

# Tree formulas, mean first passage times and Kemeny’s constant of a Markov chain

JIM PITMAN\* and WENPIN TANG\*\*

Statistics Department, University of California, Berkeley, CA 94720-3860, USA.

E-mail: \*pitman@stat.berkeley.edu; \*\*wenspintang@stat.berkeley.edu

This paper offers some probabilistic and combinatorial insights into tree formulas for the Green function and hitting probabilities of Markov chains on a finite state space. These tree formulas are closely related to *loop-erased random walks* by *Wilson’s algorithm* for random spanning trees, and to *mixing times* by the *Markov chain tree theorem*. Let  $m_{ij}$  be the mean first passage time from  $i$  to  $j$  for an irreducible chain with finite state space  $S$  and transition matrix  $(p_{ij}; i, j \in S)$ . It is well known that  $m_{jj} = 1/\pi_j = \Sigma^{(1)}/\Sigma_j$ , where  $\pi$  is the stationary distribution for the chain,  $\Sigma_j$  is the tree sum, over  $n^{n-2}$  trees  $\mathbf{t}$  spanning  $S$  with root  $j$  and edges  $i \rightarrow k$  directed towards  $j$ , of the tree product  $\prod_{i \rightarrow k \in \mathbf{t}} p_{ik}$ , and  $\Sigma^{(1)} := \sum_{j \in S} \Sigma_j$ . Chebotarev and Agaev (*Linear Algebra Appl.* **356** (2002) 253–274) derived further results from *Kirchhoff’s matrix tree theorem*. We deduce that for  $i \neq j$ ,  $m_{ij} = \Sigma_{ij}/\Sigma_j$ , where  $\Sigma_{ij}$  is the sum over the same set of  $n^{n-2}$  spanning trees of the same tree product as for  $\Sigma_j$ , except that in each product the factor  $p_{kj}$  is omitted where  $k = k(i, j, \mathbf{t})$  is the last state before  $j$  in the path from  $i$  to  $j$  in  $\mathbf{t}$ . It follows that Kemeny’s constant  $\sum_{j \in S} m_{ij}/m_{jj}$  equals  $\Sigma^{(2)}/\Sigma^{(1)}$ , where  $\Sigma^{(r)}$  is the sum, over all forests  $\mathbf{f}$  labeled by  $S$  with  $r$  directed trees, of the product of  $p_{ij}$  over edges  $i \rightarrow j$  of  $\mathbf{f}$ . We show that these results can be derived without appeal to the matrix tree theorem. A list of relevant literature is also reviewed.

**Keywords:** Cayley’s formula; Green tree formula; harmonic tree formula; Kemeny’s constant; Kirchhoff’s matrix tree theorem; Markov chain tree theorem; mean first passage times; spanning forests/trees; Wilson’s algorithm

## 1. Introduction and background

In this survey paper, we review various tree formulas of a finite Markov chain, and make connections with random spanning trees and mean first passage times in the Markov chain. Most results are known from previous work, but a few formulas and statements, e.g. the combinatorial interpretation of *Kemeny’s constant* in Corollary 1.4, and the formula (5.3), appear here for the first time. We offer a probabilistic and combinatorial approach to these results, encompassing the closely related results of Leighton and Rivest [64,65] as well as *Wilson’s algorithm* [91,107] for generation of random spanning trees.

Throughout this paper, we assume that  $S$  is a finite state space. Let  $m_{ij}$  be the *mean first passage time* from  $i$  to  $j$  for an irreducible Markov chain  $(X_n)_{n \in \mathbb{N}}$  with state space  $S$  and transition matrix  $\mathbf{P} := (p_{ij}; i, j \in S)$ . That is,

$$m_{ij} := \mathbb{E}_i T_j^+ \quad \text{for } i, j \in S,$$

where  $T_j^+ := \inf\{n \geq 1; X_n = j\}$  is the hitting time of the state  $j \in S$ , and  $\mathbb{E}_i$  is the expectation relative to the Markov chain  $(X_n)_{n \in \mathbb{N}}$  starting at  $i \in S$ .

It is well known that the irreducible chain  $(X_n)_{n \in \mathbb{N}}$  has a unique stationary distribution  $(\pi_j)_{j \in S}$  which is given by

$$\pi_j = 1/m_{jj} \quad \text{for all } j \in S.$$

See, for example, Levin, Peres and Wilmer [67], Chapter 1, or Durrett [34], Chapter 6, for background on the theory of Markov chains.

For a directed graph  $\mathbf{g}$  with vertex set  $S$ , write  $i \rightarrow j \in \mathbf{g}$  to indicate that  $(i, j)$  is a directed edge of  $\mathbf{g}$  and call

$$\Pi^{\mathbf{P}}(\mathbf{g}) := \prod_{i \rightarrow j \in \mathbf{g}} p_{ij}$$

the  $\mathbf{P}$ -weight of  $\mathbf{g}$ . Each forest  $\mathbf{f}$  with vertex set  $S$  and edges directed towards root vertices consists of some number  $r$  of trees  $\mathbf{t}_i$  whose vertex sets partition  $S$  into  $r$  non-empty disjoint subsets. Observe that if a forest  $\mathbf{f}$  consists of trees  $\mathbf{t}_1, \dots, \mathbf{t}_r$ , then  $\mathbf{f}$  has  $\mathbf{P}$ -weight

$$\Pi^{\mathbf{P}}(\mathbf{t}_1, \dots, \mathbf{t}_r) := \Pi^{\mathbf{P}}(\mathbf{f}) = \prod_{i=1}^r \Pi^{\mathbf{P}}(\mathbf{t}_i).$$

Write  $\mathbf{t} \rightarrow j$  to indicate that the edges of a tree  $\mathbf{t}$  are all directed towards a root element  $j \in \mathbf{t}$ . The formula

$$m_{jj} = \frac{1}{\pi_j} = \Sigma^{(1)}/\Sigma_j, \tag{1.1}$$

where

$$\Sigma_j := \sum_{\mathbf{t} \rightarrow j} \Pi^{\mathbf{P}}(\mathbf{t}) \quad \text{and} \quad \Sigma^{(1)} := \sum_{j \in S} \Sigma_j \tag{1.2}$$

follows readily from the *Markov chain tree theorem* [60,64,65,95,106]:

**Theorem 1.1 (Markov chain tree theorem for irreducible chains [60,95,106]).** *Assume that  $\mathbf{P}$  is irreducible or equivalently, that  $\Sigma_j > 0$  for every  $j \in S$ . Then*

$$\sum_{i \in S} \Sigma_i p_{ij} = \Sigma_j \quad \text{for all } j \in S. \tag{1.3}$$

Consequently, the unique stationary distribution for the chain is  $\pi_j = \Sigma_j / \Sigma^{(1)}$ .

Section 3 recalls the short combinatorial proof of this result due to Ventcel and Freidlin [106], where Theorem 1.1 appeared as an auxiliary lemma to study random dynamical systems. It was also formulated by Shubert [95] and Solberg [97] in the language of graph theory, and by Kolher and Vollmerhaus [60] in the context of biological multi-state systems. The name *Markov chain*

*tree theorem* was first coined by Leighton and Rivest [64,65], where they extended the result to general Markov chains which are not necessarily irreducible, see Theorem 3.1.

Later Anantharam and Tsoucas [4], Aldous [3] and Broder [17] provided probabilistic arguments by lifting the Markov chain to its spanning tree counterpart. A method to generate random spanning trees, the *Aldous–Broder algorithm*, was devised as a by-product of their proofs: see Lyons and Peres [74], Section 4.4. See also Kelner and Mądry [52], and Mądry, Straszak and Tarnawski [75] for development on fast algorithms to generate random spanning trees. Recently, Biane [10], and Biane and Chapuy [11] studied the factorization of a polynomial associated to that spanning tree-valued Markov chain. Gursoy, Kirkland, Mason and Sergeev [40] extended the Markov chain tree theorem in the max algebra setting.

As we discuss in Section 4.2, the Markov chain tree theorem is a probabilistic expression of *Kirchhoff's matrix tree theorem* [22,58,104]. See also Seneta [94], Lemma 7.1, for a weaker form of this theorem and its application to compute stationary distributions of countable state Markov chains from finite truncations. Here is a version of Kirchhoff's matrix tree theorem, essentially due to Chaiken [22] and Chen [27]. We follow the presentation of Pokarowski [90], Lemma 1.1 and 1.2.

**Theorem 1.2 (Kirchhoff's matrix forest theorem for directed graphs [22,27,90]).** *Let  $R$  be a subset of the finite state space  $S$  of a Markov chain  $(X_n)_{n \in \mathbb{N}}$  with transition matrix  $\mathbf{P}$ . Let  $\mathbf{L} := \mathbf{I} - \mathbf{P}$  where  $\mathbf{I}$  is the identity matrix on  $S$ , and let  $\mathbf{L}(R)$  be the matrix indexed by  $S \setminus R$  obtained by removing from  $\mathbf{L}$  all the rows and columns indexed by  $R$ . Then*

$$\det \mathbf{L}(R) = w(R) := \sum_{\text{ROOTS}(\mathbf{f})=R} \Pi^{\mathbf{P}}(\mathbf{f}), \quad (1.4)$$

where the sum is over all forests  $\mathbf{f}$  labeled by  $S$  whose set of roots is  $R$ . Moreover,

$$\text{if } \det \mathbf{L}(R) > 0, \quad \text{then } \mathbf{L}(R)^{-1} = \left( \frac{w_{ij}(R \cup \{j\})}{w(R)} \right)_{i,j \in S \setminus R}, \quad (1.5)$$

where

$$w_{ij}(R \cup \{j\}) := \sum_{\text{ROOTS}(\mathbf{f})=R \cup \{j\}, i \rightsquigarrow j} \Pi^{\mathbf{P}}(\mathbf{f}) \quad (1.6)$$

is the  $\mathbf{P}$ -weight of all forests  $\mathbf{f}$  with roots  $R \cup \{j\}$  in which the tree component containing  $i$  has root  $j$ .

In the above theorem, the set of roots  $R$  may include single points with no incident edges. Theodore Zhu pointed that the r.h.s. of (1.4) is the probability that the functional digraph induced by a  $\mathbf{P}$ -mapping (see Pitman [87]) is a forest with root set  $R$ . This implies that  $0 \leq \det \mathbf{L}(R) \leq 1$ .

As observed by Pokarowski [90], Theorem 1.2, the expressions of Theorem 1.2 have the following probabilistic interpretations. Assume that  $w(R) > 0$ . First of all, the matrix in (1.5) is the Green function of the Markov chain with transition matrix  $\mathbf{P}$  killed when it hits  $R$ . From this, we derive the *tree formula for the Green function of a Markov chain* or simply

**Green tree formula**

$$\mathbb{E}_i \sum_{n=0}^{T_R-1} 1(X_n = j) = \mathbf{L}(R)_{ij}^{-1} = \frac{w_{ij}(R \cup \{j\})}{w(R)} \quad \text{for } i, j \in S \setminus R, \quad (1.7)$$

where  $T_R := \inf\{n \geq 0; X_n \in R\}$  is the entry time to the set  $R$ . Summing over  $j \in S \setminus R$  gives an expression for the mean first passage time

$$\mathbb{E}_i T_R = \frac{\sum_{j \in S \setminus R} w_{ij}(R \cup \{j\})}{w(R)}. \quad (1.8)$$

The  $\mathbb{P}_i$  distribution of  $X_{T_R}$  is given by a variant of (1.7): the *tree formula for harmonic functions of a Markov chain* or simply

**Harmonic tree formula**

$$\mathbb{P}_i(X_{T_R} = j) = \frac{w_{ij}(R)}{w(R)} \quad \text{for } i \in S \text{ and } j \in R, \quad (1.9)$$

where  $w_{ij}(R) = w_{ij}(R \cup \{j\})$  is exactly as in (1.7) but now  $j \in R$  so  $R \cup \{j\} = R$ .

It is well known that the formulas (1.7)–(1.9) all follow from characterizations of the probabilistic quantities as the unique solutions of linear equations associated with the *Laplacian matrix*  $\mathbf{L}$ . For example, let

$$\mathbf{P}^{\text{HIT}}(R) := (\mathbb{P}_i(X_{T_R} = j); i \in S \setminus R \text{ and } j \in R).$$

The usual first step analysis implies that

$$\mathbf{P}^{\text{HIT}}(R) = \mathbf{P}_{(S \setminus R) \times R} + \mathbf{P}_{(S \setminus R) \times (S \setminus R)} \mathbf{P}^{\text{HIT}}(R), \quad (1.10)$$

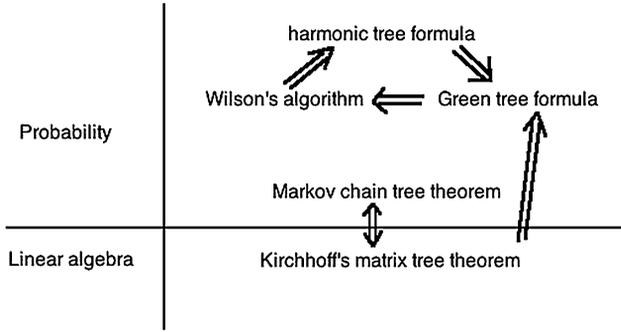
where  $\mathbf{P}_{R \times R'}$  is the restriction of  $\mathbf{P}$  to  $R \times R'$ . By letting  $\mathbf{L}(R; R')$  be obtained by removing from  $\mathbf{L}$  all the rows indexed by  $R$  and all the columns indexed by  $R'$ , the equation (1.10) is written as

$$\mathbf{L}(R) \mathbf{P}^{\text{HIT}}(R) = -\mathbf{L}(R; S \setminus R).$$

Then the harmonic tree formula (1.9) is easily deduced from the fundamental expressions (1.4) and (1.5) in Theorem 1.2.

The purpose of this work is to provide combinatorial and probabilistic meanings of these tree formulas, without appeal to linear algebra. The formulas (1.7)–(1.9) can be proved by purely combinatorial arguments. As an example, a combinatorial proof of the harmonic tree formula (1.9) is given in Section 2. In addition, the Green tree formula (1.7) and the harmonic tree formula (1.9) are closely related to Wilson's algorithm, whose original proof [107] is combinatorial. In fact:

- the Green tree formula (1.7) is derived from the harmonic tree formula (1.9), together with standard theory of Markov chains;
- the harmonic tree formula (1.9) is a consequence of the success of Wilson's algorithm;



**Figure 1.** Relations between various tree theorems/formulas.

- Wilson's algorithm follows from the Green tree formula (1.7) by a probabilistic argument due to Lawler [62].

These arguments are presented in Sections 2 and 5. We show in Section 4.2 that even the formula (1.4) can be deduced from the Markov chain tree theorem. See Figure 1 for the roadmap of relations between various tree formulas.

Theorem 1.1 provides a tree formula for the mean first passage time  $m_{ij}$  for  $i = j$ . A companion result for  $i \neq j$ , which is a reformulation of Chebotarev [25], Theorem 1, is stated as follows. The proof is deferred to Section 3.

**Theorem 1.3 (Markov chain tree formula for mean first passage times).** *Let  $\mathbf{P}$  be a transition matrix for an irreducible chain. For each  $i \neq j$ ,*

$$m_{ij} = \Sigma_{ij} / \Sigma_j, \tag{1.11}$$

where

$$\Sigma_{ij} := \sum_{\mathbf{t} \rightarrow j} \Pi^{\mathbf{P}}(\mathbf{t}) / p_{k(i,j,\mathbf{t})j}, \tag{1.12}$$

with  $k(i, j, \mathbf{t})$  being the last state before  $j$  in the path from  $i$  to  $j$  in  $\mathbf{t}$ , and  $\Sigma_j$  is defined by (1.2).

Observe that each term  $\Pi^{\mathbf{P}}(\mathbf{t}) / p_{k(i,j,\mathbf{t})j}$  on the r.h.s. of (1.12) can be written as

$$\Pi^{\mathbf{P}}(\mathbf{t}) / p_{k(i,j,\mathbf{t})j} = \Pi^{\mathbf{P}}(\mathbf{s}, \mathbf{u}),$$

where  $(\mathbf{s}, \mathbf{u})$  is the forest of two trees obtained by deleting the edge  $k(i, j, \mathbf{t}) \rightarrow j$  from  $\mathbf{t}$ . So the pair  $(\mathbf{s}, \mathbf{u})$  is a *two-component spanning forest*. It can easily be shown that the map  $\mathbf{t} \mapsto (\mathbf{s}, \mathbf{u})$  is a bijection between trees  $\mathbf{t}$  with  $\mathbf{t} \rightarrow j$  and two tree forests  $(\mathbf{s}, \mathbf{u})$  such that  $i \in \mathbf{s}$  and  $\mathbf{u} \rightarrow j$ . Thus the formula (1.12) for  $\Sigma_{ij}$  can be rewritten as

$$\Sigma_{ij} = \sum_{i \in \mathbf{s}, \mathbf{u} \rightarrow j} \Pi^{\mathbf{P}}(\mathbf{s}, \mathbf{u}). \tag{1.13}$$

Unaware of [25], Hunter [43] proposed an algorithm to compute mean first passage times in a Markov chain, and derived the instances of (1.11) for a Markov chain with two, three and four states. Note that for each  $i \in S$ , the number of terms in the sum  $\Sigma_{ij}$  is the same as the number in the sum  $\Sigma_j$ , namely  $|S|^{|S|-2}$ . Also, each term  $\Pi^{\mathbf{P}}(\mathbf{t})/p_{k(i,j,\mathbf{t})j}$  is larger than the corresponding term  $\Pi^{\mathbf{P}}(\mathbf{t})$  in  $\Sigma_j$ . To illustrate, for  $|S| = 2$  states  $\{0, 1\}$

$$m_{10} = 1/p_{10}.$$

For  $|S| = 3$  states  $\{0, 1, 2\}$

$$m_{10} = (p_{12} + p_{21} + p_{20})/(p_{12}p_{20} + p_{21}p_{10} + p_{10}p_{20}).$$

It was first observed by Kemeny and Snell [53], Corollary 4.3.6, that the quantity

$$K := \sum_{j \in S} m_{ij}/m_{jj} \quad (1.14)$$

is a constant, not depending on  $i$ . This constant associated with an irreducible Markov chain is known as *Kemeny's constant*. Since its discovery, a number of interpretations have been provided. For example, Levene and Loizou [66] interpreted Kemeny's constant as the expected distance between two typical vertices in a weighted directed graph. Lovasz and Winkler [69] rediscovered this result in their *random target lemma*, which was further developed in Aldous and Fill [2], Chapter 2.

Kemeny's constant  $K$  is closely related to the Laplacian matrix  $\mathbf{L}$ , and the *fundamental matrix*  $\mathbf{Z} := (\mathbf{L} + \mathbf{\Pi})^{-1}$  where  $\mathbf{\Pi}$  is the matrix each row of which is the stationary distribution  $\pi$ , by the following identities:

$$\begin{aligned} K &= \text{Tr} \mathbf{Z}; \\ K &= \text{Tr} [\mathbf{L}(i)^{-1}] + \frac{\mathbf{L}^{\#}_{ii}}{\pi_i}, \end{aligned}$$

where  $\mathbf{L}^{\#}$  is the *group inverse* of the Laplacian matrix  $\mathbf{L}$ . See also Doyle [33], Hunter [42], Gustafson and Hunter [41], and Catral, Kirkland, Neumann and Sze [20] for linear algebra approaches to Kemeny's constant  $K$ . The following result is a consequence of the formula (1.13) in the proof of Theorem 1.3.

**Corollary 1.4 (Combinatorial interpretation of Kemeny's constant).** *For  $r \in \mathbb{N}$ , let*

$$\Sigma^{(r)} := \sum_{\mathbf{t}_1, \dots, \mathbf{t}_r} \Pi^{\mathbf{P}}(\mathbf{t}_1, \dots, \mathbf{t}_r),$$

*where the sum is over all directed forests of  $r$  trees  $\mathbf{t}_1, \dots, \mathbf{t}_r$  spanning  $S$ . Then*

$$K = 1 + \Sigma^{(2)}/\Sigma^{(1)}. \quad (1.15)$$

Hunter [42] indicated the instances of (1.15) for a Markov chain with two and three states, but with a notation which conceals the generalization to  $n$  states. So this combinatorial interpretation of  $K$  may be new. We leave open the interpretation of  $\Sigma^{(r)}$  for  $r \geq 3$ .

### Organization of the paper:

- In Section 2, we provide a combinatorial proof of the harmonic tree formula (1.9), from which we derive the Green tree formula (1.7). We also prove Cayley's formula for enumerating spanning forests by means of the Green tree formula (1.7).
- In Section 3, we focus on the Markov chain tree theorems. We present a short proof of Theorem 1.1 and a generalization. We also provide two proofs for Theorem 1.3, one based on the formula (1.8) and the other relying on Theorem 1.1.
- In Section 4, we review Kirchhoff's matrix tree theorems. We show how to translate this graph theoretical result into the Markov chain setting. In particular, we show that the Markov chain tree theorem is derived from a version of Kirchhoff's matrix tree theorem.
- In Section 5, we explore the relation between Wilson's algorithm and various tree formulas. We also present Kassel and Kenyon's generalized Wilson's algorithm, from which we derive some cycle-rooted tree formulas.
- In Section 6, some additional notes and further references are provided.

## 2. Tree formulas and Cayley's formula

In this section, we provide combinatorial and probabilistic proofs for the Green tree formula (1.7) and the harmonic tree formula (1.9). As an application, we give a proof of Cayley's well-known formula [21] for enumerating spanning forests in a complete graph.

To begin with, we make a basic connection between results formulated for an irreducible Markov chain with state space  $S$ , and results formulated for killing of a possibly reducible Markov chain when it first hits an arbitrary subset  $R$  of its state space.

**Lemma 2.1.** *Let  $\mathbf{P}$  be a possibly reducible transition matrix indexed by a finite set  $S$ . For  $R$  a non-empty subset of  $S$ , let*

$$w(R) := \sum_{\text{ROOTS}(\mathbf{f})=R} \Pi^{\mathbf{P}}(\mathbf{f})$$

as in (1.4). The following conditions are equivalent:

- (1)  $w(R) > 0$ .
- (2) There exists at least one forest  $\mathbf{f}$  of trees spanning  $S$  with  $\text{ROOTS}(\mathbf{f}) = R$  such that the tree product  $\Pi^{\mathbf{P}}(\mathbf{f}) > 0$ , which is to say, every edge  $i \rightarrow j$  of  $\mathbf{f}$  has  $p_{ij} > 0$ .
- (3) For every  $i \in S \setminus R$  there exists a path  $\mathbf{t}$  from  $i$  to some  $r \in R$  such that  $\Pi^{\mathbf{P}}(\mathbf{t}) > 0$ .
- (4)  $\mathbf{L}(R) := (\mathbf{I} - \mathbf{P})_{(S \setminus R) \times (S \setminus R)}$  is invertible with inverse  $\mathbf{L}(R)^{-1} = \sum_{n=0}^{\infty} \mathbf{P}_{(S \setminus R) \times (S \setminus R)}^n$ , where  $\mathbf{P}_{(S \setminus R) \times (S \setminus R)}$  is the restriction of  $\mathbf{P}$  to  $(S \setminus R) \times (S \setminus R)$ .
- (5)  $\det \mathbf{L}(R) \neq 0$ .

**Proof.** Note that (1)  $\Leftrightarrow$  (2)  $\Rightarrow$  (3) is obvious. (3)  $\Rightarrow$  (2) is obtained by recursively selecting a path until it either joins an existing path leading to some  $r \in R$ , or reaches a different  $r' \in R$ . The procedure terminates when all of  $S \setminus R$  are exhausted. This point is reinforced by Wilson's algorithm, see Section 5. As for (3)  $\Leftrightarrow$  (4), this is textbook theory of absorbing Markov chains, see Kemeny and Snell [53], Theorem 3.2.1, or Seneta [94], Theorem 4.3. Finally, (4)  $\Leftrightarrow$  (5) is elementary linear algebra.  $\square$

By Kirchhoff's matrix forest theorem, Theorem 1.2, we know that  $w(R) = \det \mathbf{L}(R)$ , which is much more informative than the implication (1)  $\Leftrightarrow$  (5) of Lemma 2.1. But we are now trying to work around the matrix tree theorem, to increase our combinatorial and probabilistic understanding of its equivalence. Now we present a combinatorial proof of the harmonic tree formula.

**Proof of the harmonic tree formula (1.9).** Assume that  $w(R) > 0$ . By Lemma 2.1, for every  $i \in S$ , there exists a path  $\mathbf{t}$  from  $i$  to some  $r \in R$  such that  $\Pi^{\mathbf{P}}(\mathbf{t}) > 0$ . By finiteness of  $S$  and geometric bounds, we have  $\mathbb{P}_i(T_R < \infty) = 1$  for all  $i \in S$ . This condition implies that for each  $r \in R$ , the function

$$h_r(i) := \mathbb{P}_i(X_{T_R} = r)$$

is the unique function  $h$  such that  $h(i) = (\mathbf{P}h)(i)$  for all  $i \in S \setminus R$  with the boundary condition  $h(i) = 1(i = r)$  for  $i \in R$ , see, for example, Lyons and Peres [74], Section 2.1. Considering

$$h(i) := \frac{w_{ir}(R)}{w(R)},$$

it is obvious that this  $h$  satisfies the boundary condition, so it only remains to check that it is  $\mathbf{P}$ -harmonic on  $S \setminus R$ . After canceling the common factor of  $w(R)$  and putting all terms involving  $w_{ir}$  on the left-hand side, the harmonic equation for  $w_{ir}(R)$ ,  $i \in S \setminus R$  becomes

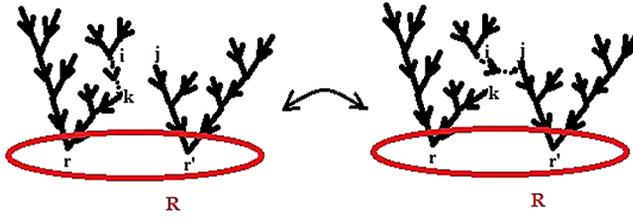
$$\left( \sum_{j \neq i} p_{ij} \right) w_{ir}(R) = \sum_{k \neq i} p_{ik} w_{kr}(R).$$

The equality of these two expressions is established by matching the terms appearing in the sums on the two sides. Specifically, for each fixed  $i \in S \setminus R$  and  $r \in R$  there is a matching

$$p_{ij} \Pi(\mathbf{f}) = p_{ik} \Pi(\mathbf{f}'), \tag{2.1}$$

where on the l.h.s.:  $j \neq i, i \overset{\mathbf{f}}{\rightsquigarrow} r$  and on the r.h.s.:  $k \neq i, k \overset{\mathbf{f}'}{\rightsquigarrow} r$  with both  $j$  and  $k$  ranging over all states in  $S$ , but always  $i \in S \setminus R$  and  $r \in R$ . If on the l.h.s. we have  $j \overset{\mathbf{f}}{\rightsquigarrow} r$ , then set  $k = j$  and  $\mathbf{f}' = \mathbf{f}$ . Then the r.h.s. conditions are met by  $(k, \mathbf{f}')$ , and (2.1) holds trivially. So we are reduced to matching, for each fixed choice of  $i \in S \setminus R$  and  $r \neq r' \in R$ , on the l.h.s.:  $j \neq i, i \overset{\mathbf{f}}{\rightsquigarrow} r, j \overset{\mathbf{f}}{\rightsquigarrow} r' \neq r$  and on the r.h.s.:  $k \neq i, k \overset{\mathbf{f}'}{\rightsquigarrow} r, i \overset{\mathbf{f}'}{\rightsquigarrow} r' \neq r$ .

Given  $(j, \mathbf{f})$  on the l.h.s., let  $i \rightarrow k$  be the edge out of  $i$  in  $\mathbf{f}$ . Create  $\mathbf{f}'$  by deleting this edge and replacing it with  $i \rightarrow j$ . Then it is easily seen that  $(k, \mathbf{f}')$  is as required on the r.h.s., and it is clear that (2.1) holds. Inversely, given  $(k, \mathbf{f}')$  as on the r.h.s., let  $i \rightarrow j$  be the edge out of  $i$  in  $\mathbf{f}'$ .



**Figure 2.** Bijection between the l.h.s.:  $i \rightsquigarrow r, j \rightsquigarrow r'$  and the r.h.s.:  $k \rightsquigarrow r, i \rightsquigarrow r'$ .

Pop this edge and replace it with  $i \rightarrow k$  to recover  $(j, \mathbf{f})$ . See Figure 2 for an illustration of the bijection.  $\square$

Next we make use of the harmonic tree formula to derive the Green tree formula. To this end, we need the following tree identity.

**Lemma 2.2.** For  $j \in S \setminus R$ ,

$$w(R \cup \{j\}) = w(R) + \sum_{k \in S \setminus R} p_{jk} w_{kj}(R \cup \{j\}).$$

**Proof.** Observe that

$$w(R \cup \{j\}) = \sum_{\text{ROOTS}(\mathbf{f})=R, \mathbf{t} \rightarrow j} \Pi^{\mathbf{P}}(\mathbf{f}) \Pi^{\mathbf{P}}(\mathbf{t}),$$

where the sum is over all forests  $\mathbf{f}$  whose set of roots is  $R$ , and all trees  $\mathbf{t}$  are directed towards  $j$ . Now for each choice of  $(\mathbf{f}, \mathbf{t})$ , we can split the product into two parts as

$$\Pi^{\mathbf{P}}(\mathbf{f}) \Pi^{\mathbf{P}}(\mathbf{t}) = \sum_{k \notin \mathbf{t}} \Pi^{\mathbf{P}}(\mathbf{f}) \Pi^{\mathbf{P}}(\mathbf{t}) p_{jk} + \sum_{k \in \mathbf{t}} \Pi^{\mathbf{P}}(\mathbf{f}) \Pi^{\mathbf{P}}(\mathbf{t}) p_{jk}.$$

The sum of the first part is evidently  $w(R)$ , with  $\Pi^{\mathbf{P}}(\mathbf{f}) \Pi^{\mathbf{P}}(\mathbf{t}) p_{jk}$  comprising those terms in  $w(R)$  indexed by forests  $\mathbf{f}'$  where the subtrees of  $\mathbf{f}'$  rooted at  $j$  equals  $\mathbf{t}$  and that subtree is attached to the remaining forest  $\mathbf{f}$  at vertex  $k \in \mathbf{f}$ . While the sum of the second part is

$$\begin{aligned} & \sum_{\text{ROOTS}(\mathbf{f})=R, \mathbf{t} \rightarrow j} \sum_{k \in \mathbf{t}} \Pi^{\mathbf{P}}(\mathbf{f}) \Pi^{\mathbf{P}}(\mathbf{t}) p_{jk} \\ &= \sum_{k \in S \setminus R} p_{jk} \left[ \sum_{\text{ROOTS}(\mathbf{f})=R, k \in \mathbf{t} \rightarrow j} \Pi^{\mathbf{P}}(\mathbf{f}) \Pi^{\mathbf{P}}(\mathbf{t}) \right] \\ &= \sum_{k \in S \setminus R} p_{jk} w_{kj}(R \cup \{j\}), \end{aligned}$$

from which the desired result follows.  $\square$

**Derivation of the Green tree formula (1.7) from the harmonic tree formula (1.9).** It follows from standard theory of Markov chains that for all  $i, j \in S \setminus R$ ,

$$\begin{aligned} \mathbb{E}_i \sum_{n=0}^{T_R-1} 1(X_n = j) &= \mathbb{P}_i(X_{T_{R \cup \{j\}}} = j) \times \mathbb{E}_j \sum_{n=0}^{T_R-1} 1(X_n = j) \\ &= \frac{w_{ij}(R \cup \{j\})}{w(R \cup \{j\})} \times \mathbb{E}_j \sum_{n=0}^{T_R-1} 1(X_n = j), \end{aligned} \quad (2.2)$$

where the last equality uses the harmonic tree formula (1.9) for  $j \in R \cup \{j\}$ . In addition,

$$\begin{aligned} \mathbb{E}_j \sum_{n=0}^{T_R-1} 1(X_n = j) &= 1 + \sum_{k \in S \setminus R} p_{jk} \times \mathbb{E}_k \sum_{n=0}^{T_R-1} 1(X_n = j) \\ &\stackrel{(2.2)}{=} 1 + \frac{\sum_{k \in S \setminus R} p_{jk} w_{kj}(R \cup \{j\})}{w(R \cup \{j\})} \times \mathbb{E}_j \sum_{n=0}^{T_R-1} 1(X_n = j), \end{aligned}$$

which together with Lemma 2.2 implies that

$$\begin{aligned} \mathbb{E}_j \sum_{n=0}^{T_R-1} 1(X_n = j) &= \frac{w(R \cup \{j\})}{w(R \cup \{j\}) - \sum_{k \in S \setminus R} p_{jk} w_{kj}(R \cup \{j\})} \\ &= \frac{w(R \cup \{j\})}{w(R)}. \end{aligned} \quad (2.3)$$

Injecting (2.3) into (2.2), we obtain the Green tree formula (1.7).

We illustrate the Green tree formula (1.7), by a derivation of Cayley's formula for the number of forests with a given set of roots. Cayley's formula is well known to be a direct consequence of Kirchhoff's matrix forest theorem, see, for example, Pitman [86], Corollary 2. The Green tree formula, while weaker than Kirchhoff's matrix forest theorem, still carries enough enumerative information about trees and forests to imply Cayley's formula.

**Corollary 2.3 (Cayley's formula [21]).** *Let  $1 \leq k \leq n$ . Then*

$$|\{\text{forests labeled by } [n] \text{ with root set } [k]\}| = kn^{n-k-1}. \quad (2.4)$$

**Proof.** Consider the Markov chain generated by an i.i.d. sequence of uniform random choices from  $S := [n]$  and run the chain until the first time it hits a state  $i \in R := [k]$ . The number of steps required is a geometric random variable  $T_{n,k}$  with mean  $n/k$ . In addition, the expectation of the intervening number of steps with mean  $(n-k)/k$  is equidistributed over the  $n-k$  other states. Thus, the expected number of visits to each of these other states prior to  $T_{n,k}$  is  $1/k$ .

Let  $\mathbf{P}_{n,k}$  be the  $(n-k) \times (n-k)$  substochastic transition matrix with all entries equal to  $1/n$ . It follows immediately from the above observation that the corresponding Green matrix

$(\mathbf{I} - \mathbf{P}_{n,k})^{-1}$  has entries  $1 + 1/k$  along the diagonal, all other entries being identically equal to  $1/k$ . To illustrate, for  $n = 6$  and  $k = 2$ ,

$$(\mathbf{I} - \mathbf{P}_{6,2})^{-1} = \begin{pmatrix} 5/6 & -1/6 & -1/6 & -1/6 \\ -1/6 & 5/6 & -1/6 & -1/6 \\ -1/6 & -1/6 & 5/6 & -1/6 \\ -1/6 & -1/6 & -1/6 & 5/6 \end{pmatrix}^{-1} = \begin{pmatrix} 3/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 3/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 3/2 & 1/2 \\ 1/2 & 1/2 & 1/2 & 3/2 \end{pmatrix}.$$

Let  $c(n, k)$  be the number of forests labeled by  $[n]$  with root set  $[k]$ . Then for  $k + 1 \leq i \neq j \leq n$ , the ratio of forest sums in the Green tree formula (1.7) is readily evaluated to give

$$\frac{1}{k} = \frac{c(n, k + 1)(k + 1)^{-1}n^{-(n-k-1)}}{c(n, k)n^{-(n-k)}}, \quad (2.5)$$

where the denominator sums the  $c(n, k)$  identical forest products  $n^{-(n-k)}$  from the  $k$ -tree forests with root set  $[k]$ , while the numerator sums the  $c(n, k + 1)/(k + 1)$  identical forest products  $n^{-(n-k-1)}$  from all the  $(k + 1)$ -tree forest products of trees with root set  $\{j\} \cup [k]$  in which  $i$  is contained in the tree with root  $j$ . Here the division by  $(k + 1)$  accounts for the fact that each tree has exactly one of  $k + 1$  distinct roots. The formula (2.5) simplifies to

$$c(n, k + 1) = \frac{(k + 1)c(n, k)}{kn} \quad \text{for } 1 \leq k \leq n - 2. \quad (2.6)$$

Since the enumerations  $c(n, n) = 1$  and  $c(n, n - 1) = n - 1$  are obvious, Cayley's formula (2.4) for  $c(n, k)$  follows immediately from (2.6).  $\square$

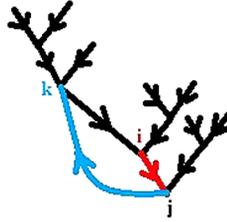
We also refer to Lyons and Peres [74], Corollary 4.5, for a proof of Cayley's formula by Wilson's algorithm, and to Pitman [88], Section 2, for that using the forest volume formula. Lyons and Peres' proof is similar in spirit to ours, and the relation between Wilson's algorithm and various tree formulas will be discussed in Section 5.

### 3. Markov chain tree theorems

In this section, we deal with the Markov chain tree theorems. To begin with, we present a three-sentence proof of Theorem 1.1, due to Ventcel and Freidlin [106], Lemma 7.1. They studied perturbed diffusion processes by Markov chain approximations, where Theorem 1.1 was used to estimate the first hitting time of the Markov chain to a set.

**Proof of Theorem 1.1.** Multiply the r.h.s. of (1.3) by  $\sum_k p_{jk} = 1$  and expand the tree sums  $\Sigma_i$  and  $\Sigma_j$  to include the extra transition factors:

$$\text{l.h.s.} = \sum_{i \in S} \sum_{\mathbf{t} \rightarrow i} p_{ij} \Pi^{\mathbf{P}}(\mathbf{t}) \quad \text{and} \quad \text{r.h.s.} = \sum_{k \in S} \sum_{\mathbf{t} \rightarrow j} \Pi^{\mathbf{P}}(\mathbf{t}) p_{jk}.$$



**Figure 3.** A cycle-rooted spanning tree whose cycle includes  $j$ .

Then both sides equal the sum of  $\Pi^{\mathbf{P}}(\mathbf{g})$  over all directed graphs  $\mathbf{g}$  which span  $S$  and contain exactly one cycle including  $j$ , see Figure 3. Such graphs  $\mathbf{g}$  are called *cycle-rooted spanning trees* (CRST), see Section 5 for definition.

The l.h.s. sum is split up according to the state  $i$  that precedes  $j$ , whereas the r.h.s. sum is split up according to the state  $k$  that follows  $j$ .  $\square$

Now let us consider a Markov chain  $(X_n)_{n \in \mathbb{N}}$  with transition matrix  $\mathbf{P}$ , which is not necessarily irreducible. Elementary considerations show that the state space  $S$  is uniquely decomposed into a list of disjoint recurrent classes  $C_1, \dots, C_k$ , and the transient states, see, for example, Feller [36], Chapter XV.6.

The general version of the Markov chain tree theorem is attributed to Leighton and Rivest [64,65]. Here we provide a probabilistic argument, see also Anantharam and Tsoucas [4] for an alternative proof.

**Theorem 3.1 (Markov chain tree theorem [64,65]).** *Assume that  $\mathbf{P}$  is a transition matrix with disjoint recurrent classes  $C_1, \dots, C_k$ . Let  $\mathcal{F}$  be a random forest picked from all forests  $\mathbf{f}$  consisting of  $k$  trees with one tree rooted in each  $C_i$ , and  $\mathbb{P}(\mathcal{F} = \mathbf{f}) = \Pi^{\mathbf{P}}(\mathbf{f})/W$  where  $W$  is the total weight of all such forests. Then*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N p_{ij}^n = \mathbb{P}(i \overset{\mathcal{F}}{\rightsquigarrow} j), \quad (3.1)$$

where the right-hand side is the probability that the tree of  $\mathcal{F}$  containing  $i$  has root vertex  $j$ .

**Proof.** If  $j$  is a transient state, then both sides of (3.1) equal to zero. Now let  $C \in \{C_1, \dots, C_k\}$  be the recurrent class containing  $j$ . According to Durrett [34], Theorem 6.6.1,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N p_{ij}^n = \mathbb{P}_i(T_C < \infty) \pi_j^C,$$

where  $T_C$  is the entry time to  $C$  and  $\pi_j^C$  is the stationary distribution in  $C$ . The result boils down to two special cases:

- (1) the ergodic case when  $\mathbf{P}$  is irreducible and all trees in  $\mathcal{F}$  have a single root, which is distributed according to the unique stationary distribution of  $\mathbf{P}$  regardless of  $i$ ;
- (2) the completely absorbing case when there is a set  $R$  of absorbing states with  $w(R) > 0$ , and  $\mathcal{F}$  is a forest whose set of roots is  $R$ .

In case (1), the conclusion reduces to Theorem 1.1, and in case (2) to the harmonic tree formula (1.9). In the general case, every possible forest  $\mathbf{f}$  consists of:

- some selection of roots  $R$ , with  $|R| = k$  and one root  $r_i$  in each  $C_i$ ,
- for each  $i$  a subtree  $\mathbf{t}_i$  spanning  $C_i$  with root  $r_i$ ,
- a collection of subtrees  $\mathbf{u}_c$  rooted at  $c \in \bigcup_{i=1}^k C_i$ .

Then an obvious factorization

$$\Pi^{\mathbf{P}}(\mathbf{f}) = \left( \prod_{i=1}^k \Pi^{\mathbf{P}}(\mathbf{t}_i) \right) \prod_{c \in \bigcup_{i=1}^k C_i} \Pi^{\mathbf{P}}(\mathbf{u}_c),$$

shows that  $\mathcal{F}$  decomposes into  $(k + 1)$  independent components,  $k$  subtrees spanning the  $C_i$ ,  $1 \leq i \leq k$  and a forest with roots  $\bigcup_{i=1}^k C_i$ , so it is easy to deduce the conclusion from the two special cases.  $\square$

The rest of this section concerns the Markov chain tree formula for mean first passage times, Theorem 1.3. First, we provide a simple proof of Theorem 1.3 by using the formula (1.8), which is derived from the Green tree formula (1.7).

**Proof of Theorem 1.3.** By setting  $R = \{j\}$  in the formula (1.8), we get:

$$m_{ij} = \frac{\sum_{k \neq j} w_{ik}(\{j, k\})}{w\{j\}}.$$

By definition,  $w(\{j\}) = \Sigma_j$ , and

$$w_{ik}(\{j, k\}) = \sum_{i \in \mathbf{s} \rightarrow k, \mathbf{u} \rightarrow j} \Pi^{\mathbf{P}}(\mathbf{s}, \mathbf{u}).$$

As a consequence,

$$\sum_{k \neq j} w_{ik}(\{j, k\}) = \sum_{i \in \mathbf{s}, \mathbf{u} \rightarrow j} \Pi^{\mathbf{P}}(\mathbf{s}, \mathbf{u}). \quad (3.2)$$

Combining (3.2) and (1.13) yields the desired result.  $\square$

As observed by Pokarowski [90], the Green tree formula (1.7) is a consequence of Kirchhoff's matrix forest theorem, Theorem 1.2. Now we give a combinatorial proof of Theorem 1.3, without appeal to the matrix tree theorem.

**Alternative proof of Theorem 1.3.** As seen in the beginning of this section, the formula (1.1) can be proved without using Kirchhoff's matrix forest theorem. We now fix  $i, j \in S$ , and apply formula (1.1) to the modified chain  $\tilde{\mathbf{P}} := (\tilde{p}_{ij}; i, j \in S)$  defined by

$$\tilde{p}_{ji} = 1, \quad \tilde{p}_{jk} = 0 \quad \text{for } k \neq i \quad \text{and} \quad \tilde{p}_{lk} = p_{lk} \quad \text{for } l \neq j.$$

So  $\tilde{\mathbf{P}}$  has a recurrent class  $\tilde{S}$  containing  $\{i, j\}$ , and it is possible that  $\tilde{S} \neq S$ . For  $k \in \tilde{S}$ , let  $\tilde{\Sigma}_k$  be the tree sum, over trees  $\tilde{\mathbf{t}} \rightarrow k$  spanning  $\tilde{S}$ , of the tree product

$$\Pi^{\tilde{\mathbf{P}}}(\tilde{\mathbf{t}}) = \prod_{i' \rightarrow j' \in \tilde{\mathbf{t}}} \tilde{p}_{i'j'}.$$

By construction,  $\tilde{m}_{jj} = 1 + m_{ij}$ . Thus,

$$m_{ij} = \tilde{m}_{jj} - 1 = \frac{\sum_{k \neq j} \tilde{\Sigma}_k}{\tilde{\Sigma}_j}. \quad (3.3)$$

Observe that the map  $\mathbf{t} \mapsto (\tilde{\mathbf{t}}, \mathbf{f})$  is a bijection between trees  $\mathbf{t}$  with  $\mathbf{t} \rightarrow j$  and tree forest pairs  $(\tilde{\mathbf{t}}, \mathbf{f})$  such that  $\tilde{\mathbf{t}} \rightarrow j$  spans  $\tilde{S}$  and  $\text{ROOTS}(\mathbf{f}) = \tilde{S}$ . This leads to a unique factorization  $\Pi^{\mathbf{P}}(\mathbf{t}) = \Pi^{\tilde{\mathbf{P}}}(\tilde{\mathbf{t}})\Pi^{\mathbf{P}}(\mathbf{f})$  for each  $\mathbf{t} \rightarrow j$ . By summing over all  $\mathbf{t} \rightarrow j$ , we get

$$\Sigma_j = \tilde{\Sigma}_j w(\tilde{S}), \quad (3.4)$$

where  $w(\tilde{S})$  is defined as in (1.4).

Further by cutting the edge  $k(i, j, \mathbf{t}) \rightarrow j$  from  $\mathbf{t}$ , the map  $\tilde{\mathbf{t}} \mapsto (\mathbf{s}, \mathbf{u})$  is a bijection between trees  $\tilde{\mathbf{t}} \rightarrow j$  spanning  $\tilde{S}$  and two tree forests  $(\mathbf{s}, \mathbf{u})$  such that  $i \in \mathbf{s}$  and  $\mathbf{u} \rightarrow j$ . Attaching  $\mathbf{u}$  to  $\mathbf{s}$  by  $j \rightarrow i$ , we define a tree  $\tilde{\mathbf{t}}' \rightarrow k(i, j, \mathbf{t})$  spanning  $\tilde{S}$ . Since  $\tilde{p}_{ji} = 1$ , we have

$$\Pi^{\tilde{\mathbf{P}}}(\tilde{\mathbf{t}})/p_{k(i,j,\mathbf{t})j} = \Pi^{\tilde{\mathbf{P}}}(\tilde{\mathbf{t}}').$$

Hence,  $\Pi^{\mathbf{P}}(\mathbf{t})/p_{k(i,j,\mathbf{t})j} = \Pi^{\tilde{\mathbf{P}}}(\tilde{\mathbf{t}}')\Pi^{\mathbf{P}}(\mathbf{f})$  for each  $\mathbf{t} \rightarrow j$ . Note that the map  $\tilde{\mathbf{t}} \mapsto \tilde{\mathbf{t}}'$  is a bijection between trees  $\tilde{\mathbf{t}} \rightarrow j$  and those  $\tilde{\mathbf{t}}' \rightarrow k \neq j$ . Again by summing over all  $\mathbf{t} \rightarrow j$ , we get

$$\Sigma_{ij} = \sum_{k \neq j} \tilde{\Sigma}_k w(\tilde{S}). \quad (3.5)$$

By injecting (3.4) and (3.5) into (3.3), we obtain the formula (1.11).  $\square$

**Chung's formula and tree identities.** Now by setting  $R = \{k\}$  in the Green tree formula (1.7), we get:

$$\mathbb{E}_i \sum_{n=0}^{T_k-1} 1(X_n = j) = \frac{w_{ij}(\{k, j\})}{\Sigma_k} \quad \text{for } i, j \neq k. \quad (3.6)$$

According to Chung [28], Theorem I.11.3, and Pitman [89], Example 4.11, for a positive recurrent chain,

$$\mathbb{E}_i \sum_{n=0}^{T_k-1} 1(X_n = j) = \frac{m_{ik} + m_{kj} - m_{ij}1(i \neq j)}{m_{jj}} \quad \text{for } i, j \neq k.$$

Further by Theorem 1.3, we have:

$$\mathbb{E}_i \sum_{n=0}^{T_k-1} 1(X_n = j) = \frac{\Sigma_{ik} \Sigma_j + \Sigma_{kj} \Sigma_k - \Sigma_{ij} \Sigma_k 1(i \neq j)}{\Sigma_k \Sigma^{(1)}} \quad \text{for } i, j \neq k. \quad (3.7)$$

By identifying (3.6) and (3.7), we obtain the following tree identities:

$$w_{ii}(\{k, i\}) \Sigma^{(1)} = \Sigma_{ik} \Sigma_i + \Sigma_{ki} \Sigma_k \quad \text{for } i \neq k, \quad (3.8)$$

$$w_{ij}(\{k, j\}) \Sigma^{(1)} + \Sigma_{ij} \Sigma_k = \Sigma_{ik} \Sigma_j + \Sigma_{kj} \Sigma_k \quad \text{for } i \neq j \text{ and } i, j \neq k. \quad (3.9)$$

It seems that these identities are non-trivial, and there is no simple bijective proof. So we leave the interpretation open for readers.

## 4. Kirchhoff's matrix tree theorems

We begin with the discussion of *Kirchhoff–Tutte's matrix tree theorem for directed graphs*. Let  $G := (V, \vec{E})$  be a directed finite graph with no multiple edges nor self loops, where  $\vec{E} \subset \{(i, j); i \neq j \in V\}$ . Equip each directed edge  $(i, j) \in \vec{E}$  with a weight  $c(i, j) \geq 0$ . The *graph Laplacian*  $\mathbf{L}^{G,c} = (l_{ij}^{G,c}; i, j \in V)$  is defined by: for  $i, j \in V$ ,

$$l_{ij}^{G,c} := \begin{cases} \sum_{k \neq i} c(i, k) & \text{if } i = j, \\ -c(i, j) & \text{if } (i, j) \in \vec{E}, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that the graph Laplacian of a directed graph is not necessarily symmetric. If we take  $c(i, j) = 1$  for all  $(i, j) \in \vec{E}$ , then  $\mathbf{L}^{G,1} = \mathbf{D}^G - \mathbf{A}^G$ , where  $\mathbf{D}^G$  is the *outer-degree matrix* of  $G$ , and  $\mathbf{A}^G$  is the *adjacency matrix* of  $G$ .

The following result, which we call *Kirchhoff–Chaiken–Chen's matrix forest theorem for directed graphs*, is due to Chaiken and Kleitman [23], Chaiken [22], and Chen [27].

**Theorem 4.1 (Kirchhoff–Chaiken–Chen's matrix forest theorem for directed graphs [22, 23,27]).** *Let  $R$  be a non-empty subset of  $V$ . Let  $\mathbf{L}^{G,c}(R)$  be the submatrix of  $\mathbf{L}^{G,c}$  obtained by removing all rows and columns indexed by  $R$ . Then*

$$\det \mathbf{L}^{G,c}(R) = \sum_{\text{ROOTS}(\mathbf{f})=R} \Pi^c(\mathbf{f}), \quad (4.1)$$

where  $\Pi^c(\mathbf{f}) := \prod_{(i,j) \in \mathbf{f}} c(i,j)$ , and the sum is over all forests whose set of roots is  $R$ .

For further discussion on forest matrices, we refer to Chebotarev and Agaev [26] and references therein.

The case  $R = \{i\}$ , which we call *Kirchhoff–Tutte’s matrix tree theorem for directed graphs*, was first proved by Tutte [104] based on the inductive argument of Brooks, Smith, Stone and Tutte [18]. It was independently discovered by Bott and Mayberry [16]. In particular,

$$\det \mathbf{L}^{G,1}(i) = |\{\text{spanning trees rooted at } i \in V\}|.$$

Orlin [83] made use of the inclusion-exclusion principle to prove this result. Zeilberger [109] gave a combinatorial proof by using cancellation arguments. For historical notes, we refer to Moon [82], Section 5.5, Tutte [105], Section VI.4, and Stanley [98], Section 5.6. There is a generalization of matrix tree theorems from graphs to simplicial complexes, initiated by Kalai [45] and developed by Duval, Klivans and Martin [35]. Lyons [72] extended the matrix tree theorem to CW-complexes. Minoux [81] studied the matrix tree theorem in the semiring setting. Masbaum and Vaintrob [79], and Abdesselam [1] considered Pfaffian tree theorems. Recently, de Tilière [32] discovered a Pfaffian half-tree theorem.

It is well known that a weighted directed graph  $(G, c)$  defines a Markov chain on the state  $S := V$ . In the sequel, let  $|S| = n$ . The transition matrix  $\mathbf{P} := (p_{ij}; i, j \in S)$  is given as

$$p_{ij} := -\frac{l_{ij}^{G,c}}{l_{ii}^{G,c}} = \frac{c(i,j)}{\sum_{k \neq i} c(i,k)} \quad \text{for } i \neq j \in S \quad \text{and} \quad p_{ii} = 0 \quad \text{for } i \in S. \quad (4.2)$$

Hence the transition matrix  $\mathbf{P}$  and the graph Laplacian  $\mathbf{L}^{G,c}$  are related by

$$\mathbf{L}^{G,c} = \mathbf{D}^{G,c}(\mathbf{I} - \mathbf{P}), \quad (4.3)$$

where  $\mathbf{D}^{G,c}$  is the diagonal matrix whose  $(i, i)$ -entry equals  $l_{ii}^{G,c}$ . Now we provide a proof of Theorem 1.2 (1.4) by using Theorem 4.1.

**Derivation of Theorem 1.2 (1.4) from Theorem 4.1.** The proof boils down to two subcases.

*Case 1.* For all  $i \in S$ ,  $p_{ii} = 0$ . By the relation (4.3), the formula (1.4) is an equivalent formulation of (4.1) in the context of Markov chains.

*Case 2.* There exists  $i \in S$  such that  $p_{ii} > 0$ . Let  $\mathcal{A} := \{i \in S; p_{ii} = 1\}$  be the set of absorbing states, and define a new chain  $\tilde{X}$  whose transition matrix  $\tilde{\mathbf{P}} = (\tilde{p}_{ij}; i, j \in S)$  is given by  $\tilde{p}_{ii} = 1$  if  $i \in \mathcal{A}$  and

$$\tilde{p}_{ij} := \frac{p_{ij}}{1 - p_{ii}} \quad \text{for } i \neq j \quad \text{and} \quad \tilde{p}_{ii} = 0, \quad \text{if } i \notin \mathcal{A}.$$

Note that the restriction of the transition matrix  $\tilde{\mathbf{P}}$  to  $(S \setminus \mathcal{A}) \times (S \setminus \mathcal{A})$  has all zeros on the diagonal. If  $\mathcal{A} \cap (S \setminus R) \neq \emptyset$ , then both sides of (1.4) equal to zero. Consider the case of  $\mathcal{A} \subset R$ . Let  $\tilde{\mathbf{L}} := \mathbf{I} - \tilde{\mathbf{P}}$ , then (1.4) holds for  $(\tilde{\mathbf{L}}, \tilde{\mathbf{P}})$ . Multiplying both sides by  $\prod_{j \notin R} (1 - p_{jj})$  and noting that  $\det \mathbf{L}(R) = \det \tilde{\mathbf{L}}(R) \prod_{j \notin R} (1 - p_{jj})$ , we obtain (1.4).

In the rest of this section, we present several applications of the matrix forest theorem.

## 4.1. A probabilistic expression for $\Sigma^{(1)}$

Recall the definition of  $\Sigma^{(1)}$  from (1.2). We present a result of Runge and Sachs [92] and Lyons [71], which expresses  $\Sigma^{(1)}$  as an infinite series whose terms have a probabilistic meaning.

By definition,

$$\det(\mathbf{I} - \mathbf{P} - t\mathbf{I}) = \det(\mathbf{L} - t\mathbf{I}) \quad \text{for all } t \in \mathbb{R}. \quad (4.4)$$

Let us look at the coefficient of  $t$  on both sides of (4.4). Let  $\text{adj}(\cdot)$  be the *adjugate matrix*. The coefficient of  $t$  in  $\det(\mathbf{L} - t\mathbf{I})$  is given by

$$-\text{Tr}[\text{adj}(\mathbf{L})] = -\sum_{i=1}^n \det \mathbf{L}(i) \stackrel{(*)}{=} -\sum_{i=1}^n \sum_{\mathbf{t} \rightarrow i} \Pi \mathbf{P}(\mathbf{t}) = -\Sigma^{(1)},$$

where  $(*)$  is obtained by applying Theorem 1.2 with  $R = \{i\}$ .

Assume that  $\mathbf{P}$  is irreducible and aperiodic. Let  $1 = \lambda_0, \lambda_1, \dots, \lambda_{n-1}$  be eigenvalues of the transition matrix  $\mathbf{P}$ . It follows from *Perron–Frobenius theory* that  $|\lambda_i| < 1$  for  $1 \leq i \leq n-1$ . See Meyer [80], Chapter 8, for development. Then the coefficient of  $t$  in  $\det(\mathbf{I} - \mathbf{P} - t\mathbf{I})$  is

$$\begin{aligned} -\prod_{i=1}^{n-1} (1 - \lambda_i) &= -\exp\left(\sum_{i=1}^{n-1} \log(1 - \lambda_i)\right) = -\exp\left(-\sum_{i=1}^{n-1} \sum_{k \geq 1} \frac{\lambda_i^k}{k}\right) \\ &= -\exp\left(-\sum_{k \geq 1} \frac{1}{k} (\text{Tr} \mathbf{P}^k - 1)\right) = -\exp\left[-\sum_{k \geq 1} \frac{1}{k} \left(\sum_{i=1}^n p_{ii}^{(k)} - 1\right)\right], \end{aligned}$$

where  $p_{ii}^{(k)}$  is the probability that the Markov chain starting at  $i$  returns to  $i$  after  $k$  steps. Therefore,

$$\Sigma^{(1)} = \exp\left[-\sum_{k \geq 1} \frac{1}{k} \left(\sum_{i=1}^n p_{ii}^{(k)} - 1\right)\right]. \quad (4.5)$$

## 4.2. Matrix tree theorems and Markov chain tree theorems

We prove that the Markov chain tree theorem for irreducible chains, Theorem 1.1, and Kirchhoff–Tutte's matrix tree theorem, Theorem 1.2 (1.4) for  $R = \{i\}$  can be derived from each other. The argument is borrowed from Leighton and Rivest [65], and Sahi [93].

**Derivation of Theorem 1.1 from Theorem 1.2 (1.4).** Observe that for an irreducible chain with Laplacian matrix  $\mathbf{L} := \mathbf{I} - \mathbf{P}$ , the stationary distribution  $\boldsymbol{\pi} = (\pi_i)_{i \in S}$  is uniquely determined by

$$\boldsymbol{\pi} \mathbf{L}(\mathbf{1}, i) = \mathbf{E}_i \quad \text{for } i \in S,$$

where  $\mathbf{L}(\mathbf{1}, i)$  is the matrix obtained from  $\mathbf{L}$  by replacing the  $i$ th column by  $\mathbf{1} := (1, \dots, 1)^T$ , and  $\mathbf{E}_i$  is the vector with a one in the  $i$ th column and zeros elsewhere. By Cramér's rule,

$$\pi_i = \frac{\det \mathbf{L}(i)}{\det \mathbf{L}(\mathbf{1}, i)} = \frac{\det \mathbf{L}(i)}{\sum_{j \in S} \det \mathbf{L}(j)} \quad \text{for } i \in S.$$

Let  $(\Sigma_i)_{i \in S}$  and  $\Sigma^{(1)}$  be defined as in (1.2). According to the formula (1.4) for  $R = \{i\}$ ,

$$\pi_i = \frac{\Sigma_i}{\Sigma^{(1)}} \quad \text{for } i \in S,$$

which leads to Theorem 1.1.

**Derivation of Theorem 1.2 (1.4) for  $R = \{i\}$  from Theorem 1.1.** By Theorem 1.1,

$$\Sigma_i = \frac{\det \mathbf{L}(i)}{\sum_{j \in S} \det \mathbf{L}(j)} \Sigma^{(1)} \quad \text{for } i \in S.$$

Observe that  $\det \mathbf{L}(i)$ ,  $\sum_{j \in S} \det \mathbf{L}(j)$ ,  $\Sigma_i$  and  $\Sigma^{(1)}$  are all homogeneous polynomials of degree  $n - 1$  in variables  $(p_{jk}; j \neq k \in S)$  with integer coefficients. Now we prove the following lemma.

**Lemma 4.2.** *For each  $i \in S$ , the polynomial  $\det \mathbf{L}(i)$  is irreducible.*

**Proof.** Identify the set  $S$  with  $\{0, 1, \dots, n - 1\}$ . By symmetry, it suffices to consider  $\det \mathbf{L}(0)$ . Note that for  $1 \leq i, j \leq n - 1$ ,

$$\mathbf{L}(0)_{ij} = \begin{cases} -p_{ij} & \text{if } i \neq j, \\ \sum_{k \neq i} p_{ik} & \text{if } i = j. \end{cases}$$

It is easy to check that  $(p_{ij}; 1 \leq i \leq n - 1, 0 \leq j \leq n - 1, i \neq j) \mapsto (\mathbf{L}(0)_{ij}; 1 \leq i, j \leq n - 1)$  is an invertible linear map. According to Bocher [13], Section 61,  $\det \mathbf{L}(0)$  is irreducible as a polynomial in the matrix entries.  $\square$

By Lemma 4.2,  $\det \mathbf{L}(i)$  and  $\sum_{j \in S} \det \mathbf{L}(j)$  do not have any common factor, since the terms of  $\det \mathbf{L}(i)$  are strictly included in the sum  $\sum_{j \in S} \det \mathbf{L}(j)$ . It follows that

$$\Sigma_i = \lambda \det \mathbf{L}(i) \quad \text{for some rational } \lambda.$$

By considering the coefficient of  $\prod_{k \neq i} p_{ki}$  on both sides, we obtain  $\lambda = 1$  as desired.

By decomposing the state space  $S$  into recurrent classes and transient sets, a similar argument as above shows that the Markov chain tree theorem, Theorem 3.1, and Kirchhoff's matrix forest theorem, Theorem 1.2 can be derived from each other.

### 4.3. Matrix tree theorem for undirected graphs

Consider the case where  $c(i, j) = c(j, i)$  for  $i \neq j \in V$ , so that the graph Laplacian  $\mathbf{L}^{G,c}$  is symmetric positive semi-definite. The following result, known as *Kirchhoff's matrix tree theorem for undirected graphs* [58,59], is easily derived from Theorem 4.1.

**Theorem 4.3 (Kirchhoff's matrix tree theorem for undirected graphs [58]).** *For  $i, j \in V$ , let  $\mathbf{L}^{G,c}(i; j)$  be the submatrix of  $\mathbf{L}^{G,c}$  obtained by removing the  $i$ th row and  $j$ th column. Then*

$$\det \mathbf{L}^{G,c}(i; j) = \sum_{t \in \text{TREES}} \Pi^c(t), \quad (4.6)$$

where  $\Pi^c(t) := \prod_{\{i', j'\} \in t} c(i', j')$ , and the sum is over all unrooted spanning trees in  $G$ . In particular,

$$\det \mathbf{L}^{G,1}(i; j) = |\{\text{unrooted spanning trees in } G\}|.$$

We refer to Moon [82], Section 5.3, in which Theorem 4.3 was proved by using the *Cauchy–Binet formula*. The most classical application of Theorem 4.3 is to count unrooted spanning trees in the complete graph  $K_n$ , known as Cayley's formula [21]. See also Pak and Postnikov [84], Section 3, for enumerating unrooted spanning trees by using the property of reciprocity for some tree-sum degree polynomial.

Theorem 4.3 states that every minor of the graph Laplacian  $\mathbf{L}^{G,c}$  is identical to the tree sum-product as in (4.6). Kelmans and Chelnikov [51] expressed these minors in terms of the eigenvalues  $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{n-1}$  of  $\mathbf{L}^{G,c}$ :

$$\det \mathbf{L}^{G,c}(i; j) = \frac{1}{n} \prod_{k=1}^{n-1} \lambda_k. \quad (4.7)$$

It was observed by Biggs [12], Corollary 6.5, that (4.7) can be derived from Temperley's identity [100]:  $\det \mathbf{L}^{G,c}(i; j) = \frac{1}{n^2} \det(\mathbf{L}^{G,c} + J)$ , where  $J$  is  $n \times n$  matrix with all entries equal to 1. Recently, Kozdron, Richards and Stroock [61], Theorem 2.2, observed that the formula (4.7) is a direct consequence of *Crámer's formula* and the *Jordan–Chevalley decomposition*, see also Stroock [99], Section 3.2.2.

Here we give a lesser known example of counting spanning trees in the complete prism. Recall that the Cartesian product  $G \square H$  of graphs  $G$  and  $H$  is the graph such that  $V(G \square H) = V(G) \times V(H)$ , and  $(u_G, u_H)$  is adjacent with  $(v_G, v_H)$  if and only if  $u_G = v_G$  and  $u_H$  is adjacent to  $v_H$  in  $H$ , or  $u_H = v_H$  and  $u_G$  is adjacent to  $v_G$  in  $G$ .

**Example 4.4 (Boesch and Prodinger [14]).** We aim at counting spanning trees in the complete prism  $K_n \square C_m$ , that is the Cartesian product of the complete graph  $K_n$  and the circulant graph  $C_m$  whose adjacency matrix is a permutation matrix. The graph Laplacian of  $K_n \square C_m$  is written as

$$\mathbf{L}^{K_n \square C_m, 1} = \mathbf{D}^{K_n \square C_m} - \mathbf{A}^{K_n \square C_m} = (n+1)I_{nm} - \mathbf{A}^{K_n} \oplus \mathbf{A}^{C_m},$$

where  $\mathbf{D}^G$  (resp.  $\mathbf{A}^G$ ) is the degree matrix (resp. the adjacency matrix) of  $G$ , and  $\oplus$  is the Kronecker sum of two matrices: if  $A$  is  $m \times m$  matrix and  $B$  is  $n \times n$  matrix, then

$$A \oplus B := A \otimes I_n + I_m \otimes B, \quad \text{where } \otimes \text{ is the usual tensor product of two matrices.}$$

Note that  $\mathbf{A}^{K_n}$  has eigenvalues  $-1$  with multiplicity  $n-1$ , and  $n-1$  with multiplicity 1, and  $\mathbf{A}^{C_m}$  has eigenvalues  $2 \cos\left(\frac{2k\pi}{m}\right)$  for  $0 \leq k \leq m-1$ . It is known that the eigenvalues of the Kronecker sum of two matrices are all possible sums of eigenvalues of the individual matrices, see Bellman [6], Chapter 12, Section 11. From Theorem 4.3 and (4.7), we deduce that

$$\begin{aligned} & |\{\text{unrooted spanning trees in } K_n \square C_m\}| \\ &= \frac{1}{nm} \left( \prod_{k=0}^{m-1} \left[ m + 2 - 2 \cos\left(\frac{2k\pi}{m}\right) \right] \right)^{n-1} \prod_{k=1}^{m-1} \left[ 2 - 2 \cos\left(\frac{2k\pi}{m}\right) \right] \\ &= mn^{n-2} \left[ U_{m-1} \left( \sqrt{\frac{n+4}{4}} \right) \right]^{2n-2}, \end{aligned}$$

where  $U_{m-1}$  is the *Chebyshev polynomial of the second kind*; that is

$$U_{m-1}(x) := -\frac{1}{2\sqrt{x^2-1}} \left[ (x + \sqrt{x^2-1})^m - (x - \sqrt{x^2-1})^m \right].$$

See Boesch and Prodinger [14], Benjamin and Yerger [7], and Zhang, Yong and Golin [110] for enumerating spanning trees in a wide class of graphs.

## 5. Wilson's algorithm and tree formulas

In this section, we explore the connections between Wilson's algorithm, the Green tree formula (1.7) and the harmonic tree formula (1.9). Wilson's algorithm [107] was originally devised to generate a random tree whose probability distribution over trees  $\mathbf{t}$  rooted at a fixed  $r \in S$  is proportional to the tree product  $\Pi(\mathbf{t})$ . The constant of normalization is  $w(\{r\})$  as in (1.4) for the case  $R = \{r\}$ . See, for example, Lyons and Peres [74], Section 4.1, and Grimmett [39], Section 2.1, for further development.

The algorithm has extensions in several directions. Marchal [77,78] provided a similar procedure to construct random Hamiltonian cycles. Gorodezky and Pak [38] gave a version of Wilson's algorithm in the hypergraph setting. Kassel and Kenyon [47] generalized Wilson's algorithm for sampling cycle-rooted spanning forests, which will be discussed later.

For any finite path  $(x_0, x_1, \dots, x_l)$  in a directed graph, its *loop erasure*  $(u_0, u_1, \dots, u_m)$  is defined by erasing cycles in chronological order. More precisely, set  $u_0 := x_0$ . If  $x_l = x_0$ , we set  $m = 0$  and terminate; otherwise, let  $u_1$  be the first vertex after the last visit to  $x_0$ , i.e.  $u_1 := x_{i+1}$ , where  $i := \max\{j; x_j = x_0\}$ . If  $x_l = u_1$ , then we set  $m = 1$  and terminate; otherwise, let  $u_2$  be the first vertex after the last visit to  $u_1$ , and so on.

In the sequel,  $(X_n)_{n \in \mathbb{N}}$  is a Markov chain with transition matrix  $\mathbf{P} := \{p_{ij}; i, j \in S\}$ . Now we describe Wilson's original cycle-popping algorithm.

**Wilson's algorithm for generating a random spanning tree rooted at  $r$ :** Start a copy of  $(X_n)_{n \in \mathbb{N}}$  at any arbitrary state  $i$ , run the chain until it hits  $r$ , and then perform a loop erasure operation to obtain a path from  $i$  to  $r$ . This path will then be the unique path in the ultimately generated tree produced by following stages of the algorithm, in which another copy of  $(X_n)_{n \in \mathbb{N}}$  is started at any arbitrary state not in this path, until it hits the path, and so on, growing an increasing family of trees which eventually span all of  $S \setminus \{r\}$ , when the algorithm terminates.

Next, we consider Wilson's algorithm to produce random spanning forests.

**Wilson's algorithm for generating a random spanning forest with roots  $R$ :** Start a copy of  $(X_n)_{n \in \mathbb{N}}$  at any arbitrary state  $i$ , run the chain until it hits  $j \in R$ , and then perform a loop erasure operation to obtain a path from  $i$  to  $j \in R$ . This path will then be the unique path in the ultimately generated forest produced by following stages of the algorithm, in which another copy of  $(X_n)_{n \in \mathbb{N}}$  is started at any arbitrary state not in this path, until it hits either the path or a different  $j' \in R$ , and so on, growing an increasing family of forests which eventually span all of  $S \setminus R$ , when the algorithm terminates.

**Proposition 5.1.** *Wilson's algorithm for generating a random forest spanning  $S$  with roots  $R$  terminates in finite time almost surely if and only if  $w(R) > 0$ , in which case Wilson's algorithm generates each forest  $\mathbf{f}$  with  $\text{ROOTS}(\mathbf{f}) = R$  with probability  $\Pi^{\mathbf{P}}(\mathbf{f})/w(R)$ .*

**Proof.** This can be proved by simply adapting the cycle-popping argument of Wilson [107] to the present case. However, the result can also be derived from the more standard case of irreducible chains as follows. Let  $\partial$  be an additional state not in  $S$ , and consider the modified Markov chain with state space  $\tilde{S} := S \cup \partial$  and transition matrix

$$\tilde{p}_{ij} := \begin{cases} p_{ij}1(j \in S) & \text{if } i \in S \setminus R, \\ 1(j = \partial) & \text{if } i \in R, \\ 1(j \in S \setminus R)/|S \setminus R| & \text{if } i = \partial. \end{cases}$$

It is straightforward that  $\tilde{\mathbf{P}} := (\tilde{p}_{ij}; i, j \in \tilde{S})$  is irreducible if and only if  $w(R) > 0$ . Also, if  $\mathbf{t}$  is a tree with root  $\partial$  and  $\mathbf{f}$  is the restriction of  $\mathbf{t}$  to  $S$ , then  $\Pi^{\tilde{\mathbf{P}}}(\mathbf{t}) = \Pi^{\mathbf{P}}(\mathbf{f})$ . Wilson's algorithm for generating a forest  $\mathbf{f}$  spanning  $S$  with root set  $R$  is now seen to be a variation of Wilson's algorithm to generate  $\mathbf{t}$  spanning  $\tilde{S}$  with root  $\partial$ .  $\square$

Avena and Gaudillière [5] were interested in random forests with random roots  $\mathcal{R}$ . They proved that under additional killing rates, the set of roots sampled by Wilson's algorithm is a determinantal process. Chang and Le Jan [24] showed how Poisson loops arise in the construction of random spanning trees by Wilson's algorithm. Though closely related to ours, their situations seem to be more complicated.

**Derivation of the harmonic tree formula (1.9) from Wilson's algorithm.** Observe that the harmonic tree formula (1.9) is a consequence of the success of Wilson's algorithm for sampling  $\mathbf{f}$  with probability proportional to  $\Pi^{\mathbf{P}}(\mathbf{f})$ . The first stage of Wilson's algorithm is to start a copy of

the Markov chain  $(X_n)_{n \in \mathbb{N}}$  at some  $i \notin R$  and run it until time  $T_R$ . For each  $r \in R$ , the eventual forest  $\mathbf{f}$  generated by Wilson's algorithm has  $r$  as the root of the tree containing  $i$  if and only if this first stage results in  $X_{T_R} = r$ . Since  $\mathbf{f}$  ends up distributed with probability proportional to  $\Pi^{\mathbf{P}}(\mathbf{f})$ , the formula (1.9) is immediate.

We have shown in Section 2 that the Green tree formula (1.7) can be deduced from the harmonic tree formula (1.9). Now we make use of the Green tree formula to prove Wilson's algorithm. These imply that Wilson's algorithm, the harmonic tree formula (1.9) and the Green tree formula (1.7) can be derived from each other.

**Derivation of Wilson's algorithm from the Green tree formula (1.7).** We borrow the argument which first appeared in Lawler [62], Section 12.2, and was further developed by Marchal [78] and Kozdron, Richards and Stroock [61]. See also Lawler and Limic [63], Section 9.7, and Stroock [99], Section 3.3.

By the strong Markov property, the probability that  $(i_1, \dots, i_K)$  with  $i_K \in R$  are successive states visited by loop-erased chain  $(X_n; 0 \leq n \leq T_R)$  can be written as follows, using the notation  $R_1 := R$  and  $R_j := R \cup \{i_1, \dots, i_{j-1}\}$  for  $2 \leq j \leq K$ .

$$\begin{aligned}
 \mathbb{P}^R(i_1, \dots, i_K) &= \sum_{m_1, \dots, m_{K-1}=0}^{\infty} \mathbb{P}_{i_1}(T_{i_1} < T_{R_1})^{m_1} p_{i_1 i_2} \cdots \mathbb{P}_{i_{K-1}}(T_{i_{K-1}} < T_{R_{K-1}})^{m_{K-1}} p_{i_{K-1} i_K} \\
 &= \prod_{k=1}^{K-1} p_{i_k i_{k+1}} \frac{1}{1 - \mathbb{P}_{i_k}(T_{i_k} < T_{R_k})} \\
 &= \prod_{k=1}^{K-1} p_{i_k i_{k+1}} \cdot \mathbb{E}_{i_k} \sum_{n=0}^{R_k} 1(X_n = i_k) \\
 &= \frac{w(R_{K-1})}{w(R)} \prod_{k=1}^{K-1} p_{i_k i_{k+1}},
 \end{aligned} \tag{5.1}$$

where the last equality follows from the Green tree formula (1.7). We initialize Wilson's algorithm by setting  $V_0 := \{r\}$ . Then define recursively  $V_l$ , the set of states visited up to  $l$ th iteration, and  $\mathbf{t}_l$  the tree branch added at  $l$ th iteration. From (5.1), we deduce the probability for a spanning tree  $\mathbf{t}$  with root  $r \in S$  generated by Wilson's algorithm

$$\prod_{l \geq 1} \frac{w(V_l)}{w(V_{l-1})} \Pi^{\mathbf{P}}(\mathbf{t}_l) = \frac{w(S)}{w(\{r\})} \Pi^{\mathbf{P}}(\mathbf{t}) = \frac{\Pi^{\mathbf{P}}(\mathbf{t})}{w(\{r\})}.$$

We conclude the section by presenting a generalized Wilson's algorithm, due to Kassel and Kenyon [47], for sampling cycle-rooted spanning forests. To proceed further, we need some definitions. Let  $G := (V, E)$  be a directed finite graph, and  $R$  be a subset of  $V$ .

- A *cycle-rooted spanning forest* (CRSF) in  $G$  is a subgraph such that each connected component is a *cycle-rooted tree*, that is containing a unique oriented cycle, and edges in the bushes (i.e. not in the cycles) being directed towards the cycle.
- An *essential cycle-rooted spanning forest* (ECRSF) of  $(G, R)$  is a subgraph such that each connected component is either a tree directed towards a vertex in  $R$ , or a *cycle-rooted tree* containing no vertices in  $R$ . In particular, an ECRSF of  $(G, \emptyset)$  is a CRSF.

Now to each cycle  $\gamma$  of the Markov chain  $(X_n)_{n \in \mathbb{N}}$ , we assign a parameter of selection  $\alpha(\gamma) \in [0, 1]$ . For an ECRSF  $\mathbf{f}_{\text{ec}}$  with vertex set  $S$ , call

$$\Pi^{\mathbf{P}, \alpha}(\mathbf{f}_{\text{ec}}) := \prod_{i \rightarrow j \in \mathbf{f}_{\text{ec}}} p_{ij} \prod_{\gamma \subset \mathbf{f}_{\text{ec}}} \alpha(\gamma)$$

the  $(\mathbf{P}, \alpha)$ -weight of  $\mathbf{f}_{\text{ec}}$ .

Kassel and Kenyon's generalized Wilson's algorithm generates a random ECRSF whose probability distribution over ECRSFs  $\mathbf{f}_{\text{ec}}$  with tree roots  $R$  is proportional to  $\Pi^{\mathbf{P}, \alpha}(\mathbf{f}_{\text{ec}})$ . The method is a refinement of the cycle-popping idea: we simply run Wilson's algorithm, and when a cycle  $\gamma$  is created, flip a coin with bias  $\alpha(\gamma) \in [0, 1]$  to decide whether to keep or to pop it.

**Kassel–Kenyon–Wilson's algorithm for generating an ECRSF with roots  $R$ .** Let  $\Xi$  be a directed subgraph, initially set to be the tree roots  $R$ . Start a copy of  $(X_n)_{n \in \mathbb{N}}$  at any arbitrary state  $i \notin \Xi$ , and run the chain until it first reaches a state  $j$ , which either belongs to  $\Xi$  or creates a loop in the path.

- If  $j \in R$ , then replace  $\Xi$  by the union of  $\Xi$  and the path which is just traced. Start another copy of  $(X_n)_{n \in \mathbb{N}}$  at  $i \notin \Xi$  and repeat the procedure.
- If a loop  $\gamma$  is created at the visit of the state  $j$ , then sample an independent  $\{0, 1\}$ -Bernoulli random variable with success probability  $\alpha(\gamma)$ .
  - If the outcome is 1, then replace  $\Xi$  by the union of  $\Xi$  and the path which is just traced. Start another copy of  $(X_n)_{n \in \mathbb{N}}$  at  $i \notin \Xi$  and repeat the procedure.
  - If the outcome is 0, then pop the loop and continue the chain until it hits  $\Xi$  or a loop is created. In this case, repeat the above procedure.

**Theorem 5.2.** [47] *Kassel–Kenyon–Wilson's algorithm for generating a random essential cycle-rooted forest spanning  $S$  with tree roots  $R$  terminates in finite time almost surely if and only if at least one cycle has a positive parameter of selection, which is equivalent to*

$$\sum_{\text{ROOTS}(\mathbf{f}_{\text{ec}})=R} \Pi^{\mathbf{P}, \alpha}(\mathbf{f}_{\text{ec}}) > 0,$$

where the sum is over all ECRSFs labeled by  $S$  with tree roots  $R$ . In this case, the algorithm generates each ECRSF  $\mathbf{f}_{\text{ec}}$  with tree roots  $R$  with probability  $\Pi^{\mathbf{P}, \alpha}(\mathbf{f}_{\text{ec}})/w^{\text{ec}}(R)$ .

When  $\alpha = 0$ , Theorem 5.2 specializes to Wilson's algorithm for generation of spanning forests, Proposition 5.1. As Wilson's algorithm is related to various matrix tree theorems, Theorem 5.2

has some affinity to *Forman–Kenyon’s matrix CRSF theorem* [37,55]. They proved that the determinant of the *line bundle Laplacian matrix* with *Dirichlet boundary* can be expressed as a certain (E)CRSF sum-product. See Kassel [46] for derivation of Forman–Kenyon’s matrix CRSF theorem from Theorem 5.2.

Observe that if we set  $\alpha(\gamma) = 1$  for all cycles  $\gamma$ , then  $\Pi^{\mathbf{P},1}(\mathbf{f}_{ec}) = \Pi^{\mathbf{P}}(\mathbf{f}_{ec})$ . For  $i \in S$  and  $j \in R$ , let

$$w_{ij}^{ec}(R) := \sum_{\text{ROOTS}(\mathbf{f}_{ec})=R, i \rightsquigarrow j} \Pi^{\mathbf{P}}(\mathbf{f}_{ec}),$$

be the  $(\mathbf{P}, 1)$ -weight of all ECRSFs  $\mathbf{f}_{ec}$  with tree roots  $R$ , in which the tree containing  $i$  has root  $j$ , and

$$w^{ec}(R) := \sum_{\text{ROOTS}(\mathbf{f}_{ec})=R} \Pi^{\mathbf{P}}(\mathbf{f}_{ec}).$$

Let  $T_{\text{loop}}$  be the first time at which a loop is created, and  $T_R$  is the entry time of the set  $R \subset S$ . As a direct application of Theorem 5.2, we derive an analog of the harmonic tree formula:

$$\mathbb{P}_i(X_{T_R \wedge T_{\text{loop}}} = j) = \frac{w_{ij}^{ec}(R)}{w^{ec}(R)} \quad \text{for } i \in S \text{ and } j \in R, \quad (5.2)$$

and by summing over  $j \in R$ ,

$$\mathbb{P}_i(T_R < T_{\text{loop}}) = \frac{\sum_{j \in R} w_{ij}^{ec}(R)}{w^{ec}(R)} \quad \text{for } i \in S. \quad (5.3)$$

## 6. Loose ends and further references

### 6.1. Spanning trees and other models

Spanning trees are closely related to various mathematical models. The connection between uniform spanning trees and loop-erased random walks has been discussed in Section 5. Here we list two more examples.

- Temperley [101,102] established a bijection between spanning trees of a square grid and perfect matchings/dimer coverings of a related square grid. This bijection was extended to general planar graphs by Burton and Pemantle [19], and Kenyon, Propp and Wilson [56]. As an analog of Kirchhoff’s matrix tree theorem for enumerating spanning trees, Kasteleyn–Temperley–Fisher’s theory [50,103] expresses the number of dimer configurations in a graph in terms of the determinant of the *Kasteleyn matrix*. See Wu [108] and Kenyon [54], Section 3, for further development.
- Dhar [31], and Majumdar and Dhar [76] constructed a bijective map between spanning trees of a graph and recurrent sandpiles on that graph. This map is not unique, and an alternative bijection was provided by Cori and Le Borgne [29]. See Járai [44] for development on sandpile models. Recently, Kassel and Wilson [49] have developed a new approach to

computing sandpile densities of planar graphs by using a two-component forest formula of Liu and Chow [68].

See also Sokal [96] for how spanning trees arise as the limit of  $q$ -Potts model as  $q \rightarrow 0$ , and Bogner and Weinzierl [15] for the use of spanning trees in the quantum field theory.

## 6.2. Kemeny's constant and enumerate of spanning forests

We start with a simple example of Corollary 1.4. Let  $(X_k)_{k \in \mathbb{N}}$  be a Markov chain with state space  $S := [n]$ , and transition matrix  $\mathbf{P} = (p_{ij}; i, j \in [n])$  defined by

$$p_{ij} := \frac{1}{n} \quad \text{for all } i, j \in [n].$$

Let  $K$  be Kemeny's constant defined by (1.14) for the chain  $(X_k)_{k \in \mathbb{N}}$ . By Corollary 1.4,

$$K = 1 + n \cdot \frac{|\{\text{rooted two-component forests spanning } S\}|}{|\{\text{rooted trees spanning } S\}|}. \quad (6.1)$$

According to Corollary 2.4 (Cayley's formula),

$$|\{\text{rooted trees spanning } S\}| = n^{n-1},$$

and

$$|\{\text{rooted two-component forests spanning } S\}| = \binom{n}{2} \cdot 2n^{n-3},$$

which yields  $K = n$ .

Generally, we consider a weighted directed graph  $(G, c)$ . Define the ratio

$$R(G, c) := \frac{\Sigma_c^{(2)}}{\Sigma_c^{(1)}}, \quad (6.2)$$

with

$$\Sigma_c^{(r)} := \sum_{\mathbf{t}_1, \dots, \mathbf{t}_r} \Pi^c(\mathbf{t}_1, \dots, \mathbf{t}_r) \quad \text{for } r \in \mathbb{N}, \quad (6.3)$$

where the sum is over all forests of  $r$  directed trees  $\mathbf{t}_1, \dots, \mathbf{t}_r$  spanning  $G$ . So

$$R(G, 1) = \frac{n-1}{n} \sim 1 \quad \text{as } n \rightarrow \infty.$$

In particular, the relation (6.1) leads to  $K = 1 + nR(G, 1) = n$ .

Similarly, for a weighted undirected graph  $(G, c)$ , we define the ratio  $R'(G, c)$  as in (6.2)–(6.3) but with the sum over all unrooted trees/forests spanning  $G$ . In this case,

$$R'(G, 1) = \frac{|\{\text{unrooted two-component forests spanning } S\}|}{|\{\text{unrooted trees spanning } S\}|}.$$

By an obvious bijection between unrooted spanning trees and rooted spanning trees with a particular root, we get

$$|\{\text{unrooted trees spanning } S\}| = n^{n-2}.$$

Following from Moon [82], Theorem 4.1, we have

$$|\{\text{unrooted two-component forests spanning } S\}| = \frac{n^{n-4}(n-1)(n+6)}{2}.$$

Consequently,

$$R'(G, 1) = \frac{(n-1)(n+6)}{2n^2} \sim \frac{1}{2} \quad \text{as } n \rightarrow \infty.$$

Recently, Kassel and Wilson [49] considered the case where  $G$  is a planar graph, and derived the asymptotic of  $R'(G, c)$  as  $n \rightarrow \infty$ . See also Kenyon and Wilson [57], and Kassel, Kenyon and Wu [48] for related results.

### 6.3. Asymptotic enumeration of spanning trees

We have seen that Kirchhoff's matrix tree theorem enumerates explicitly spanning trees in finite graphs. It is interesting to understand the asymptotics of the number of spanning trees of a sequence of finite graphs that "approach" an infinite graph.

The following result was proved by Lyons [71, 73]. Let  $(G_n; n \in \mathbb{N})$  be a sequence of finite connected graphs with bounded average degree, converging in the local weak sense of Benjamini-Schramm [9] to a probability measure  $\rho$  on an infinite graph  $G$ . Then

$$\lim_{n \rightarrow \infty} \frac{1}{|V(G_n)|} \log |\{\text{unrooted spanning trees in } G_n\}| = \mathbf{h}(\rho), \quad (6.4)$$

where  $\mathbf{h}(\rho)$  is called the *tree entropy* of  $\rho$  on  $G$ . If  $\rho$  is unimodular, then  $\mathbf{h}(\rho) = \log \det_\rho \mathbf{L}^G$ , where  $\det_\rho \mathbf{L}^G$  is the *Fuglede–Kadison determinant* of the graph Laplacian  $\mathbf{L}^G$ . The notion of Fuglede–Kadison determinant originates from von Neumann algebra, see e.g. de la Harpe [30] for a quick review.

As explained by Lyons [71], the tree entropy  $\mathbf{h}(\rho)$  appears as the entropy per vertex of a measure, which is the weak limit of the uniform spanning tree measures on  $G_n$ . We refer to Pemantle [85], Lyons [70], and Benjamini, Lyons, Peres and Schramm [8] for further discussion on the limiting uniform measures on spanning forests/trees.

## Acknowledgement

We thank Yuval Peres for a helpful suggestion, which led to our alternative proof of Theorem 1.3. We are grateful to Jeffrey Hunter for informing us of the work [43], and to Abdelmalek Abdeslam for pointing out [1]. We thank Christopher Eur and Madeline Brandt for a discussion on irreducibility of polynomials. We also thank two anonymous referees for their careful reading and valuable suggestions.

## References

- [1] Abdesselam, A. (2004). The Grassmann–Berezin calculus and theorems of the matrix-tree type. *Adv. in Appl. Math.* **33** 51–70. [MR2064357](#)
- [2] Aldous, D. and Fill, J.A. (2002). Markov Chains and random walks on graphs. Available at <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- [3] Aldous, D.J. (1990). The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.* **3** 450–465. [MR1069105](#)
- [4] Anantharam, V. and Tsoucas, P. (1989). A proof of the Markov chain tree theorem. *Statist. Probab. Lett.* **8** 189–192. [MR1017890](#)
- [5] Avena, L. and Gaudilliere, A. (2013). On some random forests with determinantal roots. Preprint. Available at [arXiv:1310.1723](https://arxiv.org/abs/1310.1723).
- [6] Bellman, R. (1997). *Introduction to Matrix Analysis. Classics in Applied Mathematics* **19**. Philadelphia, PA: SIAM. [MR1455129](#)
- [7] Benjamin, A.T. and Yerger, C.R. (2006). Combinatorial interpretations of spanning tree identities. *Bull. Inst. Combin. Appl.* **47** 37–42. [MR2225927](#)
- [8] Benjamini, I., Lyons, R., Peres, Y. and Schramm, O. (2001). Uniform spanning forests. *Ann. Probab.* **29** 1–65. [MR1825141](#)
- [9] Benjamini, I. and Schramm, O. (2001). Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6** no. 23, 13 pp. (electronic). [MR1873300](#)
- [10] Biane, P. (2015). Polynomials associated with finite Markov chains. In *In Memoriam Marc Yor – Séminaire de Probabilités XLVII. Lecture Notes in Math.* **2137** 249–262. Cham: Springer. [MR3444302](#)
- [11] Biane, P. and Chapuy, G. (2015). Laplacian matrices and spanning trees of tree graphs. Preprint. Available at [arXiv:1505.04806](https://arxiv.org/abs/1505.04806).
- [12] Biggs, N. (1993). *Algebraic Graph Theory*, 2nd ed. *Cambridge Mathematical Library*. Cambridge: Cambridge Univ. Press. [MR1271140](#)
- [13] Bôcher, M. (1964). *Introduction to Higher Algebra*. New York: Dover Publications. [MR0172882](#)
- [14] Boesch, F.T. and Prodinger, H. (1986). Spanning tree formulas and Chebyshev polynomials. *Graphs Combin.* **2** 191–200. [MR0951563](#)
- [15] Bogner, C. and Weinzierl, S. (2010). Feynman graph polynomials. *Internat. J. Modern Phys. A* **25** 2585–2618. [MR2651267](#)
- [16] Bott, R. and Mayberry, J.P. (1954). Matrices and trees. In *Economic Activity Analysis* 391–400. New York: Wiley. [MR0067067](#)
- [17] Broder, A. (1989). Generating random spanning trees. In *Proceedings of the 30th Annual Symposium on Foundations of Computer Science (SFCS'89)* 442–447. Washington, DC: IEEE Computer Society.
- [18] Brooks, R.L., Smith, C.A.B., Stone, A.H. and Tutte, W.T. (1940). The dissection of rectangles into squares. *Duke Math. J.* **7** 312–340. [MR0003040](#)
- [19] Burton, R. and Pemantle, R. (1993). Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *Ann. Probab.* **21** 1329–1371. [MR1235419](#)
- [20] Catral, M., Kirkland, S.J., Neumann, M. and Sze, N.-S. (2010). The Kemeny constant for finite homogeneous ergodic Markov chains. *J. Sci. Comput.* **45** 151–166. [MR2679794](#)
- [21] Cayley, A. (1889). A theorem on trees. *Pure Appl. Math. Q.* **23** 376–378.
- [22] Chaiken, S. (1982). A combinatorial proof of the all minors matrix tree theorem. *SIAM J. Algebr. Discrete Methods* **3** 319–329. [MR0666857](#)
- [23] Chaiken, S. and Kleitman, D.J. (1978). Matrix tree theorems. *J. Combin. Theory Ser. A* **24** 377–381. [MR0480115](#)

- [24] Chang, Y. and Jan, Y.L. (2014). Markov loops in discrete spaces. Preprint. Available at [arXiv:1402.1064](https://arxiv.org/abs/1402.1064).
- [25] Chebotarev, P. (2007). A graph theoretic interpretation of the mean first passage times. Preprint. Available at [arXiv:Math/0701359](https://arxiv.org/abs/math/0701359).
- [26] Chebotarev, P. and Agaev, R. (2002). Forest matrices around the Laplacian matrix. *Linear Algebra Appl.* **356** 253–274. [MR1945250](https://doi.org/10.1016/S0024-3795(02)00452-5)
- [27] Chen, W.K. (1976). *Applied Graph Theory*, revised ed. Amsterdam: North-Holland. [MR0472234](https://doi.org/10.1016/0012-3659(76)90022-3)
- [28] Chung, K.L. (1960). *Markov Chains with Stationary Transition Probabilities*. *Die Grundlehren der Mathematischen Wissenschaften* **104**. Berlin: Springer. [MR0116388](https://doi.org/10.1007/978-3-642-11638-8)
- [29] Cori, R. and Le Borgne, Y. (2003). The sand-pile model and Tutte polynomials. *Adv. in Appl. Math.* **30** 44–52. [MR1979782](https://doi.org/10.1016/S0195-5622(03)00025-2)
- [30] de la Harpe, P. (2012). The Fuglede–Kadison determinant, theme and variations. Preprint. Available at [arXiv:1107.1059](https://arxiv.org/abs/1107.1059).
- [31] Dhar, D. (1990). Self-organized critical state of sandpile automaton models. *Phys. Rev. Lett.* **64** 1613–1616. [MR1044086](https://doi.org/10.1103/PhysRevLett.64.1613)
- [32] de Tilière, B. (2014). Principal minors Pfaffian half-tree theorem. *J. Combin. Theory Ser. A* **124** 1–40. [MR3176189](https://doi.org/10.1214/13-AJ1189)
- [33] Doyle, P. (2009). The Kemeny constant of a Markov chain. Available at <https://math.dartmouth.edu/~doyle/docs/kc/kc.pdf>.
- [34] Durrett, R. (2010). *Probability: Theory and Examples*, 4th ed. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge Univ. Press. [MR2722836](https://doi.org/10.1017/C9780521862769)
- [35] Duval, A.M., Klivans, C.J. and Martin, J.L. (2009). Simplicial matrix-tree theorems. *Trans. Amer. Math. Soc.* **361** 6073–6114. [MR2529925](https://doi.org/10.1090/S0002-9947-09-0259925)
- [36] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. Vol. I, 3rd ed. New York: Wiley. [MR0228020](https://doi.org/10.1002/9780471472125)
- [37] Forman, R. (1993). Determinants of Laplacians on graphs. *Topology* **32** 35–46. [MR1204404](https://doi.org/10.1016/0040-914X(93)90037-8)
- [38] Gorodezky, I. and Pak, I. (2014). Generalized loop-erased random walks and approximate reachability. *Random Structures Algorithms* **44** 201–223. [MR3158629](https://doi.org/10.1112/jlms.12029)
- [39] Grimmett, G. (2010). *Probability on Graphs: Random Processes on Graphs and Lattices*. *Institute of Mathematical Statistics Textbooks 1*. Cambridge: Cambridge Univ. Press. [MR2723356](https://doi.org/10.1017/C9780521862769)
- [40] Gursoy, B.B., Kirkland, S., Mason, O. and Sergeev, S. (2013). On the Markov chain tree theorem in the max algebra. *Electron. J. Linear Algebra* **26** 15–27. [MR3028139](https://doi.org/10.1155/2013/261527)
- [41] Gustafson, K. and Hunter, J.J. (2015). Why the Kemeny Time is a constant. Preprint. Available at [arXiv:1510.00456](https://arxiv.org/abs/1510.00456).
- [42] Hunter, J.J. (2014). The role of Kemeny’s constant in properties of Markov chains. *Comm. Statist. Theory Methods* **43** 1309–1321. [MR3196194](https://doi.org/10.1080/03605310.2014.96194)
- [43] Hunter, J.J. (2016). Accurate calculations of stationary distributions and mean first passage times in Markov renewal processes and Markov chains. *Spec. Matrices* **4** 151–175. [MR3475421](https://doi.org/10.1080/21613237.2016.1191421)
- [44] Járai, A.A. (2014). Sandpile models. Preprint. Available at [arXiv:1401.0354](https://arxiv.org/abs/1401.0354).
- [45] Kalai, G. (1983). Enumeration of  $\mathbf{Q}$ -acyclic simplicial complexes. *Israel J. Math.* **45** 337–351. [MR0720308](https://doi.org/10.1080/00377218308839072)
- [46] Kassel, A. (2015). Learning about critical phenomena from scribbles and sandpiles. In *Modélisation Aléatoire et Statistique – Journées MAS 2014*. *ESAIM Proc. Surveys* **51** 60–73. Les Ulis: EDP Sci. [MR3440791](https://doi.org/10.1051/proc/201551006)
- [47] Kassel, A. and Kenyon, R. (2012). Random curves on surfaces induced from the Laplacian determinant. Preprint. Available at [arXiv:1211.6974](https://arxiv.org/abs/1211.6974).
- [48] Kassel, A., Kenyon, R. and Wu, W. (2015). Random two-component spanning forests. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 1457–1464. [MR3414453](https://doi.org/10.1214/13-AIHP653)

- [49] Kassel, A. and Wilson, D.B. (2016). The looping rate and sandpile density of planar graphs. *Amer. Math. Monthly* **123** 19–39. [MR3453533](#)
- [50] Kasteleyn, P.W. (1967). Graph theory and crystal physics. In *Graph Theory and Theoretical Physics* 43–110. London: Academic Press. [MR0253689](#)
- [51] Kelmans, A.K. and Chelnokov, V.M. (1974). A certain polynomial of a graph and graphs with an extremal number of trees. *J. Combin. Theory Ser. B* **16** 197–214. [MR0345867](#)
- [52] Kelner, J.A. and Mądry, A. (2009). Faster generation of random spanning trees. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009)* 13–21. Los Alamitos, CA: IEEE Computer Soc. [MR2648384](#)
- [53] Kemeny, J.G. and Snell, J.L. (1960). *Finite Markov Chains. The University Series in Undergraduate Mathematics*. Princeton, NJ: D. Van Nostrand Co. [MR0115196](#)
- [54] Kenyon, R. (2009). Lectures on dimers. In *Statistical Mechanics. IAS/Park City Math. Ser.* **16** 191–230. Providence, RI: Amer. Math. Soc. [MR2523460](#)
- [55] Kenyon, R. (2011). Spanning forests and the vector bundle Laplacian. *Ann. Probab.* **39** 1983–2017. [MR2884879](#)
- [56] Kenyon, R.W., Propp, J.G. and Wilson, D.B. (2000). Trees and matchings. *Electron. J. Combin.* **7** Research Paper 25, 34 pp. (electronic). [MR1756162](#)
- [57] Kenyon, R.W. and Wilson, D.B. (2015). Spanning trees of graphs on surfaces and the intensity of loop-erased random walk on planar graphs. *J. Amer. Math. Soc.* **28** 985–1030. [MR3369907](#)
- [58] Kirchhoff, G. (1847). Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird. *Ann. Phys. Chem.* **72** 497–508.
- [59] Kirchhoff, G. (1958). On the solution of the equations obtained from the investigation of the linear distribution of galvanic currents (translated by J.B. O'Toole). *IRE Trans. Circuit Theory* **5** 4–7.
- [60] Kohler, H.-H. and Vollmerhaus, E. (1980). The frequency of cyclic processes in biological multistate systems. *J. Math. Biol.* **9** 275–290. [MR0661431](#)
- [61] Kozdron, M.J., Richards, L.M. and Stroock, D.W. (2013). Determinants, their applications to Markov processes, and a random walk proof of Kirchhoff's matrix tree theorem. Preprint. Available at [arXiv:1306.2059](#).
- [62] Lawler, G.F. (1999). Loop-erased random walk. In *Perplexing Problems in Probability. Progress in Probability* **44** 197–217. Boston, MA: Birkhäuser. [MR1703133](#)
- [63] Lawler, G.F. and Limic, V. (2010). *Random Walk: A Modern Introduction. Cambridge Studies in Advanced Mathematics* **123**. Cambridge: Cambridge Univ. Press. [MR2677157](#)
- [64] Leighton, F.T. and Rivest, R.L. (1983). Estimating a probability using finite memory. In *Foundations of Computation Theory (Borgholm, 1983). Lecture Notes in Computer Science* **158** 255–269. Berlin: Springer. [MR0734725](#)
- [65] Leighton, F.T. and Rivest, R.L. (1983). The Markov chain tree theorem. Technical Report, M.I.T Laboratory for Computer Science, MIT/LCS/TM-249.
- [66] Levene, M. and Loizou, G. (2002). Kemeny's constant and the random surfer. *Amer. Math. Monthly* **109** 741–745. [MR1927624](#)
- [67] Levin, D.A., Peres, Y. and Wilmer, E.L. (2009). *Markov Chains and Mixing Times*. Providence, RI: Amer. Math. Soc. [MR2466937](#)
- [68] Liu, C.J. and Chow, Y. (1981). Enumeration of forests in a graph. *Proc. Amer. Math. Soc.* **83** 659–662. [MR0627715](#)
- [69] Lovász, L. and Winkler, P. (1998). Mixing times. In *Microsurveys in Discrete Probability (Princeton, NJ, 1997). DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* **41** 85–133. Providence, RI: Amer. Math. Soc. [MR1630411](#)
- [70] Lyons, R. (1998). A bird's-eye view of uniform spanning trees and forests. In *Microsurveys in Discrete Probability (Princeton, NJ, 1997). DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* **41** 135–162. Providence, RI: Amer. Math. Soc. [MR1630412](#)

- [71] Lyons, R. (2005). Asymptotic enumeration of spanning trees. *Combin. Probab. Comput.* **14** 491–522. [MR2160416](#)
- [72] Lyons, R. (2009). Random complexes and  $l^2$ -Betti numbers. *J. Topol. Anal.* **1** 153–175. [MR2541759](#)
- [73] Lyons, R. (2010). Identities and inequalities for tree entropy. *Combin. Probab. Comput.* **19** 303–313. [MR2593624](#)
- [74] Lyons, R. and Peres, Y. (2016). *Probability on Trees and Networks*. Cambridge: Cambridge Univ. Press. Available at <http://pages.iu.edu/~rdlyons/>.
- [75] Mađry, A., Straszak, D. and Tarnawski, J. (2015). Fast generation of random spanning trees and the effective resistance metric. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms* 2019–2036. Philadelphia, PA: SIAM. [MR3451158](#)
- [76] Majumdar, S.N. and Dhar, D. (1992). Equivalence between the Abelian sandpile model and the  $q \rightarrow 0$  limit of the Potts model. *Physica A: Statistical Mechanics and Its Applications* **185** 129–145.
- [77] Marchal, P. (1999). Cycles hamiltoniens aléatoires et mesures d'occupation invariantes par une action de groupe. *C. R. Acad. Sci. Paris Sér. I Math.* **329** 883–886. [MR1728002](#)
- [78] Marchal, P. (2000). Loop-erased random walks, spanning trees and Hamiltonian cycles. *Electron. Commun. Probab.* **5** 39–50 (electronic). [MR1736723](#)
- [79] Masbaum, G. and Vaintrob, A. (2002). A new matrix-tree theorem. *Int. Math. Res. Not.* **27** 1397–1426. [MR1908476](#)
- [80] Meyer, C. (2000). *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: SIAM. [MR1777382](#)
- [81] Minoux, M. (1997). Bideterminants, arborescences and extension of the matrix-tree theorem to semirings. *Discrete Math.* **171** 191–200. [MR1454450](#)
- [82] Moon, J.W. (1970). *Counting Labelled Trees. From Lectures Delivered to the Twelfth Biennial Seminar of the Canadian Mathematical Congress (Vancouver) 1969*. Montreal, Que.: Canadian Mathematical Congress. [MR0274333](#)
- [83] Orlin, J.B. (1978). Line-digraphs, arborescences, and theorems of Tutte and Knuth. *J. Combin. Theory Ser. B* **25** 187–198. [MR0511991](#)
- [84] Pak, I. and Postnikov, A. (1994). Enumeration of spanning trees of graphs. Available at <http://math.mit.edu/~apost/papers/tree.ps>.
- [85] Pemantle, R. (1991). Choosing a spanning tree for the integer lattice uniformly. *Ann. Probab.* **19** 1559–1574. [MR1127715](#)
- [86] Pitman, J. Enumerations and expectations related to the matrix tree expansion of a determinant. In preparation.
- [87] Pitman, J. (2001/2002). Random mappings, forests, and subsets associated with Abel–Cayley–Hurwitz multinomial expansions. *Sém. Lothar. Combin.* **46** Art. B46h, 45 pp. (electronic). [MR1877634](#)
- [88] Pitman, J. (2002). Forest volume decompositions and Abel–Cayley–Hurwitz multinomial expansions. *J. Combin. Theory Ser. A* **98** 175–191. [MR1897932](#)
- [89] Pitman, J.W. (1977). Occupation measures for Markov chains. *Adv. in Appl. Probab.* **9** 69–86. [MR0433600](#)
- [90] Pokarowski, P. (1999). Directed forests with application to algorithms related to Markov chains. *Appl. Math. (Warsaw)* **26** 395–414. [MR1737178](#)
- [91] Propp, J.G. and Wilson, D.B. (1998). How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree of a directed graph. *J. Algorithms* **27** 170–217. [MR1622393](#)
- [92] Runge, F. and Sachs, H. (1974). Berechnung der Anzahl der Gerüste von Graphen und Hypergraphen mittels deren Spektren. *Math. Balkanica* **4** 529–536. [MR0382045](#)
- [93] Sahi, S. (2014). Harmonic vectors and matrix tree theorems. *J. Comb.* **5** 195–202. [MR3254515](#)

- [94] Seneta, E. (2006). *Non-negative Matrices and Markov Chains. Springer Series in Statistics*. New York: Springer. [MR2209438](#)
- [95] Shubert, B.O. (1975). A flow-graph formula for the stationary distribution of a Markov chain. *IEEE Trans. Syst. Man Cybern. Syst.* **SMC-5** 565–566. [MR0383539](#)
- [96] Sokal, A.D. (2005). The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. In *Surveys in Combinatorics 2005. London Mathematical Society Lecture Note Series* **327** 173–226. Cambridge: Cambridge Univ. Press. [MR2187739](#)
- [97] Solberg, J.J. (1974/1975). A graph theoretic formula for the steady state distribution of finite Markov processes. *Manage. Sci.* **21** 1040–1048. [MR0441301](#)
- [98] Stanley, R.P. (1999). *Enumerative Combinatorics. Vol. 2. Cambridge Studies in Advanced Mathematics* **62**. Cambridge: Cambridge Univ. Press. [MR1676282](#)
- [99] Stroock, D.W. (2014). *An Introduction to Markov Processes*, 2nd ed. *Graduate Texts in Mathematics* **230**. Heidelberg: Springer. [MR3137424](#)
- [100] Temperley, H.N.V. (1964). On the mutual cancellation of cluster integrals in Mayer's fugacity series. *Proc. Phys. Soc.* **83** 3–16. [MR0161676](#)
- [101] Temperley, H.N.V. (1972). The enumeration of graphs on large periodic lattices. In *Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972)* 285–294. Southend-on-Sea: Inst. Math. Appl. [MR0349407](#)
- [102] Temperley, H.N.V. (1974). Enumeration of graphs on a large periodic lattice. In *Combinatorics (Proc. British Combinatorial Conf., Univ. Coll. Wales, Aberystwyth, 1973). London Math. Soc. Lecture Note Ser.* **13** 155–159. London: Cambridge Univ. Press. [MR0347616](#)
- [103] Temperley, H.N.V. and Fisher, M.E. (1961). Dimer problem in statistical mechanics – an exact result. *Philos. Mag.* (8) **6** 1061–1063. [MR0136398](#)
- [104] Tutte, W.T. (1948). The dissection of equilateral triangles into equilateral triangles. *Math. Proc. Cambridge Philos. Soc.* **44** 463–482. [MR0027521](#)
- [105] Tutte, W.T. (2001). *Graph Theory. Encyclopedia of Mathematics and Its Applications* **21**. Cambridge: Cambridge Univ. Press. [MR1813436](#)
- [106] Ventcel, A.D. and Freidlin, M.I. (1970). Small random perturbations of dynamical systems. *Uspekhi Mat. Nauk* **25** 3–55. [MR0267221](#)
- [107] Wilson, D.B. (1996). Generating random spanning trees more quickly than the cover time. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing (Philadelphia, PA, 1996)* 296–303. New York: ACM. [MR1427525](#)
- [108] Wu, F.Y. (2002). Dimers and spanning trees: Some recent results. *Internat. J. Modern Phys. B* **16** 1951–1961. [MR1912561](#)
- [109] Zeilberger, D. (1985). A combinatorial approach to matrix algebra. *Discrete Math.* **56** 61–72. [MR0808086](#)
- [110] Zhang, Y., Yong, X. and Golin, M.J. (2005). Chebyshev polynomials and spanning tree formulas for circulant and related graphs. *Discrete Math.* **298** 334–364. [MR2163455](#)

Received May 2016 and revised September 2016