RANDOMIZED VS. DETERMINISTIC DECISION TREE COMPLEXITY FOR READ-ONCE BOOLEAN FUNCTIONS

Rafi Heiman and Avi Wigderson

Abstract. We consider the deterministic and the randomized decision tree complexities for Boolean functions, denoted $DC(f)$ and $RC(f)$, respectively. A major open problem is how small $RC(f)$ can be with respect to $DC(f)$. It is well known that $RC(f) \geq DC(f)^{0.5}$ for every Boolean function f (called "0.5-exponent"). On the other hand, some Boolean function f is known to have $RC(f) = \Theta(DC(f)^{0.753\ldots})$ (or "0.753...-exponent"). It is not known whether there is a Boolean function with exponent smaller than 0.753... Likewise, no lower bound for arbitrary Boolean functions with exponent greater than 0.5 is known.

Our result is a 0.51 lower bound on the exponent for every *read-once* function. Read-once means that each input variable appears exactly once in the Boolean formula representing the function. To obtain this result we generalize an existing lower bound technique and combine it with restriction arguments. This result provides a lower bound of $n^{0.51}$ on the number of positions that have to be evaluated by any randomized α-β pruning algorithm computing the value of any two-person zero-sum game tree with n final positions.

Key words. Boolean decision trees, Randomized complexity, Read-once formulae.

Subject classifications. 68C25.

1. Introduction

The Boolean decision tree model is an extremely simple model for computing Boolean functions. In this model the algorithm has to determine a value of some known Boolean function, f, on an unknown input. The algorithm reads input variables, one at a time, in an adaptive manner (and therefore is called a tree). The computation stops as soon as sufficiently many variables have been read to determine the value of f, no matter what the values of other variables are. The time measure in this model is simply the number of input variables read (for the worst case input setting). The deterministic decision tree complexity, $DC(f)$, is the maximum number of variables that are read by the best decision tree.

A variant of this model is the randomized decision tree, first studied by Yao [?]. A randomized decision tree is a probability distribution on the deterministic decision trees that compute f. The time is measured by averaging. The randomized decision tree complexity, $RC(f)$, is the maximum (over all input settings) average (according to the distribution) number of variables read in the best distribution.

Clearly $RC(f) \leq DC(f)$. It is also well known that these two measures satisfy $RC(f) \geq [DC(f)]^{0.5}$ for every Boolean function f. (This was proved independently by several people, none of them published it. Their proof is quoted in [?].) Snir [?] showed that for g, the complete binary AND/OR-tree function, $RC(g) = O([DC(g)]^{0.753...})$. No Boolean function with lower randomized complexity is known. In contrast, no lower bound on $RC(f)$ higher than $[DC(f)]^{0.5}$ is known to hold for all Boolean functions.

While it is possible to obtain strong lower bounds on $RC(f)$ for special functions f, an important step would be to beat the 0.5 exponent for interesting families of Boolean functions. This direction was initiated by Yao [?] for the class of monotone graph properties. Yao's results were improved by King [?] and then by Hajnal [?], who obtained the current best lower bound of 2/3 in the exponent.

In this work we beat the 0.5 exponent for the class of *read-once* Boolean functions, a class that was studied in several contexts (e.g., game tree evaluation [?], [?], amplifying approximating circuits [?], [?] and learning [?]). A Boolean function is called *read-once* if it can be represented by a Boolean formula in which each input variable appears exactly once. A simple adversary argument shows that for every read-once function f, $DC(f) = n$, where n is the number of input variables. The known 0.5 lower bound states that the randomized decision tree complexity of every read-once function satisfies $RC(f) \geq \sqrt{n}$. Our main result improves this fact as follows.

THEOREM 1.1. *There exists a real number $\theta > 0.5$ such that for every read-once Boolean function f, $RC(f) \geq n^{\theta}$, where n is the number of input variables f depends on. In fact, $\theta = 0.51$ satisfies this inequality.*

The proof of Theorem 1.1 combines a generalized version of the Saks-Wigderson lower bound for the randomized decision tree complexity of read-once functions [?] with restriction arguments. In Section 2 we quote their lower bound and give a generalized version of it which is used for proving the main theorem. In that section we also point out that their lower bound is too weak to imply Theorem 1.1 (referring to Section 6), and therefore, some additional tool, e.g., restrictions, is necessary. Section 3 introduces some intuition of the main theorem's proof. Section 4 contains some inequalities on real valued functions used in the proof. Section 5 contains the rest of the proof — that part which involves restrictions.

A problem which is closely related to evaluating read-once functions by decision trees is that of evaluating two-person zero-sum game trees using α-β pruning [**?**]. Imagine a chess program making a decision about the next move in a given board position. It develops a partial tree of possible moves alternating between it and its opponent. This results in new game positions at the leaves. Each such position is assigned a real value from which the value of the root can be computed (hence also the best move under this information). This is a generalization of read-once functions to real valued inputs and MIN/MAX gates.

The standard α-β pruning algorithm evaluates this tree by evaluating *some* of the leaf positions, and is charged according to the number of leaves evaluated. Its deterministic and randomized versions are analogous to deterministic and randomized decision trees. Our theorem has as an immediate consequence a lower bound on the number of positions that have to be evaluated in any two-person game tree. Nontrivial lower bounds for this problem were previously known only for very special game trees [**?**], [**?**].

Corollary 1.2. *Any randomized α-β pruning algorithm computing the value of any two-person game tree with n final positions evaluates at least $n^{0.51}$ positions for the worst case leaf values.*

2. Preparations

Let us first quote the Saks-Wigderson lower bound on the randomized decision tree complexity of read-once Boolean formulae.

Definition 2.1. *[**?**] For a read-once Boolean formula f with AND/OR gates of fan-in 2, recursively define $R_0(f)$ and $R_1(f)$ ("lower bounds on the randomized decision tree complexities for the zeros and ones of f") by*

$$R_0(f) = \begin{cases} 1 & \text{if } f \text{ is a single variable,} \\ R_0(g) + R_0(h), & \text{if } f = g \vee h, \\ \chi(R_0(g), R_1(g), R_0(h), R_1(h)) & \text{if } f = g \wedge h, \end{cases}$$
$$R_1(f) = \begin{cases} 1 & \text{if } f \text{ is a single variable,} \\ R_1(g) + R_1(h), & \text{if } f = g \wedge h, \\ \chi(R_1(g), R_0(g), R_1(h), R_0(h)) & \text{if } f = g \vee h. \end{cases}$$

Here, for positive real numbers a, b, c and d,

$$\chi(a, b, c, d) = \min\{a + d, b + c, \frac{ab + cd + bd}{b + d}\}.$$

THEOREM 2.2. *[**?**] Let f be a read-once Boolean formula with AND/OR gates of fan-in 2. Then $RC(f) \geq R_*(f)$ where $R_*(f) = \max\{R_0(f), R_1(f)\}$.*

Using this lower-bound, it was shown in [**?**] that the complete AND/OR-tree function f satisfies also $RC(f) \geq R_*(f) = \Omega(n^{0.753\ldots})$, meeting Snir's upper-bound. For our purpose this lower bound has a weakness though. Section 6 describes an infinite family $\{F_n\}$ of read-once Boolean functions depending on n variables. That family serves to show there that the Saks-Wigderson lower bound is too weak to imply a result of the type of Theorem 1.1, since $R_*(F_n) = O(n^\theta)$ for every $\theta > \frac{1}{2}$ (see Theorem 6.4). However, a simple restriction consideration shows that F_n has high randomized decision tree complexity (Theorem 6.3). This suggests that the proof of Theorem 1.1 should involve restriction arguments. It appears also that formulae with gates of fan-in greater than two should be considered in a more delicate way than simply replacing them by binary trees and then applying the Saks-Wigderson lower bound.

The next theorem generalizes Theorem 2.2 to arbitrary fan-in gates. Note that R_0, R_1 and χ of the following definition specialize to those of Definition 2.1 when the gates are binary.

DEFINITION 2.3. *For a read-once Boolean formula f with AND/OR gates of arbitrary fan-in, recursively define the (lower bound) terms $R_0(f)$ and $R_1(f)$*

by

$$R_0(f) = \begin{cases} 1 \quad \text{if } f \text{ is a single variable,} & \\ \sum_{i=1}^k R_0(g_i) & \text{if } f = \vee_{i=1}^k g_i, \\ \chi(R_0(g_1), R_1(g_1), R_0(g_2), R_1(g_2), ..., R_0(g_k), R_1(g_k)) & \text{if } f = \wedge_{i=1}^k g_i, \end{cases}$$
$$R_1(f) = \begin{cases} 1 \quad \text{if } f \text{ is a single variable,} & \\ \sum_{i=1}^k R_1(g_i) & \text{if } f = \wedge_{i=1}^k g_i, \\ \chi(R_1(g_1), R_0(g_1), R_1(g_2), R_0(g_2), ..., R_1(g_k), R_0(g_k)) & \text{if } f = \vee_{i=1}^k g_i. \end{cases}$$

Here g_i, $1 \le i \le k$, is a Boolean function, and for positive real numbers a_i and b_i, $1 \le i \le k$,
$\chi(a_1, b_1, a_2, b_2, ..., a_k, b_k) =$

$$\min_{\emptyset \neq T = \{i_1, ..., i_t\} \subseteq \{1,2,...,k\}} [\sum_{i \notin T} b_i \; + \; \phi(a_{i_1}, b_{i_1}, a_{i_2}, b_{i_2}, ..., a_{i_t}, b_{i_t})],$$

and in the expression for χ,

$$\phi(a_1, b_1, a_2, b_2, ..., a_t, b_t) = (\sum_{j=1}^t a_j b_j + \sum_{1 \le j < h \le t} b_j b_h) / \sum_{j=1}^t b_j.$$

Theorem 2.4. *Let f be a read-once Boolean formula with AND/OR gates of arbitrary fan-in. Then $RC(f) \ge R_*(f)$ where $R_*(f) = \max\{R_0(f), R_1(f)\}$.*

The proof follows from that of Theorem 2.2 (see [?]) by treating arbitrary fan-in gates similarly to the way such gates were treated in the lower bound proof of [?]. A detailed proof appears in [?]. A corollary of Theorem 2.4 is the following simpler version which will be used in the sequel.

Definition 2.5. *For a read-once Boolean formula f with AND/OR gates of arbitrary fan-in, recursively define the (lower bound) term $R(f)$ by*

$$R(f) = \begin{cases} 1 & \text{if } f \text{ is a single variable,} \\ \psi(R(g_1), ..., R(g_k)) & \text{if } f = \vee_{i=1}^k g_i \text{ or } f = \wedge_{i=1}^k g_i. \end{cases}$$

Here, for positive real numbers a_i, $1 \le i \le k$,

$$\psi(a_1, ..., a_k) = (\sum_{1 \le i \le j \le k} a_i a_j) / (\sum_{i=1}^k a_i).$$

Theorem 2.6. *Let f be a read-once Boolean formula with AND/OR gates of arbitrary fan-in. Then $RC(f) \geq R(f)$.*

Proof. We show by induction on the structure of f that $R_0(f) \geq R(f)$ and $R_1(f) \geq R(f)$. The base case is trivial. For $f = \vee_{i=1}^{k}(g_i)$ (the $\wedge$-case is dual) we have

$$\begin{aligned} R_0(f) &= \sum R_0(g_i) \geq \sum R(g_i) \geq \psi(R(g_1),...,R(g_k)) = R(f), \text{ and} \\ R_1(f) &= \chi(R_1(g_1),R_0(g_1),R_1(g_2),R_0(g_2),...,R_1(g_k),R_0(g_k)) \\ &\geq \chi(R(g_1),R_0(g_1),R(g_2),R_0(g_2),...,R(g_k),R_0(g_k)) \\ &\geq \chi(R(g_1),R(g_1),R(g_2),R(g_2),...,R(g_k),R(g_k)) \\ &= \phi(R(g_1),R(g_1),R(g_2),R(g_2),...,R(g_k),R(g_k)) \\ &= \psi(R(g_1),R(g_2),...,R(g_k)) = R(f). \end{aligned}$$

We used here the facts that (1) for $a_i = b_i \geq 0,\ i = 1,...,k$ the minimum term in χ is when $T = \{1,...,k\}$, (2) for $a_i, b_i \geq 0$, ϕ (and therefore also χ) are monotone non-decreasing in each a_i, and (3) for $0 \leq a_i \leq b_i$, ϕ and χ are monotone non-decreasing in each b_i. For example, the numerator of $\frac{\partial\phi}{\partial b_1}(a_1,b_1,...,a_t,b_t)$ is

$$\begin{aligned} &(a_1 + \textstyle\sum_{h=2}^{t} b_h)\cdot(\sum_{j=1}^{t} b_j) - \sum_{j=1}^{t} a_j b_j - \sum_{1\leq j<h\leq t} b_j b_h &\geq \\ &(\textstyle\sum_{h=2}^{t} b_h)\cdot(\sum_{j=1}^{t} b_j) - \sum_{j=2}^{t} b_j^2 - \sum_{1\leq j<h\leq t} b_j b_h &\geq 0. \ \square \end{aligned}$$

The recursion in Definition 2.5 is quite simple. It has only a single R (replacing the two R_0 and R_1 of Definitions 2.1 and 2.3) and it contains no min operator. In addition, the function ψ it involves has the following properties.

Fact 2.7.
(i) ψ is homogeneous: $\psi(\alpha a_1,...,\alpha a_k) = \alpha\psi(a_1,...,a_k)$,
(ii) ψ is monotone non-decreasing in $\mathbb{R}_k^+$ in each of its variables.

Proof.

(i) $$\psi(\alpha a_1,...,\alpha a_k) = \frac{\sum_{1\leq i\leq j\leq k} \alpha a_i \alpha a_j}{\sum_{i=1}^{k} \alpha a_i} = \alpha\frac{\sum_{1\leq i\leq j\leq k} a_i a_j}{\sum_{i=1}^{k} a_i} = \alpha\psi(a_1,...,a_k).$$

(ii) By symmetry we show monotonicity in, say, the first argument of ψ.

$$\frac{\partial\psi}{\partial a_1}(a_1,...,a_k) = \frac{(2a_1 + \sum_{i=2}^{k} a_i)\cdot(\sum_{i=1}^{k} a_i) - \sum_{1\leq i\leq j\leq k} a_i a_j}{(\sum_{i=1}^{k} a_i)^2} \geq 0$$

for positive a_i's. $\square$

Remark 2.8. *When focusing on the lower bound R of Theorem 2.6, one gains, indeed, the simplicity which helps prove Theorem 1.1, but one also loses the ability to get an optimal lower bound. For g, the complete AND/OR-tree function,* $R(g) = \frac{3}{2}^{\log n} = n^{0.58\ldots}$*, and restrictions cannot help here. While it is clear from our proof that* $\theta = 0.51$ *is not the highest exponent this proof technique can provide, this example shows that 0.58... is an upper bound for any proof based on Theorem 2.6.*

3. Proof of the Main Theorem: Intuition

To prove Theorem 1.1, we will actually show that $R(f) \geq n^\theta$, where $R(f)$ is the lower bound for $RC(f)$ given in Definition 2.5 and Theorem 2.6. A natural way to show this is by induction on the structure of f. (Note we do not distinguish between the function and its read-once formula.) Let $f = g \diamondsuit h$, where $\diamondsuit \in \{\wedge, \vee\}$. Say h is of size αn for some $0 < \alpha \leq 0.5$, and g is of size $(1-\alpha)n$, where by the size of a function we mean the number of its input variables. We want to show that

$$R(f) \geq n^\theta.$$

By definition,

$$R(f) = \psi(R(g), R(h)).$$

By induction, we may assume that $R(h) \geq (\alpha n)^\theta$ and $R(g) \geq ((1-\alpha)n)^\theta$. Using the monotonicity of ψ, Fact 2.7, we have

$$\psi(R(g), R(h)) \geq \psi((\alpha n)^\theta, ((1-\alpha)n)^\theta).$$

Therefore we have to check whether

$$\psi((\alpha n)^\theta, ((1-\alpha)n)^\theta) \overset{??}{\geq} n^\theta,$$

or, equivalently, by the fact that ψ is homogeneous (Fact 2.7), to check whether

$$\psi(\alpha^\theta, (1-\alpha)^\theta) \overset{??}{\geq} 1.$$

However, this is not true. For every $\theta > \frac{1}{2}$, if f is very unbalanced, that is, if α is very small, $\psi(\alpha^\theta, (1-\alpha)^\theta) < 1$. This should not be a surprise — Theorem 6.4 implies that this naive approach must fail. However, Theorem 6.3 hints that we should use restrictions. Yet, we cannot simply restrict f so as to eliminate the smaller subfunction h. This way we might end up eliminating almost everything. Therefore we have to look deeper in the structure of f. We analyze this structure, and carefully restrict small parts of it. Some more details follow.

We look at g, the larger sub-function of f. If g is also very unbalanced we look at its larger sub-function too, and so on. We stop when the small sub-functions we have seen in this process accumulate together "sufficiently large" total size. Now, each such small sub-function is a child of some gate, either AND or OR. We count the total size of the small sub-functions under AND gates and compare it to the total size of those under OR gates. Say the latter is smaller. In this case we restrict f so as to eliminate exactly those small sub-functions under OR gates, leaving all other parts of f as is. We view the restricted function as a single AND, possibly of large fan-in, whose children are all the sub-functions that remain alive. On this restricted function we want to evaluate ψ. However, this is still not enough. The function ψ may still be smaller than 1 when evaluated at the θ-powers of the remaining sub-functions' sizes. This happens if the two total sizes were approximately the same.

In order to have a valid argument we distinguish between three cases of how the total size of the small sub-functions is distributed among them. These are the three cases which appear in Section 5. We then restrict f in a specific way for each case, and use a corresponding and specific inequality involving ψ. The three corresponding inequalities are stated as Lemma 4.2 in the next section.

4. A Few Inequalities

The following fact is used in the proof of the three inequalities of Lemma 4.2.

FACT 4.1.
(i) Let $0 < \theta \le 1$ *and* $a > 0$. *Then* $(1-\varepsilon)^\theta + (a\varepsilon)^\theta$ *is concave in the segment* $0 \le \varepsilon \le 1$ *and maximal at* $\varepsilon = 1/(1 + a^{\frac{\theta}{\theta-1}})$.
(ii) Let $0 < \theta \le 1$ *and* $0 \le b \le 1$. *Then* $(1-\varepsilon)^\theta + (\varepsilon - b)^\theta$ *is concave in the segment* $b \le \varepsilon \le 1$ *and maximal at* $\varepsilon = \frac{1+b}{2}$.
(iii) Let $0 < \theta \le 0.6$ *and* $0 \le b \le 1$. *Then* $(1-\varepsilon)^{2\theta} + (1-\varepsilon)^\theta(\varepsilon-b)^\theta + (\varepsilon-b)^{2\theta}$ *is concave in the segment* $b \le \varepsilon \le 1$ *and maximal at* $\varepsilon = \frac{1+b}{2}$.
(iv) Let $0 < \theta < 1$ *and* $0 \le \varepsilon \le 1$. *Then* $1 - (1-\varepsilon)^\theta \le \varepsilon\theta(1-\varepsilon)^{\theta-1}$.

PROOF. We prove (i) and (ii) together: Let $f(\varepsilon) = (1-\varepsilon)^\theta + [a(\varepsilon-b)]^\theta$. Then $\frac{\partial f}{\partial \varepsilon} = -\theta(1-\varepsilon)^{\theta-1} + a^\theta\theta(\varepsilon-b)^{\theta-1}$. This equals 0 if and only if $1-\varepsilon = a^{(\frac{\theta}{\theta-1})}(\varepsilon-b)$, i.e., $\varepsilon = (1 + ba^{\frac{\theta}{\theta-1}})/(1 + a^{\frac{\theta}{\theta-1}})$. Also,

$$\frac{\partial^2 f}{\partial \varepsilon^2} = \theta(\theta-1)(1-\varepsilon)^{\theta-2} + a^\theta \theta(\theta-1)(\varepsilon-b)^{\theta-2} \leq 0.$$

(iii): Consider $g(\varepsilon) = (1-\varepsilon)^{2\theta} + (\varepsilon-b)^{2\theta} + (1-\varepsilon)^\theta(\varepsilon-b)^\theta$. Then

$$\frac{\partial g}{\partial \varepsilon} = -2\theta(1-\varepsilon)^{2\theta-1} + 2\theta(\varepsilon-b)^{2\theta-1} - \theta(1-\varepsilon)^{\theta-1}(\varepsilon-b)^\theta + \theta(1-\varepsilon)^\theta(\varepsilon-b)^{\theta-1}.$$

First, for $b \leq \varepsilon \leq 1$, $\frac{\partial g}{\partial \varepsilon} = 0$ if and only if $1-\varepsilon = \varepsilon - b$, that is $\varepsilon = \frac{1+b}{2}$. Next,

$$\begin{array}{rcl} \frac{1}{\theta} \cdot \frac{\partial^2 g}{\partial \varepsilon^2} & = & 2(2\theta-1)(1-\varepsilon)^{2\theta-2} + 2(2\theta-1)(\varepsilon-b)^{2\theta-2} \\ & + & (\theta-1)(1-\varepsilon)^{\theta-2}(\varepsilon-b)^\theta - 2\theta(1-\varepsilon)^{\theta-1}(\varepsilon-b)^{\theta-1} \\ & + & (\theta-1)(1-\varepsilon)^\theta(\varepsilon-b)^{\theta-2} \\ & \leq & 2(2\theta-1)[(1-\varepsilon)^{2\theta-2} + (\varepsilon-b)^{2\theta-2}] \\ & + & (\theta-1)[(1-\varepsilon)^{\theta-2}(\varepsilon-b)^\theta + (1-\varepsilon)^\theta(\varepsilon-b)^{\theta-2}]. \end{array}$$

Now, $2(2\theta-1) \leq -(\theta-1)$ for $\theta \leq 0.6$, and so to show $\frac{\partial^2 g}{\partial \varepsilon^2} < 0$ we show

$$(1-\varepsilon)^{2\theta-2} + (\varepsilon-b)^{2\theta-2} \leq (1-\varepsilon)^{\theta-2}(\varepsilon-b)^\theta + (1-\varepsilon)^\theta(\varepsilon-b)^{\theta-2},$$

$$\text{i.e.,} \quad 0 \leq [(\varepsilon-b)^{\theta-2} - (1-\varepsilon)^{\theta-2}] \cdot [(1-\varepsilon)^\theta - (\varepsilon-b)^\theta].$$

Indeed, if $1-\varepsilon \geq \varepsilon - b$ the two multiplied terms are non negative, and if $1-\varepsilon \leq \varepsilon - b$ they are both non positive.
(iv): Consider $h(\varepsilon) = -(1-\varepsilon)^\theta$. Then $1-(1-\varepsilon)^\theta = h(\varepsilon) - h(0) = \varepsilon \frac{\partial h}{\partial \varepsilon}(\varepsilon')$ for some $0 \leq \varepsilon' \leq \varepsilon$, and we have $\frac{\partial h}{\partial \varepsilon}(\varepsilon') = \theta(1-\varepsilon')^{\theta-1} \leq \theta(1-\varepsilon)^{\theta-1}$. □

In Section 5 we show that Theorem 1.1 holds with $\theta = 0.51$. We use also the parameters $b = 0.001$, $c = 0.025$ and $\delta = 0.4$, and the following three inequalities.

Lemma 4.2. *Let* $\theta = 0.51$, $b = 0.001$, $c = 0.025$ *and* $\delta = 0.4$. *Then*

$$\psi((1-\varepsilon)^\theta, (\varepsilon-\varepsilon')^\theta) \geq 1 \quad \textit{for} \;\; c \leq \varepsilon \leq \frac{1+b}{2},\; 0 \leq \varepsilon' \leq b, \tag{4.1}$$

$$\psi((1-\varepsilon)^\theta, \varepsilon'^\theta) \geq 1 \quad \textit{for} \;\; \frac{1-\delta}{2-\delta}\delta b \leq \varepsilon \leq c \;\; \textit{and} \;\; \frac{1}{2-\delta}\varepsilon \leq \varepsilon' \leq \varepsilon, \tag{4.2}$$

$$\psi((1-\varepsilon)^\theta, \varepsilon'^\theta, \varepsilon''^\theta) \geq 1 \quad \textit{for} \;\; b \leq \varepsilon \leq c,\; (1-2\delta)\frac{\varepsilon}{2} \leq \varepsilon' \leq \varepsilon'',\; \varepsilon'+\varepsilon'' \geq \frac{\varepsilon}{2}. \tag{4.3}$$

PROOF. Using the monotonicity of ψ in each of its arguments (Fact 2.7), and plugging $\delta = 0.4$ we can replace the lemma by the following statement.

Let $\theta = 0.51$, $b = 0.001$ and $c = 0.025$. Then

$$\psi((1-\varepsilon)^\theta, (\varepsilon - b)^\theta) \geq 1 \quad \text{for } \ c \leq \varepsilon \leq \frac{1+b}{2}, \tag{4.4}$$

$$\psi((1-\varepsilon)^\theta, (\frac{5}{8}\varepsilon)^\theta) \geq 1 \quad \text{for } \ 0.15b \leq \varepsilon \leq c, \tag{4.5}$$

$$\psi((1-\varepsilon)^\theta, \varepsilon'^\theta, \varepsilon''^\theta) \geq 1 \quad \text{for } \ b \leq \varepsilon \leq c, \ 0.1\varepsilon \leq \varepsilon' \leq \varepsilon'', \ \varepsilon' + \varepsilon'' = \frac{\varepsilon}{2}. \tag{4.6}$$

These three inequalities involve parameters that range over the continuous domain. In the following we reduce each of the three to a single inequality or to a finite (and small) set of inequalities that each involves concrete numbers only. The check of each concrete inequality was done twice, once by a computer program and once by a "pocket" calculator.

We first prove (4.6). Denote $\varepsilon' = \rho\varepsilon$. Then $0.1 \leq \rho \leq 0.25$, $\varepsilon'' = (\frac{1}{2} - \rho)\varepsilon$ and

$$\begin{aligned} \psi((1-\varepsilon)^\theta, \varepsilon'^\theta, \varepsilon''^\theta) &= (1-\varepsilon)^\theta + \tfrac{(\varepsilon')^{2\theta} + (\varepsilon'\varepsilon'')^\theta + (\varepsilon'')^{2\theta}}{(1-\varepsilon)^\theta + (\varepsilon')^\theta + (\varepsilon'')^\theta} \\ &= (1-\varepsilon)^\theta + \tfrac{[\rho^{2\theta} + \rho^\theta(\frac{1}{2}-\rho)^\theta + (\frac{1}{2}-\rho)^{2\theta}]\varepsilon^{2\theta}}{(1-\varepsilon)^\theta + [\rho^\theta + (\frac{1}{2}-\rho)^\theta]\varepsilon^\theta}. \end{aligned}$$

We have to show that

$$(1-\varepsilon)^\theta + \frac{[\rho^{2\theta} + \rho^\theta(\frac{1}{2}-\rho)^\theta + (\frac{1}{2}-\rho)^{2\theta}]\varepsilon^{2\theta}}{(1-\varepsilon)^\theta + [\rho^\theta + (\frac{1}{2}-\rho)^\theta]\varepsilon^\theta} \geq 1 \ \text{ for } \ 0.1 \leq \rho \leq 0.25.$$

By Fact 4.1, $\rho^{2\theta} + \rho^\theta(\frac{1}{2}-\rho)^\theta + (\frac{1}{2}-\rho)^{2\theta} \geq 0.1^{1.02} + 0.04^{0.51} + 0.4^{1.02} > 0.68$, and $\rho^\theta + (\frac{1}{2}-\rho)^\theta \leq 2 \cdot 0.25^{0.51} < 1$. So it is enough to show that

$$0.68\varepsilon^{2\theta} \geq [1 - (1-\varepsilon)^\theta] \cdot [(1-\varepsilon)^\theta + \varepsilon^\theta].$$

Again, by Fact 4.1, $(1-\varepsilon)^\theta + \varepsilon^\theta \leq 0.975^{0.51} + 0.025^{0.51} < 1.14$ since $\varepsilon \leq c = 0.025$. Also $1 - (1-\varepsilon)^\theta \leq \varepsilon\theta(1-\varepsilon)^{\theta-1} \leq \varepsilon\theta \cdot 0.975^{-0.49} < \varepsilon\theta \cdot 1.013$. So it is enough to show that

$$\varepsilon^{2\theta-1} \geq \frac{0.51 \cdot 1.013 \cdot 1.14}{0.68}.$$

Indeed, $0.51 \cdot 1.013 \cdot 1.14 < 0.59$ and $\varepsilon \geq b = 0.001 > (\frac{0.59}{0.68})^{50}$.

We now prove (4.5) which states

$$(1-\varepsilon)^\theta + \frac{(0.625 \cdot \varepsilon)^{2\theta}}{(1-\varepsilon)^\theta + (0.625 \cdot \varepsilon)^\theta} \geq 1.$$

We use Fact 4.1 (iv) again, $1-(1-\varepsilon)^\theta \le \varepsilon\theta(1-\varepsilon)^{\theta-1}$, and show that

$$\varepsilon^{2\theta-1} \ge \frac{\theta(1-\varepsilon)^{\theta-1}}{0.625^{2\theta}} \cdot [(1-\varepsilon)^\theta + (0.625\cdot\varepsilon)^\theta].$$

The left hand side is monotonically increasing in ε, as is the right hand side (see part (i) of Fact 4.1). Therefore, to prove this last inequality for $0.15b \le \varepsilon \le c$, it is enough to show a sequence $0.15b = \varepsilon_0 < \varepsilon_1 < ... < \varepsilon_k = c$ such that for each i, $1 \le i \le k$, the left hand side with $\varepsilon = \varepsilon_{i-1}$ is greater than the right hand side with $\varepsilon = \varepsilon_i$,

$$\varepsilon_{i-1}^{2\theta-1} \ge \frac{\theta(1-\varepsilon_i)^{\theta-1}}{0.625^{2\theta}} \cdot [(1-\varepsilon_i)^\theta + (0.625\cdot\varepsilon_i)^\theta].$$

Indeed, the sequence $0.00015,\ 0.0006,\ 0.0034,\ 0.012,\ 0.021,\ 0.025$ satisfies this.

Finally, we prove (4.4), that is,

$$(1-\varepsilon)^{2\theta} + (1-\varepsilon)^\theta(\varepsilon-0.001)^\theta + (\varepsilon-0.001)^{2\theta} \ge (1-\varepsilon)^\theta + (\varepsilon-0.001)^\theta$$

for $c \le \varepsilon \le \frac{1+b}{2}$. By Fact 4.1 both sides of the inequality are monotone, and so, as before, we show a sequence $c = \varepsilon_0 < \varepsilon_1 < ... < \varepsilon_k = \frac{1+b}{2}$ such that for each i, $1 \le i \le k$, the left hand side with $\varepsilon = \varepsilon_{i-1}$ is greater than the right hand side with $\varepsilon = \varepsilon_i$. A sequence satisfying this is 0.025, 0.026, 0.027, ..., 0.068, 0.069, 0.07, 0.08, 0.09, ..., 0.18, 0.19, 0.2, 0.3, 0.5005, where the first ... corresponds to jumps of 0.001 and the second corresponds to jumps of 0.01. □

5. Proof of the Main Theorem: Restrictions

We now prove Theorem 1.1. As mentioned, we show that it holds with $\theta = 0.51$. We use also the parameters $b = 0.001$, $c = 0.025$ and $\delta = 0.4$, as in Lemma 4.2.

Proof. We prove by induction on n, the number of f's input variables, that

$$R(f) \ge n^\theta.$$

This suffices since R, which is defined in Definition 2.5, is by Theorem 2.6 a lower bound for RC. The case $n = 1$ is trivial. Consider $n > 1$. Assume that the formula contains no negation gates. Otherwise "sink" them and replace negated inputs by new positive variables. Assume also that all gates have fan-in two to begin with. Otherwise replace "wide" gates by binary trees.

Furthermore, consider the formula as a binary tree in which the left subtree of each node is at least as large as its right subtree with respect to the number of leaves they possess. This tree, denoted T, can be represented as

$$T = (\ (\ (T_* \diamondsuit_t T_t) ... \diamondsuit_i T_i) ... \diamondsuit_2 T_2) \diamondsuit_1 T_1,$$

where each $\diamondsuit_i$ is some gate, either AND or OR, and the T_i's and T_* are subtrees of T. Fix a specific such representation for which the two conditions below hold with the following notation.

NOTATION 5.1. *Let n_i (respectively n_*) be the number of leaves in the subtree T_i (respectively T_*). Define $\varepsilon_i = \frac{n_i}{n}$ and $\varepsilon_* = \frac{n_*}{n}$ $(= 1 - \sum \varepsilon_i)$.*

CONDITIONS 5.2. *(1) For each i, $\sum_{j=i+1}^{t} \varepsilon_j + \varepsilon_* \geq \varepsilon_i$ (the above mentioned requirement on the number of leaves in left- and right-subtrees).*
(2) The number t is minimal with $\sum_{i=1}^{t} \varepsilon_i \geq b$ (which specifies the last gate explicitly represented).

It is clear that such representation exists.

The following notation will also be used.

NOTATION 5.3. *For each i, let $S_i = \sum_{j=1}^{i} \varepsilon_j$, let $I_i^{\wedge} = \{j :\ 1 \leq j \leq i,\ \diamondsuit_j = \wedge\}$ and let $S_i^{\wedge} = \sum_{j \in I_i^{\wedge}} \varepsilon_j$. Similarly define $I_i^{\vee}$ and $S_i^{\vee}$. Note that $S_i = S_i^{\wedge} + S_i^{\vee}$ and that $S_t \geq b$. Also, for a subset $I \subseteq \{1, ..., t\}$ denote $S_I = \sum_{i \in I} \varepsilon_i$.*

Distinguish between the following three possible cases.
Case 1: $S_t \geq c$. (T_t is "very" large with respect to the other T_i's.)
Case 2: $b \leq S_t < c$ and $\varepsilon_u > \delta S_t$ for some u, $1 \leq u \leq t$. (No T_i is "very" large but one of them is "relatively" large.)
Case 3: $b \leq S_t < c$ and $\varepsilon_u \leq \delta S_t$ for all u, $1 \leq u \leq t$. (All T_i's are "small".)

In **case 1** we argue that all the other T_i's can be ignored: Consider $T_* \diamondsuit_t T_t$ which is a restriction of T. Clearly, $R(T) \geq R(T_* \diamondsuit_t T_t)$. By Definition 2.5 and by the induction hypothesis

$$R(T_* \diamondsuit_t T_t) = \psi(R(T_*), R(T_t)) \geq \psi(n_*^{\theta}, n_t^{\theta}).$$

To show that $\psi(n_*^{\theta}, n_t^{\theta}) \geq n^{\theta}$, we use the homogeneity of ψ (Fact 2.7), and show that $\psi(\varepsilon_*^{\theta}, \varepsilon_t^{\theta}) \geq 1$, i.e.,

$$\psi((1 - S_t)^{\theta}, (S_t - S_{t-1})^{\theta}) \geq 1.$$

Recall that $S_{t-1} < b$ and that $\varepsilon_t \le \frac{1}{2}(1 - S_{t-1})$ or $c \le S_t \le S_{t-1} + \frac{1-S_{t-1}}{2} = \frac{1+S_{t-1}}{2} \le \frac{1+b}{2}$. Case 1 follows from the first inequality of Lemma 4.2 with $\varepsilon = S_t$ and $\varepsilon' = S_{t-1}$.

Consider now **case 2**. The following fact states that for some prefix of the leftmost path of nonnegligible size, the total size of subtrees entering its AND gates significantly differs from that of those entering its OR gates. In this case we argue that the subtrees of smaller total size can be ignored.

FACT 5.4. *There exists an index $r \in \{u-1, u\}$ such that $S_r^{\diamondsuit_r} \ge \frac{1}{2-\delta} S_r$ and $S_r \ge \frac{1-\delta}{2-\delta}\delta b$.*

PROOF. Assume without loss of generality that $\diamondsuit_u = \wedge$. If $S_u^{\wedge} \ge \frac{1}{2-\delta} S_u$ then $r = u$ fulfills the fact since $S_u \ge \varepsilon_u \ge \delta S_t \ge \delta b$. So suppose $S_u^{\vee} (= S_{u-1}^{\vee}) \ge \frac{1-\delta}{2-\delta} S_u$. In this case

$$\frac{S_{u-1}^{\vee}}{S_{u-1}} = \frac{S_{u-1}^{\vee}}{S_u - \varepsilon_u} \ge \frac{S_{u-1}^{\vee}}{S_u - \delta S_t} \ge \frac{\frac{1-\delta}{2-\delta} S_u}{S_u - \delta S_u} = \frac{1}{2-\delta},$$

so we choose $\diamondsuit = \vee$, $r = u - 1$ and we are done since

$$S_{u-1} \ge S_{u-1}^{\vee} \ge \frac{1-\delta}{2-\delta} S_u \ge \frac{1-\delta}{2-\delta} \varepsilon_u \ge \frac{1-\delta}{2-\delta} \delta S_t \ge \frac{1-\delta}{2-\delta} \delta b. \ \square$$

Suppose, for the purpose of simplifying notation, that for the index r given in Fact 5.4, $\diamondsuit_r = \wedge$. Restrict the given formula so that each T_i with $i < r$ and with $\diamondsuit_i = \vee$ becomes '0'. Represent the remaining tree in the form

$$T_{\Box} \wedge T_{\triangle},$$

where $T_{\Box} = (...(T^* \diamondsuit_t T_t)...\diamondsuit_{r+1} T_{r+1})$ and $T_{\triangle} = \wedge_{i \in I_r^{\wedge}} T_i$. As in case 1, using the induction hypothesis and the homogeneity of ψ, it is sufficient to show

$$\psi((1 - S_r)^{\theta}, (S_r^{\wedge})^{\theta}) \ge 1.$$

This follows from the second inequality of Lemma 4.2 with $\varepsilon = S_r$ and $\varepsilon' = S_r^{\wedge}$.

Finally consider **case (3)**. Assume without loss of generality that $S_t^{\wedge} \ge S_t^{\vee}$, i.e., $S_t^{\wedge} \ge \frac{S_t}{2}$. We argue that one can ignore also in this case the subtrees entering OR gates. However one has to be more careful this time and use the fact that the subtrees entering the AND gates are small and hence can be partitioned into two sets of similar total sizes. The following fact states that there exists a partition of $I_t^{\wedge}$ into two subsets which are "fairly" balanced with respect to the total numbers of inputs in the subformula each represents.

Fact 5.5. *There exists a subset* $I \subset I_t^{\wedge}$ *such that* $S_{I_t^{\wedge}\setminus I} \geq S_I \geq (1-2\delta)\cdot\frac{S_t}{2}$.

Proof. Use the following fact with $\beta = 2\delta = 0.8$ and with some scale changing. $\square$

Fact 5.6. *Let* $0 \leq \beta \leq 1$ *and* $a_1, ..., a_k$ *such that* $0 \leq a_i \leq \beta$ *and* $\sum_{i=1}^{k} a_i \geq 1$. *Then there exists a set* $I \subseteq \{1, ..., k\}$ *such that* $\min\{\sum_{i\in I} a_i, \sum_{i\notin I} a_i\} \geq \gamma(\beta)$ *where*

$$\gamma(\beta) = \begin{cases} \frac{1-\beta}{2} & \text{if } \beta \leq \frac{1}{3}, \\ \frac{1}{3} & \text{if } \frac{1}{3} \leq \beta \leq \frac{2}{3}, \\ 1-\beta & \text{if } \beta \geq \frac{2}{3}. \end{cases}$$

Proof. If $\beta \leq \frac{1}{3}$ then any partition with sums that differ by more than β can be improved by moving some a_i from the larger part to the other. If there is some part larger than $\frac{1}{3}$ put that part in one subset and all the others in the other subset. This covers the two last cases. $\square$

Now, consider the subset I given by Fact 5.5 and restrict T to

$$T_* \wedge T_\square \wedge T_\triangle,$$

where $T_\square = \wedge_{i\in I} T_i$ and $T_\triangle = \wedge_{i\in I_t^{\wedge}\setminus I} T_i$. The third inequality of Lemma 4.2 with $\varepsilon = S_t$, $\varepsilon' = S_I$ and $\varepsilon'' = S_{I_t^{\wedge}\setminus I}$, shows that $\psi(\varepsilon_*^{\theta}, S_I^{\theta}, S_{I_t^{\wedge}\setminus I}^{\theta}) \geq 1$. This completes the third case (similarly to the previous cases) and therefore completes the proof of the main theorem. $\square$

6. The need for restrictions

We define a family $\{F_m\}$ of read-once formulae for which the Saks-Wigderson lower bound does not yield a good bound. The idea is that the formulae are very unbalanced: Each OR gate has a single variable as one child and a "large" formula as the other child. The definition is recursive. Given F_m we AND it with another copy of F_m and OR the result with a single new variable. Then we AND the result again with a copy of F_m and OR with another new variable. We repeat this "ANDing" and "ORing" m times, and call the resulting formula F_{m+1}.

The analysis deals with the terms R_0 and R_1, given in Definition 2.1 and used in the Saks-Wigderson lower bound, Theorem 2.2. Intuitively, the non-balancing of any AND gate causes its R_0 to be small relative to its R_1. Then

"ORing" this with a single variable causes the new R_0 and R_1 to be equal, and only slightly greater than the R_0 at the AND gate, no matter how great R_1 at that AND gate was.

Following is a formal definition of these formulae. First we define $F_{m,i}$, the prefix of length i of F_{m+1}, which is used in the analysis.

Definition 6.1. *For $m = 1, 2, ...$ and $1 \leq i \leq m$, define a number $n_{m,i}$ and a Boolean formula $F_{m,i}(x_1, ..., x_{n_{m,i}})$ depending on $n_{m,i}$ input variables as follows.*

For $m = i = 1$ let

$$n_{1,1} = 1, \quad F_{1,1}(x_1) = x_1.$$

For $m \geq 2$ let

$$n_{m,1} = n_{m-1,m-1}, \quad F_{m,1}(x_1, ..., x_{n_{m,1}}) = F_{m-1,m-1}(x_1, ..., x_{n_{m-1,m-1}}),$$

$$n_{m,i+1} = n_{m,i} + n_{m,1} + 1, \quad F_{m,i+1}(x_1, ..., x_{n_{m,i+1}}) =$$
$$(F_{m,i}(x_1, ..., x_{n_{m,i}}) \wedge F_{m,1}(x_{n_{m,i}+1}, ..., x_{n_{m,i}+n_{m,1}})) \vee x_{n_{m,i+1}}.$$

The next definition, which is just a renaming of the full length prefix $F_{m,m}$ as F_{m+1}, concentrates on the formulae described above.

Definition 6.2. *For $m = 1, 2, ...,$ define a subsequence $F_m(x_1, ..., x_{n_m})$ by $n_m = n_{m,1}$ and $F_m = F_{m,1}$.*

The following theorem states that the "real" randomized decision tree complexity of the functions $\{F_m(x_1, ..., x_{n_m})\}_{m=1,2,...}$ is high — linear in the number of input variables.

Theorem 6.3. *For every m, $RC(F_m) > \frac{n_m}{2}$.*

Proof. Restrict to 0 all the input variables that enter OR-gates. This kills less than half of the input-variables. The restricted resulting function is a simple AND, hence its randomized decision tree complexity is at least $\frac{n_m}{2}$. □

In contrast to Theorem 6.3, the following theorem states that the Saks-Wigderson lower bound does not imply a good lower bound for the functions $\{F_m(x_1, ..., x_{n_m})\}_{m=1,2,...}$. The term R_* is as in Theorem 2.2 and o is the "little-oh" notation.

Theorem 6.4. *For every $\theta > 0.5$, $R_*(F_m) = o((n_m)^\theta)$.*

PROOF. Let $\theta > 0.5$ be given. Clearly, $\frac{n_{m+1}}{n_m} \geq m$, and so $n_m \geq m!$. Therefore, it is sufficient to prove that $R_*(F_m) = o((m!)^\theta)$. To show this we show that $\frac{R_*(F_{m+1})}{R_*(F_m)} \leq 4\sqrt{m}$ for $m \geq 3$. The following sequence is useful for bounding this ratio.

For each m define a sequence of natural numbers $\{b_{m,i}\}_{i=1}^m$ inductively: $b_{m,1} = 1$ and $b_{m,i+1} = b_{m,i} + \frac{1}{1+b_{m,i}} + \frac{1}{R_*(F_m)}$.

The following fact asserts that $b_{m,m}$ is exactly the ratio we are interested in, $\frac{R_*(F_{m+1})}{R_*(F_m)}$. Part (i) is just to help carrying the induction.

FACT 6.5. *For $m \geq 3$ and $1 \leq i \leq m$,*

(i) $R_1(F_{m,i}) = R_0(F_{m,i})$, *hence $R_*(F_{m,i})$ equals this value.*
(ii) $\frac{R_*(F_{m,i})}{R_*(F_m)} = b_{m,i}$.

PROOF. The proof is by induction on m and i (lexicographically ordered). For all m and $i = 1$, part (ii) is trivially true; both sides of it are 1. An easy computation using the definitions of R_0 and R_1, Definition 2.1 shows that part (i) is true for $m = 3$ and $i = 1$. For $i = 1$ and $m > 3$, part (i) is valid by induction since $F_{m,1} = F_{m-1,m-1}$. This concludes the base cases.

Hereafter $G_{m,i}$ denotes $F_{m,i} \wedge F_m$. For $m \geq 3$ and any i we have by Definition 2.1 and by induction, part (i) that

$$\begin{aligned} R_1(G_{m,i}) &= R_*(F_{m,i}) + R_*(F_m) \text{ and} \\ R_0(G_{m,i}) &= \min\{R_*(F_{m,i}) + R_*(F_m),\ \tfrac{R_*(F_{m,i})^2+R_*(F_m)^2+R_*(F_{m,i})R_*(F_m)}{R_*(F_{m,i})+R_*(F_m)}\} \\ &= \tfrac{R_*(F_{m,i})^2+R_*(F_m)^2+R_*(F_{m,i})R_*(F_m)}{R_*(F_{m,i})+R_*(F_m)}. \end{aligned}$$

Also, $R_0(F_{m,i+1}) = R_0(G_{m,i}) + 1$ and

$$\begin{aligned} R_1(F_{m,i+1}) = \min\{ &R_0(G_{m,i}) + 1,\ R_1(G_{m,i}) + 1, \\ &\tfrac{R_0(G_{m,i})R_1(G_{m,i})+R_0(G_{m,i})+1}{R_0(G_{m,i})+1}\}. \end{aligned}$$

The first term is smallest if $R_0(G_{m,i}) + 1 \leq R_1(G_{m,i})$. This is indeed the case since $R_1(G_{m,i}) - R_0(G_{m,i}) = \frac{R_*(F_{m,i})R_*(F_m)}{R_*(F_{m,i})+R_*(F_m)}$ and $R_*(F_{m,i}) \geq R_*(F_m) \geq R_*(F_3) > 2$. Therefore,

$$R_1(F_{m,i+1}) = R_0(G_{m,i}) + 1 = R_0(F_{m,i+1}).$$

This value is

$$\begin{aligned} R_*(F_{m,i+1}) &= \tfrac{R_*(F_{m,i})^2+R_*(F_m)^2+R_*(F_{m,i})R_*(F_m)}{R_*(F_{m,i})+R_*(F_m)} + 1 \\ &= R_*(F_{m,i}) + \tfrac{R_*(F_m)^2}{R_*(F_{m,i})+R_*(F_m)} + 1, \end{aligned}$$

and so

$$\begin{aligned}\frac{R_*(F_{m,i+1})}{R_*(F_m)} &= \frac{R_*(F_{m,i})}{R_*(F_m)} + \frac{1}{\frac{R_*(F_{m,i})}{R_*(F_m)}+1} + \frac{1}{R_*(F_m)}\\ &= b_{m,i} + \frac{1}{1+b_{m,i}} + \frac{1}{R_*(F_m)} = b_{m,i+1}. \ \square\end{aligned}$$

Corollary 6.6. $b_{m,m} = \frac{R_*(F_{m+1})}{R_*(F_m)}$.

Proof. Follows from Fact 6.5, part (ii) with $i = m$. □

The next fact estimates this ratio. For this we use another sequence $\{a_i\}$, defined by $a_1 = 1$ and $a_{i+1} = a_i + \frac{1}{1+a_i}$.

Fact 6.7. *For every m and i, $i \le m$,*

$$\begin{array}{lrl} \text{(i)} & b_{m,i} & \ge a_i, \\ \text{(ii)} & b_{m,i} & \le a_i + \frac{i}{R_*(F_m)}, \\ \text{(iii)} & a_i & \le 2\sqrt{i}, \\ \text{(iv)} & a_i & \ge \sqrt{i}. \end{array}$$

Proof. The proof is by induction on i. For $i = 1$, $b_{m,1} = a_1 = 1$ and the four parts follow.

Assume it is true for i. Consider the following three functions defined on the positive reals, $f(x) = x + \frac{1}{1+x}$, $g(x) = +\sqrt{x^2+4}$ and $h(x) = +\sqrt{x^2+1}$. It is easy to see that $h(x) \le f(x) \le g(x)$ and that $f(x)$ is monotonically increasing. Using these properties and the induction hypothesis we have

$$\begin{array}{lrl} \text{(i)} & b_{m,i+1} & = b_{m,i} + \frac{1}{1+b_{m,i}} + \frac{1}{R_*(F_m)} = f(b_{m,i}) + \frac{1}{R_*(F_m)} \\ & & > f(b_{m,i}) \ge f(a_i) = a_{i+1}, \\ \text{(ii)} & b_{m,i+1} & = b_{m,i} + \frac{1}{1+b_{m,i}} + \frac{1}{R_*(F_m)} \le a_i + \frac{i}{R_*(F_m)} + \frac{1}{1+b_{m,i}} + \frac{1}{R_*(F_m)} \\ & & \le a_i + \frac{1}{1+a_i} + \frac{i+1}{R_*(F_m)} = a_{i+1} + \frac{i+1}{R_*(F_m)}, \\ \text{(iii)} & a_{i+1} & = f(a_i) \le f(2\sqrt{i}) \le g(2\sqrt{i}) = 2\sqrt{i+1}. \\ \text{(iv)} & a_{i+1} & = f(a_i) \ge f(\sqrt{i}) \le h(\sqrt{i}) = \sqrt{i+1}. \ \square \end{array}$$

¿From this fact we get, for $m \ge 3$, the (weak) lower bound of

$$R_*(F_m) \ge \frac{R_*(F_m)}{R_*(F_{m-1})} = b_{m-1,m-1} \overset{6.7(\text{i})}{\ge} a_{m-1} \overset{6.7(\text{iv})}{\ge} \sqrt{m-1} \ge \frac{2}{\sqrt{m}},$$

and complete the proof of Theorem 6.4 by the upper bound

$$\frac{R_*(F_{m+1})}{R_*(F_m)} \stackrel{6.6}{=} b_{m,m} \stackrel{6.7(\text{ii})}{\leq} a_m + \frac{m}{R_*(F_m)} \leq a_m + \frac{2m}{\sqrt{m}} \stackrel{6.7(\text{iii})}{\leq} 4\sqrt{m}. \square$$

Acknowledgements

This work was done while the first author was at the Weizmann Institute of Science and at DIMACS, Princeton University, and while the second author was at the Hebrew University and at Princeton University. The second author was partially supported by Israel-American Binational Science Foundation Grant No. 87-00082 and by Wolfson Research Awards, administered by the Israel Academy of Sciences and Humanities.

Special thanks are to David Harel, Nati Linial, Ilan Newman, Moshe Tennenholz, Eli Upfal and Edna Wigderson for very helpful discussions, comments and computations.

References

Manuscript received 18 December 1990

RAFI HEIMAN
Bell Communication Research
445 South Street
Morristown, New Jersey 07962-1910
rafi@thumper.bellcore.com

AVI WIGDERSON
Department of Computer Science
Princeton University
Princeton, New Jersey 08544
avi@cs.princeton.edu