The average number of spanning trees in sparse graphs with given degrees*

Catherine Greenhill
School of Mathematics and Statistics
UNSW Australia
Sydney NSW 2052, Australia
c.greenhill@unsw.edu.au

Mikhail Isaev
Research School of Computer Science
Australian National University
Canberra, ACT 2601, Australia
Moscow Institute of Physics and Technology
Dolgoprudny, 141700, Russia
isaev.m.i@gmail.com

Matthew Kwan
Department of Mathematics
ETH Zürich
Zürich, 8092, Switzerland
matthew.kwan@math.ethz.ch

Brendan D. McKay
Research School of Computer Science
Australian National University
Canberra, ACT 2601, Australia
brendan.mckay@anu.edu.au

20 February 2017

Abstract

We give an asymptotic expression for the expected number of spanning trees in a random graph with a given degree sequence $\boldsymbol{d} = (d_1, \ldots, d_n)$, provided that the number of edges is at least $n + \frac{1}{2}d_{\max}^4$, where $d_{\max}$ is the maximum degree. A key part of our argument involves establishing a concentration result for a certain family of functions over random trees with given degrees, using Prüfer codes.

1 Introduction

The number of spanning trees $\tau(G)$ in a graph G (also called the *complexity* of G) is an important graph parameter that has connections to a wide range of topics, including the study of electrical networks, algebraic graph theory, statistical physics and number theory

*Research supported by the Australian Research Council, Discovery Project DP140101519.

(see for example [1, 15, 17, 18]). These connections are largely related to the *matrix tree theorem*, which says that $\tau(G)$ is equal to any cofactor of the Laplacian matrix of G.

There is a large body of existing work concerning the approximate value of $\tau(G)$ for graphs with given degree sequences, and random graphs with given degree sequences, especially in the regular case. Let $\boldsymbol{d} = (d_1, \ldots, d_n)$ be a vector of positive integers with even sum, and let $\Gamma_{\boldsymbol{d}}$ denote the set of all graphs on the vertex set $\{1, 2, \ldots, n\}$ with degree sequence $\boldsymbol{d}$. If every entry of $\boldsymbol{d}$ equals d then we write $\Gamma_{n,d}$ for the set of all d-regular graphs on $\{1, 2, \ldots, n\}$. Let $\mathcal{G}_{\boldsymbol{d}}$ be the random graph with degree sequence $\boldsymbol{d}$, chosen uniformly at random from $\Gamma_{\boldsymbol{d}}$, and let $\mathcal{G}_{n,d}$ be the random d-regular graph on vertex set $\{1, 2, \ldots, n\}$, chosen uniformly at random from $\Gamma_{n,d}$. Unless otherwise stated, all asymptotics in this paper hold as $n \to \infty$, possibly along some infinite subsequence of $\mathbb{N}$.

The number of spanning trees in a graph is strongly controlled by its degree sequence. Let

$$d = \frac{1}{n} \sum_j d_j, \qquad \hat{d} = \biggl(\prod_{j=1}^{n} d_j \biggr)^{1/n}$$

denote the arithmetic and geometric means of the degree sequence $\boldsymbol{d}$. The best uniform upper bound for regular graphs is due to McKay [12], who proved that when $d \geq 3$,

$$\tau(G) = O(1) \biggl(\frac{(d-1)^{d-1}}{(d^2 - 2d)^{d/2-1}} \biggr)^{n} \frac{\log n}{nd \log d}$$

for all $G \in \Gamma_{n,d}$. This was proved sharp within a constant by Chung and Yau [3]. Kostochka [8] proved that

$$\bigl(\hat{d}(1-\varepsilon)\bigr)^n \leq \tau(G) \leq \frac{\hat{d}^n}{n-1}$$

for any connected $G \in \Gamma_{\boldsymbol{d}}$, where $\varepsilon = \varepsilon(\delta) > 0$ tends to zero as $\delta = \min_j d_j \to \infty$. This lower bound extended a result of Alon [1] on $\tau(G)$ in the case of d-regular graphs.

To discuss random graphs, define the random variables

$$\tau_{\boldsymbol{d}} = \tau(\mathcal{G}_{\boldsymbol{d}}) \quad \text{and} \quad \tau_{n,d} = \tau(\mathcal{G}_{n,d}).$$

That is, $\tau_{\boldsymbol{d}}$ is the number of spanning trees in $\mathcal{G}_{\boldsymbol{d}}$, and $\tau_{n,d}$ is the number of spanning trees in $\mathcal{G}_{n,d}$. McKay [11] proved that for fixed d,

$$\tau_{n,d}^{1/n} \to \frac{(d-1)^{d-1}}{(d^2 - 2d)^{d/2-1}}$$

with probability 1. An alternative proof in a much more general framework was given by Lyons in [9, Example 3.16].

McKay [10] gave the expected value $\mathbb{E}\, \tau_{\boldsymbol{d}}$ to within a constant factor, in the case that $d_j = O(1)$ for all j and the average degree is at least $2 + \varepsilon$, for some $\varepsilon > 0$. Specifically, McKay proved that under these conditions, the expected number of spanning trees is

$$\mathbb{E}\,\tau_{\boldsymbol{d}} = \Theta(1)\,\frac{1}{n}\left(\frac{\hat{d}\,(d-1)^{d-1}}{d^{d/2}(d-2)^{d/2-1}}\right)^{n}. \tag{1.1}$$

Greenhill, Kwan and Wind [6] recently found the asymptotic value of this $\Theta(1)$ factor, for random d-regular graphs with $3 \le d = O(1)$. Specifically, they proved that the $\Theta(1)$ in (1.1) is asymptotic to the constant

$$\frac{(d-1)^{1/2}}{(d-2)^{3/2}}\,\exp\left(\frac{6d^2-14d+7}{4(d-1)^2}\right). \tag{1.2}$$

(This is about $e^{3/2}/d$ for large d.) They also gave the asymptotic distribution of the number of spanning trees in a random cubic graph.

In this paper, we obtain an asymptotic expression for $\mathbb{E}\,\tau_{\boldsymbol{d}}$ for a wider range of sparse degree sequences $\boldsymbol{d}$ than in any of the above random graph results.

Our main result is the following.

Theorem 1.1. *Let $\boldsymbol{d} = \boldsymbol{d}(n) = (d_1,\ldots,d_n)$ be a vector of positive integers with even sum, for every n in some infinite subsequence of $\mathbb{N}$. Define*

$$d_{\max} = \max_j d_j, \qquad d = \tfrac{1}{n}\sum_{j=1}^{n} d_j, \qquad \hat{d} = \left(\prod_{j=1}^{n} d_j\right)^{1/n}, \qquad R = \tfrac{1}{n}\sum_{j=1}^{n}(d_j-d)^2$$

and let

$$H_{\boldsymbol{d}} = \frac{(d-1)^{1/2}}{(d-2)^{3/2}\,n}\left(\frac{\hat{d}\,(d-1)^{d-1}}{d^{d/2}\,(d-2)^{d/2-1}}\right)^{n}.$$

Suppose that $d_{\max}^4 \le (d-2)n$. Then the sequence $\boldsymbol{d}$ is graphical for sufficiently large n, and the expected number of spanning trees in $\mathcal{G}_{\boldsymbol{d}}$ is given by

$$\mathbb{E}\,\tau_{\boldsymbol{d}} = H_{\boldsymbol{d}}\,\exp\left(\frac{6d^2-14d+7}{4(d-1)^2} + \frac{R}{2(d-1)^3} + \frac{(2d^2-4d+1)R^2}{4(d-1)^4\,d^2} + O\left(\frac{d_{\max}^4}{(d-2)n} + \eta\right)\right),$$

where

$$\begin{aligned}\eta &= \min\left\{\frac{d_{\max}^4}{(d-2)^2 n},\, \frac{d_{\max}^3\log n}{(d-2)n},\, d_{\max}(d-2)\right\}\\ &= O\left(\frac{d_{\max}^4}{(d-2)n} + \frac{(\log n)^{5/2}}{n^{1/2}}\right).\end{aligned}$$

Some remarks about this result are given below.

- Due to the Erdős-Gallai Theorem, under the conditions of Theorem 1.1 the sequence $\boldsymbol{d}$ is always graphical (without any requirement for n to be large). Since this fact is not required for our asymptotic formula, we omit the proof.

- Since $d_{\max} \geq 1$, the condition $d_{\max}^4 \leq (d-2)n$ implies that $d > 2$.
- Other than the relative error term, the expression given by Theorem 1.1 matches (1.2) in the regular case, showing that the formula obtained in (1.2) for regular graphs with constant $d \geq 3$ also holds for d-regular graphs with slowly growing d (in particular, it holds when $d = o(n^{1/3})$).
- Under our assumptions, the relative error term may not be vanishing, though it is always bounded. Let $m = \frac{1}{2}\sum_{j=1}^{n} d_j$ be the number of edges in any graph in $\Gamma_{\boldsymbol{d}}$. The condition $d_{\max}^4 \leq (d-2)n$ is equivalent to the condition that $m \geq n + \frac{1}{2}d_{\max}^4$, or in other words, that there are at least $\frac{1}{2}d_{\max}^4 + 1$ more edges in any graph in $\Gamma_{\boldsymbol{d}}$ than in a tree on n vertices. For example, when $d_{\max} = 3$, our result holds with a bounded error if the number of edges exceeds $n-1$ by at least 42.

In particular, we have the following corollary when d is close to 2.

Corollary 1.2. *Suppose that $d = 2 + 2x/n$ where $\frac{1}{2}d_{\max}^4 \leq x \leq n^{1/2}$. (This corresponds to graphs with $n + x$ edges.) Then*

$$\mathbb{E}\,\tau_{\boldsymbol{d}} = \frac{1}{n}\left(\frac{e}{2}\right)^x \left(\frac{n}{2x}\right)^{3/2+x} \left(\frac{\hat{d}}{2}\right)^n \exp\left(\frac{(6+R)(2+R)}{16} + \frac{3x^2}{2n} + O\left(\frac{d_{\max}^4}{x} + \frac{x^3}{n^2}\right)\right).$$

Proof. We estimate the various terms in Theorem 1.1. First note that

$$(d-1)^{1/2} = (1 + 2x/n)^{1/2} = e^{O(x/n)}$$

and

$$\frac{1}{(d-2)^{3/2}}\left(\frac{\hat{d}}{(d-2)^{d/2-1}}\right)^n = \hat{d}^n \left(\frac{n}{2x}\right)^{3/2+x}.$$

Next, a series expansion yields

$$\begin{aligned}\log\left(\frac{(d-1)^{d-1}}{d^{d/2}}\right) &= (1+2x/n)\log\left(1 + \frac{2x}{n}\right) - (1+x/n)\log\left(2 + \frac{2x}{n}\right) \\ &= -\log 2 + (1-\log 2)\,\frac{x}{n} + \frac{3}{2}\left(\frac{x}{n}\right)^2 + O\left(\left(\frac{x}{n}\right)^3\right)\end{aligned}$$

so we have

$$\left(\frac{(d-1)^{d-1}}{d^{d/2}}\right)^n = 2^{-n}\left(\frac{e}{2}\right)^x \exp\left(\frac{3x^2}{2n} + O\left(\frac{x^3}{n^2}\right)\right).$$

Then, we can compute

$$\frac{6d^2 - 14d + 7}{4\,(d-1)^2} = \frac{3}{4} + O\left(\frac{x}{n}\right),$$

$$\frac{1}{2\,(d-1)^3} = \frac{1}{2} + O\left(\frac{x}{n}\right),$$
$$\frac{(2d^2-4d+1)}{4\,(d-1)^4\,d^2} = \frac{1}{16} + O\left(\frac{x}{n}\right).$$

So, noting that $R \le d_{\max}^2$, we have

$$\frac{6d^2-14d+7}{4\,(d-1)^2} + \frac{R}{2\,(d-1)^3} + \frac{(2d^2-4d+1)\,R^2}{4\,(d-1)^4\,d^2} = \frac{(6+R)\,(2+R)}{16} + O\left(\frac{d_{\max}^4\,x}{n}\right). \quad (1.3)$$

Finally, the error term from Theorem 1.1 is at most

$$O\left(\frac{d_{\max}^4}{(d-2)n} + (d-2)d_{\max}\right) = O\left(\frac{d_{\max}^4}{x}\right).$$

Since this error term dominates the error from (1.3) under our assumptions, the result follows. □

From Corollary 1.2 we see that when the average degree is close to but above 2, and the geometric mean $\hat{d}$ is strictly greater than 2, then $\mathbb{E}\,\tau_{\boldsymbol{d}}$ tends to infinity. This can be true even for degree sequences where the probability of connectivity tends to zero, even when Corollary 1.2 does not apply. For example, consider the degree sequence $\boldsymbol{d}$ with $n/2$ vertices of degree 5 and $n/2$ vertices of degree 1 (restricted to even n). Here $d = 3$ and $\hat{d} = \sqrt{5} > 2$. From Theorem 1.1 it follows that the expected number of spanning trees in $\mathcal{G}_{\boldsymbol{d}}$ is

$$\Theta(1/n)\left(\frac{80}{27}\right)^{n/2}$$

which tends to infinity as $n \to \infty$. However, the probability that $\mathcal{G}_{\boldsymbol{d}}$ is connected tends to zero. To see this, we work in the configuration model [2]. For ease of notation, write $n = 2t$ and let S be the set of configurations with t cells containing 5 points and t cells containing 1 point. If a configuration in S gives rise to a connected graph then every point in a cell of size 1 is paired with a point from a cell of size 5. There are at most

$$\frac{(5t)_t\,(4t)!}{2^{2t}\,(2t)!} \quad (1.4)$$

such configurations, and the probability that a random configuration in S is simple, conditioned on connectedness, is at most 1. The total number of simple configurations in S is

$$\Theta(1)\,\frac{(6t)!}{2^{3t}\,(3t)!} \quad (1.5)$$

where the $\Theta(1)$ factor is the probability that a random configuration in S is simple: this tends to a constant bounded away from zero, by [13, Theorem 4.6]. Dividing (1.4) by (1.5) gives the upper bound

$$\Theta(1)\left(\frac{5^5}{2^7\,3^3}\right)^{n/2} = o(1)$$

on the probability that a random element of $\mathcal{G}_{\boldsymbol{d}}$ is connected.

The case of dense irregular degree sequences will be treated in a separate paper.

1.1 Outline of our approach

Let $(a)_k$ denote the falling factorial $a(a-1)\cdots(a-k+1)$. We say that a sequence $\boldsymbol{x} = (x_1, \ldots, x_n)$ of positive integers is a *tree degree sequence* if the entries of $\boldsymbol{x}$ sum to $2n-2$. We say that a tree degree sequence $\boldsymbol{x}$ is a *suitable* degree sequence if $1 \le x_j \le d_j$ for all $j \in \{1, 2, \ldots, n\}$. (The intended meaning is that $\boldsymbol{x}$ is suitable as a degree sequence for a spanning tree of a graph with degree sequence $\boldsymbol{d}$.)

For a suitable degree sequence $\boldsymbol{x}$, let $\mathcal{T}_{\boldsymbol{x}}$ be the set of all trees with degree sequence $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\mathcal{T}$ be the set of all trees with vertex set $\{1, 2, \ldots, n\}$. It is well-known that

$$|\mathcal{T}_{\boldsymbol{x}}| = \binom{n-2}{x_1 - 1, \ldots, x_n - 1}. \tag{1.6}$$

(See for example [16, Theorem 3.1].) Let $\tau_{\boldsymbol{d}}(\boldsymbol{x})$ denote the number of spanning trees of $\mathcal{G}_{\boldsymbol{d}}$ with degree sequence $\boldsymbol{x}$, and denote by $P(\boldsymbol{d}, T)$ the probability that the random graph $\mathcal{G}_{\boldsymbol{d}}$ has T as a subgraph, for all $T \in \mathcal{T}$. Then the expected number of spanning trees with degree sequence $\boldsymbol{x}$ in $\mathcal{G}_{\boldsymbol{d}}$ can be written as

$$\mathbb{E}\, \tau_{\boldsymbol{d}}(\boldsymbol{x}) = \sum_{T \in \mathcal{T}_{\boldsymbol{x}}} P(\boldsymbol{d}, T) \tag{1.7}$$

and furthermore, the expected number of spanning trees (of any degree sequence) in $\mathcal{G}_{\boldsymbol{d}}$ is

$$\mathbb{E}\, \tau_{\boldsymbol{d}} = \sum_{\boldsymbol{x}} \mathbb{E}\, \tau_{\boldsymbol{d}}(\boldsymbol{x})$$

where the sum is over all suitable degree sequences $\boldsymbol{x}$.

We will estimate the summand in (1.7) using a theorem by McKay [13, Theorem 4.6], which we will restate below, including some necessary terminology, and with some minor rewording for consistency.

Theorem 1.3. [13, Theorem 4.6] *Let $\boldsymbol{g} = (g_1, \ldots, g_n)$ be a sequence of non-negative integers with even sum $2m$, and let $g_{\max} = \max\{g_1, \ldots, g_n\}$. Let X be a simple graph on the vertex set $\{1, 2, \ldots, n\}$ with degree sequence $\boldsymbol{x} = (x_1, \ldots, x_n)$, where $x_{\max} = \max\{x_1, \ldots, x_n\}$. Suppose that $g_{\max} \ge 1$ and $\hat{\Delta} \le \epsilon_1 m$, where $\epsilon_1 < 2/3$ and $\hat{\Delta} = 2 + g_{\max}(\frac{3}{2} g_{\max} + x_{\max} + 1)$. Define*

$$\lambda = \frac{1}{4m} \sum_{j=1}^{n} (g_i)_2 \quad \text{and} \quad \mu = \frac{1}{2m} \sum_{ij \in X} g_i g_j.$$

Let $N(\boldsymbol{g}, X)$ denote the number of simple graphs with degree sequence $\boldsymbol{g}$ and no edge in common with X. Then

$$N(\boldsymbol{g}, X) = \frac{(2m)!}{m!\,2^m \prod_{j=1}^n g_i!} \exp(-\lambda - \lambda^2 - \mu + O(\hat{\Delta}^2/m))$$

uniformly as $n \to \infty$.

Given a suitable degree sequence $\boldsymbol{x}$ and a tree $T \in \mathcal{T}_{\boldsymbol{x}}$, define the parameters

$$\begin{aligned}
\lambda_0 &= \frac{1}{2dn} \sum_{j=1}^n (d_j)_2, \\
\lambda(\boldsymbol{x}) &= \frac{1}{2(d-2)n+4} \sum_{j=1}^n (d_j - x_j)_2, \\
\mu(T) &= \frac{1}{(d-2)n+2} \sum_{\{i,j\} \in E(T)} (d_i - x_i)(d_j - x_j)
\end{aligned}$$

Using Theorem 1.3, we may prove the following.

Lemma 1.4. *Suppose that $\boldsymbol{x}$ is a suitable degree sequence and let $T \in \mathcal{T}_{\boldsymbol{x}}$. With notation as above, provided $d_{\max}^4 \le (d-2)n$,*

$$P(\boldsymbol{d}, T) = \frac{(dn/2)_{n-1}\, 2^{n-1}}{(dn)_{2n-2}} \prod_{j=1}^n (d_j)_{x_j} \exp\left(\lambda_0 + \lambda_0^2 - \lambda(\boldsymbol{x}) - \lambda(\boldsymbol{x})^2 - \mu(T) + O\left(\frac{d_{\max}^4}{(d-2)n}\right)\right).$$

Proof. There is a bijection between the set of graphs with degree sequence $\boldsymbol{d}$ which contain T, and those with degree sequence $\boldsymbol{d} - \boldsymbol{x}$ which contain no edges of T. Therefore, we can write

$$P(\boldsymbol{d}, T) = \frac{N(\boldsymbol{d} - \boldsymbol{x}, T)}{N(\boldsymbol{d}, \emptyset)} \tag{1.8}$$

and Theorem 1.3 to estimate the numerator and denominator.

First, consider $N(\boldsymbol{d} - \boldsymbol{x}, T)$. Let

$$\hat{\Delta} = 2 + g_{\max}\left(\tfrac{3}{2} g_{\max} + x_{\max} + 1\right)$$

where $g_{\max} = \max_{j=1,\ldots,n}(d_j - x_j)$. We require that $\hat{\Delta}^2 \le \varepsilon_1 m$, where $\varepsilon_1 < \frac{2}{3}$ is a constant and $m = \frac{1}{2}((d-2)n+2)$ is the number of edges in a graph with degree sequence $\boldsymbol{d} - \boldsymbol{x}$. By assumption,

$$m = \tfrac{1}{2}\left((d-2)n + 2\right) \ge 1 + \tfrac{1}{2} d_{\max}^4.$$

Since $g_{\max}, x_{\max} \le d_{\max}$ and $d_{\max} \ge 3$ (which follows since $d > 2$), we have

$$\frac{\hat{\Delta}}{m} \le \frac{2 + d_{\max}(\frac{5}{2} d_{\max} + 1)}{1 + \frac{1}{2} d_{\max}^4} \le \frac{55}{83}$$

which is strictly less than $\frac{2}{3}$. Observe also that $\hat{\Delta} = O(d_{\max}^2)$. Hence Theorem 1.3 applies and says that

$$N(\boldsymbol{d}-\boldsymbol{x}, T) = \frac{((d-2)n+2)!}{((d-2)n/2+1)!2^{(d-2)n/2+1}\prod_{j=1}^n (d_j - x_j)!} \times \exp\Bigl(-\lambda(\boldsymbol{x}) - \lambda(\boldsymbol{x})^2 - \mu(T) + O\Bigl(d_{\max}^4/((d-2)n)\Bigr)\Bigr).$$

Similarly, we obtain

$$N(\boldsymbol{d}, \emptyset) = \frac{(dn)!}{(dn/2)!2^{dn/2}\prod_{j=1}^n d_j!}\exp\Bigl(-\lambda_0 - \lambda_0^2 + O\Bigl(d_{\max}^4/((d-2)n)\Bigr)\Bigr), \tag{1.9}$$

noting that the value of the $\hat{\Delta}$ is smaller than in the previous application of Theorem 1.3, while the parameter m is larger. Substituting these expressions into (1.8) completes the proof. □

Observe that the only term in the argument of the exponential in Lemma 1.4 which depends on the structure of T (rather than just the degree sequence of T) is $\mu(T)$. For any suitable degree sequence $\boldsymbol{x}$ and any tree $T \in \mathcal{T}_{\boldsymbol{x}}$, define

$$f(\boldsymbol{x}) = \lambda_0 + \lambda_0^2 - \lambda(\boldsymbol{x}) - \lambda(\boldsymbol{x})^2 \tag{1.10}$$

and let

$$\beta(\boldsymbol{x}) = \frac{1}{|\mathcal{T}_{\boldsymbol{x}}|}\sum_{T\in\mathcal{T}_{\boldsymbol{x}}} e^{-\mu(T)} \tag{1.11}$$

be the average value of $e^{-\mu(T)}$ over all $T \in \mathcal{T}_{\boldsymbol{x}}$.

Combining (1.6), (1.7) and Lemma 1.4, for any suitable degree sequence $\boldsymbol{x}$ we have

$$\begin{aligned}\mathbb{E}\,\tau_{\boldsymbol{d}}(\boldsymbol{x}) &= e^{O(d_{\max}^4/((d-2)n))}\frac{(dn/2)_{n-1}\,2^{n-1}\prod_{j=1}^n d_j}{(dn)_{2n-2}}\sum_{T\in\mathcal{T}_{\boldsymbol{x}}}\Bigl(\prod_{j=1}^n (d_j-1)_{x_j-1}e^{f(\boldsymbol{x})-\mu(T)}\Bigr)\\ &= e^{O(d_{\max}^4/((d-2)n))}\frac{(dn/2)_{n-1}\,2^{n-1}\,\hat{d}^n}{(dn)_{2n-2}}(n-2)!\Bigl(\prod_{j=1}^n\binom{d_j-1}{x_j-1}\Bigr)e^{f(\boldsymbol{x})}\,\beta(\boldsymbol{x}). \end{aligned}\tag{1.12}$$

Now define

$$\bar{\mu}(\boldsymbol{x}) = \frac{1}{|\mathcal{T}_{\boldsymbol{x}}|}\sum_{T\in\mathcal{T}_{\boldsymbol{x}}}\mu(T), \tag{1.13}$$

the average value of $\mu(T)$ over $\mathcal{T}_{\boldsymbol{x}}$. By proving that $\beta(\boldsymbol{x})$ is close to $e^{-\bar{\mu}(\boldsymbol{x})}$ for each suitable degree sequence $\boldsymbol{x}$, and evaluating $\bar{\mu}(\boldsymbol{x})$, we will establish the following.

Theorem 1.5. *Suppose that the conditions of Theorem 1.1 hold and that $\boldsymbol{x}$ is a suitable degree sequence. Then with η, R and $H_{\boldsymbol{d}}$ as defined as in Theorem 1.1,*

$$\mathbb{E}\,\tau_{\boldsymbol{d}}(\boldsymbol{x}) = H_{\boldsymbol{d}} \binom{(d-1)n}{n-2}^{-1} \left(\prod_{j=1}^{n} \binom{d_j - 1}{x_j - 1} \right) \exp\biggl(\frac{(R+d^2)^2}{4d^2} - \tfrac{1}{4} - \lambda(\boldsymbol{x}) - \lambda(\boldsymbol{x})^2 \\ - \frac{1}{n} \sum_{j=1}^{n} (x_j - 1)(d_j - x_j) + O\left(\frac{d_{\max}^4}{(d-2)n} + \eta \right) \biggr).$$

The structure of the rest of the paper is as follows. In Section 3 we generalise the function μ and prove a concentration result for trees with given degrees. This proof will involve a martingale concentration result of McDiarmid [14] which we discuss in Section 2. The results of Section 3 are applied in Section 4 to prove that the average of $e^{-\mu(T)}$ over $T \in \mathcal{T}_{\boldsymbol{x}}$ is close to $\exp(-\bar{\mu}(\boldsymbol{x}))$, and hence to prove Theorem 1.5, for any suitable degree sequence $\boldsymbol{x}$. Finally, Theorem 1.1 is proved in Section 5.

Before we begin, note that we use the following conventions in our summation notation: $\sum_{i \neq j}$ will always denote a sum over all ordered pairs (i, j) with $i \neq j$ (over some appropriate range which will be clear from the context, usually $i, j = 1, \ldots, n$). On the other hand, if i is fixed and we wish to sum over all $j \neq i$ (for example, over all $j \in \{1, 2, \ldots, n\} \setminus i$) then we will write $\sum_{j : j \neq i}$.

2 Concentration results

Let $\mathcal{P} = (\Omega, \mathcal{F}, \mathbb{P})$ be a finite probability space. A sequence $\mathcal{F}_0, \ldots, \mathcal{F}_n$ of σ-subfields of $\mathcal{F}$ is a *filter* if $\mathcal{F}_0 \subseteq \cdots \subseteq \mathcal{F}_n$. A sequence $Y_0, \ldots, Y_n$ of random variables on $\mathcal{P}$ is a *martingale with respect to* $\mathcal{F}_0, \ldots, \mathcal{F}_n$ if

(i) Y_j is $\mathcal{F}_j$-measurable and has finite expectation, for $j = 0, \ldots, n$;

(ii) $\mathbb{E}(Y_j \mid \mathcal{F}_{j-1}) = Y_{j-1}$ for $j = 1, \ldots, n$.

An important example of a martingale is made by the so-called *Doob martingale process*. Suppose $X_1, X_2, \ldots, X_n$ are random variables on $\mathcal{P}$ and $f(X_1, X_2, \ldots, X_n)$ is a random variable on $\mathcal{P}$ of bounded expectation. Let $\sigma(X_1, \ldots, X_j)$ denote the σ-field generated by the random variables $X_1, \ldots, X_j$. Define the martingale $\{Y_j\}$ with respect to the filter $\{\mathcal{F}_j\}$, where for each j, $\mathcal{F}_j = \sigma(X_1, \ldots, X_j)$ and $Y_j = \mathbb{E}(f(X_1, X_2, \ldots, X_n) \mid \mathcal{F}_j)$. In particular, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $Y_0 = \mathbb{E}\, f(X_1, X_2, \ldots, X_n)$.

In this section we state some concentration results for martingales. See McDiarmid [14] for further background and for any definitions not given here. Following McDiarmid [14], for $j = 1, \dots, n$ we define the *conditional range* of Y_j as

$$\operatorname{ran}(Y_j \mid \mathcal{F}_{j-1}) = \operatorname{ess\,sup}(Y_j \mid \mathcal{F}_{j-1}) + \operatorname{ess\,sup}(-Y_j \mid \mathcal{F}_{j-1}). \tag{2.1}$$

Here "essential supremum" may be replaced by "supremum", as in [14], if the probability distribution is positive over Ω.

Our main tool is the following result from McDiarmid [14]. The tail bound on the probability is given by [14, Theorem 3.14]. The upper estimate on the moment generating function $\mathbb{E}(e^{hY_n})$ is an intermediate step of McDiarmid's proof, see [14, Section 3.5]. The lower bound on K is due to Jensen's inequality.

Theorem 2.1. ([14]) *Suppose that $\mathcal{P} = (\Omega, \mathcal{F}, \mathbb{P})$ is a finite probability space. Let $Y_0, Y_1, \dots, Y_n$ be a martingale on $\mathcal{P}$ with respect to a filter $\mathcal{F}_0, \mathcal{F}_1 \dots, \mathcal{F}_n$, where $\mathcal{F}_0 = \{\emptyset, \Omega\}$, such that*

$$\sum_{j=1}^{n} \left(\operatorname{ran}(Y_j \mid \mathcal{F}_{j-1})\right)^2 \le \hat{r}^2 \quad a.s.$$

for some real $\hat{r}$. Then

$$\mathbb{E}\, e^{Y_n} = e^{Y_0 + K}$$

where $0 \le K \le \frac{1}{8}\hat{r}^2$. Furthermore, for any real $t > 0$,

$$\Pr(|Y_n - Y_0| \ge t) \le 2\, \exp(-2t^2/\hat{r}^2).$$

As a corollary, we obtain a concentration result for functions of sets of a given size.

Corollary 2.2. *Let $\binom{[N]}{r}$ be the set of r-subsets of $\{1, \dots, N\}$ and let $h : \binom{[N]}{r} \to \mathbb{R}$ be given. Let C be a uniformly random element of $\binom{[N]}{r}$. Suppose that there exists $\alpha \ge 0$ such that*

$$|h(A) - h(A')| \le \alpha$$

for any $A, A' \in \binom{[N]}{r}$ with $|A \cap A'| = r - 1$. Then

$$\mathbb{E}\, e^{h(C)} = \exp\left(\mathbb{E}\, h(C) + K\right) \tag{2.2}$$

where K is a real constant such that $0 \le K \le \frac{1}{8}\min\{r, N - r\}\alpha^2$. Furthermore, for any real $t > 0$,

$$\Pr(|h(C) - \mathbb{E}\, h(C)| \ge t) \le 2\exp\left(-\frac{2t^2}{\min\{r, N - r\}\alpha^2}\right).$$

Proof. Let S_N denote the set of permutations of $\{1, \dots, N\}$ and $\tau = (\tau_1, \dots, \tau_N)$ be a uniform random element of S_N. Note that the set $\{\tau_1, \dots, \tau_r\}$ is a uniformly random element of $\binom{[N]}{r}$. Define $\tilde{h} : S_N \to \mathbb{R}$ by $\tilde{h}(\omega) = h(\{\omega_1, \dots, \omega_r\})$ for all $\omega \in S_N$. Then

$$|\tilde{h}(\rho) - \tilde{h}(\rho')| \le \alpha$$

for all permutations ρ, $\rho' \in S_N$ such that $\rho^{-1}\rho'$ is a transposition. Given $\omega = (\omega_1, \ldots, \omega_N) \in S_N$, for $k = 0, \ldots, N$ let

$$\tilde{h}_k(\omega) = \mathbb{E}\left(\tilde{h}(\tau) \mid \tau_j = \omega_j \text{ for } j = 1, \ldots k\right).$$

Clearly, $\tilde{h}_0(\tau), \ldots, \tilde{h}_N(\tau)$ forms a martingale: it is the result of the Doob martingale process for $\tilde{h}(\tau)$. It follows from Frieze and Pittel [5, Lemma 11] that

$$\operatorname{ran}\bigl(\tilde{h}_k(\tau) \mid \sigma(\tau_1, \ldots, \tau_{k-1})\bigr) \le \alpha.$$

Moreover, for any $\omega \in S_N$ and $k \in \{r, \ldots, N\}$, we have

$$\mathbb{E}\bigl(\tilde{h}(\tau) \mid \tau_j = \omega_j,\ 1 \le j \le k\bigr) = h(\{\omega_1, \ldots, \omega_r\}).$$

Therefore $\operatorname{ran}\bigl(\tilde{h}_k(\tau) \mid \sigma(\tau_1, \ldots, \tau_{k-1})\bigr) = 0$ for all $k > r$. Applying Theorem 2.1 to the martingale $\tilde{h}_0(\tau), \ldots, \tilde{h}_N(\tau)$, we conclude that (2.2) holds with $0 \le K \le \frac{1}{8} r\alpha^2$, and that

$$\Pr(|h(C) - \mathbb{E}\, h(C)| \ge t) \le 2 \exp\left(-\frac{2t^2}{r\alpha^2}\right).$$

If $r \le N - r$ then we are finished. Otherwise we repeat the above argument using the bijection between subsets and their complements. ☐

3 Trees with given degrees

In this section we consider sums of the form

$$F(T) = \sum_{\{j,k\} \in E(T)} \phi(j)\, \phi(k) \tag{3.1}$$

for a given function $\phi : \{1, 2, \ldots, n\} \to [\phi_{\min}, \phi_{\max}] \subset \mathbb{R}$.

Let $\bar{F}(\boldsymbol{x})$ be the average value of F over all trees with a given degree sequence $\boldsymbol{x}$:

$$\bar{F}(\boldsymbol{x}) = \frac{1}{|\mathcal{T}_{\boldsymbol{x}}|} \sum_{T \in \mathcal{T}_{\boldsymbol{x}}} F(T).$$

The goal of this section is to prove the following theorem, showing that the average of $e^{\xi F(T)}$ over $\mathcal{T}_{\boldsymbol{x}}$ is close to $e^{\xi \bar{F}(\boldsymbol{x})}$, for $\xi \in \{-1, 1\}$. We will measure this distance using the seminorm of ϕ given by

$$\|\phi\|_m = \min_{c \in \mathbb{R}} \sum_{j=1}^{n} |\phi(j) - c|. \tag{3.2}$$

Here the minimising value of c is any median of $\{\phi(j) : j = 1, \ldots, n\}$.

Theorem 3.1. *Let F satisfy (3.1). Then for any tree degree sequence $\boldsymbol{x}$ and for $\xi \in \{-1,1\}$,*

$$\frac{1}{|\mathcal{T}_{\boldsymbol{x}}|}\sum_{T\in\mathcal{T}_{\boldsymbol{x}}} e^{\xi F(T)} = \exp\bigl(\xi \bar{F}(\boldsymbol{x}) + K\bigr)$$

for some real constant K which satisfies $0 \le K \le \frac{1}{8}L_\phi$, where

$$L_\phi = (\phi_{\max} - \phi_{\min})^3 \, \min\bigl\{(\phi_{\max}-\phi_{\min})n, \ \|\phi\|_m\,(\ln n + 2)\bigr\}.$$

Furthermore, if $\widehat{T}$ is a uniformly random element of $\mathcal{T}_{\boldsymbol{x}}$ then for any real constant $t > 0$,

$$\Pr(|F(\widehat{T}) - \bar{F}(\boldsymbol{x})| > t) \le 2\,\exp\bigl(-2t^2/L_\phi\bigr).$$

First we give some explicit formulae which we will need later.

Lemma 3.2. *Let $\boldsymbol{x}$ be a tree degree sequence and consider the set $\mathcal{T}_{\boldsymbol{x}}$ of all trees with degree sequence $\boldsymbol{x}$.*

(i) *Let S be a disconnected forest with vertex set $\{1,\ldots,n\}$ and degree sequence $(s_1,\ldots,s_n)$, where $s_j \le x_j$ for $j = 1,\ldots,n$. Let $S_1,\ldots,S_r$ be the components of S. Then the probability that a uniform random tree in $\mathcal{T}_{\boldsymbol{x}}$ contains S is*

$$\frac{\prod_{i=1}^r \sum_{j\in V(S_i)}(x_j - s_j)}{(n-2)_{n-r}} \prod_{j=1}^n (x_j - 1)_{s_j-1},$$

where $(x_j-1)_{s_j-1} = x_j^{-1}$ if $s_j = 0$. In particular, for distinct $j,k \in \{1,2,\ldots,n\}$, the fraction of trees in $\mathcal{T}_{\boldsymbol{x}}$ in which vertices j,k are adjacent is

$$\frac{x_j + x_k - 2}{n-2}.$$

(ii) *The average value of F over $\mathcal{T}_{\boldsymbol{x}}$ is*

$$\bar{F}(\boldsymbol{x}) = \frac{1}{n-2}\left(\sum_{k=1}^n \phi(k)\right)\left(\sum_{j=1}^n (x_j-1)\,\phi(j)\right) - \frac{1}{n-2}\left(\sum_{j=1}^n (x_j-1)\,\phi(j)^2\right).$$

Proof. Define $\boldsymbol{x}' = (x_1',\ldots,x_r')$, where $x_i' = \sum_{j\in V(S_i)}(x_j - s_j)$ for $i = 1,\ldots,r$. If $x_i' = 0$ for any i then T cannot contain S, as S is disconnected. Hence the result holds trivially in that case, and for the remainder of the proof we may assume that all entries of $\boldsymbol{x}'$ are positive. Next, observe that the entries of $\boldsymbol{x}'$ sum to $2(r-1)$, and hence $\boldsymbol{x}'$ is a tree degree sequence.

Each tree in $\mathcal{T}_{\boldsymbol{x}}$ that contains S can be formed uniquely by the following process:

(1) Take any tree T' on the vertex set $\{1, \ldots, r\}$ with degree sequence $\boldsymbol{x}'$.

(2) For $i = 1, \ldots, r$, replace vertex i of T' by S_i and distribute the edges of T' that were incident with i amongst the vertices of S_i, so that each vertex $j \in V(S_i)$ has degree x_j in the resulting tree.

By (1.6), the number of choices for T' in Step 1 is $\binom{r-2}{x'_1-1,\ldots,x'_r-1}$, while the number of ways to distribute edges in Step 2 is

$$\prod_{i=1}^{r} \frac{x'_i!}{\prod_{j\in V(S_i)}(x_j - s_j)!}.$$

The first statement of (i) is proven by multiplying these expressions together, dividing by (1.6) and simplifying. Then taking S to be the edge jk together with $n-2$ trivial components completes the proof of (i).

Now using linearity of expectation, (3.1), and part (i), we calculate that

$$\begin{aligned}
(n-2)\,\bar{F}(\boldsymbol{x}) &= \sum_{j<k} (x_j + x_k - 2)\,\phi(j)\phi(k) &(3.3)\\
&= \sum_{j\neq k} (x_j - 1)\,\phi(j)\phi(k) &(3.4)\\
&= \sum_{j=1}^{n} (x_j - 1)\,\phi(j)\,\left(\left(\sum_{k=1}^{n}\phi(k)\right) - \phi(j)\right)\\
&= \left(\left(\sum_{k=1}^{n}\phi(k)\right)\sum_{j=1}^{n}(x_j-1)\phi(j)\right) - \left(\sum_{j=1}^{n}(x_j-1)\,\phi(j)^2\right),
\end{aligned}$$

establishing (ii). □

We complete this section with the proof of Theorem 3.1, which involves the process used to construct the *Prüfer code* of a labelled tree. The Prüfer code of a tree $T \in \mathcal{T}$ is a sequence $\boldsymbol{b} = (b_1, \ldots, b_{n-2}) \in \{1, 2, \ldots, n\}^{n-2}$. Given T, find the unique neighbour b_1 of the lowest-labelled leaf a_1. Then b_1 becomes the first entry in the Prüfer code for T. We find the next entry recursively by considering the tree $T - a_1$ with the first leaf deleted. The process stops when a single edge remains: this edge is determined by the degree sequence and does not need to be recorded in the code $\boldsymbol{b}$. We will refer to this process as the *Prüfer process* with input T. See Figure 1 for an example. The correspondence between trees and Prüfer codes is a bijection: see for example Moon [16, pp. 5-6]. This provides a proof of Cayley's formula and of (1.6).

The following useful property of the Prüfer process may be proved by induction on j.

Lemma 3.3. *Let $\boldsymbol{x}$ be a tree degree sequence and let $T \in \mathcal{T}_{\boldsymbol{x}}$. Suppose that the Prüfer process with input T produces the Prüfer code $\boldsymbol{b}$ and the sequence $(a_1, \ldots, a_{n-2})$ of "leaves".*

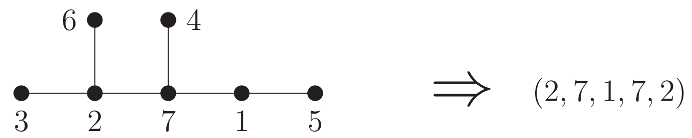

Figure 1: A tree and its corresponding Prüfer code.

For any $j = 1, \ldots, n-2$, the initial sequence $(a_1, \ldots, a_j)$ is uniquely determined by $\boldsymbol{x}$ and $(b_1, \ldots, b_{j-1})$.

When there is more than one tree under consideration we will write $a_j(T)$, $b_j(T)$ for the vertices identified at step j of the Prüfer process for the tree T. To prove Theorem 3.1 we work with a martingale defined using the Prüfer code of a tree. A martingale construction based on the Prüfer code was given by Cooper, McGrae and Zito [4], for all labelled trees. Our martingale is restricted to trees with a given degree sequence and we study a function for which it is more difficult to bound the conditional ranges.

Proof of Theorem 3.1. Suppose that T_1 and T_2 are trees on $\{1, 2, \ldots, n\}$ with the same degree sequence. For $j = 0, \ldots, n-3$, say that T_1 and T_2 are *j-equivalent*, and write $T_1 \overset{j}{\sim} T_2$, if $b_i(T_1) = b_i(T_2)$ for $i = 1, 2, \ldots, j$. By Lemma 3.3, if $T_1 \overset{j}{\sim} T_2$ then $a_i(T_1) = a_i(T_2)$ for $i = 1, \ldots, j$. The j-equivalence relation induces a partition of $\mathcal{T}_{\boldsymbol{x}}$ into equivalence classes $C_{j,1}, \ldots, C_{j,r_j}$, say, with $\cup_{\ell=1}^{r_j} C_{j,\ell} = \mathcal{T}_{\boldsymbol{x}}$.

Let $Y_{j,\ell}$ equal the average of $F(T)$ over $T \in C_{j,\ell}$, and define the function Y_j on $\mathcal{T}_{\boldsymbol{x}}$ by $Y_j(T) = Y_{j,\ell}$ if $T \in C_{j,\ell}$. Finally, define the random variable $Y_j = Y_j(\widehat{T})$, where $\widehat{T}$ is a uniformly random element of $\mathcal{T}_{\boldsymbol{x}}$. Then Y_0 is the constant function which takes the value $\mathbb{E}\, F(\widehat{T})$ everywhere, and $Y_{n-3} = F(\widehat{T})$, since each equivalence class $C_{n-3,\ell}$ is a set of size 1. Observe that $Y_0, \ldots, Y_{n-3}$ is a martingale with respect to the filter $\mathcal{F}_0, \ldots, \mathcal{F}_{n-3}$, where, for each j, $\mathcal{F}_j$ is generated by the sets $C_{j,1}, \ldots, C_{j,r_j}$. In fact, this is the Doob martingale process for the function $F(\widehat{T})$ of the random variables $b_1(\widehat{T}), \ldots, b_{n-3}(\widehat{T})$, which determine $\widehat{T}$ uniquely.

To apply Theorem 2.1 we must calculate a value for $\hat{r}^2$. Suppose that T_1 and T_2 are $(j-1)$-equivalent, where $T_1, T_2 \in \mathcal{T}_{\boldsymbol{x}}$ and $j \in \{1, \ldots, n-3\}$. Then $a_j(T_1) = a_j(T_2)$, again by Lemma 3.3. For ease of notation, write a_i instead of $a_i(T_1)$ (or $a_i(T_2)$) for $i = 1, \ldots, j$, and write b_i instead of $b_i(T_1)$ (or $b_i(T_2)$) for $i = 1, \ldots, j-1$.

For $s = 1, 2$ let T'_s be the tree (with $n - j$ vertices) obtained by deleting the vertices $a_1, \ldots, a_j$ from T_s. Both T'_1 and T'_2 have vertex set

$$V_j = \{1, 2, \ldots, n\} \setminus \{a_1, \ldots, a_j\}.$$

If $b_j(T_1) = b_j(T_2)$ then T'_1 and T'_2 have the same degree sequence, since (in this case) precisely the same edges have been deleted from T_1 and T_2. In this case, $Y_j(T_1) = Y_j(T_2)$.

Otherwise, the degree sequences of T_1', T_2' differ only for the two vertices $b_j(T_1)$ and $b_j(T_2)$. Specifically, vertex $b_j(T_1)$ has degree in T_1' which is equal to its degree in T_2' minus 1, while vertex $b_j(T_2)$ has degree in T_1' which is equal to its degree in T_2' plus 1. Hence T_1' and T_2' have the same degree on all vertices in the set

$$U_j(T_1, T_2) = V_j \setminus \{b_j(T_1),\, b_j(T_2)\}.$$

For $s = 1, 2$, let $\boldsymbol{y}_s$ be the degree sequence of T_s' (on the vertex set V_j) and let $\mathcal{T}_s'$ denote the set of all trees on the vertex set V_j with degree sequence $\boldsymbol{y}_s$. Observe that $\boldsymbol{y}_s$ and $\mathcal{T}_s'$ depend only on $(b_1, \ldots, b_{j-1}, b_j(T_s))$ and $\boldsymbol{x}$. By relabelling the equivalence classes if necessary, we may assume that $T_s \in C_{j,s}$ for $s = 1, 2$. The map $\varphi : C_{j,s} \to \mathcal{T}_s'$ which sends a tree $T \in C_{j,s}$ to $T \setminus \{a_1, \ldots, a_j\}$ is a bijection. To see this, observe that the inverse map φ^{-1} takes a tree in $\mathcal{T}_s'$, adds the vertices $a_1, \ldots, a_j$ and the edges

$$\{\{a_1, b_1\}, \ldots, \{a_{j-1}, b_{j-1}\},\ \{a_j, b_j(T_s)\}\}$$

giving a tree in $C_{j,s}$. Therefore, for $s = 1, 2$,

$$\frac{1}{|C_{j,s}|} \sum_{T \in C_{j,s}} F(T \setminus \{a_1, \ldots, a_j\}) = \frac{1}{|\mathcal{T}_s'|} \sum_{T' \in \mathcal{T}_s'} F(T').$$

Combining this with (3.1) and the definition of $Y_{j,s}$, we see that for $s = 1, 2$,

$$\begin{aligned} Y_{j,s} &= \frac{1}{|C_{j,s}|} \sum_{T \in C_{j,s}} \sum_{\{k,\ell\} \in E(T)} \phi(k)\phi(\ell) \\ &= \left(\sum_{i=1}^{j-1} \phi(a_i)\phi(b_i) \right) + \phi(a_j)\phi(b_j(T_s)) + \frac{1}{|C_{j,s}|} \sum_{T \in C_{j,s}} F(T \setminus \{a_1, \ldots, a_j\}) \\ &= \left(\sum_{i=1}^{j-1} \phi(a_i)\phi(b_i) \right) + \phi(a_j)\phi(b_j(T_s)) + \frac{1}{|\mathcal{T}_s'|} \sum_{T' \in \mathcal{T}_s'} F(T'). \end{aligned}$$

Applying Lemma 3.2(ii) gives

$$\begin{aligned} Y_{j,1} - Y_{j,2} &= (\phi(b_j(T_1)) - \phi(b_j(T_2))) \left(\phi(a_j) - \frac{1}{n-j-2} \sum_{\ell \in U_j(T_1,T_2)} \phi(\ell) \right) \\ &= \frac{\phi(b_j(T_1)) - \phi(b_j(T_2))}{n-j-2} \sum_{\ell \in U_j(T_1,T_2)} \big(\phi(a_j) - \phi(\ell) \big). \end{aligned}$$

(Note that if $T_1 \overset{j}{\sim} T_2$ then $b_j(T_1) = b_j(T_2)$ and the above equality also holds.)

Recall the definition of $\|\phi\|_m$ from (3.2), and let $c \in \mathbb{R}$ be the minimising value in this definition. By the triangle inequality,

$$\frac{1}{n-j-2} \left| \sum_{\ell \in U_j(T_1,T_2)} (\phi(a_j) - \phi(\ell)) \right| \le |\phi(a_j) - c| + \frac{1}{n-j-2} \left| \sum_{\ell \in U_j(T_1,T_2)} (c - \phi(\ell)) \right|$$

$$\leq |\phi(a_j) - c| + \frac{\|\phi\|_m}{n-j-2} \tag{3.5}$$

since $U_j(T_1, T_2)$ has $n-j-2$ elements and $j \leq n-3$. Therefore, for any equivalence class $C_{j-1,\ell}$, we have

$$\begin{aligned}
&\Bigg(\sup_{T' \in C_{j-1,\ell}} Y_j(T') + \sup_{T' \in C_{j-1,\ell}} (-Y_j(T')) \Bigg)^2 \\
&\qquad \leq \frac{(\phi_{\max} - \phi_{\min})^2}{(n-j-2)^2} \sup_{T_1, T_2 \in C_{j-1,\ell}} \Bigg(\sum_{\ell \in U_j(T_1,T_2)} (\phi(a_j) - \phi(\ell)) \Bigg)^2 \\
&\qquad \leq (\phi_{\max} - \phi_{\min})^3 \min\left\{ \phi_{\max} - \phi_{\min},\, |\phi(a_j) - c| + \frac{\|\phi\|_m}{n-j-2} \right\}. \qquad (3.6)
\end{aligned}$$

(Here we take the minimum of two possible upper bounds: the first arises from taking the worst case summand for both factors in the line above, while the second arises by applying (3.5) to one of the factors.)

Now let $C_{j-1}(\widehat{T})$ denote the random set which is the equivalence class with respect to $\overset{j-1}{\sim}$ which contains $\widehat{T}$. It follows from (3.6) that

$$\begin{aligned}
\operatorname{ran}(Y_j \mid \mathcal{F}_{j-1})^2 &= \Bigg(\sup_{T \in C_{j-1}(\widehat{T})} Y_j(T) + \sup_{T \in C_{j-1}(\widehat{T})} (-Y_j(T)) \Bigg)^2 \\
&\leq (\phi_{\max} - \phi_{\min})^3 \min\left\{ \phi_{\max} - \phi_{\min},\, |\phi(a_j(\widehat{T})) - c| + \frac{\|\phi\|_m}{n-j-2} \right\}.
\end{aligned}$$

Using the definition of c, the standard upper bound on the harmonic series and the fact that each vertex is chosen as $a_j(\widehat{T})$ at most once during the Prüfer process, we get that

$$\sum_{j=1}^{n-3} \operatorname{ran}(Y_j \mid \mathcal{F}_{j-1})^2 \leq (\phi_{\max} - \phi_{\min})^3 \min\left\{ (\phi_{\max} - \phi_{\min})n,\ \|\phi\|_m \,(\ln n + 2) \right\}.$$

Observe that the left hand side does not change if F is replaced by $-F$ (and hence, the same bound is obtained whether $\xi = 1$ or $\xi = -1$). Since $\mathbb{E}(e^{Y_{n-3}}) = \mathbb{E}(e^{F(\widehat{T})})$ and $Y_0 = \mathbb{E}(F(\widehat{T}))$, applying Theorem 2.1 completes the proof. ☐

4 Proof of Theorem 1.5

First we note the following corollary of Theorem 3.1. Recall the definition of $\beta(\boldsymbol{x})$ and $\bar{\mu}(\boldsymbol{x})$ from (1.11), (1.13), respectively.

Lemma 4.1. *Under the conditions of Theorem 1.1,*

$$\beta(\boldsymbol{x}) = \exp\left(-\bar{\mu}(\boldsymbol{x}) + O\left(\min\left\{ \frac{d_{\max}^4}{(d-2)^2 n}, \frac{d_{\max}^3 \ln n}{(d-2)n} \right\} \right) \right).$$

Proof. Set

$$\phi(j) = \frac{d_j - x_j}{\sqrt{(d-2)n+2}}$$

for $j \in \{1, 2, \ldots, n\}$, and let $\xi = -1$. We can take $\phi_{\min} = 0$ and $\phi_{\max} = d_{\max}/\sqrt{(d-2)n+2}$. Next, we bound

$$\|\phi\|_m \le \sum_{j=1}^{n} \frac{d_j - x_j}{\sqrt{(d-2)n+2}} = \sqrt{(d-2)n+2}.$$

Finally, observe that

$$\frac{d_{\max}^3(\|\phi\|_m + \phi_{\max})(\ln n + 2)}{((d-2)n+2)^{3/2}} = O\left(\frac{d_{\max}^3 \ln n}{(d-2)n} + \frac{d_{\max}^4 \ln n}{((d-2)n)^2}\right) = O\left(\frac{d_{\max}^3 \ln n}{(d-2)n}\right).$$

Now the result follows from Theorem 3.1. □

We can now prove Theorem 1.5, giving an asymptotic expression for the expected number $\mathbb{E}\,\tau_{\boldsymbol{d}}(\boldsymbol{x})$ of spanning trees in $\mathcal{G}_{\boldsymbol{d}}$ with degree sequence $\boldsymbol{x}$.

Proof of Theorem 1.5. Firstly note that, by (3.4),

$$\bar{\mu}(\boldsymbol{x}) = \frac{1}{(n-2)((d-2)n+2)} \sum_{j \ne k} (x_j - 1)(d_j - x_j)(d_k - x_k) \tag{4.1}$$

$$= \frac{1}{n} \sum_{j=1}^{n} (x_j - 1)(d_j - x_j) + O\left(\frac{d_{\max}^2}{(d-2)n}\right). \tag{4.2}$$

We rewrite (1.12) as

$$\mathbb{E}\,\tau_{\boldsymbol{d}}(\boldsymbol{x}) = e^{O(d_{\max}^4/((d-2)n))} \frac{(dn/2)_{n-1}\, 2^{n-1}\, \hat{d}^n}{(dn)_n} \binom{(d-1)n}{n-2}^{-1} \left(\prod_{j=1}^{n} \binom{d_j - 1}{x_j - 1}\right) e^{f(\boldsymbol{x})}\, \beta(\boldsymbol{x}) \tag{4.3}$$

with $f(\boldsymbol{x})$ as defined in (1.10). Applying Stirling's approximation gives

$$\frac{(dn/2)_{n-1}\, 2^{n-1}\, \hat{d}^n}{(dn)_n} = H_{\boldsymbol{d}} \left(1 + O\left(\frac{1}{(d-2)n}\right)\right)$$

where $H_{\boldsymbol{d}}$ is defined in the statement of Theorem 1.1. Combining Lemma 4.1 and (4.2) gives

$$\beta(\boldsymbol{x}) = \exp\left(-\bar{\mu}(\boldsymbol{x}) + O\left(\min\left\{\frac{d_{\max}^4}{(d-2)^2 n}, \frac{d_{\max}^3 \ln n}{(d-2)n}\right\}\right)\right)$$

$$= \exp\left(-\frac{1}{n} \sum_{j=1}^{n} (x_j - 1)(d_j - x_j)\right.$$

$$+O\left(\frac{d_{\max}^2}{(d-2)n}+\min\left\{\frac{d_{\max}^4}{(d-2)^2n},\frac{d_{\max}^3\ln n}{(d-2)n}\right\}\right)\right). \tag{4.4}$$

In some cases, when $d-2$ is small, we can obtain a smaller error bound by a different argument. Observe that

$$e^{-\bar{\mu}(\boldsymbol{x})}\le\beta(\boldsymbol{x})\le 1, \tag{4.5}$$

using Jensen's inequality for the lower bound. It follows from (4.2) that

$$\bar{\mu}(\boldsymbol{x})=O\left(\frac{d_{\max}^2}{(d-2)n}+(d-2)d_{\max}\right).$$

Hence we can replace the upper bound on $\beta(\boldsymbol{x})$ in (4.5) by $e^{-\bar{\mu}(\boldsymbol{x})}$ if we include an error term of this magnitude, leading to

$$\begin{aligned}\beta(\boldsymbol{x})&=\exp\left(-\bar{\mu}(\boldsymbol{x})+O\left(\frac{d_{\max}^2}{(d-2)n}+(d-2)d_{\max}\right)\right)\\&=\exp\biggl(-\frac{1}{n}\sum_{j=1}^{n}(x_j-1)(d_j-x_j)+O\biggl(\frac{d_{\max}^2}{(d-2)n}+(d-2)d_{\max}\biggr)\biggr)\end{aligned} \tag{4.6}$$

using (4.2). We may choose to use either this expression or (4.4), whichever gives the smaller bound. Finally, observe that

$$\lambda_0+\lambda_0^2=\frac{(R+d^2)^2}{4d^2}-\tfrac{1}{4}. \tag{4.7}$$

Combining this with (1.10), (4.3), (4.4) and (4.6) completes the proof. □

5 Proof of Theorem 1.1

In this section we prove Theorem 1.1 by summing the expression from Theorem 1.5 over all suitable degree sequences $\boldsymbol{x}$. Given a suitable degree sequence $\boldsymbol{x}$, define

$$g(\boldsymbol{x})=f(\boldsymbol{x})-\bar{\mu}(\boldsymbol{x})=\frac{(R+d^2)^2}{4d^2}-\tfrac{1}{4}-\lambda(\boldsymbol{x})-\lambda(\boldsymbol{x})^2-\bar{\mu}(\boldsymbol{x}), \tag{5.1}$$

using (4.7). By (4.2) and Theorem 1.5 we have

$$\begin{aligned}\mathbb{E}\,\tau_{\boldsymbol{d}}=H_{\boldsymbol{d}}\sum_{\boldsymbol{x}}\binom{(d-1)n}{n-2}^{-1}&\left(\prod_{j=1}^{n}\binom{d_j-1}{x_j-1}\right)\\&\times\exp\left(g(\boldsymbol{x})+O\left(\frac{d_{\max}^4}{(d-2)n}+\eta\right)\right)\end{aligned} \tag{5.2}$$

where the sum is over all suitable degree sequences $\boldsymbol{x}$. We now interpret this sum as an expected value of a function of a nonuniform distribution on suitable degree sequences.

Lemma 5.1. *Fix a partition $A_1, \ldots, A_n$ of $\{1, 2, \ldots, (d-1)n\}$ such that $|A_j| = d_j - 1$ for $j = 1, \ldots, n$, and let B be a uniformly random subset of $\{1, 2, \ldots, (d-1)n\}$ of size $n-2$. Define the random vector $\boldsymbol{X} = \boldsymbol{X}(B) = (X_1, \ldots, X_n)$ by $X_j = |A_j \cap B| + 1$. Then*

$$\mathbb{E}\, \tau_{\boldsymbol{d}} = H_{\boldsymbol{d}} \exp\left(O\left(\frac{d_{\max}^4}{(d-2)n} + \eta \right) \right) \mathbb{E}\left(e^{g(\boldsymbol{X})} \right).$$

Proof. Let $\boldsymbol{x}$ be a suitable degree sequence. Since the sets A_j are disjoint, there are $\prod_{j=1}^n \binom{d_j-1}{x_j-1}$ ways to choose a subset of $\{1, \ldots, (d-1)n\}$ with precisely $x_j - 1$ elements in A_j, for $j = 1, \ldots, n$. It follows that

$$\Pr(\boldsymbol{X} = \boldsymbol{x}) = \binom{(d-1)n}{n-2}^{-1} \prod_{j=1}^n \binom{d_j - 1}{x_j - 1}.$$

Substituting this into (5.2) completes the proof. □

Next, we prove that $\mathbb{E}\big(e^{g(\boldsymbol{X})}\big)$ can be approximated by $e^{\mathbb{E}\, g(\boldsymbol{X})}$ by applying Corollary 2.2. We say that two suitable degree sequences $\boldsymbol{x}$ and $\boldsymbol{x}'$ are *adjacent* if $\boldsymbol{x}$ and $\boldsymbol{x}'$ differ in precisely two entries, say in entries j and k, such that $x'_j = x_j + 1$ and $x'_k = x_k - 1$. Adjacent degree sequences correspond to subsets A, A' of $\{1, 2, \ldots, (d-1)n\}$ of size $n-2$ which have $n-3$ elements in common. In order to apply Corollary 2.2 to g we must bound the amount by which $g(\boldsymbol{x})$ can differ from $g(\boldsymbol{x}')$ when $\boldsymbol{x}$ and $\boldsymbol{x}'$ are adjacent.

Lemma 5.2. *Suppose that $\boldsymbol{x}$, $\boldsymbol{x}'$ are two suitable degree sequences which are adjacent. Then*

$$|g(\boldsymbol{x}) - g(\boldsymbol{x}')| = O\left(\frac{d_{\max}^2}{(d-2)n} \right).$$

Proof. Recall the definition of g in (5.1). Firstly, observe that

$$\lambda(\boldsymbol{x}')^2 - \lambda(\boldsymbol{x})^2 = (\lambda(\boldsymbol{x}') - \lambda(\boldsymbol{x}))\, (\lambda(\boldsymbol{x}') + \lambda(\boldsymbol{x})) = O(d_{\max})\, (\lambda(\boldsymbol{x}') - \lambda(\boldsymbol{x}))$$

since for any suitable $\boldsymbol{x}$ we have

$$\lambda(\boldsymbol{x}) = O\left(\frac{d_{\max}}{(d-2)n} \right) \sum_{j=1}^n (d_j - x_j) = O(d_{\max}).$$

Next we calculate that

$$|\lambda(\boldsymbol{x}') - \lambda(\boldsymbol{x})| = \frac{|(d_k - x_k) - (d_j - x_j - 1)|}{(d-2)n + 2} = O\left(\frac{d_{\max}}{(d-2)n} \right).$$

Therefore

$$\big|\lambda(\boldsymbol{x}) + \lambda(\boldsymbol{x})^2 - \big(\lambda(\boldsymbol{x}') - \lambda(\boldsymbol{x}')^2\big)\big| = O\left(\frac{d_{\max}^2}{(d-2)n} \right).$$

Now we consider $\bar{\mu}$. Suppose that $\boldsymbol{y}$ is a vector which disagrees with $\boldsymbol{x}$ in precisely one position, say $y_i = x_i + \zeta$ where $\zeta \in \{-1, 1\}$. Then using (4.1) (most conveniently in the form in (3.3)),

$$\begin{aligned}|\bar{\mu}(\boldsymbol{y}) - \bar{\mu}(\boldsymbol{x})| &\leq \frac{1}{(n-2)\,((d-2)n+2)} \sum_{j:j\neq i} (d_j - x_j)\; |(d_i - x_i) - (x_i + x_j - 2) + \zeta| \\ &= O\left(\frac{d_{\max}}{n}\right) = O\left(\frac{d_{\max}^2}{(d-2)n}\right).\end{aligned}$$

Applying this twice gives a bound of the same magnitude on $|\bar{\mu}(\boldsymbol{x}') - \bar{\mu}(\boldsymbol{x})|$, completing the proof. □

Now we apply Corollary 2.2 to prove the following.

Lemma 5.3. *Under the conditions of Theorem 1.1,*

$$\mathbb{E}\left(e^{g(\boldsymbol{X})}\right) = \exp\left(\mathbb{E}\, g(\boldsymbol{X}) + O\Big(\frac{d_{\max}^4}{(d-2)n}\Big)\right).$$

Proof. We will apply Corollary 2.2 to $h(B) = g(\boldsymbol{X}(B))$, where the random set B is defined in Lemma 5.1. We set $N = (d-1)n$ and $r = n-2$. Lemma 5.2 says that h changes by at most $\alpha = O(d_{\max}^2/((d-2)n))$ if two entries of the vector change by 1 (one increasing and one decreasing). The value of the error term given by Corollary 2.2 also depends on $\min\{r, N-r\} = \min\{n-2, (d-2)n+2\}$. We consider two cases.

If $(n-2) \leq (d-2)n + 2$ then Corollary 2.2 gives

$$\begin{aligned}\mathbb{E}(e^{g(\boldsymbol{X})}) &= \exp\left(\mathbb{E}\, g(\boldsymbol{X}) + O\Big(\frac{d_{\max}^4 (n-2)}{(d-2)^2 n^2}\Big)\right) \\ &= \exp\left(\mathbb{E}\, g(\boldsymbol{X}) + O\Big(\frac{d_{\max}^4}{(d-2)n}\Big)\right).\end{aligned}$$

(The second equality follows since in this case $d - 2 \geq 1 - \frac{4}{n} \geq \frac{1}{2}$.)

Otherwise it holds that $(d-2)n + 2 < n - 2$, and here Corollary 2.2 says that

$$\begin{aligned}\mathbb{E}(e^{g(\boldsymbol{X})}) &= \exp\left(\mathbb{E}\, g(\boldsymbol{X}) + O\Big(\frac{d_{\max}^4 ((d-2)n+2)}{(d-2)^2 n^2}\Big)\right) \\ &= \exp\left(\mathbb{E}\, g(\boldsymbol{X}) + O\Big(\frac{d_{\max}^4}{(d-2)n}\Big)\right),\end{aligned}$$

as required. □

To approximate $\mathbb{E}\, g(\boldsymbol{X})$, we need to be able to compute joint moments of the form $\mathbb{E}\left((X_j - 1)_s\, (X_k - 1)_t\right)$, where $\boldsymbol{X} = \boldsymbol{X}(B) = (X_1, \ldots, X_n)$. The random vector

$$\boldsymbol{X} - (1, 1, \ldots, 1) = (X_1 - 1, \ldots, X_n - 1)$$

has a multivariate hypergeometric distribution, and from this it follows that the entries $X_1, \ldots, X_n$ of $\boldsymbol{X} = \boldsymbol{X}(B)$ satisfy

$$\mathbb{E}((X_i - 1)_s (X_j - 1)_t) = (d_i - 1)_s (d_j - 1)_t \, \frac{(n-2)_{s+t}}{((d-1)n)_{s+t}} \tag{5.3}$$

for $i \neq j$. See for example [7, Equation (39.6)].

We now find an asymptotic expression for $\mathbb{E}(g(\boldsymbol{X}))$.

Lemma 5.4. *Under the conditions of Theorem 1.1,*

$$\mathbb{E}(g(\boldsymbol{X})) = \frac{6d^2 - 14d + 7}{4(d-1)^2} + \frac{R}{2(d-1)^3} + \frac{(2d^2 - 4d + 1)R^2}{4(d-1)^4\, d^2} + O\biggl(\frac{d_{\max}^3}{dn}\biggr).$$

Proof. First we estimate $\mathbb{E}\,\bar{\mu}(\boldsymbol{X})$ using (4.1). By (5.3) we have

$$\begin{aligned}
&\frac{\mathbb{E}((X_j - 1)(d_j - X_j)(d_k - X_k))}{(n-2)((d-2)n+2)} \\
&\quad = \frac{1}{(n-2)((d-2)n+2)} \Bigl((d_j - 2)(d_k - 1)\,\mathbb{E}(X_j - 1) \\
&\qquad\qquad - (d_j - 2)\,\mathbb{E}((X_j - 1)(X_k - 1)) \\
&\qquad\qquad - (d_k - 1)\,\mathbb{E}((X_j - 1)_2) + \mathbb{E}((X_j - 1)_2 (X_k - 1)) \Bigr) \\
&\quad = \frac{(d_j - 1)_2 (d_k - 1)}{(n-2)((d-2)n+2)} \left(\frac{n-2}{(d-1)n} - \frac{2(n-2)_2}{((d-1)n)_2} + \frac{(n-2)_3}{((d-1)n)_3} \right) \\
&\quad = \frac{(d_j - 1)_2\,(d_k - 1)\,((d-2)n+1)}{((d-1)n)_3} \\
&\quad = (d_j - 1)_2\,(d_k - 1) \left(\frac{d-2}{(d-1)^3 n^2} + O\left(\frac{1}{d^3 n^3}\right) \right).
\end{aligned}$$

Now

$$\begin{aligned}
\sum_{j \neq k} (d_j - 1)_2 (d_k - 1) &= \sum_{j=1}^{n} (d_j - 1)_2 \bigl((d-1)n - (d_j - 1)\bigr) \\
&= (d-1)(R + (d-1)_2) n^2 + O(d_{\max}^2 dn).
\end{aligned}$$

Hence the expected value of $\bar{\mu}(\boldsymbol{X})$ is given by

$$\begin{aligned}
\mathbb{E}\,\bar{\mu}(\boldsymbol{X}) &= \left(\frac{d-2}{(d-1)^3 n^2} + O\left(\frac{1}{d^3 n^3}\right) \right) \sum_{j \neq k} (d_j - 1)_2\,(d_k - 1) \\
&= \left(\frac{d-2}{(d-1)^3 n^2} + O\left(\frac{1}{d^3 n^3}\right) \right) \bigl((d-1)(R + (d-1)_2)\, n^2 + O(d_{\max}^2 dn)\bigr) \\
&= \frac{(d-2)(R + (d-1)_2)}{(d-1)^2} + O\left(\frac{d_{\max}^2}{dn}\right).
\end{aligned} \tag{5.4}$$

Next, recall that

$$\lambda(\boldsymbol{X}) = \frac{1}{2((d-2)n+2)} \sum_{j=1}^{n} (d_j - X_j)_2.$$

Applying (5.3) shows that

$$\begin{aligned}
&\frac{\mathbb{E}((d_j - X_j)_2)}{2((d-2)n+2)} \\
&\quad = \frac{1}{2((d-2)n+2)} \left((d_j-1)_2 - 2(d_j-2)\,\mathbb{E}(X_j-1) + \mathbb{E}((X_j-1)_2)\right) \\
&\quad = \frac{(d_j-1)_2}{2((d-2)n+2)} \left(1 - \frac{2(n-2)}{(d-1)n} + \frac{(n-2)_2}{((d-1)n)_2}\right) \\
&\quad = \frac{(d_j-1)_2\,((d-2)n+1)}{2((d-1)n)_2} \\
&\quad = (d_j-1)_2 \left(\frac{(d-2)}{2(d-1)^2 n} + O\left(\frac{1}{d^2n^2}\right)\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbb{E}(\lambda(\boldsymbol{X})) &= \sum_{j=1}^{n} (d_j-1)_2 \left(\frac{(d-2)}{2(d-1)^2 n} + O\left(\frac{1}{d^2n^2}\right)\right) \\
&= \frac{(d-2)(R + (d-1)_2)}{2(d-1)^2} + O\left(\frac{d_{\max}}{dn}\right).
\end{aligned} \tag{5.5}$$

The same approach works for $\mathbb{E}(\lambda(\boldsymbol{X})^2)$ but the details are a little messier. Observe that

$$\begin{aligned}
&\lambda(\boldsymbol{X})^2 \\
&= \frac{1}{4((d-2)n+2)^2} \left(\left(\sum_{j\neq k} (d_j - X_j)_2 (d_k - X_k)_2\right) + \sum_{j=1}^{n} (d_j - X_j)^2 (d_j - X_j - 1)^2\right).
\end{aligned} \tag{5.6}$$

Applying (5.3) to the off-diagonal summands gives

$$\begin{aligned}
&\frac{\mathbb{E}((d_j - X_j)_2 (d_k - X_k)_2)}{4((d-2)n+2)^2} \\
&\quad = \frac{1}{4((d-2)n+2)^2} \Big((d_j-1)_2(d_k-1)_2 - 2(d_j-1)_2(d_k-2)\,\mathbb{E}(X_k-1) \\
&\qquad - 2(d_j-2)(d_k-1)_2\,\mathbb{E}(X_j-1) + (d_j-1)_2\,\mathbb{E}((X_k-1)_2) \\
&\qquad + 4(d_j-2)(d_k-2)\,\mathbb{E}((X_j-1)(X_k-1)) + (d_k-1)_2\,\mathbb{E}((X_j-1)_2) \\
&\qquad - 2(d_j-2)\ \mathbb{E}((X_j-1)(X_k-1)_2) - 2(d_k-2)\ \mathbb{E}((X_j-1)_2(X_k-1)) \\
&\qquad\qquad + \mathbb{E}((X_j-1)_2(X_k-1)_2)\Big) \\
&\quad = \frac{(d_j-1)_2(d_k-1)_2}{4((d-2)n+2)^2} \left(1 - \frac{4(n-2)}{(d-1)n} + \frac{6(n-2)_2}{((d-1)n)_2} - \frac{4(n-2)_3}{((d-1)n)_3} + \frac{(n-2)_4}{((d-1)n)_4}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{(d_j-1)_2(d_k-1)_2\,((d-2)n+1)_3}{4((d-1)n)_4((d-2)n+2)} \\
&= (d_j-1)_2(d_k-1)_2\left(\frac{(d-2)^2}{4(d-1)^4n^2} + O\left(\frac{1}{d^3n^3}\right)\right).
\end{aligned}
$$

Next, calculate

$$
\sum_{j\neq k}(d_j-1)_2(d_k-1)_2 = (R+(d-1)_2)^2n^2 + O(d_{\max}^3 dn).
$$

Therefore the contribution to $\lambda(\boldsymbol{X})^2$ from the off-diagonal summands is

$$
\begin{aligned}
&\left(\frac{(d-2)^2}{4(d-1)^4n^2} + O\left(\frac{1}{d^3n^3}\right)\right)\sum_{j\neq k}(d_j-1)_2(d_k-1)_2 \\
&= \left(\frac{(d-2)^2}{4(d-1)^4n^2} + O\left(\frac{1}{d^3n^3}\right)\right)\left((R+(d-1)_2)^2n^2 + O\left(d_{\max}^3 dn\right)\right) \\
&= \frac{(d-2)^2(R+(d-1)_2)^2}{4(d-1)^4} + O\left(\frac{d_{\max}^3}{dn}\right).
\end{aligned}
$$

The contribution to $\mathbb{E}(\lambda(\boldsymbol{X})^2)$ from the diagonal terms of (5.6) (that is, the second summation in (5.6)) is

$$
\begin{aligned}
\frac{1}{4((d-2)n+2)^2}\sum_{j=1}^{n}\mathbb{E}((d_j-X_j)^2(d_j-X_j-1)^2) &= O\left(\frac{d_{\max}^2}{(d-2)n}\right)\mathbb{E}(\lambda(\boldsymbol{X})) \\
&= O\left(\frac{d_{\max}^3}{dn}\right),
\end{aligned}
$$

using (5.5). Therefore

$$
\mathbb{E}(\lambda(\boldsymbol{X})^2) = \frac{(d-2)^2\,(R+(d-1)_2)^2}{4(d-1)^4} + O\left(\frac{d_{\max}^3}{dn}\right). \tag{5.7}
$$

The result follows by combining (4.7), (5.4), (5.5) and (5.7), after some rearranging. □

Now we may easily prove our main theorem.

Proof of Theorem 1.1. The number of graphs with degree sequence $\boldsymbol{d}$ is positive when n is sufficiently large, by (1.9). That is, $\boldsymbol{d}$ is graphical for sufficiently large n. The claimed asymptotic expression for $\mathbb{E}\,\tau_{\boldsymbol{d}}$ then follows immediately from Lemmas 5.1, 5.3 and 5.4. We also briefly justify the bound

$$
\begin{aligned}
\eta &= \min\left\{\frac{d_{\max}^4}{(d-2)^2n},\ \frac{d_{\max}^3\log n}{(d-2)n},\ d_{\max}(d-2)\right\} \\
&= O\left(\frac{d_{\max}^4}{(d-2)n} + \frac{(\log n)^{5/2}}{n^{1/2}}\right).
\end{aligned}
$$

Note that $(d_{\max}^3\log n)/((d-2)n) \le d_{\max}^4/((d-2)n)$ if $d_{\max} \ge \log n$. When $d_{\max} \le \log n$, take the geometric mean of $(d_{\max}^3\log n)/((d-2)n)$ and $d_{\max}(d-2)$. □

Acknowledgements

We would like to thank the referees for their helpful comments.

References

[1] N. Alon, The number of spanning trees in regular graphs, *Random Structures Algorithms* **1** (1990), no. 2, 175–182.

[2] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *Europ. J. Combin.* **1** (1980), 311–316.

[3] F. Chung and S.-T. Yau, Coverings, heat kernels and spanning trees, *Electron. J. Combin.* **6** (1999), #R12.

[4] C. Cooper, A. R. A. McGrae and M. Zito, Martingales on trees and the empire chromatic number of random trees, in *Fundamentals of Computation Theory 2009* (M. Kutyłowski, M. Gębala and W. Charatonik, eds.), LNCS vol. 5699, Springer, Heidelberg, 2009, pp. 74–83.

[5] A. Frieze and B. Pittel, Perfect matchings in random graphs with prescribed minimal degree, in *Mathematics and Computer Science III: Algorithms, Trees, Combinatorics and Probabilities* (M. Drmota, P. Flajolet, D. Gardy and B. Gittenberger, eds.), Birkhäuser, Basel, 2004, pp. 95–132.

[6] C. Greenhill, M. Kwan and D. Wind, On the number of spanning trees in random regular graphs, *Electron. J. Combin.* **21** (2014) #P1.45.

[7] N. L. Johnson, S. Kotz and N. Balakrishnan, *Discrete Multivariate Distributions*, Wiley, New York, 1997.

[8] A. V. Kostochka, The number of spanning trees in graphs with a given degree sequence, *Random Structures Algorithms* **6** (1995), no. 2-3, 269–274.

[9] R. Lyons, Asymptotic enumeration of spanning trees, *Combin. Probab. Comput.* **14** (2005), no. 4, 491–522.

[10] B. D. McKay, Subgraphs of random graphs with specified degrees, *Congr. Numer.*, **33** (1981) 213–223.

[11] B. D. McKay, Spanning trees in random regular graphs, in *Proceedings of the Third Carribean Conference on Combinatorics and Computing* (C. Cadogan, ed.), Department of Mathematics, University of the West Indies, 1981, pp. 139–143.

[12] B. D. McKay, Spanning trees in regular graphs, *European J. Combin.* **4** (1983), no. 2, 149–160.

[13] B. D. McKay, Asymptotics for symmetric 0-1 matrices with prescribed row sums, *Ars Combin.*, **19A** (1985) 15–26.

[14] C. McDiarmid, Concentration, in *Probabilistic Methods for Algorithmic Discrete Mathematics* (M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, B. Reed, eds.), Springer, Berlin, 1998, pages 195–248.

[15] C. Merino and D. J. Welsh, Forests, colorings and acyclic orientations of the square lattice, *Ann. Combin.* **3** (1999), no. 2-4, 417–429.

[16] J. W. Moon, *Counting Labelled Trees*, Canadian Mathematical Monographs, Vol. 1, Canadian Mathematical Congress, Montreal, 1970.

[17] L. W. Shapiro, An electrical lemma, *Math. Mag.* **60** (1987), no. 1, 36–38.

[18] F. Wu, Number of spanning trees on a lattice, *J. Phys. A* **10** (1977), no. 6, L113.