Combinatorics, Probability and Computing (2009) 18, 423–434.
doi:10.1017/S0963548308009590 Printed in the United Kingdom

On the Difference of Expected Lengths of Minimum Spanning Trees

WENBO V. LI[†] and XINYI ZHANG

Department of Mathematical Sciences, University of Delaware, Newark DE 19716, USA
(e-mail: wli@math.udel.edu, xzhang@math.udel.edu)

Received 12 April 2007; revised 27 September 2008; first published online 30 October 2008

An exact formula for the expected length of the minimum spanning tree of a connected graph, with independent and identical edge distribution, is given, which generalizes Steele's formula in the uniform case. For a complete graph, the difference of expected lengths between exponential distribution, with rate one, and uniform distribution on the interval $(0, 1)$ is shown to be positive and of rate $\zeta(3)/n$. For wheel graphs, precise values of expected lengths are given via calculations of the associated Tutte polynomials.

1. Introduction

For a simple, finite and connected graph $G = (V(G), E(G))$ with vertex set $V(G)$ and edge set $E(G)$, we assign a non-negative independent and identical distributed (i.i.d.) random weight ξ_e with distribution F to each edge $e \in E(G)$ and denote the total length of the minimum spanning tree (MST) of the graph G by

$$L^F_{\mathrm{MST}}(G) = \sum_{e \in E(\mathrm{MST}(G))} \xi_e.$$

If $\{\xi_e, e \in E(G)\}$ follows the uniform distribution on $(0, 1)$ or the exponential distribution with rate 1, then the expectation of the length of the MST of the graph G is denoted by $\mathbb{E}L^u_{\mathrm{MST}}(G)$ or $\mathbb{E}L^e_{\mathrm{MST}}(G)$, respectively.

Frieze [10] first showed that, for a complete graph K_n with n vertices,

$$\lim_{n\to\infty} \mathbb{E}L^e_{\mathrm{MST}}(K_n) = \lim_{n\to\infty} \mathbb{E}L^u_{\mathrm{MST}}(K_n) = \zeta(3) = \sum_{k=1}^{\infty} k^{-3} = 1.202\ldots.$$

The expected lengths of the MSTs have been extensively studied since then. Steele [22] relaxed the restriction on the assumption of the edge distribution and Janson [16] proved the uniform case by a different method and showed a central limit theorem for $L^u_{\mathrm{MST}}(K_n)$.

[†] Supported in part by NSF grant no. DMS-0505805 and DMS-0805929.

The deviation properties of $L^u_{\mathrm{MST}}(K_n)$ were examined in McDiarmid [17] and Flaxman [9]. Various other special graphs have also been studied, including the bipartite graph by Frieze and McDiarmid [11], the cubic graph by Penrose [19], the 'modestly expensive' regular graph by Beveridge, Frieze and McDiarmid [4] and Frieze, Ruszinkó and Thoma [12], and the cylinder graph by Hutson and Lewis [14].

Recently, Steele [23] started the investigation on exact formulae for the expected lengths of MSTs of any simple graph, and discovered the following nice formula:

$$\mathbb{E}L^u_{\mathrm{MST}}(G) = \int_0^1 \frac{(1-t)}{t} \frac{T_x(G;1/t,1/(1-t))}{T(G;1/t,1/(1-t))}\,\mathrm{d}t, \tag{1.1}$$

where $T(G;x,y)$ is the Tutte polynomial of G and $T_x(G;x,y)$ denotes the partial derivative of $T(G;x,y)$ with respect to x. For the complete graph K_n, Fill and Steele [8] found a recursive method to compute the exact values of $\mathbb{E}L^u_{\mathrm{MST}}(K_n)$, and very recently Gamarnik [13] derived an exact formula in terms of the number of connected labelled graphs on n vertices and m edges, which reduces the computation complexity. In this paper, following the idea of Steele [23], we establish the following exact formula, which generalizes Steele's formula in the uniform case.

Theorem 1.1 (General formula). *If G is a simple, finite and connected graph and ξ_e is a positive random variable with distribution $F(x) = P(\xi_e \leqslant x)$, then*

$$\mathbb{E}L^F_{\mathrm{MST}}(G) = \int_0^\infty \frac{1-F(t)}{F(t)} \frac{T_x(G;x,y)}{T(G;x,y)}\,\mathrm{d}t, \tag{1.2}$$

where $x = 1/F(t)$, $y = 1/(1-F(t))$, and $T_x(G;x,y)$ is the partial derivative of the Tutte polynomial $T(G;x,y)$ with respect to x. In particular, for the exponential distribution with rate one, i.e., $F(x) = 1 - e^{-x}$, for $x \in (0,\infty)$,

$$\mathbb{E}L^e_{\mathrm{MST}}(G) = \int_0^1 \frac{1}{t} \frac{T_x(G;1/t,1/(1-t))}{T(G;1/t,1/(1-t))}\,\mathrm{d}t, \tag{1.3}$$

and for any connected graph G,

$$\mathbb{E}L^u_{\mathrm{MST}}(G) < \mathbb{E}L^e_{\mathrm{MST}}(G). \tag{1.4}$$

Note that (1.3) follows from (1.2) by a simple change of variable. Although it is not hard to prove (1.4) directly, a comparison of formulae (1.3) and (1.1) makes this intuitive inequality obvious, since the coefficients of $T(G;x,y)$ and $T_x(G;x,y)$ are non-negative integers. See more details in the next section.

The main result of this paper is the application of the general formula (1.2) to a comparison between $\mathbb{E}L^e_{\mathrm{MST}}(K_n)$ and $\mathbb{E}L^u_{\mathrm{MST}}(K_n)$.

Theorem 1.2. *For a complete graph K_n,*

$$0 < \mathbb{E}L^e_{\mathrm{MST}}(K_n) - \mathbb{E}L^u_{\mathrm{MST}}(K_n) = \frac{\zeta(3)}{n} + O\big(n^{-2}\log^2 n\big).$$

The basic idea of the proof is to rewrite the difference, substituting back into an integral, in terms of the expected number of components of a random graph associated with the uniform case, and then modify the arguments showing $\mathbb{E}L^u_{\mathrm{MST}}(K_n) \to \zeta(3)$ in Janson [16].

The rest of the paper is organized as follows. We prove Theorem 1.1 in Section 2 following the idea of Steele [23] in the uniform case. Basic properties of the Tutte polynomial and its connections with the rank and component functions are reviewed and analysed. They are also used in Section 4 for the derivation of the Tutte polynomials of wheel graphs. In Section 3 we prove Theorem 1.2 by decomposing the number of components of a (random) graph into the numbers of tree, unicyclic and complex components of various orders. Several technical estimates follow similar ones used by Janson [16] in the proof of the central limit theorem for $L^u_{\mathrm{MST}}(K_n)$. In the crucial case of the complex component, direct and somewhat atypical estimates are used for random graph with large edge probability. Finally, as an additional example, we examine wheel graphs in Section 4. Their Tutte polynomials are found explicitly via recursive relations, which are of independent interest. For wheel graphs, the difference of expected lengths between exponential distribution with rate one and uniform distribution on $(0, 1)$ is of rate proportional to the number of vertices, and their individual expected lengths are monotone increasing.

2. The Tutte polynomial and the Proof of Theorem 1.1

It is well known that the Tutte polynomial contains much information about the graph. Given a graph $G = (V(G), E(G))$ (not necessarily simple), one can derive its Tutte polynomial $T(G; x, y)$ by the following simple rules based on the graph structure.

(1) If G has no edges, then $T(G; x, y) = 1$.
(2) If e is an edge of G that is neither a loop nor an isthmus, then

$$T(G; x, y) = T(G'_e; x, y) + T(G''_e; x, y) \quad \text{(deletion–contraction equation)},$$

where G'_e is the graph G with the edge e *deleted* and G''_e is the graph G with the edge e *contracted*.
(3) If e is an isthmus, then $T(G; x, y) = xT(G'_e; x, y)$.
(4) If e is a loop, then $T(G; x, y) = yT(G''_e; x, y)$.

Applying these rules recursively for a graph, one may derive its Tutte polynomial explicitly. Taking the complete graph K_n, for example, one can find $T(K_2; x, y) = x$, $T(K_3; x, y) = x + x^2 + y$, and with enough patience, $T(K_4; x, y) = x^3 + y^3 + 3x^2 + 4xy + 3y^2 + 2x + 2y$. But as n grows larger, it seems much more difficult to find $T(K_n; x, y)$. In Section 4, we calculate explicitly the Tutte polynomials of wheel graphs, which seem new, based on recursive patterns.

The Tutte polynomial is mostly used by relating it to the rank function, which measures how much the graph is connected. For an edge subset $A \subseteq E(G)$, the rank of A, $r(A)$ is defined by

$$r(A) = |V(G)| - k(A),$$

where $k(A)$ is the number of components of the graph with vertex set $V(G)$ and edge set A. The Tutte polynomial of the graph G is then a two-variable polynomial defined by

$$T(G;x,y) = \sum_{A\subseteq E(G)} (x-1)^{r(E)-r(A)}(y-1)^{|A|-r(A)}. \tag{2.1}$$

Note that along the hyperbola $H_1 = \{(x,y) : (x-1)(y-1) = 1\}$, $T(G;x,y)$ is simplified to

$$T(G;x,y) = x^{|E|}(x-1)^{r(E)-|E|}.$$

For more properties of the Tutte polynomial, one may refer to [6, 25].

Since the Tutte polynomial is a sum over all the edge subsets of $E(G)$, one can connect it to a probability model and interpret it as an expectation on certain probability space, with a careful choice of x and y. Indeed,

$$T(G;x,y) = \frac{y^N}{(x-1)(y-1)^n} \sum_{A\subseteq E(G)} \left(\frac{y-1}{y}\right)^{|A|} \left(\frac{1}{y}\right)^{N-|A|} ((x-1)(y-1))^{k(A)}, \tag{2.2}$$

where $N = |E(G)|$; see Steele [23] for an excellent discussion.

Next we present a proof of Theorem 1.1 by considering a continuous-time random graph process $G_n(t)$. The edge set of $G_n(t)$ is defined to consist of all edges with weight no more than t, that is, $E(G_n(t)) = \{e : \xi_e \leqslant t\}$. In addition, let $k(t)$ be the number of components of the graph $G_n(t)$ and let $N(t)$ be the number of MST edges selected up to time t, that is,

$$N(t) = \sum_{e\in E(G_n(t))} \mathbb{I}(e \in E(\mathrm{MST}(G))) = \sum_{e\in E(\mathrm{MST}(G))} \mathbb{I}(\xi_e \leqslant t).$$

Then $k(t) = n - N(t)$, since the selection of each MST edge in the random graph process decreases the number of components by 1. Hence we have the following nice representation for the length of MST,

$$\begin{aligned} L^F_{\mathrm{MST}}(G) &= \sum_{e\in E(\mathrm{MST}(G))} \xi_e = \sum_{e\in E(\mathrm{MST}(G))} \int_0^\infty \mathbb{I}(t < \xi_e)\,\mathrm{d}t \\ &= \int_0^\infty \sum_{e\in E(\mathrm{MST}(G))} (1 - \mathbb{I}(\xi_e \leqslant t))\,\mathrm{d}t \\ &= \int_0^\infty (n - 1 - N(t))\,\mathrm{d}t \\ &= \int_0^\infty (k(t) - 1)\,\mathrm{d}t. \end{aligned}$$

Note that the integration limit is only up to $\max_{e\in E(G)} \xi_e$, since $k(t) = 1$ for $t > \max_{e\in E(G)} \xi_e$, if G is connected. Avram and Bertsimas [1] first related the length of the MST to the number of components and in [16, 23], a similar relation was used for graphs with edge weights uniformly distributed on $(0, 1)$.

Next we follow the approach of Steele [23] and connect $k(t)$ to the Tutte polynomial of the graph G. In Section 3, we need to go the opposite way. Note that the moment generating function of $k(t)$ is

$$\phi(s) = \mathbb{E}\exp(sk(t)) = \sum_{A\subseteq E(G)} (F(t))^{|A|} \cdot (1-F(t))^{N-|A|} \cdot e^{sk(A)}.$$

Hence, using (2.2), we can rewrite $\phi(s)$ in terms of the Tutte polynomial as

$$\phi(s) = e^s(1-F(t))^{N+1-n}(F(t))^{n-1} \cdot T\left(G; 1+\frac{e^s(1-F(t))}{F(t)}, \frac{1}{1-F(t)}\right).$$

Taking the derivative with respect to s, we have

$$\phi'(s) = \phi(s)\left\{1+\frac{e^s(1-F(t))}{F(t)}\frac{T_x(G;x_s,y)}{T(G;x_s,y)}\right\},$$

where $y = 1/(1-F(t))$ and $x_s = 1+e^s(1-F(t))/F(t)$. Setting $s=0$, we obtain $x = 1/F(t)$, and

$$\mathbb{E}k(t) = 1+\frac{1-F(t)}{F(t)}\frac{T_x(G;x,y)}{T(G;x,y)}, \tag{2.3}$$

which finishes the proof of the general formula (1.2). To derive formula (1.3), we have for $F(t) = 1-e^{-t}$,

$$\mathbb{E}L^e_{\mathrm{MST}}(G) = \int_0^\infty \frac{1}{e^t-1}\frac{T_x(G;e^t/(e^t-1),e^t)}{T(G;e^t/(e^t-1),e^t)}\,\mathrm{d}t.$$

A simple change of variable ends the proof of Theorem 1.1. □

3. Proof of Theorem 1.2

From (1.3) and Steele's formula (1.1), we easily see that

$$\mathbb{E}L^e_{\mathrm{MST}}(G_n) - \mathbb{E}L^u_{\mathrm{MST}}(G_n) = \int_0^1 f(t)\,\mathrm{d}t > 0,$$

where

$$f(t) = \frac{T_x(G;1/t,1/(1-t))}{T(G;1/t,1/(1-t))} = \sum_{A\subseteq E(G)} (k_A-1)t^{|A|+1}(1-t)^{N-|A|-1}, \tag{3.1}$$

by the definition of the Tutte polynomial in (2.1).

For the complete graph K_n, a key observation is that we can substitute the difference above into an integral related to the expected number of components of random graph $G_n(t)$ associated with the uniform case, $0 \leqslant t \leqslant 1$. To be more precise, by (2.3), we have

$$Q_n := \mathbb{E}L^e_{\mathrm{MST}}(K_n) - \mathbb{E}L^u_{\mathrm{MST}}(K_n) = \int_0^1 \frac{t}{1-t}(\mathbb{E}k(t)-1))\,\mathrm{d}t.$$

Note that as $n\to\infty$,

$$\mathbb{E}L^u_{\mathrm{MST}}(K_n) = \int_0^1 \mathbb{E}(k(t)-1)\,\mathrm{d}t \to \zeta(3). \tag{3.2}$$

In the following, we modify the arguments showing (3.2) in Janson [16], and prove that $Q_n = \zeta(3)n^{-1} + O(n^{-2}\log^2 n)$. The starting point is a decomposition of the number of components of a random graph $G_n(t)$ as follows:

$$k(t) = \sum_{k=1}^{n} X_{kn}(t) + \sum_{k=3}^{n} Y_{kn}(t) + Z_n(t),$$

where $X_{kn}(t)$ is the number of tree component of order k in $G_n(t)$, $Y_{kn}(t)$ is the number of unicyclic components of order k, and $Z_n(t)$ is the number of components with more than one cycle (complex components). Theorem 1.2 then follows from the next three lemmas, which treat the expected numbers of tree, unicyclic and complex components separately. As similar estimates appeared in Janson [16], we omit details and only outline the differences. In the crucial case of the complex component (see Lemma 3.3 below), full details are given, since direct and somewhat atypical estimates are used for random graphs with large edge probability.

For the rest of this section, we use C to denote a positive absolute constant whose value is not important and may change from line to line, and $(n)_k = n(n-1)\cdots(n-k+1)$ to denote the decreasing factorial.

Lemma 3.1 (Asymptotic for the tree components).

$$\sum_{k=1}^{n} \int_0^1 \frac{t}{1-t} \mathbb{E}X_{kn}(t)\,\mathrm{d}t = \frac{\zeta(3)}{n} + O(n^{-2}).$$

Proof. By Cayley's theorem [7], there are k^{k-2} spanning trees on k vertices with $k-1$ edges. In order to appear as a component in the graph $G_n(t)$, $\binom{k}{2} - k + k(n-k) + 1$ edges can not be selected. Thus, the expectation of the number of tree components in $G_n(t)$ can be expressed as

$$\mathbb{E}X_{kn}(t) = \binom{n}{k} k^{k-2} t^{k-1} (1-t)^{nk - k^2/2 - 3k/2 + 1}.$$

Therefore,

$$U_{kn} := \int_0^1 \frac{t}{1-t} \mathbb{E}X_{kn}(t)\,\mathrm{d}t = k^{k-2} \frac{(n)_k}{(nk - k^2/2 - k/2 + 1)_{k+1}}. \tag{3.3}$$

Now we are left to estimate $\sum_{k=1}^{n} U_{kn}$, where the simple bound

$$(n)_k \leqslant C(n - k/2 - 3/2)^k, \quad \text{for } \ 1 \leqslant k \leqslant n, n \geqslant 4, \tag{3.4}$$

given in Janson [16], is useful and follows from the concavity of the logarithm.

For $1 \leqslant k \leqslant \sqrt{n}$, it follows easily from Stirling's formula that

$$\log(n)_k = k \log n - \frac{k(k-1)}{2n} + O\big(k^3 n^{-2}\big),$$

$$\log\left(nk - \frac{k^2}{2} - \frac{k}{2} + 1\right)_{k+1} = (k+1)\log(nk) - \frac{(k+1)(k^2+2k-2)}{2nk} + O\big(k^3 n^{-2}\big).$$

Thus,

$$U_{kn} = \frac{k^{-3}}{n} + O\big(k^{-2}n^{-2}\big), \quad \text{for } 1 \leqslant k \leqslant \sqrt{n}. \tag{3.5}$$

For $\sqrt{n} < k \leqslant n$, we have by (3.4),

$$U_{kn} \leqslant Ck^{k-2} \frac{(n-k/2-3/2)^k}{(nk-k^2/2-3k/2)^{k+1}} \leqslant C\frac{k^{-3}}{n}. \tag{3.6}$$

Combining (3.5) and (3.6), we obtain

$$\begin{aligned}\left|\sum_{k=1}^{n} U_{kn} - \frac{1}{n}\sum_{k=1}^{\infty} k^{-3}\right| &\leqslant \sum_{1\leqslant k\leqslant\sqrt{n}} \left|U_{kn} - \frac{k^{-3}}{n}\right| + C\sum_{k>\sqrt{n}} \frac{k^{-3}}{n} \\ &\leqslant C\sum_{k=1}^{\infty} \frac{k^{-2}}{n^2} + C\sum_{k>\sqrt{n}} \frac{k^{-3}}{n} = O(n^{-2}),\end{aligned}$$

which finishes the proof of Lemma 3.1. □

Lemma 3.2 (Upper bound for the unicyclic components).

$$\sum_{k=3}^{n} \int_0^1 \frac{t}{1-t} \mathbb{E}Y_{kn}(t)\,\mathrm{d}t = O(n^{-2}).$$

Proof. Let $m(k)$ be the number of connected labelled graphs with k edges on k vertices. Then

$$\mathbb{E}Y_{kn}(t) = \binom{n}{k} m(k) t^k (1-t)^{nk-k^2+\binom{k}{2}-k}.$$

Using (3.4) and the bound $m(k) \leqslant Ck^{k-1/2}$ given in Bollobás [5, Theorem 18], we have

$$\begin{aligned}\int_0^1 \frac{t}{1-t}\mathbb{E}Y_{kn}(t)\,\mathrm{d}t &= m(k)(k+1)\frac{(n)_k}{(nk-k^2/2-k/2+1)_{k+2}} \\ &\leqslant Ck^{k-\frac{1}{2}}(k+1)\frac{(n-k/2-3/2)^k}{(nk-k^2/2-3k/2)^{k+2}} \\ &\leqslant C\frac{k^{-3/2}}{n^2},\end{aligned}$$

which finishes the proof of Lemma 3.2. □

Lemma 3.3 (Upper bound for the complex components).

$$\int_0^1 \frac{t}{1-t}\mathbb{E}(Z_n(t)-1)\,\mathrm{d}t = O\big(n^{-2}\log^2 n\big).$$

Proof. We first consider the integral over $[0,\ 3\log n/n]$. Let $\tilde{Z}_n$ be the number of times during the evolution of $G_n(t)$ that a new complex component is formed; then $Z_n(t) < \tilde{Z}_n$

for every t. Hence, using $\mathbb{E}\tilde{Z}_n = O(1)$ from Janson [15], we obtain

$$\begin{aligned}\int_0^{3\log n/n} \frac{t}{1-t}\mathbb{E}(Z_n(t)-1)\,\mathrm{d}t &\leqslant \int_0^{3\log n/n} \frac{t}{1-t}\mathbb{E}(\tilde{Z}_n+1)\,\mathrm{d}t \\ &\leqslant C\int_0^{3\log n/n} t\,\mathrm{d}t = O\big(n^{-2}\log^2 n\big).\end{aligned}$$

For the integral over $[3\log n/n,\ 1]$, we have

$$\mathbb{E}(Z_n(t)-1) \leqslant nP(Z_n(t)>1) \leqslant nP(G_n(t) \text{ is not connected}),$$

since $|Z_n(t)-1| \leqslant n$. To estimate the probability, we use the union bound

$$P(G_n(t) \text{ is not connected}) \leqslant \sum_{s=1}^{n/2}\binom{n}{s}(1-t)^{s(n-s)},$$

since $G_n(t)$ must contain a subset of vertices of size less than $n/2$, which does not connect to any of the remaining vertices. Therefore,

$$\begin{aligned}\int_{3\log n/n}^1 \frac{t}{1-t}\mathbb{E}(Z_n(t)-1)\,\mathrm{d}t &\leqslant n\sum_{s=1}^{n/2}\binom{n}{s}\int_{3\log n/n}^1 t(1-t)^{s(n-s)-1}\,\mathrm{d}t \\ &\leqslant n\sum_{s=1}^{n/2}\binom{n}{s}\int_{3\log n/n}^1 (1-t)^{s(n-s)-1}\,\mathrm{d}t \\ &= n\sum_{s=1}^{n/2}\binom{n}{s}\frac{1}{s(n-s)}\left(1-\frac{3\log n}{n}\right)^{s(n-s)} \\ &\leqslant n\sum_{s=1}^{n/2}\left(\frac{e}{s}\right)^s\frac{1}{s(n-s)}n^{-\frac{s(2n-3s)}{n}} \qquad (3.7) \\ &= \frac{en}{n-1}n^{\frac{3}{n}-2} + n\sum_{s=2}^{n/2}\left(\frac{e}{s}\right)^s\frac{1}{s(n-s)}n^{-\frac{s(2n-3s)}{n}},\end{aligned}$$

where the last inequality (3.7) follows from standard inequalities

$$\binom{n}{s} \leqslant \left(\frac{ne}{s}\right)^s \quad \text{and} \quad 1-x \leqslant e^{-x}.$$

For $2 \leqslant s \leqslant n/2$, we have $-s(2n-3s) \leqslant -2(2n-6)$. Therefore, the whole sum is majorized by $n/2$ times the first term. This shows that

$$\int_{3\log n/n}^1 \frac{t}{1-t}\mathbb{E}(Z_n(t)-1)\,\mathrm{d}t \leqslant \frac{en}{n-1}n^{\frac{3}{n}-2} + C\frac{n^2}{n-2}n^{-4+\frac{12}{n}} = O(n^{-2}),$$

which finishes the proof of Lemma 3.3. □

4. The wheel graph and its Tutte polynomial

The wheel graph $W_n = K_1 + C_n$ is defined as a join of K_1 and C_n, where K_1 is the (trivial) complete graph with 1 vertex and C_n is the cycle with n vertices. The wheel graph is an important class of planar graphs in both theory and applications, with nice properties such as self-duality; see [21, 24]. Benedict used wheel graph theory to study self-dual electric networks [2]. Note that a wheel graph W_n has $n+1$ vertices and $2n$ edges. The number of spanning trees of W_n is given by $L_{2n}-2$, where L_n is the Lucas number defined by the recursive relation

$$L_n = L_{n-1} + L_{n-2}, \quad \text{and} \quad L_1 = 1,\ L_2 = 3.$$

See [3, 18, 20] for more details about W_n. Our main result in this section is the following explicit expression for the Tutte polynomial of W_n, which seems new as far as we know. The basic idea is to use the rules that define the Tutte polynomial to obtain recursive relations among the Tutte polynomials of W_n and some of its subgraphs, and then apply the generating function techniques.

Theorem 4.1. *For the wheel graph W_n, $n \geqslant 3$, the Tutte polynomial is given by*

$$T(W_n; x, y) = -(x+y-xy+1) + \alpha^n + \beta^n,$$

where $\alpha, \beta = \frac{1}{2}(1+x+y \pm \sqrt{(1+x+y)^2 - 4xy})$. In particular,

$$T(W_n; 1/t, 1/(1-t)) = \frac{1}{t^n(1-t)^n}, \quad 0 < t < 1.$$

As an application, the expected lengths for exponential distribution with rate one, and the uniform distribution on $(0,1)$, can be found by an integration, using (1.3) and Steele's formula (1.1).

Corollary 4.2. *For $n \geqslant 3$,*

$$\mathbb{E}L^u_{\mathrm{MST}}(W_n) = \frac{(n!)^2}{(2n+1)!} + n \cdot \int_0^1 \frac{t^{n+1}(1-t)^{n+1}}{1-t(1-t)}\,\mathrm{d}t + \left(\frac{3}{2} - \frac{2\sqrt{3}}{9}\pi\right)n,$$

and

$$\mathbb{E}L^e_{\mathrm{MST}}(W_n) = \frac{n!(n-1)!}{(2n)!} + n \cdot \int_0^1 \frac{t^{n+1}(1-t)^{n}}{1-t(1-t)}\,\mathrm{d}t + \left(1 - \frac{\sqrt{3}}{9}\pi\right)n.$$

In particular, both $\mathbb{E}L^u_{\mathrm{MST}}(W_n)$ and $\mathbb{E}L^e_{\mathrm{MST}}(W_n)$ are monotone increasing with respect to n, and as $n \to \infty$

$$\frac{1}{n}\left(\mathbb{E}L^e_{\mathrm{MST}}(W_n) - \mathbb{E}L^u_{\mathrm{MST}}(W_n)\right) \to \frac{\sqrt{3}}{9}\pi - \frac{1}{2}.$$

Proof of Theorem 4.1. Using a systematic application of the deletion–contraction property of the Tutte polynomial, we find a recursive relation between the wheel graph W_{n+1} and some of the intermediate graphs obtained from deletion and contraction operations.

To define these intermediate graphs, we first label all the vertices of $W_n = K_1 + C_n$. The single vertex of K_1 is labelled with v_0. Starting with any vertex of C_n, we label it v_1. Then, in the clockwise direction, we label the other vertices of C_n as $v_2, v_3, \ldots, v_n$ sequentially.

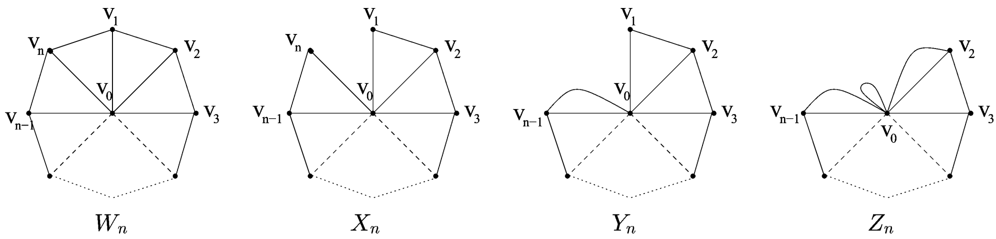

Figure 1. Wheel graph W_n and intermediate graphs X_n, Y_n, Z_n

Define X_n as the subgraph of W_n obtained by removing the edge (v_n, v_1) from the wheel W_n. Y_n is obtained by contracting the edge (v_0, v_n) in X_n, and Z_n is obtained by contracting the edge (v_0, v_1) in Y_n and attaching a loop at vertex v_0. The definition is illustrated in Figure 1.

With a slight abuse of notation, we let W_n, X_n, Y_n and Z_n represent the Tutte polynomial of the graphs W_n, X_n, Y_n and Z_n, respectively. From the rules in Section 2 that define the Tutte polynomial and the special structure of wheel graphs, we have the following recursive relations, for $n \geqslant 3$,

$$\begin{aligned} W_{n+1} &= (x+1)X_n + 2yY_n + yZ_n + W_n, \\ X_{n+1} &= (x+1)X_n + yY_n, \\ Y_{n+1} &= X_n + yY_n, \\ Z_{n+1} &= yY_n + yZ_n. \end{aligned}$$

If associated generating functions are formed as

$$F(t) = \sum_{n \geqslant 3} X_n t^n, \quad G(t) = \sum_{n \geqslant 3} Y_n t^n, \quad P(t) = \sum_{n \geqslant 3} Z_n t^n, \quad Q(t) = \sum_{n \geqslant 3} W_n t^n,$$

then by basic algebra manipulation we obtain four interrelated equations,

$$W_3 t^3 = (1-t)Q(t) - ytP(t) - 2ytG(t) - (x+1)tF(t), \tag{4.1}$$
$$X_3 t^3 = (1 - xt - t)F(t) - ytG(t), \tag{4.2}$$
$$Y_3 t^3 = -tF(t) + (1 - yt)G(t), \tag{4.3}$$
$$Z_3 t^3 = (1 - yt)P(t) - ytG(t), \tag{4.4}$$

where W_3, X_3, Y_3 and Z_3 are the Tutte polynomials for small graphs, which can be calculated directly as follows:

$$\begin{aligned} W_3 &= x^3 + y^3 + 3x^2 + 4xy + 3y^2 + 2x + 2y, \\ X_3 &= x^3 + 2x^2 + 2xy + x + y + y^2, \\ Y_3 &= x^2 + x + y + xy + y^2, \\ Z_3 &= y(x + y + y^2). \end{aligned}$$

By solving equations (4.1)–(4.4) with the initial conditions above, we find

$$Q(t) = -(x^2 + y^2 + xy + x + y)t^2 - xyt - (1 + xy - x - y) \\ - \frac{1 + x + y - xy}{1 - t} + \frac{2 - t(1 + x + y)}{1 - (t(1 + x + y) - xyt^2)}.$$

To obtain the coefficient of t^n in $Q(t)$, we factorize the denominator of the second fraction as

$$\frac{1}{1 - t(1 + x + y) + xyt^2} = \frac{a}{1 - \alpha t} + \frac{b}{1 - \beta t},$$

where

$$\alpha, \beta = \frac{1}{2}\left(1 + x + y \pm \sqrt{(1 + x + y)^2 - 4xy}\right), \quad a = \frac{\alpha}{\alpha - \beta}, \quad b = \frac{-\beta}{\alpha - \beta}.$$

Finally, one can expand the two fractions by power series to obtain

$$Q(t) = -(x^2 + y^2 + xy + x + y)t^2 - xyt - (1 + xy - x - y) \\ - (1 + x + y - xy)\sum_{n=0}^{\infty} t^n + (2 - t(1 + x + y))\sum_{n=0}^{\infty}(a\alpha^n + b\beta^n)t^n.$$

The Tutte polynomial of W_n is then the coefficient of t^n in $Q(t)$, *i.e.*,

$$\begin{aligned} T(W_n; x, y) &= -(1 + x + y - xy) \\ &\quad + \frac{\alpha^n}{\alpha - \beta}(2\alpha - (1 + x + y)) - \frac{\beta^n}{\alpha - \beta}(2\beta - (1 + x + y)) \\ &= -(x + y - xy + 1) + \alpha^n + \beta^n. \end{aligned}$$

This finishes the proof of Theorem 4.1. Corollary 4.2 is easy to verify and details are omitted. □

Acknowledgement

We thank the anonymous referees for their valuable feedback on the paper.

References

[1] Avram, F. and Bertsimas, D. (1992) The minimum spanning tree constant in geometric probability and under the independent model: A unified approach. *Ann. Appl. Probab.* **2** 113–130.

[2] Benedict, C. P. (1969) On self-dualisms in graphs and networks. PhD dissertation, University of Waterloo, Waterloo, Ontario, Canada.

[3] Benjamin, A. T. and Yerger, C. R. (2006) Combinatorial interpretations of spanning tree identities. *Bull. Inst. Combin. Appl.* **47** 37–42.

[4] Beveridge, A., Frieze, A. and McDiarmid, C. (1998) Minimum length spanning trees in regular graphs. *Combinatorica* **18** 311–333.

[5] Bollobás, B. (2001) *Random Graphs*, 2nd edn, Cambridge University Press.

[6] Bollobás, B. (2002) *Modern Graph Theory*, Springer.

[7] Cayley, A. (1889) A theorem on trees. *Quart. J. Pure Appl. Math.* **23** 376–378.

[8] Fill, J. A. and Steele, J. M. (2005) Exact expectations of minimal spanning trees for graphs with random edge weights. *Stein's method and applications, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, Singapore University Press, Singapore, **5** 169–180.
[9] Flaxman, A. D. (2007) The lower tail of the random minimum spanning tree. *Electron. J. Combin.* **14** #3.
[10] Frieze, A. M. (1985) On the value of a random minimum spanning tree problem. *Discrete Appl. Math.* **10** 47–56.
[11] Frieze, A. M. and McDiarmid, C. J. H. (1989) On random minimum length spanning trees. *Combinatorica* **9** 363–374.
[12] Frieze, A. M., Ruszinkó, M. and Thoma, L. (2000) A note on random minimum length spanning trees. *Electron. J. Combin.* **7** #5.
[13] Gamarnik, D. (2005) The expected value of random minimal length spanning tree of a complete graph. In *Proc. 16th Annual ACM–SIAM Symposium on Discrete Algorithms*, pp. 700–704.
[14] Hutson, K. and Lewis, T. M. (2007) The expected length of a minimum spanning tree of a cylinder graph. *Combin. Probab. Comput.* **16** 63–83.
[15] Janson, S. (1993) Multicyclic components in a random graph process. *Random Struct. Alg.* **4** 71–84.
[16] Janson, S. (1995) The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph. *Random Struct. Alg.* **7** 337–355.
[17] McDiarmid, C. (1989) On the method of bounded differences. In *Surveys in Combinatorics*, Vol. 141 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, pp. 148–188.
[18] Myers, B. R. (1971) Number of spanning trees in a wheel. *IEEE Trans. Circuit Theory* **CT-18** 280–282.
[19] Penrose, M. (1998) Random minimum spanning tree and percolation on the n-cube. *Random Struct. Alg.* **12** 63–82.
[20] Sedlacek, J. (1970) On the skeletons of a graph or digraph. In *Proc. Calgary International Conference on Combinatorial Structures and their Applications*, Gordon and Breach, pp. 387–391.
[21] Smith, C. A. B. and Tutte, W. T. (1950) A class of self-dual maps. *Canad. J. Math.* **2** 179–196.
[22] Steele, J. M. (1987) On Frieze's $\zeta(3)$ limit for lengths of minimal spanning trees. *Discrete Appl. Math.* **18** 99–103.
[23] Steele, J. M. (2002) Minimum spanning trees for graphs with random edge lengths. In *Mathematics and Computer Science II: Algorithms, Trees, Combinatorics and Probabilities*, Birkhäuser, pp. 223–245.
[24] Tutte, W. T. (1950) Squaring the square. *Canad. J. Math.* **2** 197–209.
[25] Welsh, D. (1999) The Tutte polynomial. *Random Struct. Alg.* **15** 210–228.