i,j) := \big(1 - p(-i)\big) p_G(i,j)$ and $p(-i) := p(x, x^-)$, where $\tau(x) = i$, for the transition probabilities in $\mathcal{T}$. The random walk on $\mathcal{T}$ starting at o is denoted by the sequence of random variables $(X_n)_{n \in \mathbb{N}_0}$. We are going to characterize recurrence and transience in terms of $\lambda^+(M)$, where

$$M = \big(m(i,j)\big)_{i,j \in G} \text{ with } m(i,j) := \frac{p(i,j)}{p(-i)}. \tag{7}$$

Observe that our notation differs from the one in [20] where the finite analogue of the matrix M is denoted by A. In [20] it is proved that for finite G the random walk on $\mathcal{T}$ is positive recurrent if $\lambda(M) < 1$, null-recurrent if $\lambda(M) = 1$, and transient if $\lambda(M) > 1$; here $\lambda(M)$ is the Perron-Frobenius eigenvalue of M. This classification was also obtained in [12] for the more general model of random strings. The matrix used in [12] is different from the one in [20] but leads to the same results.

As an additional assumption we demand that there is some $\varepsilon \in (0,1)$ such that

$$\varepsilon < p(-i) < 1 - \varepsilon \quad \text{for all } i \in G. \tag{8}$$

This assumption is necessary to exclude degenerate examples and it assures assumption (2) on the infinite matrix M.

Due to the tree structure the random walk is reversible with reversible measure m defined recursively by

$$m(o) := 1 \text{ and } m(x) := m(x^-) \frac{p(x^-, x)}{p(x, x^-)} \text{ if } x \neq o. \tag{9}$$

Hence, the random walk can be considered as an electric network, see [9] or [18] for general information on electric networks. Recall that transience of a subnetwork implies transience of the larger network, compare with Chapter 2 in [18]. Let

$$\rho(M) := \limsup_{n \to \infty} \Big(m^{(n)}(x,y)\Big)^{1/n}$$

be the spectral radius of M. If $\rho(M) > 1$ then there exists a finite set F such that $\lambda(M_F) > 1$ and hence we conclude from the results of [20] that the walk restricted to the cover of F is transient. Therefore, we have that $\rho(M) > 1$ implies transience of the random walk. The reverse is not true in general, compare with Theorem 3.3 and Remark 2.2.

Since the methods used in [12] and [20] cannot directly be generalized to the infinite setting we present a new approach for proving recurrence and transience. This provides an alternative proof in the finite case. The idea is to couple an infinite-type generalized Galton-Watson process (in continuous time) to the random walk (in discrete time) such that the Galton-Watson process dies out if and only if the random walk on $\mathcal{T}$ visits the edge (o, o). To this end observe that M is a non-negative (infinite) matrix. Thus, M can be interpreted as the first moment matrix of an infinite-type Galton-Watson process. That means that $m(i,j)$ is the mean number of particles of type j that one particle of type i produces in its lifetime. Let us define a process Z_t with types indexed by G with first moment matrix M. The process is described through the number $Z_t(i)$ of particles of type i at time t and evolves according the following rules: for each $i \in G$

$$\begin{aligned} Z_t(i) &\to Z_t(i) - 1 \text{ at rate } Z_t(i)p(-i), \\ Z_t(i) &\to Z_t(i) + 1 \text{ at rate } \sum_j Z_t(j)p(j,i). \end{aligned}$$

In words, each particle of type i dies at rate $p(-i)$ and gives birth to new particles of type j at rate $p(i,j)$. Let $Z_t = \big(Z_t(i)\big)_{i\in G}$ describe the whole population of the process. We consider the probability of (global) extinction $q := \mathbb{P}[\exists t : Z_t = \mathbf{0}]$ and make the following crucial observation:

Theorem 3.2. *The extinction probability q of the process Z_t equals the probability that the random walk on $\mathcal{T}$ visits the edge (o,o).*

Now, we use a result on infinite-type Galton-Watson processes of [3] in order to obtain the classification result:

Theorem 3.3. *The random walk $(X_n)_{n\in\mathbb{N}_0}$ on $\mathcal{T}$ is recurrent if $\lambda^+(M) < 1$ and it is transient if $\lambda^+(M) > 1$. In the critical case, $\lambda^+(M) = 1$, it may be transient or recurrent.*

Example 3.2. We consider Example 3 of [3] that was given in terms of infinite-type Galton-Watson processes. Let $G := \mathbb{N}_0$ with edges of the form $(i,i+1)$ and $(i,i-1)$ for $i \geq 1$, including the edge $(0,1)$. Let $p(-0) := 1/3,\ p(0,1) := 2/3$ and

$$p(-i) := \left(1 + \left(1 + \frac{1}{i}\right)^2 + \left(\frac{1}{3}\right)^i\right)^{-1},$$

$$p(i,i+1) := p(-i)\left(1 + \frac{1}{i}\right)^2 \text{ and } p(i,i-1) := p(-i)\left(\frac{1}{3}\right)^i \text{ for } i \geq 1$$

define the random walk on $\mathcal{T}$. Hence the values of M are $m(0,1) = 2$ and

$$m(i,i+1) = \left(1 + \frac{1}{i}\right)^2 \text{ and } m(i,i-1) = \left(\frac{1}{3}\right)^i \text{ for } i \geq 1$$

and 0 otherwise. Observe now that the function $g : \mathbb{N}_0 \to \mathbb{R}$ defined by $g(0) := 1$ and $g(i) := i/(i+1)$ for $i \geq 1$ is a solution for the inequality $Mg \geq g$. Furthermore, one can easily show by induction that $Mg \geq \lambda g$ with $\lambda > 1$ implies that either $g \equiv 0$ or $g(i) \to \infty$ for $i \to \infty$. Eventually, $\lambda^+(M) = 1$. In order to show that the random walk on $\mathcal{T}$ is transient we use a coupling argument. We compare the original process with the Markov chain on the positive integers with transition probabilities $\tilde{p}(0,1) := 1$, and $\tilde{p}(i,i+1) := p(i,i+1)$, $\tilde{p}(i,i) := p(i,i-1)$, and $\tilde{p}(i,i-1) := p(-i)$ for $i \geq 1$. Hence, using a coupling argument, the random walk on the directed cover is transient if the Markov chain $(\tilde{X}_n)_{n\in\mathbb{N}_0}$ on $\mathbb{N}_0$ with transition probabilities $\tilde{p}(\cdot,\cdot)$ is transient. To see the latter, observe that the mean drift is

$$\mu_1(i) := \mathbb{E}\big[\tilde{X}_{n+1} - \tilde{X}_n \big| \tilde{X}_n = i\big] = p(-i)\left(\frac{2}{i} + \frac{1}{i^2}\right)$$

and

$$\mu_2(i) := \mathbb{E}\big[(\tilde{X}_{n+1} - \tilde{X}_n)^2 \big| \tilde{X}_n = i\big] \leq 1.$$

Hence, for i sufficiently large we obtain

$$\mu_1(i) \geq \frac{2}{3i} \geq \frac{2\,\mu_2(i)}{3i}$$

and conclude from Theorem 3.6.1 in [10] that the random walk is transient.

Example 3.3. For any given graph G, a straightforward example for recurrence in the critical case $\lambda^+(M) = 1$ is given by $p(-i) := 1/2$ for every $i \in G$. The random walk on $\mathcal{T}$ can then be naturally projected on $\mathbb{N}_0$, that is, X_n will be projected on $|X_n|$. In this case, the Markov chain $(|X_n|)_{n\in\mathbb{N}_0}$ is null recurrent, and thus, $(X_n)_{n\in\mathbb{N}_0}$ is also null-recurrent.

3.3 Examples

3.3.1 Homesick Random Walks

A special class of random walks on trees are homesick random walks, compare with [18] and [23]. In this model the edge leading back towards the root is λ times as likely to be taken as each other edge, that is, $p(x, x^-) = \lambda/(\lambda + deg(x) - 1)$ and $p(x^-, x) = 1/(\lambda + deg(x^-) - 1)$, where $deg(x)$ is the number of edges adjacent to x. We denote the dependence of the random walk on the parameter λ by RW_λ. In [17] it is shown that the RW_λ is recurrent if $\lambda > br(\mathcal{T})$ and transient if $\lambda < br(\mathcal{T})$. This result holds for any tree $\mathcal{T}$, not necessarily directed cover.

Given an underlying simple random walk on a graph G the homesick random walk is the random walk on the directed cover with $p(-i) = \lambda/(\lambda + outdeg(i))$, where $outdeg(i) = \sum_j a(i,j)$ is the number of outgoing edges from $i \in G$. It is easy to see that $\lambda^+(M) = \lambda^+(A)/\lambda$, where A is the adjacency matrix of G and M is defined as in Equation (7). As a consequence of Theorem 3.3 we obtain the following:

Corollary 3.4. *Let A be the adjacency matrix and $\mathcal{T}$ the directed cover of G. Then the homesick random walk RW_λ on $\mathcal{T}$ is recurrent if $\lambda > \lambda^+(A)$ and transient if $\lambda < \lambda^+(A)$.*

While $\lambda = \lambda^+(A)$ implies recurrence if G is finite (compare with Theorem 3.5 in [18]), the behaviour for infinite graphs G is not known in general.

3.3.2 Directed Covers of Quasi-Transitive Graphs

Suppose we are given a locally finite graph G with bounded geometry. Denote by $\mathrm{AUT}(G)$ the group of all automorphisms γ of the vertex set of G which leave the adjacency relation invariant, that is, there is an edge from $\gamma(i)$ to $\gamma(j)$ if and only if there is an edge from $i \in G$ to $j \in G$. Assume now that G is *quasi-transitive*, that is, $\mathrm{AUT}(G)$ acts with finitely many orbits on the vertex set of G. We write $\mathrm{Orb} := \{o_1, \ldots, o_r\}$ for the set of orbits. We construct a new finite graph G' with vertex set Orb in the following natural way: there are $d(i,j) \in \mathbb{N}_0$ edges from o_i to o_j if and only if there are $d(i,j)$ edges in G from some $k \in o_i$ to some $l \in o_j$. Now we can apply the well-known results from the finite setting, since G' and G create the same (unlabelled) cover $\mathcal{T}$. Thus, $br(\mathcal{T})$, which becomes the largest eigenvalue of the adjacency matrix of G', equals $\lambda^+(A)$. Furthermore, it is easy to see that for quasi-transitive graphs we have $\lambda^+(A) = r_\infty(A)$.

Now, suppose we are given a random walk on G governed by the transition matrix P_G. The set $\mathrm{AUT}(G, P_G)$ is the group of all automorphisms γ of the vertex set of G, which leave P_G invariant, that is, $p_G\big(\gamma(i), \gamma(j)\big) = p_G(i,j)$. Then (G, P_G) is called *quasi-transitive* if $\mathrm{AUT}(G, P_G)$ acts with finitely many orbits on the vertex set of G. Assume now that (G, P_G) is quasi-transitive and that the backward probabilities $p(-i)$ are constant on the orbits of $\mathrm{AUT}(G, P_G)$. Denote by $\mathrm{Orb} := \{o_1, \ldots, o_r\}$ the orbits of $\mathrm{AUT}(G, P_G)$. We define a random walk on Orb by setting $\tilde{p}(o_i, o_j) := \sum_{l \in o_j} p_G(k,l)$, where $k \in o_i$ is arbitrary. Then $\lambda^+(M)$ is the Perron-Frobenius of the matrix $\big(\tilde{p}(o_i, o_j)/p(-o_i)\big)_{1 \le i,j \le r}$. This follows from the results of [20]. In particular, the random walk is null-recurrent if $\lambda^+(M) = 1$.

3.3.3 Directed Covers of Percolation Clusters

We consider supercritical Bernoulli(p) percolation on $\mathbb{Z}^d$, i.e., for fixed $p \in [0,1]$, each edge is kept with probability p and removed otherwise, independently of the other edges. It is well-known that there exists a critical value p_c such that for $p < p_c$ there is almost surely no infinite connected component and for $p > p_c$ there is almost surely exactly one infinite connected component. In the latter case we denote by C_ω the infinite connected component for the realization or environment ω. We refer to §6 in [18] for more information on percolation models. Let A_ω be the adjacency matrix and $\mathcal{T}_\omega$ be the directed cover of the infinite cluster C_ω with respect to some $i_0 \in C_\omega$. We get the following classification:

Theorem 3.5. *The homesick random walk RW_λ on almost every directed cover $\mathcal{T}_\omega$ is recurrent if $\lambda \ge 1/2d$ and transient if $\lambda < 1/2d$.*

3.3.4 Random Walk on Directed Covers in Random Environment

We consider the nearest neighbour random walk in random environment on $V = \mathbb{Z}$ and edge set $E = \big\{(x,y) \mid |x-y| = 1\big\}$. We choose i.i.d. random variables ω_z^+ ($z \in \mathbb{Z}$) with values in $(0,1)$ and call η the distribution of the environment with one-dimensional marginal distribution θ. For a given realization ω of this random environment, we consider the Markov chain on $\mathbb{Z}$ with transition kernel P_ω defined as

$$p_{\mathbb{Z},\omega}(z, z+1) := \omega_z^+ \quad \text{and} \quad p_{\mathbb{Z},\omega}(z, z-1) := \omega_z^- := 1 - \omega_z^+ \qquad \text{for all } z \in \mathbb{Z}.$$

We refer to [26] for details on this model. In addition, we introduce an environment that defines the backwards probabilities. So let ν_z $(z \in \mathbb{Z})$ be i.i.d. random variables with values in $(\varepsilon, 1-\varepsilon)$ for some $\varepsilon \in (0,1)$. We call $\tilde{\eta}$ the distribution of this environment and denote by $\tilde{\theta}$ its one-dimensional marginal distribution. Every given realization $\nu = (\nu_z)_{z\in\mathbb{Z}}$ determines the backwards probabilities by $p_\nu(-z) := \nu_z$. Let Θ be the corresponding product measure with one-dimensional marginal $\theta\times\tilde{\theta}$. Every given realization (ω,ν) defines a random walk on a directed cover in random environment (RWDCRE) with corresponding matrix $M_{\omega,\nu}$. The classification in recurrence and transience will be stated in terms of the top Lyapunov exponent of sequences of random matrices. For $k \in \{1,2,3,\ldots\}$, we write

$$A_k := \begin{pmatrix} \nu_k\omega_k^+ & -\frac{\omega_k^-}{\omega_k^+} \\ 1 & 0 \end{pmatrix} \text{ and } \tilde{A}_k := \begin{pmatrix} \nu_k\omega_k^- & -\frac{\omega_k^+}{\omega_k^-} \\ 1 & 0 \end{pmatrix}. \tag{10}$$

Denote by γ_1 the top Lyapunov exponent associated with the sequence $(A_k)_{k\in\mathbb{N}}$, i.e.,

$$\gamma_1 = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\big[\ln \|A_n\cdots A_1\|\big],$$

where $\|\cdot\|$ is any matrix norm. Analogously, let $\tilde{\gamma}_1$ be the top Lyapunov exponent of the sequence $(\tilde{A}_k)_{k\in\mathbb{N}}$. For sake of better readability, we set $\mu_i^- = \omega_i^-/\nu_i$ and $\mu_i^+ = \omega_i^+/\nu_i$. Combining Theorem 3.2 with Theorem 2.6 and Theorem 2.9 in [13] yields the following classification. (Observe that μ_i^0 of [13] equals 0 in our special case.)

Theorem 3.6. *We have the following classification:*

1. *If there exists no $\lambda > 0$ such that $\mu_0^-\lambda^{-1} + \mu_0^+\lambda \le 1$ Θ-almost surely, then the RWDCRE is transient Θ-almost surely.*
2. *If there exists $\lambda > 1$ such that $\mu_0^-\lambda^{-1} + \mu_0^+\lambda \le 1$ Θ-almost surely, then the RWDCRE is transient Θ-almost surely if and only if*
$$\gamma_1 < \mathbb{E}\ln\left(\frac{\mu_0^-}{\mu_0^+}\right).$$
3. *If there exists $\lambda < 1$ such that $\mu_0^-\lambda^{-1} + \mu_0^+\lambda \le 1$ Θ-almost surely, then the RWDCRE is transient Θ-almost surely if and only if*
$$\tilde{\gamma}_1 < \mathbb{E}\ln\left(\frac{\mu_0^+}{\mu_0^-}\right).$$

Remark 3.1. Observe that Theorem 3.6 gives a complete classification since $\mu_0^- + \mu_0^+ = 1/\nu_0 > 1$.

3.4 Ergodicity

In order to prove (non-)ergodicity we investigate whether the reversible measure in Equation (9) is (in)finite. This method leads to the following generalization of the finite case.

Proposition 3.7. *The random walk on $\mathcal{T}$ is non-ergodic if $\lambda^+(M) > 1$ and it is ergodic if $r_\infty(M) < 1$. If $\lambda^+(M) = 1$ and the supremum in the definition of $\lambda^+(M)$ is attained the process is non-ergodic as well.*

Since in the infinite case the random walk is *more transient* in the critical case we conjecture the random walk to be non-ergodic for $\lambda^+(M) \geq 1$.

3.5 Rate of Escape

In this section we generalize the results of [20] to covers of infinite graphs and more general length functions. In order to state the main result we need the following definitions and notations. Suppose we are given a bounded function $w : G \times G \to \mathbb{R}$ representing a weight for each edge in G. We define recursively a length function l on $\mathcal{T}$ by $l(o) := 0$ and $l(x) := l(x^-)+w\bigl(\tau(x^-),\tau(x)\bigr)$ otherwise. If $w(\cdot,\cdot) = 1$, then $l(x)$ is just the natural graph distance on $\mathcal{T}$ (that is, the number of edges that connect o with x) denoted by $|x|$. Observe that w may take negative values; in this case one can think of w describing height differences between neighbour vertices. If there is some constant $\ell \in \mathbb{R}$ such that

$$\ell = \lim_{n\to\infty} \frac{l(X_n)}{n} \quad \text{almost surely,}$$

then ℓ is called the *rate of escape* or *drift* of $(X_n)_{n\in\mathbb{N}_0}$ w.r.t. the length function l. We suppose for the rest of this section that $(X_n)_{n\in\mathbb{N}_0}$ is *transient*. Let $i \in G$, $x, y \in \mathcal{T}$ with $\tau(x) = i$ and $z \in \mathbb{C}$. We introduce the following generating functions:

$$\begin{aligned}
F(-i|z) &:= \sum_{n\geq 0} \mathbb{P}[X_n = x^-, \forall m < n : X_m \neq x^- \mid X_0 = x]\, z^n,\\
G(x,y|z) &:= \sum_{n\geq 0} \mathbb{P}[X_n = y \mid X_0 = x]\, z^n,\\
\overline{G}_i(z) &:= \sum_{n\geq 0} \mathbb{P}[X_n = x, \forall m < n : X_m \neq x^- \mid X_0 = x]\, z^n.
\end{aligned}$$

We also write $F(-i) := F(-i|1)$. Note that the definitions of $F(-i|z)$ and $\overline{G}_i(z)$ are independent of the specific choice of $x \in \mathcal{T}$ with $\tau(x) = i$. Moreover, we have the following equations:

$$\begin{aligned}
F(-i|z) &= p(-i)\, z + \sum_{j\in G} p(i,j)\, z\, F(-j|z)\, F(-i|z), \qquad (11)\\
\overline{G}_i(z) &= \frac{1}{1-\sum_{j\in G} p(i,j)\, z\, F(-j|z)}.
\end{aligned}$$

Furthermore, we have $F(-i|z) = \overline{G}_i(z)\, p(-i)\, z$. Recall that the spectral radius of $(X_n)_{n\in\mathbb{N}_0}$ is the inverse of the radius of convergence R of $G(o,o|z)$.

Define for $k \in \mathbb{N}_0$ the *exit times*

$$\mathbf{e}_k := \min\bigl\{m \in \mathbb{N}_0 \bigm| \forall m' \geq m : |X_{m'}| \geq k\bigr\} \qquad (12)$$

and write $\mathbf{W}_k := X_{\mathbf{e}_k}$. Observe that $\big(\tau(\mathbf{W}_k)\big)_{k\in\mathbb{N}_0}$ is a Markov chain with transition matrix $Q = \big(q(i,j)\big)_{i,j\in G}$ defined by

$$q(i,j) := \frac{1-F(-j)}{1-F(-i)} p(i,j)\overline{G}_i(1). \tag{13}$$

This can be easily verified analogously to [20] or [14]. Now, we can state the result about the rate of escape.

Theorem 3.8. *Suppose that Q is positive recurrent with invariant probability measure ν. Let*

$$\Lambda := \sum_{i\in G} \nu(i)\,\frac{F'(-i|1)}{F(-i)}.$$

Then the following hold::

1. *If $\Lambda < \infty$ we have for each bounded weight function w*

$$\lim_{n\to\infty} \frac{l(X_n)}{n} = \Bigg(\sum_{i,j\in G} w(i,j)\nu(i)q(i,j)\Bigg)\cdot\Bigg(\sum_{i\in G}\nu(i)\,\frac{F'(-i|1)}{F(-i)}\Bigg)^{-1} \quad \textit{almost surely.} \tag{14}$$

   In this case the rate of escape w.r.t. the natural distance exists and is positive, i.e.,

$$\lim_{n\to\infty}\frac{|X_n|}{n} > 0 \qquad \textit{almost surely.}$$

2. *If the spectral radius of $(X_n)_{n\in\mathbb{N}_0}$ is strictly smaller than 1, then $\Lambda < \infty$.*
3. *If $\Lambda = \infty$ then $\liminf_{n\to\infty}\frac{l(X_n)}{n} = 0$.*

Remark 3.2. The formula for the rate of escape (14) is of rather formal nature, since it is given in terms of the generating functions $F(-i|z)$. These generating functions are solution of the infinite system of algebraic equations (11).

Remark 3.3. Theorem 3.8.1 and 3.8.3 hold for locally infinite graphs, too.

The assumption of positive recurrence of Q in Theorem 3.8 is essential. In the following we give an example, Example 3.4, where Q governs a transient random walk such that the rate of escape is random. An example, Example 3.6, where Q is positive recurrent is given in Section 3.6.

Example 3.4. Consider $G=\mathbb{Z}$ with its usual neighbourhood relation and transition probabilities given by

$$p_G(i,i+1) = \begin{cases} p, & \text{if } i\geq 1\\ 1-q, & \text{if } i\leq -1\end{cases}, \quad p_G(i,i-1) = \begin{cases} 1-p, & \text{if } i\geq 1\\ q, & \text{if } i\leq -1\end{cases},$$

and $p_G(0,1) = p_G(0,-1) = 1/2$, where $p,q\in(1/2;1)$, $p\neq q$. We set $i_0 := 0$. Choose now c_1, c_2 with

$$0 < c_1 < c_2 < \min\Big\{1-\frac{1}{2p}, 1-\frac{1}{2q}\Big\}.$$

Consider the directed cover $\mathcal{T}$ of G, where $p(-i) := c_1$ if $i \geq 0$, and $p(-i) := c_2$ if $i < 0$. By definition of c_1 and c_2, the random walk on $\mathcal{T}$ visits only finitely many vertices of each cone type, since

$$\begin{aligned} p(i,i+1) = 1 - p(i,i-1) - p(-i) &= 1 - \big((1-c_1)(1-p) + c_1\big) > \frac{1}{2}, \quad \text{if } i \geq 0, \\ p(i,i-1) = 1 - p(i,i+1) - p(-i) &= 1 - \big((1-c_2)(1-q) + c_2\big) > \frac{1}{2}, \quad \text{if } i \leq -1. \end{aligned}$$

Thus, $\tau(X_k)$ tends either to $+\infty$ or $-\infty$. Considering the speed w.r.t. the natural graph metric in the tree, in the first case the random walk has speed

$$\mathbb{E}\big[|X_{n+1}| - |X_n| \,\big|\, \tau(X_n) = i > 0\big] = (1-c_1)\,p + (1-c_1)(1-p) - c_1 = 1 - 2c_1,$$

while in the case $\tau(X_k) \to -\infty$ the rate of escape is different, namely

$$\mathbb{E}\big[|X_{n+1}| - |X_n| \,\big|\, \tau(X_n) = i < 0\big] = (1-c_2)\,q + (1-c_2)(1-q) - c_2 = 1 - 2c_2.$$

3.6 Asymptotic Entropy and Hausdorff Dimension

A characteristic of transient random walks that is connected to the rate of escape is the asymptotic entropy of the process. Let π_n be the distribution of X_n, that is, for $x \in \mathcal{T}$ the number $\pi_n(x)$ is the probability of visiting x at time n when starting at o. If there is some non-negative number h such that

$$h = \lim_{n\to\infty} -\frac{1}{n}\mathbb{E}\big[\log \pi_n(X_n)\big],$$

then h is called the *asymptotic entropy*, introduced in [1]. For the rest of this section we assume that the transition probabilities are *bounded away from* 0, that is, $p(x,y) \geq \varepsilon_0$ for some $\varepsilon_0 > 0$ and all $x, y \in \mathcal{T}$. Under the assumption that $(X_n)_{n\in\mathbb{N}_0}$ is transient we obtain the following theorem that links the asymptotic entropy with the rate of escape.

Theorem 3.9. *Assume that Q is positive recurrent with invariant probability measure ν and that the spectral radius of $(X_n)_{n\in\mathbb{N}_0}$ is strictly smaller than 1. Let ℓ_0 be the rate of escape w.r.t. the natural graph metric. Then the entropy rate h exists and satisfies*

$$h = \ell_0 \sum_{i,j\in G} -\nu(i) q(i,j)\, \log q(i,j) > 0.$$

Remark 3.4. The assumptions of Theorem 3.9 are satisfied for transient random walks on directed covers of finite graphs, see [20, Sections 4 & 5]. Thus, for the case of covers of finite graphs we get the completely new result that the entropy exists and is strictly positive whenever the random walk is transient.

Remark 3.5. The matrix Q as defined by (13) is the transition matrix of the Markov chain $\big(\tau(X_{\mathbf{e}_k})\big)_{k\in\mathbb{N}}$, where $\mathbf{e}_k$ is a random time as defined in (12). The sum on the right hand side of Theorem 3.9 equals the entropy rate (for positive recurrent Markov chains) of $\bigl(\tau(X_{\mathbf{e}_k})\bigr)_{k\in\mathbb{N}}$ defined by

$$h_Q := \lim_{n\to\infty} -\frac{1}{n}\log \pi\bigl(\tau(X_{\mathbf{e}_1}),\dots,\tau(X_{\mathbf{e}_n})\bigr),$$

where $\pi(\tau_1,\dots,\tau_n)$ is the joint distribution of $\bigl(\tau(X_{\mathbf{e}_1}),\dots,\tau(X_{\mathbf{e}_n})\bigr)$. That is, $h=\ell\, h_Q$. For more details we refer to [7, Chapter 4].

Remark 3.6. The proof of Theorem 3.9 shows also that the entropy rate equals the rate of escape with respect to the Green metric. Recall that the distance of a vertex $x\in\mathcal{T}$ to o w.r.t. the Green metric is given by $-\log\mathbb{P}\bigl[\exists n\in\mathbb{N}_0 : X_n=x \mid X_0=o\bigr]$.

A consequence of the proof of the last theorem is the following corollary that states that one can compute the entropy also individually:

Corollary 3.10. *Under the assumptions of Theorem 3.9,*

$$h=\liminf_{n\to\infty} -\frac{\log \pi_n(X_n)}{n} \text{ almost surely.}$$

Moreover, we get the following result:

Theorem 3.11. *Under the assumptions of Theorem 3.9,*

$$-\frac{1}{n}\log \pi_n(X_n) \xrightarrow{n\to\infty} h \quad \text{in } L_1,$$

that is, $\int\bigl|-\frac{1}{n}\log\pi_n(X_n)-h\bigr| d\mathbb{P}\to 0$ *for* $n\to\infty$.

Another consequence of Theorem 3.9 follows with [15], which gives an estimation (in some special cases also an explicit formula) for the Hausdorff dimension of the harmonic measure on the boundary of $\mathcal{T}$. Since almost every random path on $\mathcal{T}$ converges to a point on the boundary $\partial\mathcal{T}$ one can investigate the image ϱ of the measure $\mathbb{P}$ under the mapping onto ∂T. Since we have a nearest neighbour random walk this image is well-defined and it is called the *harmonic measure* of $\mathbb{P}$ on $\mathcal{T}$. For $\xi_1,\xi_2\in\partial\mathcal{T}$ let $\xi_1\wedge\xi_2$ be the confluent of the geodesics from o to ξ_1 and from o to ξ_2. The Hausdorff dimension of ϱ is defined to be

$$\dim\varrho := \operatorname{ess\,sup}_{\xi\in\partial\mathcal{T}} \liminf_{k\to\infty} -\frac{\log\varrho\bigl(B_\xi^k\bigr)}{k},$$

where $B_\xi^k := \bigl\{\zeta\in\partial\mathcal{T} \bigm| |\xi\wedge\zeta|\geq k\bigr\}$ for $k\in\mathbb{N}$, $\xi\in\partial\mathcal{T}$. The following corollary follows directly with [15, Theorem 1.4.1 & 1.5.3]:

Corollary 3.12. *Under the assumptions of Theorem 3.9, we have:*

1. *The Hausdorff dimension of ϱ satisfies*

$$\sum_{i,j\in G} -\nu(i)q(i,j)\,\log q(i,j) \leq \dim\varrho \leq \frac{-\log\varepsilon_0}{\ell_0},$$

where ℓ_0 is the rate of escape w.r.t. the natural graph metric.

2. If the entropy h also satisfies $h = \lim_{n\to\infty} -\frac{1}{n} \log \pi_n(X_n)$ almost surely, then the Hausdorff dimension of ϱ is given by

$$\dim \varrho = -\sum_{i,j\in G} \nu(i) q(i,j) \log q(i,j) = \frac{h}{\ell_0}.$$

Remark 3.7. All the results in this section hold also for locally infinite graphs with finite outdegrees.

We conclude this section with two examples. While the first, Example 3.5, demonstrates how the asymptotic entropy may be calculated explicitly in the finite case, the second one, Example 3.6, gives a sufficient condition for positive recurrence of Q in the infinite setting.

Example 3.5. Consider a graph G with vertex set $V = \{i_0, i_1, i_2\}$, edge set

$$E = \big\{(i_0, i_1), (i_0, i_2), (i_1, i_0), (i_2, i_1)\big\}$$

and its directed cover $\mathcal{T}$. We define the following transition probabilities on $\mathcal{T}$:

$$p(i_0, i_1) := \tfrac{1}{3},\ p(i_0, i_2) := \tfrac{1}{3},\ p(i_1, i_0) := \tfrac{1}{2},\ p(i_2, i_1) := \tfrac{3}{4},$$
$$p(-i_0) := \tfrac{1}{3},\ p(-i_1) := \tfrac{1}{2},\ p(-i_2) := \tfrac{1}{4}.$$

We can solve the system of polynomial equations (11) with the help of MATHEMATICA and hence obtain a numerical approximation for the asymptotic entropy $h \approx 0.060499$.

Example 3.6. Suppose we are given a graph G endowed with transition probabilities such that the random walk on G is positive recurrent with invariant probability measure ν_G. Choose the backward probabilities $p(-i)$ in such a way that the following holds for every $i, j \in G$: if there are paths from i_0 to i and from i_0 to j in G of the same length, then $p(-i) = p(-j)$. This condition implies that $F\big(-\tau(y_1)\big) = F\big(-\tau(y_2)\big)$ if $y_1^- = y_2^-$. Thus, the quotient $c_i := \big(1 - F(-j)\big)/\big(1 - F(-i)\big)$ depends only on i if $p(i,j) > 0$. This yields that

$$1 = \sum_{j\in G} q(i,j) = \sum_{j\in G} c_i p(i,j)\, \overline{G}_i(1) = c_i \overline{G}_i(1)\big(1 - p(-i)\big),$$

or equivalently, $\overline{G}_i(1) = c_i^{-1}\big(1 - p(-i)\big)^{-1}$. But this implies that ν_G is also the invariant probability measure of Q.

If we have $p(-i) \leq 1/2 - \varepsilon$ for some $\varepsilon > 0$, then $G(o, o|z)$ has radius of convergence strictly greater than 1, providing that the rate of escape w.r.t $|\cdot|$ exists and is strictly positive. Furthermore, the asymptotic entropy exists and is strictly positive.

4 Proofs

4.1 Proof of Theorem 3.1

The second statement of the theorem is just the following observation. Since A has non-negative entries we find

$$\|A^n\|_\infty = \sup_{x\in\ell_\infty, \|x\|_\infty = 1} \|A^n x\|_\infty = \|A^n \mathbf{1}\|_\infty,$$

and hence $|\mathcal{T}_n| \le \|A^n\|_\infty$.

The proof of the first statement is divided into two steps:

Step 1: $\lambda^+(A) = br(\mathcal{T})$. This follows from the results on homesick random walks: Corollary 3.4 together with the fact that a homesick random walk on a tree $\mathcal{T}$ is recurrent if $\lambda > br(\mathcal{T})$ and transient if $\lambda < br(\mathcal{T})$, compare with Theorem 3.5 in [18]. For this purpose, observe that $\lambda^+(M) = \lambda^+(A)/\lambda$.

Step 2: $br(\mathcal{T}) = \underline{gr}(\mathcal{T})$. It is convenient to introduce the notation

$$\dim\sup\partial\mathcal{T} := \lim_{n\to\infty}\max_{v\in\mathcal{T}}\frac{1}{n}\log|\mathcal{T}_n^v|$$

and to follow the argumentation of §14.4 and §14.5 in [18]. We need the following lemma:

Lemma 4.1. *Let $\mathcal{T}$ be a tree of bounded geometry. Then there exists a sequence of subtrees $\mathcal{T}^j = \mathcal{T}^{v_j}$ such that*

$$\dim\partial\mathcal{T}^j \ge \left(1-\frac{1}{j}\right)\dim\sup\partial\mathcal{T}. \tag{15}$$

Proof. We use the following fact, which is (14.19) in [18].

Fact: For $j \ge 1$ there is some unit flow θ_j on some subtree $\mathcal{T}^j = \mathcal{T}^{v_j}$ such that

$$\frac{1}{|x|-|v_j|}\log\frac{1}{\theta_j(x)} \ge \left(1-\frac{1}{j}\right)\dim\sup\partial\mathcal{T} \qquad \text{for all } x \in \mathcal{T}^j.$$

This implies that for every ray $\xi = \langle \xi_1, \xi_2, \dots\rangle \in \partial\mathcal{T}^j$ we have

$$\liminf_{n\to\infty}\frac{1}{n}\log\frac{1}{\theta_j(\xi_n)} \ge \left(1-\frac{1}{j}\right)\dim\sup\partial\mathcal{T}.$$

Recall the definition of the Hölder exponent of a unit flow (or of its corresponding Borel probability measure on the boundary respectively) $\text{Hö}(\theta)(\xi) := \liminf_{n\to\infty} -\frac{1}{n}\log\theta(\xi_n)^{-1}$. With this notation we have that $\text{Hö}(\theta_j) \ge (1-1/j)\dim\sup\partial\mathcal{T}$. Those edges where the flow θ_j is positive define a subtree $\mathcal{U}^j$ of $\mathcal{T}^j$. Theorem 14.15 of [18] implies that $\text{Hö}(\theta_j) \le \dim\theta_j \le \dim\partial\mathcal{U}^j$, and hence $\text{Hö}(\theta_j) \le \dim\partial\mathcal{T}^j$. $\square$

Due to Lemma 4.1 there exists a sequence of subtrees $\mathcal{T}^j = \mathcal{T}^{v_j}$ such that

$$\dim\partial\mathcal{T}^j \ge \left(1-\frac{1}{j}\right)\dim\sup\partial\mathcal{T}.$$

Furthermore, observe that $\dim\partial\mathcal{T}^j \le \dim\partial\mathcal{T}$ and $\log\underline{gr}(\mathcal{T}) \ge \dim\partial\mathcal{T}$; compare with Equation (1). Hence,

$$\begin{aligned}
\dim\partial\mathcal{T} &\ge \left(1-\frac{1}{j}\right)\dim\sup\partial\mathcal{T}\\
&\ge \left(1-\frac{1}{j}\right)\liminf_{n\to\infty}\frac{1}{n}\log|\mathcal{T}_n|\\
&= \left(1-\frac{1}{j}\right)\log\underline{gr}(\mathcal{T})\\
&\ge \left(1-\frac{1}{j}\right)\dim\partial\mathcal{T},
\end{aligned}$$

Now, letting $j \to \infty$ yields $br(\mathcal{T}) = \underline{gr}\mathcal{T}$.

$\square$

4.2 Proof of Theorem 3.2 and 3.3

Proof of Theorem 3.2. The idea is to couple a *delayed* branching process Z_t^* (in continuous time) to the random walk X_n on $\mathcal{T}$ (in discrete time) and to show that the branching process dies out if and only if the random walk visits the loop (o, o). To this end, observe first that the rates of the branching process Z_t sum up to 1, i.e., $p(-i) + \sum_{j \in G} p(i, j) = 1$, and hence can be interpreted as probabilities. The process Z_t^* starts with one particle of type i_0. With rate $p(-i_0)$ this particle dies. Observe that the random walk started in o returns to o at the first step with probability $p(-i_0)$. The particle produces an offspring of type j with rate $p(i_0, j)$. Observe that the random walk on $\mathcal{T}$ is in a vertex $x \in \mathcal{T}$ with $|x| = 1$ and label j at time 1 with probability $p(i_0, j)$. The *delayed* (or *sleepy*) process is defined inductively. As long as one particle of type i has one offspring alive (awake or sleeping) it is *sleeping*, i.e., it does neither die nor produce offspring. If all its offspring have died it *wakes up* and either dies with rate $p(-i)$ or produces an offspring of type j with rate $p(i, j)$. If it does the latter it falls asleep again and if it dies its direct ancestor wakes up. One possibility to define Z_t^* formally is through the following rates on the state space of non-backtracking paths (including the empty path $\langle \emptyset \rangle$) of $\mathcal{T}$. Let $x_0 = o$ and

$$
\begin{aligned}
\langle x_0 \rangle &\to \langle x_0, x_1 \rangle \text{ at rate } p(\tau(x_0), \tau(x_1)), \text{ if } x_0 = x_1^-, \\
\langle x_0 \rangle &\to \langle \emptyset \rangle \text{ at rate } p(-\tau(x_0)), \\
\langle \emptyset \rangle &\to \langle \emptyset \rangle \text{ at rate } 1,
\end{aligned}
$$

and for $n \geq 1$

$$
\begin{aligned}
\langle x_0, \ldots, x_n \rangle &\to \langle x_0, \ldots, x_{n-1} \rangle \text{ at rate } p(-\tau(x_n)), \\
\langle x_0, \ldots, x_n \rangle &\to \langle x_0, \ldots, x_n, x_{n+1} \rangle \text{ at rate } p(\tau(x_n), \tau(x_{n+1})), \text{ if } x_n = x_{n+1}^-.
\end{aligned}
$$

Observe at this point that the path $\langle x_0, \ldots, x_n \rangle$ corresponds to respectively one sleeping particle of type $\tau(x_i)$ $(i < n)$ and one particle awake of type $\tau(x_n)$ in the genealogical order. The empty path $\langle \emptyset \rangle$ corresponds to the extinction of the process. Let S_n be the jump chain of Z_t^* which is the sequence of values taken by the continuous-time Markov chain Z_t^*. Define the projection ϕ from the space of paths to the set of vertices of $\mathcal{T}$ as $\phi(\langle x_0, \ldots, x_n \rangle) = x_n$ and $\phi(\langle \emptyset \rangle) = \emptyset$. Using standard arguments we can couple the two processes S_n and X_n such that $S_n = \emptyset$ if and only if $X_{n-1} = X_n = o$. We conclude that $Z_t^* = \langle \emptyset \rangle$ for some $t > 0$ if and only if the random walk X_n visits the loop (o, o). It remains to prove that $Z_t^* = \langle \emptyset \rangle$ is equivalent to the extinction of the original process Z_t. Recall the interpretation of Z_t^* as a delayed version of Z_t. Hence, both processes can be seen as functions on the same probability space and we can conclude with a standard coupling argument.

$\square$

Proof of Theorem 3.3. Theorem 3.3 is a consequence of Theorem 3.2 and the following result of [3]. Consider a branching random walk (BRW) as a continuous-time process where particles live on a countable set X. Each particle lives on a site and, independently of the others and the past history of the process, has a exponential lifetime with mean 1. A particle living at a site x gives birth to a new particle in y with exponential rate $k(x,y)$. Here $K = \big(k(x,y)\big)_{x,y\in X}$ is a matrix with non-negative entries. In [3] it is shown that there exists a critical value λ_w depending on K such that the process dies out a.s. if $\lambda_w < 1$ and survives with positive probability if $\lambda_w > 1$. Furthermore, there is the following characterization of this critical value, compare with Equation (4.11) in [3]:

$$\lambda_w = \lambda^+(K). \tag{16}$$

The statement follows now with the observation that we can scale our process by dividing each rate $(p(-i),\ p(i,j))$ at a vertex i by $p(-i)$ without influencing the survival of the process.

$\square$

4.3 Prof of Theorem 3.5

We consider the infinite Galton-Watson process Z_t with first moments $M_\omega := \lambda A_\omega$, compare with the paragraph before Theorem 3.2. Due to Theorem 3.2 it suffices to prove the following:
Claim: The process Z_t survives (globally) if and only if $\lambda > 1/2d$.
First, observe that $\rho(A_\omega) = \limsup_{n\to\infty} \left(A_\omega^{(n)}(i,j)\right)^{1/n} = 2d$. This can be seen with Equation (3) and the fact that $C(\omega)$ contains balls of arbitrary large radius as subgraphs. There are two types of survival for infinite-type Galton Watson processes. We say that the process survives globally if $Z_t > 0$ for all t and survives locally if $Z_t(i) > 0$ for infinitely many t and all ($\Leftrightarrow$ some) i, compare with [13]. Now, Corollary 2.6 in [19] implies that Z_t survives locally if and only if $\lambda > 1/\rho(A_\omega) = 1/(2d)$. Finally, it remains to show that the process survives locally if and only if it survives globally. But for $\lambda \leq 1/(2d)$, observe that $\sum_y m_\omega^{(n)}(x,y) \leq 1$ for all x and n. Hence, the expected number of particles in generation n is bounded by 1. Since Z_t either converges to 0 or ∞ we obtain that the process does not survive globally if $\lambda \leq 1/(2d)$.

$\square$

4.4 Proof of Proposition 3.7

The non-ergodicity part in the supercritical case $\lambda^+(M) > 1$ is obvious due to Theorem 3.3. Nevertheless, we give an alternative proof that uses directly the definition of $\lambda^+(M)$. This method works also for the critical case $\lambda^+(M) = 1$ when the supremum is attained. Furthermore, the proof might be useful in order to understand the behaviour between $\lambda^+(M)$ and $r_\infty(M)$.
Proof of Proposition 3.7. The first steps are quite standard and use the tree structure of our process, compare also with [20]. The random walk is positive recurrent if and only if the reversible (and stationary) measure m defined in (9) is finite, that is, if $m(\mathcal{T}) = \sum_{x\in\mathcal{T}} m(x) < \infty$. For each $i \in G$, we construct a tree $\mathcal{T}_i$: it consists of the cone $\mathcal{T}^x$ of $\mathcal{T}$ with $x \in \mathcal{T}$ and $\tau(x) = i$ which

is connected by a (nondirected) single edge from x to an additional vertex o_i. On $\mathcal{T}_i$ we consider the random walk with the same transition probabilities as in $\mathcal{T}$ but where the probability from o_i to x is 1 and from x to o_i is $p(-i)$. This random walk is obviously reversible with reversible measure m_i, which is defined analogously to Equation (9). We can express $m(\mathcal{T})$ in terms of the measures m_i of the subtrees $\mathcal{T}_i$:

$$m(\mathcal{T}) = \sum_{x\in\mathcal{T}} p(o,x) m_{\tau(x)}(\mathcal{T}_{\tau(x)}). \tag{17}$$

We approximate $\mathcal{T}_i$ with finite subtrees $\mathcal{T}_i^n$ of height n. The sequence $\big(m_i(\mathcal{T}_i^n)\big)_{n\in\mathbb{N}_0}$ is increasing with limit $m_i(\mathcal{T}_i)$. We obtain the inductive formula: $m_i(\mathcal{T}_i^0) = 1$ and for each $i \in \mathbb{N}$

$$m_i(\mathcal{T}_i^n) - 1 = \frac{1}{p(-i)}\Bigg(1 + \sum_{j\in G} p(i,j)\big(m_j(\mathcal{T}_j^{n-1}) - 1\big)\Bigg).$$

We can write the last equation in vector form using $\mathbf{m}_n := \big(m_i(\mathcal{T}_i^n) - 1\big)_{i\in G}$ and $\mathbf{p} = \big(1/p(-i)\big)_{i\in G}$ as $\mathbf{m}_n = \mathbf{p} + M\mathbf{m}_{n-1}$. Thus,

$$\mathbf{m}_n = \mathbf{p} + M\mathbf{p} + M^2\mathbf{p} + \cdots + M^{n-1}\mathbf{p}$$

and the sequence $(\mathbf{m}_n)_{n\in\mathbb{N}_0}$ will converge componentwise to a finite limit if and only if $\sum_{n\geq 0} M^n\mathbf{p}$ converges in each component. If $\lambda^+(M) > 1$ there is some $\lambda^* > 1$ and $0 < \mathbf{f}^* \in \ell_\infty$ such that $M\mathbf{f}^* \geq \lambda^*\mathbf{f}^*$. Hence, there is some $c > 0$ such that $\mathbf{p} = c\mathbf{f}^* + (\mathbf{p} - c\mathbf{f}^*)$ and $(\mathbf{p} - c\mathbf{f}^*) > 0$. At this point we need $p(-i) < 1 - \varepsilon$ of Assumption (8). Finally,

$$\sum_{n\geq 0} M^n\mathbf{p} \geq c\sum_{n\geq 0}(\lambda^*)^n\mathbf{f}^* = \infty$$

and non-ergodicity follows. If $\lambda^+(M) = 1$ and the supremum in the definition of $\lambda^+(M)$ is attained, non-ergodicity follows analogously.

In order to show that $r_\infty := r_\infty(M) < 1$ implies ergodicity we have to show that the sum $\sum_{n\geq 0} M^n\mathbf{p} < \infty$ is finite in every component if $r_\infty < 1$. Recall that $r_\infty = \lim_{n\to\infty} \sqrt[n]{\|A^n\|_\infty}$. This provides that for every small $\varepsilon > 0$ there is some $N_\varepsilon \in \mathbb{N}$ such that $1 > r_\infty + \varepsilon \geq \sqrt[n]{\|A^n\|_\infty}$ for all $n \geq N_\varepsilon$. Due to Assumption (8) we get for every $i \in G$ and for ε small enough and $n \geq N_\varepsilon$

$$\left(\sum_{n\geq 0} M^n\mathbf{p}\right)(i) \leq \left\|\sum_{n\geq 0} M^n\mathbf{p}\right\|_\infty \leq \sum_{n\geq 0}\|M^n\mathbf{p}\|_\infty \leq \sum_{n=0}^{N_\varepsilon - 1}\|M^n p\|_\infty + \|\mathbf{p}\|_\infty \sum_{n\geq N_\varepsilon}\big(r_\infty + \varepsilon\big)^n,$$

which is finite and independent of i.

□

4.5 Proof of Theorem 3.8

Proof of Theorem 3.8.1. Since the proof is an adaption of the arguments in [20] we just give a sketch. If $\Lambda < \infty$ then

$$\frac{\mathbf{e}_k}{k} \xrightarrow{k\to\infty} \Lambda = \sum_{i\in G} \nu(i)\, \frac{F'(-i|1)}{F(-i)} \quad \text{almost surely.}$$

With [20] we get $\ell_0 := \lim_{n\to\infty} |X_n|/n = \lim_{k\to\infty} k/\mathbf{e}_k = \Lambda^{-1}$ if $\Lambda < \infty$.

By assumption, the process $(\mathbf{W}_k)_{k\in\mathbb{N}_0}$ is positive recurrent with invariant probability measure ν. Thus, the process $\bigl(\tau(\mathbf{W}_{k-1}), \tau(\mathbf{W}_k)\bigr)_{k\in\mathbb{N}}$ has the invariant probability measure $\nu_1(i,j) = \nu(i)q(i,j)$. An application of the ergodic theorem for positive recurrent Markov chains yields

$$\frac{1}{n}\sum_{k=1}^{n} w\bigl(\tau(\mathbf{W}_{k-1}), \tau(\mathbf{W}_k)\bigr) = \frac{l(\mathbf{W}_n)}{n} \xrightarrow{n\to\infty} \int w(i,j)\, d\nu_1(i,j) = \sum_{i,j\in G} w(i,j)\nu_1(i,j). \tag{18}$$

Observe that the sum on the right hand side is finite, since $w(\cdot,\cdot)$ is bounded. In the case $\Lambda < \infty$ we obtain, analogously to [20], with $\mathbf{k}(n) := \max\{k \in \mathbb{N}_0 | \mathbf{e}_k \leq n\}$

$$\lim_{n\to\infty} \frac{l(X_n)}{n} = \lim_{n\to\infty} \frac{l(W_{\mathbf{k}(n)})}{\mathbf{k}(n)} \frac{\mathbf{k}(n)}{\mathbf{e}_{\mathbf{k}(n)}} \frac{\mathbf{e}_{\mathbf{k}(n)}}{n} = \ell_0 \lim_{k\to\infty} \frac{l(\mathbf{W}_k)}{k} \quad \text{almost surely,}$$

since $\mathbf{e}_{\mathbf{k}(n)}/n$ converges to 1. The claim follows now with (18). $\square$

Proof of Theorem 3.8.2. Denote by R the radius of convergence of $G(o,o|z)$. We have $R > 1$, since the spectral radius of $(X_n)_{n\in\mathbb{N}_0}$ is strictly smaller than 1 by assumption. Define for $x \in \mathcal{T}$ and $z \in \mathbb{C}$:

$$\begin{aligned} U(x,x|z) &:= \sum_{n\geq 1} \mathbb{P}\bigl[X_n = x, \forall m \in \{1,\dots,n-1\} : X_m \neq x \bigm| X_0 = x\bigr] z^n \\ &= \sum_{j\in G} p(i,j)\, z\, F(-j|z) + p(-i)\, z\, F(x^-, x|z). \end{aligned}$$

The proof splits up into the two following lemmas.

Lemma 4.2. *For $r \in [1,R)$ and all $x \in \mathcal{T}$ we have $G(x,x|r) \leq 1/(1-r/R)$.*

Proof. For every $x \in \mathcal{T}$ with $\tau(x) = i$, we have

$$\infty > G(x,x|r) = \sum_{n\geq 0} U(x,x|r)^n = \frac{1}{1-U(x,x|r)} \quad \text{for all } r \in [1,R), \tag{19}$$

that is, $U(x,x|r) < 1$ for all $r \in [1,R)$. Since $U(x,x|0) = 0$ and $U(x,x|z)$ is continuous, increasing and convex we have

$$U(x,x|r) \leq \frac{r}{R},$$

that is, $G(x,x|r) \leq 1/(1-r/R)$. $\square$

Lemma 4.3. *There is a constant C_F such that $F'(-i|1) \leq C_F$ for all $i \in G$.*

Proof. With Equation (19) we get

$$U'(x,x|z) = \frac{G'(x,x|z)}{G^2(x,x|z)} \leq G'(x,x|z).$$

Choose any $\varepsilon \in (0, R-1)$ and define

$$h(z) := \frac{1}{\varepsilon} \frac{R(z-1)}{R-1-\varepsilon}.$$

We have $h(1) = 0$ and $h(1+\varepsilon) = 1/(1-(1+\varepsilon)/R)$. Since $G(x,x|z)$ is increasing and convex in $[0, R)$ we get with Lemma 4.2 the inequality $G'(x,x|1) \leq h'(1)$, and thus $U'(x,x|1) \leq h'(1)$.

Let be $i \in G$ and choose any $x \in \mathcal{T}$ such that $p(\tau(x), i) > 0$. Then:

$$U(x,x|z) = p\big(-\tau(x)\big) z F(x^-, x|z) + \sum_{j \in G} p\big(\tau(x), j\big) z F(-j|z).$$

Differentiating yields

$$U'(x,x|z) = p\big(-\tau(x)\big)\big(F(x^-,x|z) + zF'(x^-,x|z)\big) + \sum_{j \in G} p\big(\tau(x), j\big)\big(F(-j|z) + zF'(-j|z)\big).$$

Thus,

$$U'(x,x|1) \geq p\big(\tau(x), i\big) F'(-i|1),$$

or equivalently,

$$F'(-i|1) \leq \varepsilon_0^{-1} h'(1).$$

□

Finally, $F(-i|1) \geq \varepsilon_0$ together with the last lemma imply that $\Lambda < \infty$ if $R > 1$. □

Proof of Theorem 3.8.3. We now turn to the case $\Lambda = \infty$. Define for $N \in \mathbb{N}$ the function $g_N : G \times \mathbb{N} \to \mathbb{N}$ by $g_N(i,n) := n \wedge N$. Then obviously

$$\frac{1}{k} \sum_{l=1}^{k} g_N\big(\tau(\mathbf{W}_l), \mathbf{e}_l - \mathbf{e}_{l-1}\big) \leq \frac{\mathbf{e}_k - \mathbf{e}_0}{k} = \frac{\mathbf{e}_k}{k}. \tag{20}$$

This inequality holds for every $N \in \mathbb{N}$. The process $\big(\tau(\mathbf{W}_l), \mathbf{e}_l - \mathbf{e}_{l-1}\big)_{l \in \mathbb{N}}$ is also a positive recurrent Markov chain; compare with [20]. For each N there is a constant C_N such that the left side of (20) converges to C_N for almost every realisation of $(X_n)_{n \in \mathbb{N}_0}$. The sequence $(C_N)_{N \in \mathbb{N}}$ is strictly increasing and diverges to ∞, since $\Lambda = \infty$ and $g_N(i,n) \to g(i,n)$ for $N \to \infty$ with $g(i,n) := n$. Thus, $\mathbf{e}_k/k$ tends to infinity. This yields $\liminf_{n\to\infty} |X_n|/n = 0$ almost surely since

$$0 \leq \liminf_{n \to \infty} \frac{|X_n|}{n} \leq \liminf_{n \to \infty} \frac{|X_{\mathbf{e}_{\mathbf{k}(n)}}|}{\mathbf{e}_{\mathbf{k}(n)}} \leq \liminf_{n \to \infty} \frac{|X_{\mathbf{e}_{\mathbf{k}(n)}}|}{\mathbf{k}(n)} \frac{\mathbf{k}(n)}{\mathbf{e}_{\mathbf{k}(n)}} = 0 \quad \text{almost surely.}$$

□

4.6 Proof of Theorems 3.9 and Corollary 3.10

Proof of Theorem 3.9. Define for $x, y \in \mathcal{T}$ and $z \in \mathbb{C}$

$$L(x,y|z) \;:=\; \sum_{n\geq 0} \mathbb{P}\bigl[X_n = y, \forall m \in \{1,\dots,n\} : X_m \neq x \mid X_0 = x\bigr] z^n.$$

If y is a successor of x in $\mathcal{T}$, then

$$L(x,y|z) = p(x,y)z\overline{G}_{\tau(y)}(z). \tag{21}$$

We have the following important equations, which follow by conditioning on the last visit of o, the first visit of x respectively:

$$G(o,x|z) = G(o,o|z)L(o,x|z) = F(o,x|z)G(x,x|z). \tag{22}$$

If $y \in \mathcal{T}$ lies on the unique geodesic from $x \in \mathcal{T}$ to $w \in \mathcal{T}$, then

$$L(x,w|z) = L(x,y|z)L(y,w|z).$$

Observe that the generating functions $L(\cdot,\cdot|z)$, $\overline{G}_i(z)$, $G(\cdot,\cdot|z)$ have radii of convergence of at least $R > 1$, since the spectral radius of the random walk is strictly smaller than 1. Define for $x \in \mathcal{T}$

$$l(x) := -\log L(o,x|1) = -\sum_{i=1}^{|x|} \log L(x_{i-1},x_i|1),$$

where x_i is the unique element on the geodesic from o to x at distance i from o. This length function arises from the weight function on the edges of G defined by $w(i,j) := -\log L(x,y|1)$, where $x \in \mathcal{T}$ with $\tau(x) = i$ and y is a successor of x of type j. This weight function is well-defined, since all subtrees $\mathcal{T}^{x_1}$ and $\mathcal{T}^{x_2}$ with $\tau(x_1) = \tau(x_2)$ are isomorphic as rooted trees. We claim that the rate of escape w.r.t. the length function l exists and equals the asymptotic entropy. The technique of the proof which we will give was motivated by [2], where it is shown that the asymptotic entropy of random walks on finitely generated groups equals the rate of escape w.r.t. the Green metric.

The proof of Theorem 3.9 is split up into the following lemmas.

Lemma 4.4. *$h := \lim_{n\to\infty} l(X_n)/n$ exists and is non-negative.*

Proof. Observe that $L(x,y) \geq p(x,y) \geq \varepsilon_0$ whenever $p(x,y) > 0$. If $y \in \mathcal{T}$ is a successor of $x \in \mathcal{T}$, then we obtain with Equation (21) and Lemma 4.2

$$L(x,y|1) = p(x,y)\,\overline{G}_{\tau(y)}(1) \leq p(x,y)/(1-1/R) \leq (1-\varepsilon_0)/(1-1/R).$$

Hence, the functions $L(x^-,x|1)$ are uniformly bounded in $x \in \mathcal{T} \setminus \{o\}$. Theorem 3.8 provides that the rate of escape h with respect to l exists. By (22), we get also

$$h = \lim_{n\to\infty} -\frac{1}{n}\log F(o,X_n|1) - \frac{1}{n}\log G(X_n,X_n|1) + \frac{1}{n}\log G(o,o|1).$$

With Lemma 4.2 we have $1 \le G(x,x|1) \le 1/(1-1/R)$ for every $x \in \mathcal{T}$. Since $F(o,X_n|1) \le 1$ we get

$$h = \lim_{n\to\infty} -\frac{1}{n} \log F(o,X_n|1) \ge 0.$$

□

By (22), we can rewrite h as

$$h = \lim_{n\to\infty} -\frac{1}{n} \log L(o,X_n|1) = \lim_{n\to\infty} -\frac{1}{n} \log \frac{G(o,X_n|1)}{G(o,o|1)} = \lim_{n\to\infty} -\frac{1}{n} \log G(o,X_n|1). \tag{23}$$

Since

$$G(o,X_n|1) = \sum_{m\ge 0} p^{(m)}(o,X_n) \ge p^{(n)}(o,X_n) = \pi_n(X_n),$$

we have

$$\liminf_{n\to\infty} -\frac{1}{n} \log \pi_n(X_n) \ge h. \tag{24}$$

The next aim is to prove $\limsup_{n\to\infty} -\frac{1}{n}\mathbb{E}\big[\log \pi_n(X_n)\big] \le h$. For this purpose, we need the following lemma:

Lemma 4.5. *For every $r \in (1,R)$, $x \in \mathcal{T}$ and $m \in \mathbb{N}$ we have $p^{(m)}(o,x) \le G(o,x|r)r^{-m}$.*

Proof. Denote by C_r the circle with radius r in the complex plane centered at 0. A straightforward computation shows that

$$\frac{1}{2\pi i}\oint_{C_r} z^m \frac{dz}{z} = \delta_{m,0}.$$

An application of Fubini's Theorem yields

$$\begin{aligned}\frac{1}{2\pi i}\oint_{C_r} G(o,x|z)\, z^{-m}\frac{dz}{z} &= \frac{1}{2\pi i}\oint_{C_r} \sum_{n\ge 0} p^{(n)}(o,x) z^n\, z^{-m}\frac{dz}{z}\\ &= \frac{1}{2\pi i}\sum_{n\ge 0} p^{(n)}(o,x) \oint_{C_r} z^{n-m}\frac{dz}{z} = p^{(m)}(o,x).\end{aligned}$$

Since $G(o,x|z)$ is analytic on C_r, $|G(o,x|z)| \le G(o,x|r)$ for all $|z| = r$. Thus,

$$p^{(m)}(o,x) \le \frac{1}{2\pi} r^{-m-1} G(o,x|r) 2\pi r = G(o,x|r)r^{-m}.$$

□

Let be $x \in \mathcal{T}$ and let x_i be the unique element on the geodesic from o to x at distance i from o. For $r < R$, iterated applications of equations (21) and (22) provide

$$\begin{aligned} G(o,x|r) &= G(o,o|r) \prod_{i=1}^{|x|} L(x_{i-1},x_i|r) \\ &= G(o,o|r) \prod_{i=1}^{|x|} p(x_{i-1},x_i)\, r\, \overline{G}_{\tau(x_i)}(r) \\ &\leq G(o,o|r)\big(r/(1-r/R)\big)^{|x|}. \end{aligned}$$

Thus,

$$p^{(m)}(o,X_n) \leq G(o,o|r)\Big(\frac{r}{1-r/R}\Big)^n r^{-m}. \tag{25}$$

We now need the following technical lemma:

Lemma 4.6. *Let $(A_n)_{n\in\mathbb{N}}$, $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ be sequences of positive numbers with $A_n = a_n + b_n$. Assume that $\lim_{n\to\infty} -\frac{1}{n}\log A_n = c \in [0,\infty)$ and that $\lim_{n\to\infty} b_n/q^n = 0$ for all $q \in (0,1)$. Then $\lim_{n\to\infty} -\frac{1}{n}\log a_n = c$.*

Proof. Clearly, there is some $N \in \mathbb{N}$ such that $b_n < a_n$ for all $n \geq N$. We get for all $n \geq N$:

$$\begin{aligned} -\frac{1}{n}\log(a_n+b_n) &\leq -\frac{1}{n}\log(a_n) = -\frac{1}{n}\log\Big(\frac{1}{2}a_n + \frac{1}{2}a_n\Big) \\ &\leq -\frac{1}{n}\log\Big(\frac{1}{2}a_n + \frac{1}{2}b_n\Big) \leq -\frac{1}{n}\log\left(\frac{1}{2}\right) - \frac{1}{n}\log(a_n+b_n). \end{aligned}$$

Letting $n \to \infty$ yields that $-\frac{1}{n}\log(a_n)$ tends to c. $\square$

In order to apply the last lemma let $A_n := \sum_{m\geq 0} p^{(m)}(o,X_n)$, $a_n := \sum_{m=0}^{n^2-1} p^{(m)}(o,X_n)$ and $b_n := \sum_{m\geq n^2} p^{(m)}(o,X_n)$. Thus, for $r \in (1,R)$ we get with (25)

$$b_n \leq \sum_{m\geq n^2} G(o,o|r)\Big(\frac{r}{1-r/R}\Big)^n \cdot r^{-m} = G(o,o|r)\Big(\frac{r}{1-r/R}\Big)^n \frac{r^{-n^2}}{1-1/r}.$$

Thus, b_n decays faster than any geometric sequence. Lemma 4.6 together with (23) yields

$$h = \lim_{n\to\infty} -\frac{1}{n}\log\sum_{m=0}^{n^2-1} p^{(m)}(o,X_n).$$

Since $G(o,X_n) \leq 1/(1-1/R)$ and $G(o,X_n) \geq F(o,X_n) \geq \varepsilon_0^n$ we get by an application of the

Dominated Convergence Theorem:

$$\begin{aligned}
h &= \int \lim_{n\to\infty} -\frac{1}{n} \log \sum_{m=0}^{n^2-1} p^{(m)}(o,X_n)\, d\mathbb{P} \\
&= \lim_{n\to\infty} \int -\frac{1}{n} \log \sum_{m=0}^{n^2-1} p^{(m)}(o,X_n)\, d\mathbb{P} \\
&= \lim_{n\to\infty} -\frac{1}{n} \sum_{x\in\mathcal{T}} p^{(n)}(o,x) \log \sum_{m=0}^{n^2-1} p^{(m)}(o,x).
\end{aligned}$$

Recall that Shannon's Inequality gives

$$\sum_{x\in\mathcal{T}} p^{(n)}(o,x) \log \mu(x) \leq \sum_{x\in\mathcal{T}} p^{(n)}(o,x) \log p^{(n)}(o,x)$$

for every finitely supported probability measure μ on $\mathcal{T}$. Setting $\mu(x) := n^{-2} \sum_{m=0}^{n^2-1} p^{(m)}(o,x)$ we get

$$\begin{aligned}
h &\geq \limsup_{n\to\infty} \left(-\frac{1}{n} \sum_{x\in\mathcal{T}} p^{(n)}(o,x) \log n^2 - \frac{1}{n} \sum_{x\in\mathcal{T}} p^{(n)}(o,x) \log p^{(n)}(o,x) \right) \\
&= \limsup_{n\to\infty} -\frac{1}{n} \mathbb{E}\bigl[\log \pi_n(X_n)\bigr].
\end{aligned}$$

Now we can conclude from Fatou's Lemma and (24) :

$$\begin{aligned}
h \leq \int \liminf_{n\to\infty} \frac{-\log \pi_n(X_n)}{n} d\mathbb{P} &\leq \liminf_{n\to\infty} \int \frac{-\log \pi_n(X_n)}{n} d\mathbb{P} \\
&\leq \limsup_{n\to\infty} -\frac{1}{n} \mathbb{E}\bigl[\log \pi_n(X_n)\bigr] \leq h. \qquad (26)
\end{aligned}$$

Thus, $\lim_{n\to\infty} -\frac{1}{n}\mathbb{E}\bigl[\log \pi_n(X_n)\bigr]$ exists and equals h. It remains to show:

Lemma 4.7.

$$h = \ell_0 \sum_{i,j\in G} -\nu(i)\, q(i,j) \log q(i,j).$$

Proof. For a moment let be $x \in \mathcal{T}$ with $|x| = n$ and let x_j be the element on the geodesic from

o to x at distance j from o. Then:

$$\begin{aligned}
l(x) &= -\log \prod_{j=1}^{n} L(x_{j-1}, x_j|1)\\
&= -\log \prod_{j=1}^{n} p\bigl(\tau(x_{j-1}), \tau(x_j)\bigr)\, \overline{G}_{\tau(x_{j-1})}(1) \frac{1-F\bigl(-\tau(x_j)\bigr)}{1-F\bigl(-\tau(x_{j-1})\bigr)} +\\
&\qquad\qquad\qquad + \log \frac{1-F\bigl(-\tau(x_n)\bigr)}{1-F(-i_0)} + \log \frac{\overline{G}_{i_0}(1)}{\overline{G}_{\tau(x_n)}(1)}\\
&= -\log \prod_{j=1}^{n} q\bigl(\tau(x_{j-1}), \tau(x_j)\bigr) + \log \frac{1-F\bigl(-\tau(x_n)\bigr)}{1-F(-i_0)} + \log \frac{\overline{G}_{i_0}(1)}{\overline{G}_{\tau(x_n)}(1)}. \qquad (27)
\end{aligned}$$

As $l(X_n)/n$ tends to h almost surely, the subsequence $\bigl(l(X_{\mathbf{e}_k})/\mathbf{e}_k\bigr)_{k\in\mathbb{N}}$ converges also to h. Since $1 \leq \overline{G}_i(1) \leq 1/(1-1/R)$ by Lemma 4.2, it follows with $x = X_{\mathbf{e}_k}$ in (27) that

$$\frac{1}{\mathbf{e}_k} \log \frac{\overline{G}_0(1)}{\overline{G}_{\tau(X_{\mathbf{e}_k})}(1)} \xrightarrow{k\to\infty} 0 \quad \text{almost surely.}$$

By positive recurrence of $\bigl(\tau(X_{\mathbf{e}_k})\bigr)_{k\in\mathbb{N}}$, an application of the ergodic theorem yields

$$-\frac{1}{k} \log \prod_{j=1}^{k} q\bigl(\tau(X_{\mathbf{e}_{j-1}}), \tau(X_{\mathbf{e}_j})\bigr) \xrightarrow{n\to\infty} h' := -\sum_{i,j\in G} \nu(i)\, q(i,j) \log q(i,j) \quad \text{almost surely,}$$

whenever $h' < \infty$. In the latter case, since $\lim_{k\to\infty} k/\mathbf{e}_k = \ell_0$ (see proof of Theorem 3.8) and $\liminf_{n\to\infty} -\frac{1}{\mathbf{e}_k}\bigl(1 - F(-\tau(X_{\mathbf{e}_k}))\bigr) = 0$ by ergodicity of $\bigl(\tau(X_{\mathbf{e}_k})\bigr)_{k\in\mathbb{N}}$, we have

$$h = \lim_{k\to\infty} \frac{l\bigl(X_{\mathbf{e}_k}\bigr)}{\mathbf{e}_k} = \lim_{k\to\infty} \frac{l\bigl(X_{\mathbf{e}_k}\bigr)}{k} \frac{k}{\mathbf{e}_k} = h'\ell_0.$$

In particular, $h > 0$ since $\ell_0 > 0$ by Theorem 3.8.2.

It remains to show that it cannot be that $h' = \infty$. For this purpose, assume $h' = \infty$. Let $N \in \mathbb{N}$ and define $h_N : G \times G \to \mathbb{R}$ by $h_N(i,j) := N \wedge \bigl(-\log q(i,j)\bigr)$. Then

$$-\frac{1}{k} \sum_{j=1}^{k} \log h_N\bigl(\tau(X_{\mathbf{e}_{j-1}}), \tau(X_{\mathbf{e}_j})\bigr) \xrightarrow{k\to\infty} h'_N := -\sum_{i,j\in G} \nu(i)\, q(i,j) \log h_N(i,j) \quad \text{almost surely.}$$

Since $h_N(i,j) \leq -\log q(i,j)$ and $h' = \infty$ by assumption, there is for every $M \in \mathbb{R}$ and almost every trajectory of $\bigl(\tau(X_{\mathbf{e}_k})\bigr)_{k\in\mathbb{N}_0}$ an almost surely finite random time $\mathbf{T_q} \in \mathbb{N}$ such that for all $k \geq \mathbf{T_q}$

$$-\frac{1}{k} \sum_{j=1}^{k} \log q\bigl(\tau(X_{\mathbf{e}_{j-1}}), \tau(X_{\mathbf{e}_j})\bigr) > M.$$

On the other hand there is for every $M > 0$, every small $\varepsilon > 0$ and almost every trajectory an almost surely finite random time $\mathbf{T_L}$ such that for all $k \geq \mathbf{T_L}$

$$
\begin{aligned}
-\frac{1}{\mathbf{e}_k}\sum_{j=1}^{k}\log L\big(X_{\mathbf{e}_{j-1}}, X_{\mathbf{e}_j}|1\big) &\in (h-\varepsilon, h+\varepsilon) \quad \text{and} \\
-\frac{1}{\mathbf{e}_k}\sum_{j=1}^{k}\log q\big(\tau(X_{\mathbf{e}_{j-1}}), \tau(X_{\mathbf{e}_j})\big) &= -\frac{k}{\mathbf{e}_k}\frac{1}{k}\sum_{j=1}^{k}\log q\big(\tau(X_{\mathbf{e}_{j-1}}), \tau(X_{\mathbf{e}_j})\big) > \ell_0 M - \varepsilon.
\end{aligned}
$$

Furthermore, by positive recurrence of $\big(\tau(X_{\mathbf{e}_k})\big)_{k\in\mathbb{N}_0}$ there is an almost surely finite random time $\mathbf{T} \geq \mathbf{T_L}$ such that

$$
-\frac{1}{\mathbf{e_T}}\log\frac{1-F(-\tau(X_{\mathbf{e_T}}))}{1-F(-0)} \in (-\varepsilon,\varepsilon) \quad \text{and} \quad \frac{1}{\mathbf{e_T}}\log\frac{\overline{G}_0(1)}{\overline{G}_{\tau(X_{\mathbf{e_T}})}(1)} \in (-\varepsilon,\varepsilon).
$$

Choose now $M > (h+4\varepsilon)/\ell_0$. We obtain the desired contradiction when we substitute in equality (27) the vertex x by $X_{\mathbf{e_T}}$, divide by $\mathbf{e_T}$ on both sides and see that the left side is in $(h-\varepsilon, h+\varepsilon)$ and the rightmost side is larger than $h+\varepsilon$. □

This finishes the proof of Theorem 3.9. □

Proof of Corollary 3.10. Recall Inequality (24). Integrating both sides of this inequality yields together with the inequality chain (26) that

$$
\int \liminf_{n\to\infty} -\frac{\log \pi_n(X_n)}{n} - h\, d\mathbb{P} = 0,
$$

providing that $h = \liminf_{n\to\infty} -\frac{1}{n}\log \pi_n(X_n)$ almost surely. □

4.7 Proof of Theorem 3.11

To prove the theorem we need the following lemma:

Lemma 4.8. *Under the assumptions of Theorem 3.9,*

$$
-\frac{1}{n}\log \pi_n(X_n) \xrightarrow{\mathbb{P}} h,
$$

that is, $-\frac{1}{n}\log \pi_n(X_n)$ converges in probability to the asymptotic entropy.

Proof. For every $\delta_1 > 0$ there is some index N_{δ_1} such that for all $n \geq N_{\delta_1}$

$$
\int -\frac{1}{n}\log \pi_n(X_n)\, d\mathbb{P} \in (h-\delta_1, h+\delta_1).
$$

Furthermore, due to Corollary 3.10 there is for every $\delta_2 > 0$ some index N_{δ_2} such that for all $n \geq N_{\delta_2}$

$$
\mathbb{P}\Big[-\frac{1}{n}\log \pi_n(X_n) > h - \delta_1\Big] > 1 - \delta_2. \tag{28}
$$

Since $h = \lim_{n\to\infty} \int -\frac{1}{n} \log \pi_n(X_n)\, d\mathbb{P}$ it must be that for every arbitrary but fixed $\varepsilon > \delta_1$ and for n large enough

$$\mathbb{P}\Big[-\frac{1}{n}\log \pi_n(X_n) > h - \delta_1\Big] \cdot (h-\delta_1) + \mathbb{P}\Big[-\frac{1}{n}\log \pi_n(X_n) > h + \varepsilon\Big] \cdot (\varepsilon + \delta_1) \leq h + \delta_1,$$

or equivalently,

$$\mathbb{P}\Big[-\frac{1}{n}\log \pi_n(X_n) > h + \varepsilon\Big] \leq \frac{h + \delta_1 - \mathbb{P}\Big[-\frac{1}{n}\log \pi_n(X_n) > h - \delta_1\Big] \cdot (h-\delta_1)}{\varepsilon + \delta_1}.$$

If we let $\delta_2 \to 0$, we get

$$\limsup_{n\to\infty} \mathbb{P}\Big[-\frac{1}{n}\log \pi_n(X_n) > h + \varepsilon\Big] \leq \frac{2\delta_1}{\varepsilon + \delta_1}.$$

Since we can choose δ_1 arbitrarily small we get

$$\mathbb{P}\Big[-\frac{1}{n}\log \pi_n(X_n) > h + \varepsilon\Big] \xrightarrow{n\to\infty} 0 \quad \text{for all } \varepsilon > 0.$$

But this yields convergence in probability of $-\frac{1}{n}\log \pi_n(X_n)$ to h together with (28). $\square$

For any small $\varepsilon > 0$ and $n \in \mathbb{N}$, we define the events

$$A_{n,\varepsilon} := \Big[\Big|-\frac{1}{n}\log \pi_n(X_n) - h\Big| \leq \varepsilon\Big] \text{ and } B_{n,\varepsilon} := \Big[\Big|-\frac{1}{n}\log \pi_n(X_n) - h\Big| > \varepsilon\Big].$$

There is some $N_\varepsilon \in \mathbb{N}$ such that $\mathbb{P}[B_{n,\varepsilon}] < \varepsilon$ for all $n \geq N_\varepsilon$. Since $0 \leq -\frac{1}{n}\log \pi_n(X_n) \leq \log \varepsilon_0$ we can conclude from Lemma 4.8 for $n \geq N_\varepsilon$:

$$\begin{aligned} & \int \Big|-\frac{1}{n}\log \pi_n(X_n) - h\Big|\, d\mathbb{P} \\ = & \int_{A_{n,\varepsilon}} \Big|-\frac{1}{n}\log \pi_n(X_n) - h\Big|\, d\mathbb{P} + \int_{B_{n,\varepsilon}} \Big|-\frac{1}{n}\log \pi_n(X_n) - h\Big|\, d\mathbb{P} \\ \leq & \;\varepsilon + \varepsilon \log \varepsilon_0 \xrightarrow{\varepsilon\to 0} 0. \end{aligned}$$

Thus, we have proved the theorem.

$\square$

Acknowledgement

The authors are grateful to the German Research Foundation (DFG) for supporting their projects, to which this work belongs. We acknowledge also the fruitful interplay and exchange with the Austrian Science Fund (FWF) project FWF-P19115-N18 at Graz University of Technology.

References

[1] A. Avez. Entropie des groupes de type fini. *C.R.Acad.Sci.Paris Sér. A-B*, 275:A1363–A1366, 1972.

[2] I. Benjamini and Y. Peres. Tree-indexed random walks on groups and first passage percolation. *Probab. Theory Related Fields*, 98(1):91–112, 1994.

[3] D. Bertacchi and F. Zucca. Characterization of the critical values of branching random walks on weighted graphs through infinite-type branching processes, arXiv:0804.0224v1.

[4] M. Björklund. Central limit theorems for gromov hyperbolic groups. *To appear in J. Theoret. Probab.*, DOI 10.1007/s10959-009-0230-x, 2009.

[5] S. Blachère and S. Brofferio. Internal diffusion limited aggregation on discrete groups having exponential growth. Preprint available at arXiv:math/0507582v1 [math.PR], 2005.

[6] S. Blachère, P. Haïssinsky, and P. Mathieu. Asymptotic entropy and Green speed for random walks on countable groups. *Ann. of Probab.*, 36(3):1134–1152, 2008.

[7] T. Cover and J. Thomas. *Elements of Information Theory.* Wiley & Sons, 2nd edition, 2006.

[8] Y. Derriennic. Quelques applications du théorème ergodique sous-additif. *Astérisque*, 74:183–201, 1980.

[9] P. G. Doyle and J. L. Snell. *Random walks and electric networks*, volume 22 of *Carus Mathematical Monographs*. Mathematical Association of America, Washington, DC, 1984.

[10] G. Fayolle, V. A. Malyshev, and M. V. Men′shikov. *Topics in the constructive theory of countable Markov chains.* Cambridge University Press, Cambridge, 1995.

[11] K.-H. Förster and B. Nagy. Local spectral radii and Collatz-Wielandt numbers of monic operator polynomials with nonnegative coefficients. *Linear Algebra Appl.*, 268:41–57, 1998.

[12] A. S. Gairat, V. A. Malyshev, M. V. Menshikov, and K. D. Pelikh. Classification of Markov chains that describe the evolution of random strings. *Uspekhi Mat. Nauk*, 50(2(302)):5–24, 1995.

[13] N. Gantert, S. Müller, S. Popov, and M. Vachkovskaia. Survival of branching random walks in random environment. *To appear in J. Theoret. Probab.*, DOI 10.1007/s10959-009-0227-5, 2009.

[14] L. Gilch. Rate of escape of random walks on free products. *J. of Australian Math. Soc.*, 83(1):31–54, 2007.

[15] V. Kaimanovich. Hausdorff dimension of the harmonic measure on trees. *Ergodic Theory Dynam. Systems*, 18:631–660, 1998.

[16] V.A. Kaimanovich and A.M. Vershik. Random walks on discrete groups: boundary and entropy. *Ann. of Probab.*, 11:457–490, 1983.

[17] R. Lyons. Random walks and percolation on trees. *Ann. of Probab.*, 18:931–958, 1990.

[18] R. Lyons, with Y. Peres. *Probability on Trees and Networks.* In preparation. Current version available at `http://mypage.iu.edu/~rdlyons/`, 2008.

[19] S. Müller. A criterion for transience of multidimensional branching random walk in random environment. *Elect. J. of Probab.*, 13:1189–1202, 2008.

[20] T. Nagnibeda and W. Woess. Random walks on trees with finitely many cone types. *J. Theoret. Prob.*, 15(2):383–422, 2002.

[21] H. H. Schaefer. *Banach lattices and positive operators.* Springer-Verlag, New York, 1974. Die Grundlehren der mathematischen Wissenschaften, Band 215.

[22] E. Seneta. *Non-negative matrices and Markov chains*. Springer Series in Statistics. Springer, New York, 2006. Revised reprint of the second (1981) edition [Springer-Verlag, New York; MR0719544].

[23] C. Takacs. Random walk on periodic trees. *Elect. J. of Probab.*, 2:1–16, 1997.

[24] J. von Below. An index theory for uniformly locally finite graphs. *Lin. Alg. and its Applications*, 431:1–19, 2009.

[25] W. Woess. *Random walks on infinite graphs and groups*, volume 138 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2000.

[26] O. Zeitouni. Random walks in random environment. In *Lectures on probability theory and statistics*, volume 1837 of *Lecture Notes in Math.*, pages 189–312. Springer, Berlin, 2004.