# On the Distribution of Binary Search Trees under the Random Permutation Model

**James Allen Fill***
*Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD 21218-2692*

## ABSTRACT

We study the distribution $Q$ on the set $B_n$ of binary search trees over a linearly ordered set of $n$ records under the standard random permutation model. This distribution also arises as the stationary distribution for the move-to-root (MTR) Markov chain taking values in $B_n$ when successive requests are independent and identically distributed with each record equally likely. We identify the minimum and maximum values of the functional $Q$ and the trees achieving those values and argue that $Q$ is a crude measure of the "shape" of the tree. We study the distribution of $Q(T)$ for two choices of distribution for random trees $T$; uniform over $B_n$ and $Q$. In the latter case, we obtain a limiting normal distribution for $-\ln Q(T)$. 

## 1. INTRODUCTION AND SUMMARY

Binary search trees are commonly employed and highly efficient data structures used to search for records in a file. Our aim in this paper, briefly put, is to examine certain interesting features of the distribution of binary search trees generated under the standard random permutation model. For background and a wealth of information about binary search trees in general and the random permutation model in particular, see Chapter 2 in Mahmoud [3].

In order that the present paper be self-contained, we review briefly. A *binary*

Random Structures and Algorithms, Vol. 8, No. 1 (1996)

* Research supported by NSF Grant DMS-9311367.

*tree* is a finite tree with at most two "children" for each node and in which each child is distinguished as either left or right. Recursively expressed, a binary tree either is empty or is a node (called the root) with left and right subtrees, each of which is a binary tree.

Consider an $n$-node binary tree in which the nodes are labeled distinctly with elements of some linearly ordered set, say $[n] := \{1, 2, \ldots, n\}$. If inorder transversal of the tree (left subtree then root than right subtree, applied recursively) yields the labels in natural order, the tree is called a *binary search tree*. We shall denote the set of all such trees by $B_n$. To ensure understanding, the reader is encouraged to draw the five elements of $B_3$ (see Fig. 1 in Dobrow and Fill [1]).

We next describe the distribution $Q$ on $B_n$ arising under the *random permutation model*. (Of course, $Q$ will be different for each $n$, but we will find no need to indicate this in our notation.) If the records 1 through $n$ are arranged into a uniformly random permutation and then successively inserted in the natural fashion into an initially empty search tree, the resulting trees has, by definition, the distribution $Q$. [Note that $Q$ is *not* the uniform distribution on $B_n$. For example, for $n = 3$ the permutations $(2, 1, 3)$ and $(2, 3, 1)$ both give rise to the same tree, which thus has probability twice as large as any of the other four trees in $B_3$.] For reasons discussed in Section 2.3 in Mahoud [3], the distribution $Q$ has become the standard model in computer science for random search trees.

The author's interest in $Q$ arose in connection with the move-to-root (MTR) self-organizing scheme for dynamic maintenance of binary search trees, studied by Dobrow and Fill [1, 2]. We refer the reader to the first of these papers (specifically, Sections 1–4 therein) for a more detailed background on binary search trees, self-organizing data structures, and the move-to-root Markov chain than that given here.

The probability model for MTR is as follows. At each discrete unit of time, record $i$ is requested with probability $p_i$ ($i \in [n]$), independently of past requests. *We will consider only the uniform case $p_i \equiv 1/n$ in this paper.* Upon each request, the binary search tree is then rearranged according to the *MTR rule*: The requested record is repeatedly exchanged with its parent until it has risen to the root of the tree. Each exchange is done according to the *simple exchange* (SE) rule, under which the requested record and the record stored at the parent node reverse positions while disturbing the rest of the tree only as needed to preserve the defining properties of a binary search tree. For a more careful description of the SE rule, see Dobrow and Fill [1, Section 2].

The i.i.d. model for requests induces a Markov chain on the space of binary search trees. The stationary distribution for this MTR chain is easily described (see Dobrow and Fill [1, Section 4]). In our uniform case, this distribution is the aforementioned $Q$, and Corollary 4.2 in Dobrow and Fill [1] gives the explicit formula

$$Q(T) = 1/R(T)\,, \tag{1.1}$$

where

$$R(T) := \prod_{x \in T} |T(x)|\,; \tag{1.2}$$

here $|T(x)|$ is the number of nodes in the subtree $T(x)$ of $T$ with root $x$.

In Mahmoud [3] one finds the results of various authors giving exact and asymptotic distributions for such operating characteristics as the height of a tree with distribution $Q$. In this paper we address the most basic question of all concerning $Q$: How are the numbers $Q(T)$ themselves distributed as $T$ ranges over the set $B_n$ of binary search trees? To summarize, in Section 2 we identify those trees $T \in B_n$ minimizing or maximizing $Q(T)$. The minimum value is $1/n!$, and we show how to calculate exactly and approximate asymptotically the maximum value. In Section 3 we obtain exact and asymptotic results for the first two moments of the distribution of $L_n = -\ln Q(T)$ when $T$ is given the uniform distribution on $B_n$. In particular, we establish the weak law of large numbers for $L_n$:

$$\frac{L_n}{n} \xrightarrow{P} c_1 \tag{1.3}$$

and explicitly identify the constant $c_1 \doteq 2.03$. In this sense, almost all binary search trees $T \in B_n$ satisfy $Q(T) \approx \exp(-c_1 n)$. In Section 4 we carry out a similar analysis of $Q(T)$ when $T$ is given the distribution $Q$ on $B_n$. In this case we find a linear growth in $n$ of both the mean and variance of $L_n = -\ln Q(T)$. Thus a weak law of large numbers again obtains, but with a different constant limit. Moreover, in this case we are able to establish a limiting normal law for $L_n$.

The main results of this paper are Theorems 2.2, 2.13, 3.1, and 4.1.

The results of this paper themselves provide a final motivation for studying the functional $Q$, as the following roughly stated comments demonstrate. The results of Section 2 show that the more balanced is a tree $T$ and (recursively) its subtrees, the larger is the value of $Q(T)$. So $Q$ is a crude measure of the "shape" of a tree. The maximum value of $Q(T)$, achieved by the complete tree (see Section 2.1 for a definition), is exponentially small in $n$. As the results of Sections 3 and 4 show, so are almost all values of $Q(T)$ (albeit with a different rate constant, one that depends on which of the two models of randomness is employed). On the other hand, the minimum value, $1/n!$, decays at a faster than exponential rate. Thus it might be fair to say that most binary search trees have a rather "full" shape, like the complete tree.

*Note on logarithm notation:* We use ln for natural log, lg for binary log, and log when the base doesn't matter [as in $O(\log n)$].

## 2. EXTREME VALUES OF $Q$

### 2.1. *Preliminaries and Notation*

In order to state the main result of Section 2, we need to establish some notation and a few preliminary observations. Recall that $B_n$ is the set of binary search trees with $n$ nodes, and recall the definitions of $Q$ and $R$ given in (1.1) and (1.2), with the convention $R(\emptyset) := 1$ for the empty tree $\emptyset$. The functional $R$ takes values in $\{1, 2, \ldots\}$. There is a unique way to label any binary tree so as to produce a binary search tree, so we need not be concerned any further with the labels.

Letting $\ell(T)$ and $r(T)$ denote the left and right subtrees of $T$, respectively, the following result is obvious:

**Lemma 2.1.** *If $T \neq \emptyset$, then*

$$R(T) = |T| R(\ell(T)) R(r(T)) .$$

It is clear from Lemma 2.1 that distinction of left from right plays no role in consideration of $R(\cdot)$. By this symmetry, we may, when convenient, restrict attention from $B_n$ to

$$\bar{B}_n := \{T \in B_n : |\ell(T(x))| \geq |r(T(x))| \text{ for all } x \in T\} .$$

In this notation, it is clear that the unique minimizer of $Q$ in $\bar{B}_n$ is the tree corresponding (as in Section 1) to the reversal permutation $(n, n-1, \ldots, 1)$, i.e., the tree in which $i+1$ is the parent of $i$, $i \in [n-1]$. There are $2^{n-1}$ members of $B_n$ that minimize $Q$.

At the opposite extreme from such long, stringy trees is the complete tree, which can be defined as follows. Suppose first that $n = 2^m - 1$ for integer $m$. Call the unique tree in $B_n$ with minimum possible height ($=m-1$) the *perfect tree*. For general $n$, let $m = \lfloor \lg(n+1) \rfloor$. The *complete tree* can be obtained by attaching to the perfect tree on $2^m - 1$ nodes, and as far to the left as possible, $n - 2^m + 1$ leaves at distance $m$ from the root. In particular, if $n = 2^m - 1$, the notions of perfect tree and complete tree coincide.

Define

$$R_n^* := \min_{T \in B_n} R(T) \tag{2.1}$$

and

$$B_n^* := \{T \in B_n : R(T) = R_n^*\} , \qquad \bar{B}_n^* := \bar{B}_n \cap B_n^* .$$

Write $T_n$ for the complete tree on $n$ nodes and set $R_n := R(T_n)$.

### 2.2. The Complete Tree Maximizes $Q$

We are now prepared to state the main result of Section 2:

**Theorem 2.2.** *Up to left–right symmetries, the unique maximizer of $Q$ in $B_n$ is the complete tree $T_n$. More precisely, for every $n \geq 0$,*

$$\bar{B}_n^* = \{T_n\} .$$

We will prove Theorem 2.2 in Section 2.3 and develop the asymptotics of $R_n^* = R_n$ in Section 2.4.

*Remark 2.3.* Using Lemma 2.1 we can quickly compute the following values for $R_n$, $n = 0, 1, \ldots, 10$: 1, 1, 2, 3, 8, 15, 36, 63, 192, 405, 1080.

*Remark 2.4.* For $k \geq 1$, let $2^{\pi_k}$ be the highest power of 2 that divides $k$. Granted Theorem 2.2, there are then

$$1 \le |B_n^*| = 2^{\lfloor \lg(n+1) \rfloor - \pi_{n+1}} \le \max(n, 1)$$

trees that achieve the minimum for $R(T)$ over $T \in B_n$. Two extremes:

(a) When $n = 2^m - 1$, $|B_n^*| = 1$: the perfect tree is the *only* minimizer.

(b) When $n = 2^m$, $|B_n^*| = 2^m = n$.

### 2.3. Proof of Theorem 2.2.

The proof we give for Theorem 2.2 in this subsection may seem unduly complicated to the reader. A more direct approach would be to argue that moving a single node closer to the root of a tree decreases the value of $R$, but this is not always so. For example, let $T \in B_7$ be the tree whose left subtree is $T_5$ (the complete tree on five nodes) and let $T'$ be the tree with left subtree $T_4$ and right subtree $T_2$. Then $T'$ can be obtained from $T$ by moving a single leaf one level higher, but $R(T') = 112 > 105 = R(T)$. Instead we use a "balancing act" powered by Lemmas 2.7 and 2.8.

We preface the proof of Theorem 2.2 with three useful lemmas. The first of these is immediate from the definition (2.1) of $R_n^*$, the recursion of Lemma 2.1, and left–right symmetry.

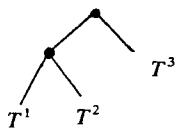
**Lemma 2.5.** *For* $n \ge 1$,

$$R_n^* = n \min\left\{ R_a^* R_b^* : a + b = n - 1 \text{ and } a \ge \left\lceil \frac{n-1}{2} \right\rceil \right\}.$$

*Remark 2.6.* Using Lemma 2.5 and Remark 2.3, it is straightforward to verify that $R_n^* = R_n$ for $0 \le n \le 10$.

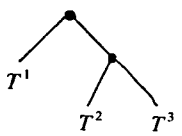
The next two lemmas describe simple operations that reduce the value of $R$. Recall the rough description of *simple exchange* given in the introduction, and see Dobrow and Fill [1, Section 2] for a more careful definition.

**Lemma 2.7.** *Consider a tree of the form*

$T =$ 

(2.2)

*and the result*

$T' =$ 

(2.3)

*of simple exchange of* root($T$) *and* root($\ell(T)$). *Then*

$$R(T') \{\lessgtr\} R(T) \text{ according as } |T^3| \{\lessgtr\} |T^1|.$$

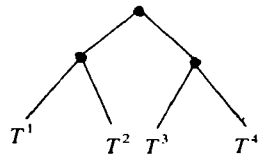
*Proof.* Write $t_i$ for $|T^i|$ and $r_i$ for $R(T^i)$ for $i = 1, 2, 3$. Then

$$R(T) = (2 + t_1 + t_2 + t_3)(1 + t_1 + t_2)r_1r_2r_3 ,$$

$$R(T') = (2 + t_1 + t_2 + t_3)(1 + t_2 + t_3)r_1r_2r_3 ,$$

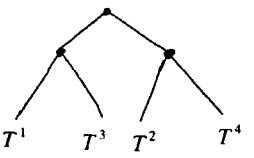
and the result follows. ■

**Lemma 2.8.** *Consider a binary tree* $T \in \bar{B}_n (n \geq 3)$ *of the form*

$T =$ 

(2.4)

*and the following modification* $T'' \in B_n$:

$T'' =$ 

(2.5)

*Then*

$$R(T'')\{\stackrel{\leq}{\geq}\}R(T) \text{ according as } |T^2|\{\stackrel{\leq}{\geq}\}|T^3| .$$

*Proof.* With notation as in the proof of Lemma 2.7,

$$R(T) = (3 + t_1 + t_2 + t_3 + t_4)(1 + t_1 + t_2)(1 + t_3 + t_4)r_1r_2r_3r_4 ,$$

$$R(T'') = (3 + t_1 + t_2 + t_3 + t_4)(1 + t_1 + t_3)(1 + t_2 + t_4)r_1r_2r_3r_4 .$$

After a little calculation, one finds that

$$R(T'')\{\stackrel{\leq}{\geq}\}R(T) \text{ according as } (t_1 - t_4)(t_2 - t_3)\{\stackrel{\leq}{\geq}\}0 .$$

If $t_2 = t_3$, then $R(T'') = R(T)$. Otherwise we claim $t_1 > t_4$, from which the other two cases follow. To $t_1 > t_4$, observe

$$2t_1 \geq t_1 + t_2 \geq t_3 + t_4 \geq 2t_4 , \tag{2.6}$$

where each of the three inequalities follows from the assumption $T \in \bar{B}_n$. If $t_1 = t_4$, then equality must hold throughout (2.6); but then $t_1 = t_2 = t_3 = t_4$, contradicting our assumption that $t_2 \neq t_3$. ■

We are now ready for the proof that $\bar{B}_n^* = \{T_n\}$.

*Proof of Theorem 2.2.* The proof is by (strong) induction on $n$. The assertion is trivially correct for $n = 0, 1, 2$. Given $n \geq 3$, we suppose that $\bar{B}_m^* = \{T_m\}$ for $0 \leq m \leq n - 1$ and prove that $\bar{B}_n^* = \{T_n\}$. To do this, we will show that if $T \in \bar{B}_n$ is *not* complete, then there exists $\tilde{T} \in B_n$ with $R(\tilde{T}) < R(T)$.

*Any* $T \in \bar{B}_n$ is of the form (2.2). If $T^3$ is empty, then (since $n \geq 3$ and $T \in \bar{B}_n$) $T^1$ is nonempty and $T'$ of (2.3) provides the required $\tilde{T}$. So we may assume that $r(T)$ is nonempty, i.e., that $T$ is of the form (2.4), whose notation we henceforth adopt.

By Lemma 2.5 and the induction hypothesis, if any of the six subtrees $\ell(T)$, $r(T)$, $T^1, T^2, T^3, T^4$ is incomplete, replacement of the subtree by the complete tree of the same size will (strictly) reduce the product $R$. So we may assume that each of the six subtrees is complete.

Put $t_i := |T^i|$ for $i = 1, 2, 3, 4$. Since $T \in \bar{B}_n$, $t_1 \geq t_2$ and $t_3 \geq t_4$. If $t_2 < t_3$, then $T''$ of (2.5) provides the required $\tilde{T}$. Thus we may assume $t_1 \geq t_2 \geq t_3 \geq t_4$. On the other hand, we may also assume $t_1 \leq t_3 + t_4 + 1$, for otherwise Lemma 2.7 can be used to reduce $R$.

Write $m_i := \lfloor \lg(t_i + 1) \rfloor$ for $i = 1, 2, 3, 4$. Then

$$2^{m_1} - 1 \leq t_1 \leq t_3 + t_4 + 1 \leq 2t_3 + 1 < 2(2^{m_3+1} - 1) + 1 = 2^{m_3+2} - 1 ,$$

so $m_1 \leq m_3 + 1$. By the completeness of $r(T)$, $m_3 \leq m_4 + 1$. If either $m_1 = m_3$ or $m_3 = m_4$, then we may conclude $m_1 \leq m_4 + 1$. Suppose instead that $m_1 = m_3 + 1$ and $m_3 = m_4 + 1$. By the completeness of $r(T)$, $T^3$ is perfect, i.e., $t_3 = 2^{m_3} - 1$. But then

$$t_3 + t_4 + 1 < t_3 + 2^{m_4+1} = t_3 + 2^{m_3} = 2^{m_3+1} - 1 = 2^{m_1} - 1 \leq t_1 ,$$

contradicting our earlier assumption that $t_1 \leq t_3 + t_4 + 1$. Therefore, $m_1 \leq m_4 + 1$. If $T$ is not complete, then $m_1 = m_4 + 1$. (This is easy to see from our assumptions about subtree completeness and the fact that $t_1 \geq t_2 \geq t_3 \geq t_4$.)

So we may assume $m_1 = m_4 + 1$, i.e., that exactly one of the three differences $m_1 - m_2$, $m_2 - m_3$, $m_3 - m_4$ equals 1, and the others vanish. We consider each of the three possibilities in turn.

*Case 1:* $m_1 = m_2 + 1$. Let $m = m_2$. If $T$ is not complete, it follows that $T^1$ is perfect with $t_1 = 2^{m+1} - 1$, $T^4$ is perfect with $t_4 = 2^m - 1$, and $T^2$ and $T^3$ are *not* perfect. Observe that $2^{m+1} \leq t_2 + t_3 \leq 2^{m+2} - 4$.

*(1a)* If $t_2 + t_3 < 3 \cdot 2^m - 1$, construct $\tilde{T}$ from $T$ by replacing $T^2$ by $T_{t_2+t_3-(2^m-1)}$ and $T^3$ by $T_{2^m-1}$; that is, let $\tilde{T} = T_n$. Then

$$R(\tilde{T}) = n[1 + t_1 + (t_2 + t_3 - (2^m - 1))][1 + (2^m - 1) + t_4]$$
$$\times R(T^1) R_{t_2+t_3-(2^m-1)} R_{2^m-1} R(T^4) .$$

Now $2^m - 1 \leq t_2 + t_3 - (2^m - 1) < 2^{m+1}$, so

$$R_{t_2+t_3-(2^m-1)} R_{2^m-1} = \frac{1}{t_2 + t_3 + 1} R_{t_2+t_3+1} \leq R_{t_2} R_{t_3} = R(T^2) R(T^3) ,$$

with the inequality holding by induction. Furthermore, $1 + t_1 + t_2 > 1 + t_3 + t_4$ and $t_3 > 2^m - 1$, so

$$[1 + t_1 + (t_2 + t_3 - (2^m - 1))][1 + (2^m - 1) + t_4] < (1 + t_1 + t_2)(1 + t_3 + t_4) .$$

Thus

$$R(\tilde{T}) < n(1 + t_1 + t_2)(1 + t_3 + t_4) R(T^1) R(T^2) R(T^3) R(T^4) = R(T) ,$$

as desired.

*(1b)* If $t_2 + t_3 \geq 3 \cdot 2^m - 1$, construct $\tilde{T}$ from $T$ by replacing $T^2$ by $T_{2^{m+1}-1}$ and

$T^3$ by $T_{t_2+t_3-(2^{m+1}-1)}$; that is, again let $\tilde{T} = T_n$. By calculations similar to those for Case 1a, $R(\tilde{T}) < R(T)$, as desired; we leave the details to the reader.

*Case 2:* $m_2 = m_3 + 1$. Let $m = m_3$. If $T$ is not complete, it follows that $T^2$ is perfect with $t_2 = 2^{m+1} - 1$, $T^4$ is perfect with $t_4 = 2^m - 1$, and $T^1$ is *not* perfect. Furthermore, $T^3$ is not perfect, for otherwise $t_3 + t_4 + 1 = (2^m - 1) + (2^m - 1) + 1 = 2^{m+1} - 1 < t_1$. We combine the analyses of Cases 2 and 3 below.

*Case 3:* $m_3 = m_4 + 1$. Let $m = m_4$. If $T$ is not complete, it follows that $T^2$ and $T^3$ are each perfect with $t_2 = t_3 = 2^{m+1} - 1$ and $T^1$ is *not* perfect. ($T^4$ may or may not be perfect.) We now combine the further analyses of Cases 2 and 3.

*Cases 2 and 3:* For both Cases 2 and 3 we have (1) $m_1 = m_2 = \mu$, where $\mu = \lfloor \lg(|r(T)| + 1) \rfloor$, (2) $T^2$ is perfect, and (3) neither $T^1$ nor $r(T)$ is perfect. For convenience we revert to the notation of (2.2) and put $t_i := |T^i|$ and $m_i = \lfloor \lg(t_i + 1) \rfloor$ for $i = 1, 2, 3$. Thus $T^1, T^2, t_1, t_2, m_1, m_2$ are as before; the new $T^3$ is $r(T)$; the new $t_3$ is what we formerly called $t_3 + t_4 + 1$; and the new $m_3$ is the former $\mu$. Since at the present stage of analysis $m_1 = m_2 = m_3$, we shall write $m \geq 1$ for their common value. In the new notation, neither $T^1$ nor $T^3$ is perfect, but $T^2$ is.

We show that the choice $\tilde{T} = T_n$ again works, i.e., that $R_n < R(T)$. To begin, observe $2^{m+1} \leq t_1 + t_3 \leq 2^{m+2} - 4$. As for Case 1, we divide the remaining analysis into two subcases.

*(a)* If $t_1 + t_3 < 3 \cdot 2^m - 2$, then

$$3 \cdot 2^m - 1 < 3 \cdot 2^m + 1 \leq n = t_1 + t_2 + t_3 + 2 = t_1 + (2^m - 1) + t_3 + 2 < 2^{m+2} - 1,$$

so

$$\begin{aligned}
R_n &= nR_{2^{m+1}-1}R_{n-2^{m+1}} = n(2^{m+1} - 1)R_{2^m-1}R_{2^m-1}R_{n-2^{m+1}} \\
&= n(2^{m+1} - 1)(n - 2^m)^{-1}R_{n-2^m}R_{2^m-1} \\
&\qquad \text{(noting } 1 \leq 2^m - 1 < n - 2^{m+1} < 2^{m+1} - 1\text{)} \\
&< n(2^{m+1} - 1)R_{t_1}R_{t_3}R_{2^m-1} \\
&\qquad \text{by induction, since } t_1 + t_3 + 1 = n - 2^m \text{ and neither } T^1 \text{ nor } T^3 \text{ is perfect} \\
&= n(2^{m+1} - 1)R_{t_1}R_{t_2}R_{t_3} \\
&\leq n(t_1 + t_2 + 1)R_{t_1}R_{t_2}R_{t_3} \qquad \text{since } t_1 > t_2 = 2^m - 1 \\
&= R(T).
\end{aligned}$$

*(b)* If $t_1 + t_3 \geq 3 \cdot 2^m - 2$, then similar calculations show again that $R_n < R(T)$; we leave the details to the reader. ■

*Remark 2.9.* It is easy to check that virtually the same proof extends Theorem 2.2 to show that the complete tree $T_n$ is the unique minimizer in $\bar{B}_n$ of any functional $g$ of the form

$$g(T) = \sum_{x \in T} f(|T(x)|)$$

with $f$ strictly increasing and strictly concave (over $\{1, 2, \ldots\}$). The theorem is the special case $f = \log$ [compare (1.2)].

If $f$ is assumed only to be nondecreasing and concave, then we still have the result that $T_n \in \bar{B}_n^*$. An example is $f(x) \equiv x$. Here uniqueness fails: It is easy to check that $T \in B_n$ minimizes $\Sigma_{x \in T} |T(x)|$ if and only if, with $m := \lfloor \lg(n+1) \rfloor$, (1) $T$ is "perfect through depth $m-1$," i.e., $T$ has $2^k$ nodes at depth $k$ for $k = 0, 1, \ldots, m-1$ [the *depth*, or *level*, of a node being the length (number of edges) in the simple path from the root to the node], and (2) $T$ has height (the maximum depth of any node) at most $m$.

## 2.4. Analysis of Modal Value of $Q$

In this subsection we investigate the asymptotic behavior of the mode of the MTR stationary distribution $Q$, or, equivalently, the minimum value $R_n^* = R_n = R(T_n)$ of $R(T) \equiv 1/Q(T)$, achieved when $T$ is the complete tree $T_n$.

*An Exact Expression for Perfect Trees and a Lower Bound in General.* Analysis of $R_n$ is most straightforward when $n = 2^m - 1$, so we begin with this perfect-tree case. For real $x \geq 1$, define

$$s(x) := \sum_{k=1}^{\infty} 2^{-k} \ln(x - 2^{-k}) \tag{2.7}$$

and for integer $n \geq 0$ define

$$\hat{R}_n := \tfrac{1}{4} \exp[(\ln 4 - |s(1)|)(n+1) - s(n+1)] . \tag{2.8}$$

We then have the following exact solution:

**Proposition 2.10.** *If* $0 \leq n = 2^m - 1$, *then* $R_n = \hat{R}_n$.

*Proof.* Writing $r_m$ for $R_n$, Lemma 2.1 gives the recurrence relation

$$r_0 = 1 , \qquad r_m = (2^m - 1) r_{m-1}^2 , \qquad m \geq 1 ,$$

whose solution,

$$r_m = \prod_{j=1}^{m} (2^j - 1)^{2^{m-j}} , \qquad m \geq 0 ,$$

can be reexpressed

$$r_m = \frac{1}{4(n+1)} \exp[(\ln 4 - |\sigma_m|)(n+1)] , \qquad m \geq 0 , \tag{2.9}$$

in terms of

$$\sigma_m := \sum_{j=1}^{m} 2^{-j} \ln(1 - 2^{-j}) , \qquad m \geq 0 ,$$

by means of straightforward calculations. But

$$2^{-j}|\ln(1-2^{-j})| = \int_{2^{-(j+1)}}^{2^{-j}} h(x)\,dx\,,$$

where

$$h(x) := \sum_{k=0}^{\infty} 2^{-k}\left[\frac{x2^{-k}}{1-x2^{-k}} - \ln(1-x2^{-k})\right],$$

and so

$$\begin{aligned}|\sigma_m| &= \sum_{j=1}^{m} \int_{2^{-(j+1)}}^{2^{-j}} h(x)\,dx = \int_{2^{-(m+1)}}^{1/2} h(x)\,dx = \int_{1/(2(n+1))}^{1/2} h(x)\,dx \\ &= |s(1)| - \int_0^{1/(2(n+1))} h(x)\,dx \\ &= |s(1)| + (n+1)^{-1}[s(n+1) - \ln(n+1)]\,. \end{aligned} \tag{2.10}$$

Combining (2.9) and (2.10) completes the proof. ■

For general $n$, $\hat{R}_n$ provides a lower bound on $R_n$:

**Lemma 2.11.** *For every $n \geq 0$, $R_n \geq \hat{R}_n$.*

*Proof.* The proof is by (strong) induction on $n$. For $n = 0$ we have equality: $R_n = 1 = \hat{R}_n$. The key to the induction step is the recurrence

$$R_n = nR_{\pi(n)}R_{\rho(n)}\,, \qquad n \geq 1\,. \tag{2.11}$$

Here $T_{\pi(n)}$ is whichever of $\ell(T_n)$ and $r(T_n)$ is perfect [and is $\ell(T_n)$ if both are], and $\rho(n)$ is the remainder $\rho(n) = n - 1 - \pi(n)$. Then, by induction,

$$\begin{aligned} R_n &\geq n\hat{R}_{\pi(n)}\hat{R}_{\rho(n)} \\ &= n \times \tfrac{1}{16} \exp\{(\ln 4 - |s(1)|)(n+1) - [s(\pi(n)+1) + s(\rho(n)+1)]\}\,. \end{aligned} \tag{2.12}$$

Now $s$ of (2.7) is concave over $[1, \infty)$, so

$$\begin{aligned} &s(\pi(n)+1) + s(\rho(n)+1) \\ &\leq 2s\left(\frac{\pi(n)+\rho(n)}{2} + 1\right) = 2s\left(\frac{n+1}{2}\right) = s(n+1) + \ln\left(\frac{n}{4}\right). \end{aligned} \tag{2.13}$$

Combining (2.12) and (2.13) completes the proof. ■

*Asymptotics.* It is simple to derive an asymptotic expansion for the function $s$ of (2.7): for any $J = 0, 1, \ldots,$

$$s(x) = \ln x - \sum_{j=1}^{J} (2^{j+1} - 1)^{-1} x^{-j} + O(x^{-(J+1)})$$

as $x \to \infty$. In particular, $s(x) = \ln x + o(1)$, and so

**Lemma 2.12.** $\hat{R}_n = (1 + o(1))\hat{\hat{R}}_n$ *as* $n \to \infty$, *where*

$$\hat{\hat{R}}_n := \frac{1}{4(n+1)} \exp[(\ln 4 - |s(1)|)(n+1)], \qquad n \geq 0.$$

According to Proposition 2.10, this lemma gives precise asymptotics for $R_n$ when $n$ is perfect, i.e., when $n$ is of the form $n = 2^m - 1$. Pinning down the asymptotics for general $n$ is not so easy. Here we will be content to establish the following result by finding a suitable upper bound on $R_n$ to serve as a companion to Lemmas 2.11 and 2.12.

**Theorem 2.13.** *Write $Q_n = 1/R_n$ for the modal value of the MTR stationary distribution $Q$. Then as $n \to \infty$,*

$$\ln R_n = (\ln 4 - |s(1)|)(n+1) + O((\log n)^2).$$

*Proof.* The key again is the recurrence (2.11). The iterates $\rho_k(\cdot)$ of the function $\rho(\cdot)$ are most easily expressed in terms of the binary expansion

$$n + 1 = 2^m + b_{m-1}2^{m-1} + \cdots + b_0 2^0 \qquad (m := \lfloor \lg(n+1) \rfloor).$$

For $0 \leq k \leq m-1$ we have

$$1 \leq \rho_k(n) = 2^{m-k} + b_{m-k-1}2^{m-k-1} + \cdots + b_0 2^0 - 1 < 2^{m-k+1} \tag{2.14}$$

and

$$\pi(\rho_k(n)) = 2^{m-k-1+b_{m-k-1}} - 1.$$

Iterating (2.11),

$$R_n = \prod_{k=0}^{m-1} \rho_k(n) \times \prod_{k=0}^{m-1} R_{\pi(\rho_k(n))}.$$

But from Proposition 2.10, (2.8), and (2.7), and a separate check for $n = 0$, it is clear that

$$R_n \leq \exp[(\ln 4 - |s(1)|)(n+1)] \qquad \text{for perfect } n \geq 0.$$

Hence (using $2^b = 1 + b$ for $b = 0, 1$)

$$\begin{aligned} \prod_{k=0}^{m-1} R_{\pi(\rho_k(n))} &\leq \exp\left[(\ln 4 - |s(1)|) \sum_{j=0}^{m-1} (1 + b_j) 2^j\right] \\ &= \exp[(\ln 4 - |s(1)|)(2^m - 1 + n + 1 - 2^m)] \\ &< \exp[(\ln 4 - |s(1)|)(n+1)]. \end{aligned}$$

Furthermore,

$$\begin{aligned} \prod_{k=0}^{m-1} \rho_k(n) &< 2^m \prod_{j=1}^{m} 2^j \qquad \text{by (2.14)} \\ &= 2^m 2^{m(m+1)/2} = 2^{m(m+3)/2} \\ &\leq 2^{m^2} \qquad \text{for } m \geq 3 \qquad (\text{i.e., for } n \geq 7). \end{aligned}$$

Therefore, for $n \geq 7$,

$$R_n \le \exp[(\ln 4 - |s(1)|)(n+1) + (\ln 2)(\lg(n+1))^2]$$
$$= \exp\left[(\ln 4 - |s(1)|)(n+1) + \frac{1}{\ln 2}(\ln(n+1))^2\right].$$

This is the needed upper bound. ■

*Remark 2.14.* The constant $|s(1)|$ is easily computed to a high degree of numerical accuracy. Rounded to seven decimal places, $\ln 4 - |s(1)| = 0.9457553$. Thus a rough summary of our analysis is that the modal value for $Q$ over $B_n$ is achieved by the complete tree $T_n$ and vanishes in the limit as $n \to \infty$ a bit more slowly than does $e^{-n}$.

## 3. THE LAW OF $Q(T)$ WHEN $T$ IS UNIFORM ON $B_n$

In Section 2 we delineated the range of values of $Q(T)$ for $T \in B_n$: The minimum value is $1/n! = \exp[-n \ln n + n + O(\log n)]$ and the maximum value is $\exp[-c_0 n + O((\log n)^2)]$ for a certain constant

$$c_0 := \ln 4 - \sum_{k=1}^{\infty} 2^{-k} |\ln(1 - 2^{-k})| \doteq 0.946 .$$

What are *typical* values of $Q(T)$? The answer depends on how one defines "typical." A natural approach is to give $T$ a probability distribution on $B_n$ and describe the induced distribution of $Q(T)$. In this section we give $T$ the uniform distribution—certainly a natural choice—and find that $Q(T) = \exp[-c_1 n + O_P((n \log n)^{1/2})]$, where $c_1 \doteq 2.03$ is a constantly explicitly defined in Theorem 3.1 below. But the choice of distribution for $T$ matters greatly. In Section 4 we will assign $T$ another natural distribution, namely, $Q$. In that case we shall find $Q(T) = \exp[-C_1 n + O_P(n^{1/2})]$ for a constant $C_1 \doteq 1.204$ smaller than $c_1$.

The following theorem summarizes the asymptotic results we shall obtain by first deriving exact formulas. Recall that the number of binary search trees on $n$ nodes ($n \ge 0$) is the Catalan number

$$\beta_n := |B_n| = \binom{2n}{n} \Big/ (n+1) . \tag{3.1}$$

The numerical values given in the theorem are correct to the accuracy stated. For example, $c_1 \in (2.025, 2.035)$.

**Theorem 3.1.** *Let $Q$ be the MTR stationary distribution defined at (1.1) and (1.2). Let $T$ be uniform distributed over the set $B_n$ of binary search trees and consider the random variable $L_n = -\ln Q(T)$ with mean $\lambda_n$ and variance $v_n$. As $n \to \infty$,*

*(a)* $\lambda_n = c_1(n+1) - c_2 n^{1/2} + O(\log n)$, *where*

$$c_1 := \sum_{j=2}^{\infty} 4^{-j} \beta_j (\ln j) \doteq 2.03$$

*and*

$$c_2 := 2\sqrt{\pi} \doteq 3.545 .$$

*(b)* $v_n = c_3 n \ln n + O(n)$, *where*

$$c_3 := 2[\pi + 4(\ln 2 - \ln(1 + 2^{-1/2})) - 2\sqrt{2}] \doteq 1.893 .$$

*(c)* $\frac{L_n}{n} \xrightarrow{P} c_1$.

*Remark 3.2.* (a) Part (c) follows immediately from parts (a) and (b).

(b) When $T$ is instead given the distribution $Q$, we will prove a central limit theorem for $L_n$. We do not know any such result in the present case.

(c) Before any calculations, we know $c_1 \geq \ln 4 (\doteq 1.386) \geq c_0$ by Jensen's inequality and the fact that $\ln \beta_n = n \ln 4 + O(\log n)$.

(d) Since the maximum value of $L_n$ is $\ln(n!) = n \ln n + O(n)$, even the fact that $\lambda_n = O(n)$ is interesting.

(e) As seen from (3.1) and Lemma 3.5, the series defining $c_1$ converges very slowly. To obtain $c_1$ to two decimal places, Mathematica was used to sum the first 500 terms (sum$\doteq$1.61), and upper and lower integral approximations were used to bound the remainder.

Our entire analysis of $\mathscr{L}(L_n)$, the distribution (or law) of $L_n$, springs from the fact that a uniformly random tree in $B_n (n \geq 1)$ can be constructed by choosing the label $j$ for the root with probability

$$p_n(j) := \frac{\beta_{j-1}\beta_{n-j}}{\beta_n}, \qquad j \in [n] , \tag{3.2}$$

and then independently choosing uniformly random trees in $B_{j-1}$ and $B_{n-j}$ for the left and right subtrees, respectively. (The labels for the right subtree are, of course, $j+1$ through $n$, not 1 through $n-j$.) As a consequence of this and the basic recursion for $Q$ [see (1.1) and Lemma 2.1],

$$\mathscr{L}(L_n) \text{ is the } p_n(\cdot)\text{-mixture of the laws } \mathscr{L}(L_{j-1} \oplus L_{n-j} + \ln n) , \tag{3.3}$$

where $\oplus$ indicates the addition of independent random variables, i.e., convolution of distributions.

### 3.1. An Exact Expression for the Expected Value $\lambda_n$

We can solve for $\lambda_n = \mathrm{E}\, L_n$ exactly in terms of the Catalan numbers (3.1):

**Proposition 3.3.** *For* $n \geq 0$,

$$\lambda_n = \sum_{j=1}^{n} \frac{\beta_j \beta_{n-j}}{\beta_n} (\ln j)(n + 1 - j) .$$

*Proof.* Let $\gamma_n := \beta_n \lambda_n$ and $\delta_n := \beta_n \ln n, n \geq 1$, and let $\gamma_0 := 0$ and $\delta_0 := 0$. Then, according to (3.3), for $n \geq 1$

$$\gamma_n = \beta_n \sum_{j=1}^{n} p_n(j)(\lambda_{j-1} + \lambda_{n-j} + \ln n)$$

$$= \delta_n + \sum_{j=1}^{n} (\gamma_{j-1}\beta_{n-j} + \beta_{j-1}\gamma_{n-j})$$

$$= \delta_n + 2\sum_{k=0}^{n-1} \beta_k \gamma_{(n-1)-k} .$$

Letting $B(x) = [1 - \sqrt{1-4x}]/(2x)$, $\Gamma(x)$, $\Delta(x)$ denote the generating functions for $(\beta_n)$, $(\gamma_n)$, $(\delta_n)$, respectively, we find $\Gamma(x) = \Delta(x) + 2xB(x)\Gamma(x)$ and hence

$$\Gamma(x) = \Delta(x)(1-4x)^{-1/2} = \Delta(x)\sum_{n=0}^{\infty} (n+1)\beta_n x^n .$$

The result now follows easily. ■

### 3.2. Asymptotics for $\lambda_n$: Proof of Theorem 3.1(a)

Our first step in proving Theorem 3.1(a) is to reduce the range of summation in Proposition 3.3 by half. Using the shorthand $u_n := \binom{2n}{n}$, elementary rearrangement—omitted here—leads to

**Lemma 3.4.** *For $n \geq 1$, $\lambda_n = (n+1)\mu_n + \ln n + \epsilon_n$, where*

$$\mu_n := \sum_{j=1}^{\lfloor (n-1)/2 \rfloor} \frac{u_j u_{n-j}}{u_n}\left(\frac{\ln j}{j+1} + \frac{\ln(n-j)}{n+1-j}\right),$$

$$\epsilon_n := \frac{2(n+1)}{n+2}\frac{u_{n/2}^2}{u_n}\ln\left(\frac{n}{2}\right) \text{ if } n \text{ is even},$$

*and $\epsilon_n := 0$ if $n$ is odd.*

Further analysis requires the following standard consequence of Stirling's formula:

**Lemma 3.5.** *The central binominal coefficient $u_n = \binom{2n}{n}$ satisfies*

$$u_n = [1 + O(n^{-1})]\hat{u}_n ,$$

*where*

$$\hat{u}_n := \frac{4^n}{\sqrt{\pi} n^{1/2}} .$$

As a first application of Lemma 3.5,

$$\epsilon_n = \frac{2(n+1)}{n+2}\frac{u_{n/2}^2}{u_n}\ln(n/2) = (1 + O(n^{-1}))4\pi^{-1/2}n^{-1/2}\ln\left(\frac{n}{2}\right)$$

if $n$ is even, and so $\epsilon_n = O(n^{-1/2}\log n)$, which is quite negligible for our purposes, in any case.

To complete the proof of Theorem 3.1(a), we need to show that $\mu_n = c_1 - c_2 n^{-1/2} + O(n^{-1}\log n)$. But as an immediate corollary of Lemma 3.5 we have

**Lemma 3.6.** *Uniformly for* $1 \le j \le \lfloor (n-1)/2 \rfloor$,

$$\frac{u_{n-j}}{u_n} = 4^{-j}\left(1 - \frac{j}{n}\right)^{-1/2}[1 + O(n^{-1})].$$

Therefore, $\mu_n = (1 + O(n^{-1}))\hat{\mu}_n$, where

$$\hat{\mu}_n := \sum_{j=1}^{\lfloor (n-1)/2 \rfloor} 4^{-j} u_j \left(\frac{\ln j}{j+1} + \frac{\ln(n-j)}{n+1-j}\right)\left(1 - \frac{j}{n}\right)^{-1/2},$$

and so we need to show that

$$\hat{\mu}_n = c_1 - c_2 n^{-1/2} + O(n^{-1}\log n).$$

We accomplish this by breaking $\hat{\mu}_n$ into $a_n + b_n + c_n$, where

$$a_n := \sum_{j=1}^{\lfloor (n-1)/2 \rfloor} 4^{-j} u_j \frac{\ln j}{j+1},$$

$$b_n := \sum_{j=1}^{\lfloor (n-1)/2 \rfloor} 4^{-j} u_j \frac{\ln j}{j+1}\left[\left(1 - \frac{j}{n}\right)^{-1/2} - 1\right],$$

$$c_n := \sum_{j=1}^{\lfloor (n-1)/2 \rfloor} 4^{-j} u_j \frac{\ln(n-j)}{n+1-j}\left(1 - \frac{j}{n}\right)^{-1/2},$$

and proving

$$a_n = c_1 - (8/\pi)^{1/2} n^{-1/2} \ln n - (2 - \ln 2)(8/\pi)^{1/2} n^{-1/2} + O(n^{-3/2}\log n), \quad (3.4)$$

$$b_n = 2(\sqrt{2} - 1)\pi^{-1/2} n^{-1/2} \ln n - J\pi^{-1/2} n^{-1/2} + O(n^{-1}\log n), \quad (3.5)$$

$$c_n = 2\pi^{-1/2} n^{-1/2} \ln n - (\pi + 2\ln 2 - 4)\pi^{-1/2} n^{-1/2} + O(n^{-1}\log n), \quad (3.6)$$

where

$$J = 4 + \pi - 4\sqrt{2} + 2(\sqrt{2} - 1)\ln 2. \quad (3.7)$$

To prove (3.4), it is enough by Lemma 3.5 to note

$$\sum_{j=\lfloor (n-1)/2 \rfloor + 1}^{\infty} j^{-3/2} \ln j = 8^{1/2} n^{-1/2} \ln n + (2 - \ln 2) 8^{1/2} n^{-1/2} + O(n^{-3/2}\log n),$$

which in turn results from upper and lower integral approximations.

To prove (3.5), we begin with

$$b_n = \pi^{-1/2} \sum_{j=1}^{\lfloor (n-1)/2 \rfloor} [1 + O(j^{-1})] j^{-3/2} (\ln j)\left[\left(1 - \frac{j}{n}\right)^{-1/2} - 1\right].$$

Here, the contribution from $O(j^{-1})$ is

$$O\Big(\sum_{j=1}^{\lfloor (n-1)/2\rfloor} j^{-5/2}(\ln j)\frac{j}{n}\Big) = O\Big(n^{-1}\sum_{j=1}^{\lfloor (n-1)/2\rfloor} j^{-3/2}(\ln j)\Big) = O(n^{-1}) .$$

The remaining contribution to $b_n$ is

$$\begin{aligned}
\tilde{b}_n &:= \pi^{-1/2}\sum_{j=1}^{\lfloor (n-1)/2\rfloor} j^{-3/2}(\ln j)\Big[\Big(1-\frac{j}{n}\Big)^{-1/2}-1\Big] \\
&= \pi^{-1/2}(\ln n)\sum_{j=1}^{\lfloor (n-1)/2\rfloor} j^{-3/2}\Big[\Big(1-\frac{j}{n}\Big)^{-1/2}-1\Big] \\
&\quad - \pi^{-1/2}\sum_{j=1}^{\lfloor (n-1)/2\rfloor} j^{-3/2}\Big|\ln\Big(\frac{j}{n}\Big)\Big|\Big[\Big(1-\frac{j}{n}\Big)^{-1/2}-1\Big] \\
&= \pi^{-1/2}n^{-1/2}(\ln n)\sum_{j=1}^{\lfloor (n-1)/2\rfloor}\Big(\frac{j}{n}\Big)^{-3/2}\Big[\Big(1-\frac{j}{n}\Big)^{-1/2}-1\Big]n^{-1} \\
&\quad - \pi^{-1/2}n^{-1/2}\sum_{j=1}^{\lfloor (n-1)/2\rfloor}\Big(\frac{j}{n}\Big)^{-3/2}\Big|\ln\Big(\frac{j}{n}\Big)\Big|\Big[\Big(1-\frac{j}{n}\Big)^{-1/2}-1\Big]n^{-1} \\
&= J_1\pi^{-1/2}n^{-1/2}(\ln n) - J_2\pi^{-1/2}n^{-1/2} + O(n^{-1}\log n) ,
\end{aligned}$$

where the final equality results from upper and lower integral approximations with

$$J_1 := \int_0^{1/2} y^{-3/2}[(1-y)^{-1/2}-1]\,dy = 2(\sqrt{2}-1)$$

$$J_2 := \int^{1/2} y^{-3/2}|\ln y|[(1-y)^{-1/2}-1]\,dy = J \text{ of (3.7)} .$$

The stated values of these integrals are the solutions to challenging calculus problems; we omit the details.

Finally, to prove (3.6), calculations like those for $b_n$ give

$$c_n = (1+O(n^{-1}))\tilde{c}_n + O(n^{-1}\log n) ,$$

where

$$\begin{aligned}
\tilde{c}_n &:= \pi^{-1/2}n^{-1}\sum_{j=1}^{\lfloor (n-1)/2\rfloor} j^{-1/2}(\ln(n-j))\Big(1-\frac{j}{n}\Big)^{-3/2} \\
&= \pi^{-1/2}n^{-1/2}(\ln n)\sum_{j=1}^{\lfloor (n-1)/2\rfloor}\Big(\frac{j}{n}\Big)^{-1/2}\Big(1-\frac{j}{n}\Big)^{-3/2}n^{-1} \\
&\quad - \pi^{-1/2}n^{-1/2}\sum_{j=1}^{\lfloor (n-1)/2\rfloor}\Big(\frac{j}{n}\Big)^{-1/2}\Big|\ln\Big(1-\frac{j}{n}\Big)\Big|\Big(1-\frac{j}{n}\Big)^{-3/2}n^{-1} \\
&= J_3\pi^{-1/2}n^{-1/2}(\ln n) - J_4\pi^{-1/2}n^{-1/2} + O(n^{-1}\log n) ,
\end{aligned}$$

with (values again by calculus)

$$J_3 := \int_0^{1/2} y^{-1/2}(1-y)^{-3/2}\,dy = 2\,,$$

$$J_4 := \int_0^{1/2} y^{-1/2}|\ln(1-y)|(1-y)^{-3/2}\,dy = \pi + 2\ln 2 - 4\,.$$

This completes the proof of Theorem 3.1(a).

### 3.3. The Variance of the Conditional Expectation of $L_n$ Given the Root Label

Let $J$ denote the root label chosen according to the probability mass function $p_n$, as at (3.2), and let

$$\theta_n := \operatorname{Var} \mathrm{E}(L_n|J) = \sum_{j=1}^{n} p_n(j)[(\lambda_{j-1} + \lambda_{n-j} + \ln n) - \lambda_n]^2\,. \tag{3.8}$$

In the next subsection we will derive an exact expression for $v_n = \operatorname{Var} L_n$ in terms of $\theta_n$. In this subsection we use Theorem 3.1(a) to obtain first-order asymptotics for $\theta_n$.

**Lemma 3.7.** *As* $n \to \infty$,

$$\theta_n = [1 + O((\log n)^{-2})]\sqrt{\pi} c_3 n^{1/2}$$

*with* $c_3$ *as defined in the statement of Theorem 3.1(b).*

*Proof.* Proceeding just as in the proof of Theorem 3.1(a), a short argument leads to $\theta_n = [1 + O(n^{-1})]\eta_n + O(n^{-1/2})$, where

$$\eta_n = 2\sum_{j=1}^{\lfloor n/2 \rfloor} 4^{-j}\beta_{j-1}\left(1 - \frac{j}{n}\right)^{-3/2}\{c_2[j^{1/2} + (n-j)^{1/2} - n^{1/2}] + O(\log n)\}^2\,,$$

with the remainder estimate holding uniformly in $j$. To analyze $\eta_n$, we first observe that the sum of terms with $j \le \lfloor (\ln n)^2 \rfloor$ equals

$$O\left((\log n)^2 \sum_{j=1}^{\infty} 4^{-j}\beta_{j-1}\right) = O((\log n)^2)\,,$$

which is negligible. It remains to show $\tilde{\eta}_n = [1 + O((\log n)^{-2})]\sqrt{\pi} c_3 n^{1/2}$, where

$$\begin{aligned}
\tilde{\eta}_n &= 2\sum_{j=\lfloor (\ln n)^2 \rfloor + 1}^{\lfloor n/2 \rfloor} 4^{-j}\beta_{j-1}\left(1 - \frac{j}{n}\right)^{-3/2}\{c_2[j^{1/2} + (n-j)^{1/2} - n^{1/2}] + O(\log n)\}^2 \\
&= [1 + O((\log n)^{-2})]\frac{1}{2}\pi^{-1/2} n^{1/2} \sum_{j=\lfloor (\ln n)^2 \rfloor + 1}^{\lfloor n/2 \rfloor} \left(\frac{j}{n}\right)^{-3/2}\left(1 - \frac{j}{n}\right)^{-3/2} \\
&\quad \times \left\{c_2\left[\left(\frac{j}{n}\right)^{1/2} + \left(1 - \frac{j}{n}\right)^{1/2} - 1\right] + O(n^{-1/2}\log n)\right\}^2 n^{-1} \\
&= [1 + O((\log n)^{-2})]2\pi^{1/2} n^{1/2} \sum_{j=\lfloor (\ln n)^2 \rfloor + 1}^{\lfloor n/2 \rfloor} \left(\frac{j}{n}\right)^{-3/2}\left(1 - \frac{j}{n}\right)^{-3/2}
\end{aligned}$$

$$\times\left[\left(\frac{j}{n}\right)^{1/2}+\left(1-\frac{j}{n}\right)^{1/2}-1\right]^2 n^{-1}+O((\log n)^2)\,.$$

But calculus gives

$$\sum_{j=\lfloor(\ln n)^2\rfloor+1}^{\lfloor n/2\rfloor}\left(\frac{j}{n}\right)^{-3/2}\left(1-\frac{j}{n}\right)^{-3/2}\left[\left(\frac{j}{n}\right)^{1/2}+\left(1-\frac{j}{n}\right)^{1/2}-1\right]^2 n^{-1}$$

$$=[1+O((\log n)^{-2})]J_5\,,$$

with

$$J_5:=\int_0^{1/2} y^{-3/2}(1-y)^{-3/2}[y^{1/2}+(1-y)^{1/2}-1]^2\,dy=\frac{c_3}{2}\,. \qquad \blacksquare$$

### 3.4. An Exact Expression for $v_n$ and Asymptotics: Proof of Theorem 3.1(b)

We can solve for $v_n=\operatorname{Var} L_n$ exactly in terms of the sequence of numbers (3.8):

**Proposition 3.8.** *For $n\geq 0$,*

$$v_n=\sum_{j=1}^{n}\frac{\beta_j\beta_{n-j}}{\beta_n}\theta_j(n+1-j)\,.$$

*Proof.* Observe

$$v_n=\operatorname{Var} L_n=\mathrm{E}\operatorname{Var}(L_n|J)+\operatorname{Var}\mathrm{E}(L_n|J)=\sum_{j=1}^{n}p_n(j)(v_{j-1}+v_{n-j})+\theta_n\,.$$

From here the proof is the same as that of Proposition 3.3, with $\ln(\cdot)$ replaced by $\theta$ throughout. $\blacksquare$

We not turn to asymptotics, proceeding as for $\lambda_n$ to obtain $v_n=[1+O(n^{-1})](n+1)w_n+O(n^{1/2})$, where

$$w_n=\sum_{j=1}^{\lfloor(n-1)/2\rfloor}4^{-j}u_j\left(\frac{\theta_j}{j+1}+\frac{\theta_{n-j}}{n+1-j}\right)\left(1-\frac{j}{n}\right)^{-1/2}\,.$$

Now if we break $w_n$ into $a_n+b_n+c_n$ just as we did for $\hat{\mu}_n$ in the proof of Theorem 3.1(a), we find by routine calculations that

$$a_n=c_3\ln n+O(1)\,,\qquad b_n=O(1)\,,\qquad c_n=O(1)\,.$$

Thus $v_n=c_3 n\ln n+O(n)$, and Theorem 3.1(b) is proved.

## 4. THE LAW OF $Q(T)$ WHEN $T$ HAS DISTRIBUTION $Q$

Theorem 3.1 gave asymptotic information about the distribution of $Q(T)$ when $T$ is uniformly distributed over $B_n$. Now we summarize the asymptotic distribution of $Q(T)$ when $T$ is given the distribution $Q$. As in Theorem 3.1, numerical values

given are correct to the accuracy stated. [However, concerning $C_2$, see Remark 4.2(c) below.]

[Technical note: As the derivations establishing Theorem 4.1 will make clear, $n+1$ is, in a certain sense, a more natural parameter than $n$. This is why the asymptotic normality in part (d) is given for $[L_n - C_1(n+1)]/[C_2(n+1)]^{1/2}$ rather than for $(L_n - C_1 n)/(C_2 n)^{1/2}$.]

**Theorem 4.1.** *Let $Q$ be the MTR stationary distribution defined at (1.1) and (1.2) and let $T$ be a random binary search tree having distribution $Q$. Consider the random variable $L_n = -\ln Q(T)$ with mean $\Lambda_n$ and variance $V_n$. As $n \to \infty$,*

(a) $\Lambda_n = C_1(n+1) - \ln n - 2 + O(n^{-1}\log n)$, *where*

$$C_1 := 2\sum_{k=2}^{\infty} \frac{\ln k}{(k+1)(k+2)} \doteq 1.204 .$$

(b) $V_n = C_2(n+1) - C_3 + O(n^{-1}(\log n)^3)$, *where*

$$C_2 := 2\sum_{k=3}^{\infty} \frac{\Theta_k}{(k+1)(k+2)} \doteq 0.168 ,$$

$$C_3 := 4 - \frac{\pi^2}{3} \doteq 0.710 ,$$

*with*

$$\Theta_n = \frac{1}{n}\sum_{j=1}^{n} [(\Lambda_{j-1} + \Lambda_{n-j} + \ln n) - \Lambda_n]^2 = C_3 + O(n^{-1}(\log n)^3) .$$

(c) $\frac{L_n}{n} \xrightarrow{P} C_1$.

(d) $[L_n - C_1(n+1)]/[C_2(n+1)]^{1/2} \xrightarrow{\mathcal{L}}$ standard normal.

*Remark 4.2.* (a) Again, part (c) follows immediately from parts (a) and (b).

(b) It is easy to see why $C_1$ of Theorem 4.1 is smaller than $c_1$ of Theorem 3.1. Indeed, writing $Q_n = \exp(-L_n)$, we have

$$\mathrm{E}_Q L_n = \sum_{T \in B_n} Q(T)L(T) = \sum_{T \in B_n} [-Q(T)\ln Q(T)] = \beta_n \mathrm{E}_U[-Q_n \ln Q_n]$$

$$\leq \beta_n[-(\mathrm{E}_U Q_n)\ln(\mathrm{E}_U Q_n)] = \beta_n[-\beta_n^{-1}\ln(\beta_n^{-1})] = \ln \beta_n = n \ln 4 + O(\log n)$$

using Jensen's inequality, so that $C_1 \leq \ln 4$. But by Remark 3.2 we have $\ln 4 \leq c_1$.

Of course we also know, before calculating $C_1$, that it is at least $c_0 \doteq 0.946$.

(c) The assertion $C_2 \in (0.1675, 0.1685)$ in part (b) is predicated on the conjecture, verified by calculations up through $n = 400$, that $\Theta_n$ is increasing in $n \geq 3$. [Note $\Theta_n = 0$ for $n = 0, 1, 2$.] Indeed, we then find numerically (with the aid of Mathematica) that $\Theta_{400} = 0.693456$, $\Sigma_{k=3}^{399}\ [\Theta_k/((k+1)(k+2))] = 0.0821197$, and

$$0.00172932 = \frac{\Theta_{400}}{401} = \Theta_{400} \sum_{k=400}^{\infty} \frac{1}{(k+1)(k+2)}$$

$$\leq \sum_{k=400}^{\infty} \frac{\Theta_k}{(k+1)(k+2)} \leq C_3 \sum_{k=400}^{\infty} \frac{1}{(k+1)(k+2)} = \frac{C_3}{401}$$

$$= 0.00177090 .$$

The paragraph immediately preceding Section 3.1 applies verbatim here, with $p_n$ replaced by the uniform distribution $u_n(j) \equiv 1/n$ on $[n]$. In particular,

$$\mathscr{L}(L_n) \text{ is the } u_n(\cdot)\text{-mixture of the laws } \mathscr{L}(L_{j-1} \oplus L_{n-j} + \ln n) . \tag{4.1}$$

### 4.1. The Solution to a Recurrence Relation and the Proof of Theorem 4.1(a)

The following standard and elementary result will thrice be used in proving Theorem 4.1.

**Lemma 4.3.** *The recurrence relation*

$$x_n = a_n + \frac{2}{n} \sum_{k=0}^{n-1} x_k , \qquad n \geq 1 ,$$

*for the sequence* $(x_n)_{n \geq 0}$ *has solution*

$$x_n = a_n + (n+1)\left[ (x_0 - a_0) + 2 \sum_{k=0}^{n-1} \frac{a_k}{(k+1)(k+2)} \right], \qquad n \geq 0 ,$$

*for arbitrarily defined* $a_0$.

Using this lemma and (4.1) we obtain immediately an exact expression for $\Lambda_n$:

**Proposition 4.4.** *We have* $\Lambda_0 = 0$ *and, for* $n \geq 1$,

$$\Lambda_n = \ln n + 2(n+1) \sum_{k=2}^{n-1} \frac{\ln k}{(k+1)(k+2)} .$$

Theorem 4.1(a) now follows simply, using $(\ln k)/((k+1)(k+2)) = [1 + O(k^{-1})](\ln k)/k^2$ and monotonicity of $(\ln k)/k^2$ to establish integral upper and lower bounds on the remainder. We omit the details.

### 4.2. Proof of Theorem 4.1(b)

As in Section 3.4, we can express $V_n = \operatorname{Var} L_n$ in terms of the sequence

$$\Theta_n := \operatorname{Var} \mathrm{E}(L_n | J) = \frac{1}{n} \sum_{j=1}^{n} [(\Lambda_{j-1} + \Lambda_{n-j} + \ln n) - \Lambda_n]^2 , \qquad n \geq 0 ,$$

where $J$ is the uniformly chosen root label. The proof is another simple application of (4.1) and Lemma 4.3.

**Proposition 4.5.** *For* $n \geq 0$,

$$V_n = \Theta_n + 2(n+1)\sum_{k=3}^{n-1} \frac{\Theta_k}{(k+1)(k+2)} .$$

The asymptotics for $V_n$ stated in Theorem 4.1(b) follow rather simply from the first-order asymptotics for $\Theta_n$ stated there. Our goal for the remainder of this subsection is therefore to establish

$$\Theta_n = C_3 + O(n^{-1}(\log n)^3) . \tag{4.2}$$

To begin, symmetry, together with a crude analysis [using Theorem 4.1(a)] of the extra middle term when $n$ is odd, gives $\Theta_n = K_n + O(n^{-1})$, where

$$K_n = \frac{2}{n}\sum_{j=1}^{\lfloor n/2 \rfloor} [(\Lambda_{j-1} + \Lambda_{n-j} + \ln n) - \Lambda_n]^2 .$$

Now, uniformly for $j \in [\lfloor n/2 \rfloor]$, we have by Theorem 4.1(a)

$$\Lambda_{j-1} + \Lambda_{n-j} + \ln n - \Lambda_n = [2\ln n - \ln j - \ln(n-j) - 2] + O(j^{-1}\log j) .$$

Squaring this out and summing, we find that $K_n = \tilde{K}_n + O(n^{-1}(\log n)^3)$, where

$$\tilde{K}_n := \frac{2}{n}\sum_{j=1}^{\lfloor n/2 \rfloor} [2\ln n - \ln j - \ln(n-j) - 2]^2$$

$$= 2\sum_{j=1}^{\lfloor n/2 \rfloor} \left[ \left|\ln\left(\frac{j}{n}\right)\right| + \left|\ln\left(1 - \frac{j}{n}\right)\right| - 2 \right]^2 n^{-1} .$$

To complete the proof of (4.2), we show that

$$\tilde{K}_n = C_3 + O(n^{-1}(\log n)^2) .$$

Indeed, it is routine to show that

$$2J(n^{-1}) + O(n^{-1}) \le \tilde{K}_n \le 2J(0) + O(n^{-1}) ,$$

where

$$J(a) := \int_a^{1/2} [|\ln y| + |\ln(1-y)| - 2]^2 \, dy .$$

Finally, another complicated exercise in calculus yields

$$J(a) = C_3/2 - O(a|\ln a|^2) .$$

### 4.3. A Central Limit Theorem: Proof of Theorem 4.1(d)

From Theorem 4.1(a) it is immediate that we may equivalently prove Theorem 4.1(d) in the form

$$\tilde{Z}_n \xrightarrow{\mathscr{L}} Z , \tag{4.3}$$

where $\tilde{Z}_n := (L_n - \Lambda_n)/[C_4(n+1)^{1/2}]$ with $C_4 := C_2^{1/2}$ and $Z$ has the standard normal distribution. We will prove (4.3) by the method of moments, i.e., by establishing the sufficient condition

$$\forall k \geq 1: \qquad \mathrm{E}\,\tilde{Z}_n^k = \mu_k + o(1) \qquad \text{as } n \to \infty\,, \tag{4.4}$$

where $\mu_k := \mathrm{E}\,Z^k$ so that $\mu_{2i+1} = 0$ and $\mu_{2i} = 2^{-i}(2i)!/i!$.

For $j \in [n]$, let

$$\begin{aligned}\alpha_{n,j} &:= \Lambda_{j-1} + \Lambda_{n-j} + \ln n - \Lambda_n \\ &= \left|\ln\left(\frac{j}{n}\right)\right| + \left|\ln\left(1 - \frac{j}{n}\right)\right| - 2 + O(j^{-1}\log j) \text{ uniformly for } j \in [\lfloor n/2 \rfloor]\,.\end{aligned}$$

Note that $\tilde{L}_n := L_n - \Lambda_n$ has the property that

$$\mathscr{L}(\tilde{L}_n) \text{ is the } u_n(\cdot)\text{-mixture of the laws } \mathscr{L}(\tilde{L}_{j-1} \oplus \tilde{L}_{n-j} + \alpha_{n,j})\,.$$

We will prove (4.4) in the equivalent form

$$\forall k \geq 2: \qquad \mu_{n,k} := \mathrm{E}\,\tilde{L}_n^k = \mu_k C_4^k (n+1)^{k/2} + o((n+1)^{k/2})\,, \tag{4.5}$$

the estimate certainly holding for $k = 0, 1$.

The case $k = 2$ is handled in Theorem 4.1(b). For larger values of $k$ we note $\mu_{n,k} = 0$ for $n = 0, 1, 2$ and $k \geq 1$ and use the recurrence

$$\begin{aligned}\mu_{n,k} &= \frac{1}{n}\sum_{j=1}^{n} \mathrm{E}(\alpha_{n,j} + \tilde{L}_{j-1} \oplus \tilde{L}_{n-j})^k \\ &= \Theta_{n,k} + \frac{2}{n}\sum_{i=0}^{n-1} \mu_{i,k}\,,\end{aligned} \tag{4.6}$$

where

$$\begin{aligned}\Theta_{n,k} &:= \mu_{n,k} - \frac{2}{n}\sum_{i=0}^{n-1} \mu_{i,k} \\ &= \frac{1}{n}\sum_{j=1}^{n} \sum_{\substack{k_1+k_2+k_3=k,\\ k_2<k,\ k_3<k}} \binom{k}{k_1, k_2, k_3} \alpha_{n,j}^{k_1} \mu_{j-1,k_2} \mu_{n-j,k_3}\,.\end{aligned} \tag{4.7}$$

According to Lemma 4.3, the recurrence relation (4.6) has the solution

$$\mu_{n,k} = \Theta_{n,k} + 2(n+1)\sum_{i=3}^{n-1} \frac{\Theta_{i,k}}{(i+1)(i+2)}\,. \tag{4.8}$$

The proof of (4.5) is now by (strong) induction on $k$. Supposing the result is true up through $k-1$, we prove it for $k$. The first main step is to get asymptotics for $\Theta_{n,k}$:

**Claim.**

$$\Theta_{n,k} = \begin{cases} O((n+1)^{(k-1)/2}) & \text{if } k \text{ is odd}\,, \\ 2^{-k/2}\left(\frac{k}{2} - 1\right)\dfrac{k!}{(\frac{k}{2}+1)!}\, C_4^k (n+1)^{k/2} + o((n+1)^{k/2}) & \text{if } k \text{ is even}\,. \end{cases} \tag{4.9}$$

We begin the proof of the claim with a (re-)symmetrization trick:

$$\Theta_{n,k} = \frac{1}{2n} \sum_{j=1}^{n} \sum_{\substack{k_1+k_2+k_3=k,\\ k_2<k,\, k_3<k}} \binom{k}{k_1, k_2, k_3} \alpha_{n,j}^{k_1} (\mu_{j-1,k_2} \mu_{n-j,k_3} + \mu_{j-1,k_3} \mu_{n-j,k_2}) .$$

Using the induction hypothesis and arguments (notably, symmetry in $j$) just like those for $\Theta_n$ in establishing (4.2) we find

$$\frac{1}{n} \sum_{j=1}^{n} \alpha_{n,j}^{k_1} (\mu_{j-1,k_2} \mu_{n-j,k_3} + \mu_{j-1,k_3} \mu_{n-j,k_2})$$

$$= 2(n+1)^{(k_2+k_3)/2} \mu_{k_2} C_4^{k_2} \mu_{k_3} C_4^{k_3} J_{k_1,k_2,k_3} + o((n+1)^{(k_2+k_3)/2}) ,$$

with

$$J_{k_1,k_2,k_3} := \int_0^{1/2} [|\ln y| + |\ln(1-y)| - 2]^{k_1} [y^{k_2/2}(1-y)^{k_3/2} + y^{k_3/2}(1-y)^{k_2/2}] \, dy$$

$$= \frac{1}{2} \int_0^1 [|\ln y| + |\ln(1-y)| - 2]^{k_1} [y^{k_2/2}(1-y)^{k_3/2} + y^{k_3/2}(1-y)^{k_2/2}] \, dy .$$

Hence

$$\Theta_{n,k} = \sum_{\substack{k_1+k_2+k_3=k,\\ k_2<k,\, k_3<k}} \binom{k}{k_1, k_2, k_3}$$

$$\times [(n+1)^{(k_2+k_3)/2} \mu_{k_2} C_4^{k_2} \mu_{k_3} C_4^{k_3} J_{k_1,k_2,k_3} + o((n+1)^{(k_2+k_3)/2})]$$

$$= \sum \binom{k}{k_1^*, k_2^*, k_3^*} (n+1)^{(k_2^*+k_3^*)/2} C_4^{k_2^*+k_3^*} \mu_{k_2^*} \mu_{k_3^*} J_{k_1^*,k_2^*,k_3^*}$$

$$+ o((n+1)^{(k_2^*+k_3^*)/2}) ,$$

where the sum is over all choices $k_1^*, k_2^*, k_3^*$ summing to $k$ with $k_2^* < k$ and $k_3^* < k$ both even that give the maximum value of $k_2^* + k_3^*$. For $k \geq 3$:

(a) *If $k$ is odd*, there are $(k+1)/2$ such choices, each with $k_2^* + k_3^* = k-1$. In that case,

$$\Theta_{n,k} = \left[ \sum_{m=0}^{(k-1)/2} \binom{k}{1, 2m, k-1-2m} \mu_{2m} \mu_{k-1-2m} J_{1,2m,k-1-2m} \right] C_4^{k-1} (n+1)^{(k-1)/2}$$

$$+ o((n+1)^{(k-1)/2}) .$$

$$= O((n+1)^{(k-1)/2}) .$$

(b) *If $k$ is even*, there are $\frac{k}{2} - 1$ such choices, each with $k_2^* + k_3^* = k$. In that case,

$$\Theta_{n,k} = \left[ \sum_{m=1}^{\frac{k}{2}-1} \binom{k}{0, 2m, k-2m} \mu_{2m} \mu_{k-2m} J_{0,2m,k-2m} \right] C_4^k (n+1)^{k/2} + o((n+1)^{k/2}) .$$

Now, writing $B$ for the beta function,

$$J_{0,2m,k-2m} = \tfrac{1}{2}[B(m+1, \tfrac{k}{2}+1-m) + B(\tfrac{k}{2}+1-m, m+1)],$$

so that

$$\sum_{m=1}^{\frac{k}{2}-1} \binom{k}{0, 2m, k-2m} \mu_{2m}\mu_{k-2m}J_{0,2m,k-2m}$$

$$= \frac{1}{2}\sum_{m=1}^{\frac{k}{2}-1} \binom{k}{2m} \mu_{2m}\mu_{k-2m}\left[B\left(m+1, \frac{k}{2}+1-m\right) + B\left(\frac{k}{2}+1-m, m+1\right)\right]$$

$$= \sum_{m=1}^{\frac{k}{2}-1} \binom{k}{2m} \mu_{2m}\mu_{k-2m}B\left(m+1, \frac{k}{2}+1-m\right)$$

$$= \sum_{m=1}^{\frac{k}{2}-1} \binom{k}{2m} \mu_{2m}\mu_{k-2m}\frac{m!(\frac{k}{2}-m)!}{(\frac{k}{2}+1)!}$$

$$= \sum_{m=1}^{\frac{k}{2}-1} \frac{k!}{(2m)!(k-2m)!} 2^{-m}\frac{(2m)!}{m!} 2^{-(\frac{k}{2}-m)}\frac{(k-2m)!}{(\frac{k}{2}-m)!}\frac{m!(\frac{k}{2}-m)!}{(\frac{k}{2}+1)!}$$

$$= 2^{-k/2}\left(\frac{k}{2}-1\right)\frac{k!}{(\frac{k}{2}+1)!},$$

and the claim about $\Theta_{n,k}$ is established.

We now conclude the proof of (4.5). Substituting (4.9) into (4.8),

(a) *If $k$ is odd,*

$$\mu_{n,k} = O((n+1)^{(k-1)/2}) + 2(n+1)\sum_{i=3}^{n-1}\frac{O(i^{(k-1)/2})}{(i+1)(i+2)}$$

$$= O((n+1)^{(k-1)/2}) + 2(n+1)\sum_{i=3}^{n-1} O(i^{(k-5)/2})$$

$$= \begin{cases} O((n+1)^{(k-1)/2}) + O(n\log n) = O(n\log n) = o(n^{3/2}) & \text{if } k=3, \\ O((n+1)^{(k-1)/2}) + O(n^{(k-1)/2}) = O(n^{(k-1)/2}) = o(n^{k/2}) & \text{if } k \geq 5 \end{cases}$$

$$= \mu_k C_4^k (n+1)^{k/2} + o((n+1)^{k/2}) \text{ in either circumstance},$$

since $\mu_k = 0$, and (4.5) is established.

(b) *If $k \geq 4$ is even,*

$$\mu_{n,k} = 2^{-k/2}\left(\frac{k}{2}-1\right)\frac{k!}{(\frac{k}{2}+1)!}C_4^k(n+1)^{k/2} + o((n+1)^{k/2})$$

$$+ (n+1)2^{-(\frac{k}{2}-1)}\left(\frac{k}{2}-1\right)\frac{k!}{(\frac{k}{2}+1)!}C_4^k\sum_{i=3}^{n-1}\frac{(i+1)^{k/2}}{(i+1)(i+2)} + \text{remainder},$$

where, recalling that $k$ is fixed, the remainder is

$$(n+1)\,\Sigma_{i=3}^{n-1}\,[\epsilon_i(i+1)^{k/2}]/[(i+1)(i+2)]$$

for a sequence $(\epsilon_n)$ with $\epsilon_n \to 0$. Since $k \geq 4$, remainder equals

$$o\Big((n+1)\sum_{i=3}^{n-1}\frac{(i+1)^{k/2}}{(i+1)(i+2)}\Big)=o\Big(n\sum_{i=3}^{n-1}i^{\frac{k}{2}-2}\Big)=o(n^{k/2})\,.$$

Moreover,

$$\sum_{i=3}^{n-1}\frac{(i+1)^{k/2}}{(i+1)(i+2)}=\sum_{i=3}^{n-1}\frac{(i+1)^{\frac{k}{2}-1}}{(i+2)}=(1+o(1))\frac{1}{\frac{k}{2}-1}(n+1)^{\frac{k}{2}-1}\,.$$

Therefore, $\mu_{n,k}=\nu_k C_4^k(n+1)^{k/2}+o((n+1)^{k/2})$, where

$$\nu_k=2^{-k/2}\frac{k!}{(\frac{k}{2}+1)!}\Big[\frac{k}{2}-1+2\Big]=2^{-k/2}\frac{k!}{(k/2)!}=\mu_k\,,$$

and again (4.5) is established. This completes the proof of Theorem 4.1(d).

## ACKNOWLEDGMENTS

The author thanks Bob Dobrow and two referees for helpful comments.

## REFERENCES

[1] R. P. Dobrow and J. A. Fill, On the Markov chain for the move-to-root rule for binary search trees, *Ann. Appl. Probab.* **5**, 1–19 (1995).
[2] R. P. Dobrow and J. A. Fill, Rates of convergence for the move-to-root Markov chain for binary search trees, *Ann. Appl. Probab.* **5**, 20–36 (1995).
[3] H. M. Mahmoud, *Evolution of Random Search Trees*, Wiley, New York, 1992.

Received December 7, 1994
Accepted August 31, 1995