Combinatorics, Probability and Computing (1998) 7, 451–463. Printed in the United Kingdom
© 1998 Cambridge University Press

Complexity and Probability of Some Boolean Formulas

PETR SAVICKÝ†

Institute of Computer Science,
Academy of Sciences of Czech Republic,
Prague, Czech Republic
(e-mail: savicky@uivt.cas.cz)

Received 26 July 1994; revised 3 November 1997

For any Boolean function f, let $L(f)$ be its formula size complexity in the basis $\{\wedge, \oplus, 1\}$. For every n and every $k \leqslant n/2$, we describe a probabilistic distribution on formulas in the basis $\{\wedge, \oplus, 1\}$ in some given set of n variables and of size at most $\ell(k) = 4^k$. Let $p_{n,k}(f)$ be the probability that the formula chosen from the distribution computes the function f. For every function f with $L(f) \leqslant \ell(k)^\alpha$, where $\alpha = \log_4(3/2)$, we have $p_{n,k}(f) > 0$. Moreover, for every function f, if $p_{n,k}(f) > 0$, then

$$(4n)^{-\ell(k)} \leqslant p_{n,k}(f) \leqslant c^{-\ell(k)^{1/4}},$$

where $c > 1$ is an absolute constant. Although the upper and lower bounds are exponentially small in $\ell(k)$, they are quasi-polynomially related whenever $\ell(k) \geqslant \ln^{\Omega(1)} n$. The construction is a step towards developing a model appropriate for investigation of the properties of a typical (random) Boolean function of some given complexity.

1. Introduction

Probabilistic methods appear to be very powerful in combinatorics and computer science. A natural point of view on these methods is that we investigate the properties of a typical object chosen from a set. One of the very first facts proved on Boolean functions is that a typical Boolean function chosen from the set of all functions has exponential complexity in any reasonable computation model. In particular, for Boolean formulas the result may be found in [5]. Hence, the properties of functions chosen uniformly from the set of all functions cannot say much about functions of moderate complexity.

In this situation, it is natural to ask what are the typical properties of functions chosen among the functions of some given complexity rather than among all functions. One

† This research was supported by GA CR, Grant No. 201/95/0976, and by Heinrich-Hertz-Stiftung while visiting Universität Dortmund, FB Informatik, LS2.

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

possibility for constructing a probabilistic distribution on functions of limited complexity is to describe the distribution in terms of their representations. In this case, it is easy to guarantee the complexity bound using only representations of an appropriate size. However, if the distribution is defined only in terms of syntactic properties of the representations, it may easily be the case that the distribution is concentrated on a small set of functions: for instance, on the two constant functions.

In the present paper a syntactically defined probabilistic model of Boolean formulas is described. The model is constructed by iterating the 4-ary Boolean operation $x_1x_2 \oplus x_3 \oplus x_4$ starting from a simple distribution on variables, their negations and the constants. After k iterations, the model generates a distribution on functions of the formula size complexity in the basis $\{\wedge, \oplus, 1\}$ bounded by $\ell(k) = 4^k$. The set of functions having nonzero probability contains all functions of complexity at most $\ell(k)^\alpha$, where $\alpha = \log_4(3/2)$. An upper bound on the probability of each of the functions with a positive probability is given. The upper bound is quasi-polynomially related to a trivial lower bound on this probability (Theorem 3.2). It follows that the distribution is not concentrated on any small set of functions.

The result is proved for a particular case of the model investigated in [6] and [7]. For this particular case, compared with the bounds from [6] and [7], much stronger bounds on the probability of single functions are obtained.

A similar model based on balanced formulas built up from the NAND operation (or equivalently from alternating levels of ANDs and ORs) and with randomly chosen literals was suggested by Friedman [2] in order to get information on Boolean complexity. Friedman suggested studying the distributions using their moments, and presented an application of this method to iterated AND, namely to random 1-SAT and random 2-SAT.

Formulas with a fixed tree of connectives and with the leaves assigned to variables or some other simple functions at random were also used for some other more specific purposes. Let us mention the construction of a monotone formula of size $O(n^{5.3})$ presented in [10] and the proof of existence of, for instance, Ramsey graphs on 2^n vertices, whose adjacency matrix is representable by a Boolean formula of polynomial size in n (see [4, 8]).

A different model of random Boolean formulas based on the uniform distribution on all AND/OR formulas of size tending to infinity was investigated in [3]. There it is proved that the distributions on functions obtained in this way converge to a limit distribution, in which the probability of every function f is positive and related to the complexity of f as follows. If $L'(f) \geqslant \Omega(n^3)$, then the probability $p(f)$ of f in the limit distribution satisfies

$$(8n)^{-L'(f)-2} \leqslant p(f) \leqslant c_1^{-L'(f)/n^3},$$

where $c_1 > 1$ is an absolute constant and $L'(f)$ is the formula size complexity of f in the basis $\{\wedge, \vee, \neg\}$. The existence of a limit distribution with all probabilities positive was investigated also for a more general model of random trees (see [12]).

The number $B(n, \ell)$ of distinct Boolean functions of n variables expressible by an AND/OR formula of size at most ℓ is estimated in [9]. For a wide range of the values of ℓ, matching lower and upper bounds on $B(n, \ell)$ are proved. Namely, if both $\alpha(n)$ and

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

$\beta(n)$ tend to infinity with n and $\alpha(n) \leqslant \ell \leqslant 2^n/n^{\beta(n)}$, then $B(n,\ell) = ((c_2 - o(1))n)^\ell$, where $c_2 = 2/(\ln 4 - 1)$.

2. The probability model

Let $n \geqslant 2$ be a fixed natural number throughout the paper. The Boolean functions of n variables are the functions $\{0,1\}^n \to \{0,1\}$. Since n is fixed, we call them simply Boolean functions. The projection functions are denoted x_i for $i = 1,\ldots,n$ as usual. The negation of x_i is denoted as $\bar{x}_i$. The conjunction is denoted like multiplication, that is, without any operation symbol. Recall that $\oplus$ is addition mod 2.

For any Boolean function u let $u^{-1}(1)$ be the set of $a \in \{0,1\}^n$ for which $u(a) = 1$. Moreover, let $|u| = |u^{-1}(1)|$. For arbitrary Boolean functions u, v let

$$\langle u, v\rangle = \bigoplus_{a\in\{0,1\}^n} u(a)v(a).$$

If $A \subseteq \{0,1\}^n$ and g is a Boolean function, we denote as $g|_A$ the restriction of g to the set A. Let X_A be the characteristic function of A.

For any nonconstant function f, let $L(f)$ be the formula size complexity of f in the basis $\{\wedge, \oplus, 1\}$, that is, the minimum number of occurrences of the variables in a formula expressing f in the given basis. Moreover, let $L(f) = 1$, if f is a constant function.

The probability distributions studied in the present paper are defined as follows.

Definition. Let $\tilde{g}_{n,0} \in \{0, 1, x_1, ..., x_n, \bar{x}_1, \bar{x}_2 ..., \bar{x}_n\}$ be a random Boolean function such that $\Pr(\tilde{g}_{n,0} = 0) = \Pr(\tilde{g}_{n,0} = 1) = 1/4$ and each of the literals occurs as $\tilde{g}_{n,0}$ with probability $1/(4n)$. For every $k \geqslant 0$ let $\tilde{g}_{n,k+1} = \tilde{g}_{n,k,1}\tilde{g}_{n,k,2} \oplus \tilde{g}_{n,k,3} \oplus \tilde{g}_{n,k,4}$, where $\tilde{g}_{n,k,j}$ are independent realizations of $\tilde{g}_{n,k}$. Finally, for every $k \geqslant 0$, let $p_{n,k}(f) = \Pr(f = \tilde{g}_{n,k})$.

For the purpose of the present paper, $\tilde{g}_{n,k}$ is defined to be a Boolean function. Clearly, the definition of this function implicitly describes a Boolean formula expressing $\tilde{g}_{n,k}$, which contains $\ell(k) = 4^k$ occurrences of variables, their negations and constants. Hence, $p_{n,k}(f) > 0$ implies $L(f) \leqslant \ell(k)$.

The distribution is chosen so that if a and $\bar{a}$ are complementary points in $\{0,1\}^n$, that is, they have Hamming distance n, then $\tilde{g}_{n,0}(a)$ and $\tilde{g}_{n,0}(\bar{a})$ are independent random variables. This simplifies the analysis of the distribution of $\tilde{g}_{n,k}$ for small k in the proof of Lemma 4.5.

For every f and k, if $p_{n,k}(f) > 0$, then also $p_{n,k+1}(f) > 0$, since, for instance, $f = 1 \cdot f \oplus 0 \oplus 0$ and the constants occur as $\tilde{g}_{n,k}$ with positive probability for all n and k. Moreover, every Boolean function f, that depends essentially only on variables $x_1, x_2, ..., x_i$, may be expressed as $f = x_i f_1 \oplus f_2 \oplus 0$, where f_1, f_2 do not depend on x_i. Extending this by induction and using the fact that $p_{n,0}(f) > 0$ for every f depending on at most one variable, one can prove that for every Boolean function f we have $p_{n,n-1}(f) > 0$.

By the well-known relationship between the size and the depth complexity of the Boolean formulas in any complete basis (see, for instance, [11]), one can prove that for any Boolean function f we have $p_{n,k}(f) > 0$ for some $k = O(\log L(f))$. In fact, we can

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

prove a better estimate using the method of balancing formulas in the basis $\{\wedge, \oplus\}$ from [1]. For the reader's convenience, we present the proof from [1] adapted to our model. For a comparison, recall that $p_{n,k}(f) > 0$ is possible only if $L(f) \leqslant \ell(k)$.

Theorem 2.1. *For every n and k and for every Boolean function f that satisfies $L(f) \leqslant \ell(k)^{\log_4(3/2)}$, we have $p_{n,k}(f) > 0$.*

Proof. Let $D(f)$ be the smallest k for which $p_{n,k}(f) > 0$. Recall that, for any function f, $L(f) \geqslant 1$. By induction on $L(f)$, an upper bound $D(f) \leqslant \log_{3/2} L(f)$ will be proved. This bound implies the theorem, since $L(f) \leqslant \ell(k)^{\log_4(3/2)}$ is equivalent to $\log_{3/2} L(f) \leqslant k$.

If $L(f) = 1$, then f is a constant, a variable or negation of a variable. Hence, $D(f) = 0$. Let $L(f) \geqslant 2$ and let the upper bound on $D(f)$ be true for all functions of complexity less than $L(f)$. We will find functions f_1, f_2 and f_3 such that $f = f_1 f_2 \oplus f_3$ and $L(f_j) \leqslant 2/3 \cdot L(f)$ for $j = 1, 2, 3$. Then, since $f = f_1 f_2 \oplus f_3 \oplus 0$, we have $D(f) \leqslant \max(D(f_1), D(f_2), D(f_3)) + 1$. By the induction hypothesis and the bound on $L(f_j)$, this implies $D(f) \leqslant \log_{3/2} L(f)$.

For any formula ϕ, let *size*(ϕ) denote the number of occurrences of variables in ϕ. In order to find f_1, f_2 and f_3, consider a formula ϕ expressing f satisfying size$(\phi) = L(f)$. This is possible, since f is a nonconstant function. Find the smallest subformula ϕ' of ϕ that satisfies size$(\phi') > 2/3 \cdot L(f)$. This subformula is either of the form $\alpha \oplus \beta$ or $\alpha\beta$, where α and β are subformulas of ϕ. It is easy to see that $1/3 \cdot L(f) \leqslant$ size$(\alpha) \leqslant 2/3 \cdot L(f)$. The same inequality also holds for size(β). Let $\psi(y)$ be the formula ϕ, where α is replaced by a new variable y. For every input, α evaluates to either 0 or 1. Hence, using the properties of $\oplus$, $\phi \equiv \psi(\alpha) \equiv \alpha \cdot (\psi(0) \oplus \psi(1)) \oplus \psi(0)$.

Now, let f_1, f_2 and f_3 be the functions expressed by the formulas α, $\psi(0) \oplus \psi(1)$ and $\psi(0)$, respectively. This choice immediately implies $f = f_1 f_2 \oplus f_3$. Since α contains at least one occurrence of a variable, we have $L(f_1) \leqslant$ size$(\alpha) \leqslant 2/3 \cdot L(f)$. Since the size of $\psi(0)$ is size$(\phi) -$ size$(\alpha) > 0$, we also have $L(f_3) \leqslant$ size$(\psi(0)) \leqslant 2/3 \cdot L(f)$.

It remains to show that $L(f_2) \leqslant 2/3 \cdot L(f)$. To this end, we first construct two sequences of formulas. Let $\psi_0(y)$ be the formula $\psi(y)$. This formula consists of two subformulas. Denote the two subformulas $\psi_1(y)$ and γ_1, where $\psi_0(y)$ is the one containing the unique occurrence of y. By repeating this decomposition and assuming without loss of generality that, in each step of the process, y is in the left-hand side subformula, we obtain sequences $\psi_0(y), \ldots, \psi_m(y)$ and $\gamma_1, \ldots, \gamma_m$ of subformulas of $\psi(y)$, such that $\psi_0(y) = \psi(y)$, $\psi_m(y) = y$ and, for all $i = 1, \ldots, m$, $\psi_{i-1}(y)$ is either $\psi_i(y) \oplus \gamma_i$ or $\psi_i(y)\gamma_i$. Let I be the set of those is for which $\psi_{i-1}(y) \equiv \psi_i(y)\gamma_i$.

Clearly, for any assignment of the values of the variables $x_1, x_2, \ldots, x_n$, a change of the value of y propagates to the root of the formula $\psi(y)$ if and only if the value of γ_i for all $i \in I$ is equal to 1 for the given assignment. Hence,

$$\psi(0) \oplus \psi(1) \equiv \bigwedge_{i \in I} \gamma_i.$$

Since γ_i are disjoint subformulas of $\psi(0)$, we have size$(\bigwedge_{i \in I} \gamma_i) \leqslant 2/3 \cdot L(f)$. Since $L(f) \geqslant 2$, this implies $L(f_2) \leqslant 2/3 \cdot L(f)$ even if f_2 is a constant function. □

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

Lemma 2.2. *For every $k \geqslant 0$ we have*

$$\max_f \Pr(\tilde{g}_k = f) \geqslant \max_f \Pr(\tilde{g}_{k+1} = f)$$

Proof. Let $\tilde{g}_{k+1} = \tilde{h}_1\tilde{h}_2 \oplus \tilde{h}_3 \oplus \tilde{h}_4$, where the $\tilde{h}_j$ are independent realizations of $\tilde{g}_k$. Because of the independence of the $\tilde{h}_j$, we obtain for any Boolean function f

$$\Pr(f = \tilde{g}_{k+1}) = \sum_{f_1} \Pr(f = \tilde{h}_1\tilde{h}_2 \oplus \tilde{h}_3 \oplus f_1) \cdot \Pr(\tilde{h}_4 = f_1).$$

By using the properties of $\oplus$, this is at most

$$\sum_{f_1} \Pr(f_1 = \tilde{h}_1\tilde{h}_2 \oplus \tilde{h}_3 \oplus f) \cdot \max_{f_2} \Pr(\tilde{h}_4 = f_2) = \max_{f_2} \Pr(\tilde{g}_k = f_2).$$

This finishes the proof of the lemma. □

3. The result

For the proof of the bounds on $p_{n,k}(f)$, we use the fact that the distribution of $\tilde{g}_k|_A$ for a fixed $A \subseteq \{0,1\}^n$ tends to the uniform distribution on the functions $A \to \{0,1\}$ when k tends to infinity. Moreover, we need an explicitly given estimate of the distance of the distribution of $\tilde{g}_k|_A$ from the uniform one depending on A and k. Such an estimate is given in the following theorem, which is proved in the next section. For any at least two-element subset A of the Boolean cube, let $\mu(A)$ be the minimum of $\rho(x,y)$, where ρ is the Hamming distance and $x, y \in A$ are distinct. If $|A| = 1$ then let $\mu(A) = n$. Moreover, for every nonzero function w, let $\mu(w)$ be defined as $\mu(w^{-1}(1))$.

Theorem 3.1. *There exists a constant $K \geqslant 0$ such that, for every n, every nonempty subset $A \subseteq \{0,1\}^n$, every $f : A \to \{0,1\}$ and every $k \geqslant k_0(A) = 2\log_2|A| + \log_2(n/\mu(A)) + K$, we have*

$$\left|\Pr(\tilde{g}_k|_A = f) - \frac{1}{2^{|A|}}\right| \leqslant \frac{1}{2^{|A|}\,2^{k-k_0(A)}}.$$

Note that the number of different functions that can appear as a realization of $\tilde{g}_{n,k}$ does not exceed $(2n+2)^{4^k}$. Hence, for $k < \log_4|A| - \log_4\log_2(2n+2)$ not every function $f : A \to \{0,1\}$ has a positive probability. This gives a lower bound on the values of $k_0(A)$ that satisfy the statement of Theorem 3.1. If $|A|$ is at least $n^{\Omega(1)}$, this lower bound and the value of $k_0(A)$ for which Theorem 3.1 is actually proved differ by at most a multiplicative constant.

The proof of Theorem 3.1 is given in Section 4. Now we will apply the theorem to derive the following bound.

Theorem 3.2. *There exists a constant $c > 1$ such that, for every sufficiently large n, every k that satisfies $0 \leqslant k \leqslant n/2$ and every Boolean function f, we have either $p_{n,k}(f) = 0$ or*

$$(4n)^{-\ell(k)} \leqslant p_{n,k}(f) \leqslant c^{-\ell(k)^{1/4}},$$

where $\ell(k) = 4^k$.

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

Proof. Clearly, every realization of $\tilde{g}_{n,0}$ has a probability at least $1/(4n)$. For every k, $\tilde{g}_{n,k}$ is a combination of $\ell(k)$ independent realizations of $\tilde{g}_{n,0}$. This implies the lower bound.

Let A_0 be any maximal subset of $\{0,1\}^n$ that satisfies $\mu(A_0) > n/6$. Then, every point of the Boolean cube is within Hamming distance at most $n/6$ from A_0. Hence,

$$|A_0| \sum_{j=0}^{\lfloor n/6 \rfloor} \binom{n}{j} \geqslant 2^n.$$

By using the estimate

$$\sum_{j=0}^{d} \binom{n}{j} \leqslant \left(\frac{n}{d}\right)^d \left(\frac{n}{n-d}\right)^{n-d} \leqslant \left(\frac{ne}{d}\right)^d$$

for $d = \lfloor n/6 \rfloor$ and using the fact that the estimate is even larger with $d = n/6$, we obtain

$$|A_0| \geqslant 2^n (6e)^{-n/6} \geqslant 2^{n/4}.$$

Let $K' = \log_2 6 + K$, where K is the constant from Theorem 3.1. Clearly, for every nonempty $A \subseteq A_0$, we have $K' \geqslant \log_2(n/\mu(A_0)) + K \geqslant \log_2(n/\mu(A)) + K$. The desired bound on $\Pr(\tilde{g}_k = f)$ will be proved separately for $0 \leqslant k \leqslant K' + 4$ and $K' + 4 \leqslant k \leqslant n/2$, starting with the latter range.

If $K' + 4 \leqslant k \leqslant n/2$, choose any subset $A \subseteq A_0$ such that

$$2^{(k-K')/2} - 1 \leqslant |A| \leqslant 2^{(k-K')/2}.$$

The upper bound on $|A|$ in this requirement implies $k \geqslant k_0(A)$, where $k_0(A)$ is the number from Theorem 3.1. Hence, for every Boolean function f, we have by Theorem 3.1

$$\Pr(\tilde{g}_k = f) \leqslant \Pr(\tilde{g}_k|_A = f|_A) \leqslant \frac{2}{2^{|A|}}.$$

Hence, assuming $\log_2 0 = -\infty$ and using the lower bound on $|A|$, we have

$$\log_2 \Pr(\tilde{g}_k = f) \leqslant 1 - |A| \leqslant 2 - 2^{(k-K')/2} \leqslant -2^{(k-K')/2-1}.$$

In other words,

$$\log_2 \Pr(\tilde{g}_k = f) \leqslant -2^{-K'/2-1} \ell(k)^{1/4}. \tag{3.1}$$

If $0 \leqslant k \leqslant K' + 4$, use Lemma 2.2 and the fact that, for every f, $\Pr(\tilde{g}_0 = f) \leqslant 1/4$ to derive

$$\log_2 \Pr(\tilde{g}_k = f) \leqslant -2 \leqslant -2^{(k-K')/2-1} = -2^{-K'/2-1} \ell(k)^{1/4}.$$

This, together with (3.1), implies the upper bound in the theorem, if we choose $c = 2^{2^{-K'/2-1}}$. □

4. Convergence to the uniform distribution

In this section we prove Theorem 3.1. To this end, we use the discrete Fourier transform of the distribution of $\tilde{g}_k$. For a random Boolean function with an arbitrary distribution, the discrete Fourier transform is defined as follows.

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

Definition. If w is a Boolean function and $\tilde{g}$ a random Boolean function, then let $\Delta(\tilde{g}, w) = \mathrm{E}(-1)^{\langle \tilde{g}, w\rangle}$.

One may easily see that $|\Delta(\tilde{g}, w)| \leqslant 1$ and $\Delta(\tilde{g}, 0) = 1$ holds for any $\tilde{g}$ and w. Moreover, it is easy to verify the following formula for the inverse of the Fourier transform.

Lemma 4.1. *For every $A \subseteq \{0,1\}^n$ and every $f : A \to \{0,1\}$, we have*

$$\Pr(\tilde{g}|_A = f) = \frac{1}{2^{|A|}} \sum_{w \leqslant X_A} \Delta(\tilde{g}, w)(-1)^{\langle f, w\rangle}. \qquad \square$$

Notice that, if A is a nonempty subset of $\{0,1\}^n$ and $\tilde{g}|_A$ has the uniform distribution on the functions $A \to \{0,1\}$, then $\Delta(\tilde{g}, X_A) = 0$. On the other hand, using Lemma 4.1, one can see that, if $\Delta(\tilde{g}, X_B) = 0$ for every nonempty subset $B \subseteq A$, then $\tilde{g}|_A$ is uniformly distributed on the functions $A \to \{0,1\}$.

In the following theorem, we express the Fourier coefficients of the distribution of a parity and of a conjunction of two independent random Boolean functions. Note that, for two random Boolean functions $\tilde{h}_1$ and $\tilde{h}_2$, $\Delta(\tilde{h}_1, \tilde{h}_2)$ is a random variable depending on the distribution of $\tilde{h}_1$ and on the actual value of $\tilde{h}_2$. It does not depend on the actual value of $\tilde{h}_1$. In the context of random Boolean formulas, the identity (4.1) for the parity was already used in [4].

Lemma 4.2. *Let $\tilde{h}_1, \tilde{h}_2$ be independent random Boolean functions. Then we have*

$$\Delta(\tilde{h}_1 \oplus \tilde{h}_2, w) = \Delta(\tilde{h}_1, w)\Delta(\tilde{h}_2, w) \tag{4.1}$$

$$\Delta(\tilde{h}_1\tilde{h}_2, w) = \mathrm{E}_{\tilde{h}_2}\Delta(\tilde{h}_1, \tilde{h}_2 w) \tag{4.2}$$

Proof. Owing to the independence of $\tilde{h}_1, \tilde{h}_2$, we have

$$\mathrm{E}(-1)^{\langle \tilde{h}_1 \oplus \tilde{h}_2, w\rangle} = \mathrm{E}(-1)^{\langle \tilde{h}_1, w\rangle}(-1)^{\langle \tilde{h}_2, w\rangle} = \mathrm{E}(-1)^{\langle \tilde{h}_1, w\rangle}\mathrm{E}(-1)^{\langle \tilde{h}_2, w\rangle}.$$

This proves (4.1). Since $\langle \tilde{h}_1\tilde{h}_2, w\rangle = \langle \tilde{h}_1, \tilde{h}_2 w\rangle$, by using the expansion to the conditional expectations, we obtain

$$\mathrm{E}(-1)^{\langle \tilde{h}_1\tilde{h}_2, w\rangle} = \sum_{v \leqslant w} \mathrm{E}\left((-1)^{\langle \tilde{h}_1, \tilde{h}_2 w\rangle} | \tilde{h}_2 w = v\right) \Pr(\tilde{h}_2 w = v).$$

Since $\tilde{h}_1$ and $\tilde{h}_2$ are independent, the distribution of $\tilde{h}_1$ under the condition $\tilde{h}_2 w = v$ is the same as the unconditional distribution of $\tilde{h}_1$. Hence, the conditional expectation in the sum above is equal to $\Delta(\tilde{h}_1, v)$. Hence,

$$\Delta(\tilde{h}_1\tilde{h}_2, w) = \sum_{v \leqslant w} \Delta(\tilde{h}_1, v)\Pr(\tilde{h}_2 w = v).$$

This implies (4.2). $\square$

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

For simplicity, let us use the abbreviation $\Delta_k(w) = \Delta(\tilde{g}_k, w)$. Let $\tilde{h}_j$ for $j = 1, 2, 3, 4$ be independent realizations of $\tilde{g}_k$. Then $\tilde{g}_{k+1} = \tilde{h}_1\tilde{h}_2 \oplus \tilde{h}_3 \oplus \tilde{h}_4$ and, by (4.1),

$$\Delta_{k+1}(w) = \Delta(\tilde{h}_1\tilde{h}_2, w)\Delta_k(w)^2. \tag{4.3}$$

In particular, since $|\Delta(\tilde{h}_1\tilde{h}_2, w)|$ is always at most one, we have

$$|\Delta_{k+1}(w)| \leqslant |\Delta_k(w)|^2. \tag{4.4}$$

This implies that, by increasing k, $|\Delta_k(w)|$ can be made arbitrarily small provided that it is initially strictly less than one.

In the proof of Theorem 3.1, we use a real number v that satisfies $1 < v < 3/2$. It is very natural to present the proof with such a general value of this parameter, although the theorem is finally proved by setting $v = \sqrt{2}$.

Theorem 3.1 is proved at the end of this section as a consequence of an upper bound on the Fourier coefficients of the distribution of $\tilde{g}_k$. The upper bound will have the form

$$|\Delta(\tilde{g}_k, w)| \leqslant q^{|w| - v^{k-r}m_0}$$

for all w, $0 \neq w \leqslant X_A$, and all $k \geqslant r$, where a real number $q > 1$ and integers r and m_0 are appropriately chosen. Extending an estimate in this form from any $k \geqslant r$ to $k+1$ instead of k is guaranteed by Lemma 4.6 on the assumption that m_0 is large enough. The number r for which the bound is true for the required m_0 and $k = r$ is found using Lemma 4.5. Let us start with two auxiliary statements.

Lemma 4.3. *Let $\tilde{h}$ be a random Boolean function. Let $q \geqslant 1$ and m be some real numbers, let A be a subset of the Boolean cube and let $|\Delta(\tilde{h}, u)| \leqslant q^{|u|-m}$ be satisfied for every nonzero u, $u \leqslant X_A$. Then, for every function $f : A \to \{0, 1\}$ we have*

$$\left|\Pr(\tilde{h}|_A = f) - \frac{1}{2^{|A|}}\right| \leqslant \left(\frac{q+1}{2}\right)^{|A|} q^{-m}.$$

Proof. Since $\Delta(\tilde{h}, 0) = 1$, we have by Lemma 4.1 that the LHS of the inequality in the lemma is at most

$$\frac{1}{2^{|A|}} \sum_{0 \neq u \leqslant X_A} |\Delta(\tilde{h}, u)| \leqslant \frac{1}{2^{|A|}} \sum_{u \leqslant X_A} q^{|u|-m} = \frac{1}{2^{|A|}}(1+q)^{|A|}q^{-m}. \qquad \square$$

Theorem 4.4. *Let $\tilde{h}$, q, m and A be as in Lemma 4.3. Let $|\Delta(\tilde{h}, u)| \leqslant q^{|u|-m}$ be satisfied for every nonzero u, $u \leqslant X_A$. Let $\tilde{h}_1, \tilde{h}_2$ be independent realizations of $\tilde{h}$. Then, for every Boolean function w satisfying $w \leqslant X_A$, we have*

$$|\Delta(\tilde{h}_1\tilde{h}_2, w)| \leqslant \left(\frac{1}{2}\right)^{|w|} + 2\left(\frac{q+1}{2}\right)^{|w|} q^{-m} + \left(\frac{q^2+1}{2}\right)^{|w|} q^{-2m}.$$

Proof. Let a Boolean function $w \leqslant X_A$ be given. First, let us prove the inequality

$$\mathrm{E}\left[q^{|\tilde{h}w|}\right] \leqslant \left(\frac{q+1}{2}\right)^{|w|} + \left(\frac{q^2+1}{2}\right)^{|w|} q^{-m}. \tag{4.5}$$

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

Let $v \leqslant X_A$ be any function and denote $B = v^{-1}(1)$. Note that the assumption of Lemma 4.3 is satisfied also with B instead of A. Since $v \leqslant \tilde{h}$ is equivalent to $\tilde{h}|_B \equiv 1$, Lemma 4.3 implies

$$\Pr(v \leqslant \tilde{h}) \leqslant \frac{1}{2^{|v|}} + \left(\frac{q+1}{2}\right)^{|v|} q^{-m}.$$

Hence, we have

$$\begin{aligned} \mathrm{E}\left[q^{|\tilde{h}w|}\right] &= \mathrm{E}\left[\sum_{v \leqslant \tilde{h}w} (q-1)^{|v|}\right] = \sum_{v \leqslant w} (q-1)^{|v|} \Pr(v \leqslant \tilde{h}) \\ &\leqslant \sum_{v \leqslant w} \left(\left(\frac{q-1}{2}\right)^{|v|} + \left(\frac{q^2-1}{2}\right)^{|v|} q^{-m}\right) \\ &= \left(1 + \frac{q-1}{2}\right)^{|w|} + \left(1 + \frac{q^2-1}{2}\right)^{|w|} q^{-m}. \end{aligned}$$

This proves (4.5).

By (4.2) we have $|\Delta(\tilde{h}_1\tilde{h}_2, w)| \leqslant \mathrm{E}_{\tilde{h}_2}|\Delta(\tilde{h}_1, \tilde{h}_2 w)|$. Using the bound $|\Delta(\tilde{h}_1, 0)| \leqslant 1$ and for $v \neq 0$ the bound $|\Delta(\tilde{h}_1, v)| \leqslant q^{|v|-m}$, we obtain

$$\begin{aligned} |\Delta(\tilde{h}_1\tilde{h}_2, w)| &\leqslant \sum_{v \leqslant w} |\Delta(\tilde{h}_1, v)| \cdot \Pr(\tilde{h}_2 w = v) \\ &\leqslant \Pr(\tilde{h}_2 w = 0) + \sum_{v \leqslant w} q^{|v|-m} \cdot \Pr(\tilde{h}_2 w = v) \\ &= \Pr(\tilde{h}_2 w = 0) + \mathrm{E}\left[q^{|\tilde{h}_2 w|-m}\right]. \end{aligned}$$

By Lemma 4.3 used for the set $A = w^{-1}(1)$, we have

$$\Pr(\tilde{h}_2 w = 0) = \Pr(\tilde{h}_2|_A \equiv 0) \leqslant \frac{1}{2^{|w|}} + \left(\frac{q+1}{2}\right)^{|w|} q^{-m}.$$

By (4.5),

$$\mathrm{E}\left[q^{|\tilde{h}_2 w|-m}\right] \leqslant \left(\frac{q+1}{2}\right)^{|w|} q^{-m} + \left(\frac{q^2+1}{2}\right)^{|w|} q^{-2m}.$$

By combining these two contributions, we obtain the theorem. □

Lemma 4.5. *For every real number q that satisfies $1 \leqslant q \leqslant 3$, every natural number $m \geqslant 3$, every nonzero Boolean function w and every integer $r \geqslant \log_2(n/\mu(w)) + 5(m-3) + 1$, we have $|\Delta_r(w)| \leqslant q^{|w|-m}$.*

Proof. Throughout the proof, we assume $q = 3$. Clearly, proving this case is sufficient, since the inequality is weaker if $q < 3$ and $|w| \leqslant m$. If $|w| \geqslant m$, the inequality is trivially satisfied.

Let us fix some nonzero function w. In order to prove the lemma, we prove a slightly stronger statement. Namely, we prove that, for every $m \geqslant 3$ and every integer $r \geqslant$

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

$\log_2(n/\mu(w)) + 5(m-3) + 1$, we have

$$|\Delta_r(v)| \leqslant q^{|v|-m} \tag{4.6}$$

for every nonzero v satisfying $v \leqslant w$.

The estimate is proved by induction on m. Let us start with $m = 3$. In this case, $|\Delta_r(v)| \leqslant q^{|v|-3}$ is trivially satisfied for every v satisfying $|v| \geqslant 3$ and every $r \geqslant 0$. For the cases $|v| = 1$ and $|v| = 2$ assume that $v \leqslant w$.

If $|v| = 1$, then $v = X_{\{a\}}$ for some $a \in \{0,1\}^n$. By definition,

$$\Delta_0(v) = \mathrm{E}(-1)^{\tilde{g}_{n,0}(a)} = \Pr(\tilde{g}_{n,0}(a) = 0) - \Pr(\tilde{g}_{n,0}(a) = 1) = 0.$$

Using (4.4), $|\Delta_r(v)| = 0 \leqslant q^{|v|-3}$ for all $r \geqslant 0$.

If $|v| = 2$, then $v = X_{\{a,b\}}$ for some distinct $a, b \in \{0,1\}^n$. By definition,

$$\Delta_0(v) = \mathrm{E}(-1)^{\tilde{g}_{n,0}(a) \oplus \tilde{g}_{n,0}(b)} = 1 - 2\Pr(\tilde{g}_{n,0}(a) \neq \tilde{g}_{n,0}(b)).$$

The event $\tilde{g}_{n,0}(a) \neq \tilde{g}_{n,0}(b)$ takes place if and only if $\tilde{g}_{n,0}$ is equal to x_i or $\bar{x}_i$ for some i such that $a_i \neq b_i$. Since each value $i = 1, \ldots, n$ appears with the probability $1/(2n)$, we have $\Delta_0(v) = 1 - \rho(a,b)/n$, where ρ denotes the Hamming distance. Since $v \leqslant w$, we have $\rho(a,b) = \mu(v) \geqslant \mu(w)$. Hence, $0 \leqslant \Delta_0(v) \leqslant 1 - \mu(w)/n$.

Summarizing this and using (4.4), we obtain

$$|\Delta_r(v)| \leqslant |\Delta_0(v)|^{2^r} \leqslant \left(1 - \mu(w)/n\right)^{2^r} \leqslant e^{-2^r \mu(w)/n}.$$

It follows that for every $r \geqslant \log_2(n/\mu(w)) + 1$ and every v satisfying $|v| = 2$ and $v \leqslant w$ we have

$$|\Delta_r(v)| \leqslant e^{-2} \leqslant q^{-1} \leqslant q^{|v|-3}.$$

Now, let $m \geqslant 4$ and $r \geqslant \log_2(n/\mu(w)) + 5(m-3) + 1$. It is sufficient to prove (4.6) for all v that also satisfy $1 \leqslant |v| \leqslant m-1$. The induction hypothesis says that, for all nonzero v, $v \leqslant w$ and every $s \geqslant \log_2(n/\mu(w)) + 5(m-4) + 1$, we have $|\Delta_s(v)| \leqslant q^{|v|-m+1}$.

If $1 \leqslant |v| \leqslant m-2$, then (4.4) and the induction hypothesis with $s = r-5$ imply $|\Delta_r(v)| \leqslant |\Delta_s(v)|^{32} \leqslant |\Delta_s(v)|^2 \leqslant q^{2(|v|-m+1)}$. This is at most $q^{|v|-m}$ by comparing the exponents. Let $|v| = m-1$. By (4.4) we obtain $|\Delta_r(v)| \leqslant |\Delta_{s+1}(v)|^{16}$. Moreover, using (4.3) and the induction hypothesis, together with Theorem 4.4 used with $m-1$ instead of m, and with $A = w^{-1}(1)$, we obtain

$$|\Delta_{s+1}(v)| \leqslant \left(\frac{1}{2}\right)^{m-1} + 2\left(\frac{q+1}{2q}\right)^{m-1} + \left(\frac{q^2+1}{2q^2}\right)^{m-1}.$$

By a routine calculation, one can verify that, for $q = 3$ and $m \geqslant 4$, this implies

$$|\Delta_r(v)| \leqslant |\Delta_{s+1}(v)|^{16} \leqslant q^{-1} = q^{|v|-m}.$$

This completes the proof of (4.6) and hence also of the lemma. □

Lemma 4.6. *For every ν, $1 < \nu < 3/2$, there exists a real number $q > 1$ and a natural number $m_0 \geqslant 1$ such that, for every $k \geqslant 0$, every real $m \geqslant m_0$ and every $A \subseteq \{0,1\}^n$, the following is true: if for every w, $0 \neq w \leqslant X_A$, we have $|\Delta_k(w)| \leqslant q^{|w|-m}$, then for every w, $0 \neq w \leqslant X_A$, we have $|\Delta_{k+1}(w)| \leqslant q^{|w|-\nu m}$.*

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

Proof. Let δ be such that $0 < \delta < \min(3/2 - v, 1/6)$. Moreover, let us prove that there is a number $q > 1$ satisfying

$$\frac{q+1}{2} < q^{1/2+\delta} \quad \text{and} \quad \frac{q^2+1}{2} < q^{1+2\delta}. \tag{4.7}$$

By taking the logarithm of both sides of both inequalities, the existence of such a number q follows from the facts that

$$\lim_{q\to 1+} \frac{\ln \frac{q^2+1}{2}}{\ln q^2} = \lim_{q\to 1+} \frac{\ln \frac{q+1}{2}}{\ln q} = \frac{1}{2} < \frac{1}{2} + \delta.$$

Let $q > 1$ be such that (4.7) and $q < 1.618$ is satisfied. Note that $1 + q > q^2$. Moreover, let m_0 be a large integer specified later according to v, δ and q. Let m be such that $m \geqslant m_0$. We are going to formulate conditions which imply

$$|\Delta_{k+1}(w)| \leqslant q^{|w|-vm} \tag{4.8}$$

for all w, $0 \neq w \leqslant X_A$, provided $|\Delta_k(w)| \leqslant q^{|w|-m}$ holds for all w, $0 \neq w \leqslant X_A$.

To this end, we consider three cases: $1 \leqslant |w| \leqslant m/2$, $m/2 \leqslant |w| \leqslant m$ and $m \leqslant |w| \leqslant vm$. In the remaining case, $|w| > vm$, (4.8) is trivially satisfied.

If $1 \leqslant |w| \leqslant m/2$, then $2(|w| - m) \leqslant |w| - vm$ and hence, using (4.4), we obtain $|\Delta_{k+1}(w)| \leqslant |\Delta_k(w)|^2 \leqslant q^{2(|w|-m)} \leqslant q^{|w|-vm}$.

In the two remaining cases we use the identity (4.3). To obtain a bound on the first factor of its RHS, we simplify the bound from Theorem 4.4 in the range of $|w|$ and q now considered. Since $q + 1 \geqslant q^2$ and $|w| \geqslant m/2$, we have $(q+1)^{|w|} q^{-m} \geqslant 1$. Hence, also using (4.7),

$$\begin{aligned} |\Delta(\tilde{h}_1\tilde{h}_2, w)| &\leqslant 3\left(\frac{q+1}{2}\right)^{|w|} q^{-m} + \left(\frac{q^2+1}{2}\right)^{|w|} q^{-2m} \\ &\leqslant 3q^{(1/2+\delta)|w|-m} + q^{(1+2\delta)|w|-2m}. \end{aligned} \tag{4.9}$$

In order to prove (4.8), we first derive a bound on the ratio of its LHS and RHS in both cases now considered. If $m/2 \leqslant |w| \leqslant m$, we use (4.3), (4.9) and $|\Delta_k(w)| \leqslant q^{|w|-m}$ to obtain

$$|\Delta_{k+1}(w)| q^{-|w|+vm} \leqslant 3q^{(3/2+\delta)|w|+(v-3)m} + q^{2(1+\delta)|w|+(v-4)m}.$$

The RHS of this is increasing in $|w|$. Hence, we obtain an upper bound by setting $|w| = m$. Thus

$$|\Delta_{k+1}(w)| q^{-|w|+vm} \leqslant 3q^{(v-3/2+\delta)m} + q^{(v-2+2\delta)m}. \tag{4.10}$$

If $m \leqslant |w| \leqslant vm$, we only know $|\Delta_k(w)| \leqslant 1$. Hence, by (4.3) and (4.9) we get

$$|\Delta_{k+1}(w)| q^{-|w|+vm} \leqslant 3q^{(-1/2+\delta)|w|+(v-1)m} + q^{2\delta|w|+(v-2)m}.$$

We derive an upper bound on the RHS of this inequality by substituting an appropriate value of $|w|$ in each of its two terms. In the first one we substitute the smallest value of the range ($|w| = m$) and in the second one the largest value ($|w| = vm$). Hence, we obtain

$$|\Delta_{k+1}(w)| q^{-|w|+vm} \leqslant 3q^{(v-3/2+\delta)m} + q^{(v-2+2\delta v)m}. \tag{4.11}$$

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

Because of our choice of δ, the RHS of both (4.10) and (4.11) converge to zero, if $m \to \infty$. Since v, δ and q are now fixed, it is possible to take a natural number m_0 large enough to guarantee that (4.10) and (4.11) are both at most 1 and, hence, (4.8) is satisfied. This completes the proof of the lemma. □

Finally, here is the convergence result that we have been aiming toward.

Proof of Theorem 3.1. Let v, $1 < v < 3/2$, and a nonempty subset $A \subseteq \{0,1\}^n$ be given. Let q and m_0 be some numbers for which the conclusion of Lemma 4.6 holds. As a basis for an iterative use of Lemma 4.6, we need a number r such that for all w, $0 \neq w \leqslant X_A$, the inequality

$$|\Delta_r(w)| \leqslant q^{|w|-m_0}$$

is satisfied. Lemma 4.5 guarantees that this is true for some $r = \log_2(n/\mu(A)) + O(1)$. For every $s \geqslant 0$, by using the inequality from Lemma 4.6 s times, we obtain

$$|\Delta_{r+s}(w)| \leqslant q^{|w|-v^s m_0}$$

for every w, $0 \neq w \leqslant X_A$. Since $m_0 \geqslant 1$, assuming $s = \lceil \log_v |A| + \log_v 2 \rceil$, we obtain

$$|\Delta_{r+s}(w)| \leqslant q^{|w|-2|A|}$$

for every w, $0 \neq w \leqslant X_A$. Let t be such that $q^{2^t} \geqslant 2$ and $u \geqslant 0$. By using (4.4) $t+u$ times, we obtain

$$|\Delta_{r+s+t+u}(w)| \leqslant \left(q^{2^{t+u}}\right)^{|w|-2|A|} \leqslant \left(2^{2^u}\right)^{|w|-2|A|}$$

for every w, $0 \neq w \leqslant X_A$. Let $k_0(A) = r + s + t$, $k \geqslant k_0(A)$ and let $u = k - k_0(A)$. By Lemma 4.3 used with $q = 2^{2^u}$ and $m = 2|A|$, we have

$$\left|\Pr(\tilde{g}_k|_A = f) - \frac{1}{2^{|A|}}\right| \leqslant \frac{1}{2^{|A|}} \left(1 + 2^{2^u}\right)^{|A|} \left(2^{2^u}\right)^{-2|A|} \leqslant \left(2^{-2^u}\right)^{|A|}.$$

Since $r = \log_2(n/\mu(A)) + O(1)$, $s = \log_v |A| + O(1)$ and $t = O(1)$, Theorem 3.1 is proved, if we assume $v = \sqrt{2}$. □

References

[1] Buss, S. R. and Bonet, M. L. (1994) Size–depth tradeoffs for Boolean formulae. *IPL* **11** 151–155.

[2] Friedman, J. (1992) Probabilistic spaces of Boolean functions of a given complexity: generalities and random k-SAT coefficients. Research Report CS-TR-387-92, Princeton University.

[3] Lefmann, H. and Savický, P. (1997) Some typical properties of large AND/OR Boolean formulas. *Random Struct. Alg.* **10** 337–351.

[4] Razborov, A. A. (1988) Bounded-depth formulae over $\{\wedge, \oplus\}$ and some combinatorial problems. In *Complexity of Algorithms and Applied Mathematical Logic* (in Russian), Ser. *Voprosy Kibernetiky* (Problems in Cybernetics), (S. I. Adian, ed.), Moscow, pp. 149–166.

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

[6] Savický, P. (1990) Random Boolean formulas representing every Boolean function with asymptotically equal probability. *Discrete Math.* **83** 95–103.

[7] Savický, P. (1995) Bent functions and random Boolean formulas. *Discrete Math.* **147** 211–234.

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press

[8] Savický, P. (1995) Improved Boolean formulas for the Ramsey graphs. *Random Struct. Alg.* **6** 407–415.
[9] Savický, P. and Woods, A. R. The number of Boolean functions computed by formulas of a given size. To appear in *Random Struct. Alg.*
[10] Valiant, L. G. (1984) Short monotone formulae for the majority function. *J. Algorithms* **5** 363–366.
[11] Wegener, I. (1987) *The Complexity of Boolean Functions*, Wiley-Teubner Series in Computer Science.
[12] Woods, A. R. (1997) Coloring rules for finite trees, and probabilities of monadic second order sentences. *Random Struct. Alg.* **10** 453–485.

https://doi.org/10.1017/S0963548398003630 Published online by Cambridge University Press