Combinatorics, Probability and Computing (2005) 14, 491–522. © 2005 Cambridge University Press
doi:10.1017/S096354830500684X Printed in the United Kingdom

Asymptotic Enumeration of Spanning Trees

RUSSELL LYONS†

Department of Mathematics, Indiana University, Bloomington, IN 47405-5701, USA
(e-mail: rdlyons@indiana.edu, http://mypage.iu.edu/~rdlyons)

Received 17 March 2003; revised 8 December 2003

We give new formulas for the asymptotics of the number of spanning trees of a large graph. A special case answers a question of McKay [*Europ. J. Combin.* **4** 149–160] for regular graphs. The general answer involves a quantity for infinite graphs that we call 'tree entropy', which we show is a logarithm of a normalized determinant of the graph Laplacian for infinite graphs. Tree entropy is also expressed using random walks. We relate tree entropy to the metric entropy of the uniform spanning forest process on quasi-transitive amenable graphs, extending a result of Burton and Pemantle [*Ann. Probab.* **21** 1329–1371].

1. Introduction

Methods of enumeration of spanning trees in a finite graph G and relations to various areas of mathematics and physics have been investigated for more than 150 years. The number of spanning trees is often called the *complexity* of the graph, denoted here by $\tau(G)$. The best-known formula for the complexity, proved in every basic text on graph theory, is called the 'Matrix-Tree Theorem', which expresses it as a determinant. One is often interested in asymptotics of the complexity of a sequence of finite graphs that approach (in some sense) some infinite graph. Use of the Matrix-Tree Theorem typically involves calculating eigenvalues and their asymptotics.[1] This often leads to a formula of the form

$$\lim_{n\to\infty} \frac{1}{|\mathsf{V}(G_n)|} \log \tau(G_n) = \int \log f$$

† Research partially supported by NSF grants DMS-0103897 and DMS-0231224.

[1] It appears not to have been noticed before that in the case of Euclidean lattice graphs, one can dispense with eigenvalues and go directly to the limit by invoking Szegő's limit theorem. In one dimension, this theorem appears in, *e.g.*, Grenander and Szegő [20, p. 44]. In higher dimensions, it is due to Helson and Lowdenslager [24], but the form in which they present it does not involve determinants. See the last part of the proof of Theorem 5.11 in Lyons and Steif [30] to see how to transform their result to give the asymptotics of determinants. In any case, use of our formula is simpler still.

for some function f on the unit cube of some Euclidean space, where $\mathsf{V}(G_n)$ is the vertex set of G_n. One of the cases that is most well known, owing to its connection with domino tilings, is that where the graphs G_n approach the usual square lattice, *i.e.*, the nearest-neighbour graph on $\mathbb{Z}^2$. In this case, we obtain

$$\lim_{n\to\infty} |\mathsf{V}(G_n)|^{-1} \log \tau(G_n) = \int_0^1 \int_0^1 \log(4 - 2\cos 2\pi x - 2\cos 2\pi y)\, dx\, dy = 4\mathrm{G}/\pi,$$

where $\mathrm{G} := \sum_{k=0}^{\infty} (-1)^k/(2k+1)^2 = 0.9160^-$ is *Catalan's constant*. See, *e.g.*, Burton and Pemantle [11] and Shrock and Wu [41] and the references therein for this and several other such examples. We shall reformulate the Matrix-Tree Theorem as an infinite series whose terms have a probabilistic meaning (see Proposition 3.1). In fact, the terms of the series are 'local' quantities. This will permit us to find the asymptotic complexity in a very general setting (see Theorem 3.2). Other than assuming a bound on the average degree of the finite graphs of whose complexity we find the asymptotics, our results are essentially the most general possible. In the simplest case, where the finite graphs G_n tend to a fixed transitive infinite graph G, we shall prove that

$$\lim_{n\to\infty} |\mathsf{V}(G_n)|^{-1} \log \tau(G_n) = \mathbf{h}(G) := \log \deg_G(o) - \sum_{k\geqslant 1} p_k(o; G)/k,$$

where o is a fixed vertex of G and $p_k(o; G)$ is the probability that simple random walk started at o on G is again at o after k steps. We term $\mathbf{h}(G)$ the *tree entropy*[2] of G. We have chosen this terminology to reflect both its provenance as a limit of entropies per vertex of uniform spanning trees of finite graphs as well as its agreement (up to normalization) with the metric entropy of uniform spanning forests in certain infinite graphs (see Theorem 5.1), as we shall discuss shortly.

Our result allows us to find asymptotics that were not known previously. For a simple example, when the graphs G_n approach the regular tree of degree 4, then

$$\lim_{n\to\infty} |\mathsf{V}(G_n)|^{-1} \log \tau(G_n) = 3\log(3/2);$$

see Example 1. This answers a question of McKay [32], who had shown such a result only under stronger conditions on the sequence of graphs G_n, and who had noted (in the case of regular graphs) that, if sufficient, our conditions would be weakest possible.

In the past, it has been observed by direct calculation that one has the same asymptotic complexity both of rectangular portions of $\mathbb{Z}^d$ as well as of d-dimensional tori. Our first main result, Theorem 3.2, shows that this stability phenomenon is quite general; see Corollary 3.8.

Our infinite series representation of the asymptotic complexity has an additional benefit. Namely, even when an integral representation is known, it turns out that for numerical estimation, our infinite series can be much better than the integral[3]; see, *e.g.*, Felker and Lyons [17] for examples comparing the two approaches.

[2] Another possible name, 'combinatorial entropy', is already in use with a variety of meanings.

[3] In the $\mathbb{Z}^d$-transitive case, it is not hard to pass directly between our series representation and the standard integral representation.

In order to state our second main result, we recall the Matrix-Tree Theorem. Given a graph G, define the following matrices indexed by the vertices of G. Let D_G be the diagonal matrix whose (x,x)-entry equals the degree of x: this is the *degree matrix*. Let A_G be the matrix whose (x,y)-entry equals the number of edges joining x and y: this is the *adjacency matrix*. Finally, let $\Delta_G := D_G - A_G$, the *graph Laplacian matrix*. The Matrix-Tree Theorem says that given any finite graph G, every cofactor of Δ_G is the same, namely, $\tau(G)$ (see, *e.g.*, Godsil and Royle [19, Lemma 13.2.3]). Another version of the Matrix-Tree Theorem states that when G is connected,

$$\tau(G) = |\mathsf{V}(G)|^{-1}\det{}'\Delta_G, \tag{1.1}$$

where $\det{}'A$ denotes the product of the nonzero eigenvalues of a matrix A (see, *e.g.*, Godsil and Royle [19, Lemma 13.2.4]). Thus, the asymptotic complexity of a sequence of graphs is a limit of the logarithm of the determinant of the graph Laplacians, appropriately normalized. Our second main aim is to give meaning to and to prove the formula

$$\mathbf{h}(G) = \log \operatorname{Det} \Delta_G$$

directly in terms of a normalized determinant of the graph Laplacian for infinite graphs. We shall also use this formula to prove inequalities for tree entropy and to calculate easily and quickly the classical tree entropy for Euclidean lattices. Indeed, this result, Theorem 4.1, provides a quick way to derive virtually all the prior asymptotics of this type in the literature, while Theorem 3.2 gives the rest (and more).

We have alluded to the fact that tree entropy also arises as the entropy per vertex of a measure that is the weak limit of the uniform measures on spanning trees of finite graphs. To state this precisely, we must first recall the work of Pemantle [35]. He proved a conjecture of the present author, namely, that if an infinite connected graph G is exhausted by a sequence of finite connected subgraphs G_n, then the weak limit of the uniform spanning tree measures on G_n exists. However, it may happen that the limit measure is not supported on trees, but on forests. This limit measure is now called the *free uniform spanning forest* on G, denoted FSF, or FSF_G when we want to indicate the graph G. If G is itself a tree, then this measure is trivial, namely, it is concentrated on $\{G\}$. Therefore, Häggström [23] introduced another limit that had been considered on $\mathbb{Z}^d$ more implicitly by Pemantle [35] and explicitly by Häggström [22], namely, the weak limit of the uniform spanning tree measures on G_n^*, where G_n^* is the graph G_n with its boundary identified to a single vertex. As Pemantle [35] showed, this limit also always exists on any graph and is now called the *wired uniform spanning forest*, denoted WSF or WSF_G. It is clear that both FSF and WSF are concentrated on the set of spanning forests[4] of G that are *essential*, meaning that all their trees are infinite. Furthermore, since the limits exist regardless of the exhaustion chosen, both FSF_G and WSF_G are invariant under any automorphisms that G may have. As shown by Häggström [22], $\mathsf{FSF}_G = \mathsf{WSF}_G$ on every amenable transitive graph G such as $\mathbb{Z}^d$. In fact, the proof of this result given by Benjamini, Lyons, Peres and Schramm [7], hereinafter referred to as BLPS [7], is easily

[4] By a 'spanning forest', we mean a subgraph without cycles that contains every vertex.

modified to show that the same holds for every quasi-transitive amenable graph, using the notion of natural frequency distribution recalled below in Section 5. Both FSF and WSF are important in their own right; see Lyons [29] for a survey and BLPS [7] for a comprehensive treatment.

Suppose now that Γ is an amenable group that acts quasi-transitively on a graph G. Since FSF_G is defined as a limit of uniform measures, it is natural to expect that it has maximum (metric) entropy in an appropriate class of Γ-invariant measures. Since the set of essential spanning forests is closed, one generally considers the class of Γ-invariant measures on this set. Furthermore, one might expect that the entropy of FSF_G is related in a simple way to the exponential growth rate of the complexity of the finite subgraphs G_n exhausting G. Finally, one might wonder whether FSF_G is unique as a measure of maximal entropy on the essential spanning forests.

The case of $\Gamma = \mathbb{Z}^d$ was considered by Burton and Pemantle [11]. Their Theorem 6.1 gave a positive answer to all three questions implicit in the preceding paragraph. However, it turns out that one of the claims, the uniqueness of the measure of maximal entropy, received an incorrect proof. The correct proofs of the other claims relied to some extent on the natural tiling of $\mathbb{Z}^d$ by large boxes, which is not available in general amenable quasi-transitive graphs. The third aim of this paper is to extend these two other claims to amenable quasi-transitive graphs, that is, to prove that the formula for entropy is correct and to prove that this entropy is maximal (see Theorem 5.1). We have not been able to prove uniqueness of the measure of maximal entropy. Therefore, the deduction by Burton and Pemantle [11] of the uniqueness of the measure of maximal entropy for domino tilings in $\mathbb{Z}^2$ (their Theorem 7.2) remains incomplete. However, a new and more general proof for the uniqueness of the measure of maximal entropy for domino tilings has been found by Sheffield [40]. The arguments in Burton and Pemantle [11] can then be used to deduce the uniqueness of the measure of maximal entropy for spanning trees in $\mathbb{Z}^2$ (but not in higher dimensions, which remains open).

We give general background and definitions related to graphs, groups, and entropy in Section 2. In Section 3, we give our reformulation of the Matrix-Tree Theorem and its asymptotics, together with several examples. We give the formula for tree entropy as a determinant on infinite graphs in Section 4, with applications. We prove the assertions on metric entropy in Section 5.

Note added in proof. It has just been discovered that there is an error in [2]. Namely, there is a gap in the proof that all unimodular random rooted graphs are random weak limits of finite graphs. This makes the proofs of certain results here incomplete. Remark 7, Proposition 4.3, and Theorem 4.4 will be complete when restricted to random weak limits of finite graphs, rather than claimed for all unimodular random rooted graphs.

2. Background

We denote a (multi-)graph G with vertex set V and edge set E by $G = (\mathsf{V}, \mathsf{E})$. When there is more than one graph under discussion, we write $\mathsf{V}(G)$ or $\mathsf{E}(G)$ to avoid ambiguity. We denote the degree of a vertex x in a graph G by $\deg_G(x)$. Unless stated otherwise, we assume all degrees finite. *Simple random walk* on G is the Markov chain whose state space

is V and whose transition probability from x to y equals the number of edges joining x to y divided by $\deg_G(x)$.

An infinite path in a tree that starts at any vertex and does not backtrack is called a *ray*. Two rays are *equivalent* if they have infinitely many vertices in common. An equivalence class of rays is called an *end*. More generally, an *end* of an infinite graph G is an equivalence class of infinite simple paths in G, where two paths are equivalent if, for every finite subgraph $K \subset G$, there is a connected component of $G \setminus K$ that intersects both paths.

Let G be a graph. For a subgraph H, let its (*internal*) *vertex boundary* $\partial_{\mathsf{V}} H$ be the set of vertices of H that are adjacent to some vertex not in H. We say that G is *amenable* if there is an exhaustion $H_1 \subset H_2 \subset \cdots \subset H_n \subset \cdots$ with $\bigcup_n H_n = G$ and

$$\lim_{n\to\infty} \frac{|\partial_{\mathsf{V}} H_n|}{|\mathsf{V}(H_n)|} = 0.$$

Such an exhaustion (or the sequence of its vertex sets) is called a *Følner sequence*. A finitely generated group is *amenable* if its Cayley graph is amenable. For example, every finitely generated abelian group is amenable. For more on amenability of graphs and groups, see Benjamini, Lyons, Peres and Schramm [6], hereinafter referred to as BLPS [6].

A *homomorphism* $\varphi : G_1 \to G_2$ from one graph $G_1 = (\mathsf{V}_1, \mathsf{E}_1)$ to another $G_2 = (\mathsf{V}_2, \mathsf{E}_2)$ is a pair of maps $\varphi_{\mathsf{V}} : \mathsf{V}_1 \to \mathsf{V}_2$ and $\varphi_{\mathsf{E}} : \mathsf{E}_1 \to \mathsf{E}_2$ such that φ_{V} maps the endpoints of e to the endpoints of $\varphi_{\mathsf{E}}(e)$ for every edge $e \in \mathsf{E}_1$. When both maps $\varphi_{\mathsf{V}} : \mathsf{V}_1 \to \mathsf{V}_2$ and $\varphi_{\mathsf{E}} : \mathsf{E}_1 \to \mathsf{E}_2$ are bijections, then φ is called an *isomorphism*. When $G_1 = G_2$, an isomorphism is called an *automorphism*. The group of all automorphisms of G will be denoted by $\mathrm{Aut}(G)$. The action of a group Γ on a graph G by automorphisms is said to be *transitive* if there is only one Γ-orbit in $\mathsf{V}(G)$ and to be *quasi-transitive* if there are only finitely many orbits in $\mathsf{V}(G)$. A graph G is *transitive* or *quasi-transitive* according as whether the corresponding action of $\mathrm{Aut}(G)$ is. For example, every Cayley graph is transitive.

The action of a group on a set is called *free* if the stabilizer of each element of the set is just the identity element of the group. For example, every group acts on itself freely by multiplication.

A locally compact group is called *unimodular* if its left Haar measure is also right-invariant. In particular, every discrete countable group is unimodular. We call a graph G *unimodular* if $\mathrm{Aut}(G)$ is unimodular, where $\mathrm{Aut}(G)$ is given the weak topology generated by its action on G. Every Cayley graph and, as Soardi and Woess [42] and Salvatori [39] proved, every quasi-transitive amenable graph is unimodular. See BLPS [6] for more details on unimodular graphs.

We now recall some definitions pertaining to entropy. For simplicity, we restrict our discussion to groups acting on graphs, which is enough for our purposes. Suppose first that μ is a probability measure on a finite or countable set, X. The *entropy* of μ is

$$\mathbf{H}(\mu) := -\sum_{x\in X} \mu(x) \log \mu(x).$$

For example, suppose that μ is a probability measure on the Borel sets of $2^{\mathsf{E}} := \{0,1\}^{\mathsf{E}}$ in the product topology. As usual, we identify an element of 2^{E} with the subset (or

'configuration') of edges where the value 1 is taken. If H is a finite subgraph of G, write $\mu\restriction H$ for the restriction of μ to the σ-field generated by the restrictions to $\mathsf{E}(H)$. That is, $\mu\restriction H$ is defined on the finite set $2^{\mathsf{E}(H)}$ and has an entropy as above.

We shall twice use the following well-known lemma (see, *e.g.*, Lemma 6.2 of Burton and Pemantle [11] and p. 11 of Bollobás [10] for the bound on the binomial coefficient sum).

Lemma 2.1. *Let Y be a finite set and m be a positive integer. Write $\alpha := m/|Y|$. Suppose that μ is a probability measure on $2^Y \times 2^Y$ that is supported on the set of pairs (ω_1, ω_2) with $|\omega_1 \triangle \omega_2| \leqslant m$. Let μ_1 and μ_2 be the coordinate marginals of μ. Then*

$$|\mathbf{H}(\mu_1) - \mathbf{H}(\mu_2)| \leqslant \log \sum_{k=0}^{m} \binom{|Y|}{k} \leqslant |Y|\big(-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)\big).$$

Suppose that the finite group Γ acts on X and preserves the measure μ. Then the *entropy* of the pair (μ, Γ) is

$$\mathbf{H}(\mu, \Gamma) := |\Gamma|^{-1}\mathbf{H}(\mu).$$

Finally, suppose that Γ is a countable amenable finitely generated subgroup of $\mathrm{Aut}(G)$. Let μ be a probability measure on 2^{E} that is preserved by Γ. Let $\langle \Gamma_n \rangle$ be a Følner sequence in Γ and H be a finite subgraph of G such that $\Gamma H = G$, provided such an H exists. Then the (*metric*) *entropy* of the pair (μ, Γ), also called the Γ-*entropy* of μ, is

$$\mathbf{H}(\mu, \Gamma) := \lim_{n\to\infty} |\Gamma_n|^{-1}\mathbf{H}\big(\mu\restriction(\Gamma_n H)\big).$$

This does not depend on the choice of H. See Ornstein and Weiss [34] for more details on entropy.

3. Asymptotic complexity

Recall that $\tau(G)$ denotes the complexity of the graph G, *i.e.*, the number of spanning trees of G. Let $p_k(x; G)$ denote the probability that simple random walk on G started at x is back at x after k steps.

We begin with a formula for the complexity of finite graphs.

Proposition 3.1. *Suppose that G is a finite connected graph. Then*

$$\log \tau(G) = -\log(2|\mathsf{E}(G)|) + \sum_{x\in\mathsf{V}(G)} \log \deg_G(x) - \sum_{k\geqslant 1} \frac{1}{k}\left(\sum_{x\in\mathsf{V}(G)} p_k(x; G) - 1\right).$$

Proof. Let P be the transition matrix for simple random walk on G and I be the identity matrix of the same size. As shown by Runge and Sachs [36], we may rewrite (1.1) as

$$\tau(G) = \frac{\prod_{x\in\mathsf{V}(G)} \deg_G(x)}{\sum_{x\in\mathsf{V}(G)} \deg_G(x)} \det{}'(I - P)$$

(the proof follows from looking at the coefficient of t in $\det(I - P - tI) = (\det D_G)^{-1} \det(\Delta_G - tD_G)$ and using the Matrix-Tree Theorem in its original form with cofactors). Since the sum of the degrees of a graph equals twice the number of its edges, we obtain

$$\log \tau(G) = -\log(2|\mathsf{E}(G)|) + \sum_{x \in \mathsf{V}(G)} \log \deg_G(x) + \log \det{}'(I - P). \tag{3.1}$$

Let Λ be the multiset of eigenvalues of P other than 1 (with multiplicities). Since $\Lambda \subset [-1, 1)$, we may rewrite the last term of (3.1) as

$$\begin{aligned} \log \det{}'(I - P) &= \sum_{\lambda \in \Lambda} \log(1 - \lambda) = -\sum_{\lambda \in \Lambda} \sum_{k \geqslant 1} \lambda^k / k \\ &= -\sum_{k \geqslant 1} \sum_{\lambda \in \Lambda} \lambda^k / k = -\sum_{k \geqslant 1} \frac{1}{k} (\mathrm{tr} P^k - 1), \end{aligned}$$

where in the last step, we have used the fact that the eigenvalue 1 of P has multiplicity 1 since G is connected. Since $\mathrm{tr} P^k = \sum_{x \in \mathsf{V}(G)} p_k(x; G)$, the desired formula now follows from this and (3.1). □

Remark 1. There is an extension of Proposition 3.1 to any irreducible Markov chain with transition matrix P. In this case, a spanning tree of the associated directed graph, with all edges leading towards a single vertex, is often called a 'spanning arborescence'. Let $\tau'(P)$ be the sum over all spanning trees (with all possible roots) of the product of $P(e)$ over all e in the tree. The analogue of Proposition 3.1 states that

$$\log \tau'(P) = -\sum_{k \geqslant 1} \frac{1}{k} \left(\sum_{x \in \mathsf{V}(G)} p_k(x; G) - 1 \right).$$

This is reminiscent of a formula that appears in Lind and Tuncel [27]. In fact, combining their formula with ours, one gets an expression for the derivative at 1 of the reciprocal of the so-called stochastic zeta function of P.

A *rooted graph* (G, x) is a graph G with a distinguished vertex x of G, called the *root*. A *rooted isomorphism* of rooted graphs is an isomorphism of the underlying graphs that takes the root of one to the root of the other. We shall use the following notion of random weak convergence introduced by Benjamini and Schramm [5] and studied further by Aldous and Steele [3] and Aldous and Lyons [2]. Given a positive integer R, a finite rooted graph H, and a probability distribution ρ on rooted graphs, let $p(R, H, \rho)$ denote the probability that H is rooted isomorphic to the ball of radius R about the root of a graph chosen with distribution ρ. If (G, ρ) is a graph with probability distribution ρ on its vertices, then ρ induces naturally a distribution on rooted graphs, which we also denote by ρ; namely, the probability of (G, x) is $\rho(x)$. For a finite graph G, let $U(G)$ denote the distribution of rooted graphs obtained by choosing a uniform random vertex of G as root of G. Suppose that $\langle G_n \rangle$ is a sequence of finite graphs and that ρ is a probability measure on rooted infinite graphs; in most practical cases, ρ will be induced by a probability distribution on the vertices of a fixed infinite graph. We say the *random*

weak limit of $\langle G_n \rangle$ is ρ if, for any positive integer R and any finite graph H, we have $\lim_{n\to\infty} p\big(R, H, U(G_n)\big) = p(R, H, \rho)$. More generally, if G_n are random finite graphs, then we say the *random weak limit* of $\langle G_n \rangle$ is ρ if for any positive integer R, any finite graph H, and any $\epsilon > 0$, we have $\lim_{n\to\infty} \mathbf{P}\big[|p\big(R, H, U(G_n)\big) - p(R, H, \rho)| > \epsilon\big] = 0$. Note that only the component of the root matters for convergence to ρ. Thus, we may and shall assume that ρ is concentrated on connected graphs. If ρ is induced by a distribution on the vertices of a fixed transitive graph G, then the random weak limit depends only on G and not on the root. In this case, we say that the *random weak limit* of $\langle G_n \rangle$ is G.

Given $R > 0$ and a finite graph G, let $\nu_R(G)$ be the distribution of the number of edges in the ball of radius R about a random vertex of G. Call a collection of finite graphs G *tight* if for each R, the collection of corresponding distributions $\nu_R(G)$ is tight. Note that any tight collection of finite graphs has a subsequence that possesses a random weak limit.

Define the *expected degree* of a probability measure ρ on rooted graphs to be

$$\overline{\deg}(\rho) := \int \deg_G(x)\, d\rho(G, x).$$

When the following integral converges, define the *tree entropy* of ρ to be

$$\mathbf{h}(\rho) := \int \left(\log \deg_G(x) - \sum_{k \geqslant 1} \frac{1}{k} p_k(x; G) \right) d\rho(G, x).$$

Our terminology is justified by Theorems 3.2 and 5.1 below. The integral converges, for example, when ρ has finite expected degree, *i.e.*, $\overline{\deg}(\rho) < \infty$, by virtue of the inequality between the arithmetic and geometric means; see also Corollary 3.5 and Proposition 4.3 below. If ρ is induced by a fixed transitive graph, G, of degree d, we write

$$\mathbf{h}(G) := \mathbf{h}(\rho) = \log d - \sum_{k \geqslant 1} \frac{1}{k} p_k(o; G)$$

for the tree entropy of G, where o is any vertex of G.

Our main theorem in this section is the following. It also suggests thinking of $\mathbf{h}(\rho)$ as a normalized logarithm of the determinant of the Laplacian. See Theorem 4.1 for a more direct reason for thinking thus.

Theorem 3.2. *If G_n are finite connected graphs with bounded average degree whose random weak limit is a probability measure ρ on infinite rooted graphs, then*

$$\lim_{n\to\infty} \frac{1}{|\mathsf{V}(G_n)|} \log \tau(G_n) = \mathbf{h}(\rho).$$

The same limit holds in probability when G_n are random with bounded expected average degree.

Remark 2. If there is a homomorphism that is not an isomorphism from one transitive graph G_1 onto another G_2 of the same degree, then it is clear that $p_k(o; G_1) \leqslant p_k(o; G_2)$ for every k, with strict inequality for some k. Therefore, $\mathbf{h}(G_1) > \mathbf{h}(G_2)$. In particular, among

all transitive graphs G of degree d, the maximum of $\mathbf{h}(G)$ is achieved uniquely for G the regular tree of degree d. This maximum value is calculated explicitly in Example 1. The uniqueness of the maximum implies that if a sequence of finite regular graphs G_n of degree d does not tend to the d-regular tree, T, then $\liminf_n \log \tau(G_n)/|\mathsf{V}(G_n)| < \mathbf{h}(T)$; this is Theorem 4.5 of McKay [32].

The only difficulty in deducing Theorem 3.2 from Proposition 3.1 is the interchange of limit and summation. In order to accomplish this, we shall use the following lemmas.

Lemma 3.3. *Let P be the transition matrix of a Markov chain. For $\alpha \in [0,1]$, define the transition matrix $Q := \alpha I + (1-\alpha)P$. For a state x, let $p_k(x)$ and $q_k(x)$ denote the return probabilities to x after k steps when the Markov chains start at x, where the transition matrices are P and Q, respectively. We have*

$$\sum_k q_k(x)/k = -\log(1-\alpha) + \sum_k p_k(x)/k. \tag{3.2}$$

Proof. Write $(\cdot,\cdot)$ for the ordinary inner product in $\ell^2(V)$, where V is the state space. Given $x \in V$ and $z \in (0,1)$, we have

$$\begin{aligned}
\sum_k q_k(x) z^k/k &= -\Big(\big[\log(I - zQ)\big]\mathbf{1}_{\{x\}}, \mathbf{1}_{\{x\}}\Big) \\
&= -\Big(\big[\log\big((1-z\alpha)I - z(1-\alpha)P\big)\big]\mathbf{1}_{\{x\}}, \mathbf{1}_{\{x\}}\Big) \\
&= -\left(\left[\log\left(I - \frac{z(1-\alpha)}{1-z\alpha}P\right)\right]\mathbf{1}_{\{x\}}, \mathbf{1}_{\{x\}}\right) - \log(1-z\alpha)\big(\mathbf{1}_{\{x\}}, \mathbf{1}_{\{x\}}\big) \\
&= -\log(1-z\alpha) + \sum_k p_k(x)\left(\frac{z(1-\alpha)}{1-z\alpha}\right)^k \frac{1}{k}.
\end{aligned}$$

Letting $z \uparrow 1$, we obtain the desired equation. □

Next we give a universal bound for the rate of convergence of $p_k(x;G)$ to the stationary probability $\deg_G(x)/2|\mathsf{E}(G)|$ for finite graphs and to 0 for infinite graphs. In the case of infinite graphs, such a result is first due to Carlen, Kusuoka and Stroock [12]. Our argument is a modification of that of Coulhon [14] and seems not to be written anywhere for the case of finite graphs, although there is some overlap with the treatment of the special case in Example 2.3.1 of Saloff-Coste [38]. For sharper bounds that depend on more information about the graph, see, *e.g.*, Theorems 3.3.11 and 2.3.1 of Saloff-Coste [38] for the finite case and Barlow, Coulhon and Grigor'yan [4] for the infinite case. See also Morris and Peres [33].

Lemma 3.4. *Suppose that Q is a transition matrix of a Markov chain that is reversible with respect to a positive measure π. If π is finite, then we assume that π is normalized to be a probability measure. Assume that $c := \inf\{\pi(x)Q(x,y)\,;\ x \neq y$ and $Q(x,y) > 0\} > 0$ and*

that $a := \inf_x Q(x,x) > 0$. *For all states* x *and all* $k \geqslant 0$, *we have*

$$|Q^k(x,x)/\pi(x) - 1| \leqslant \min\left\{ \frac{1}{ac\sqrt{k+1}},\ \frac{1}{2a^2c^2(k+1)} \right\} \tag{3.3}$$

if π *is finite and*

$$Q^k(x,x)/\pi(x) \leqslant \frac{1}{ac\sqrt{k+1}} \tag{3.4}$$

if π *is infinite.*

Proof. Write V for the state space and E for the set of pairs (x,y) where $Q(x,y) > 0$ and $x \neq y$. Write

$$c_2(x,y) := \pi(x)Q^2(x,y)$$

and note that for $(x,y) \in E$, we have

$$c_2(x,y) \geqslant \pi(x)[Q(x,x)Q(x,y) + Q(x,y)Q(y,y)] \geqslant 2ac.$$

Define the inner product $(f_1, f_2)_\pi := \sum_x f_1(x)f_2(x)\pi(x)$ on $\ell^2(V,\pi)$.

The case where π is infinite is simpler, so we treat that first. Let f be a function on V with finite support. Let x_0 be a vertex where $|f|$ achieves its maximum. Then

$$\|f\|_\infty = |f(x_0)| \leqslant \frac{1}{2}\sum_{(x,y)\in E} |f(x) - f(y)| \leqslant \frac{1}{2}\sum_{x,y\in V} c_2(x,y)|f(x) - f(y)|/(2ac), \tag{3.5}$$

where the factor $1/2$ arises from counting each pair (x,y) in each order. Applying (3.5) to the function f^2, we obtain

$$\begin{aligned}
(2ac)^2\|f\|_\infty^4 &\leqslant \left(\frac{1}{2}\sum_{x,y\in V} c_2(x,y)|f(x) - f(y)| \cdot |f(x) + f(y)| \right)^2 \\
&\leqslant \left(\frac{1}{2}\sum_{x,y\in V} c_2(x,y)[f(x) - f(y)]^2 \right) \left(\frac{1}{2}\sum_{x,y\in V} c_2(x,y)[f(x) + f(y)]^2 \right) \\
&= \big((I - Q^2)f, f\big)_\pi \big((I + Q^2)f, f\big)_\pi
\end{aligned}$$

by the Cauchy–Schwarz inequality and some straightforward algebra. Therefore, if $(f,f)_\pi \leqslant 1$, we have $2(ac)^2\|f\|_\infty^4 \leqslant \big((I - Q^2)f, f\big)_\pi$. Apply this inequality to $Q^l f$ for $l = 0, \dots, k$ and sum these inequalities to obtain

$$\begin{aligned}
(k+1)2(ac)^2\|Q^k f\|_\infty^4 \leqslant 2(ac)^2 \sum_{l=0}^{k} \|Q^l f\|_\infty^4 &\leqslant \sum_{l=0}^{k} \big((I - Q^2)Q^l f, Q^l f\big)_\pi \\
&= \sum_{l=0}^{k} \big((I - Q^2)Q^{2l} f, f\big)_\pi = \big((I - Q^{2k+2})f, f\big)_\pi \leqslant 1
\end{aligned}$$

for $(f,f)_\pi \leqslant 1$. This shows that the norm of $Q^k : \ell^2(V,\pi) \to \ell^\infty(V)$ is bounded by

$$\beta_k := [(2k+2)(ac)^2]^{-1/4}.$$

The same bound holds for $Q^k : \ell^1(V,\pi) \to \ell^2(V,\pi)$ by duality. Therefore, considering $Q^{2k} = Q^k \circ Q^k$, we find that the norm of $Q^{2k} : \ell^1(V,\pi) \to \ell^\infty(V)$ is at most β_k^2, while the norm of $Q^{2k+1} : \ell^1(V,\pi) \to \ell^\infty$ is at most $\beta_k\beta_{k+1}$. Applying these inequalities to $f := \mathbf{1}_x/\pi(x)$ gives (3.4).

The case of finite π is quite similar. The essential difference is that we work with $\ell_0^2(V,\pi)$, the orthogonal complement of the constants in $\ell^2(V,\pi)$. Note that $\mathbf{1}$ is an eigenfunction of Q and that $\ell_0^2(V,\pi)$ is invariant under Q. We may still conclude (3.5) for all f that have at least one nonnegative value and at least one nonpositive value, such as all $f \in \ell_0^2(V,\pi)$.

Take $f \in \ell_0^2(V,\pi)$. Notice that $\sum_{x,y\in V} c_2(x,y) = \sum_{x\in V} \pi(x) = 1$. Thus, we have from (3.5) that

$$(2ac)^2\|f\|_\infty^2 \leqslant \frac{1}{2}\sum_{x,y\in V} c_2(x,y)[f(x)-f(y)]^2 = \big((I-Q^2)f,\, f\big)_\pi. \tag{3.6}$$

Alternatively, we may apply (3.5) to the function $\operatorname{sgn}(f)f^2$. Using the trivial inequality

$$|\operatorname{sgn}(s)s^2 - \operatorname{sgn}(t)t^2| \leqslant |s-t|\cdot\big(|s|+|t|\big),$$

valid for any real numbers s and t, we obtain that

$$\begin{aligned}(2ac)^2\|f\|_\infty^4 &\leqslant \left(\frac{1}{2}\sum_{x,y\in V} c_2(x,y)|f(x)-f(y)|\cdot\big(|f(x)|+|f(y)|\big)\right)^2\\ &\leqslant \big((I-Q^2)f,\, f\big)_\pi\big((I+Q^2)|f|,\, |f|\big)_\pi.\end{aligned}$$

Putting both these estimates together, we get

$$2(ac)^2\max\{2\|f\|_\infty^2,\, \|f\|_\infty^4\} \leqslant \big((I-Q^2)f,\, f\big)_\pi$$

for $(f,\, f)_\pi = 1$. As before, this implies that

$$(k+1)2(ac)^2\max\{2\|Q^kf\|_\infty^2,\, \|Q^kf\|_\infty^4\} \leqslant 1,$$

which shows that the norm of $Q^k : \ell_0^2(V,\pi) \to \ell^\infty(V)$ is bounded by

$$\alpha_k := \min\{[(2ac)^2(k+1)]^{-1/2},\, [(2k+2)(ac)^2]^{-1/4}\}.$$

Let $T : \ell^2(V,\pi) \to \ell_0^2(V,\pi)$ be the orthogonal projection $Tf := f - (f,\, \mathbf{1})_\pi\mathbf{1}$. Given what we have shown, we see that the norm of $Q^kT : \ell^2(V,\pi) \to \ell^\infty(V)$ is bounded by α_k. By duality, the same bound holds for $TQ^k : \ell^1(V,\pi) \to \ell^2(V,\pi)$. As before, we deduce that the norm of $Q^kTQ^k : \ell^1(V,\pi) \to \ell^\infty(V)$ is at most α_k^2 and the norm of $Q^kTQ^{k+1} : \ell^1(V,\pi) \to \ell^\infty$ is at most $\alpha_k\alpha_{k+1}$. Applying these inequalities to $f := \mathbf{1}_x/\pi(x)$ gives (3.3). □

Remark 3. In the infinite case, we do not actually need to assume that $a > 0$. That is, suppose that Q is a transition matrix of a Markov chain that is reversible with respect to a positive measure π. Assume that $c := \inf\{\pi(x)Q(x,y)\,;\ x \neq y \text{ and } Q(x,y) > 0\} > 0$. If π is infinite, then

$$Q^k(x,x)/\pi(x) \leqslant \frac{4}{c\sqrt{k+1}}.$$

To see this, we have only to establish that

$$|f(x_0)| \leqslant (1/c) \sum_{x,y \in V} c_2(x,y)|f(x) - f(y)|$$

as a substitute for (3.5). Let $E_2 := \{(x,y)\,;\ Q^2(x,y) > 0\}$. Because π is infinite, for any finite set S of states, there is some y such that $Q(y,S)Q(y,S^c) > 0$, where $Q(y,A) := \sum_{z \in A} Q(y,z)$. Note that $Q(y,S) + Q(y,S^c) = 1$. Either $Q(y,S)$ or $Q(y,S^c)$ is at least $1/2$ and $\pi(y)$ times the other is at least c, whence

$$\pi(y)Q(y,S)Q(y,S^c) \geqslant c/2.$$

It follows that

$$\begin{aligned}\sum_{x \in S} \sum_{z \notin S} c_2(x,z) &= \sum_{x \in S} \pi(x) \sum_{z \notin S} Q^2(x,z) = \sum_{x \in S} \pi(x) \sum_{y} \sum_{z \notin S} Q(x,y)Q(y,z) \\ &= \sum_{y} \pi(y)Q(y,S)Q(y,S^c) \geqslant c/2.\end{aligned}$$

That is, for any cutset of edges $e \in E_2$ that separates x_0 from infinity in the graph (V, E_2), the sum of $c_2(e)$ over the cutset is at least $c/2$. Therefore, the max-flow min-cut theorem provides a flow θ from x_0 to infinity of value $c/2$ that is bounded by c_2 on each edge in E_2. This yields

$$\begin{aligned}|f(x_0)|c &= \left| \sum_{(x,y) \in E_2} \big[f(x) - f(y)\big]\theta(x,y) \right| \leqslant \sum_{(x,y) \in E_2} |f(x) - f(y)| \cdot |\theta(x,y)| \\ &\leqslant \sum_{(x,y) \in E_2} |f(x) - f(y)| c_2(x,y),\end{aligned}$$

as desired. (Recall that each pair (x,y) is counted twice in the sum.)

Remark 4. The proofs as given show the very same bounds on the more general quantities $|Q^k(y,x)/\pi(y) - 1|$ or $Q^k(y,x)/\pi(y)$ for all states x, y and all $k \geqslant 0$.

Corollary 3.5. *If ρ is a probability measure on infinite rooted graphs with finite expected degree, then $\mathbf{h}(\rho)$ is finite.*

Proof. For a graph G with transition matrix P, let $Q := (I + P)/2$. Then Q is the transition matrix of the graph G' obtained from G by adding loops to each vertex so as to double its degree. Write $q_k(x;G) := p_k(x;G')$. Lemma 3.3 tells us that

$$\sum_k q_k(x;G)/k = \log 2 + \sum_k p_k(x;G)/k.$$

In addition, Q is reversible with respect to the measure $x \mapsto \deg_G(x)$, so that Lemma 3.4 applies with $a \geqslant 1/2$ and $c \geqslant 1/2$ (equality holds in both cases when G has no loops) to yield

$$q_k(x;G) \leqslant \frac{4}{\sqrt{k+1}} \deg_G(x).$$

Therefore,

$$\begin{aligned}\mathbf{h}(\rho) &= \log 2 + \int \left(\log \deg_G(x) - \sum_{k \geqslant 1} q_k(x; G)/k \right) d\rho(G, x) \\ &\geqslant \log 2 + \int \left(\log \deg_G(x) - \sum_{k \geqslant 1} \frac{4}{k\sqrt{k+1}} \deg_G(x) \right) d\rho(G, x) \\ &= \log 2 + \int \log \deg_G(x)\, d\rho(G, x) - \overline{\deg}(\rho) \sum_{k \geqslant 1} \frac{4}{k\sqrt{k+1}}.\end{aligned}$$

This gives the corollary by the inequality between the arithmetic and geometric means. □

The following lemma is well known.

Lemma 3.6. *Suppose that Y_n are real-valued random variables that converge in distribution to Y and that $\sup_n \mathbf{E}\big[|Y_n|\big] < \infty$. Then for all continuous functions $f : \mathbb{R} \to \mathbb{R}$ such that $\lim_{|x| \to \infty} |f(x)|/|x| = 0$, we have $\lim_{n \to \infty} \mathbf{E}\big[f(Y_n)\big] = \mathbf{E}\big[f(Y)\big]$.*

Proof. The hypotheses easily imply that $f(Y_n)$ form a uniformly integrable set of random variables. The continuity of f ensures that $f(Y_n)$ converge in distribution to $f(Y)$. Together, these imply the conclusion. □

Proof of Theorem 3.2. We claim that the more general second statement of the theorem follows from the first statement. The space of rooted-isomorphism classes of graphs is complete, separable and metrizable (see Aldous and Steele [3] for some of the details). Thus, under the hypothesis of the second statement, we may assume by Skorohod's theorem (see, *e.g.*, Theorem 11.7.2 of Dudley [15]) that G_n are defined on a common probability space such that a.s., G_n has a random weak limit ρ. Therefore, if the first statement holds, then so does the second.

We now prove the first statement. Double the degree of each vertex in G_n by adding loops to give graphs G'_n. These new graphs have transition matrices $Q_n := (I + P_n)/2$, where P_n is the transition matrix of G_n. Furthermore, the random weak limit of G'_n is ρ', where ρ' is obtained from ρ by doubling the degree of each vertex by adding loops. By Lemma 3.3, we have $\mathbf{h}(\rho') = \mathbf{h}(\rho)$. Also, $\tau(G'_n) = \tau(G_n)$, so it suffices to show that

$$\lim_{n \to \infty} \frac{1}{|\mathsf{V}(G'_n)|} \log \tau(G'_n) = \mathbf{h}(\rho').$$

Let d be an upper bound for the average degree of G'_n, *i.e.*, for all n,

$$2|\mathsf{E}(G'_n)| \leqslant d|\mathsf{V}(G'_n)|, \tag{3.7}$$

so that $|\mathsf{V}(G'_n)|^{-1} \log\big(2|\mathsf{E}(G'_n)|\big) \to 0$ as $n \to \infty$. Since the degree of a random vertex in G'_n converges in distribution to the degree of the root under ρ', it follows that

$$\frac{1}{|\mathsf{V}(G'_n)|} \sum_{x \in \mathsf{V}(G'_n)} \log \deg_{G'_n}(x) \to \int \log \deg_G(x)\, d\rho'(G, x)$$

by Lemma 3.6 (setting $f := \log^+$ and Y_n to be the degree of a uniform vertex in G'_n). Thus, in using Proposition 3.1, we need only show that

$$\lim_{n\to\infty} \sum_{k\geqslant 1} |\mathsf{V}(G'_n)|^{-1} \frac{1}{k} \left(\sum_{x\in\mathsf{V}(G'_n)} p_k(x; G'_n) - 1 \right) = \int \sum_{k\geqslant 1} \frac{1}{k} p_k(x; G)\, d\rho'(G, x).$$

By definition and the hypothesis, we have for each k that

$$\lim_{n\to\infty} |\mathsf{V}(G'_n)|^{-1} \left(\sum_{x\in\mathsf{V}(G'_n)} p_k(x; G'_n) - 1 \right) = \int p_k(x; G)\, d\rho'(G, x).$$

Lemma 3.4 applies to G'_n with stationary probability measure $x \mapsto \deg_{G'_n}(x)/\big[2|\mathsf{E}(G'_n)|\big]$ and constants $a \geqslant 1/2$, $c \geqslant 1/[4|\mathsf{E}(G'_n)|]$ to yield

$$|\mathsf{V}(G'_n)|^{-1} \left| \sum_{x\in\mathsf{V}(G'_n)} p_k(x; G'_n) - 1 \right| \leqslant \frac{4d}{\sqrt{k+1}}.$$

Hence Weierstrass's M-test justifies the interchange of limit and summation and we are done. □

Remark 5. A similar result holds for weighted graphs. That is, given a graph G whose edges are assigned positive weights, write $\tau(G)$ for the sum of the weights of its spanning trees, where the weight of a spanning tree is the product of the weights of its edges. Let the weight of a vertex be the sum of the weights of its incident edges. The random walk corresponding to a weighted graph has transition probability from x to y equal to the sum of the weights of the edges joining x to y divided by the weight of x. If a sequence of weighted connected finite graphs with weights bounded above and away from 0 and with bounded average vertex weight has a random weak limit ρ on weighted rooted infinite graphs, then the conclusion of Theorem 3.2 holds, where $\mathbf{h}(\rho)$ is defined using the weight of the root in place of its degree and by using the weighted random walk on the limit graph.

We now illustrate some of the consequences of Theorem 3.2. Our first result concerns the stability of the asymptotic complexity, for which we prepare with a lemma related to tightness.

Let $B_R(x) = \big(\mathsf{V}_R(x), \mathsf{E}_R(x)\big)$ denote the ball of radius R about a vertex x.

Lemma 3.7. *Let $\langle G_n \rangle$ be a sequence of finite graphs with vertex subsets $W_n \subset \mathsf{V}(G_n)$ satisfying $\lim_{n\to\infty} |W_n|/v_n = 0$, where $v_n := |\mathsf{V}(G_n)|$. For $R > 0$, let $s_n(R, t) := |\{x \in \mathsf{V}(G_n)\,;\ |\mathsf{V}_R(x)| > t\}|$. Let also $w_n(R) := |\{x \in \mathsf{V}(G_n)\,;\ \mathsf{V}_R(x) \cap W_n \neq \emptyset\}|$. If for each $R > 0$ we have*

$$\lim_{t\to\infty} \limsup_{n\to\infty} s_n(R, t)/v_n = 0, \tag{3.8}$$

then for each $R > 0$ we have $\lim_{n\to\infty} w_n(R)/v_n = 0$.

Proof. We have that $\mathsf{V}_R(x) \cap W_n \neq \emptyset$ if and only if x lies in a ball of radius R about some vertex of W_n. If we partition W_n in two, one part consisting of those vertices x with

$|\mathsf{V}_R(x)| \leqslant t$ and the other part consisting of the rest, then we deduce that

$$w_n(R) \leqslant t|W_n| + s_n(2R, t)$$

for any $R > 0$ and $t > 0$. Thus,

$$\limsup_{n\to\infty} w_n(R)/v_n \leqslant \limsup_{n\to\infty} s_n(2R, t)/v_n.$$

If we now let $t \to \infty$, we obtain the desired result. □

Note that (3.8) holds if $\langle G_n \rangle$ has a random weak limit. In fact, (3.8) is just slightly weaker than tightness, since this condition counts vertices, while tightness counts edges.

Corollary 3.8. *Let $\langle G_n \rangle$ be a tight sequence of finite connected graphs with bounded average degree such that $\lim_{n\to\infty} |\mathsf{V}(G_n)|^{-1} \log \tau(G_n) = h$. If $\langle G'_n \rangle$ is a sequence of connected subgraphs of $\langle G_n \rangle$ such that*

$$\lim_{n\to\infty} |\mathsf{V}(G_n)|^{-1} |\{x \in \mathsf{V}(G'_n)\ ;\ \deg_{G'_n}(x) = \deg_{G_n}(x)\}| = 1,$$

then $\lim_{n\to\infty} |\mathsf{V}(G'_n)|^{-1} \log \tau(G'_n) = h$.

Proof. By taking a subsequence, if necessary, we may assume (by tightness) that $\langle G_n \rangle$ has a random weak limit, ρ. By Theorem 3.2, we have $h = \mathbf{h}(\rho)$. Let $W_n := \{x \in \mathsf{V}(G'_n)\ ;\ \deg_{G'_n}(x) \neq \deg_{G_n}(x)\}$. Then (3.8) holds because of tightness, whence $\langle G'_n \rangle$ also has the random weak limit ρ by Lemma 3.7. Hence Theorem 3.2 applies again to give the desired conclusion. □

We next illustrate the flexibility of Theorem 3.2 by considering hybrid graphs as follows.

Corollary 3.9. *Let $\langle G_n \rangle$ and $\langle G'_n \rangle$ be tight sequences of finite connected graphs with bounded average degree such that*

$$\lim_{n\to\infty} |\mathsf{V}(G_n)|^{-1} \log \tau(G_n) = h \qquad \textit{and} \qquad \lim_{n\to\infty} |\mathsf{V}(G'_n)|^{-1} \log \tau(G_n) = h'.$$

Suppose that

$$\lim_{n\to\infty} \frac{|\mathsf{V}(G_n)|}{|\mathsf{V}(G_n)| + |\mathsf{V}(G'_n)|} = \alpha \in [0, 1].$$

Let H_n be formed by connecting disjoint copies of G_n and G'_n with $o\big(|\mathsf{V}(G_n)| + |\mathsf{V}(G'_n)|\big)$ edges in any manner that gives a sequence of connected graphs. Then

$$\lim_{n\to\infty} \frac{1}{|\mathsf{V}(H_n)|} \log \tau(H_n) = \alpha h + (1 - \alpha) h'.$$

Proof. By taking subsequences, if necessary, we may assume that $\langle G_n \rangle$ and $\langle G'_n \rangle$ have random weak limits, ρ and ρ'. By Theorem 3.2, we have $h = \mathbf{h}(\rho)$ and $h' = \mathbf{h}(\rho')$. By Lemma 3.7, the random weak limit of H_n is $\alpha\rho + (1 - \alpha)\rho'$. Thus Theorem 3.2 gives the desired conclusion. □

Given probability measures ρ_n and ρ on rooted graphs, we say that ρ_n *converges weakly* to ρ if $p(R,H,\rho_n) \to p(R,H,\rho)$ as $n \to \infty$ for any positive integer R and any finite graph H. It is not hard to show that tree entropy is a continuous functional when one bounds the expected degree.

Proposition 3.10. *If ρ_n converges weakly to ρ as $n \to \infty$ with $\sup_n \overline{\deg}(\rho_n) < \infty$, then $\mathbf{h}(\rho_n) \to \mathbf{h}(\rho)$ as $n \to \infty$.* □

Proof. As in the proof of Theorem 3.2, we may double the degree of each vertex by adding loops without changing the tree entropies. By Lemma 3.6 and our assumption of bounded expected degree, we have $\int \log \deg_G(x)\, d\rho_n(G,x) \to \int \log \deg_G(x)\, d\rho(G,x)$. Weak convergence itself already guarantees that $\int p_k(x;G)\, d\rho_n(G,x) \to \int p_k(x;G)\, d\rho(G,x)$ for each k. The bounded expected degree and Lemma 3.4 allow us to apply Weierstrass's M-test to get the desired conclusion. □

We now give several explicit examples illustrating the use of tree entropy, beginning with the transitive case.

In order to evaluate the infinite sum appearing in $\mathbf{h}(G)$, the following integral is sometimes useful. Let $\mathscr{G}(z) := \mathscr{G}(z;G) := \sum_{k\geqslant 0} p_k(o;G)z^k$ be the *return probability-generating function* of the graph G. Then clearly

$$\sum_{k\geqslant 1} \frac{1}{k} p_k(o;G) = \int_0^1 \frac{\mathscr{G}(z)-1}{z}\, dz. \tag{3.9}$$

Example 1. For a group Γ with a given generating set, let $\ell(\Gamma)$ denote the length of the smallest (nonempty) reduced word in the generating elements that represents the identity, *i.e.*, the girth of the Cayley graph of Γ. Suppose that Γ_n are finite groups, each generated by s elements, such that $\lim_{n\to\infty} \ell(\Gamma_n) = \infty$. Then the Cayley graphs G_n of Γ_n have a random weak limit equal to the usual Cayley graph G of the free group Γ on s letters, *i.e.*, the regular tree of degree $2s$. By Theorem 3.2, it follows that

$$\lim_{n\to\infty} |\Gamma_n|^{-1} \log \tau(G_n) = \mathbf{h}(G),$$

independently of the particular choice of Γ_n. To evaluate $\mathbf{h}(G)$, we use the fact that the return series is

$$\mathscr{G}(z;G) = \frac{1 - s + \sqrt{s^2 - (2s-1)z^2}}{1 - z^2},$$

a result of Kesten [25]. The integrand in (3.9) then has an 'elementary' antiderivative, which yields

$$\mathbf{h}(G) = \log \frac{(2s-1)^{2s-1}}{[4s(s-1)]^{s-1}}.$$

For example, for $s = 2$, we find $\mathbf{h}(G) = \log(3/2)^3$. More generally, when G is the regular tree of degree d, we have

$$\mathscr{G}(z; G) = \frac{2(d-1)}{d - 2 + \sqrt{d^2 - 4(d-1)z^2}}$$

(see, *e.g.*, Lemma 1.24 of Woess [45]), whence

$$\mathbf{h}(G) = \log \frac{(d-1)^{d-1}}{[d(d-2)]^{d/2-1}}.$$

For example, $\mathbf{h}(G) = \log(4/\sqrt{3})$ if $d = 3$. As we mentioned in the introduction, this calculation of the asymptotic complexity of regular graphs with girth tending to infinity was first done by McKay [32] under additional hypotheses on the graphs. For the case of (uniformly) random d-regular graphs, where it is easy to see that they have a random weak limit equal to the d-regular tree, G, we obtain that the asymptotic complexity tends in probability to $\mathbf{h}(G)$; this was also shown by McKay [32], who showed in [31] that random regular graphs satisfy his extra hypotheses.

Example 2. The usual Erdős–Rényi model of random graphs, $\mathscr{G}(n, p)$, is a graph on n vertices, each pair of which is connected by an edge with probability p, independently of other edges. Another name for this is Bernoulli(p) bond percolation on the complete graph K_n. Fix $c > 1$. The well-known fact that the entire graph $\mathscr{G}(n, c/n)$ has a random weak limit $\mathsf{PGW}(c)$ is proved explicitly in Aldous [1], where $\mathsf{PGW}(c)$ is the law of a rooted Galton–Watson tree with Poisson(c) offspring distribution. It is well known that with probability approaching 1 as $n \to \infty$, there is a unique connected component, called the *giant component*, of $\mathscr{G}(n, c/n)$, that has $\Omega(n)$ vertices. See, *e.g.*, Bollobás [10]. Also, the giant component has a random weak limit $\mathsf{PGW}^*(c)$, which is $\mathsf{PGW}(c)$ conditioned on nonextinction. This limit of the giant component is folklore and seems not to be written anywhere. Let $f(c) := \mathbf{h}\big(\mathsf{PGW}^*(c)\big)$. We also define $\mathsf{PGW}^*(1)$ to be the weak limit of $\mathsf{PGW}^*(c)$ as $c \downarrow 1$. Since $\mathsf{PGW}^*(1)$ is the random weak limit of trees (more specifically, of the uniform spanning tree on K_n) by Grimmett [21], we have $f(1) = 0$ (which also follows from Theorem 4.4 below). By Proposition 3.10, f is continuous on $[1, \infty)$. We wonder whether f is strictly increasing on $[1, \infty)$ and real analytic on $(1, \infty)$. The fact that $f(c) > 0$ for $c > 1$ follows from Theorem 4.4 below, together with the well-known fact that $\mathsf{PGW}^*(c)$ has infinitely many ends a.s. for $c > 1$. It would be interesting to see an explicit formula for f.

An additional useful tool for calculation is explained in Section 9 of Woess [45]. Namely, as explained there, if we let $\mathbf{r}$ be the radius of convergence of $\mathscr{G}(\cdot)$, then there is a strictly increasing function $\Phi := \Phi_G : \big[0, \mathbf{r}\mathscr{G}(\mathbf{r})\big) \to \mathbb{R}$ such that

$$\mathscr{G}(z) = \Phi\big(z\mathscr{G}(z)\big) \tag{3.10}$$

on $\big[0, \mathbf{r}\mathscr{G}(\mathbf{r})\big)$. We are grateful to Woess for pointing out to us the remainder of this paragraph. In many cases, it is easier to find Φ than to (solve and) find $\mathscr{G}$. Note that from

(3.10) we have

$$\mathscr{G}(1) = \Phi\big(\mathscr{G}(1)\big); \tag{3.11}$$

it can be much easier to solve for $\mathscr{G}(1)$ and use that in (3.13) below than it is to solve (3.10) for $\mathscr{G}$ and use that in (3.9). Substitute (3.10) in the right-hand side of (3.9) to obtain

$$\int_0^1 \frac{\mathscr{G}(z)-1}{z}dz = \int_0^1 \frac{\Phi\big(z\mathscr{G}(z)\big)-1}{z}dz. \tag{3.12}$$

Now use the change of variable $t := z\mathscr{G}(z)$. This gives us

$$\begin{aligned}\int_0^1 \frac{\Phi\big(z\mathscr{G}(z)\big)-1}{z}dz &= \int_0^{\mathscr{G}(1)} \frac{\big(\Phi(t)-1\big)\big(\Phi(t)-t\Phi'(t)\big)}{t\Phi(t)}dt \\ &= \int_0^{\mathscr{G}(1)} \frac{\Phi(t)-1}{t}dt - \Phi\big(\mathscr{G}(1)\big) + 1 + \log\Phi\big(\mathscr{G}(1)\big) \\ &= \int_0^{\mathscr{G}(1)} \frac{\Phi(t)-1}{t}dt - \mathscr{G}(1) + 1 + \log\mathscr{G}(1).\end{aligned}$$

Thus, we have

$$\mathbf{h}(G) = \log d - \int_0^{\mathscr{G}(1)} \frac{\Phi(t)-1}{t}dt + \mathscr{G}(1) - 1 - \log\mathscr{G}(1). \tag{3.13}$$

If we are interested only in the asymptotic complexity of finite graphs, for which ρ would we want to calculate $\mathbf{h}(\rho)$? This is answered in Section 4, where we shall see that all ρ whose underlying graph is a fixed Cayley graph are included, among others.

Example 3. Suppose that G is the free product $K_{s_1} * \cdots * K_{s_n}$ of the complete graphs $K_{s_1}, \ldots, K_{s_n}$ for some integers $s_j \geqslant 2$ with $\sum_j s_j \geqslant 5$. In other words, G is the Cayley graph of the free product of groups of order s_j with respect to the generating set corresponding to every element of the factors other than the identities. In order to calculate $\mathbf{h}(G)$, we shall find it easier to work with the graph G', in which we have added n loops to each vertex of G. By Lemma 3.3, we have $\mathbf{h}(G') = \mathbf{h}(G)$. Let $d := \sum_{j=1}^n s_j$ be the degree of G'. Now by Woess [44], we have

$$\Phi_{G'}(z) = 1 + \frac{z-n}{2} + \frac{1}{2}\sum_{j=1}^n \sqrt{(1-s_jz/d)^2 + 4z/d}.$$

For example, if $n = 2$, then (3.11) gives that $\mathscr{G}(1, G') = s_1s_2/(s_1s_2 - s_1 - s_2)$. We then find via (3.13) applied to G' that

$$\mathbf{h}(G) = \left(1 - \frac{1}{s_1}\right)\log(s_1 - 1) + \left(1 - \frac{1}{s_2}\right)\log(s_2 - 1) + \left(1 - \frac{1}{s_1} - \frac{1}{s_2}\right)\log\frac{s_1 + s_2}{s_1s_2 - s_1 - s_2}.$$

For example, if $s_1 = 2$ and $s_2 = 3$, when G is a Cayley graph of the modular group $\mathrm{PSL}(2, \mathbb{Z})$ (use the generators $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$), then $\mathbf{h}(G) = \log\big(2^{2/3}5^{1/6}\big)$. As examples of other tree entropies that one may find by similar means, we mention that $\mathbf{h}(G) = \log(16/3)$ if $n = 3$ with $s_1 = s_2 = s_3 = 3$, while

$$\mathbf{h}(G) = \log(61 + 9\sqrt{57}) + \frac{1}{6}\log(317 - 33\sqrt{57}) - \frac{7}{2}\log 2 - \log 7 = 1.190^+$$

if $n = 3$ with $s_1 = s_2 = 2$ and $s_3 = 3$. As an application, suppose that H_n is a random 3-regular 3-uniform hypergraph on n vertices. Let G_n be the associated graph in which every hyperedge of H_n is replaced by a clique on its vertices. Then G_n tends weakly to $K_3 * K_3 * K_3$, whence $n^{-1} \log \tau(G_n) \to \log(16/3)$.

Example 4. Choose a ray $\langle x_m \,;\, m \geqslant 0\rangle$ in the regular tree T of degree 3. Let G_m be the ball in T of radius m about x_m. Remove the edge $[x_m, x_{m+1}]$ from G_m; let T_m be the connected component of x_m that remains. Thus, $T_m \subset T_{m+1}$ for all m. Let $G := \bigcup_{m\geqslant 1} T_m$. The random weak limit of $\langle G_m\rangle$ is ρ, where $\rho(G, x_m) := 2^{-m-1}$ for $m \geqslant 0$. Since $\tau(G_m) = 1$, Theorem 3.2 tells us that $\mathbf{h}(\rho) = 0$, *i.e.*,

$$\sum_{k\geqslant 1} \frac{1}{k} \sum_{m\geqslant 0} 2^{-m-1} p_k(x_m; G) = \log \sqrt{3}.$$

We comment finally on situations where the average degree of G_n is not bounded. We suspect that the following holds for 'lazy' simple random walk $Q_G := (I + P_G)/2$ on any *simple* (unweighted) graph for some universal constant C:

$$\forall k \geqslant 2 \qquad |\mathsf{V}(G)|^{-1}\big(\mathrm{tr} Q_G^k - 1\big) \leqslant \frac{C}{\log^2 k}. \tag{3.14}$$

(It may even be true with $\log^2 k$ replaced by something like $k^{1/3}$, but as Ben Morris has pointed out, nothing better than $k^{1/3}$ is possible, as shown by the example of two cliques on n vertices joined by a path of length n.) If (3.14) holds, this would replace Lemma 3.4 (except in the proof of Corollary 3.5) and allow results still more general than Theorem 3.2 for simple graphs.

As we have not been able to establish (3.14), we consider instead sequences $\langle G_n\rangle$ that are an expanding family, meaning that the second largest eigenvalue $\lambda_2(G_n)$ of P_{G_n} is bounded away from 1. In this case, we do not need Lemma 3.4, since if $Q_n = Q_{G_n}$ is the transition matrix in the proof of Theorem 3.2, we have

$$|\mathsf{V}(G_n)|^{-1}\big(\mathrm{tr} Q_n^k - 1\big) \leqslant \left(\frac{1 + \lambda_2(G_n)}{2}\right)^k$$

for all $n, k \geqslant 1$. In addition, we do not need Lemma 3.3; that can be replaced with the corresponding statement for finite Markov chains,

$$\sum_{k\geqslant 1} (\mathrm{tr} Q^k - 1)/k = -\log(1-\alpha) + \sum_{k\geqslant 1} (\mathrm{tr} P^k - 1)/k,$$

where P and Q are as in Lemma 3.3.

If the average degree of $\langle G_n\rangle$ is unbounded, we must consider a different normalization of the complexity. Let us assume that the limit (as $n \to \infty$) of the return probability after k steps of simple random walk started at a random vertex of G_n exists; denote it by p_k. For example, if $\langle G_n\rangle$ has a random weak limit ρ that is concentrated on infinite graphs of finite degree, then $p_k = \int p_k(o; G)\, d\rho(G, o)$. For another example, $p_k = 0$ when $k \geqslant 1$ for simple graphs whose minimum degree tends to infinity. The proof of Theorem 3.2

(as modified above) shows that

$$\lim_{n\to\infty} |\mathsf{V}(G_n)|^{-1}\left[\log\tau(G_n) - \sum_{x\in\mathsf{V}(G_n)} \log\deg x\right] = -\sum_{k\geqslant 1} p_k/k.$$

For a particular example, if G_n is the giant component of the random graph $\mathscr{G}(n,p_n)$ with $np_n - \log n \to \infty$ (which has n vertices with probability tending to 1; see, *e.g.*, Theorem 9 of Chapter VII in Bollobás [9]), one has

$$\lim_{n\to\infty}\left(n^{-1}\log\tau(G_n) - \log(p_n n)\right) = 0$$

in probability. To facilitate comparison to Cayley's theorem (that $\tau(K_n) = n^{n-2}$), we may state this as $\tau(G_n) = \left([1+o(1)]p_n n\right)^n$. (The fact that $\langle G_n\rangle$ is an expanding family a.s. is probably folklore. It can be proved as follows. First, the proof that $\mathscr{G}(n,p_n)$ is connected with probability approaching 1 is easily modified to show that its isoperimetric constant, also called conductance, is bounded away from 0 a.s. Second, a well-known inequality relating this constant to the second largest eigenvalue gives the result; see, *e.g.*, Chung [13].)

4. Tree entropy as log determinant

Under certain assumptions, Theorem 3.2 shows that if ρ is a random weak limit of finite connected graphs, G_n, then its tree entropy $\mathbf{h}(\rho)$ is a limit of the logarithm of the determinant of the graph Laplacians of G_n, normalized by omitting the zero eigenvalue and by dividing by the number of vertices of G_n. In fact, one may give a formula for $\mathbf{h}(\rho)$ directly in terms of a normalized determinant of the graph Laplacian for infinite graphs. This is our main purpose in the present section. We shall also use this formula to prove inequalities for tree entropy and to calculate easily and quickly the classical tree entropy for Euclidean lattices.

We first discuss the class of probability measures ρ to which our formula will apply. This class, the class of ρ that arise as limits of finite graphs, is the class of unimodular ρ, defined as follows. Given a rooted graph (G,x) and an edge e incident to x, define the involution $\imath(G,x,e) := (G,y,e)$, where y is the other endpoint of e. Given a probability measure ρ on rooted graphs, define the probability measure $\widehat{\rho}$ to be the law of the isomorphism class of (G,x,e), where (G,x) is chosen according to ρ and e is then chosen uniformly among the edges incident to x. Also, define $\widetilde{\rho}$ to be the (non-probability) measure that is the result of biasing $\widehat{\rho}$ by the degree of the root; that is, the Radon–Nikodým derivative of $\widetilde{\rho}$ with respect to $\widehat{\rho}$ at the isomorphism class of (G,x,e) is $\deg_G(x)$. (If the expected degree of the root is finite, one could obtain a probability measure from $\widetilde{\rho}$ by dividing by the expected degree; but this is not always the case.) The involution $\imath$ induces a pushforward map $\widetilde{\rho} \mapsto \imath_*\widetilde{\rho}$. We say that ρ is *unimodular* or *involution invariant* if $\imath_*\widetilde{\rho} = \widetilde{\rho}$. It is easy to see that any ρ that is a random weak limit of finite graphs is unimodular, as observed by Aldous and Steele [3], who introduced the notion of involution invariance; essentially the same observation occurs in Benjamini and Schramm [5]. The converse is much harder, but is established in Aldous and Lyons [2]. Intuitively, unimodularity means that, up to isomorphism, all vertices are equally likely to be the root. See Aldous and Lyons [2] for a comprehensive treatment of unimodular random networks. In particular, it is shown

there that a transitive graph is unimodular if and only if it is unimodular as a rooted random graph.

The preceding definitions and results extend easily to the class of rooted weighted graphs or multi-graphs (G, o), where $G = \big(\mathsf{V}(G), \mathsf{E}(G), w\big)$ and $w : \mathsf{E}(G) \to [0, \infty)$ is a weight function as in Remark 5. For $x \neq y \in \mathsf{V}(G)$, let $\Delta_G(x, y) := -\sum_e w(e)$, where the sum is over all the edges between x and y, and $\Delta_G(x, x) := \sum_e w(e)$, where the sum is over all non-loop edges incident to x. We assume that $\Delta_G(x, x) < \infty$ for all x. The associated random walk has the transition probability from x to y of $-\Delta_G(x, y)/\Delta_G(x, x)$. Extend the definition of *tree entropy* to probability measures on rooted weighted graphs by

$$\mathbf{h}(\rho) := \int \left(\log \Delta_G(o, o) - \sum_{k \geqslant 1} \frac{1}{k} p_k(o; G) \right) d\rho(G, o)$$

whenever this integral converges, where $p_k(o; G)$ is the return probability after k steps for the associated random walk.

The (*graph*) *Laplacian* Δ_G just defined determines an operator

$$f \mapsto \left(x \mapsto \sum_{y \in \mathsf{V}} \Delta_G(x, y) f(y) \right)$$

for functions f on V with finite support. This operator extends by continuity to a bounded linear operator on all of $\ell^2(\mathsf{V})$ when $\sup_x \Delta_G(x, x) < \infty$. If we do not have such a uniform bound, we proceed as follows. Let U_e be a uniform $[0, 1]$-valued random variable chosen independently for all e. Given $M \in \mathbb{Z}^+$, let G' be the random weighted graph obtained from G by letting the weight of e be $w(e)(1 - U_e/M)$. Now let G_M be the weighted graph formed from G' by changing the weight to 0 of those edges e whose weights are greater than M or which are not among the M largest weights of the edges incident to (or equal to) e. Clearly the matrix Δ_{G_M} converges to Δ_G entrywise a.s. as $M \to \infty$. Since Δ_{G_M} is a bounded self-adjoint nonnegative-definite operator, the operator $\log(\Delta_{G_M} + \epsilon I)$ is bounded for any $\epsilon > 0$, where I denotes the identity operator on $\ell^2(\mathsf{V})$. Let ρ_M be the law of (G_M, o) when (G, o) has the law of ρ. If ρ is unimodular, then so is ρ_M.

Now the logarithm of the determinant of a matrix equals the trace of the logarithm of the matrix. Furthermore, one usually *defines* the determinant via this equality when one has a trace on a von Neumann algebra. This is the approach we take.

The trace we use is defined by Aldous and Lyons [2], which we review here. Suppose that ρ is a unimodular probability measure on rooted weighted graphs. Let $T : (G, o) \mapsto T_{G,o}$ be a measurable assignment of bounded linear operators $T_{G,o} : \ell^2\big(\mathsf{V}(G)\big) \to \ell^2\big(\mathsf{V}(G)\big)$ with finite μ-essential supremum of the norms $\|T_{G,o}\|$. Let Alg be the von Neumann algebra of such operators T that are equivariant in the sense that for all isomorphisms $\gamma : G_1 \to G_2$ and all $o, x, y \in \mathsf{V}(G)$, we have $(T_{G_2,\gamma o}\mathbf{1}_{\{\gamma x\}}, \mathbf{1}_{\{\gamma y\}}) = (T_{G_1,o}\mathbf{1}_{\{x\}}, \mathbf{1}_{\{y\}})$. Since $T_{G,o}$ does not depend on $o \in G$ for $T \in \mathsf{Alg}$, we shall write T_G in place of $T_{G,o}$ for $T \in \mathsf{Alg}$. We define the *trace* of $T \in \mathsf{Alg}$ to be

$$\mathrm{Tr}(T) := \mathrm{Tr}_\rho(T) := \mathbf{E}\big[(T_G \mathbf{1}_{\{o\}}, \mathbf{1}_{\{o\}})\big] := \int (T_G \mathbf{1}_{\{o\}}, \mathbf{1}_{\{o\}})\, d\rho(G, o).$$

For self-adjoint operators A and B, recall that $A \leqslant B$ means that $\big((B-A)v, v\big) \geqslant 0$ for all vectors v. Our trace has the following properties: $\mathrm{Tr}(\cdot)$ is linear, $\mathrm{Tr}(T) \geqslant 0$ for $T \geqslant 0$, and $\mathrm{Tr}(ST) = \mathrm{Tr}(TS)$. In addition, for any increasing function $f : \mathbb{R} \to \mathbb{R}$ and any $S \leqslant T$, we have $\mathrm{Tr}\big(f(S)\big) \leqslant \mathrm{Tr}\big(f(T)\big)$.

In effect, we now show the trace formula $\mathbf{h}(\rho) = \log \mathrm{Det}_\rho \Delta_G := \mathrm{Tr}_\rho(\log \Delta_G)$. The determinant here is a so-called Fuglede–Kadison determinant. See Fuglede and Kadison [18].

Theorem 4.1. *If ρ is a unimodular probability measure on rooted infinite weighted graphs with*

$$\int |\log \Delta_G(o,o)| \, d\rho(G,o) < \infty \tag{4.1}$$

and

$$\int \frac{\Delta_G(o,o)}{\inf\{-\Delta_G(x,y)\,;\ x \neq y \text{ and } \Delta_G(x,y) \neq 0\}} d\rho(G,o) < \infty, \tag{4.2}$$

then

$$\mathbf{h}(\rho) = \lim_{M\to\infty} \lim_{\epsilon \downarrow 0} \mathrm{Tr}_\rho\big[\log(\Delta_{G_M} + \epsilon I)\big]. \tag{4.3}$$

Proof. It is obvious that $\Delta_{G_M} + \epsilon I$ is monotone increasing in ϵ. Furthermore, it is easy to check that Δ_{G_M} is monotone increasing in M. Since log is an increasing function, it follows that the limits exist in (4.3).

The condition (4.1) and Lebesgue's Dominated Convergence Theorem guarantee that

$$\lim_{M\to\infty} \int \log \Delta_G(o,o) \, d\rho_M(G,o) = \int \log \Delta_G(o,o) \, d\rho(G,o).$$

Clearly, for all $k > 0$,

$$\lim_{M\to\infty} \int p_k(o;G) \, d\rho_M(G,o) = \int p_k(o;G) \, d\rho(G,o).$$

By Remark 3 and (4.2), it follows that

$$\lim_{M\to\infty} \int \sum_{k\geqslant 1} \frac{1}{k} p_k(o;G) \, d\rho_M(G,o) = \int \sum_{k\geqslant 1} \frac{1}{k} p_k(o;G) \, d\rho(G,o).$$

Therefore, we have

$$\lim_{M\to\infty} \mathbf{h}(\rho_M) = \mathbf{h}(\rho). \tag{4.4}$$

Let D_G be the diagonal matrix that has the same diagonal as Δ_G. For $\epsilon > 0$, define $P_{G,\epsilon}$ by

$$(D_G + \epsilon I)(I - P_{G,\epsilon}) = \Delta_G + \epsilon I.$$

In other words, $P_{G,\epsilon}(x,x) = 0$ and $P_{G,\epsilon}(x,y) = -\Delta_G(x,y)/(D_G(x,x) + \epsilon)$ for $x \neq y$. The matrix $P_{G,\epsilon}$ defines a killed random walk that, at x, is killed (sent to an absorbing cemetery state outside the graph G) with probability $\epsilon/(D_G(x,x) + \epsilon)$ and transits to y

with probability $P_{G,\epsilon}(x,y)$. Let $p_k(o;G,\epsilon)$ be the return probability after k steps for the killed random walk. It is clear that as $\epsilon \downarrow 0$, we have $p_k(o;G,\epsilon) \uparrow p_k(o;G)$. If the degrees of G are bounded, then the norm of $P_{G,\epsilon}$ is less than 1, whence

$$\log(I - P_{G,\epsilon}) = -\sum_{k\geqslant 1} \frac{1}{k} P_{G,\epsilon}^k;$$

in particular,

$$-\sum_{k\geqslant 1} \frac{1}{k} p_k(o;G,\epsilon) = \big(\log(I - P_{G,\epsilon})\mathbf{1}_{\{o\}}, \mathbf{1}_{\{o\}}\big).$$

By a theorem of Fuglede and Kadison [18], we have therefore for any M,

$$\begin{aligned}
\mathbf{h}(\rho_M) &= \int (\log D_G \mathbf{1}_{\{o\}}, \mathbf{1}_{\{o\}}) - \sum_{k\geqslant 1} \frac{1}{k} p_k(o;G)\, d\rho_M(G,o) \\
&= \lim_{\epsilon\downarrow 0} \int \big(\log(D_G + \epsilon I)\mathbf{1}_{\{o\}}, \mathbf{1}_{\{o\}}\big) - \sum_{k\geqslant 1} \frac{1}{k} p_k(o;G,\epsilon)\, d\rho_M(G,o) \\
&= \lim_{\epsilon\downarrow 0} \int \big(\log(D_G + \epsilon I)\mathbf{1}_{\{o\}}, \mathbf{1}_{\{o\}}\big) + \big(\log(I - P_{G,\epsilon})\mathbf{1}_{\{o\}}, \mathbf{1}_{\{o\}}\big)\, d\rho_M(G,o) \\
&= \lim_{\epsilon\downarrow 0} \mathrm{Tr}_{\rho_M}\big[\log(D_G + \epsilon I) + \log(I - P_{G,\epsilon})\big] \\
&= \lim_{\epsilon\downarrow 0} \mathrm{Tr}_{\rho_M}\Big[\log\big((D_G + \epsilon I)(I - P_{G,\epsilon})\big)\Big] \\
&= \lim_{\epsilon\downarrow 0} \mathrm{Tr}_{\rho_M}\big[\log(\Delta_G + \epsilon I)\big].
\end{aligned}$$

Putting together this limit relation with that of (4.4), we obtain (4.3). □

Remark 6. As Δ_G is unchanged by the addition or deletion of loops, we see that neither is $\mathbf{h}(\rho)$.

As one indication of the usefulness of Theorem 4.1 beyond its theoretical interest, we show how it leads immediately to calculation of the classically known tree entropy for the nearest-neighbour graph on $\mathbb{Z}^d$. In this case, the space $\ell^2(\mathbb{Z}^d)$ is isometrically isomorphic to $L^2\big([0,1]^d\big)$ (with Lebesgue measure) via the Fourier transform. Under this isomorphism, the Laplace operator $\Delta_{\mathbb{Z}^d}$ becomes the operator of multiplication by the function $(s_1,\dots,s_d) \mapsto 2d - 2\sum_{i=1}^d \cos(2\pi s_i)$, the vector $\mathbf{1}_{\{o\}}$ becomes the constant function $\mathbf{1}$, and thus

$$\mathbf{h}(\mathbb{Z}^d) = \int_{[0,1]^d} \log\left(2d - 2\sum_{i=1}^d \cos(2\pi s_i)\right) ds.$$

More generally, suppose that G is a graph with vertex set $\mathbb{Z}^d \times K$ for a finite set K and with edge set that is invariant under the natural action of $\mathbb{Z}^d$. That is, for each $x \in \mathbb{Z}^d$, there is a $K \times K$ matrix L^x such that for $x, y \in \mathbb{Z}^d$ and $u, v \in K$,

$$\Delta_G\big((x,u),(y,v)\big) = L^{y-x}(u,v).$$

Such graphs G are called 'periodic' by Burton and Pemantle [11]. Consider the measure ρ that puts equal mass on each $\big(G,(\mathbf{0},u)\big)$ $(u \in K)$. We may regard Δ_G operating on $\ell^2(\mathbb{Z}^d \times K)$ as an operator T on $\ell^2\big(\mathbb{Z}^d;\ell^2(K)\big)$, that is, on the space of vector-valued functions $f : \mathbb{Z}^d \to \ell^2(K)$ with $\sum_{x\in\mathbb{Z}^d} \|f(x)\|^2_{\ell^2(K)} < \infty$. Under the Fourier isomorphism with $L^2\big([0,1]^d;\ell^2(K)\big)$, the operator T becomes the operator of multiplication by the matrix-valued function

$$M : (s_1, s_2, \dots, s_d) \mapsto \sum_{x\in\mathbb{Z}^d} L^x e^{2\pi i x\cdot s} \qquad (s = (s_1, s_2, \dots, s_d) \in [0,1]^d),$$

the vector $\mathbf{1}_{\{(\mathbf{0},u)\}}$ becomes the constant function $\mathbf{1}_{\{u\}} \in \ell^2(K)$, and thus

$$\begin{aligned}
\mathbf{h}(\rho) &= |K|^{-1} \sum_{u\in K} \big((\log \Delta_G)\mathbf{1}_{\{(\mathbf{0},u)\}},\ \mathbf{1}_{\{(\mathbf{0},u)\}}\big) \\
&= |K|^{-1} \sum_{u\in K} \int_{[0,1]^d} \Big(\big(\log M(s)\big)\mathbf{1}_{\{u\}},\ \mathbf{1}_{\{u\}}\Big)\, ds \\
&= |K|^{-1} \int_{[0,1]^d} \operatorname{tr}\big(\log M(s)\big)\, ds \\
&= |K|^{-1} \int_{[0,1]^d} \log\det M(s)\, ds.
\end{aligned}$$

This is a slightly simpler version of the formula in Theorem 6.1(b) of Burton and Pemantle [11].

We consider next some inequalities. If ρ_1 and ρ_2 are two probability measures on rooted weighted graphs, let us say that ρ_1 is *edge-dominated* by ρ_2 if there exists a probability measure ν on rooted graphs (G,o) with two sets of weights (w_1, w_2) such that for all edges e, we have $w_1(e) \leqslant w_2(e)$ and such that the law of (G,o,w_i) is ρ_i for $i = 1,2$. We call ν a *monotone coupling* of ρ_1 and ρ_2. When the weight of an edge is 0, one can regard it as being absent.

Theorem 4.2. *If $\rho_1 \neq \rho_2$ are unimodular probability measures on rooted weighted infinite loopless graphs that both satisfy* (4.1) *and* (4.2) *and ρ_1 is edge-dominated by ρ_2, then* $\mathbf{h}(\rho_1) < \mathbf{h}(\rho_2)$.

The proof relies on the following notion. A continuous function $f : (0,\infty) \to \mathbb{R}$ is called *operator monotone on* $(0,\infty)$ if, for any bounded self-adjoint operators A, B with spectrum in $(0,\infty)$ and $A \leqslant B$, we have $f(A) \leqslant f(B)$. Löwner [28] proved that the logarithm is an operator monotone function on $(0,\infty)$ (see also Chapter V of Bhatia [8]).

Proof. As in the proof of Theorem 4.1, we have that $\Delta_{G_M,w_1} \leqslant \Delta_{G_M,w_2}$. Since log is an operator monotone function on $(0,\infty)$, it follows that $\operatorname{Tr}_{\rho_1}\big(\log(\Delta_{G_M,w_2} + \epsilon I)\big) \leqslant \operatorname{Tr}_{\rho_2}\big(\log(\Delta_{G_M,w_1} + \epsilon I)\big)$, so that $\mathbf{h}(\rho_1) \leqslant \mathbf{h}(\rho_2)$ by Theorem 4.1. If $\mathbf{h}(\rho_1) = \mathbf{h}(\rho_2)$, then by Theorem 4.1, we have

$$\lim_{M\to\infty} \lim_{\epsilon\downarrow 0} \big[\operatorname{Tr}_{\rho_2}\big(\log(\Delta_{G_M,w_2} + \epsilon I)\big) - \operatorname{Tr}_{\rho_1}\big(\log(\Delta_{G_M,w_1} + \epsilon I)\big)\big] = 0.$$

Let ν be a monotone coupling of ρ_1 and ρ_2. Since log is an operator monotone function on $(0,\infty)$ with $\lim_{t\downarrow 0} t\log t = 0$, we may apply a result from Aldous and Lyons [2] to deduce that $\Delta_{G_M,w_1} = \Delta_{G_M,w_2}$ ν-a.s., *i.e.*, that $\rho_1 = \rho_2$. □

Remark 7. If there is a unimodular monotone coupling ν (via marked graphs) of probability measures $\rho \neq \rho'$ on rooted infinite graphs that have finite expected degree, where ρ is edge-dominated by ρ', then one can prove an explicit lower bound for the difference $\mathbf{h}(\rho') - \mathbf{h}(\rho)$. As shown in Aldous and Lyons [2], there is then a sequence $\langle (G_n, G_n') \rangle$ of pairs of finite connected graphs on the same vertex sets and with the edge set of G_n contained in the edge set of G_n' such that G_n (resp., G_n') has a random weak limit ρ (resp., ρ') with the average degree of G_n' tending to the ρ'-expected degree of the root. A counting argument then shows that $\mathbf{h}(\rho') - \mathbf{h}(\rho) \geqslant c^2(\log 2)/[d(d+1)^2]$, where $c := \nu\left[(G, G', o)\,;\ \deg_{G'}(o) \neq \deg_G(o)\right]$ and $d := \overline{\deg}(\rho')$.

Corollary 3.5 shows that $\mathbf{h}(\rho) > -\infty$ as long as ρ has finite expected degree. As an example where $\mathbf{h}(\rho) < 0$, consider ρ to be concentrated on the single rooted graph $(\mathbb{N}, 0)$. However, this is not possible in the unimodular case of unweighted graphs.

Proposition 4.3. *If ρ is a unimodular probability measure on rooted infinite (unweighted) graphs that has finite expected degree, then* $\mathbf{h}(\rho) \geqslant 0$.

Proof. Under these hypotheses, Aldous and Lyons [2] establish that ρ is the random weak limit of a sequence of finite connected graphs with bounded average degree. Thus, we may apply Theorem 3.2. □

Naturally, we wish to know when the tree entropy is 0.

Theorem 4.4. *If ρ is a unimodular probability measure on rooted infinite (unweighted) graphs that has finite expected degree, then* $\mathbf{h}(\rho) = 0$ *iff* $\overline{\deg}(\rho) = 2$ *iff ρ-a.s. G is a locally finite tree with* 1 *or* 2 *ends.*

Proof. The last equivalence is proved in Aldous and Lyons [2]. To prove the first equivalence, let G_n be finite connected graphs whose random weak limit is ρ and with bounded average degree. Let T_n be a spanning tree of G_n. Since $\langle G_n \rangle$ is a tight sequence, so is $\langle T_n \rangle$. Therefore, by taking a subsequence if necessary, we may assume that $\langle T_n \rangle$ has a random weak limit ρ_0. Clearly, ρ_0 is edge-dominated by ρ and $\mathbf{h}(\rho_0) = 0$ (since $\tau(T_n) = 1$). Aldous and Lyons [2] proved that $\overline{\deg}(\rho_0) = 2$. Thus, $\overline{\deg}(\rho) = 2$ iff $\rho = \rho_0$ iff $\mathbf{h}(\rho) = \mathbf{h}(\rho_0)$ by Theorem 4.2. □

5. Metric entropy

Suppose that G is an infinite quasi-transitive amenable connected graph. Choose one element o_i from each vertex orbit. It is shown in BLPS [6], Proposition 3.6, that there is a probability measure ρ on the set $\{o_i\}$ such that for any Følner sequence $\langle H_n \rangle$, the relative

frequency of vertices in H_n that are in the same orbit as o_i converges to $\rho(o_i)$. We call this measure ρ the *natural frequency distribution* of G.

Theorem 5.1. *Let G be an infinite quasi-transitive amenable connected graph with natural frequency distribution ρ. Let G_n be finite connected Følner subgraphs of G. Then*

$$\lim_{n\to\infty} \frac{1}{|\mathsf{V}(G_n)|} \log \tau(G_n) = \sum_{x\in\mathsf{V}(G)} \rho(x) \log \deg_G(x) - \sum_{k\geqslant 1} \frac{1}{k} \sum_{x\in\mathsf{V}(G)} \rho(x) p_k(x;G) = \mathbf{h}(G,\rho). \quad (5.1)$$

If $\Gamma \subseteq \mathrm{Aut}(G)$ is a countable group acting freely on $\mathsf{V}(G)$ with a finite number I of orbits, then

$$\mathbf{H}(\mathsf{WSF}_G, \Gamma) = I\mathbf{h}(G,\rho). \quad (5.2)$$

Furthermore, the Γ-entropy of any invariant measure on essential spanning forests of G is at most $I\mathbf{h}(G,\rho)$.

We shall need several lemmas to prove Theorem 5.1.

Lemma 5.2. *Let G be an infinite quasi-transitive unimodular graph. If G has 2 ends, then $\mathfrak{F}$ is a tree with exactly 2 ends* WSF*-a.s., while otherwise, for* WSF*-a.e. $\mathfrak{F}$, each component tree of $\mathfrak{F}$ has exactly one end.*

Proof. Pemantle [35] established this one-endedness for $G = \mathbb{Z}^d$ when $2 \leqslant d \leqslant 4$. BLPS [7] proved this in the general case that G is transitive (and unimodular). This latter proof is long, but not too hard to modify so as to apply to quasi-transitive graphs. The case where G has 2 ends or is recurrent needs only a few simple modifications that we do not detail. The major changes needed in the transient case with other than 2 ends are as follows. Let $\{o_i\}$ be a set of representatives of the orbits of V and let w_i be the reciprocal of the Haar measure of the stabilizer of o_i, where we normalize Haar measure so that $\sum_i w_i = 1$. (In the amenable case, we have $w_i = \rho(o_i)$ by BLPS [6], Proposition 3.6.) Consider any $\mathrm{Aut}(G)$-invariant probability measure $\mathbf{P}$ on $2^{\mathsf{E}(G)}$. For any subgraph ω of G and any vertex x, write $D(x)$ for the degree of x in ω. Let A_i be the event that the component of o_i in ω is infinite and $p_{\infty,i} := \mathbf{P}[A_i]$. Then according to Theorem 6.4 of BLPS [6], we have

$$\sum_i w_i \mathbf{E}[D(o_i)\,;\, A_i] \geqslant \sum_i 2 w_i p_{\infty,i}. \quad (5.3)$$

We may use this to prove an analogue of Theorem 7.2 in BLPS [6], namely, that if some component of ω has at least 3 ends with positive probability, then strict inequality holds in (5.3). The next step is to combine the proof of Theorem 6.5 of BLPS [7] with Corollary 3.5 of BLPS [6] to show that when $\mathbf{P} = \mathsf{WSF}_G$, we have $\sum_i w_i \mathbf{E}[D(o_i)] = 2$, so that equality holds in (5.3). Therefore, each tree has at most 2 ends WSF-a.s. The rest of the proof needs simple obvious modifications only, except for the crucial 'trunk' lemma, *i.e.*, Lemma 10.5 of BLPS [7]. Almost all of this proof can also be used word for word. The only significant change needed is that if o_i is on the trunk, then the shift along the trunk should bring to o_i the next vertex in the orbit of o_i in the direction of the orientation of the trunk. □

For our other lemmas, we shall find the following notation convenient. Given a finite subgraph H of a graph G and a configuration ω of $\mathsf{E}(G)$, let $H(\omega)$ denote the cylinder event consisting of those configurations of $\mathsf{E}(G)$ that agree with ω on $\mathsf{E}(H)$. Given also a configuration ω of $\mathsf{E}(G) \setminus \mathsf{E}(H)$, we define two finite graphs from certain vertex identifications on H: Write $H \circ \omega$ for the graph obtained by identifying all vertices of H that are connected to each other in the graph $\big(\mathsf{V}(G), \omega\big)$. Write $H * \omega$ for the graph obtained by identifying all vertices of H that are connected to each other in the graph $\big(\mathsf{V}(G), \omega\big)$ and by identifying all vertices of H that belong to any infinite connected component in $\big(\mathsf{V}(G), \omega\big)$. Note that in $H * \omega$, each finite component of $\big(\mathsf{V}(G), \omega\big)$ yields a separate identification, while the infinite components of $\big(\mathsf{V}(G), \omega\big)$ yield together one single identification.

The next lemma provides another justification for the adjective 'uniform' in 'wired uniform spanning forest', similar to a Gibbs specification. However, it does not hold for all graphs.

Lemma 5.3. *Let G be an infinite quasi-transitive unimodular graph and let H be a finite connected subgraph of G. If G has 2 ends, then*

$$\mathsf{WSF}\big(H(\mathfrak{F}) \mid \mathfrak{F}\restriction\big(\mathsf{E} \setminus \mathsf{E}(H)\big)\big) = \tau\big(H \circ \big(\mathfrak{F}\restriction\big(\mathsf{E} \setminus \mathsf{E}(H)\big)\big)\big)^{-1} \qquad \mathsf{WSF}\text{-}a.s.,$$

while otherwise,

$$\mathsf{WSF}\big(H(\mathfrak{F}) \mid \mathfrak{F}\restriction\big(\mathsf{E} \setminus \mathsf{E}(H)\big)\big) = \tau\big(H * \big(\mathfrak{F}\restriction\big(\mathsf{E} \setminus \mathsf{E}(H)\big)\big)\big)^{-1} \qquad \mathsf{WSF}\text{-}a.s.$$

Proof. The case where G has 2 ends is similar, though simpler, than the other case, so we give the details only for the case where G has other than 2 ends. Let Z be the event that each tree of $\mathfrak{F}$ has exactly one end. By Lemma 5.2, we have $\mathsf{WSF}(Z) = 1$.

Let B_R denote the ball of radius R about some fixed vertex of G. Choose R_H so that $H \subset B_{R_H}$. Let A_R be the event that for all $x, y \in \partial_{\mathsf{V}} H$ and $z, w \in \partial_{\mathsf{V}} B_R$, if in $\mathfrak{F}\restriction\big(\mathsf{E}(B_R) \setminus \mathsf{E}(H)\big)$

- x is connected to z,
- y is connected to w, and
- x is not connected to w,

then x and y are not connected in $\mathfrak{F}\restriction\mathsf{E}(H)$. Thus, $A_R \subseteq A_{R+1}$ for all $R \geqslant R_H$ and

$$Z \subseteq \bigcup_R A_R,$$

whence $\lim_{R\to\infty} \mathsf{WSF}(A_R) = 1$. Since

$$\mathsf{WSF}(A_R) = \int \mathsf{WSF}\big(A_R \mid \mathfrak{F}\restriction\big(\mathsf{E}(B_R) \setminus \mathsf{E}(H)\big)\big)\, d\mathsf{WSF},$$

it follows that for all large R we have

$$\mathsf{WSF}(C_R) \geqslant 1 - \sqrt{1 - \mathsf{WSF}(A_R)},$$

where

$$C_R := \Big\{\mathfrak{F}\,;\ \mathsf{WSF}\big(A_R \mid \mathfrak{F}\restriction\big(\mathsf{E}(B_R) \setminus \mathsf{E}(H)\big)\big) \geqslant 1 - \sqrt{1 - \mathsf{WSF}(A_R)}\Big\}.$$

In particular, $\mathsf{WSF}(\limsup_{R\to\infty} C_R) = 1$.

By definition,

$$\mathsf{WSF}\big(H(\mathfrak{F}) \mid \mathfrak{F}\restriction\big(\mathsf{E}\setminus\mathsf{E}(H)\big)\big) = \lim_{R\to\infty} \mathsf{WSF}\big(H(\mathfrak{F}) \mid \mathfrak{F}\restriction\big(\mathsf{E}(B_R)\setminus\mathsf{E}(H)\big)\big) \quad \mathsf{WSF}\text{-a.s.}$$

Fix a forest $\omega \in Z \cap \limsup_{R\to\infty} C_R$ for which the limit above holds. Choose $\epsilon > 0$ arbitrarily small. Choose $R \geqslant R_H$ so large that

$$|\mathsf{WSF}\big(H(\omega) \mid \omega\restriction\big(\mathsf{E}\setminus\mathsf{E}(H)\big)\big) - \mathsf{WSF}\big(H(\omega) \mid \omega\restriction\big(\mathsf{E}(B_R)\setminus\mathsf{E}(H)\big)\big)| < \epsilon, \tag{5.4}$$

that $\omega \in C_R$, that $\sqrt{1-\mathsf{WSF}(A_R)} < \epsilon$, and that each vertex in $\partial_{\mathsf{V}} H$ that is connected in $\omega\restriction\big(\mathsf{E}(B_R)\setminus\mathsf{E}(H)\big)$ to $\partial_{\mathsf{V}} B_R$ belongs to an infinite component in $\omega\restriction\big(\mathsf{E}\setminus\mathsf{E}(H)\big)$. This last requirement, in combination with $\omega \in Z$, implies that $\omega \in A_R$. Consider the cylinder set

$$D := \big(B_R\setminus\mathsf{E}(H)\big)(\omega) = \{\mathfrak{F}\,;\ \mathfrak{F}\restriction\big(\mathsf{E}(B_R)\setminus\mathsf{E}(H)\big) = \omega\restriction\big(\mathsf{E}(B_R)\setminus\mathsf{E}(H)\big)\}.$$

Let μ_N be the uniform spanning tree measure on B_N^*. By definition,

$$\mathsf{WSF}\big(H(\omega) \mid D\big) = \lim_{N\to\infty} \mu_N\big(H(\omega) \mid D\big)$$

and

$$\mathsf{WSF}(A_R \mid D) = \lim_{N\to\infty} \mu_N(A_R \mid D).$$

Since $\omega \in C_R$ and $\sqrt{1-\mathsf{WSF}(A_R)} < \epsilon$, we have $\mathsf{WSF}(A_R \mid D) > 1-\epsilon$. Thus, we may choose $N > R$ so large that

$$|\mathsf{WSF}\big(H(\omega) \mid D\big) - \mu_N\big(H(\omega) \mid D\big)| < \epsilon$$

and $\mu_N(A_R \mid D) > 1-\epsilon$. Since

$$\mu_N\big(H(\omega) \mid D\big) = \mu_N(A_R \mid D)\mu_N\big(H(\omega) \mid A_R\cap D\big) + \mu_N(A_R^c \mid D)\mu_N\big(H(\omega) \mid A_R^c\cap D\big),$$

we have

$$|\mu_N\big(H(\omega) \mid D\big) - \mu_N\big(H(\omega) \mid A_R\cap D\big)| < 2\epsilon.$$

Given $A_R \cap D$, the configurations inside H and outside B_R are μ_N-independent. Since $\omega \in A_R$, it follows that

$$\mu_N\big(H(\omega) \mid A_R\cap D\big) = \tau(H*\omega)^{-1},$$

and so

$$|\mathsf{WSF}\big(H(\omega) \mid D\big) - \tau(H*\omega)^{-1}| < 3\epsilon.$$

Therefore,

$$|\mathsf{WSF}\big(H(\omega) \mid \omega\restriction\big(\mathsf{E}\setminus\mathsf{E}(H)\big)\big) - \tau(H*\omega)^{-1}| < 4\epsilon$$

by (5.4). Since ϵ is arbitrary and ω is an arbitrary element of a set of measure 1, the result follows. □

Lemma 5.4. *Let H be a finite connected graph and W be a subset of vertices of H. Let H' be any graph obtained from H by making certain identifications of the vertices in W with*

each other. Write $\alpha := (|W| - 1)/|\mathsf{E}(H)|$. *Then*

$$|\log \tau(H) - \log \tau(H')| \leqslant |\mathsf{E}(H)|\big(-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)\big).$$

Proof. Let μ be the uniform spanning tree measure on H and μ' the uniform spanning tree measure on H'. It follows from Feder and Mihail [16] that μ stochastically dominates μ'. By Strassen's theorem [43], this means that there is a probability measure on pairs (T, T') so that the law of T is μ, the law of T' is μ', and $T \supseteq T'$ a.s. Now $|\mathsf{E}(T)| = |\mathsf{V}(H)| - 1$ and $|\mathsf{E}(T')| = |\mathsf{V}(H')| - 1$. We deduce that a.s.

$$|\mathsf{E}(T) \triangle \mathsf{E}(T')| = |\mathsf{V}(H)| - |\mathsf{V}(H')| \leqslant |W| - 1.$$

It now follows from Lemma 2.1 that

$$|\mathbf{H}(\mu) - \mathbf{H}(\mu')| \leqslant |\mathsf{E}(H)|\big(-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)\big).$$

Since $\mathbf{H}(\mu) = \log \tau(H)$ and $\mathbf{H}(\mu') = \log \tau(H')$, this is the same as the desired inequality. □

Lemma 5.5. *Let G be an infinite quasi-transitive unimodular graph and H be a finite connected subgraph of G. Write* $\alpha := (|\partial_{\mathsf{V}} H| - 1)/|\mathsf{E}(H)|$. *Then for* WSF*-a.e.* $\mathfrak{F}$,

$$|\log \mathsf{WSF}\big(H(\mathfrak{F})\big) - \log \tau(H)^{-1}| \leqslant |\mathsf{E}(H)|\big(-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)\big).$$

Proof. Write $\beta := \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}$. According to Lemma 5.4, we have $\tau(H')^{-1} \in I_H$, where

$$I_H := \big[\tau(H)^{-1}\beta^{-|\mathsf{E}(H)|},\ \tau(H)^{-1}\beta^{|\mathsf{E}(H)|}\big]$$

and $H' := H \circ \big(\mathfrak{F}\restriction(\mathsf{E} \setminus \mathsf{E}(H))\big)$ if G has 2 ends, while $H' := H * \big(\mathfrak{F}\restriction(\mathsf{E} \setminus \mathsf{E}(H))\big)$ otherwise. Combining the preceding relation with Lemma 5.3, we obtain

$$\mathsf{WSF}\big(H(\mathfrak{F})\big) = \mathbf{E}\big[\mathsf{WSF}\big(H(\mathfrak{F}) \mid \mathfrak{F}\restriction(\mathsf{E} \setminus \mathsf{E}(H))\big)\big] \in I_H$$

a.s. This is the same as the desired inequality. □

Lemma 5.6. *Let H be a finite connected graph and W be a subset of vertices of H. Let $\tau_{\mathrm{e}}(H, W)$ be the number of spanning forests of H such that each tree contains at least one vertex of W. Write* $\alpha := (|W| - 1)/|\mathsf{E}(H)|$. *Then*

$$\log \tau_{\mathrm{e}}(H, W) \leqslant \log \tau(H) + |\mathsf{E}(H)|\big(-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)\big).$$

Proof. Let μ be the uniform measure on spanning forests of H such that each tree contains at least one vertex of W. Let ν be obtained from μ by choosing a spanning forest $\mathfrak{F}$ with distribution μ and then randomly adding to $\mathfrak{F}$ enough edges of H so that a spanning tree of H results. By Lemma 2.1, we have

$$|\mathbf{H}(\mu) - \mathbf{H}(\nu)| \leqslant |\mathsf{E}(H)|\big(-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)\big)$$

since at most $|W| - 1$ edges are added. Since $\mathbf{H}(\mu) = \log \tau_{\mathrm{e}}(H, W)$ and $\mathbf{H}(\nu) \leqslant \log \tau(H)$, the desired inequality follows. □

Proof of Theorem 5.1. We first prove that (5.1) holds. By definition of ρ, the graphs G_n have a random weak limit (G, ρ). Thus, (5.1) is a consequence of Theorem 3.2.

We next show (5.2). Choose a ball $B_R(o)$ of vertices and edges such that $\Gamma B_R(o) = G$. Let

$$\Gamma_n := \{\gamma \in \Gamma \,;\, \gamma B_R(o) \cap G_n \neq \emptyset\}$$

and put

$$G'_n := \Gamma_n B_R(o).$$

Since $\langle G_n \rangle$ is a Følner sequence in G, it follows that $\langle \Gamma_n \rangle$ is a Følner sequence in Γ. Therefore,

$$\mathbf{H}(\mathsf{WSF}, \Gamma) = -\lim_{n\to\infty} |\Gamma_n|^{-1} \log \mathsf{WSF}\big(G'_n(\mathfrak{F})\big)$$

in $L^1(\mathsf{WSF})$ by the generalized Shannon–McMillan Theorem of Kieffer [26]. Now

$$\lim_{n\to\infty} |\mathsf{V}(G'_n)|/|\Gamma_n| = I$$

since Γ acts freely on V. Hence

$$\mathbf{H}(\mathsf{WSF}, \Gamma) = -\lim_{n\to\infty} I\,|\mathsf{V}(G'_n)|^{-1} \log \mathsf{WSF}\big(G'_n(\mathfrak{F})\big) \tag{5.5}$$

in $L^1(\mathsf{WSF})$. As we recalled in Section 2, every quasi-transitive amenable graph is unimodular. The result now follows from Lemma 5.5 in conjunction with (5.5).

We finally show that no measure μ on essential spanning forests of G has larger entropy. Using some of the same reasoning as above, we have that

$$\mathbf{H}(\mu, \Gamma) = \lim_{n\to\infty} |\Gamma_n|^{-1} \mathbf{H}(\mu \restriction G'_n).$$

Because μ is concentrated on essential spanning forests, the number of elements of the partition generated by G'_n that have positive μ-measure is at most $\tau_{\mathrm{e}}(G'_n, \partial_{\mathsf{V}} G'_n)$, whence $\mathbf{H}(\mu \restriction G'_n) \leqslant \log \tau_{\mathrm{e}}(G'_n, \partial_{\mathsf{V}} G'_n)$. We may apply Lemma 5.6 to obtain the desired conclusion. □

Acknowledgements

I am grateful to Wolfgang Woess, Thierry Coulhon, Ben Morris, Yuval Peres, and Scott Sheffield for useful discussions and references. Thanks are due to Benny Sudakov for asking about the asymptotics for graphs whose degree tends to infinity.

References

[1] Aldous, D. (1998) Tree-valued Markov chains and Poisson–Galton–Watson distributions. In *Microsurveys in Discrete Probability* (D. Aldous and J. Propp, eds), Vol. 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, AMS, Providence, RI, pp. 1–20.

[2] Aldous, D. J. and Lyons, R. (2004) Processes on unimodular random networks. In preparation.

[3] Aldous, D. J. and Steele, J. M. (2004) The objective method: Probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures* (H. Kesten, ed.), Vol. 110 of *Encyclopaedia Math. Sci.* Springer, pp. 1–72.

[4] Barlow, M., Coulhon, T. and Grigor'yan, A. (2001) Manifolds and graphs with slow heat kernel decay. *Invent. Math.* **144** 609–649.

[5] Benjamini, I. and Schramm, O. (2001) Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6** # 23 (electronic).

[6] Benjamini, I., Lyons, R., Peres, Y. and Schramm, O. (1999) Group-invariant percolation on graphs. *Geom. Funct. Anal.* **9** 29–66.

[7] Benjamini, I., Lyons, R., Peres, Y. and Schramm, O. (2001) Uniform spanning forests. *Ann. Probab.* **29** 1–65.

[8] Bhatia, R. (1997) *Matrix Analysis*, Vol. 169 of *Graduate Texts in Mathematics*, Springer, New York.

[9] Bollobás, B. (1998) *Modern Graph Theory*, Vol. 184 of *Graduate Texts in Mathematics*, Springer, New York.

[10] Bollobás, B. (2001) *Random Graphs*, Vol. 73 of *Cambridge Studies in Advanced Mathematics*, second edition, Cambridge University Press, Cambridge.

[11] Burton, R. M. and Pemantle, R. (1993) Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *Ann. Probab.* **21** 1329–1371.

[12] Carlen, E. A., Kusuoka, S. and Stroock, D. W. (1987) Upper bounds for symmetric Markov transition functions. *Ann. Inst. H. Poincaré Probab. Statist.* **23** 245–287.

[13] Chung, F. R. K. (1996) Laplacians of graphs and Cheeger's inequalities. In *Combinatorics, Paul Erdős is Eighty*, Vol. 2, (D. Miklós, V. T. Sós and T. Szőnyi, eds), Vol. 2 of *Bolyai Soc. Math. Stud.*, János Bolyai Math. Soc., Budapest, pp. 157–172.

[14] Coulhon, T. (2000) Random walks and geometry on infinite graphs. In *Lecture Notes on Analysis on Metric Spaces, Trento, CIRM, 1999* (L. Ambrosio and F. S. Cassano, eds), Scuola Normale Superiore di Pisa, pp. 5–30.

[15] Dudley, R. M. (1989) *Real Analysis and Probability*, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA.

[16] Feder, T. and Mihail, M. (1992) Balanced matroids. In *Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing*, ACM, New York, pp. 26–38.

[17] Felker, J. L. and Lyons, R. (2003) High-precision entropy values for spanning trees in lattices. *J. Phys. A* **36** 8361–8365.

[18] Fuglede, B. and Kadison, R. V. (1952) Determinant theory in finite factors. *Ann. of Math.* **55** 520–530.

[19] Godsil, C. and Royle, G. (2001) *Algebraic Graph Theory*, Vol. 207 of *Graduate Texts in Mathematics*, Springer, New York.

[20] Grenander, U. and Szegő, G. (1984) *Toeplitz Forms and Their Applications*, second edition, Chelsea Publishing Co., New York.

[21] Grimmett, G. R. (1980/81) Random labelled trees and their branching networks. *J. Austral. Math. Soc. Ser. A* **30** 229–237.

[22] Häggström, O. (1995) Random-cluster measures and uniform spanning trees. *Stochastic Process. Appl.* **59** 267–275.

[23] Häggström, O. (1998) Uniform and minimal essential spanning forests on trees. *Random Struct. Alg.* **12** 27–50.

[24] Helson, H. and Lowdenslager, D. (1958) Prediction theory and Fourier series in several variables. *Acta Math.* **99** 165–202.

[25] Kesten, H. (1959) Symmetric random walks on groups. *Trans. Amer. Math. Soc.* **92** 336–354.

[26] Kieffer, J. C. (1975) A generalized Shannon–McMillan theorem for the action of an amenable group on a probability space. *Ann. Probab.* **3**, 1031–1037.

[27] Lind, D. and Tuncel, S. (2001) A spanning tree invariant for Markov shifts. In *Codes, Systems, and Graphical Models* (*Minneapolis, MN, 1999*) (B. Marcus and J. Rosenthal, eds), Vol. 123 of *IMA Vol. Math. Appl.*, Springer, New York, pp. 487–497.

[28] Löwner, K. (1934) Über monotone Matrixfunktionen. *Math. Z.* **38** 177–216.

[29] Lyons, R. (1998) A bird's-eye view of uniform spanning trees and forests. In *Microsurveys in Discrete Probability* (D. Aldous and J. Propp, eds), Vol. 41 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, AMS, Providence, RI, pp. 135–162.

[30] Lyons, R. and Steif, J. E. (2003) Stationary determinantal processes: Phase multiplicity, Bernoullicity, entropy, and domination. *Duke Math. J.* **120** 515–575.

[31] McKay, B. D. (1981) Spanning trees in random regular graphs. In *Proceedings of the Third Caribbean Conference on Combinatorics and Computing* (C. C. Cadogan, ed.), pp. 139–143.

[32] McKay, B. D. (1983) Spanning trees in regular graphs. *Europ. J. Combin.* **4** 149–160.

[33] Morris, B. and Peres, Y. (2004) Evolving sets, mixing and heat kernel bounds. *Probab. Theory Rel. Fields*, to appear.

[34] Ornstein, D. S. and Weiss, B. (1987) Entropy and isomorphism theorems for actions of amenable groups. *J. Analyse Math.* **48** 1–141.

[35] Pemantle, R. (1991) Choosing a spanning tree for the integer lattice uniformly. *Ann. Probab.* **19** 1559–1574.

[36] Runge, F. and Sachs, H. (1974) Berechnung der Anzahl der Gerüste von Graphen und Hypergraphen mittels deren Spektren. In *Papers presented at the Fifth Balkan Mathematical Congress*, Vol. 4, Beogradski Izdavačko-Grafički Zavod, Belgrade, pp. 529–536. *Math. Balkanica* **4** (1974).

[37] Saloff-Coste, L. (1995) Isoperimetric inequalities and decay of iterated kernels for almost-transitive Markov chains. *Combin. Probab. Comput.* **4** 419–442.

[38] Saloff-Coste, L. (1997) Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics* (*Saint-Flour, 1996*) (P. Bernard, ed.), Vol. 1665 of *Lecture Notes in Math.*, Springer, Berlin, pp. 301–413.

[39] Salvatori, M. (1992) On the norms of group-invariant transition operators on graphs. *J. Theoret. Probab.* **5** 563–576.

[40] Sheffield, S. (2003) Random surfaces: Large deviations principles and gradient Gibbs measure classifications. PhD thesis, Stanford University.

[41] Shrock, R. and Wu, F. Y. (2000). Spanning trees on graphs and lattices in d dimensions. *J. Phys. A* **33** 3881–3902.

[42] Soardi, P. M. and Woess, W. (1990) Amenability, unimodularity, and the spectral radius of random walks on infinite graphs. *Math. Z.* **205** 471–486.

[43] Strassen, V. (1965) The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–439.

[44] Woess, W. (1984) A random walk on free products of finite groups. In *Probability Measures on Groups VII* (*Oberwolfach, 1983*) (H. Heyer, ed.), Vol. 1064 of *Lecture Notes in Mathematics*, Springer, Berlin, pp. 467–470.

[45] Woess, W. (2000) *Random Walks on Infinite Graphs and Groups*, Vol. 138 of *Cambridge Tracts in Mathematics*, Cambridge University Press, Cambridge.