On the Expected Depth of Random Circuits

Sunil Arya* Mordecai J. Golin† Kurt Mehlhorn‡

April 3, 1998

Abstract: In this paper we analyze the expected depth of random circuits of fixed fanin f. Such circuits are built a gate at a time, with the f inputs of each new gate being chosen randomly from among the previously added gates. The depth of the new gate is defined to be one more than the maximal depth of its input gates. We show that the expected depth of a random circuit with n gates is bounded from above by $ef\ln n$ and from below by $2.04\ldots f\ln n$.

1 Introduction

Recently, Diaz, Serna, Spirakis and Toran [1], motivated by the problem of discovering how quickly a circuit could be evaluated in parallel, posed the question of calculating the depth of a random circuit. They described a model for the generation of random circuits in which each gate has fanin two and showed an upper bound of $O(\log^3 n)$ on the circuit's expected depth, where n is the number of nodes in the circuit. In this paper we extend their problem by investigating random circuits of fixed fanin f where f can be any integer greater than 0 and tighten their bound by showing the existence of constants c_1 and c_2 such that

$$c_1 f\ln n \leq \text{Expected Depth} \leq c_2 f\ln n.$$

We use straightforward methods to prove that $c_2 \leq e$. To evaluate c_1 we develop a new technique that constructs a Markov chain that approximates (and understates) the growth of the circuit and are able to lower bound c_1 as a function of the values of the stationary distribution of the chain. Using this technique we can analytically bound $1.39\ldots \leq c_1$. The technique allows us to do better, though. By constructing a more complicated Markov chain and using

*Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. arya@cs.ust.hk. The work of this author partially supported by HKUST grant DAG95/96.EG01.

†Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, golin@cs.ust.hk. The work of this author partially supported by HK RGC CRG grants 181/93E and 652/95E.

‡Max-Planck-Institut für Informatik, D-66123 Saarbrücken, Germany, mehlhorn@mpi-sb.mpg.de.

Process $\mathcal{P}$:

- $d_0 = 0$
- For each n choose f integers $i_n^1, \ldots, i_n^f$ where each i_n^j is chosen randomly from $\{0, 1, 2, \ldots, n-1\}$.
  1. The multiset $S_n = \{i_n^1, \ldots, i_n^f\}$ contains the *inputs* to gate n.
  2. $d_n = 1 + \max_{j \le f} d_{i_n^j}$ is the *depth* of gate n.

Figure 1: The random process $\mathcal{P}$ that generates random circuits

a linear equation solver to find its stationary distribution we are able to prove that $2.0479\ldots \le c_1$.

As an interesting philosophical aside we point out that our technique also seems to provide a good indication (although not a formal proof) that $2.44 \le c_1$. The idea is this. We show that c_1 is bounded from below by a function of the values of a stationary distribution of an extremely complicated Markov chain, a chain so large that with our computer resources we were unable to explicitly calculate its stationary distribution. We were easily able to simulate the action of this chain, though, and did so for a very large number of steps. The number of times a state was reached divided by the number of steps the process was run should, with high confidence, closely approach the stationary distribution of the state. Plugging the values found into our equations we are thus able to say that we believe with high confidence that $2.44 \le c_1$.

1.1 Problem Definition and History

For any integer $f \ge 1$, a *circuit of fanin f* is an acyclic graph where the indegree of each node is bounded by f. Circuits are among the most useful structures in computer science. Depending on the application, different structural parameters of circuits are important. For example, when a circuit is used as a computational device the length of a longest and a shortest path from a source (= node with indegree 0) to a node measures the minimum and maximum delay of an input signal to reach the node.

We study structural parameters of *random circuits.* A random circuit of maximal fanin f is constructed node-by-node by starting with node 0 and defining its *depth* or *level,* $d_0 = 0$. The f inputs of the nth[1] gate are chosen at random (with repetition) from the first $n-1$ gates; the depth of the nth gate is defined to be 1 plus the maximum depth of its inputs. Formally, random circuits are built a gate at a time by the process $\mathcal{P}$ defined in Figure 1.

In what follows we will use the following notation:

[1]The *nth gate* refers to the nth gate added to the circuit. Because gate numbering starts at 0, the circuit actually contains the $n+1$ gates $0, 1, \ldots, n$ after the nth gate is added.

Figure 2: Random circuit of fanin two. Black node 6 has inputs 0 and 4 and therefore has depth $d_6 = \max(d_0, d_4) + 1 = 3$.

Definition 1

$$
\begin{aligned}
L(d,n) &= \text{Number of nodes at depth } d \text{ at time } n \\
&= |\{i \,:\, i \leq n,\, d_i = d\}| \\
E(d,n) &= \mathtt{E}\,(L(d,n)) \\
\text{Depth}_n(P) &= \text{Depth of the circuit at time } n \;=\; \max_{i \leq n} d_i
\end{aligned}
$$

Note that if $f = 1$ the random circuit process reduces to picking a node at random and making the new node its child. This is a very standard and well studied technique for creating random trees. The expected depth of a random tree of n nodes was shown by Pittel [7] to be $e \ln n$. We refer the reader to papers by Mahmoud and Smythe [3], Meir and Moon [4], Moon [5], Na and Rapaport [6], and Szymanski [8] for more results on random trees.

In section 2 we present an upper bound on the expected depth of random circuits of $ef\ln n = 2.718\ldots f\ln n$, where f is the fanin. Our approach is based on first computing an upper bound on the expected number of nodes at any level and then arguing that the probability that the tree is so deep that the expected number of nodes at the deepest level is less than one is small. Note that for $f = 1$, Pittel's result implies that this bound is tight. In section 3 we present lower bounds on the expected depth. The first lower bound we find, one which we present for pedagogical reasons because it illustrates the technique used, is $1.39\ldots f\ln n$. To compute the lower bound, we define a modified random circuit which is guaranteed to always have a depth smaller than the original circuit. The advantage of the modified circuit is that it is mathematically easier to analyze than the original circuit. We model the modified circuit by a Markov chain whose analysis yields the above mentioned lower bound. By using a more accurate model, we are able to prove a better lower bound of $2.04\ldots f\ln n$. As we attempt to use still more accurate models, it turns out that the Markov chain has too many states and we are not able to solve the equations due to the large computational resources needed. We are, however, able to simulate the Markov chain directly leading to a possible lower bound of $2.44\ldots f\ln n$.

Note: Tsukiji and Xhafa have just released a new paper [9] in which they analyze the same problem and actually prove a stronger result, namely that the circuit depth actually converges to $ef\ln n$ in probability. Their paper, though uses much more sophisticated and specialized tools as "subroutines", namely Kingman's theorem and Pittel's analysis [7] of the height of random trees (actually the case $f = 1$). This paper, while presenting a weaker result, does so in a fully self-contained manner, its main tool being the convergence theorem for homogeneous Markov chains.

2 An Upper Bound on the Expected Circuit Depth

In this section we derive an upper bound on the expected circuit depth. We start by establishing a recurrence inequality describing the growth of $E(d,n)$, the expected number of nodes on level d at time n.

Lemma 1 *Consider a random circuit of fanin f. Then*

$$\begin{array}{rcll} E(0,n) & = & 1 & \text{for } n \geq 0 \\ E(d,0) & = & 0 & \text{for } d \geq 1 \\ E(d,n) & \leq & E(d,n-1) + f \cdot \frac{E(d-1,n-1)}{n} & \text{for } n \geq 1, d \geq 1 \end{array} \tag{1}$$

Proof The first and the second statements follow directly from the way the random circuit is constructed. To prove the third equation note that

$$L(d,n) = L(d,n-1) + \delta(d,n), \tag{2}$$

where $\delta(d,n) = 1$, if $d_n = d$ and 0 otherwise. Taking expectations of both sides of the equation yields

$$\begin{aligned} E(d,n) &= E(d,n-1) + \Pr\left(\delta(d,n) = 1\right) \\ &= E(d,n-1) + \\ &\quad \sum_{i \geq 1} \Pr\left(L(d-1,n-1) = i\right) \cdot \Pr\left(\delta(d,n) = 1 \,\middle|\, L(d-1,n-1) = i\right) \end{aligned}$$

Observe that if $d_n = d$ this implies that at least one of the nth gate's f inputs is at depth $d-1$. Thus

$$\Pr\left(\delta(d,n) = 1 \,\middle|\, L(d-1,n-1) = i\right) \leq 1 - (1 - i/n)^f \leq fi/n. \tag{3}$$

Substituting this back we have

$$
\begin{aligned}
E(d,n) &\le E(d,n-1) + \sum_{i\ge 1} \Pr\left(L(d-1,n-1)=i\right)\cdot fi/n \\
&\le E(d,n-1) + fE(d-1,n-1)/n, \qquad (4)
\end{aligned}
$$

establishing the third equation of the Lemma. $\square$

The inequalities just derived can now be used to upper bound $E(d,n)$. In what follows $H_n = \sum_{1\le i\le n} 1/i$ is the nth *harmonic number.*

Lemma 2

$$E(d,n) \le (f\cdot H_n)^d/d!.$$

Proof Define $\overline{E}(d,n)$ using the same relation as in the statement of Lemma 1 except with the inequality in the third equation being replaced by equality. We compute an upper bound for $\overline{E}(d,n)$ defined by this new relation; clearly this upper bound also applies to $E(d,n)$.

Let $\overline{E}_d(x) = \sum_{n\ge 0} \overline{E}(d,n)x^n$ be the generating function for the $\overline{E}(d,n)$. Then

$$
\begin{aligned}
\overline{E}_0(x) &= 1+x+x^2+\cdots = \frac{1}{1-x} && (5)\\
\overline{E}_d(x) &= \frac{f}{1-x}\int \overline{E}_{d-1}(x)dx \qquad \text{for } d\ge 1. && (6)
\end{aligned}
$$

The latter equation follows easily from the recurrence relation on $\overline{E}(d,n)$ (multiply both sides by x^n and then sum over n). The constant of integration in Eq. (6) must be chosen to satisfy the boundary condition $\overline{E}_d(0)=0$. It is now easy to prove by induction that $\overline{E}_d(x)$ is given by

$$\overline{E}_d(x) = \frac{f^d}{d!}\cdot\frac{1}{1-x}\cdot(-\ln(1-x))^d. \qquad (7)$$

We can extract $\overline{E}(d,n)$, the coefficient of x^n in the series expansion of $\overline{E}_d(x)$, as follows. Using the expansion $-\ln(1-x) = \sum_{i\ge 1} x^i/i$ and the convolution theorem of ordinary generating functions it is seen that the coefficient of x^r in $(-\ln(1-x))^d$ is given by

$$\sum_{\substack{i_1,\ldots,i_d\ge 1\\ i_1+\cdots+i_d=r}} \frac{1}{i_1 i_2\cdots i_d}. \qquad (8)$$

The coefficient of x^n in $(1/(1-x))\cdot(-\ln(1-x))^d$ can be computed by summing the coefficients of x^r in $(-\ln(1-x))^d$ over r ranging from 0 to n. Thus, $\overline{E}(d,n)$, the coefficient of x^n in $\overline{E}_d(x)$, is given by

$$\overline{E}(d,n) = \frac{f^d}{d!}\left(\sum_{\substack{i_1,\ldots,i_d\geq 1\\ i_1+\cdots+i_d\leq n}} \frac{1}{i_1 i_2 \cdots i_d}\right). \tag{9}$$

We can compute an upper bound on $\overline{E}(d,n)$ by relaxing the summation,

$$\overline{E}(d,n) \leq \frac{f^d}{d!}\left(\sum_{1\leq i_1,\ldots,i_d\leq n} \frac{1}{i_1 i_2 \cdots i_d}\right) = \frac{f^d}{d!}\left(\sum_{1\leq i\leq n}\frac{1}{i}\right)^d = \frac{(f\cdot H_n)^d}{d!}. \tag{10}$$

Since this is also an upper bound on $E(d,n)$ this completes the proof. $\square$

Theorem 1 *The expected depth of the random circuit after n gates are added satisfies*

$$\mathtt{E}\left(\mathrm{Depth}_n(P)\right) \leq f\cdot e\cdot H_n + O(1).$$

Proof Observe that for any d, $\mathrm{Depth}_n(P)\geq d$ implies $L(d,n)\geq 1$. Thus

$$\Pr\left(\mathrm{Depth}_n(P)\geq d\right) \leq E(d,n) \leq \frac{(f\cdot H_n)^d}{d!} \tag{11}$$

where the last inequality follows from Lemma 2. Inserting Stirling's formula $d! = (d/e)^d\sqrt{2\pi d}(1+O(1/d))$, into this equation yields

$$\Pr\left(\mathrm{Depth}_n(P)\geq d\right) \leq \left(\frac{feH_n}{d}\right)^d \frac{1}{\sqrt{2\pi d}}\left(1+O\left(\frac{1}{d}\right)\right). \tag{12}$$

Thus, for any $x>0$,

$$\begin{aligned}
\mathtt{E}\left(\mathrm{Depth}_n(P)\right) &\leq x + \sum_{d>x}\Pr\left(\mathrm{Depth}_n(P)\geq d\right) \qquad (13)\\
&\leq x + \sum_{d>x}\left(\frac{feH_n}{d}\right)^d \frac{1}{\sqrt{2\pi d}}\left(1+O\left(\frac{1}{d}\right)\right)
\end{aligned}$$

We will now show that for any $x>0$

$$\sum_{d>x}\left(\frac{x}{d}\right)^d = O(1). \tag{14}$$

Setting $x = \lceil feH_n\rceil$ and inserting the last inequality into the one that preceeds it will conclude the proof of the lemma. To show (14) note that

$$\begin{aligned}
\sum_{d>x}\left(\frac{x}{d}\right)^d &\leq \sum_{0\leq i\leq x}\left(\frac{1}{1+\frac{i}{x}}\right)^x + \sum_{d>2x}\left(\frac{x}{d}\right)^d\\
&\leq \sum_{0\leq i\leq x} e^{-i/2} + \sum_{d>2x}\left(\frac{1}{2}\right)^d = O(1)
\end{aligned}$$

Process $\mathcal{P}'$:

- $d'_0(0) = 0$
- For each n

1. Choose f integers $i_n^1, \ldots, i_n^f$ where each i_n^j is chosen randomly from $\{0, 1, 2, \ldots, n-1\}$.
2. The multiset $S_n = \{i_n^1, \ldots, i_n^f\}$ contains the *inputs* to gate n.
3. Set the depth of the gate $d'_n(n-1)$ to $1 + \max_{j \le f} d'_{i_n^j}(n-1)$.
4. For all $i \le n$ set $d'_i(n)$ to some value $\le d'_i(n-1)$ using some arbitrary rule.

Figure 3: A process $\mathcal{P}'$ that generates modified random circuits

where the second inequality follows from the fact that $(1+u)^{-1} \le e^{-u/2}$ for $0 \le u \le 1$. $\square$

3 Lower Bounds on the Expected Circuit Depth

In this section we describe how to derive lower bounds on the expected depths of random circuits. Our approach will be to construct various modified random circuits P', such that $\mathtt{E}(\mathrm{Depth}_n(P')) \le \mathtt{E}(\mathrm{Depth}_n(P))$. Because of the way that they are constructed we will easily be able to lower bound $\mathtt{E}(\mathrm{Depth}_n(P'))$, thus effectively bounding $\mathtt{E}(\mathrm{Depth}_n(P))$ as well.

More specifically we will show how to construct a Markov chain that approximates P' where the state of the Markov chain is the number of gates on the deepest level of P'. Calculating the stationary probabilities of the Markov chain will then permit finding out how often a new level is started, letting us in turn calculate the expected depth of P'.

This section is split into three parts. In the first we prove two lemmas that will be crucial for establishing the lower bound. In the second we describe a family of modified circuit processes, P_ϕ, $\phi > 1$ and analytically calculate $\mathtt{E}(\mathrm{Depth}_n(P_\phi))$ as a function of ϕ. We then figure out the best lower bound that this approach can yield. In the third part we extend our techniques to examine more complicated modified circuits, again parameterized by some variable.

3.1 Basic Lemmas

In this subsection we present two lemmas upon which our analysis will be based. We say that P' is a process that generates a *modified random circuit* if it is defined as in Figure 3.

Note that this process generates random circuits a gate at a time in the same way that $\mathcal{P}$ does. The difference between $\mathcal{P}$ and $\mathcal{P}'$ is that $\mathcal{P}'$ sometimes decreases the depths of some gates using some arbitrary rule. Since gates can have their depths decreased at any time we can no longer talk about the absolute depth d_i of gate i but instead must specify $d'_i(n)$, the depth of gate i after the nth step of the modified process.

Because gates can only have their depths decreased and not increased it is natural to suspect that, for any modified circuit P', $\mathrm{Depth}(P') \le \mathrm{Depth}(P)$. In fact this is true, as stated in the next lemma.

Lemma 3

$$\mathtt{E}\left(\mathrm{Depth}_n(P')\right) \le \mathtt{E}\left(\mathrm{Depth}_n(P)\right).$$

Proof

Let $C = (S_1, S_2, \ldots, S_n)$ be a sequence of multisets in which $S_i \in \{0, 1, \ldots, i-1\}^f$. Fixing C actually fixes the action of the first n steps of both $\mathcal{P}$ and $\mathcal{P}'$. We can therefore, for $j \le n$, talk about $d_i(C)$ and $d'_i(j, C)$, the depths of the gates (at time j) given C, and similarly, $\mathrm{Depth}_j(P, C)$ and $\mathrm{Depth}_j(P', C)$. By induction it is clear that

$$\forall j \le n, \forall i \le j, \quad d'_i(j, C) \le d_i(C). \tag{15}$$

This is obviously true for the base case $n = 1$. Assume then it is true for $C = (S_1, S_2, \ldots, S_{n-1})$. To prove it is true for $C = (S_1, S_2, \ldots, S_n)$ first note that the induction hypothesis immediately tells us that

$$\forall j \le n-1, \forall i \le j, \quad d'_i(j, C) \le d_i(C).$$

Thus

$$\forall i \le n-1, \quad d'_i(n, C) \le d'_i(n-1, C) \le d_i(C)$$

and

$$d'_n(n, C) = 1 + \max_{j \le f} d'_{i_n^j}(n-1, C) \le 1 + \max_{j \le f} d_{i_n^j}(C) = d_n(n, C)$$

proving (15). This equation then implies

$$\mathrm{Depth}_n(P', C) \le \mathrm{Depth}_n(P, C).$$

Let $\mathcal{C}$ be the collection of all possible sequences C. Then

$$\begin{aligned} \mathtt{E}\left(\mathrm{Depth}_n(P')\right) &= \frac{1}{|\mathcal{C}|} \sum_{C \in \mathrm{C}} \mathrm{Depth}_n(P', C) \\ &\le \frac{1}{|\mathcal{C}|} \sum_{C \in \mathrm{C}} \mathrm{Depth}_n(P, C) = \mathtt{E}\left(\mathrm{Depth}_n(P)\right). \end{aligned}$$

□

The second lemma we will need generalizes the ergodic theory of homogeneous Markov chains to special time-dependent ones.

In what follows if $M = (m_{i,j})$ is a matrix then $|M| = \sup_j \sum_i |m_{i,j}|$, i.e., the supremum of the column (absolute value) sums and, if $\pi = (\pi(1), \pi(2), \ldots)$ is a vector, then $|\pi| = \sup_j \pi(j)$.

Lemma 4 *Let M be a Markov-chain with time-dependent transition probabilities. Let A_r be the transition matrix for the r-th step. Suppose A is the transition matrix for some irreducible aperiodic time-independent chain and A has a stationary distribution π. Suppose further that $|A_r - A| = O(\phi^{-r})$ for some $\phi > 1$. Let π_0 be any distribution and for $r > 0$, set $\pi_r = A_{r-1}\pi_{r-1}$. Then*

$$|\pi_r - \pi| \to 0$$

Proof We will prove the lemma by showing that $\forall \epsilon > 0$, $\exists m, K$ such that $\forall r > m + K$, $|\pi_r - \pi| < \epsilon$.

Let $\epsilon > 0$. First note that since A is irreducible and has a stationary distribution, standard theorems on Markov chains [2, p. 208] imply that all the states in A are non-null and persistent. Now, since A is non-null persistent, irreducible and aperiodic, other standard theorems, e.g., [2, p. 214, note (c)], then tell us that given any initial distribution $\overline{\pi}$, $A^k\overline{\pi} \to \pi$ so

$$|A^k(\overline{\pi} - \pi)| \to 0 \tag{16}$$

Next note that, $\forall m > 0$,

$$\begin{aligned}
\pi_{m+1} - \pi &= A_m\pi_m - A\pi \\
&= A_m\pi_m - A\pi_m + A\pi_m - A\pi \\
&= (A_m - A)\pi_m + A(\pi_m - \pi).
\end{aligned}$$

Iterating this equation yields

$$\pi_{m+k} - \pi = \sum_{l=0}^{k-1} A^{k-1-l}(A_{m+l} - A)\pi_{m+l} + A^k(\pi_m - \pi). \tag{17}$$

We now show that the sum on the right side of this equation tend, componentwise, to 0 as $k \to \infty$:

First suppose r is fixed and $B = A_r - A$; denote the components of B by $b_{i,j}$. The lemma condition $|A_r - A| = O(\phi^{-r})$ translates into $\forall j$, $\sum_i |b_{i,j}| = O(\phi^{-r})$, where the constant implicit in the $O()$ is independent of j. Now, given *any* distribution $\pi' = (\pi'(1), \pi'(2) \ldots)$ set $\sigma = (A_r - A)\pi' = B\pi'$. We then find that

$$\begin{aligned}
\sum_i |\sigma(i)| &\le \sum_{i,j} |b_{i,j}|\pi'(j) \\
&= \sum_j \pi'(j) \left(\sum_i |b_{i,j}| \right) \\
&= \sum_j \pi'(j) O(\phi^{-r}) \\
&= O(\phi^{-r})
\end{aligned}$$

where the last equality follows from the fact that $\sum_j \pi'(j) = 1$.

Now let $r = m + l$, $\pi' = \pi_{m+l}$ and set

$$\delta = A^{k-1-l}(A_{m+l} - A)\pi' = A^{k-1-l}\sigma.$$

Since every entry in A^{k-1-l} has absolute value at most 1 we find that $\forall j$, $\delta(j) \le \sum_i |\sigma(i)| = O(\phi^{-(m+l)})$. Thus

$$\begin{aligned}
\left| \sum_{l=0}^{k-1} A^{k-1-l}(A_{m+l} - A)\pi_{m+l} \right| &\le \sum_{l=0}^{k-1} \left| A^{k-1-l}(A_{m+l} - A)\pi_{m+l} \right| \\
&= \sum_{l=0}^{k-1} O(\phi^{-(m+l)}) = O(\phi^{-m})
\end{aligned}$$

where the constant implicit in the $O(\phi^{-m})$ term is *independent of k.* Thus

$$|\pi_{m+k} - \pi| \le O(\phi^{-m}) + |A^k(\pi_m - \pi)|. \tag{18}$$

Given ϵ we can therefore find m such that the $O(\phi^{-m})$ term is less that $\epsilon/2$. Once m is fixed equation (16) tells us that $\exists K$ such that $\forall k > K$, $|A^k(\pi_m - \pi)| < \epsilon/2$. Thus $\forall r > m + K$, $|\pi_r - \pi| < \epsilon/2 + \epsilon/2 = \epsilon$ proving the theorem. $\square$

3.2 The First Lower Bound

In this subsection we describe our first method of calculating lower bounds. We create a family of processes $\mathcal{P}_\phi$, $\phi > 1$ that create modified random circuits P_ϕ and bound the expected depths of the P_ϕ. This will permit us to prove the following theorem:

Theorem 2 *Let P be a random circuit grown by process $\mathcal{P}$. Then, for every $\epsilon > 0$,*

$$\mathtt{E}\left(\mathrm{Depth}_n(P)\right) \ge \left(\frac{f}{e-2} - \epsilon \right) \ln n + O(1).$$

Note that $1/(e-2) \sim 1.39\ldots$.

For a given fixed $\phi > 1$ the process $\mathcal{P}_\phi$ works in stages, during stage r adding gates n, $\lceil \phi^r \rceil < n \leq \lceil \phi^{r+1} \rceil$. Within a stage, while adding the gates, process $\mathcal{P}_\phi$ behaves exactly the same as process $\mathcal{P}$, picking the inputs of the gates at random and placing a gate one level deeper than the maximum depth of its parents. It is only at the end of a stage that the process will modify the circuit, decreasing the depths of some gates using a rule to be described below. The essential fact about the process P_ϕ is that its 'level-decreasing' rule will be defined so that between any two consecutive stages exactly one of the following transitions must occur:

T1 The depth of P_ϕ increases by one and the new deepest level contains only one gate.

T2 The depth of P_ϕ remains the same and the number of gates on the deepest level increases by one.

T3 The depth of P_ϕ remains the same and the number of gates on the deepest level does not change.

Because of this fact we will be able to define a Markov chain whose state at time r is the number of nodes on the deepest level of P_ϕ after the conclusion of stage r. Calculating the stationary distribution of this chain will permit us to calculate how often a new level is added to P_ϕ, letting us lower bound $\mathtt{E}\left(\mathrm{Depth}_n(P_\phi)\right)$ and thus via Lemma 3.1, lower bounding $\mathtt{E}\left(\mathrm{Depth}_n(P)\right)$ as well. We now describe the process P_ϕ in detail.

It starts in stage 0 by placing gate 0 at depth 0 and gate 1 at depth 1.

At the end of stage $r-1$ the process knows three pieces of information describing the result of stage $r-1$: $(i)\, l_{r-1} = \mathrm{Depth}_{\lceil \phi^r \rceil}(P_\phi)$, $(ii)\, S_{r-1} = \{i \leq \lceil \phi^r \rceil \,:\, d_i'(\lceil \phi^r \rceil) = l_{r-1}$, and (iii) some gate g_{r-1} such that $d'_{g_{r-1}}(\lceil \phi^r \rceil) = l_{r-1}-1$

Before starting stage $r=1$ the process thus sets $l_0 = 0$, $g_0 = 0$ and $S_0 = \{1\}$.

Assume now that the process knows l_{r-1}, g_{r-1} and S_{r-1} and is ready to start stage r. It then adds gates n, $\lceil \phi^r \rceil < n \leq \lceil \phi^{r+1} \rceil$ in the same way as $\mathcal{P}$ would. When it completes the stage there are three possibilities

1. Some gate $g' \in S_{r-1}$ was chosen as the parent of one of the newly created gates, call it g'', added during the stage. In this case, the depth of g'' is at least $l_{r-1}+1$.

2. The first case did not hold but g_{r-1} was the parent of some gate, call it g'. Note that the depth of g' is at least l_r.

3. Neither of the first two cases occur.

The corresponding actions taken by $\mathcal{P}_\phi$ are

1. Move g'' to depth $l_{r-1}+1$. Set $l_r = l_{r-1}+1$, $g_r = g'$, and $S_r = \{g''\}$.

2. Move g' to depth l_{r-1}. Set $l_r = l_{r-1}$ and $S_r = S_{r-1} \cup \{g'\}$.

3. Nothing.

After completing these actions the process then moves *all* gates at level l_r or greater except for those in S_r to level $l_r - 1$. Note that after these modifications the circuit P_ϕ knows the correct values for l_r, g_r, and S_r as defined in (i), (ii) and (iii). Also note that, by definition, the process $\mathcal{P}_\phi$ makes exactly one of the transitions **T1**, **T2**, or **T3** each time it advances from one state to the next.

We now define L to be the random process $L = (l_0, l_1, l_2, l_3, \ldots)$ where $l_r = \text{Depth}_{\lceil \phi^{r+1} \rceil}(P_\phi)$ as defined above. We also set $|S_r| = i$ and

$$\begin{aligned}
p_i^r &= \Pr\left(\text{Transition } \mathbf{T1} \text{ occurs after stage } r+1 \;\middle|\; |S_r| = i\right) \\
q_i^r &= \Pr\left(\text{Transition } \mathbf{T2} \text{ occurs after stage } r+1 \;\middle|\; |S_r| = i\right) \\
z_i^r &= \Pr\left(\text{Transition } \mathbf{T3} \text{ occurs after stage } r+1 \;\middle|\; |S_r| = i\right)
\end{aligned}$$

and let

$$\pi_i^r = \Pr\left(|S_r| = i\right).$$

Since the depth of P_ϕ increases (by one) during stage r if and only if transition **T1** occurs this means

$$\Pr\left(P_\phi \text{ increases by one during stage } r\right) = \sum_{j \geq 1} \pi_j^{r-1} p_j^{r-1}.$$

But because

$$\mathtt{E}\left(\text{Depth}_n(P)\right) \geq \mathtt{E}\left(\text{Depth}_{\phi^{\lfloor \log_\phi n \rfloor}}(P_\phi)\right)$$

this implies that

$$\mathtt{E}\left(\text{Depth}_n(P)\right) \geq \sum_{1 \leq r \leq \log_\phi n} \left(\sum_{j \geq 1} \pi_j^{r-1} p_j^{r-1}\right). \tag{19}$$

The remainder of this section is devoted to proving Theorem 2 by evaluating the right hand side of the last equation. We start by showing that p_i^r, q_i^r and z_r^i converge rather quickly to constants.

Lemma 5

$$\begin{aligned}
p_i^r &= 1 - \phi^{-if} + O\left(\phi^{-r}\right) && (20) \\
q_i^r &= \phi^{-if}\left(1 - \phi^{-f}\right) + O\left(\phi^{-r}\right) && (21) \\
z_i^r &= \phi^{-(i+1)f} + O\left(\phi^{-r}\right) && (22)
\end{aligned}$$

where the constants implicit in the $O()$ notation are independent of i.

Proof First note that $|S_r| \le r+1$ because $|S_0| = 1$ and the value of S_r can only increase by 1 between a stage and the one following it. This means that p_i^r, q_i^r and z_i^r are undefined for $i = |S_r| > r+1$ and we are able to set their values to whatever we want. For convenience, we set their values to respectively be $1-\phi^{-if}$, $\phi^{-if}\left(1-\phi^{-f}\right)$, and $\phi^{-(i+1)f}$.

Thus we will only be interested in analyzing situations in which $i \le r+1$. For technical reasons our proofs will require i/ϕ^r and i^2/ϕ^r to be bounded by a constant, say $1/2$. To ensure this we will therefore assume that r is taken large enough so that $\frac{i^2}{\phi^r} \le \frac{(r+1)^2}{\phi^r} < \frac{1}{2}$.

To prove (20) it now suffices to note that if $i = |S_r|$ then **T1** *does not* occur only if all of the i gates on level l_r are never chosen during the $(r+1)$st stage. Thus

$$1 - p_i^r = \prod_{\lceil \phi^{r+1} \rceil < m \le \lceil \phi^{r+2} \rceil} \left(1 - \frac{i}{m}\right)^f = \exp\left(\sum_m f \cdot \ln\left(1 - \frac{i}{m}\right)\right) \tag{23}$$

But

$$\begin{aligned} \sum_m \ln(1 - i/m) &= -\sum_m i/m + O\left(i^2/m^2\right) \\ &= -i\ln\frac{\lceil \phi^{r+2} \rceil}{\lceil \phi^{r+1} \rceil + 1} + O\left(i^2\phi^{-r}\right) = -i\ln\phi + O\left(i^2\phi^{-r}\right) \end{aligned}$$

where the constant implicit in the $O()$ on the right side of the first equality is independent of i because $i/m < 1/2$.

Thus

$$1 - p_i^r = \exp\left(-if\ln\phi + O\left(fi^2\phi^{-r}\right)\right) = \phi^{-if}\left[1 + O\left(fi^2\phi^{-r}\right)\right] = \phi^{-if} + O\left(\phi^{-r}\right) \tag{24}$$

where the second equality follows from (recall our assumption) $i^2\phi^{-r} < 1/2$ so $e^{O\left(fi^2\phi^{-r}\right)} = 1 + O\left(fi^2\phi^{-r}\right)$ with the constant implicit in the $O()$ on the right hand side independent of r; the last equality follows from the fact that $\phi^{-if}i^2 f = O(1)$.

To prove (22) a similar calculation shows that

$$z_i^r = \prod_{\lceil \phi^{r+1} \rceil < m \le \lceil \phi^{r+2} \rceil} \left(1 - \frac{i+1}{m}\right)^f = \phi^{-(i+1)f} + O\left(\phi^{-r}\right).$$

Equation (21) then follows from $q_i^r = 1 - p_i^r - z_i^r$. $\square$

Next observe that p_i^r, q_i^r, z_i^r do *not* depend on the total circuit P_ϕ under construction but only upon its state i. Thus

$$\pi_i^{r+1} = \begin{cases} \sum_{j\ge1} \pi_j^r p_j^r + \pi_1^r z_1^r & \text{for } i = 1 \\ \pi_{i-1}^r q_{i-1}^r + \pi_i^r z_i^r & \text{for } i \ge 2 \end{cases} \tag{25}$$

and process L is a time-dependent Markov chain with transition probabilities given by (25) and initial distribution $\pi_0 = (1, 0, 0, \ldots)$. Let A_r be the transition matrix for L at time r.

Now define

$$p_i = 1 - \phi^{-if}, \quad q_i = \phi^{-if}\left(1 - \phi^{-f}\right), \quad z_i = \phi^{-(i+1)f},$$

and let M be the homogeneous (time-independent) Markov chain with transition matrix A given by

$$a_{ij} \quad = \quad \begin{cases} p_i & \text{if } j = 1 \\ q_i & \text{if } j = i+1 \\ z_i & \text{if } j = i \\ 0 & \text{otherwise.} \end{cases} \tag{26}$$

If we let $B = A_r - A$ and set the entries of B to be $B_{i,j}$ then we find from Lemma 5 that

$$\sum_j |b_{i,j}| = (p_i - p_i^r) + (q_i - q_i^r) + (z_i - z_i^r) = O(\phi^{-r}).$$

This simply says that

$$|A_r - A| = O(\phi^{-r}).$$

We now show, by calculation, that A has a stationary distribution. We will then use this fact to permit the application of Lemma 4. Suppose then that A has some stationary distribution, π, i.e., there exists a distribution $\pi = (\pi_1, \pi_2, \pi_3, \ldots)$ satisfying

$$\pi_1 = \sum_{j \geq 1} \pi_j p_j + \pi_1 z_1 \tag{27}$$

and

$$\pi_i \quad = \quad \pi_{i-1} q_{i-1} + \pi_i z_i \quad \text{for } i \geq 2. \tag{28}$$

Then, for $i \geq 2$,

$$\pi_i = \frac{\pi_{i-1} q_{i-1}}{1 - z_i} = \frac{s^{i-1}(1-s)}{1 - s^{i+1}} \pi_{i-1} \tag{29}$$

where $s = \phi^{-f}$. We unwind this last equation to find

$$\pi_i = \frac{s^{i-1}(1-s)}{1-s^{i+1}} \cdot \frac{s^{i-2}(1-s)}{1-s^i} \cdots \frac{s(1-s)}{1-s^3} \cdot \pi_1, \tag{30}$$

which can be rewritten as,

$$\pi_i = s^{(i^2-i)/2} \left[\frac{1-s}{1-s^{i+1}} \cdot \frac{1-s}{1-s^i} \cdots \frac{1-s}{1-s^3} \right] \cdot \pi_1. \tag{31}$$

Recalling that $\sum_i \pi_i = 1$ we solve to find

$$\pi_1 = \left[1 + \sum_{i\geq 2}\left(s^{(i^2-i)/2}\left[\frac{1-s}{1-s^{i+1}}\cdot\frac{1-s}{1-s^i}\cdots\frac{1-s}{1-s^3}\right]\right)\right]^{-1}$$

and for $i \geq 2$ define π_i by equation (31). By substitution we find that this distribution satisfies (27) and (28) and is therefore a stationary one for A.

We have just proven that A has a stationary distribution. It is easy to see that A is irreducible and aperiodic and thus we may apply Lemma 4 to find that $|\pi^r - \pi| \to 0$.

Now set

$$c_\phi = \pi_1 - \pi_1 z_1 = \sum_{j\geq 1}\pi_j p_j.$$

Note that $\sum_{j\geq 1}\pi_j p_j \leq 1$, and $\forall r\, p_j,\, p_j^{r-1} \leq 1$. Furthermore $\forall j,\, p_j^r \to p_j$ and, because $|\pi^r - \pi| \to 0$, $\forall j,\, \pi_j^r \to \pi_j$. Thus $\pi_j^{r-1}p_j^{r-1} \to \pi_j p_j$ and the Lesbegue dominated convergence theorem implies that

$$\lim_{r\to\infty}\left(\sum_{j\geq 1}\pi_j^{r-1}p_j^{r-1}\right) = \left(\sum_{j\geq 1}\pi_j p_j\right) = c_\phi.$$

Plugging back into equation (19) we find that for every $\epsilon > 0$

$$\mathtt{E}\left(\text{Depth}_n(P)\right) \geq \frac{\ln n}{\ln\phi}(c_\phi - \epsilon) + O(1). \tag{32}$$

To prove Theorem 2 it will therefore suffice to show that $\lim_{\phi\to 1}(c_\phi/\ln\phi) = f/(e-2)$.

Let $\tilde{\pi}_i = \lim_{\phi\to 1}\pi_i$ (we omit the straightforward but tedious formal justification of the existence of $\tilde{\pi}_i$). Taking the limit as $\phi \to 1$ of equation (31) yields

$$\tilde{\pi}_i = \left[\frac{1}{i+1}\cdot\frac{1}{i}\cdots\frac{1}{3}\right]\cdot\tilde{\pi}_1 = \frac{2}{(i+1)!}\cdot\tilde{\pi}_1. \tag{33}$$

Equating the sum of the steady state probabilities to one, we get

$$2\tilde{\pi}_1\sum_{i\geq 1}\frac{1}{(i+1)!} = 1 \tag{34}$$

or

$$\tilde{\pi}_1 = \frac{1}{2(e-2)}. \tag{35}$$

For any specific ϕ we can rewrite Eq. (27) as follows

$$c_\phi = \sum_{j\geq 1}\pi_j p_j = \pi_1(1-z_1) = \pi_1(1-s^2). \tag{36}$$

Thus, recalling that $s = \phi^{-f}$ we find that

$$\lim_{\phi \to 1} \frac{c_\phi}{\ln \phi} = \lim_{\phi \to 1} \pi_1 \frac{1 - s^2}{\ln \phi} = -f \tilde{\pi}_1 \lim_{s \to 1} \frac{1 - s^2}{\ln s} = 2 f \tilde{\pi}_1 = \frac{f}{e - 2} \tag{37}$$

completing the proof of the theorem.

3.3 Improved Lower Bounds

The modified random circuit process $\mathcal{P}_\phi$ analyzed in the previous subsection keeps track of the number of gates at the deepest level of the random circuit $\mathcal{P}$; the number of gates at the second deepest level is essentially assumed to be one; and the gates at all other levels are ignored. In the event that a new level is not added during a given stage, this has the effect of underestimating the likelihood that a new gate is added to the deepest level; in turn, this results in underestimating the likelihood of adding a new level in the next stage. This suggests that we might obtain a better lower bound on the expected circuit depth by considering modified random circuits that keep track of the number of gates at two, three, or more of the deepest levels.

We first illustrate the idea for a modified process with two levels; the generalization to more than two levels will be obvious. As before, the process $\mathcal{P}_\phi$ works in stages, operating exactly like process $\mathcal{P}$, within a stage, but activating level-decreasing rules at the end of each stage. These ensure that between two stages, exactly one of the following transitions occur:

T1 The depth of P_ϕ increases by one and the new deepest level contains only one gate. The number of gates on the 2nd deepest level (the old deepest one) remains the same.

T2 The depth of P_ϕ remains the same and the number of gates on the deepest level increases by one. The number of gates on the 2nd deepest level remains the same.

T3 The depth of P_ϕ and the number of gates on the deepest level remains the same and the number of gates on the second deepest level increases by one.

T4 The depth of P_ϕ and the number of gates on the deepest and second deepest levels all remain the same.

As before, we can model process $\mathcal{P}_\phi$ by a Markov chain. The state of the Markov chain at time r is given by the number of nodes at the deepest and second deepest levels of P_ϕ at the end of stage r. The details are similar to those developed for the one level Markov chain and are given in the Appendix.

It follows that for every $\epsilon > 0$

$$\mathrm{E}\left(\mathrm{Depth}_n(P)\right) \geq \frac{\ln n}{\ln \phi}(c_\phi - \epsilon) + O(1), \tag{38}$$

B_0	B_1	B_2	B_3	B_4	Lower Bound	B_0	B_1	B_2	B_3	B_4	Lower Bound
1	1	1	1	1	$1.0000 f \ln n$	5	5	4	1	1	$1.9072 f \ln n$
2	1	1	1	1	1.3333... ”	5	6	4	1	1	1.9295... ”
2	2	1	1	1	1.4545... ”	5	6	5	1	1	1.9491... ”
3	2	1	1	1	1.5673... ”	6	6	5	1	1	1.9652... ”
3	3	1	1	1	1.6510... ”	6	7	5	1	1	1.9817... ”
4	3	1	1	1	1.6896... ”	6	7	6	1	1	1.9970... ”
4	4	1	1	1	1.7337... ”	6	7	6	2	1	2.0120... ”
4	4	2	1	1	1.7896... ”	6	7	7	2	1	2.0231... ”
4	4	3	1	1	1.8201... ”	6	7	7	3	1	2.0352... ”
4	5	3	1	1	1.8479... ”	6	8	7	3	1	2.0479... ”
5	5	3	1	1	1.8819... ”						

Table 1: Lower bounds on expected circuit depth. B_i is the maximum number of gates allowed on depth l_{r-i} in the modified random circuit process.

where c_ϕ is a function of the steady state probabilities of this chain. This can be done not only for the two level process but also for processes with three, four, or more levels. However, unlike the Markov chain for the one level process, we do not know how to analytically compute the steady state probabilities for the infinite Markov chains corresponding to two or more levels as functions of ϕ and to take the limit as $\phi \to 1$. Our approach therefore is to modify the process further by limiting the maximum number of gates allowed at each of the deepest levels. The resulting process can be modelled by a finite Markov chain. We are then able to compute the steady state probabilities of the finite Markov chain using an equation solver (Maple). In fact, we can directly obtain the value of $\lim_{\phi \to 1}(c_\phi / \ln \phi)$ by solving the equations.

We used Maple to solve the equations for a finite Markov chain corresponding to a five level process for various numbers of states permitted at each level. The results are shown in Table 1. Each entry of the table shows the maximum number of gates allowed at the five deepest levels along with the lower bound achieved. To choose the number of gates permitted on each level, we started with at most one gate allowed on each of the five deepest levels. Each subsequent entry was obtained by looking at the previous entry and attempting to increase the maximum number of gates allowed by one in exactly one out of the five deepest levels. The choice that led to the best lower bound is the one shown in the table. The last row of the table shows the best lower bound of $2.04 \ldots f \ln n$ obtained by using this approach. Of course, by building larger Markov chains we could improve this lower bound at the expense of requiring more computer time to solve the equations. We note that every "solution" is actually a proof of a lower bound. The way we build the Markov chains tells us that the lower

Levels	Simulated Lower Bound
1	$1.39 f \ln n$
2	1.78 "
3	2.08 "
4	2.26 "
5	2.37 "
10	2.44 "
20	2.44 "

Table 2: Simulated lower "bounds" on expected circuit depth. These "bounds" have not been proven but are suggested by the results of Markov chain simulations.

bound to the circuit depth can actually be expressed as a function of the steady state probabilities of the Markov chain. Solving the equations to find the steady state probabilities is therefore actually a *proof* of a lower bound. The results in the table can therefore be viewed as the results of a heuristic search for better proofs.

Since the computer time needed to solve equations quickly becomes enormous, we also directly simulated the Markov chains to compute the steady state probabilities. This was done as follows. Since the best lower bound is obtained in the limiting case as $\phi \to 1$, we chose $s = \phi^{-f} = 0.999$. We started the simulation from the state of the Markov chain that corresponds to having one gate at each of the deepest levels. At each step, using a random number generator, we carried out a transition to another state in accordance with the probabilities of the Markov chain. This step was repeated 10^8 times to get rid of any biases introduced by the initial probability distribution. Then the process was continued for another 10^8 steps, during which the time spent in each state was recorded; the time spent in a state as a fraction of the total number of steps should approach the steady state probability. In order to obtain the best results, we did not place any limit on the maximum number of gates allowed at each level. We did the simulation for Markov chains corresponding to 1, 2, 3, 4, 5, 10 and 20 level processes. The corresponding lower bounds are shown in Table 2. The best lower bound we obtained is $2.44 \ldots f \ln n$ for a 10 level Markov chain; using more levels than 10 barely improved the lower bound. Note that these are not really lower bounds; these are the simulation results for the Markov chains which should be telling us with great accuracy (but not total certainty) the values of the stationary probability distribution, which in turn give us the lower bound. If the stationary distribution values are accurate, then using a Markov chain equation solver will yield these values, formally proving the lower bounds.

References

[1] J. Diaz, M. J. Serna, P. Spirakis and J. Toran. On the expected depth of Boolean circuits. Report LSI-94-7-R, Departament de Llenguatges i Sistemes, Universitat Politècnica Calalunya Pau Gargallo 5, Barcelona, 1994.

[2] G. R. Grimmett and D.R. Stirzaker. *Probability and Random Processes (2nd ed.)*. Clarendon Press - Oxford, Oxford. 1992.

[3] H. Mahmoud and R. T. Smythe. On the distribution of leaves in rooted subtrees of recursive trees. *Ann. Appl. Probab.*, 1:406–418, 1991.

[4] A. Meir and J. Moon. Recursive trees with no nodes of outdegree one. *Congr. Numer.*, 66:49–62, 1974.

[5] J. Moon. The distance between nodes in recursive trees. In *London Math. Soc. Lecture Notes, Ser 13*, pages 125–132, Cambridge University Press, London, 1974.

[6] H. Na and A. Rapaport. Distribution of nodes of a tree by degree. *Math. Biosci.*, 6:313–329, 1970.

[7] B. Pittel. Notes on the Heights of Random Recursive Trees and Random M-ary Search Trees, *Random Structures and Algorithms*, 5(2):337–347, 1994.

[8] J. Szymanski. On the maximum degree and the height of a random recursive tree. In M. Karonski, J. Jaworski, and A. Rucinski, editors, *Random Graphs '87*, pages 431–524, Wiley, New York, 1990.

[9] T. Tsukiji and F. Xhafa On the Depth of Randomly Generated Circuits. Appearing in the European Symposium on Algorithms (ESA96) 1996. Extended version appearing as an unpublished manuscript.

4 Appendix

We present more details on how to model the two level modified random circuit process using a Markov chain (see Section 3.3). The process $\mathcal{P}_\phi$ knows four pieces of information at the end of stage $r-1$: the depth of the modified circuit l_{r-1}; sets $S_{1,r-1}$ and $S_{2,r-1}$ of gates at depths l_{r-1} and $l_{r-1}-1$, respectively; some gate g_{r-1} at depth $l_{r-1}-2$. On completing a stage there are four possibilities:

1. Some gate $g' \in S_{1,r-1}$ was chosen as the parent of one of the newly created gates, call it g'', added during the stage. In this case, the depth of g'' is at least $l_{r-1}+1$.
2. The first case did not hold but some gate $g' \in S_{2,r-1}$ was chosen as the parent of one of the newly created gates, call it g'', added during the stage. In this case the depth of g'' is at least l_{r-1}.
3. The first two cases did not hold but g_{r-1} was the parent of some gate, call it g'. Note that the depth of g' is at least $l_{r-1}-1$.
4. None of the first three cases occur.

The corresponding actions taken by $\mathcal{P}_\phi$ are

1. Move g'' to depth $l_{r-1}+1$. Set $l_r = l_{r-1}+1$, g_r to be any gate from $S_{2,r-1}$, $S_{2,r} = S_{1,r-1}$ and $S_{1,r} = \{g''\}$.
2. Move g'' to depth l_{r-1}. Set $l_r = l_{r-1}$ and $S_{1,r} = S_{1,r-1} \cup \{g''\}$.
3. Move g' to depth $l_{r-1}-1$. Set $l_r = l_{r-1}$ and $S_{2,r} = S_{2,r-1} \cup \{g'\}$.
4. Nothing.

After completing these actions the process then moves *all* gates at level $l_r - 1$ or greater except for those in $S_{1,r}$ and $S_{2,r}$ to level $l_r - 2$. The probabilities of the transitions are defined as:

$$\begin{aligned}
p^r_{i,j} &= \Pr\left(|S_{1,r}| = i, |S_{2,r}| = j \text{ and transition } \mathbf{T1} \text{ occurs after stage } r+1\right) \\
q^r_{i,j} &= \Pr\left(|S_{1,r}| = i, |S_{2,r}| = j \text{ and transition } \mathbf{T2} \text{ occurs after stage } r+1\right) \\
t^r_{i,j} &= \Pr\left(|S_{1,r}| = i, |S_{2,r}| = j \text{ and transition } \mathbf{T3} \text{ occurs after stage } r+1\right) \\
z^r_{i,j} &= \Pr\left(|S_{1,r}| = i, |S_{2,r}| = j \text{ and transition } \mathbf{T4} \text{ occurs after stage } r+1\right)
\end{aligned}$$

and let

$$\pi^r_{i,j} = \Pr\left(|S_{1,r}| = i, |S_{2,r}| = j\right).$$

We can derive the transition probabilities using the same method as for the one level Markov chain.

Lemma 6

$$
\begin{aligned}
p_{i,j}^{r} &= 1-\phi^{-if}+O\left(\phi^{-r}\right) && (39)\\
q_{i,j}^{r} &= \phi^{-if}\left(1-\phi^{-jf}\right)+O\left(\phi^{-r}\right) && (40)\\
t_{i,j}^{r} &= \phi^{-(i+j)f}\left(1-\phi^{-f}\right)+O\left(\phi^{-r}\right) && (41)\\
z_{i,j}^{r} &= \phi^{-(i+j+1)f}+O\left(\phi^{-r}\right) && (42)
\end{aligned}
$$

We let $p_{i,j}, q_{i,j}, t_{i,j}$, and $z_{i,j}$ denote the limit values of these transition probabilities. Arguing exactly as for the one level Markov chain we can show that for every $\epsilon > 0$

$$\mathtt{E}\left(\mathrm{Depth}_n(P)\right) \geq \frac{\ln n}{\ln\phi}(c_\phi-\epsilon)+O(1), \tag{43}$$

where $c_\phi = \sum_{i,j}\pi_{i,j}p_{i,j}$.