Connectivity of Addable Monotone Graph Classes ⋆

L. Addario-Berry, C. McDiarmid

Department of Statistics, 1 South Parks Road, Oxford, OX1 3TG, UK

B. Reed

School of Computer Science, McGill University, 3480 University Street, Montreal, Quebec, H3A 2A7, Canada

Abstract

A class $\mathcal{A}$ of labelled graphs is *weakly addable* if if for all graphs G in $\mathcal{A}$ and all vertices u and v in distinct connected components of G, the graph obtained by adding an edge between u and v is also in $\mathcal{A}$; the class $\mathcal{A}$ is *monotone* if for all $G \in \mathcal{A}$ and all subgraphs H of G, we have $H \in \mathcal{A}$. We show that for any weakly addable, monotone class $\mathcal{A}$ whose elements have vertex set $\{1, \ldots, n\}$, the probability that a uniformly random element of $\mathcal{A}$ is connected is at least $(1 - o_n(1))e^{-0.540760}$, where $o_n(1) \to 0$ as $n \to \infty$. Furthermore, if every element of $\mathcal{A}$ has girth at least $g > 1$, then the probability that $\mathcal{A}$ is connected is at least $(1 - o_g(1)))e^{-1/2}$. The latter result establishes a conjecture of McDiarmid et al. (2006) for graphs of large girth.

1. Introduction

Given a class $\mathcal{A}$ of labelled graphs, we say that $\mathcal{A}$ is *weakly addable* (or *bridge-addable*) if for all graphs G in $\mathcal{A}$ and all vertices u and v in distinct connected components of G, the graph obtained by adding an edge between u and v is also in $\mathcal{A}$. The concept of addability was introduced in McDiarmid et al. (2005) in the course of studying random planar graphs. It was shown there that, for a uniformly random element of a finite non-empty weakly addable class $\mathcal{A}$ of graphs, the probability that it is connected is at least e^{-1}.

⋆ This research was supported by the Banff International Research station and the McGill Bellairs Research Institute

Email addresses: addario@stats.ox.ac.uk (L. Addario-Berry,), cmcd@stats.ox.ac.uk (C. McDiarmid), breed@cs.mcgill.ca (B. Reed).

As well as the class of planar graphs, other examples of weakly addable graph classes include forests, series-parallel graphs, graphs with tree-width at most k, graphs embeddable on any fixed surface, and more generally any minor-closed graph class with cut-point-free excluded minors; triangle-free graphs, and more generally H-free graphs for any two-edge-connected graph H; and k-colourable graphs.

It is well-known (Cayley, 1889) that there are n^{n-2} trees on n labelled vertices. Together with a result of Rényi and Szekeres (1967) (see also Moon (1970)) that the corresponding number of forests is asymptotic to $e^{\frac{1}{2}}n^{n-2}$, we see that, for a uniformly random forest on n labelled vertices, the probability that it is connected is asymptotically $e^{-\frac{1}{2}}$.

In McDiarmid et al. (2006) it was suggested that if connectivity is desired then the worst possible example of a weakly addable graph class is the class of forests. More precisely, they conjectured that the lower bound of e^{-1} on the probability of connectedness for a weakly addable graph class can be improved to $(1 + o(1))e^{-\frac{1}{2}}$. (They assumed also that the graph class was closed under isomorphism.) Recently, Balister et al. (2007) took a first step towards proving this conjecture, proving an asymptotic lower bound of $e^{-0.7983}$.

Observe that the examples above of weakly addable graph classes, and many other interesting examples of such graph classes, also satisfy the property of being *monotone* (decreasing); that is, given a graph G in $\mathcal{A}$, each graph obtained by deleting edges from G is also in $\mathcal{A}$. In this paper we investigate the probability of connectivity of a uniformly random element of a monotone weakly addable graph class; for such graph classes, we improve the bounds of Balister et. al. and obtain a bound which is quite close to the conjectured lower bound. Our method relies upon a reduction to weighted forests; when the class contains two-connected components our reduction yields vertices of weight larger than one, and these vertices prevent us from proving the desired lower bound exactly. However, we find that we can approach arbitrarily close to that bound if the graphs have sufficiently high girth.

For several weakly addable graph classes, including some of those mentioned above, the asymptotic probability of connectedness has recently been determined – see the last section below.

2. Main results

A *bridge* in a graph G is an edge e such that $G - e$ has strictly more connected components than G. We say that a class $\mathcal{A}$ of graphs is *bridge-alterable* if for any graph G and bridge e in G, G is in $\mathcal{A}$ if and only if $G - e$ is. We remark that if $\mathcal{A}$ is bridge-alterable then it is weakly addable, and if it is both weakly addable and monotone then it is bridge-alterable.

For any class $\mathcal{A}$ of graphs, let $\mathcal{A}_n$ denote the set of graphs in $\mathcal{A}$ on the vertex set $[n] = \{1, \dots, n\}$. Let

$$\beta = \sum_{i=1}^{5} \frac{i^{i-2}}{i!e^i} + \sum_{i=6}^{\infty} \frac{1}{i^2 e} = \frac{\pi^2}{6e} - \sum_{i=1}^{5} \left(\frac{1}{i^2 e} - \frac{i^{i-2}}{i!e^i} \right) \approx 0.540760.$$

We shall show that, if $\mathcal{A}$ is bridge-alterable and $\mathbf{G}$ is a uniformly random element of $\mathcal{A}_W$, then $\mathbf{P}\{\mathbf{G} \text{ is connected}\} \geq e^{-\beta+o(1)}$, where $o(1) \to 0$ as $W \to \infty$. More precisely, we shall prove:

Theorem 1 *For any $\varepsilon > 0$ there exists W_0 such that, if $W \geq W_0$ and $\mathcal{A}$ is a non-empty bridge-alterable class of graphs on $\{1, \dots, W\}$, and if $\mathbf{G}$ is a uniformly random element of $\mathcal{A}$, then*

$$\mathbf{P}\{\mathbf{G} \text{ is connected}\} \geq (1-\varepsilon)e^{-\beta}. \tag{1}$$

This is our main theorem. As a biproduct of its proof, we find that we obtain a lower bound close to $e^{-\frac{1}{2}}$ given sufficiently high girth. Recall that the *girth* of a graph G is the minimum length of a cycle (and is ∞ if G is a forest).

Theorem 2 *For any $\varepsilon > 0$ there exist W_0 and g_0 such that, if $W \geq W_0$ and $\mathcal{A}$ is a non-empty bridge-alterable class of graphs on $\{1, \dots, W\}$ in which each graph has girth at least g_0, and if $\mathbf{G}$ is a uniformly random element of $\mathcal{A}$, then*

$$\mathbf{P}\{\mathbf{G} \text{ is connected}\} \geq (1-\varepsilon)e^{-\frac{1}{2}}. \tag{2}$$

In the next section we show that Theorem 1 will follow from Claim 4 concerning random weighted forests, and we introduce Lemma 6. In the following two sections we show that Lemma 6 will yield Claim 4, and then prove Lemma 6. At this stage we will have completed the proof of Theorem 1. Following that we quickly deduce Theorem 2 from three of the lemmas proved earlier, and finally we make some concluding remarks.

3. A reduction to weighted forests

We shall prove Theorem 1 by partitioning $\mathcal{A}_W$ and showing that an inequality such as (1) holds for a uniformly random element of each block of the partition. (This step of our proof is essentially Lemma 2.1 of Balister et al. (2007).)

Definition 3 *Given a graph G, let $b(G)$ be the graph obtained by removing all bridges from G. We say G and G' are* equivalent *if $b(G) = b(G')$, and in this case write $G \sim G'$. For a graph G, let $[G]$ be the set of graphs G' for which $b(G') = b(G)$.*

It is easily seen that $\sim$ is an equivalence relation on graphs, and thus we always have $[G] = [b(G)]$. Furthermore, if $G \in \mathcal{A}_W$ then as $\mathcal{A}$ is closed under deleting bridges, $b(G) \in \mathcal{A}_W$, and as $\mathcal{A}$ is weakly addable, we have $[G] \subseteq \mathcal{A}_W$. It follows that $\mathcal{A}_W$ can be written as a union of some set of disjoint equivalence classes $[G_1], \dots, [G_{m_W}]$. To prove Theorem 1, we will in fact prove:

Claim 4 *For any graph $G \in \mathcal{A}_W$, if H is a uniformly random element of $[G]$ then*

$$\mathbf{P}\{H \text{ is connected}\} \geq e^{-\beta + o(1)},$$

where $o(1) \to 0$ as $W \to \infty$.

Clearly, Theorem 1 immediately follows from Claim 4, and it thus remains to prove Claim 4. Fix a bridgeless graph G on vertex set $[W]$ and let $\mathcal{B} = [G]$. Write $C_1, \dots, C_n$ for the components of G, and let $w_i = |V(C_i)|$ for $i = 1, \dots, n$, so $W = \sum_{i=1}^n w_i$. We remark that since the components $C_1, \dots, C_n$ are bridgeless, either $w_i = 1$ or $w_i \geq 3$ for all $i \in \{1, \dots, n\}$. Indeed, if G has girth g (that is, g is the minimum length of a cycle), then each $w_i \neq 1$ satisfies $w_i \geq g$. We denote by $\overrightarrow{w}$ the vector $(w_1, \dots, w_n)$. We use $\overrightarrow{w}$ to define a probability measure on the set $\mathcal{F}_n$ of labeled forests with vertex set $\{1, \dots, n\}$. Given $F \in \mathcal{F}_n$, let

$$mass(F) = mass_{\overrightarrow{w}}(F) = \prod_{i=1}^{n} w_i^{d_F(i)},$$

let $K = \sum_{F \in \mathcal{F}_n} mass(F)$, and let $\mathbf{F}$ be a random element of $\mathcal{F}_n$ with $\mathbf{P}\{\mathbf{F} = F\} = mass(F)/K$ for all $F \in \mathcal{F}_n$. We say $\mathbf{F}$ is *distributed according to* $\overrightarrow{w}$; when we wish to highlight the distribution of $\mathbf{F}$, we will sometimes write $\mathbf{F}_{\overrightarrow{w}}$ in place of $\mathbf{F}$. For our purposes, the key fact about such a random forest is the following:

Lemma 5 *Given a uniformly random element* $\mathbf{H}$ *of* $\mathcal{B}$,

$$\mathbf{P}\{\mathbf{H} \textit{ is connected}\} = \mathbf{P}\{\mathbf{F} \textit{ is connected}\}.$$

PROOF. We construct a flow from $\mathcal{B}$ to $\mathcal{F}_n$ in the following fashion: given $G \in \mathcal{B}$, let $f(G)$ be the graph obtained from G by contracting C_i to a single point for each $i = 1, \ldots, n$; then $f(G) \in \mathcal{F}_n$, and for each $F \in \mathcal{F}_n$, the set $f^{-1}(F)$ has cardinality precisely $\prod_{i=1}^n w_i^{d_F(i)}$. Since $G \in \mathcal{B}$ in connected if and only if $f(G)$ is connected, it follows that

$$\begin{aligned}\mathbf{P}\{\mathbf{H} \text{ is connected}\} &= \frac{|\{G \in \mathcal{B} : G \text{ is connected}\}|}{|\mathcal{B}|} \\ &= \frac{\sum_{\{F \in \mathcal{F}_n : F \text{ is connected}\}} |f^{-1}(F)|}{\sum_{F \in \mathcal{F}_n} |f^{-1}(F)|} \\ &= \frac{\sum_{\{F \in \mathcal{F}_n : F \text{ is connected}\}} mass(F)}{K} \\ &= \mathbf{P}\{\mathbf{F} \text{ is connected}\}. \qquad \Box\end{aligned}$$

To prove Claim 4, it therefore suffices to show that for such a random forest $\mathbf{F}$, $\mathbf{P}\{\mathbf{F} \text{ is connected}\} \geq e^{-\beta+o(1)}$, where $o(1)$ tends to zero uniformly in W.

For $i = 1, \ldots, n$, let $\mathcal{F}_{n,i}$ be the set of elements of $\mathcal{F}_n$ with i components (so $F \in \mathcal{F}_{n,1}$ precisely if F is connected). For larger i set $\mathcal{F}_{n,i} = \emptyset$. It turns out that bounds on $\mathbf{P}\{\mathbf{F} \text{ is connected}\}$ follow from bounds on the ratio between $\mathbf{P}\{F \in \mathcal{F}_{n,2}\}$ and $\mathbf{P}\{F \in \mathcal{F}_{n,1}\}$. More precisely, Claim 4 follows from Lemma 5 and the following lemma.

Lemma 6 *For all* $\epsilon > 0$, *for* W *sufficiently large, for all* $\overrightarrow{w} = (w_1, \ldots, w_n)$ *with* $\sum_{j=1}^n w_j = W$,

$$\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,2}\} \leq (1+\epsilon)\beta\, \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\}. \tag{3}$$

In Section 4 we explain how to use Lemma 6 to prove Claim 4; in Section 5 we prove Lemma 6.

4. Proof of Claim 4 assuming Lemma 6

The proof of Claim 4 proceeds somewhat differently depending on the value of the ratio of n and W. When W is much larger than n, the proof is rather straightforward, and in fact does not require Lemma 6 at all, but rather Lemma 9, below. In both cases, however, we construct a flow φ on $\mathcal{F}_n$ and using it to analyze the probability mass of $\mathcal{F}_{n,i}$ relative to that of $\mathcal{F}_{n,i+1}$ (for $i = 1, \ldots, n-1$). Given a graph G, let $c(G)$ be the set of connected components of G. Given a forest $F \in \mathcal{F}_n$ and $T \in c(F)$, let $w(T) = \sum_{i \in V(T)} w_i$. Consider forests $F, F' \in \mathcal{F}_n$ such that F' can be obtained from F by the addition of an edge e. Writing $T \neq T' \in c(F)$ as shorthand for $\{\{T, T'\} \subseteq c(F) : T \neq T'\}$, we let

$$\varphi(F',F) = \frac{mass(F')}{\sum_{T \neq T' \in c(F)} w(T)w(T')}. \tag{4}$$

For all other pairs F, F', we let $\varphi(F', F) = 0$.

Lemma 7 *For all $i = 1, \ldots, n-1$*

$$\sum_{F' \in \mathcal{F}_{n,i}} \sum_{F \in \mathcal{F}_{n,i+1}} \varphi(F', F) = \sum_{F \in \mathcal{F}_{n,i+1}} \mathrm{mass}(F) = K \cdot \mathbf{P}\left\{\mathbf{F} \in \mathcal{F}_{n,i+1}\right\}.$$

PROOF. Given $i \in \{1, \ldots, n-1\}$ and $F \in \mathcal{F}_{n,i+1}$, if $F' \in \mathcal{F}_{n,i}$ is obtained from F by the addition of edge uv, then $mass(F') = mass(F) \cdot w_u \cdot w_v$. We thus have

$$\begin{aligned}
\sum_{F' \in \mathcal{F}_{n,i}} \varphi(F', F) &= \left(\sum_{T \neq T' \in c(F)} \sum_{u \in V(T), v \in V(T')} mass(F) \cdot w_u \cdot w_v \right) \\
&\qquad \cdot \left(\frac{1}{\sum_{T \neq T' \in c(F)} w(T)w(T')} \right) \\
&= \frac{mass(F)}{\sum_{T \neq T' \in c(F)} w(T)w(T')} \cdot \left(\sum_{T \neq T' \in c(F)} \sum_{u \in V(T), v \in V(T')} w_u \cdot w_v \right) \\
&= mass(F). \qquad \square
\end{aligned}$$

We remark that we have actually proved the stronger fact that for all $i = 1, \ldots, n-1$ and all $F \in \mathcal{F}_{n,i+1}$, $\sum_{F' \in \mathcal{F}_{n,i}} \varphi(F', F) = mass(F)$. Using Lemma 7, we can straightforwardly prove a first bound on the ratio of the mass of $\mathcal{F}_{n,i}$ and that of $\mathcal{F}_{n,i+1}$.

Lemma 8 *For all positive integers W and all positive integer weight vectors $\overrightarrow{w} = (w_1, \ldots, w_n)$ with $\sum_{j=1}^n w_j = W$,*

$$\mathbf{P}\left\{\mathbf{F} \in \mathcal{F}_{n,i+1}\right\} \leq \frac{\mathbf{P}\left\{\mathbf{F} \in \mathcal{F}_{n,i}\right\}(n/W)}{i}, \tag{5}$$

PROOF. [Proof of Lemma 8] Fix i with $1 \leq i \leq n-1$. By the definition of φ, for all $F' \in \mathcal{F}_{n,i}$ we have

$$\sum_{F \in \mathcal{F}_{n,i+1}} \varphi(F', F) = mass(F') \cdot \sum_{e \in E(F')} \frac{1}{\sum_{T \neq T' \in c(F'-e)} w(T)w(T')}. \tag{6}$$

We assert that for any set of positive integers $a_1, \ldots, a_{i+1}$ with $\sum_{j=1}^{i+1} a_j = W$,

$$\sum_{1 \leq j < k \leq i+1} a_j a_k \geq i(W-i) + \binom{i}{2} \tag{7}$$

(where we let $\binom{1}{2} = 0$). To see this, if $a_1 \geq a_2 \geq 2$ then let $a_1' = a_1 + 1$, $a_2' = a_2 - 1$ and $a_j' = a_j$ for each $j \geq 3$. Then with sums as above, $\sum_{j<k} a_j' a_k' - \sum_{j<k} a_j a_k = a_1' a_2' - a_1 a_2 = -a_1 + a_2 - 1 < 0$. Hence the sum is minimised when there are i entries 1 and one entry $W - i$.

In particular, for any $F' \in \mathcal{F}_{n,i}$ and any $e \in E(F')$, we have

$$\sum_{T \neq T' \in c(F'-e)} w(T)w(T') \geq i(W-i) + \binom{i}{2}. \tag{8}$$

Since, if $F' \in \mathcal{F}_{n,i}$ then F' has exactly $n-i$ edges, it follows from (6) and (8) that for all $F' \in \mathcal{F}_{n,i}$

$$\begin{aligned}\sum_{F \in \mathcal{F}_{n,i+1}} \varphi(F', F) &\leq mass(F') \cdot (n-i) \cdot \left(\frac{1}{i(W-i) + \binom{i}{2}} \right) \\ &\leq mass(F') \cdot \frac{1}{i} \cdot \frac{n}{W},\end{aligned}$$

so

$$\begin{aligned}\sum_{F' \in \mathcal{F}_{n,i}} \sum_{F \in \mathcal{F}_{n,i+1}} \varphi(F', F) &\leq \sum_{F' \in \mathcal{F}_{n,i}} mass(F') \cdot \frac{1}{i} \cdot \frac{n}{W} \\ &= K \cdot \mathbf{P}\left\{\mathbf{F} \in \mathcal{F}_{n,i}\right\} \cdot \frac{1}{i} \cdot \frac{n}{W}, \end{aligned} \tag{9}$$

and (5) follows by combining Lemma 7 and (9). □

As a consequence of Lemma 8, we have

Lemma 9 *For all positive integers W and all weight vectors $\overrightarrow{w} = (w_1, \ldots, w_n)$ with $\sum_{j=1}^{n} w_j = W$, $\mathbf{P}\{\mathbf{F}$ is connected$\} > e^{-n/W}$.*

PROOF. It follows immediately from (5) that

$$\mathbf{P}\left\{\mathbf{F} \in \mathcal{F}_{n,i+1}\right\} \leq \frac{\mathbf{P}\left\{\mathbf{F} \in \mathcal{F}_{n,1}\right\} (n/W)^i}{i!}, \tag{10}$$

and (10) implies that

$$1 = \sum_{i=0}^{n-1} \mathbf{P}\left\{\mathbf{F} \in \mathcal{F}_{n,i+1}\right\} < \sum_{i \geq 0} \frac{\mathbf{P}\left\{\mathbf{F} \in \mathcal{F}_{n,1}\right\} (n/W)^i}{i!} = e^{n/W} \mathbf{P}\left\{\mathbf{F} \in \mathcal{F}_{n,1}\right\}. \qquad \square$$

When $n/W \leq \beta$, Claim 4 follows immediately from Lemma 5 and Lemma 9. To explain why Lemma 6 implies Claim 4 when n is not much smaller than W, it turns out to be useful to prove a slightly more general implication. For each finite non-empty set V of positive integers, let $\mathcal{G}(V)$ be the set of all graphs on the vertex set V, and let $\mathcal{G}^k(V)$ be the set of all graphs in $\mathcal{G}(V)$ with exactly k components. Also, write $\mathcal{G}_n$ for $\mathcal{G}(\{1, \ldots, n\})$, and $\mathcal{G}_n^k$ for $\mathcal{G}^k(\{1, \ldots, n\})$. For each positive integer n, let μ_n be a measure on the set of all graphs with vertex set a subset of $\{1, \ldots, n\}$, which is multiplicative on components (that is, if G has components $H_1, \ldots, H_k$, then $\mu_n(G) = \prod_{i=1}^{k} \mu_n(H_i)$).

Lemma 10 *Suppose there exist $x > 0$ and integers $n \geq m_0 \geq 1$ such that*

$$\mu_n(\mathcal{G}^2(V)) \leq x\mu_n(\mathcal{G}^1(V)) \qquad \text{for all } V \subseteq \{1, \ldots, n\} \text{ with } |V| \geq m_0. \tag{11}$$

Let k be a positive integer and suppose that $n \geq km_0$. Then

$$\mu_n(\mathcal{G}_n^{k+1}) \leq \frac{x}{k} \mu_n(\mathcal{G}_n^k). \tag{12}$$

PROOF. Let $\mathcal{A}$ be the collection of all sets $\{H_1, \ldots, H_{k-1}\}$ of $k-1$ connected graphs such that the vertex sets $V(H_i)$ are pairwise disjoint subsets of $\{1, \ldots, n\}$ and

$$\max_{1 \le i \le k-1} |V(H_i)| < n - \sum_{i=1}^{k-1} |V(H_i)|.$$

Let $\mathcal{H} = \{H_1, \ldots, H_{k-1}\} \in \mathcal{A}$, let $V_{\mathcal{H}} = \{1, \ldots, n\} \setminus \left(\bigcup_{i=1}^{k-1} V(H_i)\right)$, and note that $|V_{\mathcal{H}}| > \max_{1 \le i \le k-1} |V(H_i)|$ and $|V_{\mathcal{H}}| \ge m_0$. For $j = k$ and $k+1$, let $\mathcal{G}_n^j(\mathcal{H})$ denote the set of all graphs G in $\mathcal{G}_n^j$ such that $H_1, \ldots, H_{k-1}$ are each components of G. Then, letting $\alpha = \prod_{i=1}^{k-1} \mu_n(H_i)$, by the multiplicativity of μ_n and by (11) we have

$$\mu_n(\mathcal{G}_n^{k+1}(\mathcal{H})) = \alpha \cdot \mu_n(\mathcal{G}^2(V_{\mathcal{H}})) \le x\alpha \cdot \mu_n(\mathcal{G}^1(V_{\mathcal{H}})) = x \cdot \mu_n(\mathcal{G}_n^k(\mathcal{H})). \tag{13}$$

Next, consider any graph $G \in \mathcal{G}_n^{k+1}$, and suppose that G has components $G_1, \ldots, G_{k+1}$, where $|V(G_1)| \le \ldots \le |V(G_{k+1})|$. For each set $\mathcal{H}$ formed by picking any $k-1$ of the graphs $G_1, \ldots, G_k$, we have $\mathcal{H} \in \mathcal{A}$ and $G \in \mathcal{G}_n^{k+1}(\mathcal{H})$. It follows that

$$k \cdot \mu_n(\mathcal{G}_n^{k+1}) \le \sum_{\mathcal{H} \in \mathcal{A}} \mu_n(\mathcal{G}_n^{k+1}(\mathcal{H})). \tag{14}$$

Applying (13) to bound the right-hand side of (14), we obtain

$$k \cdot \mu_n(\mathcal{G}_n^{k+1}) \le x \cdot \sum_{\mathcal{H} \in \mathcal{A}} \mu_n(\mathcal{G}_n^k(\mathcal{H})). \tag{15}$$

Furthermore, the sets $\{\mathcal{G}_n^k(\mathcal{H}) : \mathcal{H} \in \mathcal{A}\}$ are pairwise disjoint subsets of $\mathcal{G}_n^k$, so $\sum_{\mathcal{H} \in \mathcal{A}} \mu_n(\mathcal{G}_n^k(\mathcal{H})) \le \mu_n(\mathcal{G}_n^k)$, which combined with (15) yields that

$$k \cdot \mu_n(\mathcal{G}_n^{k+1}) \le x \cdot \mu_n(\mathcal{G}_n^k). \qquad \square$$

Lemma 10 allows us to use Lemma 6 to derive bounds on the ratio between $\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,i+1}\}$ and $\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,i}\}$ for $i > 1$.

Lemma 11 *Suppose that there exist γ with $0 < \gamma < 1$, m_0 with $m_0 > 0$, and a non-empty set of positive integers $\mathcal{W}$ such that for any positive integer weights $\overrightarrow{w} = (w_1, \ldots, w_n)$ with $w_i \in \mathcal{W}$ for $i = 1, \ldots, n$ and with $\sum_{k=1}^n w_k \ge m_0$, $\mathbf{P}\{\mathbf{F}_{\overrightarrow{w}} \in \mathcal{F}_{n,2}\} \le \gamma \mathbf{P}\{\mathbf{F}_{\overrightarrow{w}} \in \mathcal{F}_{n,1}\}$. Fix any positive integer j. Then for W sufficiently large, for all integers i with $1 \le i \le j$ and any positive integer weights $\overrightarrow{w} = \{w_1, \ldots, w_n\} \in \mathcal{W}^n$ with $\sum_{k=1}^n w_k = W$,*

$$\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,i+1}\} \le \frac{\gamma \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,i}\}}{i}. \tag{16}$$

PROOF. Suppose γ, m_0, and $\mathcal{W}$ satisfy the hypotheses of the lemma, and fix some positive integer j. Observe first that, by Lemma 8, the inequality in Lemma 6 holds if $n \le \gamma W$. We may thus assume that $n > \gamma W$.

Let $n \ge m_0$ and consider any weights $w_1, \ldots, w_n$. Define $\mu_n(G)$ for each graph G with vertex set $V \subseteq \{1, \ldots, n\}$ by setting $\mu_n(G) = \prod_{i \in V} w_i^{d_G(i)}$ if G is a forest and $\mu_n(G) = 0$ otherwise. Then μ_n is multiplicative on components, and by the hypotheses of the lemma, for each $V \subseteq \{1, \ldots, n\}$ with $\sum_{i \in V} w_i \ge m_0$ we have

$$\mu_n(\mathcal{G}^2(V)) \le \gamma \mu_n(\mathcal{G}^1(V));$$

Now we may use Lemma 10 to obtain

$$\mu_n(\mathcal{G}_n^{k+1}) \le \frac{\gamma}{k}\mu_n(\mathcal{G}_n^k)$$

whenever $n \ge km_0$. Since $n \ge km_0$ whenever $W \ge km_0/\gamma$, Lemma 11 follows. □

PROOF. [Proof of Claim 4 assuming Lemma 6]

Fix α with $0 < \alpha < 1$, and choose j large enough that $2/j! \le \alpha/2$. Let $\epsilon > 0$ be small enough that $(1-\alpha/2)/(1+\epsilon)^j \ge 1-\alpha$. We apply Lemma 11 hold with $\mathcal{W}$ the set of positive integers and with $\gamma = (1+\epsilon)\beta$ (Lemma 6 guarantees that there exists $m_0 > 0$ such that the hypotheses of Lemma 11 hold with this choice of m_0, γ, and $\mathcal{W}$). It follows that for W large enough, for all i with $1 \le i \le j$ we have

$$\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,i+1}\} \le (1+\epsilon)^i \frac{\beta^i \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\}}{i!}.$$

Furthermore, writing $\kappa(F)$ for the number of connected components of F,

$$\begin{aligned} 1 &= \sum_{i=0}^{n-1} \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,i+1}\} \\ &\le (1+\epsilon)^j \sum_{i=0}^{j-1} \frac{\beta^i \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\}}{i!} + \mathbf{P}\{\kappa(\mathbf{F}) \ge j+1\}. \end{aligned} \tag{17}$$

By Lemma 8, for all $i \ge 1$,

$$\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,i+1}\} \le \frac{(n/W)^i}{i!} \le \frac{1}{i!},$$

from which it follows that for all $k \ge 1$,

$$\mathbf{P}\{\kappa(\mathbf{F}) \ge k+1\} \le \sum_{i \ge k} \frac{1}{i!} \le \frac{2}{k!}.$$

Combining the latter equation with (17) yields that

$$1 \le (1+\epsilon)^j e^{\beta} \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\} + \alpha/2,$$

so

$$\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\} \ge \frac{1-\alpha/2}{(1+\epsilon)^j e^{\beta}} \ge \frac{1-\alpha}{e^{\beta}}. \tag{18}$$

As $\alpha > 0$ was arbitrary, (18) implies that $\mathbf{P}\{\mathbf{F} \text{ is connected}\} \ge e^{-\beta+o(1)}$ which combined with Lemma 5 proves Claim 4. □

5. Proof of Lemma 6

As already noted, Lemma 6 follows immediately from Lemma 5 and Lemma 8 when $n \le \beta W$, and we hereafter focus on the case that $n > \beta W$. For the remainder of the paper let W be a positive integer and fix a positive integer weight vector $\overrightarrow{w} = (w_1, \ldots, w_n)$ with $n > \beta W$. Given a tree T with vertex set $\{1, \ldots, n\}$ and an edge $e \in T$, we denote by $s(T, e)$ the smaller weight component of $T - e$, or the component of T containing 1 if the components have equal weights. We call the components of $T - e$ *pendant subtrees of* T. Given $S \subset \{1, \ldots, n\}$, by $T|_S$ we mean the subgraph of T induced by S; the graph $T|_S$ is not necessarily connected. For $i = 1, \ldots, \lfloor W/2 \rfloor$, denote by $c(T, i)$ the quantity $|\{e \in T : w(s(T, e)) = i\}|$.

Let $K' = \sum_{T \in \mathcal{F}_{n,1}} mass(T) = K \cdot \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\}$, and let $\mathbf{T}$ be a random tree with vertex set $\{1, \ldots, n\}$ and such that

$$\mathbf{P}\{\mathbf{T} = T\} = \frac{mass(T)}{K'}. \tag{19}$$

The following lemma provides an identity that lies at the heart of our proof of Lemma 6, and of Theorem 2.

Lemma 12

$$\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,2}\} = \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\} \cdot \sum_{i=1}^{\lfloor W/2 \rfloor} \frac{\mathbf{E}c(\mathbf{T}, i)}{i(W-i)}. \tag{20}$$

PROOF. We recall the definition of the flow φ given in (4). It follows from this definition that

$$\begin{aligned}\sum_{F' \in \mathcal{F}_{n,1}} \sum_{F \in \mathcal{F}_{n,2}} \varphi(F', F) &= \sum_{T \in \mathcal{F}_{n,1}} mass(T) \cdot \sum_{e \in T} \frac{1}{s(T,e)(W - s(T,e))} \\ &= \sum_{i=1}^{\lfloor W/2 \rfloor} \frac{1}{i(W-i)} \sum_{T \in \mathcal{F}_{n,1}} mass(T) \cdot c(T, i). \end{aligned} \tag{21}$$

Furthermore, by Lemma 7 applied with $i = 1$, we have that

$$\sum_{F' \in \mathcal{F}_{n,1}} \sum_{F \in \mathcal{F}_{n,2}} \varphi(F', F) = K \cdot \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,2}\},$$

which combined with (21) yields that

$$K \cdot \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,2}\} = \sum_{i=1}^{\lfloor W/2 \rfloor} \frac{1}{i(W-i)} \sum_{T \in \mathcal{F}_{n,1}} mass(T) \cdot c(T, i). \tag{22}$$

For each $i = 1, \ldots, \lfloor W/2 \rfloor$, we then have

$$\begin{aligned}\sum_{T \in \mathcal{F}_{n,1}} mass(T) \cdot c(T, i) &= K \cdot \sum_{T \in \mathcal{F}_{n,1}} \mathbf{P}\{\mathbf{T} = T\} \cdot \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\} \cdot c(T, i) \\ &= K \cdot \mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\} \cdot \mathbf{E}c(\mathbf{T}, i), \end{aligned} \tag{23}$$

and combining (22) and (23) proves the lemma. □

Lemma 12 allows us to understand the ratio between $\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,2}\}$ and $\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,1}\}$ by studying the values $\mathbf{E}c(\mathbf{T}, i)$ for $1 \leq i \leq \lfloor W/2 \rfloor$; it is by studying these values that we shall prove Lemma 6.

5.1. *Pendant subtrees of weight one vertices*

For integers $k \geq 1$ let $c^*(\mathbf{T}, k)$ be the number of pendant subtrees of $\mathbf{T}$ consisting precisely of k weight one vertices. By definition, $c^*(\mathbf{T}, k) \leq c(\mathbf{T}, k)$, and if $w_i = 1$ for all i then $c^*(\mathbf{T}, k) = c(\mathbf{T}, k)$ for all k. Also, let

$$\rho_k = \frac{k^{k-2}}{k!e^k}$$

for $k = 1, 2, \ldots$. We shall prove:

Lemma 13 *Given any fixed positive integer k,*

$$\frac{\mathbf{E}c^*(\mathbf{T}, k)}{k(W-k)} \leq (1 + o(1))\rho_k.$$

Before proving this lemma, let us give one easy result which shows that the values ρ_k are the 'right' targets to aim for here. We consider the case when all the weights in $\overrightarrow{w}$ are 1 (corresponding to the case when the original class $\mathcal{A}$ consisted only of forests) and note that in this case we have $W = n$ and $c^*(\mathbf{T}, k) = c(\mathbf{T}, k)$ for all k.

Proposition 14 *Suppose that all the weights in $\overrightarrow{w}$ are 1. Given any fixed positive integer k,*

$$\frac{\mathbf{E}c(\mathbf{T}, k)}{k(n-k)} \to \rho_k \quad \textit{as } n \to \infty.$$

PROOF. Let k be fixed and let $n > 2k$. By Cayley's formula, double counting, and symmetry, we have

$$\begin{aligned}
\mathbf{E}c(\mathbf{T}, k) &= n^{-(n-2)} \sum_{\{\text{Trees } T \text{ on } [n]\}} \sum_{e \in T} \mathbf{1}_{[|s(T,e)|=k]} \\
&= n^{-(n-2)} \sum_{1 \leq i \neq j \leq n} \sum_{\{\text{Trees } T \text{ on } [n]\}} \mathbf{1}_{[ij \in E(T)]} \mathbf{1}_{[|s(T,ij)|=k]} \mathbf{1}_{[i \in s(T,ij)]} \\
&= n^{-(n-2)} n(n-1) \sum_{\text{Trees } T \text{ on } [n]} \mathbf{1}_{[12 \in E(T)]} \mathbf{1}_{[1 \in s(T,12)]} \mathbf{1}_{[|s(T,12)|=k]}.
\end{aligned}$$

We thus have

$$\begin{aligned}
\mathbf{E}c(\mathbf{T}, k) &= n^{-(n-2)} \; n(n-1) \binom{n-2}{k-1} k^{k-2} (n-k)^{n-k-2} \\
&\sim \frac{k^{k-2}}{(k-1)!} \; \frac{n^{k+1} \; (n-k)^{n-k-2}}{n^{n-2}} \\
&\sim \frac{k^{k-1}}{k!} e^{-k} n
\end{aligned}$$

which yields the desired result. □

We will prove Lemma 13 by comparing the probability that $\mathbf{T}$ contains many pendant subtrees composed of k weight-one vertices (i.e., that $c(\mathbf{T}, k)$ is large) to the probability that $\mathbf{T}$ contains a large pendant subtree composed entirely of weight-one vertices. We will use the following definitions throughout Section 5.

Definition 15 *Given $S \subseteq \{1, \ldots, n\}$, let P_S be the event that $\mathbf{T}$ contains a pendant subtree T with $V(T) = S$, and let L_S be the event that for all $v \in S$, v is a leaf of $\mathbf{T}$.*

PROOF. [Proof of Lemma 13] Fix a positive integer k. Given an integer m with $1 \leq m \leq \lfloor W/2k \rfloor$, let

$$\mathcal{S}_{m,k} = \{S \subseteq [n] \ : \ |S| = mk, \ w_i = 1 \ \forall i \in S\}.$$

Given $S \in \mathcal{S}_{m,k}$, let $\mathcal{Q}_{S,k}$ be the set of partitions of S into sets $S_1, \ldots, S_m$, ordered lexicographically, such that for all $i \in [m]$, $|S_i| = k$. For all $S \in \mathcal{S}_{m,k}$,

$$|\mathcal{Q}_{S,k}| = \frac{(mk)!}{m!(k!)^m}. \tag{24}$$

Next, for any $S \in \mathcal{S}_{m,k}$ and $Q \in \mathcal{Q}_{S,k}$, let

$$\mathcal{T}_Q = \{m\text{-tuples } (T_1, \ldots, T_m) : T_i \text{ is a tree and } V(T_i) = S_i \ \forall i = 1, \ldots, m\};$$

then $|\mathcal{T}_Q| = (k^{k-2})^m$. Given S and Q as above, let $L_{S,Q}$ be the event that for all $S' \in Q$, T contains a pendant subtree T with $V(T) = S'$. For a given $S \in \mathcal{S}_{m,k}$, the following procedure constructs all trees T^* containing a pendant subtree T with $V(T) = S$.

(i) Choose a tree T with $V(T) = S$.
(ii) Choose a tree T' with $V(T') = [n] - S$.
(iii) Choose $r_1 \in S$, $r_2 \in [n] - S$, and add an edge between r_1 and r_2.

A tree constructed in this manner has mass

$$mass(T) \cdot mass(T') \cdot w_{r_1} \cdot w_{r_2}.$$

The sum of $w_{r_1} w_{r_2}$ over all possible choices of r_1, r_2 is $mk(W - mk)$; since $mass(T) = 1$ for all $(mk)^{mk-2}$ possible choices of T, it follows that

$$\begin{aligned} K' \cdot \mathbf{P}\{P_S\} &= \sum_{\text{Trees } T' \text{ on } [n]-S} mass(T') \cdot (mk)^{mk-2} \cdot [(mk)(W - mk)] \\ &= (mk)^{mk-1}(W - mk) \cdot \sum_{\text{Trees } T' \text{ on } [n]-S} mass(T'). \end{aligned} \tag{25}$$

Furthermore, given S and $Q = (S_1, \ldots, S_m)$ as above, the following procedure constructs all trees contributing to the event $L_{S,Q}$ (i.e. all trees T such that if $\mathbf{T} = T$ then $L_{S,Q}$ holds)

(i) Choose a tree T' with $V(T') = [n] - S$.
(ii) Choose an m-tuple $(T_1, \ldots, T_m) \in \mathcal{T}_Q$.
(iii) For $i = 1, \ldots, m$, choose $r_1^{(i)}$ in S_i, choose $r_2^{(i)}$ in $[n] - S$, and add an edge between $r_1^{(i)}$ and $r_2^{(i)}$.

The sum of $\prod_{i=1}^k w_{r_1^{(i)}} w_{r_2^{(i)}}$ over all possible choices of $r_1^{(1)}, \ldots, r_1^{(m)}$ and $r_2^{(1)}, \ldots, r_2^{(m)}$ is $[k(W - mk)]^m$. Since $\prod_{i=1}^m mass(T_i) = 1$ for all $(k^{k-2})^m$ of the m-tuples in $\mathcal{T}_Q$, it follows that

$$K' \cdot \mathbf{P}\{L_{S,Q}\} = \sum_{\text{Trees } T' \text{ on } [n]-S} mass(T') \cdot [k^{k-1}(W - mk)]^m, \tag{26}$$

so by (24) we obtain

$$K' \cdot \sum_{Q \in \mathcal{Q}_{S,k}} \mathbf{P}\left\{L_{S,Q}\right\} = \frac{(mk)!}{m!(k!)^m} \cdot [k^{k-1}(W-mk)]^m \cdot \sum_{\text{Trees } T' \text{ on } [n]-S} mass(T'). \quad (27)$$

Combining (25) and (27), it follows that

$$\mathbf{P}\left\{P_S\right\} = \frac{(mk)^{mk-1}(W-mk)}{[k^{k-1}(W-mk)]^m} \cdot \frac{m!(k!)^m}{(mk)!} \cdot \sum_{Q \in \mathcal{Q}_{S,k}} \mathbf{P}\left\{L_{S,Q}\right\}. \quad (28)$$

Choose $m = m(W)$ so that $m(W) \to \infty$ as $W \to \infty$ and so that $m = o(W)$. We have

$$n \geq \mathbf{E}\left\{c^*(\mathbf{T}, mk)\right\} = \sum_{S \in S_{m,k}} \mathbf{P}\left\{P_S\right\},$$

which combined with (28) gives

$$\begin{aligned} n &\geq \frac{(mk)^{mk-1}(W-mk)}{[k^{k-1}(W-mk)]^m} \cdot \frac{m!(k!)^m}{(mk)!} \cdot \sum_{S \in S_{m,k}} \sum_{Q \in \mathcal{Q}_{S,k}} \mathbf{P}\left\{L_{S,Q}\right\} \\ &= \left[\frac{k!}{k^{k-1}(W-mk)}\right]^m \cdot \frac{m!(mk)^{mk-1}}{(mk)!} \cdot (W-mk) \cdot \mathbf{E}\left\{\binom{c^*(\mathbf{T},k)}{m}\right\} \\ &\geq \left(\frac{k!}{k^{k-1}W}\right)^m \cdot \frac{m!(mk)^{mk-1}}{(mk)!} \cdot W \cdot \mathbf{E}\left\{\binom{c^*(\mathbf{T},k)}{m}\right\}. \end{aligned} \quad (29)$$

Since

$$\binom{c^*(\mathbf{T},k)}{m} = \frac{c^*(\mathbf{T},k)(c^*(\mathbf{T},k)-1)\ldots(c^*(\mathbf{T},k)-m+1)}{m!} \geq \frac{(c^*(\mathbf{T},k)-m)_+^m}{m!},$$

by Jensen's inequality (29) yields

$$n \geq W\left[\frac{(\mathbf{E}c^*(\mathbf{T},i)-m)_+ \cdot k!}{k^{k-1}W}\right]^m \cdot \frac{(mk)^{mk-1}}{(mk)!},$$

which by Stirling's formula gives

$$\frac{n}{W} \geq \left(\frac{1}{\sqrt{2\pi}} + o(1)\right) \cdot \frac{1}{(mk)^{3/2}} \cdot \left[\frac{(\mathbf{E}c^*(\mathbf{T},i)-m)_+ \cdot e^k k!}{k^{k-1}W}\right]^m, \quad (30)$$

where $o(1)$ tends to zero as $m = m(W)$ tends to infinity. Since $m(W) \to \infty$ and $n \leq W$, (30) yields that

$$\mathbf{E}c^*(\mathbf{T},k) - m \leq (1+o(1))W \cdot \frac{k^{k-1}}{e^k k!}\left(\frac{\sqrt{2\pi}n(mk)^{3/2}}{W}\right)^{1/m} \leq (1+o(1))W \cdot \frac{k^{k-1}}{e^k k!},$$

so $\mathbf{E}c^*(\mathbf{T},k) \leq (1+o(1))W \cdot k^{k-1}/(e^k k!) + m = (1+o(1)W \cdot k^{k-1}/(e^k k!)$, as claimed. □

In order to prove Lemma 6, we extend the technique we used to prove Lemma 13 in two distinct directions, both relatively straightforward. In Section 5.2, we use this technique to study pendant subtrees containing vertices of weight greater than one; in Section 5.3, we prove Lemma 19, which can be viewed as a strengthening of both Lemma 13 and of Lemma 16 from Section 5.2, in the special case that the pendant subtree under consideration is a single vertex.

5.2. *Pendant subtrees containing heavier vertices*

In the last subsection we considered pendant subtrees containing only weight one vertices. In this subsection, we present and prove a similar lemma for pendant subtrees containing heavier vertices. More precisely, we prove bounds on $\mathbf{E}c(\mathbf{T}, k)$, which we shall want for large but bounded k. Recall that we have fixed a positive integer W and a positive integer weight vector $\overrightarrow{w} = (w_1, \ldots, w_n)$ with $n > \beta W$.

Lemma 16 *Given any fixed positive integer k,*

$$\mathbf{E}c(\mathbf{T}, k) \leq (1 + o(1))\frac{W}{ek}, \tag{31}$$

where $o(1) \to 0$ as $W \to \infty$.

The proof of Lemma 16 proceeds by comparing the probabilities of P_S and $\bigcup_{Q \in \mathcal{Q}_{S,k}} L_{S,Q}$. Without the restriction that all vertices in the pendant subtrees have weight one, we are unable to compare these probabilities directly and our bounds are correspondingly weaker; however, the flavour of the proof is by and large the same as that of Lemma 13.

PROOF. [Proof of Lemma 16] Fix a positive integer k as above and $\epsilon > 0$. For any m with $1 \leq m \leq \lfloor W/2k \rfloor$, let

$$\mathcal{S}_{m,k} = \{S \subseteq \{1, \ldots, n\} : w(S) = mk\}$$

(this notation differs slightly from that in Lemma 13). Given $S \in \mathcal{S}_{m,k}$, let $\mathcal{Q}_{S,k}$ be the set of partitions of S into sets $S_1, \ldots, S_m$, ordered lexicographically, such that for all $i \in \{1, \ldots, m\}$, $w(S_i) = k$; of course, $\mathcal{Q}_{S,k}$ may be empty, for example, if there is $j \in S$ with $w_j > k$. Next, for any $S \in \mathcal{S}_{m,k}$ and any $Q \in \mathcal{Q}_{S,k}$, let

$$\mathcal{R}_{S,Q} = \{\text{Trees } T' : V(T') = S, \forall S' \in Q, T'|_{S'} \text{ is connected}\},$$

and let

$$\mathcal{T}_Q = \{m\text{-tuples } (T_1, \ldots, T_m) : T_i \text{ is a tree and } V(T_i) = S_i \ \forall i = 1, \ldots, m\}.$$

Given S and Q as above, let $R_{S,Q}$ be the event that $\mathbf{T}$ contains a pendant subtree $T \in \mathcal{R}_{S,Q}$. We remark that $\bigcup_{Q \in \mathcal{Q}_{S,k}} R_{S,Q} \subseteq P_S$. Also, let $L_{S,Q}$ be the event that for all $S' \in Q$, $\mathbf{T}$ contains a pendant subtree T with $V(T) = S'$.

Given $S \in \mathcal{S}_{m,k}$ and $Q \in \mathcal{Q}_{S,k}$, we have

$$\begin{aligned} & K' \cdot \mathbf{P}\{R_{S,Q}\} \\ = & \sum_{\{\text{Trees } T \text{ on } [n]-S\}} \sum_{T' \in \mathcal{R}_{S,Q}} \sum_{u \in T, v \in T'} mass(T)mass(T') \cdot w_u w_v \\ = & \sum_{\{\text{Trees } T \text{ on } [n]-S\}} \sum_{T' \in \mathcal{R}_{S,Q}} mass(T)mass(T') \cdot km(W - km). \end{aligned} \tag{32}$$

Given S and $Q = (S_1, \ldots, S_m)$ as above, the following procedure constructs all trees $T' \in \mathcal{R}_{S,Q}$.

(i) For each $i = 1, \ldots, m$, choose a tree T_i with vertex set S_i.
(ii) Choose a labelled tree T^* with m vertices, say $v_1, \ldots, v_m$.
(iii) For each edge $e = v_i v_j$ of T^*, choose $u_e \in V(T_i)$, $v_e \in V(T_j)$, and add edge $u_e v_e$ between T_i and T_j.

The total mass of all trees $T' \in \mathcal{R}_{S,Q}$ corresponding to a given choice of $T_1, \ldots, T_m$ and T^* is

$$k^{2(m-1)} \cdot \prod_{i=1}^{m} mass(T_i);$$

the factor $k^{2(m-1)}$ is the contribution of all possible choices for edges between the trees $T_1, \ldots, T_m$. Furthermore, there are m^{m-2} ways to choose the tree T^*. We therefore have

$$\sum_{T' \in R(S,Q)} mass(T') = \sum_{(T_1,\ldots,T_m) \in \mathcal{T}_Q} m^{m-2} k^{2(m-1)} \prod_{i=1}^{m} mass(T_i). \tag{33}$$

Let

$$m(S,Q) = \sum_{\{\text{Trees } T \text{ on } [n]-S\}} \sum_{(T_1,\ldots,T_m) \in \mathcal{T}_Q} mass(T) \cdot \prod_{i=1}^{m} mass(T_i).$$

Combining (32) and (33), we have

$$\begin{aligned} K' \cdot \mathbf{P}\{R_{S,Q}\} = & \sum_{\{\text{Trees } T \text{ on } [n]-S\}} \sum_{(T_1,\ldots,T_m) \in \mathcal{T}_Q} \\ & \quad mass(T) \cdot \prod_{i=1}^{m} mass(T_i) \cdot m^{m-1} k^{2m-1} (W - mk) \\ = & \, m(S,Q) \cdot m^{m-1} k^{2m-1} (W - mk). \end{aligned} \tag{34}$$

Further, by almost the same argument we used to establish (26) in the proof of Lemma 13, it can be seen that

$$\begin{aligned} K' \cdot \mathbf{P}\{L_{S,Q}\} = & \sum_{\{\text{Trees } T \text{ on } [n]-S\}} \sum_{(T_1,\ldots,T_m) \in \mathcal{T}_Q} \\ & \quad mass(T) \cdot \prod_{i=1}^{m} mass(T_i) \cdot [k(W - mk)]^m \\ = & \, m(S,Q) \cdot [k(W - mk)]^m. \end{aligned} \tag{35}$$

Combining (34) and (35), we obtain

$$\mathbf{P}\{R_{S,Q}\} = \mathbf{P}\{L_{S,Q}\} \cdot \frac{m^{m-1} k^{2m-1} (W - mk)}{[k(W - mk)]^m} = \mathbf{P}\{L_{S,Q}\} \cdot \left[\frac{mk}{W - mk}\right]^{m-1}.$$

Since $\bigcup_{Q \in \mathcal{Q}_{S,k}} R_{S,Q} \subseteq P_S$, and $[mk/(W - mk)]^{m-1} \geq [mk/W]^{m-1}$, we therefore have

$$\mathbf{P}\{P_S\} \geq \sum_{Q \in \mathcal{Q}_{S,k}} \mathbf{P}\{R_{S,Q}\} \geq \left[\frac{mk}{W}\right]^{m-1} \cdot \sum_{Q \in \mathcal{Q}_{S,k}} \mathbf{P}\{L_{S,Q}\}. \tag{36}$$

The remainder of the proof follows from (36) just as Lemma 13 followed from (28). □

5.3. *Light pendant leaves*

We call vertex i *light* (with respect to $\overrightarrow{w}$) if $w_i \leq 5$, and let $\ell(\mathbf{T})$ be the total weight of all light leaves; we remark that $\ell(\mathbf{T}) \geq c(\mathbf{T}, 1)$, and if $w_i = 1$ for all i then $\ell(\mathbf{T}) = c(\mathbf{T}, 1)$. For the remainder of Section 5.3, let $\mathcal{S} = \{v \in [n] : w_v \leq 5\}$, and let

$$\mathcal{S}(\mathbf{T}) = \{v \in \mathcal{S} : v \text{ is a leaf of } \mathbf{T}\}.$$

For $1 \le i \le \lfloor W/5 \rfloor$, let $\mathcal{S}_i = \{S \subset \mathcal{S} : |S| = i\}$, and define $\mathcal{S}_i(\mathbf{T})$ accordingly. In the special case that we are only considering leaves (i.e. that $k = 1$) the identity in the derivation of (29) is based on the fact that

$$\sum_{S \in S_{m,1}} \sum_{Q \in \mathcal{Q}_{S,1}} \mathbf{1}_{[L_{S,Q}]} = \binom{c^*(\mathbf{T},1)}{m} = \binom{c(\mathbf{T},1)}{m}. \tag{37}$$

The analogous identity does not quite hold when we replace $c(\mathbf{T},1)$ by $\ell(\mathbf{T})$; however, it turns out that it is not too far from holding. As before, for all $S \in \mathcal{S}$, let L_S be the event that for all $v \in S$, v is a leaf of $\mathbf{T}$. We can prove:

Lemma 17 *For any positive integer $k \le w(\mathcal{S})$, if $\ell(\mathbf{T}) \ge k$ then*

$$\left(1 - \frac{10k^2}{\ell(\mathbf{T})}\right) \cdot \binom{\ell(\mathbf{T})}{k} \le \sum_{S \in \mathcal{S}_k} \mathbf{1}_{[L_S]} \prod_{j \in S} w_j \le \binom{\ell(\mathbf{T})}{k}$$

PROOF. [Proof of Lemma 17] We think of the set $\{1, \ldots, \ell(\mathbf{T})\}$ as split into blocks $B_1, \ldots, B_{|\mathcal{S}(\mathbf{T})|}$, where $B_1 = \{1, \ldots, w_1\}$ and for each $j = 2, \ldots, |\mathcal{S}(\mathbf{T})|$,

$$B_j = \left\{ \left(\sum_{m=1}^{j-1} w_m \right) + 1, \ldots, \left(\sum_{m=1}^{j-1} w_m \right) + w_j \right\}.$$

Observe that $\sum_{S \in \mathcal{S}_k} \mathbf{1}_{[L_S]} \prod_{j \in S} w_j = \sum_{S \in \mathcal{S}_k(\mathbf{T})} \prod_{j \in S} w_j$. Furthermore, the sum $\sum_{S \in \mathcal{S}_k(\mathbf{T})} \prod_{j \in S} w_j$ is simply the number of ways of choosing k elements from $[\ell(\mathbf{T})]$, each from a distinct block; the second inequality follows immediately. This also establishes the first inequality in the case $k = 1$.

To see that the first inequality holds when $k \ge 2$, we first note that since $k \le \ell(\mathbf{T})$,

$$\binom{\ell(\mathbf{T}) - 2}{k - 2} = \binom{\ell(\mathbf{T})}{k} \cdot \frac{k(k-1)}{\ell(\mathbf{T})(\ell(\mathbf{T}) - 1)} \le \binom{\ell(\mathbf{T})}{k} \frac{k^2}{\ell(\mathbf{T})^2},$$

so the number of ways of choosing k elements from $[\ell(\mathbf{T})]$ with at least two elements from some block is at most

$$|\mathcal{S}(\mathbf{T})| \cdot \binom{5}{2} \cdot \binom{\ell(\mathbf{T}) - 2}{k - 2} \le 10|\mathcal{S}(\mathbf{T})| \cdot \frac{k^2}{\ell(\mathbf{T})^2} \cdot \binom{\ell(\mathbf{T})}{k},$$

so

$$\binom{\ell(\mathbf{T}))}{k} \le \sum_{S \in \mathcal{S}_k(\mathbf{T})} \prod_{j \in S} w_j + \frac{10|\mathcal{S}(\mathbf{T})|k^2}{\ell(\mathbf{T})^2} \cdot \binom{\ell(\mathbf{T})}{k} \le \sum_{S \in \mathcal{S}_k(\mathbf{T})} \prod_{j \in S} w_j + \frac{10k^2}{\ell(\mathbf{T})} \cdot \binom{\ell(\mathbf{T})}{k},$$

and the first inequality follows by rearrangement. □

The following easy inequality will also be useful:

Lemma 18 *Given any non-empty set $S \subset \mathcal{S}$,*

$$\sum_{\text{Trees } T \text{ on } S} \text{mass}(T) \ge |S|^{|S|-2} \cdot \frac{(\prod_{v \in S} w_v)^2}{25}$$

PROOF. Let $\mathbf{T}^*$ be a uniformly random tree with vertex set S. By Jensen's inequality, we have

$$
\begin{aligned}
log\mathbf{E}\{mass(\mathbf{T}^*)\} &\geq \mathbf{E}\{\log mass(\mathbf{T}^*)\} \\
&= \mathbf{E}\left\{\sum_{v\in S} d_{\mathbf{T}^*}(v)\log w_v\right\} \\
&= \sum_{v\in S}(\log w_v)\cdot \mathbf{E}d_{\mathbf{T}^*}(v) \qquad (38)
\end{aligned}
$$

By symmetry, $\mathbf{E}\{d_{\mathbf{T}^*}(v)\} = 2 - 2/|S|$ for all $v \in S$, so as $w_v^2 \leq 25$ for all $v \in S$, (38) yields

$$
\mathbf{E}\{mass(\mathbf{T}^*)\} \geq \prod_{v\in S} w_v^{2-2/|S|} \geq \frac{\prod_{v\in S} w_v^2}{25}.
$$

As $\mathbf{E}\{mass(\mathbf{T}^*)\} = \sum_{\text{Trees } T \text{ on } S} mass(T)/|S|^{|S|-2}$, the fact follows. □

Using Lemmas 17 and 18, we can now prove the key lemma of this subsection.
Lemma 19 $\mathbf{E}\ell(\mathbf{T}) \leq \frac{(1+o(1))W}{e}$.

PROOF. Let $m(S) = \sum_{\text{Trees } T \text{ on } [n]-S} mass(T)$. For any $1 \leq i \leq \lfloor W/5 \rfloor$ and any $S \in \mathcal{S}_i$ we have

$$
\begin{aligned}
&K' \cdot \mathbf{P}\{P_S\} \\
&= \sum_{\text{Trees } T \text{ on } [n]-S} \; \sum_{\text{Trees } T' \text{ on } S} \; \sum_{u\in S, v\in[n]-S} mass(T)mass(T')w_u w_v. \qquad (39)
\end{aligned}
$$

By Lemma 18 and (39), we have

$$
\begin{aligned}
K' \cdot \mathbf{P}\{P_S\} &\geq \sum_{\{\text{Trees } T \text{ on } [n]-S\}} \; \sum_{u\in S, v\in[n]-S} mass(T)\cdot i^{i-2}\cdot \frac{\prod_{v\in S} w_v^2}{25}\cdot w_u w_v. \\
&= \sum_{\text{Trees } T \text{ on } [n]-S} mass(T) i^{i-2}[i(W-i)]\frac{\prod_{v\in S} w_v^2}{25} \\
&= m(S)\cdot [i^{i-1}(W-i)]\frac{\prod_{v\in S} w_v^2}{25}. \qquad (40)
\end{aligned}
$$

Let $\mathcal{M}_S$ be the set of vectors $M = (m_1, \ldots, m_i)$ with elements from $[n] - S$. List the elements of S in increasing order of weight as $(s_1, \ldots, s_i)$. We may uniquely describe a tree T with vertex set $\{1, \ldots, n\}$ for which the vertices of S are all leaves by first specifying a tree T' with vertex set $[n] - S$, then choosing $M = (m_1, \ldots, m_i) \in \mathcal{M}_S$ and connecting m_j to s_j, for $j = 1, \ldots, i$. Such a tree has mass $mass(T') \cdot \prod_{j=1}^{i} w_{m_j} w_{s_j}$; it follows that

$$
\begin{aligned}
K' \cdot \mathbf{P}\{L_S\} &= \sum_{\{\text{Trees } T' \text{ on } [n]-S\}} \sum_{M \in \mathcal{M}_S} mass(T') \prod_{m \in M} w_m \cdot \prod_{v \in S} w_v. \\
&= \sum_{\text{Trees } T' \text{ on } [n]-S} mass(T')(W-i)^i \prod_{v \in S} w_v \\
&= m(S) \cdot (W-i)^i \cdot \prod_{v \in S} w_v. \qquad (41)
\end{aligned}
$$

Combining (40) and (41) yields that for any $S \in \mathcal{S}_i$,

$$
\begin{aligned}
\mathbf{P}\{P_S\} &\geq \mathbf{P}\{L_S\} \cdot \frac{i^{i-1}(W-i)}{(W-i)^i} \cdot \frac{\prod_{v \in S} w_v}{25} \\
&\geq \left(\frac{i}{W}\right)^{i-1} \cdot \frac{\mathbf{E}\{\mathbf{1}_{[L_S]}\} \cdot \prod_{v \in S} w_v}{25}. \qquad (42)
\end{aligned}
$$

Choose $i = i(W)$ so that $i(W) \to \infty$ as $W \to \infty$ and so that $i \leq W^{1/4}$. We have

$$
n \geq \mathbf{E}c(\mathbf{T}, i) = \sum_{S \in \mathcal{S}_i} \mathbf{P}\{P_S\}, \qquad (43)
$$

which combined with (42) gives

$$
\begin{aligned}
25n &\geq \left(\frac{i}{W}\right)^{i-1} \cdot \sum_{S \in \mathcal{S}_i} \mathbf{E}\{\mathbf{1}_{[L_S]}\} \cdot \prod_{v \in S} w_v \\
&= \left(\frac{i}{W}\right)^{i-1} \cdot \mathbf{E}\left\{\sum_{S \in \mathcal{S}_i} \mathbf{1}_{[L_S]} \cdot \prod_{v \in S} w_v\right\}. \qquad (44)
\end{aligned}
$$

By (44) and Lemma 17, it follows that

$$
25n \geq \left(\frac{i}{W}\right)^{i-1} \cdot \mathbf{E}\left\{\binom{\ell(\mathbf{T})}{i}\left(1 - \frac{10i^2}{\ell(\mathbf{T})}\right)\right\} \qquad (45)
$$

where $\binom{x}{i} = 0$ if $x < i$. If $\ell(\mathbf{T}) \geq i^3$ then $(1 - 10i^2/\ell(T)) \geq (1 - 10/i)$. If $\ell(\mathbf{T}) < i^3$ then

$$
\binom{\ell(\mathbf{T})}{i} \leq \left(\frac{\ell(\mathbf{T})e}{i}\right)^i \leq \ell(\mathbf{T}) \cdot i^{2i-3}e^i,
$$

so

$$
\binom{\ell(\mathbf{T})}{i}\left(\frac{10i^2}{\ell(\mathbf{T})}\right) \leq 10e^i i^{2i-1}.
$$

It follows from these inequalities and from (45) that

$$
25n \geq \left(\frac{i}{W}\right)^{i-1} \cdot \mathbf{E}\left\{\binom{\ell(\mathbf{T})}{i}\left(1 - \frac{10}{i}\right) - 10e^i i^{2i-1}\right\}. \qquad (46)
$$

Expanding $\binom{\ell(\mathbf{T})}{i}$ and applying Jensen's inequality just as in the proof of Lemma 13 we obtain

$$25n \geq \left(\frac{i}{W}\right)^{i-1} \cdot \left(\frac{(\mathbf{E}\ell(\mathbf{T})-i)_+^i}{i!}\left(1-\frac{10}{i}\right) - 10e^i i^{2i-1}\right),$$
$$= (1+o(1)) \cdot \frac{W}{i} \cdot \frac{i^i}{i!} \cdot \left(\frac{(\mathbf{E}\ell(\mathbf{T})-i)_+}{W}\right)^i - \frac{10e^i i^{3i-2}}{W^{i-1}}. \tag{47}$$

Since $i \leq W^{1/4}$ and $i(W) \to \infty$, $10e^i i^{3i-2}/W^{i-1} = o(1)$, so by by Stirling's formula (47) gives

$$\frac{25n + o(1)}{W} \geq \left(\frac{1+o(1)}{\sqrt{2\pi}}\right) i^{-3/2} \left(\frac{e \cdot (\mathbf{E}\ell(\mathbf{T})-i)_+}{W}\right)^i. \tag{48}$$

Since $i(W) \to \infty$ and $n \leq W$, by rearrangement (48) yields that

$$(\mathbf{E}\ell(\mathbf{T}) - i) \leq (1+o(1))\frac{W}{e}\left(\frac{(25+o(1))\sqrt{2\pi}ni^{3/2}}{W}\right)^{1/i}$$
$$\leq (1+o(1)\frac{W}{e}, \tag{49}$$

so $\mathbf{E}\ell(\mathbf{T}) \leq i + (1+o(1))W/e = (1+o(1))W/e$ as claimed. $\square$

With Lemmas 13, 16, and 19 under our belt, we are finally ready for the proof of Lemma 6.

5.4. *Putting it all together*

PROOF. [Proof of Lemma 6] Fix $\epsilon > 0$. We recall the definition of ρ_i from Section 5.1: for integers $i \geq 1$, $\rho_i = i^{i-2}/(i!e^i)$. Let $x(k)$ be the expected number of leaves of $\mathbf{T}$ of weight k, for $k = 1, \ldots, 5$. More generally, let $x(a^i b^j)$ denote the expected number of pendant subtrees of $\mathbf{T}$ which consist of i vertices of weight a and j of weight b, that is pendant $a^i b^j$-subtrees, and so on. In particular, $x(1^i)$ is precisely $\mathbf{E}c^*(\mathbf{T}, i)$. With this notation we have:

$$\sum_{i=1}^{5} \frac{\mathbf{E}c(\mathbf{T},i)}{i(W-i)} = \frac{\mathbf{E}\ell(\mathbf{T}) - 3x(3) - 4x(4) - 5x(5)}{W-1}$$
$$+\frac{x(1^2)}{2(W-2)} + \frac{x(1^3)+x(3)}{3(W-3)} + \frac{x(1^4)+x(13)+x(4)}{4(W-4)}$$
$$+\frac{x(1^5)+x(1^2 3)+x(14)+x(5)}{5(W-5)}.$$

We now prove upper bounds on $x(13)$, $x(14)$ and $x(1^2 3)$. We may assume that $n \geq 6$. Thus for any tree T on n vertices, any leaf v of T, and any $k = 1, 2, 3$, there is at most one edge e of T such that the component of $T - e$ containing v has order k. It follows that the expected number of pendant subtrees in $\mathbf{T}$ of a given order at most 3 which have a leaf of weight j is at most $x(j)$.

Let k be 3 or 4, and let a and b be vertices with $w_a = 1$ and $w_b = k$. Let p_a be the probability that $\mathbf{T}$ has a pendant subtree on a, b with a the leaf (in $\mathbf{T}$), and let p_b be the corresponding probability with b the leaf; the latter situation is depicted in Figure 1. Then $p_a = kp_b$, and so the probability that $\mathbf{T}$ has a pendant subtree on a, b is $(k+1)p_b$. Hence $x(1k)$ is $(k+1)$ times the expected number of pendant $1k$-subtrees with the weight k vertex the leaf, and so $x(1k) \le (k+1)x(k)$. Thus $x(13) \le 4x(3)$ and $x(14) \le 5x(4)$.

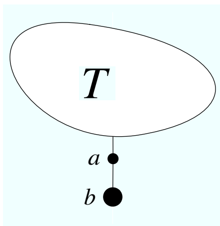

Fig. 1. The event with probability p_b.

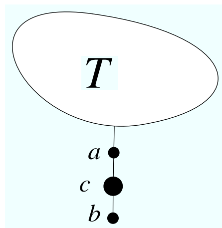

Fig. 2. The event with probability q_c.

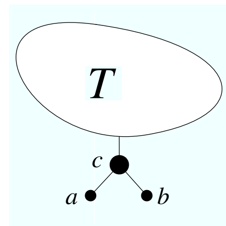

Fig. 3. The event with probability $3q_c$.

We argue similarly for $x(1^2 3)$. Now let a, b, c be distinct vertices with $w_a = w_b = 1$ and $w_c = 3$. Let q_b be the probability that $\mathbf{T}$ has a pendant subtree the path abc rooted at a, and let q_c be the probability that $\mathbf{T}$ has a pendant subtree the path acb rooted at a – then $q_c = 3q_b$ (the latter event is depicted in Figure 2). Furthermore, the probability that $\mathbf{T}$ has a pendant subtree on a, b, c and c has degree 1 is $4q_b$, as the pendant subtree containing a, b, and c may be either abc or bac, and in each case the subtree may be rooted at either a or b. Similarly, the probability that $\mathbf{T}$ has a pendant subtree on a, b, c and c has degree 2 is $4q_c$, and the probability that $\mathbf{T}$ has a pendant subtree on a, b, c and c has degree 3 is $3q_c$ (this last event is shown in Figure 3). Thus the probability that $\mathbf{T}$ has a pendant subtree on a, b, c is

$$4q_b + 4q_c + 3q_c = 25q_b = \frac{25}{4}\mathbf{P}\left\{\mathbf{T} \text{ has a pendant subtree on } a, b, c \text{ and } c \text{ is a leaf}\right\}.$$

Hence $x(1^2 3)$ is $\frac{25}{4}$ times the expected number of pendant $1^2 3$-subtrees with the weight 3 vertex a leaf, and so $x(1^2 3) \le \frac{25}{4}x(3)$.

Using these upper bounds, we see that for W sufficiently large,

$$\begin{aligned}
\sum_{i=1}^{5} \frac{\mathbf{E}c(\mathbf{T}, i)}{i(W-i)} \le \frac{\mathbf{E}\ell(\mathbf{T})}{W-1} + \sum_{i=2}^{5} \frac{x(1^i)}{i(W-i)} \\
&+x(3)\left(-\frac{3}{W-1} + \frac{1}{3(W-3)} + \frac{1}{W-4} + \frac{5}{4(W-5)}\right) \\
&+x(4)\left(-\frac{4}{W-1} + \frac{1}{4(W-4)} + \frac{1}{W-5}\right) \\
&+x(5)\left(-\frac{5}{W-1} + \frac{1}{5(W-5)}\right) \\
&\le \frac{\mathbf{E}\ell(\mathbf{T})}{W-1} + \sum_{i=2}^{5} \frac{x(1^i)}{i(W-i)}. \qquad (50)
\end{aligned}$$

Given any $\delta > 0$, for W sufficiently large, by Lemma 19 and since $\rho_1 = 1/e$ we have

$$\frac{\mathbf{E}\ell(\mathbf{T})}{W-1} \le (1+\delta)\,\rho_1.$$

Also, for each $i = 2, \ldots, 5$, for W sufficiently large, by Lemma 13 we have

$$\frac{x(1^i)}{i(W-i)} \leq (1+\delta)\,\rho_i.$$

Combining the two preceding equations and (50), we obtain

$$\sum_{i=1}^{5} \frac{\mathbf{E}c(\mathbf{T}, i)}{i(W-i)} \leq (1+\delta) \sum_{i=1}^{5} \rho_i, \tag{51}$$

for all W sufficiently large. Next, choose $\delta > 0$ small enough that $(1+\delta)^2 \leq (1+\epsilon/2)$, and choose $k_0 \geq 4/\epsilon\beta$ – then for all $W \geq 2k_0$,

$$\sum_{i=k_0}^{\lfloor W/2 \rfloor} \frac{\mathbf{E}c(\mathbf{T}_{\overrightarrow{w}})}{i(W-i)} \leq \frac{n-1}{k_0(W-k_0)} \leq \frac{2}{k_0} \leq \frac{\epsilon\beta}{2}. \tag{52}$$

Choose $W_0 \geq 2k_0$ large enough that $W_0/(W_0 - k_0) \leq 1 + \delta$ and large enough that for all weight vectors $\overrightarrow{w} = (w_1, \ldots, w_n)$ with $\sum_{i=1}^n w_i = W$ and $n > \beta W$, (51) holds and additionally, for all $i \in \{1, \ldots, k_0\}$, $\mathbf{E}c(\mathbf{T}_{\overrightarrow{w}}, i) \leq (1+\delta)W/ie$; such a choice of W_0 exists by Lemma 16. For any weight vector as above, since $\mathbf{E}c(\mathbf{T}_{\overrightarrow{w}}, i) \leq W$ for all i, by Lemma 12 and by (52), we have

$$\begin{aligned}
\frac{\mathbf{P}\left\{\mathbf{F}_{\overrightarrow{w}} \in \mathcal{F}_{n,2}\right\}}{\mathbf{P}\left\{\mathbf{F}_{\overrightarrow{w}} \in \mathcal{F}_{n,1}\right\}} &= \sum_{i=1}^{\lfloor W/2 \rfloor} \frac{\mathbf{E}c(\mathbf{T}_{\overrightarrow{w}}, i)}{i(W-i)} \\
&\leq \sum_{i=1}^{5} \frac{\mathbf{E}c(\mathbf{T}_{\overrightarrow{w}}, i)}{i(W-i)} + \sum_{i=6}^{k_0} \frac{\mathbf{E}c(\mathbf{T}_{\overrightarrow{w}}, i)}{i(W-i)} + \sum_{i=k_0+1}^{\lfloor W/2 \rfloor} \frac{W}{i(W-i)} \\
&\leq \sum_{i=1}^{5} \frac{\mathbf{E}c(\mathbf{T}_{\overrightarrow{w}}, i)}{i(W-i)} + \sum_{i=6}^{k_0} \frac{\mathbf{E}c(\mathbf{T}_{\overrightarrow{w}}, i)}{i(W-i)} + \frac{\epsilon\beta}{2}.
\end{aligned} \tag{53}$$

By our choice of W_0 and by (51) and Lemma 16, we thus have

$$\begin{aligned}
\frac{\mathbf{P}\left\{\mathbf{F}_{\overrightarrow{w}} \in \mathcal{F}_{n,2}\right\}}{\mathbf{P}\left\{\mathbf{F}_{\overrightarrow{w}} \in \mathcal{F}_{n,1}\right\}} &\leq (1+\delta)\left(\sum_{i=1}^{5} \rho_i + \sum_{i=6}^{k_0} \frac{W}{i^2 e(W-i)}\right) + \frac{\epsilon\beta}{2} \\
&< (1+\delta)^2 \left(\sum_{i=1}^{5} \rho_i + \sum_{i=1}^{k_0} \frac{1}{i^2 e}\right) + \frac{\epsilon\beta}{2} \\
&< (1+\delta)^2 \beta + \frac{\epsilon\beta}{2} \\
&\leq \beta\left(1 + \frac{\epsilon}{2}\right) + \frac{\epsilon\beta}{2} = (1+\epsilon)\beta,
\end{aligned}$$

which proves (3) and completes the proof of Lemma 6. □

6. Proof of Theorem 2

In this section we use Lemmas 11, 12, and 13 to prove Theorem 2.

Let $\varepsilon > 0$, let $\delta = \ln(2/(2-\varepsilon))$ and let $k = \lceil 2/\delta \rceil$. Let $\mathcal{W} = \{1, k+1, k+2, \ldots\}$, the set of positive integers aside less $\{2, \ldots, k\}$. Observe that given any weights $\overrightarrow{w} = (w_1, \ldots, w_n) \in \mathcal{W}^n$, any pendant subtree of weight at most k consists only of weight 1 vertices, so $\mathbf{E}c(\mathbf{T}_{\overrightarrow{w}}, i) = \mathbf{E}c^*(\mathbf{T}_{\overrightarrow{w}}, i)$ for each $i = 1, \ldots, k$. Recall from Bollobás (2001), Page 109, that $\sum_{i=1}^{\infty} \rho_i = \frac{1}{2}$. For $W \geq 2k$, by Lemma 13 and arguing as in (52), it therefore follows that

$$\begin{aligned}\sum_{i=1}^{\lfloor W/2 \rfloor} \frac{\mathbf{E}c(\mathbf{T}, i)}{i(W-i)} &\leq (1+o(1)) \sum_{i=1}^{k} \rho_i + \sum_{i=k+1}^{\lfloor W/2 \rfloor} \frac{\mathbf{E}c(\mathbf{T}, i)}{i(W-i)} \\ &\leq (1+o(1)) \sum_{i=1}^{\infty} \rho_i + \frac{2}{k} \\ &\leq \frac{1}{2} + \delta\end{aligned}$$

for W sufficiently large. Thus, for this choice of $\mathcal{W}$, by Lemma 12 the assumptions of Lemma 11 are satisfied with $\gamma = 1/2 + \delta$. By Lemma 11, therefore, for any fixed integer $j > 0$, for W sufficiently large and for all integers i with $1 \leq i \leq j$, we have

$$\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,i+1}\} \leq \left(\frac{1}{2} + \ln\left(\frac{2}{2-\varepsilon}\right)\right)^i \frac{\mathbf{P}\{\mathbf{F} \in \mathcal{F}_{n,i}\}}{i!}. \tag{54}$$

Just as Claim 4 followed from Lemma 6 (at the end of Section 4), it follows from (54) that for W sufficiently large,

$$\begin{aligned}\mathbf{P}\{\mathcal{F}_{\overrightarrow{w}} \text{ is connected}\} &\geq (1+o(1))e^{-(1/2+\delta)} \\ &\geq (1+o(1))\left(1 - \frac{\varepsilon}{2}\right)e^{-1/2} \\ &\geq (1-\varepsilon)e^{-1/2}.\end{aligned} \tag{55}$$

Finally, consider a bridge-alterable class $\mathcal{A}$ of graphs on $\{1, \ldots, W\}$ in which each graph has girth greater than k and where W is large enough that (55) holds. Then for each equivalence class $[G]$, each co-ordinate w in the corresponding weight vector $\overrightarrow{w}$ is in $\mathcal{W}$, and so by (55), $\mathbf{P}\left\{\mathcal{F}_{\overrightarrow{w}_i} \text{ is connected}\right\} \geq (1-\varepsilon)e^{-1/2}$. Theorem 2 then follows by Lemma 5.

7. Concluding remarks

Recently, a substantial amount of work has gone into counting the number of random graphs in a variety of graph classes that are addable; in some cases, this has also led to precise estimates on the probability of connectedness. Giménez and Noy (2005) have shown that for a uniformly random planar graph, the probability of connectedness is approximately 0.963253 (correct to 6 decimal places, as are all the figures in this paragraph). Similarly, Bodirsky et al. (2005) have shown that for series-parallel graphs and outerplanar graphs, the probabilities of connectedness are approximately 0.889038 and 0.862082, respectively; and Gerke et al. (2006) have shown that for random $K_{3,3}$-minor-free graphs, this probability is approximately 0.963253. For further discussion of the sizes of minor-closed graph families, see Bernardi et al. (2007); Giménez et al. (2007). All these results agree with the conjecture from McDiarmid et al. (2006) that the class of forests is asymptotically the worst possible example of an addable graph class, from the point of view of connectivity. (The conjectured value, $e^{-1/2}$, is approximately 0.606531; our bound of $e^{-\beta}$ is approximately 0.582306.) We venture the following, stronger conjecture.

Conjecture 20 *For any n and any non-empty weakly addable set $\mathcal{A}$ of graphs on $\{1, \ldots, n\}$, if $\mathbf{G}$ is a uniformly random element of $\mathcal{A}$ and $\mathbf{F}$ is a uniformly random element of $\mathcal{F}_n$, then*

$$\mathbf{P}\{\mathbf{G} \textit{ is connected}\} \geq \mathbf{P}\{\mathbf{F} \textit{ is connected}\}. \tag{56}$$

This conjecture would of course yield the conjecture in McDiarmid et al. (2006). By Lemma 5, Conjecture 20 above would imply the following weaker conjecture, which would in turn still yield the conjecture in McDiarmid et al. (2006) in the special case when $\mathcal{A}$ is bridge-alterable rather than just bridge-addable.

Conjecture 21 *For any positive integer weights $\overrightarrow{w} = (w_1, \ldots, w_n)$*

$$\mathbf{P}\{F_{\overrightarrow{w}} \textit{ is connected }\} \geq \mathbf{P}\{\mathbf{F} \textit{ is connected}\}$$

where $\mathbf{F}$ is a uniformly random element of $\mathcal{F}_n$.

References

P. Balister, B. Bollobás, and S. Gerke. Connectivity of addable graph classes. In preparation, August 2007.

O. Bernardi, M. Noy, and D. Welsh. On the growth rate of minor-closed classes of graphs. arXiv:0710.2995v1, October 2007.

M. Bodirsky, O. Giménez, M. Kang, and M. Noy. On the number of series parallel and outerplanar graphs. In *Proceedings of the European Conference on Combinatorics, Graph Theory, and Applications (EuroComb '05)*, DMTCS Proceedings Series, pages 383–388, 2005.

B. Bollobás. *Random Graphs.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, second edition, 2001.

A. Cayley. A theorem on trees. *Quart. J. Math*, 23:376–378, 1889.

S. Gerke, O. Giménez, M. Noy, and A. Weissl. On the number of $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs. Prepublicacions 727, Centre de Recerca Mathematica, December 2006.

O. Giménez and M. Noy. Asymptotic enumeration and limit laws of planar graphs. arXiv:math.CO/0501269v2, 2005.

O. Giménez, M. Noy, and J.J. Rué. Graph classes with given 3-connected components: asymptotic counting and critical phenomena. *Electronic Notes in Discrete Mathematics*, 29:521–529, 2007.

C. McDiarmid, A. Steger, and D. Welsh. Random planar graphs. *Journal of Combinatorial Theory, Series B*, 93:187–206, 2005.

C. McDiarmid, A. Steger, and D. Welsh. Random graphs from planar and other addable classes. In M. Klazar, J. Kratochvil, M. Loebl, J. Matousek, R. Thomas, and P. Valtr, editors, *Topics in Discrete Mathematics, dedicated to Jaric Nesetril on the Occasion of his 60th birthday*, volume 26 of *Algorithms and Combinatorics*, pages 231–246. Springer-Verlag, 2006.

J.W. Moon. *Counting labelled trees.* Number 1 in Canadian Mathematical Monographs. Canadian Mathematical Monographs, 1970.

A. Rényi and G. Szekeres. On the height of trees. *Journal of the Australian Mathematical Society*, 7:497–507, 1967.