dnf.rev.tex

The Shortest Disjunctive Normal Form of a Random Boolean Function

Nicholas Pippenger*
(nicholas@cs.ubc.ca)

Department of Computer Science
The University of British Columbia
Vancouver, British Columbia V6T 1Z4
CANADA

Abstract: This paper gives a new upper bound for the average length $\overline{\ell}(n)$ of the shortest disjunctive normal form for a random Boolean function of n arguments, as well as new proofs of two old results related to this quantity. We consider a random Boolean function of n arguments to be uniformly distributed over all 2^{2^n} such functions. (This is equivalent to considering each entry in the truth-table to be 0 or 1 independently and with equal probabilities.) We measure the length of a disjunctive normal form by the number of terms. (Measuring it by the number of literals would simply introduce a factor of n into all our asymptotic results.) We give a short proof using martingales of Nigmatullin's result that almost all Boolean functions have the length of their shortest disjunctive normal form asymptotic to the average length $\overline{\ell}(n)$. We also give a short information-theoretic proof of Kuznetsov's lower bound $\overline{\ell}(n) \geq \big(1+o(1)\big)\ 2^n/\log n\ \log\log n$. (Here log denotes the logarithm to base 2.) Our main result is a new upper bound $\overline{\ell}(n) \leq \big(1+o(1)\big)\ H(n)\ 2^n/\log n\ \log\log n$, where $H(n)$ is a function that oscillates between $1.38826\ldots$ and $1.54169\ldots$. The best previous upper bound, due to Korshunov, had a similar form, but with a function oscillating between $1.581411\ldots$ and $2.621132\ldots$. The main ideas in our new bound are (1) the use of Rödl's "nibble" technique for solving packing and covering problems, (2) the use of correlation inequalities due to Harris and Janson to bound the effects of weakly dependent random variables, and (3) the solution of an optimization problem that determines the sizes of "nibbles" and larger "bites" to be taken at various stages of the construction.

* The work reported here was supported by a Canada Research Chair and an NSERC Research Grant.

1. Introduction

A *Boolean function* (of n arguments) is a map $\mathbf{B}^n \to \mathbf{B}$, where $\mathbf{B} = \{0, 1\}$ is the set of Boolean values. A Boolean function can be expressed (usually in many different ways) in *disjunctive normal form*; that is, as the disjunction (logical OR) of zero or more *terms*, each of which is the conjunction (logical AND) of zero or more *literals*, each of which is either an argument or the complement (logical NOT) of an argument. (Disjunctive normal form is also referred to as "alternative" normal form, or as a "sum-of-products" expansion.) Let $\ell(f)$ denote the minimum possible number of terms in a disjunctive normal form for f, and let $L(f)$ denote the minimum possible number of literals in a disjunctive normal form for f. Algorithms for finding minimal expressions in either of these senses have been given by Quine [Q1, Q2] and McCluskey [M], who identified the central problem as that of covering the set $U_f = f^{-1}(1)$ of 1s in the truth-table of f with *implicants* (that is, with subcubes of the n-cube $\mathbf{B}^n$ all of whose vertices correspond to 1s in the truth-table of f). All known algorithms for these minimization problems are plagued by the need to consider exponentially many cases. (The word "exponentially" here means exponentially in a fractional power of 2^n, which is the size of the input to such a minimization problem when the function is specified by a truth-table.) Indeed, Glagolev [G] has shown that, for every $\varepsilon > 0$ and all sufficiently large n, there are at least $2^{2^{(1-\varepsilon)n}}$ "terminal" forms, which are local minima for the minimization problem.

A Boolean function can also be expressed in *conjunctive normal form* (also referred to as a "product-of-sums" expansion); that is, as the conjunction of zero or more *clauses*, each of which is the disjunction of zero or more literals. A conjunctive normal form for f can be obtained by exchanging conjunctions and disjunctions in a disjunctive normal form for the *dual* function f^* of f, given by $f^*(x_1, \ldots, x_n) = \overline{f(\overline{x_1}, \ldots, \overline{x_n})}$. Let $\ell^*(f)$ denote the minimum possible number of clauses in a conjunctive normal form for f, and let $L^*(f)$ denote the minimum possible number of literals in a conjunctive normal form for f. Then we have $\ell^*(f) = \ell(f^*)$ and $L^*(f) = L(f^*)$. This correspondence allows us to confine our attention without loss of generality to disjunctive normal forms.

Let $\ell(n)$ and $L(n)$ denote the maxima of $\ell(f)$ and $L(f)$, respectively, over all Boolean functions f of n arguments. It is easy to see that $\ell(n) = 2^{n-1}$ and $L(n) = n2^{n-1}$. To see this, not that any Boolean function of n arguments can be decomposed, according to the values of the first $n-1$ arguments, into 2^{n-1} subfunctions of the remaining argument, and each of these subfunctions contributes at most one term and at most n literals to a disjunctive normal form. Thus the expressions given are upper bounds. To see that they are also lower bounds, we note that they are achieved by the parity function, $f(x_1, \ldots, x_n) =$

$x_1 \oplus \cdots \oplus x_n$. For this function, U_f consists of 2^{n-1} points, and the expressions stated are achieved as upper bounds by allocating a separate term to each of these points. To see that these expressions are lower bounds, we observe that no two points of U_f are adjacent in the n-cube $\mathbf{B}^n$, so no term can cover more than one point, and no term can have fewer than n literals.

In order to discuss random Boolean functions, we shall establish a probability distribution on the Boolean functions of n arguments by taking all such functions to be equally likely (that is, by assigning probability $1/2^{2^n}$ to each such function, or equivalently by considering each entry in the truth-table of the function to be independently equally likely to be 0 or 1). In order to make asymptotic statements, we shall consider a sequence $f_1, f_2, \ldots$ of random Boolean functions, where f_n is a function of n arguments distributed as described above. When we say that a statement about a function holds for *almost all* Boolean functions, we shall mean that the probability that the statement holds for f_n tends to 1 as $n \to \infty$. When we make an asymptotic statement such as $\phi(f) \sim \psi(f)$ for almost all Boolean functions, we shall mean that, for every $\varepsilon > 0$, $(1-\varepsilon)\,\psi(f) \leq \phi(f) \leq (1+\varepsilon)\,\psi(f)$ holds for almost all Boolean functions.

It is an easy observation (the origin of which we have not succeeded in identifying) that $L(f) \sim n\,\ell(f)$ for almost all Boolean functions. Indeed, we have $L(f) \leq n\,\ell(f)$ for all Boolean functions. It is easy to show that, if $k = \lceil 2 \log n \rceil$, then almost all Boolean functions f have no subcube of dimension as large as k in U_f. (There are $\binom{n}{k} 2^{n-k} = 2^{n+O((\log n)^2)}$ subcubes of dimension k in $\mathbf{B}^n$, and the probability that any one of them appears in U_f is $2^{-2^k} \leq 2^{-n^2}$. Therefore the probability that some such subcube appears in U_f is at most $2^{-n^2+O(n)}$.) Thus almost all Boolean functions can have no term with as few as $n-k$ literals in a disjunctive normal form, so we have $L(f) \geq (n-k)\,\ell(f) \sim n\,\ell(f)$ for almost all Boolean functions. This fact allows us to transfer probabilistic assertions about random Boolean functions between $\ell(f)$ and $L(f)$ by adding or deleting a factor of n.

Let $\overline{\ell}(n)$ and $\overline{L}(n)$ denote the averages of $\ell(f)$ and $L(f)$, respectively, over all Boolean functions f of n arguments. It is a peculiarity of disjunctive normal form that $\overline{\ell}(n)$ and $\overline{L}(n)$ grow less rapidly than $\ell(n)$ and $L(n)$. We shall see below that $\overline{\ell}(n) = O(2^n/\log n\ \log\log n)$ and $\overline{L}(n) = O(n\ 2^n/\log n\ \ \log\log n)$. By contrast, if we consider $\overline{L}_3(n)$ and $L_3(n)$, the average and maximum number of literals in the shortest three level (for example, sum-of-products-of-sums) formula for a Boolean function of n arguments, then we have $\overline{L}_3(n) \sim L_3(n) \sim 2^n/\log n$. Indeed, the lower bound $L_3(f) \geq (1-\varepsilon)2^n/\log n$ for almost all Boolean functions was established by Riordan and Shannon [R1] in 1942, while the upper bound

$L_3(f) \leq (1+\varepsilon)2^n/\log n$ for all Boolean functions was established by Lupanov [L] in 1961. Most complexity measures for Boolean functions, such as those defined by other kinds of formulas, or by circuits rather than formulas, follow the same pattern as formulas of depth 3, with the average (and in fact the value for almost all functions) being asymptotic to the maximum (and with it being an unsolved problem to identify an explicit function achieving the maximum). Formulas of depth two appear as an isolated exception, with the average growing more slowly than the maximum (and with the functions achieving the maximum being obvious).

In 1967, Nigmatullin [N2] showed that $\ell(f) \sim \overline{\ell}(n)$ for almost all Boolean functions of n arguments. In Section 2, we shall present a new proof of Nigmatullin's result. This new proof uses Azuma's inequality [A] on martingales, which has become a widely used tool for establishing concentration phenomena (see Spencer [S2] for examples and a short proof), in place of a difficult isoperimetric inequality used by Nigmatullin [N1].

In 1967, Glagolev [G] gave the lower bound

$$\overline{\ell}(n) \geq \frac{2^n}{2\log n\,\log\log n}\bigl(1+o(1)\bigr),$$

and the upper bound

$$\overline{\ell}(n) \leq \frac{2^n\,\ln\log n}{\log n}\bigl(1+o(1)\bigr),$$

where $\ln x$ denotes the natural logarithm of x. In 1969, Korshunov [K1] improved the upper bound to

$$\overline{\ell}(n) \leq \frac{4\cdot 2^n}{\log n}\bigl(1+o(1)\bigr),$$

and in 1972, Sapozhenko [S1] improved it further to

$$\overline{\ell}(n) \leq \frac{2^n}{\ln n}\bigl(1+o(1)\bigr).$$

In 1980, Kuznetsov [K4] improved the lower bound to

$$\overline{\ell}(n) \geq \frac{2^n}{\log n\,\log\log n}\bigl(1+o(1)\bigr).$$

In Section 3, we shall present a new proof of Kuznetsov's lower bound. Kuznetsov's proof uses a counting argument. Our information-theoretic proof uses entropy, and thus is in a sense merely a recasting of Kuznetsov's, but is nevertheless somewhat simpler.

In 1981, Korshunov [K2] finally established the order of growth of $\overline{\ell}(n)$ within constant factors by giving the upper bound

$$\overline{\ell}(n) \leq \frac{F(n)\, 2^n}{\log n\, \log\log n}\big(1+o(1)\big).$$

To describe $F(n)$ (and other functions that will be introduced below), we use the generating function

$$P(x) = \sum_{k\geq 0} x^{2^k}$$

for the integral powers of 2, and its derivative

$$P'(x) = \sum_{k\geq 0} 2^k\, x^{2^k-1}.$$

We note that $P'(\frac{1}{2}) = 2.562988\ldots$ and $P'(\frac{1}{4}) = P'(\frac{1}{2}) - 1 = 1.562988\ldots$. We then have

$$F(n) = 2^{\vartheta-1}\,\big(2+P'(\tfrac{1}{2})\big)\,\ln 2,$$

where

$$\vartheta = \{\log\log n + \log\log\log n\}$$

and $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part of x. The function $F(n)$ varies between a minimum of $\frac{1}{2}\big(2+P'(\frac{1}{2})\big)\,\ln 2 = 1.581411\ldots$ for $\vartheta = 0$ and a supremum of $\big(2+P'(\frac{1}{2})\big)\,\ln 2 = 3.162822\ldots$ for $\vartheta \to 1$. In 1983, Korshunov [K3] improved this upper bound to

$$\overline{\ell}(n) \leq \frac{G(n)\, 2^n}{\log n\, \log\log n}\big(1+o(1)\big),$$

where

$$G(n) = \Big(2^{2\vartheta-1} + 2^{\vartheta-2^\vartheta}\,\big(2+P'(\tfrac{1}{2})\big)\Big)\,\ln 2.$$

The function $G(n)$ varies between the same minimum of $\frac{1}{2}\big(2+P'(\frac{1}{2})\big)\,\ln 2 = 1.581411\ldots$ for $\vartheta = 0$ and a supremum of $\frac{1}{2}\big(5+P'(\frac{1}{2})\big)\,\ln 2 = 2.621132\ldots$ for $\vartheta \to 1$.

In Section 4 we shall present a new upper bound,

$$\overline{\ell}(n) \leq \frac{H(n)\, 2^n}{\log n\, \log\log n}\big(1+o(1)\big),$$

where

$$\begin{aligned} H(n) &= \Big(2^{\vartheta-1} + 2^{\vartheta-2^\vartheta}\; P'\left(2^{-2^\vartheta}\right)\,\ln 2\Big) \\ &= 2^\vartheta \left(\frac{1}{2} + \left(\sum_{k\geq 0} 2^k 2^{-2^k 2^\vartheta}\right) \ln 2\right). \end{aligned}$$

The function $H(n)$ varies between a minimum of $\frac{1}{2}\left(1 + P'(\frac{1}{2})\ln 2\right) = 1.38826\ldots$ for $\vartheta = 0$ and a supremum of $\left(1 + \frac{1}{2}P'(\frac{1}{4})\ln 2\right) = 1.54169\ldots$ for $\vartheta \to 1$. We note that the supremum of $H(n)$ is smaller than the minimum of $F(n)$ and $G(n)$. The main idea leading to our improved upper bound is the use of "nibbles" rather than "bites" in the covering problem. (This technique for solving packing and covering problems was introduced by Rödl [R2], and has since been developed by several other authors; see Frankl and Rödl [F2] and Pippenger and Spencer [P] for examples.) Our proof also uses two new ideas to simplify the analysis. First we use disjoint sets of directions for each nibble or bite, increasing the independence among various random variables in the proof. Second, we use correlation inequalities (Harris's inequality [H] (which is a special case of the FKG inequality of Fortuin, Kasteleyn and Ginibre [F1]), and Janson's inequality [J1] (see also Boppana and Spencer [B]) to bound the interactions of weakly dependent random variables. Finally, we use the solution of an optimization problem to determine the sizes of the nibbles and bites to be taken at various stages of the construction.

The following metaphorical version of the optimization problem is not exactly equivalent to it, but should convey the essential idea. Suppose that you wish to drive a long distance along a road at which fuel stations are closely spaced. If fuel sells at the same price at all stations, you will minimize your cost by buying at each station just enough fuel to get you to the next station, for buying more wastes money transporting fuel. This correspond to taking nibbles. If from time to time you encounter a sign saying that henceforth the cost of fuel will be doubled, it will be profitable for you to take on more fuel at the last station before the doubling. The amount of fuel taken on must be chosen carefully; too little will have you buying expensive fuel too soon; but too much will again waste money by transporting fuel. This corresponds to taking bites. Thus the solution to the optimization problem consists almost entirely of a long sequence of nibbles, inturrupted by occasional bites. (There will also be an extra large bite (a "chomp") at the end, followed by a final stage that "finishes up the crumbs".)

Before proceeding to the proofs of these results, we feel obliged to explicitly disclaim any practical significance for our new upper bound. Indeed, the error factor $1 + o(1)$ is actually $1+O(\log\log\log\log n/\log\log n)$. To have $\log\log\log\log n/\log\log n \leq 1/4$, we must have $n \geq 2^{16} = 65536$. The size of truth-table of the function is then at least 2^{65536}, and the number of literals in the shortest disjunctive normal form is even larger. The error factor can be improved, with the consequence that the impracticality of the result is less spectacular, but in truth none of the asymptotic results in this area are suited to functions with fewer than about 20 arguments, and "random" functions of this size are at present invariably computed simply by table look-up in a read-only memory.

2. Concentration

In this section we shall give a simple proof of the following result.

Theorem 2.1: (R. G. Nigmatullin [N2]) For almost all Boolean functions f of n arguments,

$$\ell(f) \sim \overline{\ell}(n).$$

Proof: Let $k = \lceil 2 \log n \rceil$, and let $\ell^+(f)$ denote the length of the shortest disjunctive normal form for f in which no term has as few as $n-k$ literals. Clearly $\ell^+(f) \geq \ell(f)$. Furthermore, we have $\ell^+(f) = \ell(f)$ unless U_f contains a subcube of dimension k. There are $\binom{n}{k} 2^{n-k} = 2^{n+O((\log n)^2)}$ subcubes of dimension k in $\mathbf{B}^n$, and the probability that any one of them appears in U_f is at most $2^{-2^k} \leq 2^{-n^2}$. The probability that some such subcube appears in U_f is therefore at most $p = 2^{-n^2+O(n)}$. Thus we have

$$\ell^+(f) = \ell(f) \tag{2.1}$$

for almost all Boolean functions f.

Let $\overline{\ell^+}(n)$ denote the average of $\ell^+(f)$ over all Boolean functions of n arguments. For any f, $\ell^+(f) - \ell(f)$ is at most the number of points in U_f that are contained in subcubes of dimension k in U_f. The expected number of such points is at most $2^k p = 2^{-n^2+O(n)}$. Thus we have

$$\overline{\ell^+}(n) \sim \overline{\ell}(n).$$

Combining this with (2.1), we see that to prove Theorem 2.1 it will suffice to show that

$$\ell^+(f) \sim \overline{\ell^+}(n), \tag{2.2}$$

for almost all Boolean functions f of n arguments.

A sequence $X_0, X_1, \ldots, X_N$ of random variables is called a *martingale* if

$$\mathrm{Ex}[X_{i+1} \mid X_i] = X_i$$

for $0 \leq i \leq N-1$. Azuma's inequality [A] (see Spencer [S2] for a short proof) states that if $0 = X_0, X_1, \ldots, X_N$ is a martingale, and if $|X_{i+1} - X_i| \leq C$ for $0 \leq i \leq N-1$, then

$$\Pr\left[\, |X_N| > \lambda \,\right] \leq 2 \exp(-\lambda^2 / 2C^2 N).$$

To use Azuma's inequality, we define a martingale as follows. Let $N = 2^n$, and let $v_1, \ldots, v_n$ denote the vertices of cube $\mathbf{B}^n$ (in some arbitrary order). Let $X_i = Y_i - \overline{\ell^+}(n)$, where Y_i is the expectation of $\ell^+(f)$ conditioned on the values of f at the points $v_1, \ldots, v_i$. Then $X_0 = 0$ and $X_N = \ell^+(f) - \overline{\ell^+}(n)$. We may take for C an upper bound to $|\ell^+(f) - \ell^+(g)|$ over all functions f and g that differ in just one entry of their truth-tables. If we change a 0 to a 1 in the truth-table, we need add at most one term to the disjunctive normal form. If we change a 1 to a 0, we may need to add more terms, but we can bound the number as follows. The changed point lies in at most $\binom{n}{k} \leq n^{2\log n + 2}$ terms of a shortest disjunctive normal form. We can omit these terms, replacing each of them by at most k new terms that cover the points, other than the changed one, covered by the old term. Thus we need to add at most $(k-1)n^{2\log n+2} \leq n^{2\log n+3}$ terms in this case.

We can now apply Azuma's inequality with $C = n^{2\log n+3}$ and $\lambda = 2^{2n/3}$, and conclude that

$$\Pr\left[|\ell^+(f) - \overline{\ell^+}| > 2^{2n/3}\right] \leq 2\exp\left(-2^{n/3+O\left((\log n)^2\right)}\right),$$

which establishes (2.2) and thus completes the proof of Theorem 2.1. $\triangle$

3. Lower Bound

In this section we shall give a simple proof of the following result.

Theorem 3.1: (S. E. Kuznetsov [K4]) For $n \to \infty$,

$$\overline{\ell}(n) \geq \frac{2^n}{\log n \, \log\log n}\left(1 + O\left(\frac{\log\log\log n}{\log\log n}\right)\right).$$

Proof: Let f be a random Boolean function of n arguments, and let Φ be some disjunctive normal form for f having $\ell(f)$ terms. Let A denote the set of 3^n terms, each of which is a conjunction in which each argument can appear in direct form, in complemented form, or not at all. For each $w \in A$, let the random variable ϕ_w assume the value 1 if the term w appears in Φ, and the value 0 otherwise. Let p_w denote the expectation of ϕ_w (that is, the probability that ϕ_w assumes the value 1). Then we have

$$\overline{\ell}(n) \geq \sum_{w \in A} p_w. \tag{3.1}$$

If X is a random variable that assumes values $1, \ldots, N$ with probabilities $q_1, \ldots, q_N$, respectively, the *entropy* $H(X)$ of X is defined by

$$H(X) = \sum_{1 \leq M \leq N} \eta(q_M),$$

where

$$\eta(q) = -q \log q.$$

In particular, if X assumes just two values, with probabilities q and $1-q$, then

$$H(X) = h(q),$$

where

$$h(q) = -q \log q - (1-q) \log(1-q).$$

The random variable f assumes 2^{2^n} values, each with probability $1/2^{2^n}$, so $H(f) = 2^n$. Since entropy is non-decreasing (that is, since $H(X,Y) \geq H(X)$; see for example Jelinek [J2], Lemmas 4.12 and 4.13), and since $\phi = \{\phi_w\}_{w \in A}$ determines f, we have $H(\phi) \geq H(f) = 2^n$. Since entropy is subadditive (that is, $H(X,Y) \leq H(X) + H(Y)$; see for example Jelinek [J2], Lemma 4.14), we have $H(\phi) \leq \sum_{w \in A} H(\phi_w) = \sum_{w \in A} h(p_w)$. Combining these results gives

$$\sum_{w \in A} h(p_w) \geq 2^n. \tag{3.2}$$

For each term $w \in A$, let $|w|$ denote the number of variables appearing in w (so that $n - |w|$ is the number of literals not appearing in w). Since a term w omitting k literals can appear in Φ only if U_f contains the subcube of dimension k corresponding to w, and this event occurs with probability 2^{-2^k}, we have

$$0 \leq p_w \leq 2^{-2^{n-|w|}}. \tag{3.3}$$

Let $P = \#A$ denote the number of elements in A. Since $h(q)$ is a concave function of q, we have

$$\sum_{w \in A} h(p_w) \leq P \; h\left(\frac{1}{P} \sum_{w \in A} p_w\right). \tag{3.4}$$

Since (3.3) implies that $0 \leq p_w \leq 1/2$, and since $h(q)$ is an increasing function of q for $0 \leq q \leq 1/2$, (3.1) yields

$$P \; h\left(\frac{1}{P} \sum_{w \in A} p_w\right) \leq P \; h\left(\frac{\overline{\ell}(n)}{P}\right). \tag{3.5}$$

Since $-(1-q)\log(1-q) \leq q \log e$ (the graph of the concave function $-(1-q)\log(1-q)$ lies below that of $q \log e$, its tangent at $q = 0$), we obtain

$$P \; h\left(\frac{\overline{\ell}(n)}{P}\right) \leq \overline{\ell}(n) \; \log\left(\frac{eP}{\overline{\ell}(n)}\right). \tag{3.6}$$

Combining (3.2), (3.4), (3.5) and (3.6) yields

$$\overline{\ell}(n) \geq \frac{2^n}{\log\left(\frac{eP}{\overline{\ell}(n)}\right)}. \tag{3.7}$$

Substituting $P = 3^n$ and the trivial lower bound $\overline{\ell}(n) \geq 1$ into the right-hand side of (3.7) gives the preliminary lower bound

$$\overline{\ell}(n) = \Omega\left(\frac{2^n}{n}\right). \tag{3.8}$$

Let $k = \lceil \log\log n + \log\log\log n \rceil + 1$. Write $A = B \cup C$, where B contains the terms containing more than $n - k$ literals (corresponding to subcubes of dimension less than k), and C contains the terms containing at most $n - k$ literals (corresponding to subcubes of dimension at least k). The inequalities (3.1) and (3.2) will remain valid if for $w \in C$ we assign the terms in (3.1) their smallest possible values (according to (3.3)), and assign them in (3.2) their largest possible values. This yields

$$\overline{\ell}(n) \geq \sum_{w \in B} p_w \tag{3.9}$$

and

$$\sum_{w \in B} h(q_w) + \sum_{w \in C} h(2^{-2^{n-|w|}}) \geq 2^n. \tag{3.10}$$

To bound the last sum in (3.10), we note that C contains $\binom{n}{j} 2^{n-j}$ terms w with $n - |w| = j$. Furthermore, for $0 \leq q \leq 1/2$ we have $\eta(q) \leq \eta(1-q)$, so that $h(q) \leq 2\eta(q) = -2q \log q$. Thus we have

$$\sum_{w \in C} h(2^{-2^{|w|}}) = 2 \cdot 2^n \sum_{l \leq j \leq n} \binom{n}{j} 2^{-2^j} \leq 2 \cdot 2^n \; n^{k+1} 2^{-2^k} = O\left(\frac{2^n}{n}\right),$$

where we have bounded the second sum by the number of terms (at most n) times the largest term (the one for $j = k$). Substituting this bound in (3.10) yields

$$\sum_{w \in B} h(q_w) \geq 2^n \left(1 + O\left(\frac{1}{n}\right)\right). \tag{3.11}$$

We can now use (3.9) and (3.11) in the same way we used (3.1) and (3.2). Let $Q = \#B$ denote the number of elements in B. By the same reasoning that led to (3.7), we obtain

$$\overline{\ell}(n) \geq \frac{2^n}{\log\left(\frac{eQ}{\overline{\ell}(n)}\right)} \left(1 + O\left(\frac{1}{n}\right)\right). \tag{3.12}$$

To bound Q, we note that B contains $\binom{n}{j}2^{n-j}$ terms w with $n - |w| = j$. Thus

$$Q = \sum_{0 \le j < k} \binom{n}{j} 2^{n-j} \le 2^n \, n^k, \tag{3.13}$$

where we have bounded the sum by the number of terms (at most n) times the largest term (the one for $j = k - 1$). Substituting this bound for Q and the bound (3.8) for $\overline{\ell}(n)$ into the right-hand side of (3.12) completes the proof of Theorem 2.1. $\triangle$

4. Upper Bound: Algorithm

In this section and the next four we shall prove the following result.

Theorem 4.1: For $n \to \infty$,

$$\overline{\ell}(n) \le \frac{H(n)\, 2^n}{\log n \, \log\log n} \left(1 + O\left(\frac{\log\log\log\log n}{\log\log n}\right)\right), \tag{4.1}$$

where

$$H(n) = 2^{\vartheta} \left(\frac{1}{2} + \left(\sum_{k \ge 0} 2^k 2^{-2^k 2^{\vartheta}}\right) \ln 2\right) \tag{4.2}$$

and

$$\vartheta = \{\log\log n + \log\log\log n\}. \tag{4.3}$$

Proof: We shall give a specific algorithm for constructing a disjunctive normal form for an arbitrary Boolean function, and analyze the behaviour this algorithm for a random Boolean function.

Let f be a random Boolean function of n arguments. We shall construct a disjunctive normal form Φ for f in S stages. For $1 \le R \le S$, stage R will use a parameter k_R. The significance of the parameter k_R is that all terms added to Φ during stage R will correspond to subcubes of dimension k_R. For $1 \le R \le S$, stage R will use a parameter m_R and a set M_R containing m_R arguments. The significance of M_R is that all terms added to Φ during stage R will correspond to subcubes having their dimensions in directions corresponding to arguments in M_R.

We shall choose the sets $M_1, \ldots, M_S$ to be pairwise disjoint; that is, for $1 \le Q < R \le S$ we have

$$M_Q \cap M_R = \emptyset. \tag{4.4}$$

To ensure that we may choose $M_1, \ldots, M_S$ in this way, we shall choose $m_1, \ldots, m_S$ to satisfy the condition

$$\sum_{1 \leq R \leq S} m_R \leq n. \tag{4.5}$$

Since both the probability distribution on f and the algorithm for constructing Φ (to be described below) are invariant under the group of symmetries of $\mathbf{B}^n$, and since this group includes the subgroup induced by permutations of the arguments, it is immaterial which arguments are assigned to which of the sets $M_1, \ldots, M_S$, as long as these sets have the correct numbers $m_1, \ldots, m_S$ of elements and the disjointness condition (4.4) is satisfied.

For $1 \leq R \leq S$, we shall say that a subcube C of $\mathbf{B}^n$ is *considered* at stage R if (1) C has dimension k_R, and (2) C has all its dimensions in directions corresponding to arguments in M_R. There are $2^{n-k_R}\binom{m_R}{k_R}$ subcubes considered in stage R.

For $1 \leq R \leq S+1$, we shall say that a point v in $\mathbf{B}^n$ *remains to be covered* at stage R if (1) v belongs to U_f and (2) v was not covered by any of the terms added to Φ in stages $1, \ldots, R-1$. All the points in U_f remain to be covered at stage 1; to ensure the correctness of our algorithm, we must arrange that no points remain to be covered at stage $S+1$.

The algorithm for constructing Φ can now be stated as follows. In stage R, we add to Φ all terms corresponding to subcubes that

(I) are considered in stage R, and

(II) are such that all their points remain to be covered at stage R.

By virtue of condition (II), every term added to Φ is an implicant if f, and thus the final form Φ implies f. Furthermore, shall choose

$$k_S = 0, m_S = 0 \text{ and } M_S = \emptyset. \tag{4.6}$$

These conditions mean that in the last stage we add to Φ a term covering a single point for each point of U_f that has not been covered in any previous stage. This ensures that f implies Φ, and thus that the final form Φ is indeed a disjunctive normal form for f.

5. Upper Bound: Independence

Consider the probability that point v remains to be covered at stage R. Since both the probability distribution on f and the algorithm for constructing Φ are invariant under the group of symmetries of $\mathbf{B}^n$, and since this group acts transitively on the points of $\mathbf{B}^n$, this probability is the same for all points v. It will be denoted p_R. We have $p_1 = 1/2$ and $p_{S+1} = 0$.

For each point v in $\mathbf{B}^n$ and each stage R, let $B_{v,R}$ be the event "v remains to be covered at stage R". In general the events $\{B_{v,R} : v \in \mathbf{B}^n\}$ are not independent, since a subcube added to Φ in one of the stages $1, \ldots, R-1$ may cover several of these points. Our goal in this section, however, is to show that we do have independence in some special cases, as a consequence of our disjointness assumption (4.4).

Say that two points v and w in $\mathbf{B}^n$ are R-*equivalent* if their coordinates differ only in arguments belonging to $M_1 \cup \cdots \cup M_{R-1}$.

Lemma 5.1: Let $v_1, \ldots, v_N$ be points in $\mathbf{B}^n$. If no two distinct points among $v_1, \ldots, v_N$ are R-equivalent, the events $B_{v_1,R}, \ldots, B_{v_N,R}$ are independent.

Proof: For $0 \le R \le S$, we shall construct a dependence graph Γ_R as follows. The vertices of Γ_R will be the points of $\mathbf{B}^n$. The graph Γ_1 will have no edges. For $2 \le R \le S$, the graph Γ_R will be obtained by adding to Γ_{R-1} edges joining every pair of points that both belong to a subcube considered in stage $R-1$.

If $v_1, \ldots, v_N$ are all contained in distinct connected components of Γ_R, then the events $B_{v_1,R}, \ldots, B_{v_N,R}$ are independent. We prove this by induction on R. The assertion is true for stage $R = 1$, since a point remains to be covered at stage 1 if and only if it belongs to U_f, and all points independently belong to U_f with probability 1/2. Furthermore, the truth of this assertion is maintained for each stage $R \ge 2$, since $B_{v,R}$ depends only on the events $\{B_{w,R-1} : w \in V\}$, where V is the set of points contained in subcubes containing v and considered in stage $R-1$, and edges joining v to all other vertices in V are added to Γ_{R-1} to obtain Γ_R.

Finally, we observe that R-equivalence is the transitive closure of Γ_R, since pairs of points that both belong to a subcube considered in stage R differ only in arguments belonging to M_R. $\triangle$

Corollary 5.2: Let v be a point in $\mathbf{B}^n$, and let W be the set of points contained in subcubes containing v and considered in stage R. Then the events $\{B_{w,R} : w \in W\}$ are independent.

Proof: Any two distinct points in W must differ in some argument, and they can differ only in arguments belonging to M_R. Since M_R is disjoint from $M_1 \cup \cdots M_{R-1}$ by the condition (4.4), no two points in W can be R-equivalent. The corollary thus follows from Lemma 5.1. $\triangle$

Corollary 5.3: Let C be any subcube considered in stage R. The the events $\{B_{w,R} : w \in C\}$ are independent.

Proof: We have $C \subseteq W$ for any point $v \in C$ in Corollary 5.2. $\triangle$

Let E_R denote the expected number of terms added to Φ during stage R. The upper bound we seek can be written

$$\overline{\ell}(n) \leq \sum_{1 \leq R \leq S} E_R. \tag{5.1}$$

In much of what follows we shall be discussing some stage R, for $1 \leq R \leq S$. To avoid a surfeit of subscripts, we shall write k, m, p and E for k_R, m_R, p_R and E_R.

Corollary 5.3 allows us to write an expression for E. There are $2^{n-k}\binom{m}{k}$ subcubes considered in stage R. Each of these adds a term to Φ if an only if each of its 2^k points remains to be covered at stage R. These 2^k events each occur with probability p, and by Corollary 5.3 they are independent. Thus we have

$$E = 2^{n-k}\binom{m}{k}p^{2^k}. \tag{5.2}$$

6. Upper Bound: Correlations

The terms added to Φ in different stages do not overlap, so there is no duplication of effort between distinct stages. The terms added within a single stage may overlap, however, and we shall need to carefully estimate the overlaps among such terms. To do this we shall use two correlation inequalities. The first of these is Harris's inequality [H]. Let Z be a finite set. Let Y be a random subset of Z such that each element $z \in Z$ appears independently in Y with some probability p_z. Let $X_1, \ldots, X_N$ be subsets of Z, and let $A_1, \ldots, A_N$ be the corresponding events: A_M occurs if and only if $X_M \subseteq Y$. Then Harris's inequality asserts that

$$\Pr\left[\bigcap_{1 \leq M \leq N} \overline{A_M}\right] \geq \prod_{1 \leq M \leq N} \Pr\left[\,\overline{A_M}\,\right]. \tag{6.1}$$

The second correlation inequality that we shall use is Janson's inequality [J1] (see also Boppana and Spencer [B]). Suppose that $\Pr[A_M] \leq 1/2$ for $1 \leq M \leq N$. Construct a graph Γ on the vertices $1, \ldots, N$ by taking $LM \in \Gamma$ if and only if $X_L \cap X_M \neq \emptyset$. Then Janson's inequality asserts that

$$\Pr\left[\bigcap_{1 \leq M \leq N} \overline{A_M}\right] \leq \left(\prod_{1 \leq M \leq N} \Pr\left[\,\overline{A_M}\,\right]\right) \exp\left(2 \sum_{LM \in \Gamma} \Pr[A_L \cap A_M]\right). \tag{6.2}$$

Since we shall be discussing a single stage R, we shall write k, m, p and E for k_R, m_R, p_R and E_R as in the preceding section, and also write p' for p_{R+1}.

Lemma 6.1 We have

$$p' \geq p\left(1 - p^{2^k-1}\right)^{\binom{m}{k}}.$$

Proof: Consider a point v. The probability that v remains to be covered at stage R is p. If v remains to be covered at stage R, it will be covered during stage R unless each of the $\binom{m}{k}$ subcubes containing v considered during stage R contains a point that does not remain to be covered at stage R, so that none of the terms corresponding to these subcubes is added to Φ during stage R. We shall assume that v remains to be covered during stage R, and tacitly condition all other probabilities on this event.

We shall apply Harris's inequality (6.1) with Z being the set containing all points in subcubes containing v and considered during stage R. Let Y be the subset of Z comprising those points that remain to be covered at stage R. By Corollary 5.2, every point of Z independently appears in Y with probability p. Let $N = \binom{m}{k}$, let $X_1, \ldots, X_N$ be the subcubes containing v and considered during stage R, and for $1 \leq M \leq N$, let A_M the event $X_M \subseteq Y$, so that A_M occurs if and only if the term corresponding to X_M is added during stage R. Then we have

$$\Pr[A_M] = 1 - p^{2^k-1}.$$

Using Harris's inequality (6.1), we have

$$\Pr\left[\bigcap_{1 \leq M \leq N} \overline{A_M}\right] \leq \left(1 - p^{2^k-1}\right)^{\binom{m}{k}}.$$

It follows that

$$p' \geq p\left(1 - p^{2^k-1}\right)^{\binom{m}{k}},$$

which completes the proof of the lemma. $\triangle$

Lemma 6.2: We have

$$p' \leq p\left(1 - p^{2^k-1}\right)^{\binom{m}{k}} \exp\left(2 \sum_{1 \leq j \leq k-1} \binom{m}{k}\binom{k}{j}\binom{m-k}{k-j} p^{2\cdot 2^k - 2^j - 1}\right).$$

Proof: We shall proceed as in the proof of Lemma 6.1, but use Janson's inequality instead of Harris's inequality. We observe that the inequality $\Pr[A_M] \leq 1/2$ follows from $p \leq 1/2$

and $k \geq 1$. The graph Γ has vertices corresponding to the $\binom{m}{k}$ subcubes containing v and considered during stage R, and an edge joining vertices L and M if the corresponding subcubes have an intersection of dimension strictly greater than 0 (that is, have an intersection strictly larger than $\{v\}$). If these two subcubes have an intersection of dimension j, then their union contains $2 \cdot 2^k - 2^j$ points, and thus contains $2 \cdot 2^k - 2^j - 1$ points other than v. By Corollary 5.2, each of these points independently remains to be covered with probability p, so we have

$$\Pr[A_L \cap A_M] = p^{2 \cdot 2^k - 2^j - 1}.$$

There are $\binom{m}{k}$ choices for L and, for each of these, $\binom{k}{j}\binom{m-k}{k-j}$ choices for a subcube M that has an intersection with L of dimension j. Thus we have

$$\sum_{LM \in \Gamma} \Pr[A_L \cap A_M] = \sum_{1 \leq j \leq k-1} \binom{m}{k}\binom{k}{j}\binom{m-k}{k-j} p^{2 \cdot 2^k - 2^j - 1}.$$

Substituting this expression into Janson's inequality (6.2) yields

$$p' \leq p \left(1 - p^{2^k - 1}\right)^{\binom{m}{k}} \exp\left(2 \sum_{1 \leq j \leq k-1} \binom{m}{k}\binom{k}{j}\binom{m-k}{k-j} p^{2 \cdot 2^k - 2^j - 1}\right),$$

which completes the proof of the lemma. $\triangle$

7. Upper Bound: Estimates

In this section we shall describe how the values of k and m are chosen for each stage except the last. The parameter k will depend only on p, which was defined in Section 5. The parameter m will depend on p and k, but also on a new parameter q, which will be chosen in the next section. The significance of q is that it represents the factor by which would like to reduce p during the stage; that is, we hope to have p' very nearly equal to pq.

We shall use these choices of parameters to obtain estimates for the expressions appearing in (5.2) and in Lemmas 6.1 and 6.2. These estimates will allow us to prove that p' is indeed very nearly equal to pq, in a sense that will be made precise below. Finally, we shall use these estimates to verify the condition (4.5).

In what follows, the constants implicit in O-notation will always be absolute, independent of n and any other parameters. When we assert inequalities involving n and any other parameters, we shall always tacitly make the reservation that the inequalities are asserted only for all sufficiently large n.

Given p, we shall choose

$$k = \lfloor \log\log n + \log\log\log n \rfloor - \left\lceil \log\log\left(\frac{1}{p}\right) - \{\log\log n + \log\log\log n\} \right\rceil . \tag{7.1}$$

In the estimates that follow, we shall always have

$$\frac{1}{(\log\log n)^4} \leq p \leq \frac{1}{2}. \tag{7.2}$$

This implies that

$$2 \leq \frac{1}{p} \leq (\log\log n)^4$$

and

$$0 \leq \log\log\left(\frac{1}{p}\right) \leq \log\log\log\log n + 2. \tag{7.3}$$

Substituting these bounds in (7.1) yields

$$\log\log n + 1 \leq k \leq 2\log\log n. \tag{7.4}$$

Furthermore, we shall always choose q so that

$$\frac{1}{(\log n)^2} \leq q \leq 1 - \frac{1}{\log\log n}. \tag{7.5}$$

This implies that

$$\left(1 - \frac{1}{\log\log n}\right)^{-1} \leq \frac{1}{q} \leq (\log n)^2$$

and, since $\ln(1+x) \leq x$,

$$\frac{1}{\log\log n} \leq \ln\left(\frac{1}{q}\right) \leq 2\ln\log n. \tag{7.6}$$

Given p, k and q, we shall choose m to be the smallest integer such that

$$\binom{m}{k} \geq \left(\frac{1}{p}\right)^{2^k - 1} \ln\left(\frac{1}{q}\right). \tag{7.7}$$

It follows from this choice of m that

$$\binom{m-1}{k} < \left(\frac{1}{p}\right)^{2^k - 1} \ln\left(\frac{1}{q}\right). \tag{7.8}$$

First, we shall prove

$$m \le \frac{n}{(\log\log n)^2}. \tag{7.9}$$

To do this, we first note that (7.1) implies

$$k \le \log\log n + \log\log\log n - \log\log\left(\frac{1}{p}\right),$$

which implies

$$\left(\frac{1}{p}\right)^{2^k} \le n^{\log\log n}.$$

On the other hand, the lower bound in (7.4) implies

$$n^{k-1} \ge n^{\log\log n},$$

so that we have

$$n^{k-1} \ge \left(\frac{1}{p}\right)^{2^k-1}.$$

This, together with $\binom{m}{k} \ge (m/k)^k$ and the upper bounds in (7.2), (7.4) and (7.6), allows us to verify that (7.7) holds for $m = \lfloor n/(\log\log n)^2 \rfloor \ge n/2(\log\log n)^2$:

$$\binom{\left\lfloor \frac{n}{(\log\log n)^2} \right\rfloor}{k} \ge \left(\frac{n}{2k(\log\log n)^2}\right)^k \ge \left(\frac{n}{4(\log\log n)^3}\right)^k \ge n^{k-1}(2\ln\log n)$$
$$\ge \left(\frac{1}{p}\right)^{2^k-1}(2\ln\log n) \ge \left(\frac{1}{p}\right)^{2^k-1}\ln\left(\frac{1}{q}\right).$$

This completes the verification of (7.9).

Next, we shall prove

$$m \ge n^{1/8}. \tag{7.10}$$

To do this, we first note that (7.1) implies

$$k \ge \log\log n + \log\log\log n - \log\log\left(\frac{1}{p}\right) - 1,$$

which implies

$$\left(\frac{1}{p}\right)^{2^k} \ge n^{(\log\log n)/2}.$$

This, together with $\binom{m}{k} \le m^k$ and the upper bound in (7.4) and the lower bounds in (7.2) and (7.6), allows us to verify that (7.7) does not hold for $m = \lfloor n^{1/8} \rfloor \le n^{1/8}$:

$$\binom{\lfloor n^{1/8} \rfloor}{k} \le n^{k/8} \le n^{(\log\log n)/4} < \frac{n^{(\log\log n)/2}}{(\log\log n)^5} \le \left(\frac{1}{p}\right)^{2^k-1} \ln\left(\frac{1}{q}\right).$$

This completes the verification of (7.10).

From (7.8), we have

$$\binom{m}{k} = \binom{m-1}{k}\frac{m}{m-k} < \left(\frac{1}{p}\right)^{2^k-1} \ln\left(\frac{1}{q}\right)\left(1 + \frac{k}{m-k}\right).$$

Combining this with (7.10) and the upper bound in (7.4) yields

$$\binom{m}{k} = \left(\frac{1}{p}\right)^{2^k-1} \ln\left(\frac{1}{q}\right)\left(1 + O\left(\frac{\log\log n}{n^{1/8}}\right)\right). \tag{7.11}$$

Our first application of (7.11) is to estimate the right-hand side of (5.2):

$$\begin{aligned} E &= 2^{n-k}\binom{m}{k}p^{2^k} \\ &= 2^{n-k}\, p \ln\left(\frac{1}{q}\right)\left(1 + O\left(\frac{\log\log n}{n^{1/8}}\right)\right). \end{aligned} \tag{7.12}$$

Our next application of (7.11) is to show that p' is indeed very nearly equal to pq.

Proposition 7.1: Suppose that p and q satisfy (7.2) and (7.5), and that k and m are chosen according to (7.1) and (7.7). Then

$$p' = pq\left(1 + O\left(\frac{1}{n^{1/16}}\right)\right).$$

Proof: From Lemma 6.1 we have

$$p' \ge pI, \tag{7.13}$$

where

$$\begin{aligned} I &\ge \left(1 - p^{2^k-1}\right)^{\binom{m}{k}} \\ &= \exp\left(\binom{m}{k} \ln\left(1 - p^{2^k-1}\right)\right). \end{aligned} \tag{7.14}$$

From (7.11), we have

$$\binom{m}{k} p^{2^k-1} = \ln\left(\frac{1}{q}\right)\left(1+O\left(\frac{\log\log n}{n^{1/8}}\right)\right) \tag{7.15}$$

and, using the upper bound in (7.6), $k \geq 2$ and (7.10),

$$\begin{aligned} p^{2^k-1} &= \ln\left(\frac{1}{q}\right)\left(1+O\left(\frac{\log\log n}{n^{1/8}}\right)\right) \Big/ \binom{m}{k} \\ &= O\left(\frac{1}{n^{1/8}}\right). \end{aligned}$$

Using these estimates, together with $\ln(1+x) = x + O(x^2)$, in (7.14) yields

$$I = q\left(1+O\left(\frac{\log\log n}{n^{1/8}}\right)\right), \tag{7.16}$$

and substituting this in (7.13) yields

$$p' \geq pq\left(1+O\left(\frac{\log\log n}{n^{1/8}}\right)\right). \tag{7.17}$$

From Lemma 6.2 we have

$$p' \leq pI \exp J, \tag{7.18}$$

where

$$J = 2 \sum_{1\leq j\leq k-1} \binom{m}{k}\binom{k}{j}\binom{m-k}{k-j} p^{2\cdot 2^k - 2^j - 1}. \tag{7.19}$$

We shall now prove

$$J = O\left(\frac{1}{n^{1/16}}\right). \tag{7.20}$$

We begin by using the estimates $\binom{k}{j} \leq 2^k$ and

$$\binom{m-k}{k-j} \leq \binom{m-j}{k-j} = \binom{m}{k}\frac{k(k-1)\cdots(k-j+1)}{m(m-1)\cdots(m-j+1)} \leq \binom{m}{k}\left(\frac{k}{m}\right)^j$$

in (7.19), together with (7.11) and the upper bounds in (7.4) and (7.6), to obtain

$$\begin{aligned} J &\leq 2\cdot 2^k \left(\ln\left(\frac{1}{q}\right)\right)^2 \left(1+O\left(\frac{\log\log n}{n^{1/8}}\right)\right) \sum_{1\leq j\leq k-1} \left(\frac{k}{m}\right)^j \left(\frac{1}{p}\right)^{2^j-1} \\ &= O\left((\log n)^2 (\log\log n)^2 \sum_{1\leq j\leq k-1} \left(\frac{k}{m}\right)^j \left(\frac{1}{p}\right)^{2^j-1}\right). \end{aligned} \tag{7.21}$$

Writing $T_j = (k/m)^j (1/p)^{2^j-1}$ for the general term in the last sum, we observe that $\log T_j = j \log(k/m) + (2^j - 1)\log(1/p)$ is a convex function of j, so that the largest term in the sum is either the first or the last. In either case, we can bound the sum from above by the number $k-1$ of terms times the largest term. If the first term is the largest, we have

$$\begin{aligned} (k-1)T_1 &= \frac{(k-1)k}{mp} \\ &\leq \frac{4(\log\log n)^6}{n^{1/8}}, \end{aligned} \tag{7.22}$$

using (7.10), the lower bound in (7.2), and the upper bound in (7.4). If the last term is the largest, we use (7.7) and the inequality $\binom{m}{k} \leq (4m/k)^k$ to obtain

$$\begin{aligned} \left(\frac{1}{p}\right)^{2^{k-1}-1} &\leq \left(\frac{1}{p}\right)^{(2^k-1)/2} \\ &= \binom{m}{k}^{1/2} \left(\ln\left(\frac{1}{q}\right)\right)^{-1/2} \\ &\leq \left(\frac{4m}{k}\right)^{k/2} \left(\ln\left(\frac{1}{q}\right)\right)^{-1/2}. \end{aligned}$$

Thus we have, for $k \geq 4$,

$$\begin{aligned} (k-1)T_{k-1} &\leq (k-1)4^{k/2} \left(\frac{k}{m}\right)^{k/2-1} \left(\ln\left(\frac{1}{q}\right)\right)^{-1/2} \\ &\leq \frac{(k-1)k2^k}{m} \left(\ln\left(\frac{1}{q}\right)\right)^{-1/2}. \\ &\leq \frac{4(\log\log n)^{5/2}(\log n)^2}{n^{1/8}}, \end{aligned} \tag{7.23}$$

using $k \geq 4$, the upper bound in (7.4), the lower bound in (7.6), and (7.10). Using (7.22) or (7.23) to estimate the sum in (7.21) yields (7.20). Using (7.16) and (7.20) in (7.18) yields

$$p' \leq pq\left(1 + O\left(\frac{1}{n^{1/16}}\right)\right).$$

Finally, combining this estimate with (7.17) completes the proof of the lemma. △

Let us now verify the condition (4.5). Apart from the last stage, each stage reduces p by a factor of

$$q\left(1+O\left(\frac{1}{n^{1/16}}\right)\right) \leq \left(1-\frac{1}{\log\log n}\right)\left(1+O\left(\frac{1}{n^{1/16}}\right)\right)$$
$$\leq \left(1-\frac{1}{2\log\log n}\right),$$

by Theorem 7.1 and the upper bound in (7.5). During all but the last two of these stage, p is reduced from its initial value $1/2$ to a value at least $1/(\log\log n)^4$, by the lower bound in (7.2). Thus

$$S \leq 2 + \left\lceil \frac{\log\left(\frac{1}{(\log\log n)^4}\right)}{\log\left(1-\frac{1}{\log\log n}\right)} \right\rceil .$$

It follows that

$$S = O(\log\log n \ \log\log\log n) \tag{7.24}$$

bounds the number of terms in the sum in (4.5). Since each term is bounded by (7.9), the verification of (4.2) is complete.

8. Upper Bound: Optimization

In this section we shall estimate the sum (5.1), which will complete the proof of Theorem 4.1.

We shall define

$$i = \lfloor \log\log n + \log\log\log n \rfloor$$

and

$$j = \left\lceil \log\log\left(\frac{1}{p}\right) - \{\log\log n + \log\log\log n\} \right\rceil, \tag{8.1}$$

so that (7.1) can be written

$$k = i - j.$$

Since p decreases from the initial value $1/2$, k decreases from the initial value i, with k decreasing by 1 (from $i-j$ to $i-j-1$) each time p passes through one of the thresholds

$$\beta_j = 2^{-2^{\vartheta+j}}, \tag{8.2}$$

for $j = 0, 1, \ldots$. We shall divide all the stages except for the last two in *phases*, where phase j comprises all the stages in which $k = i - j$.

We can now describe how q is chosen in each stage but the last. If

$$p \le \frac{4}{(\log\log n)^4},$$

then we shall take

$$q = \frac{1}{(\log n)^2},$$

and this stage, called the *chomp*, will be the last stage but one. Otherwise, if the current stage is in phase j and

$$p \le \beta_j \left(1 - \frac{2}{\log\log n}\right),$$

then we shall take

$$q = \frac{1}{2},$$

and this stage, called a *bite*, will the the last stage of the current phase. Otherwise, we shall take

$$q = 1 - \frac{1}{\log\log n}, \tag{8.3}$$

and this stage will be called a *nibble*. Each phase, except the first and last, comprises a number of nibbles, which together we shall call a *fresser*, followed by a bite. (The first phase, in which $j = 0$, may have no fresser (if ϑ is sufficiently close to 0). The last phase may have no fresser (if one of the β_j is sufficiently close to $4/(\log\log n)^4$), and the last phase is terminated by the chomp rather than by a bite.) From (8.1) and the upper bound in (7.2), we have

$$j \le \log\log\log\log n + 3, \tag{8.4}$$

so in particular there are $O(\log\log\log\log n)$ phases.

Let us consider the contribution to the sum in (5.1) from a fresser that reduces p from α_j to a value at least β_j, in a phase with

$$\begin{aligned} k &= i - j \\ &= \log\log n + \log\log\log n - \vartheta - j. \end{aligned}$$

We have

$$2^{n-k} = \frac{2^n \, 2^{\vartheta + j}}{\log n \, \log\log n}.$$

For each nibble of the fresser we have $q = 1 - 1/\log\log n$, which implies

$$\frac{1}{1-q} \ln\left(\frac{1}{q}\right) = 1 + O\left(\frac{1}{\log\log n}\right). \tag{8.5}$$

By induction using Proposition 7.1, we have

$$p = \alpha_j q^{R-Q} \left(1 + O\left(\frac{1}{n^{1/16}}\right)\right)^{R-Q}$$

for stage R in the fresser, where stage Q is the first stage in the fresser. Since $R - Q \le S = O(\log\log n\ \log\log\log n)$ by (7.24), this estimate becomes

$$p = \alpha_j q^{R-Q} \left(1 + O\left(\frac{\log\log n\ \log\log\log n}{n^{1/16}}\right)\right).$$

Summing (7.12) over the stages in the fresser and using (8.5), we obtain

$$\frac{2^n\ 2^{\vartheta+j}(\alpha_j - \beta_j)}{\log n\ \log\log n} \left(1 + O\left(\frac{1}{\log\log n}\right)\right) \tag{8.6}$$

as an upper bound to the expected number of terms added to Φ during the fresser.

Let us now consider the bite in such a phase. As before, we have

$$2^{n-k} = \frac{2^n\ 2^{\vartheta+j}}{\log n\ \log\log n}.$$

Since $q = 1/2$, by Proposition 7.1 the bite reduces p from a value

$$\beta_j \left(1 + O\left(\frac{1}{\log\log n}\right)\right)$$

to a value

$$\frac{\beta_j}{2} \left(1 + O\left(\frac{1}{\log\log n}\right)\right).$$

From (7.12), we have

$$\frac{2^n\ 2^{\vartheta+j}\beta_j}{\log n\ \log\log n} \left(\ln 2 + O\left(\frac{1}{\log\log n}\right)\right) \tag{8.7}$$

as an upper bound to the expected number of terms added to Φ during the bite.

Let us now sum the contributions (8.6) of the fressers over the various phases. We have $\alpha_0 = 1/2$. Since for $j \ge 1$ we have

$$\alpha_j = \frac{\beta_{j-1}}{2} \left(1 + O\left(\frac{1}{\log\log n}\right)\right),$$

the sum telescopes to

$$\frac{2^n\,2^\vartheta}{\log n\ \log\log n}\left(\frac{1}{2}+O\left(\frac{\log\log\log\log n}{\log\log n}\right)\right), \tag{8.8}$$

where the numerator in the error term reflects the fact that there are $O(\log\log\log\log n)$ phases.

Let us now sum the contributions (8.7) of the bites over the various phases. Using (8.2), these contributions sum to at most

$$\frac{2^n\,2^\vartheta}{\log n\ \log\log n}\left(\left(\sum_{j\geq 0} 2^j\ 2^{-2^{\vartheta+j}}\right)\ln 2+O\left(\frac{\log\log\log\log n}{\log\log n}\right)\right). \tag{8.9}$$

The contributions (8.8) and (8.9) together give the bound in Theorem 4.1, but we have still to account for the last two stages. By Proposition 7.1, the last stage but one (the chomp) reduces p from a value in the range

$$\frac{1}{(\log\log n)^4}\leq p\leq\frac{4}{(\log\log n)^4},$$

to a value at most

$$\begin{aligned} q\left(1+O\left(\frac{1}{n^{1/16}}\right)\right) &= \frac{1}{(\log n)^2}\left(1+O\left(\frac{1}{n^{1/16}}\right)\right) \\ &\leq \frac{8}{(\log n)^2\ (\log\log n)^4}. \end{aligned} \tag{8.10}$$

By the lower bound in (7.4), we have

$$2^{n-k}=O\left(\frac{2^n}{\log n}\right).$$

Using the lower bound in (7.2) and the upper bound in (7.6) in (7.12), we obtain the bound

$$2^{n-k}p\ln\left(\frac{1}{q}\right)=O\left(\frac{2^n}{\log n\ (\log\log n)^3}\right) \tag{8.11}$$

for the expected number of terms added to Φ in the last stage but one.

Finally, for the last stage we have

$$p=O\left(\frac{1}{(\log n)^2\ (\log\log n)^4}\right)$$

from (8.10). For this stage, we add a term for each point that remains to be covered, so we use $p2^n$ rather than (7.12) to estimate the contribution. Thus we have bound

$$p2^n = O\left(\frac{2^n}{(\log n)^2\,(\log\log n)^4}\right) \tag{8.12}$$

for the expected number of terms added to Φ in the last stage. The contributions (8.11) and (8.12) are accommodated by the error term in Theorem 4.1, which completes the proof of the theorem. $\triangle$

9. Conclusion

We have presented the best lower bound known and a new upper bound for the length of the shortest disjunctive normal form for a random Boolean function. These bounds have the same order of growth, but a gap of a bounded oscillating factor remains between them. The obvious remaining problem is to close this gap. This will entail determining whether $\overline{\ell}(n) \;/\; \left(2^n/\log n\;\log\log n\right)$ tends to a limit independent of n as $n \to \infty$ (as would be the case if the upper bound were tightened to match the current best lower bound) or not (as would be the case if the lower bound were tightened to match the the upper bound of this paper).

10. References

[A] K. Azuma, "Weighted Sums of Certain Dependent Random Variables", *Tôhoku Math. J.*, 19 (1967) 357–367.

[B] R. Boppana and J. Spencer, "A Useful Elementary Correlation Inequality", *J. Combinatorial Theory A*, 50 (1989) 305–307.

[F1] C. M. Fortuin, P. W. Kasteleyn and J. Ginibre, "Correlation Inequalities on Some Partially Ordered Sets", *Commun. Math. Phys.*, 22 (1971) 89–103.

[F2] P. Frankl and V. Rödl, "Near Perfect Coverings in Graphs and Hypergraphs", *European J. Combinatorics*, 6 (1985) 317–326.

[G] V. V. Glagolev, "Nekotorye Otsenki Dizyunktivnykh Normalnykh Form Funktsiĭ Algebry Logiki", *Problemy Kibernetiki*, 19 (1967) 75–95.

[H] T. E. Harris, "A Lower Bound for the Critical Probability in a Certain Percolation Process", *Proc. Cambridge Phil. Soc.*, 56 (1960) 13–20.

[J1] S. Janson, T. Łuczak and A. Ruciński, "An Exponential Bound for the Probability of Nonexistence of a Specified Subgraph in a Random Graph", in: M. Karónski, J. Jaworski and A. Ruciński (Editors), *Random Graphs '87*, John Wiley and Sons, Chichester, 1990.

[J2] F. Jelinek, *Probabilistic Information Theory*, McGraw-Hill, New York, 1968.

[K1] A. D. Korshunov, "Verkhnyaya Otsenka Slozhnosti Kratchaĭshikh DNF Pochti Vsekh Bulevykh Funktsiĭ", *Kibernetika*, 6 (1969) 1–8.

[K2] A. D. Korshunov, "O Slozhnosti Kratchaĭshikh Dizyunktivnykh Normalnykh Form Bulevykh Funktsiĭ", *Metody Diskretnogo Analiza*, 37 (1981) 9–41.

[K3] A. D. Korshunov, "O Slozhnosti Kratchaĭshikh Dizyunktivnykh Normalnykh Form Sluchaĭnykh Bulevykh Funktsiĭ", *Metody Diskretnogo Analiza*, 40 (1983) 25–53.

[K4] S. E. Kuznetsov, "O Nizhneĭ Otsenke Dliny Kratchaĭsheĭ DNF Pochti Vsekh Bulevykh Funktsiĭ", *Veroyatnoste Metody i Kibernetiki*, 19 (1983) 44–47.

[L] O. B. Lupanov, "O Realizatsiĭ Funktsiĭ Algebry Logiki Formulami iz Konechnykh Klassov (Formulami Ogranichennoĭ Glubiny) v Bazise &, $\vee$, $^{-}$", *Problemy Kibernetiki*, 6 (1961) 5–14.

[M] E. J. McCluskey, Jr., "Minimization of Boolean Functions", *Bell System Tech. J.*, 35 (1956) 1417–1444.

[N1] R. G. Nigmatullin, "Nekotorye Metricheskie Sootnosheniya v Edinichnom Kube", *Diskretnyĭ Analiz*, 9 (1967) 47–58.

[N2] R. G. Nigmatullin, "Variatsionnyĭ Printsip v Algebre Logiki", *Diskretnyĭ Analiz*, 10 (1967) 69–89.

[P] N. Pippenger and J. Spencer, "Asymptotic Behavior of the Chromatic Index for Hypergraphs", *J. Combinatorial Theory A*, 51 (1989) 24–42.

[Q1] W. V. Quine, "The Problem of Simplifying Truth Functions", *Amer. Math. Monthly*, 59 (1952) 521–531.

[Q2] W. V. Quine, "A Way to Simplify Truth Functions", *Amer. Math. Monthly*, 62 (1955) 627–631.

[R1] J. Riordan and C. E. Shannon, "The Number of Two-Terminal Series-Parallel Networks", *J. Math. Phys.*, 21 (1942) 83–93.

[R2] V. Rödl, "On a Packing and Covering Problem", *European J. Combinatorics*, 5 (1985) 69–78.

[S1] A. A. Sapozhenko, "O Slozhnosti Dizyunktivnykh Normalnykh Form, Poluchaemykh s Pomoshchyu Gradientnogo Algoritma", *Diskretnyĭ Analiz*, 21 (1972) 62–71.

[S2] J. Spencer, *Ten Lectures on the Probabilistic Method*, Society for Industrial and Applied Mathematics, Philadelphia, 1987.