The Boolean Functions Computed by Random Boolean Formulas
OR
How to Grow the Right Function

Alex Brodsky Nicholas Pippenger
Department of Computer Science,
University of British Columbia,
201-2366 Main Mall, Vancouver,
BC, Canada, V6R 1J9
{abrodsky,nicholas}@cs.ubc.ca

Abstract

Among their many uses, growth processes (probabilistic amplification), were used for constructing reliable networks from unreliable components, and deriving complexity bounds of various classes of functions. Hence, determining the initial conditions for such processes is an important and challenging problem. In this paper we characterize growth processes by their initial conditions and derive conditions under which results such as Valiant's[Val84] hold. First, we completely characterize growth processes that use linear connectives. Second, by extending Savický's [Sav90] analysis, via "Restriction Lemmas", we characterize growth processes that use monotone connectives, and show that our technique is applicable to growth processes that use other connectives as well. Additionally, we obtain explicit bounds on the convergence rates of several growth processes, including the growth process studied by Savický (1990).
Keywords: Computational and structural complexity, growth processes, probabilistic amplification

1 Introduction

The notion of a random Boolean function occurs many times, both implicitly and explicitly, in the literature of theoretical computer science. Not long after Shannon [Sha38] pointed out the relevance of Boolean algebra to the design of switching circuits, Riordan and Shannon [RS42] obtained a lower bound to the complexity (the size of series-parallel relay circuits, or of formulas with the connectives "and", "or" and "not") of "almost all" Boolean functions, and this bound can naturally be applied to a "random" Boolean function when all 2^{2^n} Boolean functions of n arguments are assumed to occur with equal probability. Lupanov [Lup61] later showed that Riordan and Shannon's lower bound is matched asymptotically by an upper bound that applies to all Boolean functions, so in this situation the average case is asymptotically equivalent to the worst case. This asymptotic equivalence of average and worst cases also holds in many other situations involving circuits or formulas. There are some complexity measures, however, such as the length of the shortest disjunctive-normal-form formula, for which the average case behaves quite differently from the worst case (see Glagolev [Gla67], for example), and the complexity of a random Boolean function remains a challenging open problem. In these cases, probability distributions other than the uniform distribution have also been considered; for example, one may assume that each entry in the truth-table is independently 1 with probability p and 0 with probability $1-p$, so that the uniform distribution is the special case $p = 1/2$ (see Andreev [And84]).

Another approach to the study of random Boolean functions is to put a probability distribution on formulas, and let that induce a probability distribution on the functions that they compute. This may be done by using a "growth process" (defined below) to grow random formulas.

Valiant [Val84] considered such a growth process, and showed that the resulting probability distribution tends to the distribution concentrated on a single function: the threshold function that assumes the value 1 if and only at least $n/2$ of its n arguments assume the value 1. This result was used to obtain a non-constructive upper bound on the minimum possible size of a formula for computing this threshold function. This argument in fact gives the best upper bound currently known for this and similar threshold functions.

The choice of the initial probability distribution on formulas dictates the probability distribution on functions. To facilitate the design and use of growth process, as in the case above, deriving a characterization based on the initial conditions is an important problem.

One such result in this framework is due to Savický [Sav90]. Savický formulated broad conditions under which the distribution of the random function computed by a formula of depth i tends to the uniform distribution on all Boolean functions of n variables as $i \to \infty$.

Savický [Sav95a] has also shown that in some cases the rate of approach of the probability of computing a particular function f to the uniform probability 2^{-2^n} gives information about f: it is fastest for the linear functions, and slowest for the "bent" functions (which are furthest, in Hamming distance, from the linear functions). For some other models of random formulas, Lefmann and Savický [LS97] and Savický [Sav98] have shown that the logarithm of the probability of computing a particular function is related to the complexity of that function (as measured by the size of the smallest formula computing that function). Finally, we should mention that Razborov [Raz88] has used random formulas in yet another model to show that some large graphs with Ramsey properties have representations by formulas of exponentially smaller size. This result, which has been improved quantitatively by Savický [Sav95b], shows that Ramsey properties are possessed by graphs that are far from random.

Our goal in this paper is to determine under what circumstances results like Valiant's and Savický's hold. We show that for many growth processes, the probability distribution on the computed function tends to the uniform distribution on some set of functions (which may range in size from a single function, as in Valiant's result, to all functions, as in Savický's).

2 Definitions

Let $\mathcal{F}_n$ denote the family of n-adic Boolean functions, let $\mathcal{M}_n$ denote the family of n-adic monotone Boolean functions, and let $\mathcal{L}_n$ denote the family of n-adic linear functions. The set B_n denotes Boolean cube of size n.

Let k be a positive integer and α be a k-adic Boolean function, which we call the **connective**. Let $A_0 = \{x_1, x_2, ..., x_n, \bar{x}_1, ..., \bar{x}_n, 0, 1\}$ be the set comprising the projection functions, their negations, and the constant functions, and let $A_i = \{\alpha(v_1, v_2, ..., v_k) \mid v_i \in A_{i-1}\}$ be the set comprising the formulas composed from A_{i-1}. A **growth process** is denoted by a pair (μ, α), where μ is a distribution on A_0 and α is a connective; μ is called the **initial distribution**. A growth process gives rise to a probability distribution π_i on A_i for each $i \geq 0$ in the following way. We take $\pi_0 = \mu$. For $i \geq 1$, we take $\pi_i(f)$ to be the probability that $\alpha(g_1, \ldots, g_k) = f$, where $g_1, \ldots, g_k$ are independent random functions distributed according to π_{i-1} on A_{i-1}.

We shall assume that μ is a uniform distribution on a subset of A_0. This subset will always contain the n projections; it may or may not contain their n negations; and it may contain neither, one, or both of the two constants. All of our results could be extended to more general distributions μ, but these assumptions allow us to present the most interesting results with a minimum of notation. They also cover the results of Valiant [Val84] and Savický [Sav90]. (Valiant's proof actually uses a non-uniform distribution, but the same bound can be obtained by a simple modification using a uniform distribution on the projections.)

The **support** of a probability distribution π, denoted $supp(\pi)$, is the set $\{f \mid \pi(f) > 0\}$. The **support** of a growth process is the set of all functions $f \in \mathcal{F}_n$ for which $\pi_i(f) > 0$ for some $i > 0$: $\cup_i supp(\pi_i)$.

We are particularly interested in cases in which π_i tends to a **limiting distribution** π as $i \to \infty$. (There are also cases in which π_{2i} and π_{2i+1} tend to distinct **alternating limiting distributions**.) When a limiting distribution exists, we can have $\pi(f) > 0$ only for f in the support of the growth process. As Valiant's result indicates, however, there may be functions in the support for which $\pi_i(f) \to 0$, so that $\pi(f) = 0$. The **asymptotic support** of a growth process with a limiting distribution π is the set of functions $f \in \mathcal{F}_n$ for which $\pi(f) > 0$.

Additionally, we investigate how quickly the distribution π_i approaches the limiting distribution as i approaches infinity. Namely, for some $\varepsilon > 0$, the size of i such that $\max_f |\pi(f) - \pi_i(f)| < \varepsilon$. Almost all growth processes that we study share the important characteristic: for any $\varepsilon > 0$, $\max_f |\pi(f) - \pi_{O(\log(n))}(f)| < \varepsilon$. Note, unless otherwise stated, the base of the logarithm is assumed to be 2.

Growth processes in which the limiting distribution is concentrated on one function are used extensively in probabilistic amplification methods and can be analyzed by studying the properties of the corresponding "characteristic polynomial". Let $\{X_1, X_2, ..., X_n\}$ be a set of random independent binary variables that are 1 with probability p and let each X_i represent the input x_i. The **characteristic polynomial** of f is defined by $A_f(p) = \Pr[f(X_1, X_2, ...X_n) = 1]$ and is given by

$$A_f(p) = \sum_{i=0}^{n} \beta_i \binom{n}{i} p^i (1-p)^{n-i}$$

where β_i is the fraction of assignments of weight i for which f is true. The characteristic polynomial was used by von Neumann [vN56] and by Moore and Shannon [MS56] to study reliable computation with unreliable components, as well as by Valiant [Val84] (see also Boppana [Bop85, Bop89, DZ97]).

To analyze growth processes whose limiting distribution is uniform over a set of functions, we use a Fourier transform technique. The Fourier transform Δ_i of a probability distribution π_i is defined by

$$\Delta_i(f) = \sum_{g \in \mathcal{F}_n} (-1)^{\langle f,g \rangle} \pi_i(g) \tag{1}$$

where $\pi_i(g)$ is the probability of selecting g from A_i. For convenience, the inner product $\langle f,g \rangle = \sum_i f_i g_i$ is defined to be over the integers, rather than over $\mathbb{Z}_2$. Unless otherwise noted, Boolean n-adic functions are represented as Boolean vectors from B_{2^n}. The inverse Fourier transform is defined by

$$\pi_i(g) = \frac{1}{2^{2^n}} \sum_{f \in \mathcal{F}_n} (-1)^{\langle f,g \rangle} \Delta_i(f). \tag{2}$$

The Fourier transform was used by Razborov [Raz88] to derive his results on Ramsey graphs, as well as by Savický [Sav90].

The Fourier transform plays a role in many of our results, but it needs to be adapted in various ways to suit different cases. When dealing with linear functions, for example, we will have to represent the functions f and g in definition 1 not as Savický does, by their truth-tables, but rather by their coefficients as multivariate polynomials over $GF(2)$. In other cases, when establishing a limiting distribution that is uniform over a proper subset of $\mathcal{F}_n$, we shall need to use what we call "restriction lemmas", which assert relationships that hold among the values of the Fourier transform.

3 Growing Linear Functions

A function f is linear if it is of the form $f(x_1, \ldots, x_n) = c_0 \oplus c_1 x_1 \oplus \cdots \oplus c_n x_n$ for some constants $c_0, c_1, \ldots, c_n \in GF(2)$. We may assume without loss of generality that α depends on all its arguments, so that $\alpha(y_1, \ldots, y_k) =$

$c \oplus y_1 \oplus \cdots \oplus y_k$, where $k \geq 2$. The result of the growth process depends on the support of of the initial distribution μ, the parity of k, and the constant term c.

To prove this we derive a recurrence for the Fourier coefficients of the respective probability distribution π_i, from which we derive the limiting distribution. Since compositions of linear functions are themselves linear, we represent the linear functions by their vector $(c_0, c_1, \ldots, c_n)$ of coefficients, and the following summations range over $\mathcal{L}_n$. Finally, let w_1 denote the constant function 1 ($w_1 = 100\ldots0$), whereas $\mathbf{1} = 11\ldots1$.

Proposition 3.1 *Let α be a linear connective as described above and let $w \in \mathcal{L}_n$. The Fourier coefficients of the probability distribution π_i of the corresponding growth process are described by the recurrence relation*

$$\Delta_{i+1}(w) = (-1)^{c\langle w_1, w\rangle} \Delta_i(w)^k.$$

Proof:

$$\begin{aligned}
\Delta_{i+1}(w) &= \sum_{f \in \mathcal{L}_n} \pi_{i+1}(f)(-1)^{\langle f,w\rangle} = \sum_{f \in \mathcal{L}_n} \sum_{\substack{\mathbf{g} \in \mathcal{L}_n^k \\ \alpha(\mathbf{g}) = f}} \prod_{j=1}^{k} \pi_i(\mathbf{g}_j)(-1)^{\langle f,w\rangle} \\
&= \sum_{\mathbf{g} \in \mathcal{L}_n^k} \prod_{j=1}^{k} \pi_i(\mathbf{g}_j)(-1)^{\langle \alpha(\mathbf{g}),w\rangle} = \sum_{\mathbf{g} \in \mathcal{L}_n^k} \prod_{j=1}^{k} \pi_i(\mathbf{g}_j)(-1)^{\langle cw_1 \oplus \bigoplus_{j=1}^{k} \mathbf{g}_j, w\rangle} \\
&= \sum_{\mathbf{g} \in \mathcal{L}_n^k} \prod_{j=1}^{k} \pi_i(\mathbf{g}_j)(-1)^{\langle cw_1,w\rangle \oplus \bigoplus_{j=1}^{k} \langle \mathbf{g}_j, w\rangle} = (-1)^{\langle cw_1,w\rangle} \sum_{\mathbf{g} \in \mathcal{L}_n^k} \prod_{j=1}^{k} \pi_i(\mathbf{g}_j)(-1)^{\langle \mathbf{g}_j,w\rangle} \\
&= (-1)^{\langle cw_1,w\rangle} \Delta_i(w)^k
\end{aligned}$$

■

Using proposition 3.1, the following theorems classify the growth processes on linear connectives.

Theorem 3.2 *Let $\alpha(y) = c \oplus y_1 \oplus \cdots \oplus y_k$, $k > 1$, be a linear k-adic connective, as defined above, and assume that the support of μ does not contain negations of the projections.*

1. *If $\{0,1\} \cap supp(\mu) \neq \{0,1\}$, k is odd and $c = 1$, then the growth process has alternating limiting distributions, each of which is uniform over one half of the support of the growth process (which consists of all linear functions for which $\bigoplus_{j=1}^{n} c_j = 1$).*
2. *In all other cases, the limiting distribution is uniform over the support of the growth process (which depends on k, c, and the presence of constants in the support).*

Proof: Two facts are key to this theorem: first, that $|\Delta_i(w)| \leq 1$, and second, that if $|\Delta_i(w)| < 1$, then $\lim_{i\to\infty} \Delta_i(w) = 0$. Only the nonzero (magnitude 1) coefficients contribute to limiting distribution (equation 2); fortunately, these are determined solely by the support of the initial distribution. Depending on which constants are part of the support, there are either one, two, or four magnitude 1 coefficients:

$$\begin{aligned}
\{0,1\} \cap supp(\mu) = \{0,1\} &\Rightarrow \Delta_0(0) = 1, \\
\{0,1\} \cap supp(\mu) = \{0\} &\Rightarrow \Delta_0(0) = \Delta_0(w_1) = 1, \\
\{0,1\} \cap supp(\mu) = \{1\} &\Rightarrow \Delta_0(0) = 1,\ \Delta_0(\mathbf{1}) = -1, \\
\{0,1\} \cap supp(\mu) = \emptyset &\Rightarrow \Delta_0(0) = \Delta_0(w_1) = 1,\ \Delta_0(\mathbf{1}) = \Delta_0(w_1 \oplus \mathbf{1}) = -1.
\end{aligned}$$

If k is odd and $c = 1$, the recurrence from Proposition 3.1 implies that $\Delta_{i+1}(w_1) = -\Delta_i(w_1)$ and $\Delta_{i+1}(\mathbf{1}) = -\Delta_i(\mathbf{1})$. Hence, the limiting distribution is alternating. In the case where one of the constants is missing from the support only two coefficients are magnitude 1, hence the alternating distribution is uniform over half of $\mathcal{L}_n$. In the case where both constants are missing the alternating distribution is uniform over two quarters of $\mathcal{L}_n$, specifically, parity of an odd number of variables and their negations.

If $c = 0$, k is even, or $\{0,1\} \cap supp(\mu) = \{0,1\}$, the limiting distribution exists because the sign of the magnitude 1 coefficients does not alternate. We can read off the limiting distribution from the Fourier coefficients. If both constants are in the support, then the limiting distribution is uniform over $\mathcal{L}_n$. If only one of the constants is present, then the distribution is uniform over half of $\mathcal{L}_n$, and if neither is present, then the distribution will be uniform over a quarter of $\mathcal{L}_n$. ∎

If the support of μ contains negations, then using the same proof technique yields the following theorem.

Theorem 3.3 *Let $\alpha(y) = c \oplus y_1 \oplus \cdots \oplus y_k$ be a linear k-adic connective, as defined above, and assume that the support of μ contains negations of the projections.*

1. *If $\{0,1\} \cap supp(\mu) = \emptyset$ and k is odd then the limiting distribution is uniform over all linear functions of odd number of variables.*
2. *If $\{0,1\} \cap supp(\mu) = \emptyset$ and k is even then the limiting distribution is uniform over all linear functions of even number of variables.*
3. *Otherwise, the limiting distribution is uniform over all of $\mathcal{L}_n$.*

Proof: If $\{0,1\} \cap supp(\mu) \neq \emptyset$, then there is only one coefficient of magnitude 1, $\Delta_0(\mathbf{z}) = 1$, implying the latter case.

Otherwise, there is one other magnitude 1 coefficient, $\Delta_0(\mathbf{1} \oplus w_1) = -1$. If k is odd, then $\Delta_{i+1}(\mathbf{1} \oplus w_1) = \Delta_i(\mathbf{1} \oplus w_1)^k = -1$, implying the first case of the theorem. If k is even, then $\Delta_{i+1}(\mathbf{1} \oplus w_1) = \Delta_i(\mathbf{1} \oplus w_1)^k = 1$, implying the second case. ∎

Note, that if negations are present, no alternating distribution can occur. To bound the convergence of π_i to π we use the inverse Fourier transform.

Theorem 3.4 *Let α be a k-adic linear connective, $k > 1$, of a linear process on n variables that has a limiting distribution π. There exists a constant c_α, such that for all $n > 0$, if $i > \frac{2\log(n)}{\log(k)} + c_\alpha$, then for any linear function f, $|\pi(f) - \pi_i(f)| < 2^{-n}$.*

Proof: Let $D = \{w \ : \ |\Delta_0(w)| < 1\}$, then $\pi_i(f)$ may be written as:

$$\begin{aligned}
\pi_i(f) &= 2^{-n-1} \sum_{w \in B_{n+1}} (-1)^{\langle w,f \rangle} \Delta_i(w) \\
&= 2^{-n-1} \sum_{w \notin D} (-1)^{\langle w,f \rangle} \Delta_i(w) + 2^{-n-1} \sum_{w \in D} (-1)^{\langle w,f \rangle} \Delta_i(w) \\
&= \pi(f) + 2^{-n-1} \sum_{w \in D} (-1)^{\langle w,f \rangle} \Delta_i(w).
\end{aligned}$$

Thus, for any linear function f,

$$|\pi(f) - \pi_i(f)| = |2^{-n-1} \sum_{w \in D} (-1)^{\langle w,f \rangle} \Delta_i(f)| \leq \max_{w \in D} |\Delta_0(w)|^{k^i} \leq (1 - n^{-1})^{k^i}.$$

Solving inequality $(1 - n^{-1})^{k^i} < 2^{-n}$, in terms of i, yields: $i > \frac{2\log(n)}{\log(k)} + \frac{1}{\log(k)}$. ∎

4 Growing Self-Dual Functions

Savický [Sav90] showed that if the connective is balanced (that is, if it assumes the value 1 for just one-half of the combinations of argument values) and non-linear, and the support of μ is all of A_0, then the limiting distribution will be uniform over all of $\mathcal{F}_n$. If we remove the constants from the support of μ and assume the connective α is self-dual (that is, satisfies $\alpha(y_1,\dots y_k) = \overline{\alpha(\bar{y}_1,\dots,\bar{y}_k)}$), then the support of the growth process is the set of all self-dual functions. In this case the limiting distribution of the growth process is uniform over this support.

Theorem 4.1 *If the connective is non-linear and self-dual, and the support of μ comprises the projections and their negations, then the limiting distribution will be uniform over the family of self-dual n-adic functions.*

Proof: Observe that there is a bijection between the set of all functions on n variables and the set of self-dual functions on $n+1$ variables, for example, the map

$$f(x_1,x_2,\dots,x_n) \mapsto f(x_1,x_2,\dots,x_n)x_{n+1} \vee \overline{f(\bar{x}_0,\bar{x}_1,\dots,\bar{x}_n)}\bar{x}_{n+1}.$$

The result follows. ■

5 Growing Monotone Functions

We now focus on growth processes that use monotone connectives. For the rest of this section we assume that α is monotone and the support of μ contains only monotone functions from A_0 (that is, projections and possibly constants). We first investigate unbalanced connectives.

5.1 Using Unbalanced Connectives

Growth processes that use unbalanced monotone connectives concentrate probability on a threshold function; the type of threshold function depends on the connective and the support. A threshold function $T_k(x_1,\dots,x_n)$ assumes the value 1 if and only if at least k of its n arguments assume the value 1. We consider constant functions $T_{n+1} = 0$ and $T_0 = 1$ to be special cases of threshold functions. There are two cases to consider: first, when the characteristic polynomial of α, $A_\alpha(p)$, has no fixed-point on the open interval $(0,1)$, and second, when $A_\alpha(p)$ has a fixed-point on $(0,1)$.

Proposition 5.1 *If α is a monotone connective whose characteristic polynomial, $A(p)$, has no fixed-point on the interval $(0,1)$, then the limiting distribution will be concentrated on a threshold function.*

Proof: Since $A(p)$ has no fixed-point on $(0,1)$, either $A_\alpha(p) < p$ throughout $(0,1)$, or $A_\alpha(p) > p$ throughout $(0,1)$. If $A_\alpha(p) > p$ throughout $(0,1)$, then by the standard amplification argument, the limiting distribution is concentrated on T_1 (disjunction of all variables) or, T_0 if 1 is in the support of μ. Similarly, if $A_\alpha(p) < p$ throughout $(0,1)$, then the limiting distribution is concentrated on T_n (conjunction of all variables) or T_{n+1} if 0 is in the support of μ. ■

Furthermore, all connectives whose characteristic polynomials have no fixed-point on $(0,1)$, are either of the form $\alpha(x) = x_i \vee \alpha'(x)$ (when $A_\alpha(p) > p$) or $\alpha(x) = x_i \wedge \alpha'(x)$ (when $A_\alpha(p) < p$). If $\alpha(x) \neq x_i \vee \alpha'(x)$, then $A_\alpha(p) = O(p^2)$ which implies that there exists a positive constant ε_0 such that for all $0 < \varepsilon < \varepsilon_0$, $A_\alpha(\varepsilon) < \varepsilon$. Similarly, if $\alpha(x) \neq x_i \wedge \alpha'(x)$, then by duality, $1 - A_\alpha(1-p) = O(p^2)$, which means that $A_\alpha(1-\varepsilon) > 1-\varepsilon$ for all $0 < \varepsilon < \varepsilon_1$ for some $\varepsilon_1 > 0$. Since $A_\alpha(p)$ is continuous, there must exist a fixed-point in $(0,1)$, which is a contradiction.

In the second case, where $A_\alpha(p)$ has a fixed-point in $(0,1)$, Moore and Shannon [MS56] have shown that this fixed-point is unique. Not surprisingly, the limiting distribution depends on the fixed-point. Thus, we first derive two facts about the fixed-point of the characteristic polynomial, to deal with the second case.

Lemma 5.2 *The characteristic polynomial $A(p)$ has a fixed-point of $\frac{1}{2}$ if and only if the connective α is balanced.*

Proof: By definition $\sum_{i=0}^{n}\beta_i\binom{n}{i}$ is the number of assignments for which α is true. If $A_\alpha(\frac{1}{2})=\frac{1}{2}$, then $A_\alpha(\frac{1}{2})=\sum_{i=0}^{n}\beta_i\binom{n}{i}(\frac{1}{2})^i(\frac{1}{2})^{n-i}=\frac{1}{2^n}\sum_{i=0}^{n}\beta_i\binom{n}{i}=\frac{1}{2}$. Hence, $\sum_{i=0}^{n}\beta_i\binom{n}{i}=2^{n-1}$ which means that α is balanced. Conversely, if α is balanced, then $A_\alpha(\frac{1}{2})=\frac{1}{2}$. ∎

Lemma 5.3 *If α is a monotone, non-projection connective, then any fixed-point of $A_\alpha(p)$ on $(0,1)$ is either irrational or $\frac{1}{2}$.*

Proof: By contradiction; without loss of generality assume that the fixed-point $p_0=\frac{r}{s}<\frac{1}{2}$ and $\gcd(r,s)=1$. Hence,

$$A_\alpha\left(\frac{r}{s}\right)=\sum_{j=0}^{k}\beta_j\binom{k}{j}\left(\frac{r}{s}\right)^j\left(\frac{s-r}{s}\right)^{k-j}=\frac{r}{s}.$$

Multiplying both sides by s^k, noting that $\beta_k=\beta_{k-1}=1$, and evaluating the result modulo $(s-r)^2$ yields

$$\begin{aligned} rs^{k-1} &\equiv \sum_{j=0}^{k}\beta_j\binom{k}{j}r^j(s-r)^{k-j}\equiv r^k+kr^{k-1}(s-r)+(s-r)^2\sum_{j=0}^{k-2}\beta_j\binom{k}{j}r^j(s-r)^{k-j-2}\\ &\equiv r^k+kr^{k-1}(s-r) \bmod (s-r)^2. \end{aligned}$$

Evaluating the left side modulo $(s-r)^2$ yields

$$\begin{aligned} rs^{k-1} &\equiv r(r+(s-r))^{k-1}\equiv r\sum_{i=0}^{k-1}\binom{k-1}{i}r^i(s-r)^{k-1-i}\\ &\equiv rr^{k-1}+r(k-1)r^{k-2}(s-r)+r(s-r)^2\sum_{j=0}^{k-3}\binom{k-1}{j}r^j(s-r)^{k-3-j}\\ &\equiv r^k+(k-1)r^{k-1}(s-r) \bmod (s-r)^2. \end{aligned}$$

Therefore,

$$r^{k-1}(s-r)\equiv 0 \bmod (s-r)^2.$$

Since $\gcd(r,s)=\gcd(r,(s-r)^2)=1$, $r^{k-1}\not\equiv 0 \bmod (s-r)^2$; this is a contradiction. ∎

Theorem 5.4 *Let α be a monotone unbalanced connective whose characteristic polynomial has a fixed-point $t\in(0,1)$, and let the support of μ contain only the projections. The limiting distribution of the growth process is concentrated on the threshold function $T_{\lceil tn\rceil}$.*

Proof: Since α is unbalanced and has a fixed-point on $(0,1)$, by Lemma 5.2, the fixed-point is not $\frac{1}{2}$. Hence, by Lemma 5.3, the fixed-point is irrational. Since the fraction of variables set to true in any assignment is by definition rational, the fraction will always be strictly greater or strictly less than the fixed-point t. Hence, by the standard amplification argument, the limiting distribution will be concentrated on the threshold function $T_{\lceil tn\rceil}$. ∎

Theorem 5.4 can easily be modified to cover the cases in which one or both constants are in the support of μ. Combining proposition 5.1 and theorem 5.4 proves the initial claim.

Theorem 5.5 *If α is a monotone unbalanced connective and the support of μ does not contain the negations of projections, then the limiting distribution will be concentrated on a threshold function.*

5.1.1 Convergence Bounds

Except in one case, all these growth processes converge very quickly to their limiting distribution: in $O(\log(n))$ iterations. In the exceptional case the convergence requires $O(n^k)$ iterations where k is the arity of the connective α; we provide specific criteria that determine whether a process will converge quickly or not. There are two main cases: either $A_\alpha(p)$ has a fixed-point, or not. We first derive bounds for the latter case, and then for the former. Unless explicitly stated, we assume that constants are not in $supp(\mu)$, however, the following analysis changes little if constants are in $supp(\mu)$.

In the first case, either $A_\alpha(p) > p$ for $p \in (0,1)$, and $A_\alpha(p) < p$, for $p \in (0,1)$. Since, the two cases are symmetric, the same bounds apply to both. Hence, without loss of generality assume that $A_\alpha(p) < p$ on the interval $(0,1)$.

Lemma 5.6 *If α is a monotone connective such that $A_\alpha(p) < p$ on the interval $(0,1)$ and, $A_\alpha(p)$ has degree $k > 2$ and $\beta_{k-1} \leq \frac{k-2}{k}$, then for all positive $\varepsilon < \varepsilon_k = \frac{1}{k2^{k+1}}$,*

$$\frac{35}{24} < A'_\alpha(1-\varepsilon)$$

Proof: See Appendix. ■

Lemma 5.7 *If α is a monotone connective such that $A_\alpha(p) < p$ on the interval $(0,1)$ and, $A_\alpha(p)$ has degree $k > 2$ and $\beta_{k-1} = \frac{k-1}{k}$, then, for all positive $\varepsilon < k^{-1}$,*

$$1+\varepsilon^k < A'_\alpha(1-\varepsilon) \leq (1-\varepsilon)^{k-2}(k(k-2)\varepsilon+1)$$

Proof: See Appendix. ■

Lemma 5.8 *If α is a monotone connective that is not of the form $\alpha(x) = x_i \vee \alpha'(x)$, then on the interval $(0,1)$, $A_\alpha(p) < (\binom{k}{2}+1)p^2$.*

Proof: See Appendix. ■

Theorem 5.9 *Let α be a k-adic monotone connective such that $A_\alpha(p) < p$ on the interval $(0,1)$, $k > 2$ and $\beta_{k-2} \leq \frac{k-2}{k}$. There exists a constant c_α, such that for all $n > 0$, if $i \geq 3\log(n) + c_\alpha$, then for all f, $|\pi_i(f) - \pi(f)| < 2^{-n}$.*

Proof: Let $\tilde{f}_i$ be a random variable with the distribution π_i. Using an argument similar to Valiant's [Val84], we claim that if $i \geq 3\log(n) + c_\alpha$, then for $|x| = n$, $P[\tilde{f}_i(x) = 0] = 0$, and for all x such that $|x| < n$, $P[\tilde{f}_i(x) = 1] < 2^{-2n}$. The former follows from the monotonicity of α; regardless of the number of iterations, a false negative will never occur.

In the latter case, assuming that all variables are independent, if $|x| < n$, $P[\tilde{f}_0(x) = 1] = |x|/n \leq 1 - n^{-1}$. For $i > 0$, $P[\tilde{f}_i(x) = 1] = A^i_\alpha(p)$, where A^i_α denotes the ith composition of A_α with itself. Expanding $A_\alpha(p)$ around 1,

$$A_\alpha(p) = A_\alpha(1) + A'_\alpha(1)(p-1) + O((p-1)^2),$$

yields:

$$A_\alpha(1-\varepsilon) = 1 - \varepsilon A'_\alpha(1) + O(\varepsilon^2).$$

From Lemma 5.6, let $\gamma = 35/24$ and let $\varepsilon_k = \frac{1}{k2^{k+1}}$. There exists an $\varepsilon_0 < \varepsilon_k$ such that for all $\varepsilon < \varepsilon_0$

$$A_\alpha(1-\varepsilon) < 1 - \varepsilon\gamma.$$

Since $P[\tilde{f}_0(x) = 1] \leq 1 - n^{-1}$, for $i \geq 2\log(n) + 2\log(\varepsilon_0) > (\log(n) + \log(\varepsilon_0))/\log(35/24)$,

$$A^i_\alpha(1-\varepsilon) < 1 - \varepsilon\gamma^i < 1 - \varepsilon_0.$$

An additional constant number of iterations, say d_α, yields

$$A_\alpha^{d_\alpha}(1-\varepsilon_0) < c.$$

By Lemma 5.8, $A_\alpha(p) < k^2p^2$, thus we fix $c < \frac{1}{2k^2}$ and let $j = \log(n)+1$. Hence,

$$A_\alpha^j(c) < (k^2c)^{2^j} < 2^{-2^j} = 2^{-2n}.$$

Therefore, for $i \geq 3\log(n) + 2\log(\varepsilon_0) + d_\alpha + 1$ and all x such that $|x| < n$, $P[\tilde{f}_i(x) = 1] < 2^{-2n}$, implying that $|\pi_i(f) - \pi(f)| < 2^{-n}$. ■

Unfortunately, if $\beta = \frac{k-1}{k}$, convergence takes time polynomial in n. If $|x| = n-1$ then $P[\tilde{f}_0(x) = 1] = 1-n^{-1}$. Furthermore, by Lemma 5.7, for sufficiently large n, $A'_\alpha(1-n^{-1}) < (1-n^{-1})^{k-2}(k(k-2)n^{-1}+1)$. Since $\gamma < A'_\alpha(1-n^{-1})$, therefore

$$\log(\gamma) < (k-2)\log(1-n^{-1}) + \log(k(k-2)n^{-1}+1),$$

implying that

$$\log(\gamma)^{-1} > \left((k-2)\log(1-n^{-1}) + \log(k(k-2)n^{-1}+1)\right)^{-1} > \frac{n}{k^2-3k+2} + O(1).$$

Thus, if $A_\alpha^i(1-n^{-1}) < 1-\varepsilon_0$, then for sufficiently large n, $A_\alpha^i(1-n^{-1}) < \varepsilon_0$ implies that $i > (k-1)(k-2)2n(\log(n)+\log(\varepsilon_0))$. In fact, this is the best case. If $\alpha(x) = \vee_{i=2}^k(x_1 \wedge x_i)$, then $A_\alpha(p) = p - p(1-p)^{k-1}$. By Lemma 5.7, $\gamma < 1 + n^{2-k}(k-1-kn^{-1})$, implying that $\log(\gamma) < \log(1+n^{2-k}(k-1-kn^{-1}))$, and

$$\log(\gamma)^{-1} > \Big(\log(1+n^{2-k}(k-1-kn^{-1}))\Big)^{-1} > \frac{n^{k-2}}{k-1}.$$

Thus, if $A_\alpha^i(1-n^{-1}) < 1-\varepsilon_0$, then for sufficiently large n, $i > \frac{n^{k-2}}{k-1}(\log(n)+\log(\varepsilon_0))$. Consequently, connectives whose characteristic polynomial has no fixed-point can be classified as either quickly converging or slowly converging, with the value of the second last coefficient, β_{k-1}, determining rate of convergence!

When the characteristic polynomial $A_\alpha(p)$ does have a fixed-point on the interval $(0,1)$, a similar analysis is used.

Lemma 5.10 *Let $A_\alpha(p)$ be the characteristic polynomial of any k-adic monotone connective. If $A_\alpha(p)$ has a fixed-point $s \in (0,1)$, then $A'_\alpha(s) \geq 1 + \frac{k-2}{2^{k-2}}$.*

Proof: See Appendix. ■

Theorem 5.11 *Let α be a k-adic monotone connective such that $A_\alpha(s) = s \in (0,1)$. There exists a constant c_α, such that for all $n > 0$, if $i \geq k2^k\log(n) + c_\alpha$, then for all functions f, $|\pi_i(f) - \pi(f)| < 2^{-n}$.*

Proof: Let $\tilde{f}_i$ be a random variable with the distribution π_i. Using an argument similar to Valiant's [Val84], we claim that if $i \geq k2^k\log(n) + c_\alpha$ then for all x such that $|x| < sn$, $P[\tilde{f}_i(x) = 1] < 2^{-2n}$, and for all x such that $|x| > sn$, $P[\tilde{f}_i(x) = 0] < 2^{-2n}$. We first argue the former.

Assuming that all variables are independent, if $|x| < sn$, $P[\tilde{f}_0(x) = 1] \leq s - n^{-1}\varepsilon_\alpha(n)$, where $\varepsilon_\alpha(n) = \min_{j\in\mathbb{Z}}|s - \frac{j}{n}| = |s - \frac{j_0}{n}|$. Since s is an algebraic of degree k, by Liouville's Approximation Theorem [Apo97]

$$\varepsilon_\alpha(n) = \left|s - \frac{j_0}{n}\right| > \frac{e_\alpha}{n^k},$$

where the constant e_α depends only on the connective.

For $i > 0$, $P[\tilde{f}_i(x) = 1] = A^i_\alpha(p)$, where A^i_α denotes the ith composition of A_α with itself. Expanding $A_\alpha(p)$ around s,

$$A_\alpha(p) = A_\alpha(s) + A'_\alpha(s)(p-s) + O((p-s)^2),$$

yields:

$$A_\alpha(s-\varepsilon) = s - \varepsilon A'_\alpha(s) + O(\varepsilon^2).$$

By Lemma 5.10, fix $\gamma = 1 + 2^{-k+1}$; there exists an ε_0 such that for all $\varepsilon < \varepsilon_0$, $A_\alpha(s-\varepsilon) < s - \varepsilon\gamma$. Since $P[\tilde{f}_0(x) = 1] \le s - n^{-1}\varepsilon_\alpha(n)$, if

$$i \ge \log(n\, \varepsilon_\alpha(n)^{-1}\varepsilon_0)/\log(\gamma) \ge \log(n^{k+1}e_\alpha^{-1}\varepsilon_0)/\log(\gamma),$$

then

$$A^i_\alpha(s-\varepsilon) < s - \varepsilon\gamma^j < s - \varepsilon_0.$$

An additional constant number of iterations, say d_α, yields

$$A^{d_\alpha}_\alpha(s-\varepsilon_0) < c;$$

By Lemma 5.8, $A_\alpha(p) < k^2p^2$, thus, we fix $c < \frac{1}{2k^2}$ and let $j = \log(n) + 1$. Hence,

$$A^j_\alpha(c) < (k^2c)^{2^j} < 2^{-2^j} = 2^{-2n}.$$

Therefore, if

$$i \ge k2^k\log(n) + \frac{\log(e_\alpha^{-1}\varepsilon_0)}{\log(\gamma)} + d_\alpha + 1,$$

for all x such that $|x| < sn$, $P[\tilde{f}_i(x) = 1] < 2^{-2n}$.

By the same argument, if $|x| > sn$, $P[\tilde{f}_i(x) = 0] < 2^{-2n}$. Since $|x| > sn$, for $P[\tilde{f}_0(x) = 0] \le 1 - s - n^{-1}\varepsilon_\alpha(n)$, $P[\tilde{f}_1(x) = 0] = \bar{A}_\alpha(p) = 1 - A_\alpha(1-p)$, and $P[\tilde{f}_j(x) = 0] = \bar{A}^j_\alpha(p)$, $j > 0$. Just as in the preceding case, the composition of $\bar{A}_\alpha$ with itself first yields a first order divergence from $1-s$, followed by a second order convergence towards zero. Therefore, $|\pi_i(f) - \pi(f)| < 2^{-n}$. ■

To reduce the constant in front of the log term, one solution is to use a non-uniform initial distribution, as was done by Valiant [Val84].

5.2 Using Balanced Connectives

In this subsection, it will be convenient to start by assuming that the support of μ contains both constants, as well as the projections, and to deal later with the cases in which one or both constants are missing from the support of μ. If the connective is balanced, then by Lemma 5.2, its characteristic polynomial has a fixed-point of $\frac{1}{2}$. If the number n of variables is odd, then the fraction of inputs that are true for any assignment is bounded away from $\frac{1}{2}$, that is, for any $j \in \{1,2,..,n+1\}, |\frac{1}{2} - \frac{j}{n+2}| \ge \frac{1}{2n+4}$. Hence, by the standard amplification argument, the limiting distribution will be concentrated on the n-adic majority function $T_{\lceil n/2 \rceil}$. In fact, the convergence to the majority function is logarithmic in n; by Theorem 5.11, if $i \ge k2^k\log(n) + O(1)$, the $|\pi_i(f) - \pi(f)| < 2^{-n}$. When the number of variables is even, however, something completely different happens.

The family of slice functions, denoted $\mathcal{S}_{m,n}$ and defined by Berkowitz [Ber82], are monotone n-adic functions that assume the value 1 for all assignments of weight greater than m, assume the value 0 for all assignments of weight less than m, and may take on either value for assignments of weight m. Unlike other growth processes where the distribution is either concentrated on a single function or is uniform on the support of the growth process, the growth processes we are about to deal with have a limiting distribution that is uniform on $\mathcal{S}_{n/2,n}$. This set includes a large number, $2^{\binom{n}{n/2}}$, of functions; but according to a result of Korshunov [Kor80], includes only a tiny fraction, less than $\exp-(\binom{n}{n/2+1}2^{-n/2})$, of the support of the growth process, which is the set $\mathcal{M}_n$ of all monotone functions.

Define the n-adic functions

$$\chi_n(x) = \begin{cases} 1, & |x| = \frac{n}{2} \\ 0, & \textit{otherwise} \end{cases} \quad \text{and} \quad \upsilon_n(x) = \begin{cases} 1, & |x| > \frac{n}{2} \\ 0, & \textit{otherwise} \end{cases} .$$

Claim 5.12 *The Fourier coefficients of the probability distribution* π *that is uniform on the slice functions in* $\mathcal{S}_{n/2,n}$ *are given by*

$$\Delta(f) = \begin{cases} 0, & \langle f, \chi_n \rangle \neq 0 \\ (-1)^{\langle f, \upsilon_n \rangle}, & \langle f, \chi_n \rangle = 0 \end{cases}$$

Proof: Let $c = |\mathcal{S}_{\frac{n}{2},n}|^{-1} = 2^{-\binom{n}{n/2}}$. If $\langle f, \chi_n \rangle = 0$, then

$$\Delta(f) = \sum_{g \in \mathcal{F}_n} (-1)^{\langle f,g \rangle} \pi(g) = c \sum_{g \in \mathcal{S}_{\frac{n}{2},n}} (-1)^{\langle f,g \rangle} = c \sum_{g \in \mathcal{S}_{\frac{n}{2},n}} (-1)^{\langle f,\upsilon_n \rangle} = (-1)^{\langle f,\upsilon_n \rangle}.$$

Otherwise let w be a singleton such that $w \leq f \wedge \chi_n$ and let $W = \{g \in \mathcal{S}_{\frac{n}{2},n} \mid g \geq w\}$. Then

$$\Delta(f) = c \sum_{g \in \mathcal{S}_{\frac{n}{2},n}} (-1)^{\langle f,g \rangle} = c \sum_{g \in W} (-1)^{\langle f,g \rangle} + (-1)^{\langle f, g \oplus w \rangle} = 0.$$

■

We shall need to combine amplification with Fourier methods to obtain our result in this case. The following "Restriction Lemma" is the key to doing this.

Claim 5.13 *Let* α *be a balanced monotone connective. Then if* $f(x) = 1$ *for some* x *with* $|x| < n/2$*, or if* $f(x) = 0$ *for some* x *with* $|x| > n/2$*, then* $\lim_{i\to\infty} \pi_i(f) = 0$.

Proof: This follows from Theorem 5.11. ■

Lemma 5.14 (The Restriction Lemma)
If α *is a balanced monotone connective, then for all* $w \in \mathcal{F}_n$,

$$\lim_{i\to\infty} \Delta_i(w) = (-1)^{\langle \upsilon_n, w \rangle} \lim_{i\to\infty} \Delta_i(w \wedge \chi_n).$$

Proof: We begin with the definition

$$\Delta_i(w) = \sum_{v \in B_n} (-1)^{\langle v,w \rangle} \pi_i(v),$$

then rewrite the equation as

$$\Delta_i(w) = \sum_{v \in B_n} (-1)^{\langle v,w \rangle} \pi_i(v) = \sum_{t \leq \chi_n} \sum_{u \leq \overline{\chi_n}} (-1)^{\langle t \vee u, w \rangle} \pi_i(t \vee u),$$

and consider the restriction of w to the slice $\frac{n}{2}$, that is, $w \wedge \chi_n$. Since $\lim_{i\to\infty}\pi_i(t \vee u) = 0$ if $u \neq \upsilon_n$, $\lim_{i\to\infty}\Delta_i(w \wedge \chi_n)$ can be rewritten as

$$\begin{aligned}
\lim_{i\to\infty}\Delta_i(w\wedge\chi_n) &= \lim_{i\to\infty}\sum_{v\in B_n}(-1)^{\langle v,w\wedge\chi_n\rangle}\pi_i(v)\\
&= \lim_{i\to\infty}\sum_{t\leq\chi_n}\sum_{u\leq\overline{\chi_n}}(-1)^{\langle t\vee u,w\wedge\chi_n\rangle}\pi_i(t\vee u)\\
&= \sum_{t\leq\chi_n}\sum_{u\leq\overline{\chi_n}}(-1)^{\langle t\vee u,w\wedge\chi_n\rangle}\lim_{i\to\infty}\pi_i(t\vee u)\\
&= \sum_{t\leq\chi_n}(-1)^{\langle t\vee\upsilon_n,w\wedge\chi_n\rangle}\lim_{i\to\infty}\pi_i(t\vee\upsilon_n)\\
&= \sum_{t\leq\chi_n}(-1)^{\langle t,w\wedge\chi_n\rangle}(-1)^{\langle\upsilon_n,w\wedge\chi_n\rangle}\lim_{i\to\infty}\pi_i(t\vee\upsilon_n)\\
&= \sum_{t\leq\chi_n}(-1)^{\langle t,w\wedge\chi_n\rangle}\lim_{i\to\infty}\pi_i(t\vee\upsilon_n)\\
&= \sum_{t\leq\chi_n}(-1)^{\langle t,w\rangle}\lim_{i\to\infty}\pi_i(t\vee\upsilon_n).
\end{aligned}$$

This, in conjunction with

$$\begin{aligned}
\lim_{i\to\infty}\Delta_i(w) &= \lim_{i\to\infty}\sum_{t\leq\chi_n}\sum_{u\leq\overline{\chi_n}}(-1)^{\langle t\vee u,w\rangle}\pi_i(t\vee u) = \sum_{t\leq\chi_n}\sum_{u\leq\overline{\chi_n}}(-1)^{\langle t\vee u,w\rangle}\lim_{i\to\infty}\pi_i(t\vee u)\\
&= \sum_{t\leq\chi_n}(-1)^{\langle t\vee\upsilon_n,w\rangle}\lim_{i\to\infty}\pi_i(t\vee\upsilon_n) = \sum_{t\leq\chi_n}(-1)^{\langle t,w\rangle}(-1)^{\langle\upsilon_n,w\rangle}\lim_{i\to\infty}\pi_i(t\vee\upsilon_n)\\
&= (-1)^{\langle\upsilon_n,w\rangle}\sum_{t\leq\chi_n}(-1)^{\langle t,w\rangle}\lim_{i\to\infty}\pi_i(t\vee\upsilon_n) = (-1)^{\langle\upsilon_n,w\rangle}\lim_{i\to\infty}\Delta_i(w\wedge\chi_n)
\end{aligned}$$

yields the identity. ■

Hence, all we need to show is that $\lim_{i\to\infty}\Delta_i(w) = 0$ for w such that $0 < w \leq \chi_n$. To do this we use Savický's [Sav90] argument, which uses induction on the weight of w together with the recurrence

$$\Delta_{i+1}(w) = \sum_{j=1}^{k} a_j(w)\Delta_i(w)^j + y_i(w)$$

where

$$\begin{aligned}
a_j(w) &= \sum_{\substack{t\in B_k\\ |t|=j}} S_\alpha(t)^{|w|},\\
y_i(w) &= \sum_{\substack{\mathbf{v}\in\mathcal{F}_n^k\\ 0<\mathbf{v}_j<w}}\prod_{\substack{a\in B_n\\ w(a)=1}} S_\alpha(\mathbf{v}(a))\prod_{j=1}^{k}\Delta_i(\mathbf{v}_j),\\
S_\alpha(t) &= \frac{1}{2^k}\sum_{r\in B_k}(-1)^{\langle r,t\rangle}(-1)^{\alpha(r)}.
\end{aligned}$$

Since the connective is not linear, this recurrence is much more complicated than the one in proposition 3.1. The result is the following proposition.

Proposition 5.15 *Let α be a monotone balanced non-projection connective, n be even and the support of μ comprise the projection functions and constants. If $0 < w \leq \chi_n$, then $\lim_{i\to\infty}\Delta_i(w) = 0$.*

Proof: Let $w \leq \chi_n$ and recall equation 1:

$$\Delta_0(w) = \sum_{f \in F_n} \pi_0(f)(-1)^{\langle f,w \rangle} = \frac{1}{n+2} \sum_{f \in A_0} (-1)^{\langle f,w \rangle}.$$

To prove this proposition we need only show that $\Delta_0(w) = 0$ if $|w| = 1$, and that $|\Delta_0(w)| < 1$ if $|w| = 2$. In the first case, since $w \leq \chi_n$, w is true on a single assignment of weight $n/2$. Hence, $\langle w, x_i \rangle = 1$ for exactly half the projections, where x_i is the ith projection function. Hence, the projections cancel each other out. Similarly, the two constants annihilate one another. Hence, $\Delta_0(w) = 0$ if $|w| = 1$.

The latter case, $|w| = 2$, is only slightly harder. Since the constant 0 is part of the support, there will at least one positive contribution, $-1^{\langle \mathbf{0},w \rangle}\pi_0(\mathbf{0}) = \frac{1}{n+2}$. Hence, in order for $|\Delta_0(w)| = 1$ all other contributions must also be positive, specifically, $\langle w, x_i \rangle = 0$ for all x_i; by the pigeonhole principle this is not possible. Hence, $|\Delta_0(w)| < 1$ if $|w| = 2$. Substituting these base cases into Theorem 5.3 of [Sav90] yields the result. ■

This proposition, together with Claim 5.12, yields the one of our main results.

Theorem 5.16 *Let α be a monotone balanced non-projection connective, n be even and the support of μ comprise the projection functions and constants. Then the limiting distribution is uniform on the functions in $\mathcal{S}_{n/2,n}$.*

5.2.1 Convergence Bounds

To bound the convergence within the slice we use a theorem of Savický [Sav95a]; the conditions of the theorem are verified in Theorem 5.15.

Theorem 5.17 ((Savický, 4.8 in [Sav95a])) *If α is balanced and nonlinear, $\Delta_0(w) = 0$ for every w such that $|w| = 1$, $\Delta_0(w) < 1$ for every w such that $|w| = 2$, and there exists a w such that $|w| = 2$ and $\Delta_0(w) > 0$, then*

$$\max_{f \in \mathcal{F}_n} |\pi_i(f) - \pi(f)| = e^{-\Omega(i)}$$

A more explicit bound, in terms of the number of arguments and the arity of the connective, is possible. We use a more explicit version of Lemma 5.14 and bound the convergence of the growth processes characterized by Theorem 5.3 in [Sav90]. A corollary of Lemma 5.18 is that the same bound also applies to the growth processes on monotone formulas whose limiting distribution is uniform over the slice functions.

Lemma 5.18 *If α is a balanced monotone connective, then for all $w \in \mathcal{F}_n$,*

$$\Delta_i(w) = O(\varepsilon^{2^i}) + (-1)^{\langle \upsilon_n, w \rangle} \Delta_i(w \wedge \chi_n).$$

Proof: See appendix. ■

Theorem 5.19 ((Savickỳ, 5.3 in [Sav90])) *Let $\mathbf{0} < w \in \mathcal{F}_n$, let α be a k-adic nonlinear balanced connective, and assume that the initial distribution is uniform over the projections, negations, and constants. The* $\lim_{i\to\infty} \Delta_i(w) = 0$.

Lemma 5.20 *Assume that the conditions of Theorem 5.19 are satisfied and let $a = \sum_{t \in B_k} |S_\alpha(t)|^3 < 1 - 2^{-k}$. If $|w| = d \geq 2$ and*

$$i_d = n2^k \log(a^{-1}) + \sum_{j=3}^{d} \frac{(k+1)^j j}{\log(a^{-1})}, \tag{3}$$

then $|\Delta_i(w)| \leq a^{i-i_d} b_d(i)$, where $b_d(i) = (i - i_2 + 2)^{(k+1)^{d-3}}$, and $b_2(i) = 1$.

Proof: See appendix. ■

Theorem 5.21 *Assume that the conditions of Theorem 5.19 are satisfied, let a be as in Lemma 5.20, and let*

$$I = n2^k \log(a^{-1}) + \frac{2^{2n}(k+1)^{2^n}}{\log(a^{-1})}.$$

For any positive $c < 1$, if $w \neq \mathbf{0}$ and

$$i \geq \frac{\log(c)}{\log(a)} + (k+1)^{2^n} \frac{\log(i - I + 2)}{\log(a^{-1})} + I, \tag{4}$$

then $|\Delta_i(w)| \leq c$.

Proof: By Lemma 5.20 the coefficient of weight 2^n has the greatest converging bound:

$$|\Delta_{i+1}(w)| \leq a^{i-i_{2^n}}(i - n2^k \log(a^{-1}) + 2)^{(k+1)^{2^n}}$$

where

$$\begin{aligned} i_{2^n} &= n2^k \log(a^{-1}) + \sum_{j=3}^{2^n} \frac{(k+1)^j j}{\log(a^{-1})} \\ &\leq n2^k \log(a^{-1}) + \frac{2^{2n}(k+1)^{2^n}}{\log(a^{-1})} = I \end{aligned}$$

Solving for i in the inequality

$$a^{i-i_{2^n}}(i - n2^k \log(a^{-1}) + 2)^{(k+1)^{2^n}} < c$$

completes the proof. ■

Thus, by equation 2

$$\begin{aligned} \pi_i(g) &= \frac{1}{2^{2^n}} \sum_{f \in \mathcal{F}_n} (-1)^{\langle f,g \rangle} \Delta_i(f) \\ &= \frac{1}{2^{2^n}} + \frac{1}{2^{2^n}} \sum_{f \in \mathcal{F}_n \setminus \mathbf{0}} (-1)^{\langle f,g \rangle} \Delta_i(f) \\ &\leq \frac{1}{2^{2^n}} + \frac{1}{2^{2^n}} \sum_{f \in \mathcal{F}_n \setminus \mathbf{0}} |\Delta_i(f)| \\ &\leq \frac{1}{2^{2^n}} + \max_{f \in \mathcal{F}_n \setminus \mathbf{0}} |\Delta_i(f)|, \end{aligned}$$

implying that for all $g \in \mathcal{F}_n$, $|\pi(g) - \pi_i(g)| \leq c$ if i satisfies equation 4.

5.2.2 Varying the Initial Distribution

Theorem 5.15 can easily be modified to cover the cases in which one of the constants is missing from the support of μ: in these cases there is concentration on a single function when n is even and uniform distribution on a set of slice functions when n is odd. When the support of μ consists only of the projection functions, however, the situation can be more complicated. If α is not self-dual or n is odd, the result is the same as when both constants are present. If α is self-dual and n is even, however, the result is given by the following theorem.

Theorem 5.22 *Let α be a monotone self-dual non-projection connective, n be even and the support of μ comprise the projection functions. Then the limiting distribution is uniform on the self-dual functions in $\mathcal{S}_{n/2,n}$.*

Proof: Same as theorem 4.1. ■

We note that there are $2^{\frac{1}{2}\binom{n}{n/2}}$ self-dual functions in $\mathcal{S}_{n/2,n}$. According to a result of Sapozhenko [Sap89], this is only a tiny fraction, less than $\exp -(\binom{n}{n/2+1}2^{-n/2-1})$, of the support of the growth process, which is the set of self-dual monotone functions.

6 Growing Other Functions

We can use the same method to analyze other growth processes. For example, the uniform distribution on the set of bi-preserving functions (that is, those functions satisfying $f(0,\ldots,0) = 0$ and $f(1,\ldots,1) = 1$) can be generated by a growth process that uses the bi-preserving selection connective $\alpha(x,y,z) = xy \vee \bar{x}z$ and an initial distribution that is uniform on the projection functions. The same technique as in the monotone case is sufficient to prove this; the corresponding restriction lemma yields the identity

$$\lim_{i\to\infty} \Delta_i(w) = (-1)^{\langle w,\eta_n\rangle}\Delta(w \wedge \kappa_n)$$

where $\eta_n = \wedge_{j=1}^n x_j$ and $\kappa_n = \vee_{j=1}^n x_j \bigwedge \overline{\wedge_{j=1}^n x_j}$. Similar analysis for the 0-preserving and 1-preserving functions follows easily.

7 Conclusion

In this paper we have developed techniques for analyzing growth processes when the limiting distribution is uniform over a set of Boolean functions. In particular, we can deal with situations in which this set comprises neither a single function nor all Boolean functions.

We believe that straightforward extensions of the techniques used here will yield a classification of all situations leading to uniform distributions over sets of functions. The step that remains to be taken is the classification of all sets of functions that can be computed by formulas that are complete k-ary trees built from a single connective. There is some work by Kudryavtsev [Kud60b, Kud60a] on this problem, but it stops short of a complete classification.

A more adventurous direction for further work is to deal with situations leading to non-uniform distributions. Empirical computations show that these situations can be quite complicated, and we are not yet able to formulate a conjecture that covers all our data.

References

[And84] A. E. Andreev. On the problem of minimizing disjunctive normal forms. *Soviet Math. Doklady*, 29:32–36, 1984.

[Apo97] T. Apostol. *Modular Functions and Dirichlet Series in Number Theory*. Springer, 2nd edition, 1997.

[Ber82] S. Berkowitz. On some relationships between monotone and nonmonotone circuit complexity. Technical report, Department of Computer Science, University of Toronto, Canada, Toronto, Canada, 1982.

[Bop85] R. Boppana. Amplification of probabilistic Boolean formulas. In *26th Annual Symposium on Foundations of Computer Science*, pages 20–29, Oct 1985.

[Bop89] R. Boppana. Amplification of probabilistic Boolean formulas. *Advances in Computing Research 5: Randomness and Computation*, pages 27–45, 1989.

[DZ97] M. Dubiner and U. Zwick. Amplification by read-once formulas. *SIAM Journal on Computing*, 26(1):15–38, January 1997.

[Gla67] V. V. Glagolev. Nekotorye otsenki dizyunktivnykh normalnykh form funktsiĭ algebry logiki. *Problemy Kibernetiki*, 19:75–95, 1967.

[Kor80] A. D. Korshunov. O chisle monotonnykh bulevykh funktsiĭ. *Problemy Kibernetiki*, 38:5–100, 1980.

[Kud60a] V. B. Kudryavcev. Completeness theorem for a class of automata without feedback couplings. *Soviet Math. Doklady*, 1:537–539, 1960.

[Kud60b] V. B. Kudryavcev. Problems of completeness for automatic machine systems. *Soviet Math. Doklady*, 1:146–149, 1960.

[LS97] H. Lefmann and P. Savický. Some typical properties of large and/or Boolean functions. *Random Structures and Algorithms*, 10:337–351, 1997.

[Lup61] O. B. Lupanov. O realizatsiĭ funktsiĭ algebry logiki formulami iz konechnykh klassov (formulami ogranichennoĭ glubiny) v bazise &, $\vee$, $^{-}$. *Problemy Kibernetiki*, 6:5–14, 1961.

[MS56] E. Moore and C. Shannon. Reliable circuits using less reliable relays. *Journal of Franklin Institute*, 262(3):191–208, 1956.

[Raz88] A. Razborov. Formulas of bounded depth in basis $(\wedge, \oplus)$ and some combinatorial problems. *Voprosy Kibernetiky, USSR*, pages 149–166, 1988.

[RS42] J. Riordan and C. E. Shannon. The number of two-terminal series-parallel networks. *J. Math. Phys.*, 21:83–93, 1942.

[Sap89] A. A. Sapozhenko. O chisle antitsepeĭ v mnogosloĭnykh ranzhirovannykh mnozhestvakh. *Diskretnaya Matematika*, 1:110–128, 1989.

[Sav90] P. Savický. Random Boolean formulas representing any Boolean function with asymptotically equal probability. *Discrete Mathematics*, 83:95–103, 1990.

[Sav95a] P. Savický. Bent functions and random Boolean formulas. *Discrete Mathematics*, 147:211–234, 1995.

[Sav95b] P. Savický. Improved Boolean formulas for the Ramsey graphs. *Random Structures and Algorithms*, 6:407–415, 1995.

[Sav98] P. Savický. Complexity and probability of some Boolean formulas. *Combinatorics, Probability and Computing*, 7(4):451–463, 1998.

[Sha38] C. E. Shannon. A symbolic analysis of relay and switching circuits. *AIEE Trans.*, 57:713–723, 1938.

[Val84] L. Valiant. Short monotone formulae for the majority function. *Journal of Algorithms*, 5:363–366, 1984.

[vN56] J. von Neumann. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*. Princeton University Press, 1956.

A Proofs of Lemmas 5.6, 5.7, 5.8, 5.10, and 5.18

Fact A.1 *If α is a monotone connective, then:*

1. $A_\alpha(p) > p$ *on the interval* $(0,1)$ *if and only if* $\alpha(x) = x_i \vee \alpha'(x)$ *if and only if* $\beta_1 > 0$.
2. $\beta_{k-1} = \frac{a}{k}$, $a \in \{0,1,\ldots,k-1\}$.
3. $\beta_0 = 0$ *and* $\beta_k = 1$.
4. $A_\alpha(s) = s \in (0,1)$ *if and only if* $A_\alpha(p) < p,\ p \in (0,s)$ *and* $A_\alpha(p) > p,\ p \in (s,1)$ *if and only if* $A'_\alpha(0) = A'_\alpha(1) = 0$ *if and only if* $\beta_0 = \beta_1 = 0$ *and* $\beta_k = \beta_{k-1} = 1$.

A.1 Proof of Lemma 5.6

Proof: Begin by differentiating $A_\alpha(p)$

$$
\begin{aligned}
A'_\alpha(p) &= \frac{d}{dp}\left(\sum_{i=0}^{k} \beta_i \binom{k}{i} p^i (1-p)^{k-i}\right) \\
&= \frac{d}{dp}\left(p^k + \beta_{k-1} k p^{k-1}(1-p) + \sum_{i=0}^{k-2} \beta_i \binom{k}{i} p^i (1-p)^{k-i}\right) \\
&= kp^{k-1} + \beta_{k-1} k p^{k-2}(k - kp - 1) - \sum_{i=0}^{k-2} \beta_i \binom{k}{i} p^{i-1}(1-p)^{k-i-1}(pk - i),
\end{aligned}
$$

and evaluate at $1-\varepsilon$

$$
\begin{aligned}
A'_\alpha(1-\varepsilon) &= k(1-\varepsilon)^{k-2}(1-\varepsilon+\beta_{k-1}k\varepsilon-\beta_{k-1}) - \sum_{i=0}^{k-2} \beta_i \binom{k}{i}(1-\varepsilon)^{i-1}(\varepsilon)^{k-i-1}(k-i-k\varepsilon) \\
&> k(1-\varepsilon)^{k-2}(1-\varepsilon+\beta_{k-1}k\varepsilon-\beta_{k-1}) - \sum_{i=0}^{k-2} \beta_i \binom{k}{i}(1-\varepsilon)^{i-1}(\varepsilon)^{k-i-1}k \\
&> k(1-\varepsilon)^{k-2}(1-\varepsilon+\beta_{k-1}k\varepsilon-\beta_{k-1}) - \sum_{i=0}^{k-2} \beta_i \binom{k}{i}(\varepsilon)^{k-i-1}k \\
&= k(1-\varepsilon)^{k-2}(1-\varepsilon+\beta_{k-1}k\varepsilon-\beta_{k-1}) - k\varepsilon\sum_{i=0}^{k-2} \beta_i \binom{k}{i}(\varepsilon)^{k-i-2} \\
&> k(1-\varepsilon)^{k-2}(1-\varepsilon+\beta_{k-1}k\varepsilon-\beta_{k-1}) - k\varepsilon\sum_{i=0}^{k} \binom{k}{i} \\
&> k(1-\varepsilon)^{k-2}(1-\varepsilon+\frac{k-2}{k}(k\varepsilon-1)) - k\varepsilon 2^k \\
&= (1-\varepsilon)^{k-2}(2+k\varepsilon(k-3)) - k\varepsilon 2^k.
\end{aligned}
$$

For $k > 2$ and $\varepsilon < \varepsilon_k$,

$$\begin{aligned} h'(1-\varepsilon) &= (1-\varepsilon)^{k-2}(2+k\varepsilon(k-3)) - k\varepsilon 2^k \\ &> (1-\varepsilon)^{k-2}(2+k\varepsilon(k-3)) - \frac{1}{2} \\ &> (1-\frac{1}{48})2 - \frac{1}{2} \\ &= \frac{35}{24} \end{aligned}$$

■

A.2 Proof of Lemma 5.7

Proof: For the lower bound, observe that since

$$A'_\alpha(p) = kp^{k-1} - \sum_{i=0}^{k-1} \beta_i \binom{k}{i} p^{i-1}(1-p)^{k-i-1}(pk-i),$$

for all $p > 1-k^{-1}$, minimizing $A'_\alpha(p)$, maximizes the coefficients β_i, for all $i = 2\ldots k-2$. Since the connective is of the form $\alpha(x) = x_j \wedge \alpha'(x)$, $\beta_i \le \binom{k-1}{i}\binom{k}{i}^{-1} = \frac{k-i}{k}$. Hence,

$$A'_\alpha(p) \ge kp^{k-1} - \sum_{i=0}^{k-1} \binom{k-1}{i} p^{i-1}(1-p)^{k-i-1}(pk-i) = 1+(1-p)^{k-2}(kp-1).$$

Thus, for all $\varepsilon < k^{-1}$

$$A'_\alpha(1-\varepsilon) \ge 1+\varepsilon^{k-2}(k-1-k\varepsilon) > 1+\varepsilon^{k-2} > 1+\varepsilon^k.$$

For the upper bound we minimize all β_i for $i = 2\ldots k-2$, i.e., $\beta_i = 0$. Thus, for all $p > 1-k^{-1}$,

$$A'_\alpha(p) \le kp^{k-1} - (k-1)p^{k-2}(pk-k+1) = p^{k-2}(k(k-2)(1-p)+1),$$

implying that for all $\varepsilon < k^{-1}$,

$$A'_\alpha(1-\varepsilon) \le (1-\varepsilon)^{k-2}(k(k-2)\varepsilon+1).$$

■

A.3 Proof of Lemma 5.8

Proof: Since $A_\alpha(p) \not> p$ on all $(0,1)$, hence α is not of the form $\alpha(x) = x_i \vee \alpha'(x)$. Thus, the first two coefficient of A_α are $\beta_0 = \beta_1 = 0$. Thus, on the interval $(0,1)$,

$$A_\alpha(p) = \beta_2 \binom{k}{2} p^2 + B(p)p^3 \le \binom{k}{2} p^2 + B(p)p^2 < (\binom{k}{2}+1)p^2$$

since $B(p) < 1$. ■

A.4 Proof of Lemma 5.10

Proof: Consider a projection connective, say $\chi(x) = x_b$, $b \in [1,n]$. The corresponding characteristic polynomial is

$$A_\chi(p) = p = p^k + p(1-p)^{k-1} + \sum_{i=2}^{k-1} \binom{k-1}{i} p^i (1-p)^{k-i},$$

whose fixed-point is everywhere and whose slope is 1. Note that $\beta_1 = 1$ and $\beta_{k-1} = \frac{k-1}{k}$. Let

$$\eta(x) = \wedge_{i=2}^k x_i \bigvee \vee_{i=2}^k (x_1 \wedge x_i),$$

be a k-adic monotone connective. The corresponding characteristic polynomial

$$\begin{aligned} A_\eta(p) &= p - p(1-p)^{k-1} + p^{k-1}(1-p) \\ &= A_\chi(p) - p(1-p)^{k-1} + p^{k-1}(1-p) \\ &= A_\chi(p) + (A_\eta - A_\chi(p)) \end{aligned}$$

has a fixed-point at $\frac{1}{2}$. Not surprisingly, this is almost $A_\chi(p)$ except that $\beta_1 = 0$ and $\beta_{k-1} = 1$, i.e., the difference is just two terms. We claim that $A'_\eta(\frac{1}{2}) \le A'_\alpha(s)$.

The claim is proved by contradiction. Assume that $A'_\eta(\frac{1}{2})$ is not the minimum slope at a fixed-point, then there exists a k-adic monotone connective ζ, whose degree k characteristic polynomial $A_\zeta(p)$ has a fixed-point $t \in (0,1)$, such that $A'_\zeta(t) < A'_\eta(\frac{1}{2})$. Since $\beta_1 = 0$ and $\beta_{k-1} = 1$ must hold for A_ζ, we write $A_\zeta(p)$ in a manner similar to $A_\eta(p)$: $A_\zeta(p) = A_\eta(p) + (A_\zeta(p) - A_\eta(p))$. Specifically we are interested in the differences between $A_\zeta(p)$ and $A_\eta(p)$. In fact,

$$\begin{aligned} A_\zeta(p) &= A_\eta(p) + \left[p^{i_0}(1-p)^{k-i_0} + p^{i_1}(1-p)^{k-i_1} \ldots\right] \\ &\quad - \left[p^{j_0}(1-p)^{k-j_0} + p^{j_1}(1-p)^{k-j_1} \ldots\right], \end{aligned}$$

where $i_l \le i_{l+1} \le k-2$ and $j_l \ge j_{l+1} \ge 2$.

Since $A_\zeta(p) \ne A_\eta(p)$, and $A_\zeta(p)$ has a fixed-point on $(0,1)$, either i_0 or j_0 must exist. Without loss of generality assume that i_0 exists and consider the characteristic polynomial $A_\delta(p) = A_\zeta(p) - p^{i_0}(1-p)^{k-i_0}$. Since the connective corresponding to $A_\zeta(p)$ is monotone, $j_0 < kt < i_0$. which implies that the derivative $p^{i_0-1}(1-p)^{k-i_0-1}(i_0 - kp)$ of the term $p^{i_0}(1-p)^{k-i_0}$ is positive for all $u \le t$. Since the fixed-point u of $A_\delta(p)$ is in the interval $(\frac{j_0}{k}, t)$, $A'_\delta(u) < A'_\zeta(t)$, which is a contradiction! In fact, iteratively subtracting the terms $p^{i_l}(1-p)^{k-i_l}$ and adding the terms $p^{j_l}(1-p)^{k-j_l}$, reduces $A_\zeta(p)$ to $A_\eta(p)$!.

Evaluating $A'_\eta(\frac{1}{2}) = 1 + \frac{k-2}{2^{k-2}}$ completes the proof. ■

A.5 Proof of Lemma 5.18

Proof: Following the proof Lemma 5.14 begin with the definition

$$\Delta_i(w) = \sum_{v \in B_n} (-1)^{\langle v,w \rangle} \mu_i(v),$$

and rewrite the equation as

$$\Delta_i(w) = \sum_{v \in B_n} (-1)^{\langle v,w \rangle} \mu_i(v) = \sum_{t \le \chi_n} \sum_{u \le \overline{\chi_n}} (-1)^{\langle t \vee u, w \rangle} \mu_i(t \vee u).$$

Consider the restriction of w to the slice $\frac{n}{2}$, that is, $w \wedge \chi_n$. By Theorem 5.11, after $O(\log(n))$ iterations $\mu_i(t \vee u) = O(\varepsilon^{2^i})$, if $u \neq \upsilon_n$. Hence, for $i > O(\log(n))$, $\Delta_i(w \wedge \chi_n)$ can be rewritten as

$$\begin{aligned}
\Delta_i(w \wedge \chi_n) &= \sum_{v \in B_n} (-1)^{\langle v, w \wedge \chi_n \rangle} \mu_i(v) \\
&= \sum_{t \leq \chi_n} \sum_{u \leq \overline{\chi_n}} (-1)^{\langle t \vee u, w \wedge \chi_n \rangle} \mu_i(t \vee u) \\
&= \sum_{t \leq \chi_n} (-1)^{\langle t \vee \upsilon_n, w \wedge \chi_n \rangle} \mu_i(t \vee \upsilon_n) + \sum_{t \leq \chi_n} \sum_{u \leq \overline{\chi_n}, u \neq \upsilon_n} (-1)^{\langle t \vee u, w \wedge \chi_n \rangle} \mu_i(t \vee u) \\
&= \sum_{t \leq \chi_n} (-1)^{\langle t \vee \upsilon_n, w \wedge \chi_n \rangle} \mu_i(t \vee \upsilon_n) + \sum_{t \leq \chi_n} \sum_{u \leq \overline{\chi_n}, u \neq \upsilon_n} O(\varepsilon^{2^i}) \\
&= O(\varepsilon^{2^i}) + \sum_{t \leq \chi_n} (-1)^{\langle t, w \rangle} \mu_i(t \vee \upsilon_n).
\end{aligned}$$

This, in conjunction with

$$\begin{aligned}
\Delta_i(w) &= \sum_{t \leq \chi_n} \sum_{u \leq \overline{\chi_n}} (-1)^{\langle t \vee u, w \rangle} \mu_i(t \vee u) \\
&= \sum_{t \leq \chi_n} (-1)^{\langle t \vee \upsilon_n, w \rangle} \mu_i(t \vee \upsilon_n) + \sum_{t \leq \chi_n} \sum_{u \leq \overline{\chi_n}, u \neq \upsilon_n} (-1)^{\langle t \vee u, w \rangle} \mu_i(t \vee u) \\
&= \sum_{t \leq \chi_n} (-1)^{\langle t \vee \upsilon_n, w \rangle} \mu_i(t \vee \upsilon_n) + \sum_{t \leq \chi_n} \sum_{u \leq \overline{\chi_n}, u \neq \upsilon_n} O(\varepsilon^{2^i}) \\
&= O(\varepsilon^{2^i}) + (-1)^{\langle \upsilon_n, w \rangle} \sum_{t \leq \chi_n} (-1)^{\langle t, w \rangle} \mu_i(t \vee \upsilon_n) \\
&= O(\varepsilon^{2^i}) + (-1)^{\langle \upsilon_n, w \rangle} \Delta_i(w \wedge \chi_n)
\end{aligned}$$

yields the identity. ■

B Proof of Lemma 5.20

Lemma B.1 ((Savicky, 1990, Lemma 5.1 in [Sav90]))
Let $x_{i+1} = \sum_{j=1}^{k} a_j x_i^j$, such that $|x_0| < 1$, $\sum_{j=1}^{k} |a_j| \leq 1$, and $|a_1| < 1$. Then $|x_{i+1}| \leq a|x_i|$, where $a \leq |a_1| + |x_0|(1 - |a_1|)$. Hence, $|x_i| \leq a^i |x_0|$.

Lemma B.2 ((Savicky, 1990, Lemma 4.7 in [Sav90]))
For any k-adic connective α, $\sum_{t \in B_k} S_\alpha(1)^2 = 1$, and α is balanced and nonlinear if and only if $S_\alpha(\mathbf{0}) = 0$ and for all $s \in B_k$, $|S_\alpha(s)| < 1$.

Lemma B.3 *Let α be a balanced nonlinear k-adic connective and let $|w| = 2$. For any positive $c < 1$, if*

$$i > \log(c^{-1}) n 2^k,$$

then $|\Delta_i(w)| \leq c$.

Proof: By Theorem 5.3 in [Sav90],

$$\Delta_i(w) = \sum_{j=1}^{k} a_j(w) \Delta_{i-1}(w)^j,$$

where

$$a_j(w) \;=\; \sum_{\substack{t\in B_k\\ |t|=j}} S_\alpha(t)^{|w|} \;=\; \sum_{\substack{t\in B_k\\ |t|=j}} S_\alpha(t)^2.$$

By corollary B.1, $|\Delta_i(w)| \le a^i|\Delta_0(w)|$, where $a \le |a_1(w)+\Delta_0(w)(1-a_1(w))$. By Theorem 5.3 in [Sav90], $\Delta_0(w) < \frac{n-1}{n+1}$, hence,

$$a \;\le\; |a_1(w)+\frac{n-1}{n+1}(1-a_1(w)) \;=\; 1-\frac{2}{n+1}(1-a_1(w)),$$

and since $a_1(w) \le 1-2^{-k}$,

$$a \le 1-\frac{2}{n+1}2^{-k}.$$

Solving for i in the inequality $a^i|\Delta_0(w)| \le c$ yields:

$$i \ge \frac{\log(c)-\log|x_0|}{\log(a)},$$

and substituting for a on the right:

$$\begin{aligned}
\frac{\log(c)-\log|x_0|}{\log(a)} &= \frac{\log(c)-\log|x_0|}{\log\left(1-\frac{2}{n+1}2^{-k}\right)}\\
&< ((n+1)2^{k-1}+\frac{1}{2})(\log(c^{-1})+\log|x_0|)\\
&< ((n+1)2^{k-1}+\frac{1}{2})\log(c^{-1})\\
&< n2^k\log(c^{-1})
\end{aligned}$$

Thus, for $i > \log(c^{-1})n2^k$, $|\Delta_i(w)| \le c$. ■

Lemma B.4 ((Savicky, 1990, Lemma 5.2 in [Sav90]))
Let $x_{i+1} = y_i + \sum_{j=1}^k a_j x_i^j$, such that $|x_i| \le 1$, and $a = \sum_{j=1}^k |a_j| < 1$. Then $|x_{i+1}| \le a|x_i| + y_i$.

Corollary B.5 *If $y_i < (i-l+2)^k a^{i-l+1} < a$ for some $k \ge 0$ and $l > 0$, then $|x_i| \le (i-l+1)^{k+1}a^{i-l}$ for all $i \ge l$.*

Proof: By Lemma B.4, $|x_{i+1}| \le a|x_i| + y_i$. Hence, by induction on i

$$\begin{aligned}
|x_{i+l}| &\le a^i(|x_l| + (i-l+2)^k a^l)\\
&\le a^{i-l}\left(1+\sum_{j=l}^{i}(i-l+2)^k\right)\\
&\le a^{i-l}(i-l+2)^{k+1}
\end{aligned}$$

■

B.1 Proof of Lemma 5.20

Proof: The proof is by induction on j. By Theorem 5.3 in [Sav90],

$$\Delta_i(w) \;=\; y_i(w) + \sum_{j=1}^{k} a_j(w)\Delta_{i-1}(w)^j,$$

where

$$a_j(w) = \sum_{\{t\in B_k:\ |t|=j\}} S_\alpha(t)^{|w|}, = \sum_{\{t\in B_k:\ |t|=j\}} S_\alpha(t)^d,$$

and

$$y_i(w) = \sum_{\mathbf{v}\in G_w^k} \left(\prod_{a\in B_n|w(a)=1} S_\alpha(\mathbf{v}(a))\right)\left(\prod_{j=1}^{k}\Delta_i(\mathbf{v}_j)\right),$$

where $G_w = \{v \in B_{2^n} \mid \mathbf{0} < v < w\}$.

Since $\sum_{t\in B_k} S_\alpha(1)^2 = 1$ and $|S_\alpha(s)| < 1$ for all $s \in B_k$,

$$\begin{aligned}
a(w) &= \sum_{j=1}^{k} a_j(w) \;=\; \sum_{j=1}^{k}\sum_{t\in B_k:|t|=j} S_\alpha(t)^{|w|}\\
&= \sum_{t\in B_k} S_\alpha(t)^{|w|} \;\le\; \sum_{t\in B_k} |S_\alpha(t)|^{|w|}\\
&\le \sum_{t\in B_k} |S_\alpha(t)|^3 \;<\; \sum_{t\in B_k} S_\alpha(t)^2 = 1,
\end{aligned}$$

The base case, $d = 2$, is proved by Lemma B.3. For the base case, $d = 3$, we first bound $y_i(w)$. By Theorem 5.3 in [Sav90], $\Delta_i(w) = 0$ if $|w| = 1$ and by Lemma B.3, $\Delta_i(w) \le a^{i-n2^k\log(a^{-1})}$ if $|w| = 2$. Hence, for all $w \in G_w$, $\Delta_i(w) \le a^{i-n2^k\log(a^{-1})}$. If we let $i_2 = n2^k\log(a^{-1})$ then $\Delta_i(w) \le a^{i-i_2}b_2(i)$ and,

$$\begin{aligned}
y_i(w) &= \sum_{\mathbf{v}\in G_w^k} \left(\prod_{a\in B_n|w(a)=1} S_\alpha(\mathbf{v}(a))\right)\left(\prod_{j=1}^{k}\Delta_i(\mathbf{v}_j)\right)\\
&< \sum_{\mathbf{v}\in G_w^k}\prod_{j=1}^{k}\Delta_i(\mathbf{v}_j)\\
&< \sum_{\mathbf{v}\in G_w^k}\left(a^{i-i_{d-1}}b_{d-1}(i)\right)^k\\
&= \left(\sum_{j=1}^{d-1}\binom{d}{j}\right)^k \left(a^{i-i_{d-1}}b_{d-1}(i)\right)^k\\
&< \left(2^d a^{i-i_{d-1}}b_{d-1}(i)\right)^k.
\end{aligned}$$

Solving for i the inequality

$$\left(2^d a^{i-i_{d-1}}b_{d-1}(i)\right)^k \;<\; a$$

yields

$$\frac{d+\log(b_{d-1}(i))}{\log(a^{-1})} + k^{-1} + i_{d-1} \;\le\; \frac{(k+1)^d d}{\log(a^{-1})} + i_{d-1} \;=\; i_d \;\le\; i,$$

which equals to equation 3 when evaluated at $d = 3$. Hence, for $i \geq i_d$

$$y_i(w) \ < \ \left(2^d a^{i-i_{d-1}} b_{d-1}(i)\right)^k \ < \ b_{d-1}(i)^k a^{i-i_d} < a,$$

and by Lemma B.4 and Corollary B.5 we complete the base case:

$$\begin{aligned}
|\Delta_{i+1}(w)| &\leq a|\Delta_i(w)| + y_i(w) \\
&\leq a|\Delta_i(w)| + a^{i-i_d} b_{d-1}(i)^k \\
&\leq a^{i-i_d}\left(1 + \sum_{j=i_d}^{i} 1\right) \\
&\leq a^{i-i_d}(i - i_d + 2) \\
&\leq a^{i-i_d}(i - i_2 + 2) \\
&= a^{i-i_d} b_d(i).
\end{aligned}$$

Assume that the hypothesis holds for all w of weight less than some fixed d and let $|w| = d$. Repeating the above calculations in terms of d we get an identical bound for $y_i(w)$

$$y_i(w) \ < \ \left(2^d a^{i-i_{d-1}} b_{d-1}(i)\right)^k \ < \ b_{d-1}(i)^k a^{i-i_d} < a,$$

where, by the inductive hypothesis,

$$\begin{aligned}
i_d &= \frac{(k+1)^d d}{\log(a^{-1})} + i_{d-1} \\
&= \frac{(k+1)^d d}{\log(a^{-1})} + n2^k \log(a^{-1}) + \sum_{j=3}^{d-1} \frac{(k+1)^j j}{\log(a^{-1})} \\
&= n2^k \log(a^{-1}) + \sum_{j=3}^{d} \frac{(k+1)^j j}{\log(a^{-1})}
\end{aligned}$$

and hence

$$\begin{aligned}
|\Delta_{i+1}(w)| &\leq a|\Delta_i(w)| + y_i(w) \\
&\leq a|\Delta_i(w)| + a^{i-i_d} b_{d-1}(i)^k \\
&\leq a^{i-i_d}\left(1 + \sum_{j=i_d}^{i} b_{d-1}^k\right) \\
&\leq a^{i-i_d}\left(1 + \sum_{j=i_d}^{i} (i - i_2 + 2)^{(k+1)^{d-1-3}k}\right) \\
&\leq a^{i-i_d}(i - i_d + 1)(i - i_2 + 2)^{(k+1)^{d-1-3}k} \\
&\leq a^{i-i_d}(i - i_2 + 2)(i - i_2 + 2)^{(k+1)^{d-1-3}k} \\
&= a^{i-i_d}(i - i_2 + 2)^{(k+1)^{d-3}} \\
&= a^{i-i_d} b_d(i).
\end{aligned}$$

completing inductive step. ■