ABSTRACT: Fix a sequence $c = (c_1, \ldots, c_n)$ of non-negative integers with sum $n - 1$. We say a rooted tree $T$ has child sequence $c$ if it is possible to order the nodes of $T$ as $v_1, \ldots, v_n$ so that for each $1 \leq i \leq n$, $v_i$ has exactly $c_i$ children. Let $T$ be a plane tree drawn uniformly at random from among all plane trees with child sequence $c$. In this note we prove sub-Gaussian tail bounds on the height (greatest depth of any node) and width (greatest number of nodes at any single depth) of $T$. These bounds are optimal up to the constant in the exponent when $c$ satisfies $\sum_{i=1}^n c_i^2 = O(n)$; the latter can be viewed as a “finite variance” condition for the child sequence.

© 2012 Wiley Periodicals, Inc. Random Struct. Alg., 41, 253–261, 2012

Keywords: random trees; height; width; configuration model

1. INTRODUCTION

For a positive integer $n$, let $c = (c_i)_{i=1}^n$ be a sequence of non-negative integers whose sum is $n - 1$ (we call such a sequence a child sequence). In this paper we consider the random plane tree $\mathcal{T}_c$, chosen uniformly at random from the set of plane trees (rooted ordered trees) $T$ with $n$ nodes for which, for some ordering $v_1, \ldots, v_n$ of the nodes of $T$, node $v_i$ has $c_i$ children, for each $i \in [n] = \{1, \ldots, n\}$. The number of such trees is

$$\frac{1}{n!} \frac{n!}{\prod_{k=1}^n n_k!},$$

(1)
are the following sub-Gaussian tail bounds on the width and height of $T$:

For a given tree $T$ and call the invariants $F$ or any $n$ from

Theorem 1.

Theorem 1. For any $n \geq 1$ and all $m \geq 1$,

$$\Pr \left( w(T_c) \geq m + 2 \right) \leq 3e^{-m^2/(152|c|^2)} \quad \text{and} \quad \Pr \left( h(T_c) \geq m \right) \leq 7e^{-m^2/(608|c|^2|\epsilon|^2)}.$$

Remarks.

- When $|c|^2 = O(n)$, this result is best possible up to the constants in the exponents. For the width, this follows from a connection, explained below, between the width and the fluctuations of random lattice paths. For the height, consider for example the special case where $n = 2m + 1$ and $c$ consists of $m$ twos and $m + 1$ zeros. Then $T_c$ is a uniformly random binary plane tree, and in this case our bound (and the fact that it is tight) is a well-known result of Flajolet, Gao, Odlyzko and Richmond [5, Theorem 1.3].
- A result related to Theorem 1 appears in [1]. Fix a random variable $B$ with $\mathbb{E}B = 1$ and $0 < \text{Var}[B] < \infty$. Then, for $n \geq 1$, let $T_n$ be a Galton–Watson tree with offspring distribution $B$, conditioned to have total progeny $n$. [1, Theorems 1.1 and 1.2] then state that, for some $\epsilon > 0$ not depending on $n$, $\Pr \{ w(T_n) > t \} \leq \exp(-\epsilon t^2/n)$, and if additionally $\text{Var}[B] > 0$ then $\Pr \{ h(T_n) > t \} \leq \exp(-\epsilon t^2/n)$. The requirement that $\text{Var}[B] > 0$ excludes the degenerate case where $\Pr \{ B = 1 \} = 1$. Note that if $B_1, \ldots, B_n$ are independent copies of $B$ then $\mathbb{E} \left[ \sum_{i=1}^n B_i^2 \right] = n \cdot (\text{Var}[B] + 1)$, and so the finite variance condition would roughly correspond in our setting to a requirement that $|c|^2 = O(n)$. Now temporarily write $C_1, \ldots, C_n$ for the numbers of children of the nodes of $T_n$ (note that $C_1, \ldots, C_n$ are exchangeable, but are not independent—their sum is $n - 1$—and are not distributed as $B$). We conjecture that in fact $\Pr \left[ \sum_{i=1}^n C_i^2 - n \cdot (\text{Var}[B] + 1) \right]$ has Gaussian tails. A proof of this would show that the main results of [1] can be recovered from Theorem 1.
- • In forthcoming work [3], Broutin and Marckert use the tail bound for the height in Theorem 1 as an ingredient in proving that, under suitable conditions on the child sequence, $c$, the tree $T_c$ converges in distribution to a Brownian continuum random tree after suitable rescaling.
- • In [2], a bound very similar to the second bound of Theorem 1 was required, for the height of a uniformly random labelled rooted tree of a fixed size. This bound was a key step in establishing the existence of a distributional Gromov–Hausdorff scaling limit.
for the sequence of rescaled components of a critical Erdős–Rényi random graph $G_{n,p}$ when $p = p(n)$ is in the critical window $p - 1/n = O(n^{-2/3})$. The results of this paper may thus be seen as a step towards establishing that the same scaling limit obtains for the sequence of components of a critical random graph with a given degree sequence [6, 7, 11]. This is a line of enquiry that we shall pursue in a future paper.

- The appearance of the term $1/\epsilon$ in the bound on the height is necessary. For example, the sequence $c = (1, 1, \ldots, 1, 0)$ corresponds to a unique rooted plane tree, of height $n$. (For technical convenience, we exclude this unique, degenerate case from consideration for the remainder of the paper. Note that for any other child sequence $c$, we have $|c| \geq n$.) More generally, given $c$, define the one-reduced sequence $c^*$, obtained by suppressing all entries of $c$ which are equal to one. If $c$ has $k$ entries which are equal to one, then a tree with distribution $T_c$ can then be generated from the tree $T_c^*$ by repeatedly choosing a node $v$ uniformly at random, then subdividing the edge between $v$ and its parent (or, if $v$ happens to be the root, then adding a new node above $v$ and rerooting at this new node). Under this construction, each edge in $T_c^*$ is subdivided $k/(n - k)$ times on average, and this is precisely the factor encoded by $1/\epsilon$.

The remainder of the note is devoted to proving Theorem 1. We first briefly describe a family of bijective correspondences between rooted plane trees and certain lattice paths; these correspondences allow us to prove bounds for the height and width by studying the fluctuations of a certain martingale. We accomplish this bounding by using a martingale concentration result of McDiarmid [9], which appears as Theorem 2, below. This immediately yields the first bound in Theorem 1; the second requires a little further thought, and the use of a negative association result of Dubhashi [4]. Forthwith the details.

2. THE ULAM–HARRIS TREE, BREADTH-FIRST SEARCH, LEX-DFS AND REV-DFS

Below is a brief review of some basic connections between rooted plane trees and lattice paths. An excellent and detailed reference, with proofs, is [8]. The Ulam–Harris tree $U$ is the tree with root $\emptyset$ whose non-root nodes correspond to finite sequences of positive integers $v_1 \ldots v_k$, with $v_1 \ldots v_k$ having parent $v_1 \ldots v_{k-1}$ and children $\{v_1 \ldots v_k i : i \in \{1, 2, \ldots\}\}$. For a node $v$ of $U$ we think of $vi$ as the $i$th child of $v$. Any rooted plane tree $T$ in which all nodes have at most countably many children can be viewed as a subtree of $U$ by sending the root of $T$ to the root $\emptyset$ of $U$ and using the ordering of children in $T$ to recursively define an embedding of $T$ into $U$.

Having viewed $T$ as a subtree of $U$, we now define three orderings on the nodes of $T$:

1. breadth-first search (or BFS) order lists the nodes of $T$ in increasing order of depth, and for nodes of the same depth, in lexicographic order (so, for example, node 2, 3 would appear before 3, 1 but after 1, 7);
2. lexicographic depth-first search (or lex-DFS) order lists the nodes of $T$ in lexicographic order;
3. reverse lexicographic depth-first search (or rev-DFS) is most easily described informally. Let $T^*$ be the mirror-image of $T$, and list the nodes of $T$ in the order they (their mirror images) appear in a lexicographic depth-first search of $T^*$.

The use of rev-DFS to bound heights of trees was introduced in [1]. Each of these orders have the property that when a node $v$ appears in the order, its parent in $T$ has already appeared.
For such orders, we may define a queue process, as follows. Given the order \( u_1, \ldots, u_n \) of the nodes of \( T \). Let \( Q_0 = 1 \) and, for \( 1 \leq i \leq n \), let \( Q_i = Q_{i-1} + c_{u_i} \), where \( c_{u_i} \) is the number of children of \( u_i \) in \( T \). Then \( Q_i \) is the number of nodes \( u \) of the tree whose parent is among \( u_1, \ldots, u_i \) but who are not themselves among \( u_1, \ldots, u_i \). We will thus always have \( Q_i > 0 \) for \( i < n \) and \( Q_n = 0 \). We write \( (Q^i(T))_{i=1}^n \) for the queue process on the BFS order of \( T \), and likewise define \( (Q^b(T))_{i=1}^n \) and \( (Q^l(T))_{i=1}^n \) for the lex-DFS and rev-DFS orders, respectively.

Given the tree \( T \), the preceding three processes are uniquely specified. Conversely, given any of the three sequences \( (Q^b(T))_{i=1}^n \), \( x \in b, l, r \), the tree \( T \) can be recovered. For each \( x \in b, l, r \), this provides a bijection between rooted plane trees with \( n \) nodes, on the one hand, and child sequences \( (c_i)_{i=1}^n \) with \( \sum_{1 \leq i \leq k} (c_i - 1) \geq 0 \) for all \( 1 \leq k < n \). Call such sequences tree sequences.

Given a sequence \( c = (c_i)_{i=1}^n \), set \( S_0 = 1 \) and \( S_i = S_i(c) = 1 + \sum_{j=1}^i (c_j - 1) \) for \( i \in [n] \). Also, given a permutation \( \sigma : [n] \to [n] \), write \( \sigma(c) \) for the sequence \( (c_{\sigma(i)})_{i=1}^n \). For a given sequence \( c = (c_1, \ldots, c_n) \) of non-negative integers with sum \( n - 1 \), there is a unique cyclic permutation \( \sigma = \sigma_c : [n] \to [n] \) for which the sequence of partial sums \( \sigma(c) \) forms a tree sequence. (This fact yields a one-line proof of (1), above, by considering the number of permutations leaving \( c \) unchanged.) To be precise, \( \sigma \) is the cyclic permutation sending \( k \) to \( n \), where \( k \) is the least index at which the sequence \( (S_i(c))_{i=0}^n \) achieves its minimum overall value. Fix \( x \in \{b, l, r\} \) and write \( T^x(c) \) for the tree \( T \) corresponding to \( \sigma_x(c) \) under the \( x \)-bijection. It follows that letting \( \tau \) be a uniformly random permutation of \([n]\), the tree \( T^\tau(x(c)) \) is a uniformly random tree with child sequence \( c \). Conversely, if \( T \) is a uniformly random tree, then \( (Q^b(T))_{i=1}^n \) is distributed as \( (S_i(\sigma_C(C(c))))_{i=0}^n \), where \( C = \tau(x) \) and \( \tau \) is a uniformly random permutation, independent of \( C \). In particular, \( \sigma_C(C) = \sigma_x(c) \).

### 3. Martingales for the Queue Processes

We will use a martingale inequality that can be found in [9]. Let \( \{X_i\}_{i=0}^n \) be a bounded martingale adapted to a filtration \( \{\mathcal{F}_i\}_{i=0}^n \). Let \( V = \sum_{i=0}^n \text{Var}[X_{i+1}|\mathcal{F}_i] \), where

\[
\text{Var}[X_{i+1}|\mathcal{F}_i] := E[(X_{i+1} - X_i)^2|\mathcal{F}_i] = E[X_{i+1}^2|\mathcal{F}_i] - X_i^2
\]

is the predictable quadratic variation of \( X_{i+1} \). Define

\[
v = \text{ess sup} V, \quad \text{and} \quad b = \max_{0 \leq i \leq n-1} \text{ess sup}(X_{i+1} - X_i|\mathcal{F}_i).
\]

Then we have the following bound.

**Theorem 2** ([9], Theorem 3.15). For any \( t \geq 0 \),

\[
P\left( \max_{0 \leq i \leq n} X_i \geq t \right) \leq \exp\left( -\frac{t^2}{2v(1 + bt/(3v))} \right).
\]

In [9], this result is stated for \( P[X_n \geq t] \) rather than for the supremum of the \( X_i \) as above. However, as noted by McDiarmid, the proof is based on bounding \( E[e^{hX_n}] \) for suitably chosen \( h > 0 \). Since \( e^{hX_i}, 0 \leq i \leq n \) is a submartingale, the version for the supremum in fact holds by a simple application of Doob’s inequality.
Now, fix a child sequence \( c = (c_1, \ldots, c_n) \) and \( x \in \{b, l, r\} \). Let \( \tau : [n] \to [n] \) be a uniformly random permutation, and write \( C = (C_1, \ldots, C_n) = (\tau(c)) \).

For \( 0 \leq k \leq n - 1 \) let \( n_k^i = \# \{ j : C_j = k \} = n_i(c) \). For \( i > 0 \) and \( 0 \leq k \leq n - 1 \), define

\[
n_k^i = n_k^i (C) = \begin{cases} n_{k-1}^i & \text{if } C_i \neq k, \\ n_{k-1}^i - 1 & \text{if } C_i = k, \end{cases}
\]

and note that \( n_k^i = \# \{ j > i : C_j = k \} \). Then for all \( 0 \leq i \leq n \), \( \sum_{k=0}^{n-1} n_k^i = n - i \).

Also, for each \( 1 \leq i \leq n \), there is a single \( k \) with \( n_k^i \neq n_k^i - 1 \), and furthermore, for this \( k \),

\[
S_i(C) = S_{i-1}(C) + k - 1. \quad \text{Thus, for all } 0 \leq i \leq n,
\]

\[
\sum_{k=0}^{n-1} kn_k^i + S_i = \sum_{k=0}^{n-1} kn_k^i - i = n - 1 - i.
\]

Writing \( \mathcal{F}_i \) for the sigma-field generated by \( S_0, \ldots, S_i \), we then have

\[
\mathbb{E}[S_{i+1} | \mathcal{F}_i] = S_i + \sum_{j=0}^{n-1} (j-1) \frac{n_j^i}{n-i} = S_i - S_i + 1 \quad \text{and}
\]

\[
\mathbb{E}\left[ S_{i+1}^2 | \mathcal{F}_i \right] = \sum_{k=0}^{n-1} (S_k + (k-1))^2 \frac{n_k^i}{n-i} = S_i^2 - 2S_i(S_i+1) \frac{n_i^i}{n-i} + \sum_{k=0}^{n-1} (k-1)^2 \frac{n_k^i}{n-i}.
\]

At this point it would be natural to turn to the study of the martingale whose value at time \( i \) is \( S_i + \sum_{j=0}^{i-1} (S_j + 1)/(n-j) \), or in other words to subtract off the predictable part. However, this would require us to separately bound the sums of the \( (S_j + 1)/(n-j) \), and a more direct route is to simply bound these summands directly. From the preceding equations, for \( i < n \) we have

\[
\mathbb{E}\left[ \frac{S_{i+1} + 1}{n-(i+1)} \right| \mathcal{F}_i \right] = \frac{S_i + 1}{n-(i+1)} - \frac{S_i + 1}{(n-i)(n-(i+1))} = \frac{S_i + 1}{n-i}.
\]

Here we take \( 0/0 = 1 \) by convention to deal with the term \( i = n - 1 \). Thus, \( M_i = (S_i + 1)/(n-i) \) is an \( \mathcal{F}_i \)-martingale. Since \( S_{i+1} \geq S_i - 1 \) for each \( i < n \), for \( i < [n/2] \) we have

\[
M_{i+1} = \frac{S_{i+1} + 1}{n-(i+1)} = \frac{S_i + 1}{n-(i+1)} - \frac{2}{n}
\]

\[
= \frac{S_i + 1}{n-i} - \frac{S_i + 1}{(n-i)(n-(i+1))} - \frac{2}{n} \geq \frac{S_i + 1}{n-i} - \frac{4}{n} = M_i - \frac{4}{n},
\]

which we will use below when applying Theorem 2. We also have

\[
\text{Var}[M_{i+1} | \mathcal{F}_i] = \mathbb{E}[M_{i+1}^2 | \mathcal{F}_i] - M_i^2
\]

\[
= \frac{1}{(n-(i+1))^2} \left( \mathbb{E}[S_{i+1}^2 | \mathcal{F}_i] + 2\mathbb{E}[S_{i+1} | \mathcal{F}_i] + 1 \right) - \left( \frac{S_i + 1}{n-i} \right)^2.
\]
Writing \( a = \sum_{i=1}^{\lfloor n/2 \rfloor} c_i^2 / n \), for \( i < \lfloor n/2 \rfloor \) we obtain the bound

\[
\text{Var}[M_{i+1} | \mathcal{F}_i] \leq \frac{8a}{n^2},
\]

and so

\[
\sum_{i=1}^{\lfloor n/2 \rfloor} \text{Var}[M_i | \mathcal{F}_{i-1}] \leq \frac{8a}{n}.
\]

It follows by applying Theorem 2 to \( \{-M_i\}_{i=0}^{\lfloor n/2 \rfloor} \) that for all \( t \geq 0 \),

\[
P\left\{ \min_{0 \leq i \leq \lfloor n/2 \rfloor} S_i \leq -(t + 1) \right\} \leq P\left\{ \min_{0 \leq i \leq \lfloor n/2 \rfloor} \frac{S_i + 1}{n - i} \leq -\frac{t}{n} \right\} \leq \exp\left(-\frac{t^2}{164an + 8t/3}\right). \tag{2}
\]

Recall that \( \sigma_c \) is the unique cyclic permutation \( \sigma \) which makes \( \sigma(c) \) a tree sequence. We are now prepared for our principal bound on the fluctuations of \( \{S_i(\sigma(c)), 0 \leq i \leq n\} \).

**Theorem 3.** For any non-negative integer \( m \),

\[
P\left\{ \max_{0 \leq i \leq n} S_i(\sigma_c(c)) \geq m + 2 \right\} \leq 3 \exp\left(-\frac{m^2}{152|c|^2}\right).
\]

**Proof.** We will prove the theorem by showing that

\[
P\left\{ \max_{0 \leq i \leq n} S_i(\sigma_c(c)) \geq m + 2 \right\} \leq 3P\left\{ \min_{0 \leq i \leq \lfloor n/2 \rfloor} S_i \leq -(m/3 + 1) \right\}. \tag{3}
\]

Assuming that this bound holds, by (2) we then have that

\[
P\left\{ \max_{0 \leq i \leq n} S_i(\sigma_c(c)) \geq m + 2 \right\} \leq 3 \exp\left(-\frac{m^2}{144an + 8m}\right).
\]

But \( 8m < 8(n - 3) < 8|c|^2 \) and \( 144an = 144|c|^2 \), and the result follows.

It therefore remains to prove (3). Since \( \sigma_c(c) = \sigma_c(C) \), it suffices to bound the probability \( P\{\max_{0 \leq i \leq n} S_i(\sigma(C)) \geq m + 2\} \), which is what we shall do. Also, for \( m \geq n - 3 \) the event under consideration can never occur, so we may and shall assume \( m < n - 3 \). Finally, for this proof, by \( S_i \) we mean \( S_i(C) \) unless an argument is provided.

First note that if \( \max_{0 \leq i \leq n} S_i(\sigma(C)) = m + 2 \), then

\[
\max_{0 \leq i \leq n} S_i - \min_{0 \leq i \leq n} S_i = m + 3. \tag{4}
\]
If (a) does not occur then \( \{S_i, 0 \leq i \leq n\} \) drops in value significantly. Let \( m_0 = \max_{0 \leq i \leq \lfloor n/2 \rfloor} S_i \), and consider the following two events.

a. \( \min_{0 \leq i \leq \lfloor n/2 \rfloor} S_i \leq -(m + 3)/3 \)

b. \( S_{\lfloor n/2 \rfloor} \leq m_0 - (m + 3)/3 \).

If (a) does not occur then \( \{S_0, S_1, \ldots, S_{\lfloor n/2 \rfloor}\} \subset (- (m + 3)/3, m_0) \). Thus, if neither (a) nor (b) occur then for (4) to hold one of the following must take place.

c. \( m_0 > 2(m + 1)/3 \),

d. \( \max_{\lfloor n/2 \rfloor < i \leq n} S_i > 2(m + 3)/3 \),

e. \( \min_{\lfloor n/2 \rfloor < i \leq n} S_i < m_0 - (m + 3) \).

If (b) does not occur but (c) occurs then \( S_n - S_{\lfloor n/2 \rfloor} < -(m + 3)/3 \). If (d) occurs then since \( S_n = -1, S_n - \max_{\lfloor n/2 \rfloor < i \leq n} S_i < -2(m + 3)/3 \).

Now note that if either (a) or (b) occurs then

\[
\min_{0 \leq i \leq \lfloor n/2 \rfloor} (S_{\lfloor n/2 \rfloor} - S_{\lfloor n/2 \rfloor - i}) \leq -(m + 3)/3,
\]

and if (b) does not occur but one of (c), (d) does then \( S_n - \max_{\lfloor n/2 \rfloor < i \leq n} S_i < -(m + 3)/3 \), and so

\[
\min_{0 \leq i \leq \lfloor n/2 \rfloor} (S_n - S_{n-i}) < -(m + 3)/3.
\]

Finally, if (b) does not occur but (e) occurs then since \( S_{\lfloor n/2 \rfloor} > m_0 - (m + 3)/3 \), we have

\[
\min_{\lfloor n/2 \rfloor < i \leq n} (S_i - S_{\lfloor n/2 \rfloor}) < -2(m + 3)/3.
\]

Since \( (S_{\lfloor n/2 \rfloor} - S_{\lfloor n/2 \rfloor - i}, 0 \leq i \leq \lfloor n/2 \rfloor) \) has the same distribution as \( (S_i, 0 \leq i \leq \lfloor n/2 \rfloor) \), and \( (S_i - S_{\lfloor n/2 \rfloor}, \lfloor n/2 \rfloor < i \leq n) \) and \( (S_n - S_{n-i}, 0 \leq i \leq \lfloor n/2 \rfloor) \) both have the same distribution as \( (S_i, 0 \leq i \leq \lfloor n/2 \rfloor) \), (3) follows from the preceding three offset equations.

**4. BOUNDING THE WIDTH AND THE HEIGHT**

The bounds of Theorem 1 follow straightforwardly from Theorem 3. Let \( T_c \) be a uniformly random tree with child sequence \( c \). As noted earlier, \( (Q^V(T_c))_{v=0} \) is distributed as \( (S_i(\sigma_C(C)))_{v=0} \), where \( C = \tau(c) \) and \( \tau \) is a uniformly random permutation, independent of \( c \). Furthermore, when the breadth-first exploration has just finished exploring all the nodes at depth \( k \), the queue length is precisely the number of nodes at depth \( k - 1 \). Recalling that \( \sigma_C(c) = \sigma_C(C) \), it follows by Theorem 3 that

\[
P[w(T_c) \geq m + 2] \leq P \left\{ \max_{0 \leq i \leq n} Q^i(T_c) \geq m + 2 \right\} \leq 3 \exp \left( \frac{-m^2}{152|c|^2} \right),
\]

proving the bound for the width. (Also, if at some point the queue length is at least \( m \) then \( w(T_c) \geq m/2 \), from which the optimality of the with bound when \( |c|^2 = O(n) \) follows straightforwardly.)

*Random Structures and Algorithms* DOI 10.1002/rsa
In bounding the height, we assume that \( m \geq 6\sqrt{n} \), or else the bound follows trivially since \(|e|^2 \geq n\). First suppose that \( e \) is one-reduced (so has no entries equal to one). For any node \( u \in T \), let \( \lambda(u) \) (resp. \( \rho(u) \)) be the index of \( u \) when the nodes of \( T \) are listed in lex-DFS order (resp. rev-DFS order). Since \( e \) is one-reduced, each ancestor of \( u \) in \( T \) has at least one child that is not an ancestor of \( u \), and so either \( Q^e_{\lambda(u)}(T) \) or \( Q^e_{\rho(u)}(T) \) is at least \(|u|/2\). It follows that when \( e \) is one-reduced,

\[
P\{h(T_c) \geq m + 4\} \leq P\left\{ \max_{0 \leq i \leq n} Q^e_i(T_c) \geq \lfloor m/2 \rfloor + 2 \right\} + P\left\{ \max_{0 \leq i \leq n} Q^e_i(T_c) \geq \lceil m/2 \rceil + 2 \right\} \\
\leq 6 \exp \left( -\frac{m^2}{602|e|^2} \right),
\]

proving the bound in this case.

More generally, write \( e^* \) for the one-reduced version of \( e \), obtained from \( e \) by removing all entries that are equal to one, and let \( n^* \) be the length (number of elements) of \( e^* \). Also, write \( T^* \) for the tree obtained from \( T \), by replacing each maximal path whose internal nodes all have exactly one child, by a single edge. List the edges of \( T^* \) according to some fixed rule as \( (e_1, \ldots, e_{n^*-1}) \). Note that we always have \( n^* \leq n - 1 \). Each edge \( e_i \) corresponds to some path in \( T \), and we write \( s_i \) for the number of internal nodes of this path (i.e. the total number of nodes, minus two). Then \( T^* \) is distributed as a uniformly random tree with child sequence \( e^* \), and, independently of \( T^* \), \( (s_1, \ldots, s_{n^*-1}) \) is a uniformly random element of the set of vectors of non-negative integers of length \( n^* - 1 \) with sum \( n - n^* - 1 \). From Theorem 3 we thus have

\[
P\{h(T^*) \geq m + 4\} \leq 6 \exp \left( -\frac{m^2}{608|e^*|^2} \right) \leq 6 \exp \left( -\frac{m^2}{608|e|^2} \right). \tag{5}
\]

If \( n^* \geq n - \sqrt{n} \) then \( h(T_c) \leq h(T^*) + \sqrt{n} \), and in this case the required bound follows (recall that we have shown we may assume \( m \geq 6\sqrt{n} \)). In what follows we thus assume \( n^* < n - \sqrt{n} \).

By Proposition 5 of [4], the entries of \( (s_1, \ldots, s_{n^*-1}) \) are negatively correlated and thus standard Chernoff bounds apply to any restricted sum of elements of \( (s_1, \ldots, s_{n^*-1}) \). In particular, for any node \( v \) of \( T^* \),

\[ J_v = \{ i : e_i \text{ is an edge of the path from } v \text{ to the root of } T^* \}. \]

We always have \( J_v \leq h(T^*) - 1 \), and thus by a Chernoff bound (e.g., [9], Theorem 2.2),

\[
P\left\{ S_{J_v} \geq (1 + x)(m + 2) \frac{n - n^* - 1}{n^* - 1} \min_{h(T^*) \leq m + 3} \right\} \leq \exp \left( -2x^2(m + 2)\frac{n - n^* - 1}{n^* - 1} \right).
\]

To get a clean final bound, we choose \( x \) so that \((1 + x)(m + 2) = 2m\). It then follows by a union bound that

\[
P\{ \exists v \in T^* : S_{J_v} \geq 2m \frac{n - n^* - 1}{n^* - 1} \min_{h(T^*) \leq m + 3} \right\} \leq \exp \left( \left( \log(n^* - 1) - 2(m - 2)\frac{n - n^* - 1}{n^* - 1} \right) \right).
\]

Random Structures and Algorithms DOI 10.1002/rsa
\begin{align*}
\leq & \exp\left(-(m-2)^2 \frac{n-n^*-1}{n^*-1}\right) \\
\leq & \exp\left(-\frac{m^2(n-n^*-1)}{9|c|^2}\right)
\end{align*}

the second inequality holding since \((m-2)^2 \geq n^*-1\) and \(n-n^*-1 \geq \sqrt{n}-1 \geq \log(n^*-1)\), and the third holding since \(|c|^2 \geq n > n^*-1\) and \((m-2) \geq m/3\). Since \(m+4 \leq 2m = 2m(n^*-1)/(n^*-1)\), it then follows from (5) that

\[
P\left\{h(T_c) \geq 4m \frac{n-2}{n^*-1}\right\} \leq 7 \exp\left(-\frac{m^2}{608|c|^2}\right).
\]

But \((n-2)/(n^*-1) = 1_c\), and the result follows.

ACKNOWLEDGMENT

The author thanks two anonymous referees for their careful, swift reading of this paper.

REFERENCES

[1] L. Addario-Berry, L. Devroye, and S. Janson, Sub-Gaussian tail bounds for the width and height of conditioned Galton–Watson trees, Ann Prob, (in press).
[2] L. Addario-Berry, N. Broutin, and C. Goldschmidt, The continuum limit of critical random graphs, Prob Theory Relat Fields, (in press).
[3] N. Broutin and J. F. Marckert, Asymptotics for trees with a prescribed degree sequence, and applications, arXiv:1110.5203v2 [math.PR], October 2011.
[4] D. Dubhashi and D. Ranjan, Balls and bins: A study in negative dependence, Random Struct Algorithm 13 (1998), 99–124.
[5] P. Flajolet, Z. Gao, A. Odlyzko, and B. Richmond, The distribution of heights of binary trees and other simple trees, Combinator Probab Comput 2 (1993), 145–156.
[6] H. Hatami and M. Molloy, The scaling window for a random graph with a given degree sequence, Random Struct Algorithm 41 (2012).
[7] A. Joseph, The component sizes of a critical random graph with a given degree sequence, arXiv:1012.2352v2 [math.PR], December 2011.
[8] J. F. Le Gall, Random trees and applications, Prob Survey 2 (2005), 245–311.
[9] C. McDiarmid, Concentration, In M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, and B. Reed (Editors), Probabilistic methods for algorithmic discrete mathematics, Springer Verlag, New York, 1998, pp. 195–248.
[10] J. W. Moon, Counting labelled trees, Canadian Mathematical Monographs, No. 1, Canadian Mathematical Congress, Montreal, Que., 1970, p. 113.
[11] O. Riordan, The phase transition in the configuration model, arXiv:1104.0613v1 [math.PR], April 2011.