Journal of Prime Research in Mathematics Vol. **7** *(2011), 09-18*

OUTPUTS IN RANDOM f-ARY RECURSIVE CIRCUITS

MEHRI JAVANIAN[1], MOHAMMAD Q. VAHIDI-ASL[2]

ABSTRACT. This paper extends the study of outputs for random recursive binary circuits in Tsukiji and Mahmoud (Algorithmica 31(2001), 403). We show via martingales that a suitably normalized version of the number of outputs in random f-ary recursive circuits converges in distribution to a normal random variate.

Key words: analysis of algorithm, recursive circuit, martingale central limit theorem.
AMS SUBJECT CLASSIFICATION 2010: Primary 05C05.

1. Introduction

The underlying graph of any circuit is a directed acyclic graph. We call the nodes of indegree 0 in a directed acyclic graph *inputs*, and we call the nodes of outdegree 0 *outputs*. We study circuits with fixed indegree and unbounded outdegree.

Consider a circuit of size n where nodes are labeled with $\{1, \ldots, n\}$ in such a way that labels increase along every input-output path. In this case the circuit is called *recursive circuit*. Various models in electro and neural sciences adopt recursive circuits as their graph-theoretic backbone. Some examples are boolean, algebraic and VLSI circuits in complexity theory [1], and neural computing networks in artificial intelligence [7].

For any integer $f \geq 1$, an f-ary recursive circuit is a recursive circuit where the indegree of all non-input nodes is f.

The growth of an f-ary recursive circuit is as follows. The circuit starts out with $a \geq f$ isolated inputs, labeled $1, \ldots, a$, and evolves in stages. After $n-1$ stages, a circuit RC_{n-1} has grown. At the nth stage, f *distinct* nodes

[1]Department of Statistics, Zanjan University, Zanjan, Iran. Email: javanian_m@yahoo.com
[2]Department of Statistics, Shahid Beheshti University, Tehran, Iran. Email: m-vahidi@sbu.ac.ir.

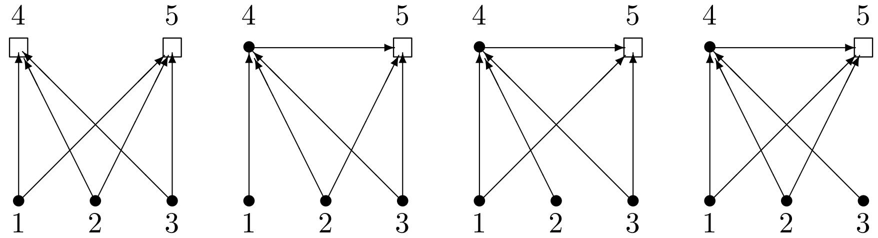

FIGURE 1. All ternary circuits of size 5 grown from three inputs.

are chosen from RC_{n-1} as parents for a new entrant labeled $n+a$. The new node is joined to the circuit with edges directed from the f parents to it, and is given 0 outdegree forming the circuit RC_n. A recursive forest corresponds to the case $f = 1$ (see Balińska et al. [2]). The building block of a recursive forest is the recursive tree, which grows out of a single node. The recursive trees has been a popular topic in both probability and computer science (see Smythe and Mahmoud [5] and the many references therein).

Figure 1 below shows all possible ternary circuits after two insertion steps into an initial graph of three isolated nodes (outputs are illustrated as boxes and non-outputs as bullets).

We impose a probability distribution induced by growth process that chooses f distinct parents *uniformly at random* from all existing nodes. It can be easily argued that the growth after n insertions according to this stochastic view is equivalent to a sample space of all recursive circuits of size $n + a$, where each circuit is equally likely. So a *random* f-ary recursive circuit having a input nodes and n non-input nodes, is one chosen with equal probability from the space of all such circuits.

At the nth stage of the growth of a random f-ary recursive circuit having a input nodes, let $L_{n,a}$ be the number of outputs. The number of inputs a will be held fixed throughout. So we can drop it from all notation and think of it as implicit. For example, we write L_n for $L_{n,a}$, and so on.

Throughout, we shall use the following notation. We shall denote the normally distributed random variate with mean 0 and variance σ^2 by $\mathcal{N}(0, \sigma^2)$. We shall use the symbols $\xrightarrow{\mathcal{D}}$ and $\xrightarrow{P}$ for convergence in distribution and in probability, respectively. The notation $f(n) \sim g(n)$ will be used occasionally to denote that

$$\lim_{n\longrightarrow\infty} \frac{f(n)}{g(n)} = 1.$$

The notation $O_{\mathcal{L}_1}(g(n))$ will stand for a random variable that is $O(g(n))$ in $\mathcal{L}_1$ norm.

In electrical engineering the number of outputs may have many implications concerning the amount of output currents down and in boolean circuits they stand for how many "answers" are derived from a given inputs. Via martingale difference formulation, Tsukiji and Mahmoud [6] find a central limit tendency for the number of outputs in random binary circuit. Mahmoud and Tsukiji [4] studied the joint probability distribution of the number of nodes of outdegree k in random binary circuit.

The current paper considers the study of the number of outputs of random f-ary recursive circuits, for arbitrary $f \geq 2$. The main result of this investigation is to prove the central limit tendency:

$$\frac{L_n - \frac{1}{f+1}n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0\,,\ \frac{f^2}{(2f+1)(f+1)^2}\Big).$$

This central limit theorem is derived in Section 2 via a matingale difference theorem.

2. Central Limit Law

There are $F = \min\{f, L_{n-1}\}$ ways for choosing f parents for the nth insertion from the nodes of RC_{n-1} which may be numbered from 1 to F. For the kth way, the k parents of the nth insertion are outputs in RC_{n-1}, and the other $f-k$ parents for n are not. Let $I_{k,n}$ be the indicator of this event (the kth way). In this case, k outputs of RC_{n-1} are turned into non-output nodes, and a new output appears in RC_n, a net gain of $-k+1$. The change in the number of outputs can be written conditionally as

$$L_n = L_{n-1} - \sum_{k=0}^{\min\{f,L_{n-1}\}} k\, I_{k,n} + 1. \tag{1}$$

If we let $\mathcal{F}_n$ be the sigma filed generated by the first n steps, we have the conditional expectation

$$\mathbf{E}[L_n \,|\, \mathcal{F}_{n-1}] = L_{n-1} - \sum_{k=0}^{\min\{f,L_{n-1}\}} k\, \mathbf{E}[I_{k,n} \,|\, \mathcal{F}_{n-1}] + 1. \tag{2}$$

According to the definition of the indicators, we have

$$\mathbf{E}[I_{k,n} \,|\, \mathcal{F}_{n-1}] = \frac{\binom{L_{n-1}}{k}\binom{n+a-1-L_{n-1}}{f-k}}{\binom{n+a-1}{f}},\ \ 0 \leq k \leq \min\{f, L_{n-1}\}. \tag{3}$$

Plugging (3) into the conditional equation (2),

$$\mathbf{E}[L_n \,|\, \mathcal{F}_{n-1}] = \frac{n+a-f-1}{n+a-1}\, L_{n-1} + 1. \tag{4}$$

The recurrence (4) can be "martingalized:" Appropriate factors b_n and c_n can be chosen so that $b_nL_n + c_n$ is a martingale. We develop this useful lead next.

We let $M_n = b_n L_n + c_n$ and seek b_n and c_n so as to satisfy

$$\mathbf{E}[M_n \,|\, \mathcal{F}_{n-1}] = M_{n-1}.$$

Lemma 1. *The random variable*

$$M_n = \frac{(a-f)!}{a!}(n+a-1)(n+a-2)\cdots(n+a-f)\,L_n - (a-f)$$
$$-\frac{(a-f)!}{a!}\sum_{k=1}^{n}(k+a-1)(k+a-2)\cdots(k+a-f),$$

is a martingale with respect to the sigma fields $\mathcal{F}_n$.

Proof. Let $M_n = b_n L_n + c_n$, for yet-to-be-determined constants b_n and c_n, that render M_n a martingale sequence, with respect to the sigma fields $\mathcal{F}_n$. These constants must then satisfy

$$b_n \mathbf{E}[L_n \,|\, \mathcal{F}_{n-1}] + c_n = \mathbf{E}[M_n \,|\, \mathcal{F}_{n-1}] = M_{n-1} = b_{n-1}L_{n-1} + c_{n-1}.$$

Using the recurrence (4), we obtain

$$b_n\left(\frac{n+a-f-1}{n+a-1}\,L_{n-1} + 1\right) + c_n = b_{n-1}L_{n-1} + c_{n-1}, \tag{5}$$

for every $n \geq 1$. This is possible, if

$$b_n = \frac{n+a-1}{n+a-f-1}\,b_{n-1},$$

which unwinds in

$$b_n = \frac{(a-f)!\,b_1}{a!}(n+a-1)(n+a-2)\cdots(n+a-f),$$

for arbitrary constant b_1. Equating the free terms in (5), we get

$$c_n = c_{n-1} - b_n,$$

yielding

$$c_n = c_0 - \sum_{k=1}^{n} b_k = c_0 - \frac{(a-f)!\,b_1}{a!}\sum_{k=1}^{n}(k+a-1)(k+a-2)\cdots(k+a-f).$$

We also want $\mathbf{E}[M_1] = 0$, requiring that $c_0 = (f-a)b_1$, ($\mathbf{E}[L_1] = a-f+1$). Hence, for arbitrary b_1,

$$M_n = \frac{(a-f)!\,b_1}{a!}(n+a-1)(n+a-2)\cdots(n+a-f)\,L_n - b_1(a-f)$$
$$-\frac{(a-f)!\,b_1}{a!}\sum_{k=1}^{n}(k+a-1)(k+a-2)\cdots(k+a-f)$$

is a martingale. □

The main result of this section is presented next.

Theorem 2. *Let L_n be the number of outputs in a random f-ary recursive circuit after the insertion of n nodes. The outputs follow the central limit law:*

$$\frac{L_n - \frac{1}{f+1} n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0\,,\ \frac{f^2}{(f+1)^2} \cdot \frac{1}{2f+1}\Big).$$

Proof. The expected number of outputs after n random insertions is immediate from Lemma 1, as the associated martingale has 0 mean. That is,

$$\mathbf{E}[L_n] = \frac{1}{(n+a-1)(n+a-2)\cdots(n+a-f)} \Bigg[\frac{a!}{(a-f-1)!} + \sum_{k=1}^{n} (k+a-1)(k+a-2)\cdots(k+a-f) \Bigg]. \tag{6}$$

On the other hand

$$\begin{aligned}
&\sum_{k=f+1}^{n} (k+a-1)(k+a-2)\cdots(k+a-f) \\
&\quad = f! \sum_{k=f+1}^{n} \binom{k+a-1}{f} \\
&\quad = f! \sum_{k=f+1}^{n} \left[\binom{k+a}{f+1} - \binom{k+a-1}{f+1} \right] \\
&\quad = f! \binom{n+a}{f+1} - f! \binom{f+a}{f+1} \\
&\quad = \frac{(n+a)(n+a-1)\cdots(n+a-f)}{f+1} - f! \binom{f+a}{f+1}.
\end{aligned}$$

Therefore for large values of n,

$$\begin{aligned}
&\sum_{k=1}^{n} (k+a-1)(k+a-2)\cdots(k+a-f) \\
&\qquad \sim \sum_{k=f+1}^{n} (k+a-1)(k+a-2)\cdots(k+a-f) \\
&\qquad \sim \frac{(n+a)(n+a-1)\cdots(n+a-f)}{f+1}.
\end{aligned} \tag{7}$$

From which we conclude that

$$\sum_{k=1}^{n} k^f \sim \sum_{k=1}^{n} (k+a-1)(k+a-2)\cdots(k+a-f) \sim \frac{n^{f+1}}{f+1}. \tag{8}$$

It is immediate from (6) and (7), that

$$\mathbf{E}[L_n] \sim \frac{1}{f+1}\, n. \tag{9}$$

Now we obtain the growth rate of the variance. Starting with squaring both sides of the recurrence in (1), one derives

$$\begin{aligned}
\mathbf{E}[L_n^2 \,|\, \mathcal{F}_{n-1}] &= L_{n-1}^2 + 1 + 2L_{n-1} \\
&\quad + \sum_{k=0}^{\min\{f, L_{n-1}\}} (k^2 - 2kL_{n-1} - 2k)\, \mathbf{E}[I_{k,n} \,|\, \mathcal{F}_{n-1}] \\
&= L_{n-1}^2 + 1 + 2L_{n-1} \\
&\quad + \frac{fL_{n-1}}{n+a-1} \left(\frac{n+a-1-L_{n-1}-f+fL_{n-1}}{n+a-2} \right) \\
&\quad - 2L_{n-1} \frac{fL_{n-1}}{n+a-1} - 2\, \frac{fL_{n-1}}{n+a-1},
\end{aligned} \tag{10}$$

where the relations $I_{k,n}^2 = I_{k,n}$, for all k, and $I_{k,n} I_{j,n} = 0$, for all $k \neq j$ are applied. Take expectations and simplify to get

$$\begin{aligned}
\mathbf{E}[L_n^2] &= \frac{(n+a-f-1)(n+a-f-2)}{(n+a-1)(n+a-2)}\, \mathbf{E}[L_{n-1}^2] \\
&\quad + \frac{(n+a-f-1)\big(2(n+a-2)+f\big)}{(n+a-1)(n+a-2)}\, \mathbf{E}[L_{n-1}] + 1.
\end{aligned} \tag{11}$$

Taking expectations, then squaring of the recurrence (4), we get

$$\mathbf{E}^2[L_n] = \left(\frac{n+a-f-1}{n+a-1} \right)^2 \mathbf{E}^2[L_{n-1}] + 2\left(\frac{n+a-f-1}{n+a-1} \right) \mathbf{E}[L_{n-1}] + 1. \tag{12}$$

Subtracting (12) from (11), we obtain

$$\begin{aligned}
\mathbf{Var}[L_n] &= \frac{(n+a-f-1)(n+a-f-2)}{(n+a-1)(n+a-2)}\, \mathbf{Var}[L_{n-1}] \\
&\quad + \frac{f(n+a-f-1)}{n+a-2} \left(\frac{\mathbf{E}[L_{n-1}]}{n+a-1} \right) \left(1 - \frac{\mathbf{E}[L_{n-1}]}{n+a-1} \right).
\end{aligned} \tag{13}$$

From (13), one can recursively conclude

$$\begin{aligned}
\mathbf{Var}[L_n] &= \sum_{k=1}^{n-1} \frac{f(k+a-f)}{k+a-1} \left(\frac{\mathbf{E}[L_k]}{k+a} \right) \left(1 - \frac{\mathbf{E}[L_k]}{k+a} \right) \\
&\quad \times \prod_{j=k+1}^{n-1} \frac{(j+a-f)(j+a-f-1)}{(j+a)(j+a-1)},
\end{aligned}$$

where the product is interpreted as 1 when the range of j is empty.

However, $L_k \le k+a$, and by (9) there exist a positive constant C such that for all k

$$1 - \frac{\mathbf{E}[L_k]}{k+a} - \frac{f}{f+1} \le \left|1 - \frac{\mathbf{E}[L_k]}{k+a} - \frac{f}{f+1}\right| \le C;$$

therefore

$$\begin{aligned}\mathbf{Var}[L_n] &\le \sum_{k=1}^{n-1} \frac{f(k+a-f)}{k+a-1}\Big(C + \frac{f}{f+1}\Big) \prod_{j=k+1}^{n-1} \frac{(j+a-f)(j+a-f-1)}{(j+a)(j+a-1)} \\ &\le f(C+1)\sum_{k=1}^{n-1} \frac{k+a-f}{k+a-1}.\end{aligned}$$

It follows that

$$\mathbf{Var}[L_n] = O(n). \tag{14}$$

The growth rates of the mean and the variance give a concentration law. It is immediate from (9) and (14), by Chebyshev's inequality, that

$$\frac{L_n}{n} \xrightarrow{P} \frac{1}{f+1}.$$

Further, let $\bigtriangledown M_k = M_k - M_{k-1}$. For any constant factors A_n, $A_n \bigtriangledown M_k$ is a martingale difference sequence, with respect to the sigma fields $\mathcal{F}_k$. The factor $A_n = n^{-(2f+1)/2}$ suits our purpose. We verify martingale central limit theorem for the martingale difference $n^{-(2f+1)/2} \bigtriangledown M_k$. It suffices to check the conditional Lindeberg condition and the conditional variance condition on the martingale differences (see [3]). The conditional Lindeberg condition requires that, for all $\varepsilon > 0$,

$$U_n \overset{\text{def}}{=} \sum_{k=1}^{n} \mathbf{E}\left[\left(\frac{\bigtriangledown M_k}{n^{(2f+1)/2}}\right)^2 \mathbf{1}_{\{|n^{-(2f+1)/2}\bigtriangledown M_k| > \varepsilon\}} \,\middle|\, \mathcal{F}_{n-1}\right] \xrightarrow{P} 0.$$

We have

$$\begin{aligned}\bigtriangledown M_k = \frac{(a-f)!}{a!}&(k+a-2)(k+a-3)\cdots(k+a-f) \\ &\times\big[(k+a-1)(L_k - L_{k-1} - 1) + fL_{k-1}\big].\end{aligned}$$

However, $L_{k-1} - f + 1 \le L_k \le L_{k-1} + 1$, and $L_k \le k+a$; therefore

$$|\bigtriangledown M_k| \le \frac{2f(a-f)!}{a!}(k+a-1)(k+a-2)\cdots(k+a-f).$$

The square differences $(\bigtriangledown M_k)^2$ are therefore $O(k^{2f})$, and it follows that the sets $\{|\bigtriangledown M_k|^2 > \varepsilon^2 n^{2f+1}\}$ are empty for large n, and every $k \le n$. Deterministically, $U_n \longrightarrow 0$; the conditional Lindeberg condition has been verified.

A Z- conditional variance condition requires that

$$V_n \stackrel{\text{def}}{=} \sum_{k=1}^{n} \mathbf{E}\left[\left(\frac{\triangledown M_k}{n^{(2f+1)/2}}\right)^2 \,\middle|\, \mathcal{F}_{n-1}\right] \xrightarrow{P} Z,$$

for the random variable Z. In our case, it will turn out that $Z = \left(\frac{(a-f)!}{a!} \cdot \frac{f}{f+1}\right)^2 \cdot \frac{1}{2f+1}$. According to the martingale formulation of Lemma 1, we have

$$\begin{aligned}
&\mathbf{E}[(\triangledown M_k)^2 \,|\, \mathcal{F}_{n-1}] \\
&= \left[\frac{(a-f)!}{a!}(k+a-2)(k+a-3)\cdots(k+a-f)\right]^2 \mathbf{E}[(k+a-1)^2 L_k^2 \\
&+(k+a-f-1)^2 L_{k-1}^2 + (k+a-1)^2 - 2(k+a-1)(k+a-f-1)L_k L_{k-1} \\
&-2(k+a-1)^2 L_k + 2(k+a-1)(k+a-f-1)L_{k-1} \,|\, \mathcal{F}_{n-1}].
\end{aligned}$$

Substituting (4) and (10), one derives

$$\begin{aligned}
\mathbf{E}[(\triangledown M_k)^2 \,|\, \mathcal{F}_{n-1}] &= \left[\frac{(a-f)!}{a!}(k+a-2)(k+a-3)\cdots(k+a-f)\right]^2 \\
&\times \frac{k+a-f-1}{k+a-2}\left[\left(f + O\left(\frac{1}{k}\right)\right) k L_{k-1} - f L_{k-1}^2\right]. \qquad (15)
\end{aligned}$$

From the asymptotics of the variance we have

$$\mathbf{E}\left[\left(L_k - \frac{k}{f+1}\right)^2\right] = \mathbf{Var}[L_k] + \left(\mathbf{E}[L_k] - \frac{k}{f+1}\right)^2 = O(k), \qquad (16)$$

so, from the Cauchy-Schwartz inequality one derives

$$\mathbf{E}\left[\left|L_k - \frac{k}{f+1}\right| \cdot 1\right] \le \sqrt{E\left[\left(L_k - \frac{k}{f+1}\right)^2\right]} = O(\sqrt{k}).$$

The latter inequality gives

$$L_k = \frac{k}{f+1} + O_{\mathcal{L}_1}(\sqrt{k}). \qquad (17)$$

Further, by the Cauchy-Schwartz inequality we have

$$\begin{aligned}
\mathbf{E}\left[\left|L_k^2 - \frac{k^2}{(f+1)^2}\right|\right] &= \mathbf{E}\left[\left|L_k + \frac{k}{f+1}\right|\left|L_k - \frac{k}{f+1}\right|\right] \\
&\le \sqrt{\mathbf{E}\left[\left(L_k + \frac{k}{f+1}\right)^2\right]\mathbf{E}\left[\left(L_k - \frac{k}{f+1}\right)^2\right]}.
\end{aligned}$$

We can use (16) to bound $\sqrt{\mathbf{E}\Big[\Big(L_k - \frac{k}{f+1}\Big)^2\Big]}$ by $O(\sqrt{k})$, and use the obvious $O(k)$ bound $L_k + k/(f+1)$, to obtain

$$\mathbf{E}\bigg[\bigg|L_k^2 - \frac{k^2}{(f+1)^2}\bigg|\bigg] = O(k^{3/2}).$$

We thus have

$$L_k^2 = \frac{k^2}{(f+1)^2} + O_{\mathcal{L}_1}(k^{3/2}). \tag{18}$$

In view of (17) and (18), (15) can be rewritten as

$$\begin{aligned}
&\mathbf{E}[(\triangledown M_k)^2 \,|\, \mathcal{F}_{n-1}] \\
&= \bigg[\frac{(a-f)!}{a!}(k+a-3)(k+a-4)\cdots(k+a-f)\bigg]^2 (k+a-2)(k+a-f-1) \\
&\qquad \times \bigg[fk\Big(\frac{k-1}{f+1} + O_{\mathcal{L}_1}(k^{1/2})\Big) + O(k) - f\Big(\Big(\frac{k-1}{f+1}\Big)^2 + O_{\mathcal{L}_1}(k^{3/2})\Big)\bigg] \\
&= \bigg[\frac{(a-f)!}{a!}(k+a-3)(k+a-4)\cdots(k+a-f)\bigg]^2 (k+a-2)(k+a-f-1) \\
&\qquad \times \bigg[\frac{f^2}{(f+1)^2}k^2 + O_{\mathcal{L}_1}(k^{3/2})\bigg] \\
&= \Big(\frac{(a-f)!}{a!} \cdot \frac{f}{f+1}\Big)^2 k^{2f} + O(k^{2f-1}) + O_{\mathcal{L}_1}(k^{\frac{4f-1}{2}}) \\
&= \Big(\frac{(a-f)!}{a!} \cdot \frac{f}{f+1}\Big)^2 k^{2f} + O_{\mathcal{L}_1}(k^{\frac{4f-1}{2}}).
\end{aligned}$$

When this is summed from 1 to n and normed by $n^{-(2f+1)}$, by (8), the conditional variance V_n approaches $\Big(\frac{(a-f)!}{a!}\cdot\frac{f}{f+1}\Big)^2\cdot\frac{1}{2f+1}$ in $\mathcal{L}_1$, and the Z conditional variance has been verified.

Finally, by martingale central limit theorem,

$$\sum_{k=1}^{n} \frac{\triangledown M_k}{n^{(2f+1)/2}} = \frac{M_n}{n^{(2f+1)/2}} \xrightarrow{\mathcal{D}} \mathcal{N}\bigg(0\,,\ \Big(\frac{(a-f)!}{a!} \cdot \frac{f}{f+1}\Big)^2 \cdot \frac{1}{2f+1}\bigg).$$

□

References

[1] J. Balcázar, J. Diaz, J. Gabarró: Structural complexity I, Springer-Verlag, Berlin, 1995.
[2] K. Balinska, L. Quintas, J. Szymanski: Random recursive forests, Random Structures Algorithms 5(1994), 3–12.
[3] P. Hall, C. Heyde: Martingale Limit Theory and Its Applications, Academic Press, New York, 1980.
[4] H. Mahmoud, T. Tsukiji: On the internal structure of random recursive circuits, Journal of Computatinoal and Applied Mathematics 142(2002), 155–171.
[5] R. Smythe, H. Mahmoud: A survey of recursive trees, Theory of Probability and Mathematical Statistics 51(1996), 1–29.
[6] T. Tsukiji, H. Mahmoud: A limit law for outputs in random circuits, Algorithmica 31(2001), 403–412.
[7] L. Valiant: Circuits of the Mind, Oxford University Press, Oxford, 1994.