The Block Connectivity of Random Trees

Andrew R. A. McGrae Michele Zito

Department of Computer Science,
University of Liverpool
Ashton Street, Liverpool L69 3BX (UK).
e-mail: (andrew|michele)@liverpool.ac.uk

Submitted: Aug 27, 2008; Accepted: Dec 29, 2008; Published: Jan 7, 2009
Mathematics Subject Classification: 05C80

Abstract

Let r, m, and n be positive integers such that $rm = n$. For each $i \in \{1, \ldots, m\}$ let $B_i = \{r(i-1)+1, \ldots, ri\}$. The r-block connectivity of a tree on n labelled vertices is the vertex connectivity of the graph obtained by collapsing the vertices in B_i, for each i, to a single (pseudo-)vertex v_i. In this paper we prove that, for fixed values of r, with $r \geq 2$, the r-block connectivity of a random tree on n vertices, for large values of n, is likely to be either $r-1$ or r, and furthermore that $r-1$ is the right answer for a constant fraction of all trees on n vertices.

1 Introduction

A random tree on n vertices, $\mathcal{T}_n$, is the typical element of the probability space defined over the n^{n-2} labelled trees on n vertices, assigning the uniform measure to each of its elements (see for instance [5] and references therein).

Let r be a positive integer and define m such that $rm = n$. Consider the partition $B_1, \ldots, B_m$ of $\{1, \ldots, n\}$ such that $B_i = \{r(i-1)+1, \ldots, ri\}$, for each $i \in \{1, \ldots, m\}$. If T_n is a tree on n vertices then its *r-reduced graph* $R_r(T_n)$ is a graph on m vertices labelled $1, 2, \ldots, m$ having an edge connecting vertices i and j for each edge in T_n connecting a vertex $u \in B_i$ to a vertex $v \in B_j$. We will also say that T_n *r-reduces to graph* G if $G = R_r(T_n)$. Note that $R_r(T_n)$, in general, may contain cycles. In what follows, $R_r(\mathcal{T}_n)$ will denote the r-reduced graph of a random tree on n vertices. In a companion paper [4] we used random trees and their reduced graphs to study particular random instances of the *empire colouring problem* (a variant of the classical planar graph colouring problem defined by Heawood [2]). Here we are interested in the vertex connectivity of these graphs, for each fixed value of $r \geq 2$. Following [1, p. 9 – 10], a non-empty graph G is connected if any two of its vertices are linked by a path and for any positive integer d, it is d-connected if $|V(G)| > d$ and $G - X$ is connected for any set $X \subseteq V(G)$ with $|X| < d$. Obviously, for each $r \geq 1$, $R_r(\mathcal{T}_n)$ is connected since the original tree is connected and any path in that graph is preserved in $R_r(\mathcal{T}_n)$. In fact, for $r \geq 2$, there may be multiple paths between any two vertices $u, v \in R(\mathcal{T}_n)$ since each consists of r vertices in the underlying tree and each of the r^2 pairs of vertices $u_i \in B_u$, $v_i \in B_v$ are connected by a path in the tree. This suggests that $R_r(\mathcal{T}_n)$ will be d-connected for some d depending on r. In the rest of this paper we make this intuition more precise. More specifically, first we argue that *asymptotically almost surely* (a.a.s.), that is with probability tending to one as n tends to infinity, the connectivity of $R_r(\mathcal{T}_n)$ cannot be more than r. Then we show that, for each fixed $r \geq 2$, the probability that $R_r(\mathcal{T}_n)$ is not r-connected is lower bounded by a quantity that approaches a positive value, dependent on r (but independent of n), as n tends to infinity. Finally we argue that, a.a.s. the connectivity of $R_r(\mathcal{T}_n)$ is at least $r-1$. The first two results follow from a careful analysis of the properties of the vertices of degree r in $R_r(\mathcal{T}_n)$. The last one, for $r \geq 3$, will be proved by estimating the number of trees whose r-reduced graph would be disconnected by the removal of a set S of (at most) $r-2$ vertices and showing that this number is $o(n^{n-2})$.

2 Connectivity Upper Bounds

Let v be a vertex in $R_r(\mathcal{T}_n)$, we call v a *funny* vertex if the degree of v is r and v is incident to a pair of edges incident to the same two distinct vertices of $R_r(\mathcal{T}_n)$ (one of them being v). From now on any such pair of edges will be called a *double edge*. In this section we study the number of vertices of degree r in $R_r(\mathcal{T}_n)$, and among those, the number of funny vertices. Notice that the existence of a vertex v of degree r (resp. a funny vertex) immediately implies that the connectivity of $R_r(\mathcal{T}_n)$ cannot be larger than r (resp. $r-1$) as in such case the removal of the neighbours of v would leave v as an isolated vertex.

2.1 The Number of Vertices of Degree r in $R_r(\mathcal{T}_n)$

In this section we will use Chebyshev's inequality to prove that, for fixed values of $r \geq 1$, $N_{r,n}$ the number of vertices of degree r in $R_r(\mathcal{T}_n)$ is well approximated by its expected value, in the sense that the probability that the (random) value of $N_{r,n}$ is far from $\mathbf{E}N_{r,n}$ tends to zero, as n tends to infinity.

Lemma 1 *Let positive integers r, and n be given, with $r \leq n$. Then*

$$\begin{aligned} \mathbf{E}N_{r,n} &= \frac{n}{r}\left(1 - \frac{r}{n}\right)^{n-2} \\ \mathbf{E}N_{r,n}^2 &= \mathbf{E}N_{r,n} + \left[\left(\frac{n}{r}\right)^2 - \frac{n}{r}\right]\left(1 - \frac{2r}{n}\right)^{n-2}. \end{aligned}$$

Proof. Let $\mathcal{E}_r(v)$ denote the event "$\deg_{R_r(\mathcal{T}_n)}(v) = r$". We can write

$$N_{r,n} = \sum_{v=1}^{m} I_{\mathcal{E}_r(v)}$$

where $I_{\mathcal{E}_r(v)}$ is the random indicator for $\mathcal{E}_r(v)$. The number of labelled trees on n vertices in which r specific vertices are leaves is exactly $(n-r)^{n-2}$ (this is easily understood, say, via the correspondence between labelled trees and Prüfer codes [5]). Thus $\Pr[I_{\mathcal{E}_r(v)} = 1] = \left(1 - \frac{r}{n}\right)^{n-2}$. The result on $\mathbf{E}N_{r,n}$ follows.

Also we may write

$$\mathbf{E}N_{r,n}^2 = \sum_{u,v=1}^{m} \Pr[I_{\mathcal{E}_r(v)} = 1, I_{\mathcal{E}_r(u)} = 1].$$

Given two distinct vertices $u, v \in V(R_r(\mathcal{T}_n))$, the number of trees reducing to graphs in which both u and v have degree r is equal to $(n-2r)^{n-2}$. Hence

$$\mathbf{E}N_{r,n}^2 = \mathbf{E}N_{r,n} + \left[\left(\frac{n}{r}\right)^2 - \frac{n}{r}\right]\left(1 - \frac{2r}{n}\right)^{n-2}.$$

■

Theorem 1 *Let r be a fixed positive integer. Then*

$$N_{r,n} = \frac{n}{r}\left(1 - \frac{r}{n}\right)^{n-2} + o(n) \quad a.a.s.$$

Proof. For positive integers r and n, with $n > r$,

$$\left(1 - \frac{2r}{n}\right)^{n-2} \leq \left(1 - \frac{r}{n}\right)^{2(n-2)}.$$

From this and Lemma 1, it is easy to see that

$$\mathrm{Var}N_{r,n} \leq \mathbf{E}N_{r,n}.$$

The result now follows readily from Chebyshev's inequality. ■

2.2 The Asymptotic Distribution of the Funny Vertices in $R_r(\mathcal{T}_n)$

The result in Theorem 1 implies that the connectivity of $R_r(\mathcal{T}_n)$ is a.a.s. at most r. In this section we will further improve on this. Using the method of moments (see [3]) we find the asymptotic distribution of the number of funny vertices in $R_r(\mathcal{T}_n)$. This, in turns, implies that a simple upper bound on the probability that $R_r(\mathcal{T}_n)$ is r-connected approaches, as n tends to infinity, a positive value smaller than one, dependent on r (but independent of n).

We start by stating a simple Lemma regarding the exponential function that will be used later.

Lemma 2 *For any real number z such that $|z| \leq \frac{4}{7}$,*

$$\mathrm{e}^{z-z^2} \leq 1 + z \leq \mathrm{e}^z.$$

Proof. The upper bound is obvious. For the other inequality, if $z > 0$, note that, since $z < 1$,

$$\begin{aligned} \mathrm{e}^z &\leq 1 + z + \frac{z^2}{2} + z^3\left(\frac{1}{3!} + \frac{1}{4!} + \dots\right) \\ &= 1 + z + \frac{z^2}{2} + z^3\left(\mathrm{e} - \left(1 + \frac{1}{1!} + \frac{1}{2!}\right)\right). \end{aligned}$$

Hence,

$$\begin{aligned} \mathrm{e}^z &\leq 1 + z + \frac{z^2}{2} + (\mathrm{e} - 2.5)z^3 \\ &\leq (1+z) + (1+z)\frac{z^2}{2}. \end{aligned}$$

If $z < 0$ we can write e^z as $\mathrm{e}^{-|z|}$ and set $x = |z|$. We can thus write

$$\mathrm{e}^{-x} \leq 1 - x + \frac{x^2}{2} - \frac{x^3}{3!} + \frac{x^4}{4!}.$$

Now, since $1 - \frac{x}{4} \geq \frac{3}{4}$ for $x < 1$, we can write

$$\frac{x^3}{3!} - \frac{x^4}{4!} = \frac{x^3}{3!}\left(1 - \frac{x}{4}\right) \geq \frac{x^3}{8}.$$

Therefore,

$$\mathrm{e}^{-x} \leq 1 - x + \frac{x^2}{2} - \frac{x^3}{8} = 1 - x + x^2\left(\frac{1}{2} - \frac{x}{8}\right).$$

Since $x \leq \frac{4}{7}$, $\left(\frac{1}{2} - \frac{x}{8}\right) \leq (1-x)$ and so

$$\mathrm{e}^{-x} \leq 1 - x + x^2(1-x).$$

In other words, since $z = -x$, we can write $\mathrm{e}^{-x} = \mathrm{e}^z \leq (1+z)(1+z^2)$. Along with the upper bound, this implies that $\mathrm{e}^z \leq (1+z)\mathrm{e}^{z^2}$. The lower bound follows by dividing both sides by e^{z^2} ∎

In what follows, $\mathrm{Po}(\lambda)$ will denote the Poisson probability law with average λ, and expressions like $S_n \stackrel{\mathrm{D}}{\to} \mathrm{Po}(\lambda)$, describe the fact that, as n tends to infinity, the random variable S_n *converges in distribution* to (a random variable with distribution) $\mathrm{Po}(\lambda)$. Let $F_{r,n}$ be the number of funny vertices in $R_r(\mathcal{T}_n)$. The following Lemma is key to our tight estimates of the t^{th} factorial moment of $F_{r,n}$, which in turn allow us to estimate the full distribution of $F_{r,n}$.

Lemma 3 *Let r be a fixed positive integer with $r \geq 2$. For any set of $t > 0$ vertices $v^1, \ldots, v^t \in V(R_r(\mathcal{T}_n))$, the probability that $v^1, \ldots, v^t$ are funny vertices is*

$$\binom{r}{2}^t r^t \frac{(n-rt)^{n-2-t}}{n^{n-2}}(1+o(1))$$

as n tends to infinity.

Proof. For a vertex in $R_r(\mathcal{T}_n)$ to have minimum degree, each of its component tree vertices must be a leaf in $\mathcal{T}_n$, the number of trees in which $v^1, \ldots, v^t$ are funny is therefore equal to the number of trees on $n - rt$ vertices

$$(n-rt)^{n-rt-2}$$

multiplied by the number of ways to add t groups of r vertices as leaves such that in each group two vertices have parents in the same empire. For each group there are $\binom{r}{2}$ choices for the two vertices that are to have parents in the same empire, and $r(n-rt)$ choices for the parent vertices. For each $1 \leq j \leq t$ the number of ways to choose the vertices within v^j and their parents is therefore

$$\binom{r}{2} r(n-rt).$$

We now must count the number of ways to choose parents for the remaining $r-2$ vertices in each group, we can give an upper bound by allowing any remaining vertex in v^j to choose any of the $(n-rt)$ vertices in the tree as its parent, giving a total of

$$\binom{r}{2} r(n-rt)^{r-1} \tag{1}$$

choices. This however, may overcount by counting trees more than once if there is more than one double edge incident to v^j. We therefore give a lower bound by counting only trees in which there is only one double edge and all other vertices have parents in different empires

$$\binom{r}{2} r(n-rt)\prod_{l=1}^{r-2}(n-rt-l) = \binom{r}{2} r(n-rt)^{r-1}(1+o(1)). \tag{2}$$

It follows from (1) and (2) that the number of ways to add the rt vertices such that $v^1, \ldots, v^t$ are funny is

$$\binom{r}{2}^t r^t (n-rt)^{rt-t}(1+o(1)),$$

the result follows by multiplying this by the number of trees on $n-rt$ vertices and dividing by n^{n-2}. ∎

We are ready to state the main result of this section.

Theorem 2 *Let r be a fixed positive integer, with $r \geq 2$. Then*

$$F_{r,n} \xrightarrow{\mathrm{D}} \mathrm{Po}\left(\binom{r}{2}\mathrm{e}^{-r}\right)$$

as n tends to infinity.

Proof. For integer $t \geq 1$, let $\mathbf{E}(F_{r,n})_t$ be the t^{th} factorial moment of $F_{r,n}$. Then

$$\mathbf{E}(F_{r,n})_t = \sum_{v^1,\ldots,v^t}^{*} \Pr[v^1,\ldots,v^t \text{ are funny vertices}],$$

where the sum is over all t-tuples of distinct vertices $v^1,\ldots,v^t \in V(R_r(\mathcal{T}_n))$. We can see that the number of ordered t-tuples is $\left(\frac{n}{r}\right)_t$, and by Lemma 3 the probability that all vertices are funny is

$$\binom{r}{2}^t r^t \frac{(n-rt)^{n-2-t}}{n^{n-2}}(1+o(1)).$$

The t^{th} factorial moment of $F_{r,n}$ is therefore

$$\mathbf{E}(F_{r,n})_t = \binom{r}{2}^t \frac{(n-rt)^{n-2-t}}{n^{n-2-t}}(1+o(1)). \tag{3}$$

Using Lemma 2, we can bound (3) above by

$$\left(\binom{r}{2}\mathrm{e}^{-r}\right)^t (1+o(1)),$$

and below by

$$\left(\binom{r}{2}\mathrm{e}^{-r-\frac{r^2}{n}}\right)^t (1+o(1)).$$

The result follows, recalling that for a random variable S_n depending on n, if $\lambda \geq 0$ is such that

$$\mathbf{E}(S_n)_t \to \lambda^t,$$

for all $t \geq 1$ as n tends to infinity, then $S_n \xrightarrow{\mathrm{D}} \mathrm{Po}(\lambda)$ (see [3, Corollary 6.8]). ∎

Let $\phi_{r,n}(k)$ denote the probability that $R_r(\mathcal{T}_n)$ contains $k \geq 0$ funny vertices. If $R_r(\mathcal{T}_n)$ contains one or more funny vertices, then the removal of the $r-1$ neighbours of one of these vertices would disconnect the graph. The probability that $R_r(\mathcal{T}_n)$ is r-connected can therefore be bounded above by $\phi_{r,n}(0)$. The following result now is a direct consequence of Theorem 2.

Corollary 1 *Let r be a fixed positive integer with $r \geq 2$, and n be a positive integer. Then the probability that the graph $R_r(\mathcal{T}_n)$ is r-connected is at most $\phi_{r,n}(0)$ and furthermore*

$$\phi_{r,n}(0) \to \exp\left\{-\binom{r}{2}\mathrm{e}^{-r}\right\}$$

as n tends to infinity.

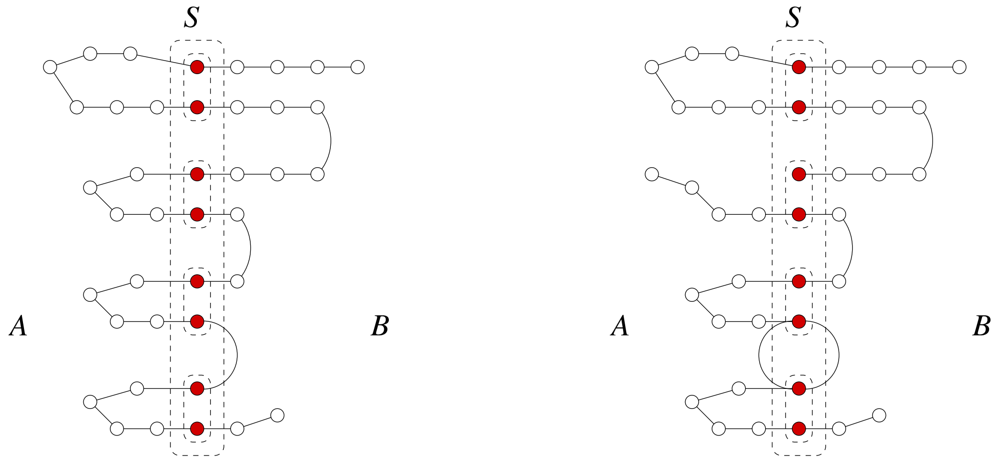

Figure 1: An example of (A, B, S)-arborescence (left) obtained from the r-reduced graph of a tree, with $r = 2$, $d = 4$, and $m = 20$. The sets of blocks S, A, and B are represented as sets of vertices. To avoid cluttering the picture only the four blocks of S have been represented as rectangles enclosing two vertices each. The vertices of all blocks in A are to the left of S, those blocks in B are to the right of S. The example on the right-hand side describes a more general case in which $F_A \cup F_B$ is not a tree.

3 Connectivity Lower Bound

Let m, r and d be fixed positive integers and set $n = mr$. Let G be a connected graph[1] on m vertices. If, for some $d < m - 1$, G is not $(d + 1)$-connected, then there exists a partition of $V(G)$ into non-empty sets A, B, and S, such that[2] $|S| = d$ and all edges in the graph are either internal to one of the blocks or join a vertex in S to a vertex in either A or B. If G is the r-reduced graph of some graph H (the definition of r-reduced graph, given for trees in Section 1, readily generalizes to arbitrary graphs on n vertices) and G is not $d + 1$-connected, the subgraph of H induced by the vertices in the blocks in $A \cup S$ (resp. $B \cup S$) will be denoted by F_A (resp. F_B) and will be such that each of its components contains at least one vertex in one of the blocks of S. Note that $F_A \cup F_B$ is not necessarily either connected or simple (see example on the right-hand side of Figure 1), however if $F_A \cup F_B$ is a tree (this is the case when $G = R_r(T_n)$), we call the pair (F_A, F_B) an (A, B, S)-*arborescence*. We obtain an upper bound on the number of trees on n vertices whose r-reduced graph would be disconnected by the removal of a set S of vertices of size d by estimating the total number of (A, B, S)-arborescences definable on a set of n vertices.

Given positive integers d, k, and n, and positive integers $c_1, \ldots, c_k$ with $\sum_{i=1}^{k} c_i = d$,

[1]As we mentioned before, the r-reduced graph of a tree is always connected.

[2]We will only consider sets S containing exactly d vertices since if there is a smaller cut-set then any set formed from this by adding more vertices to it while leaving A and B non-empty will also disconnect the graph.

let $h_{n,d}(c_1, \ldots, c_k)$ be the number of forests spanning a set V of $n + d$ vertices with k components such that, for each $i \in \{1, \ldots, k\}$, the i^{th} component contains $c_i > 0$ vertices in a given set $S \subseteq V$ of size d and x_i other vertices in $V \setminus S$. Then

$$h_{n,d}(c_1, \ldots, c_k) \leq d! \sum_{x_1,\ldots,x_k} \left(\binom{n}{x_1, \ldots, x_k} \prod_{i=1}^{k} (x_i + c_i)^{x_i+c_i-2} \right), \tag{4}$$

where the sum is over all k-tuples of non-negative integers $x_1, \ldots, x_k$ summing to n.

The total number of (A, B, S)-arborescences on a set of n vertices is at most

$$Z_{n,r,d} = \sum_{b=1}^{\frac{1}{2}\lfloor \frac{n}{r} - d \rfloor} \left(\binom{\frac{n}{r}}{\frac{n}{r} - b - d, b, d} \sum_{\mathbf{c}^A, \mathbf{c}^B} h_{n-br-dr,dr}(c_1^A, \ldots, c_{k_A}^A) h_{br,dr}(c_1^B, \ldots, c_{k_B}^B) \right), \tag{5}$$

where the inner sum is over all ways to choose two non-empty sequences of positive integers $c_1^A, \ldots, c_{k_A}^A$ and $c_1^B, \ldots, c_{k_B}^B$ adding up to dr. In the next section we will prove an upper bound on this quantity that is valid for fixed values of $r \geq 2$ and $d < r$, and sufficiently large values of n. This in turn leads to the following result, bounding the number of trees on $n = mr$ vertices whose r-reduced graph is $(r-1)$-connected. Its proof is deferred to the end of the forthcoming section.

Theorem 3 *Let r be a fixed positive integer with $r > 1$. There exists a positive constant $C \leq ((r-2)!)^2 2^{2r(r-2)} (r-1)^{(r-1)r-2} r^{(r-1)^2-1}$ such that, for any fixed $\epsilon \in \left(0, \frac{r-1}{r}\right)$, if n is sufficiently large, then the number of trees T_n for which $R_r(T_n)$ is not $(r-1)$-connected is at most $Cn^{n-3+\epsilon}$.*

From Theorem 3, our result on the typical connectivity of $R_r(\mathcal{T}_n)$ follows as a simple corollary.

Corollary 2 *For any fixed integer $r > 1$, the r-reduced graph of a random tree on n vertices is a.a.s. $(r-1)$-connected.*

Proof. By the previous Theorem, the number of trees on n vertices with r-reduced graphs that are not $(r-1)$-connected is at most $Cn^{n-3+\epsilon}$ for some constant C. The probability that a random tree will have an $(r-1)$-connected r-reduced graph is therefore at least

$$1 - \frac{C}{n^{1-\epsilon}}.$$

■

3.1 Approximations and Proof Details

To complete our proofs we need to work on $h_{n,d}(c_1, \ldots, c_k)$ first.

Lemma 4 *Let k and d be fixed positive integers. Then for any positive integer n, for all positive integers $c_1, \ldots, c_k$ with $\sum_{i=1}^{k} c_i = d$,*

$$\sum_{x_1,\ldots,x_k} \binom{n}{x_1,\ldots,x_k} \prod_{i=1}^{k} (x_i + c_i)^{x_i+c_i-2} \leq (n+d)^{n+d-2}.$$

Proof. Consider the sets of vertices $W = \{w_1, \ldots, w_n\}$, $S = \{u_1, \ldots, u_d\}$ and for $0 \leq i \leq d$ let $d_i = \sum_{j=1}^{i} c_j$. Then,

$$\binom{n}{x_1,\ldots,x_k} \prod_{i=1}^{k} (x_i + c_i)^{x_i+c_i-2}$$

counts the number of trees $T_1, \ldots, T_k$, where for $1 \leq i \leq k$, the tree T_i contains the vertices $u_{d_{i-1}+1}, \ldots, u_{d_i}$ and all vertices in W_i, given some arbitrary partition of W into k (possibly empty) subsets W_i. By summing over all $x_1, \ldots, x_k$ we consider all such partitions.

We can connect this sequence of trees by adding an edge $(u_{d_{i-1}+1}, u_{d_i+1})$ for every $1 \leq i \leq k-1$ to obtain a tree T with $n+d$ vertices. By construction, a different sequence of trees $T_1, \ldots, T_k$ leads to a different tree T . Thus we obtain that the number of different sequences of such trees $T_1, \ldots, T_k$ is less than or equal to the number of different trees T with $n+d$ vertices, which is

$$(n+d)^{n+d-2}.$$

■

Let r and d be fixed positive integers, with $r > 1$. For any positive integer n define

$$Y_{n,r,d}(a,b) = \binom{\frac{n}{r}}{a, b, d} (ar + dr)^{ar+dr-2} (br + dr)^{br+dr-2}.$$

By (4) and Lemma 4,

$$\left(\binom{\frac{n}{r}}{\frac{n}{r} - b - d, b, d} \sum_{\mathbf{c}^A, \mathbf{c}^B} h_{n-br-dr,dr}(c_1^A, \ldots, c_{k_A}^A) h_{br,dr}(c_1^B, \ldots, c_{k_B}^B) \right)$$

is at most $(d!)^2 Y_{n,r,d}(a,b)$. In what follows we will consider $Y_{n,r,d}(a,b)$ as defined on the set of positive integers a and b satisfying $a + b = \frac{n}{r} - d$.

Lemma 4 enables us to simplify our counting. $Z_{n,r,d}$ can be bounded above by $X_{n,r,d}\left(1, \frac{n}{2r} - \frac{d}{2}\right)$ where

$$X_{n,r,d}(b_1, b_2) = (d!)^2 C \sum_{b=b_1}^{b_2} Y_{n,r,d}\left(\frac{n}{r} - b - d, b\right),$$

and the positive constant C is the number of ways to choose two non-empty sequences of positive integers $\mathbf{c}^A, \mathbf{c}^B$ each summing to dr, this is 2^{2dr} by the binomial theorem. The remainder of our argument is a proof that this quantity is small compared with n^{n-2}.

To prove Theorem 3 we will split $X_{n,r,r-2}\left(1, \frac{1}{2}\lfloor \frac{n}{r} - r + 2 \rfloor\right)$ into two parts:

$$X_{n,r,r-2}\left(1, \frac{1}{2}\left\lfloor \frac{n}{r} - r + 2 \right\rfloor\right) \leq X_{n,r,r-2}(1, \lfloor n^{\epsilon} \rfloor) + X_{n,r,r-2}\left(\lfloor n^{\epsilon} \rfloor, \frac{1}{2}\left\lfloor \frac{n}{r} - r + 2 \right\rfloor\right)$$

for some $\epsilon \in (0,1)$ to be chosen later. The following lemma shows that, for sufficiently large n, $Y_{n,r,d}(a,b)$ is maximised when either a or b is as large as possible. This fact will be used in turns to prove upper bounds on the two parts of $X_{n,r,r-2}\left(1, \frac{1}{2}\lfloor \frac{n}{r} - r + 2 \rfloor\right)$.

Lemma 5 *Let n be a positive integer, and d and r be fixed positive integers with $r > 1$. Then,*

$$Y_{n,r,d}(a+1, b-1) > Y_{n,r,d}(a,b)$$

for any integer a and b with $a > b \geq 1$, such that $a + b = \frac{n}{r} - d$.

Proof. For a fixed positive d,

$$\begin{aligned}
\frac{Y_{n,r,d}(a+1,b-1)}{Y_{n,r,d}(a,b)} &= \frac{\binom{\frac{n}{r}}{a+1,b-1,d}}{\binom{\frac{n}{r}}{a,b,d}} \frac{(ar+dr+r)^{ar+dr+r-2}}{(ar+dr)^{ar+dr-2}} \frac{(br+dr-r)^{br+dr-r-2}}{(br+dr)^{br+dr-2}} \\
&= \frac{1}{a+1} \frac{(a+d+1)^{(a+1)r+dr-2}}{(a+d)^{ar+dr-2}} b \frac{(b+d-1)^{(b-1)r+dr-2}}{(b+d)^{br+dr-2}}. \qquad (6)
\end{aligned}$$

Define the function $f(x)$, for $x \geq 0$, as

$$f(x) = \frac{1}{x+1} \frac{(x+d+1)^{(x+1)r+dr-2}}{(x+d)^{xr+dr-2}},$$

then (6) is equal to

$$f(a) f(b-1)^{-1}.$$

The statement of this Lemma therefore holds if $f(x)$ is strictly monotone increasing for $x > 0$. The first derivative of $f(x)$ has the same sign as

$$r \log\left(1 + \frac{1}{x+d}\right) + \frac{1 - d - dx - x^2}{(x+d+1)(x+d)(x+1)}.$$

Using Lemma 2 we can bound this below by

$$\frac{(r-1)x^3 + (r + 2(r-1)d)x^2 + (2d-1)(r-1)x + (d^2-1)r + d}{(x+d+1)(x+d)(x+1)}.$$

For positive x, d and r, every bracketed term in the last expression is non-negative and so $f'(x) > 0$ for all $x > 0$. Hence $f(x)$ is strictly monotone increasing for $x > 0$ and the result follows. ■

Proof of Theorem 3. For $r = 2$ the result is obvious since trees are connected graphs. For $r > 2$, we give an upper bound on $Z_{n,r,d}$ and hence on the number of trees T_n for which the vertex set of $R_r(T_n)$ can be split in three sets A, B and S with $|S| = r-2$, $|B| = b$ for some $b \in \left\{1, \ldots, \frac{1}{2}\lfloor\frac{n}{r} - r + 2\rfloor\right\}$, and $|A| = \frac{n}{r} - b - (r-2)$, and such that there are no edges connecting A to B. First note that

$$X_{n,r,r-2}\left(1, \frac{1}{2}\left\lfloor\frac{n}{r} - r + 2\right\rfloor\right) \leq X_{n,r,r-2}(1, \lfloor n^\epsilon\rfloor) + X_{n,r,r-2}\left(\lfloor n^\epsilon\rfloor, \frac{1}{2}\left\lfloor\frac{n}{r} - r + 2\right\rfloor\right).$$

Lemma 5 allows us to bound $X_{n,r,r-2}(1, \lfloor n^\epsilon\rfloor)$ above by making A as large as possible in each term

$$X_{n,r,r-2}(1, \lfloor n^\epsilon\rfloor) \leq Cn^\epsilon\left(\binom{\frac{n}{r}}{\frac{n}{r} - r + 1, 1, r-2}(n-r)^{n-r-2}((r-1)r)^{(r-1)r-2}\right).$$

The multinomial coefficient $\binom{\frac{n}{r}}{\frac{n}{r}-r+1,1,r-2}$ is at most $\left(\frac{n}{r}\right)^{r-1}$, thus

$$\begin{aligned} X_{n,r,r-2}(1, \lfloor n^\epsilon\rfloor) &\leq ((r-2)!)^2 2^{2r(r-2)} n^\epsilon\left(\left(\frac{n}{r}\right)^{r-1}(n-r)^{n-r-2}((r-1)r)^{(r-1)r-2}\right) \\ &\leq C' n^{n-3+\epsilon} \end{aligned} \tag{7}$$

for some constant $0 < C' \leq ((r-2)!)^2 2^{2r(r-2)}(r-1)^{(r-1)r-2} r^{(r-1)^2-1}$.

Next we look at $X_{n,r,r-2}\left(\lfloor n^\epsilon\rfloor, \frac{1}{2}\lfloor\frac{n}{r} - r + 2\rfloor\right)$. This part of $X_{n,r,r-2}\left(1, \frac{1}{2}\lfloor\frac{n}{r} - r + 2\rfloor\right)$ still contains a large number of terms, but each term is relatively small. By Lemma 5 moving vertices from B to A will increase the size of $Y_{n,r}\left(\frac{n}{r} - b - (r-2), b, r-2\right)$. We can therefore bound $X_{n,r,r-2}\left(\lfloor n^\epsilon\rfloor, \frac{1}{2}\lfloor\frac{n}{r} - r + 2\rfloor\right)$ above by

$$\begin{aligned} &((r-2)!)^2 2^{2r(r-2)}\frac{n}{2r}\times \\ &\quad\left(\binom{\frac{n}{r}}{\frac{n}{r} - r - \lfloor n^\epsilon\rfloor + 2, \lfloor n^\epsilon\rfloor, r-2}(n - r\lfloor n^\epsilon\rfloor)^{n-r\lfloor n^\epsilon\rfloor-2}(r\lfloor n^\epsilon\rfloor + r^2 - r)^{r\lfloor n^\epsilon\rfloor + r^2 - r - 2}\right). \end{aligned}$$

In the expression above, the multinomial coefficient is at most $\left(\frac{n}{r}\right)^{\lfloor n^\epsilon\rfloor + r - 2}$, and thus we get (for n sufficiently large)

$$X_{n,r,r-2}\left(\lfloor n^\epsilon\rfloor, \frac{1}{2}\left\lfloor\frac{n}{r} - r + 2\right\rfloor\right) \leq ((r-2)!)^2 2^{2r(r-2)} r^{r^2} r^{(r-1)n^\epsilon} n^{n-2+\epsilon(r^2-r-2)-(r-1-\epsilon r)n^\epsilon}.$$

For $r \geq 2$ and $\epsilon < \frac{r-1}{r}$, this means that

$$X_{n,r,r-2}\left(\lfloor n^\epsilon\rfloor, \frac{1}{2}\left\lfloor\frac{n}{r} - r + 2\right\rfloor\right) \leq C'' n^{n-3}. \tag{8}$$

for some constant $0 < C'' \leq ((r-2)!)^2 2^{2r(r-2)} r^{r^2}$. The result follows by adding together (7) and (8). ■

Acknowledgements. The authors are endebted to the anonymous referees for pointing out a mistake in our original counting of the number of (A, B, S)-arborescences, and, in general, for the thorough reading of the manuscript that lead to a number of comments that have helped to significantly improve the quality of the presentation.

References

[1] R. Diestel. *Graph Theory* 2nd edition, Springer-Verlag, New York, 2000.

[2] P. J. Heawood. Map colour theorem, *Quarterly Journal of Pure and Applied Mathematics*, 24:332–338, 1890.

[3] S. Janson, T. Luczak, A. Ruciński. *Random Graphs*, John Wiley & Sons, 2000.

[4] A. R. A. McGrae, M. Zito. Colouring Random Empire Trees, In E. Ochmański and J. Tyszkiewicz, editors, *Mathematical Foundations of Computer Science 2008*, volume 5162 of *Lecture Notes in Computer Science*, pages 515–526. Springer-Verlag, 2008.

[5] J. W. Moon. *Counting Labelled Trees*, volume 1 of *Canadian Mathematical Monographs*, Canadian Mathematical Congress, 1970.