JOURNAL OF COMPUTER AND SYSTEM SCIENCES **46**, 129–154 (1993)

Threshold Circuits of Bounded Depth

ANDRÁS HAJNAL

Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago and Mathematical Institute of the Hungarian Academy of Sciences, Budapest

WOLFGANG MAASS*

Institute for Theoretical Computer Science, Technische Universität Graz, Austria, and Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago

PAVEL PUDLÁK

Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago and Mathematical Institute, Czechoslovakian Academy of Sciences, Prague

MÁRIÓ SZEGEDY

Department of Computer Science, University of Chicago

AND

GYÖRGY TURÁN

Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago and Automata Theory Research Group of the Hungarian Academy of Sciences, Szeged

Received December 8, 1988; revised January 5, 1991

We examine a powerful model of parallel computation: polynomial size threshold circuits of bounded depth (the gates compute threshold functions with polynomial weights). Lower bounds are given to separate polynomial size threshold circuits of depth 2 from polynomial size threshold circuits of depth 3 and from probabilistic polynomial size circuits of depth 2. With regard to the unreliability of bounded depth circuits, it is shown that the class of functions computed reliably with bounded depth circuits of unreliable $\wedge$, $\vee$, $\neg$ gates is narrow. On the other hand, functions computable by bounded depth, polynomial-size threshold circuits can also be computed by such circuits of unreliable threshold gates. Furthermore we examine to what extent imprecise threshold gates (which behave unpredictably near the threshold value) can compute nontrivial functions in bounded depth and a bound is given for the permissible amount of imprecision. We also discuss threshold quantifiers and prove an undefinability result for graph connectivity.

* Supported in part by NSF Grant CCR-8703889.

0022-0000/93 $5.00

1. Introduction

Circuits with bounded depth and unbounded fanin were studied recently for several different reasons. These circuits are related to relativized separation of complexity classes (Furst *et al.* [13]) and definability problems (Ajtai [1]); they also form a model of parallel computation (Chandra *et al.* [8]).

The results of Ajtai [1], Furst *et al.* [13], Yao [34], Hastad [15], Razborov [25], and Smolensky [31] give superpolynomial lower bounds for the size of bounded depth circuits with $\wedge$, $\vee$, $\neg$ gates computing PARITY, and of bounded depth circuits with $\wedge$, $\vee$, $\neg$ and PARITY gates computing MAJORITY. Thus with respect to bounded depth, polynomial size reducibility ($\leqslant_{\mathrm{bp}}$) [8], $\emptyset <_{\mathrm{bp}}$ PARITY $<_{\mathrm{bp}}$ MAJORITY.

The next step in extending this hierarchy would be to study bounded depth, polynomial size circuits where MAJORITY is also allowed as a single gate.

We consider here these circuits and the complexity class TC^0 of Boolean functions computed by them. This class was first studied by Parberry and Schnitger [22]. The definition we use allows gates computing *threshold functions* of the form

$$T_k^{\boldsymbol{\alpha}}(y_1, \ldots, y_m) = 1 \qquad \text{iff} \quad \sum_{i=1}^{m} \alpha_i y_i \geqslant k \qquad (\text{where } \boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)).$$

Circuits with threshold gates are also related to several computational brain models (e.g., Boltzmann machines [27]) and to other models related to pattern recognition and learning, some of which have been studied since the 1950's (e.g., perceptrons [20, 27]).

We show that depth-2 threshold circuits computing INNER PRODUCT MOD 2 must have exponential size. The proof is based on the notion of a discriminator and a lemma of Lindsey [3, 9, 10] about Hadamard matrices. As INNER PRODUCT MOD 2 can be computed by polynomial size depth-3 threshold circuits and by polynomial size depth-2 probabilistic threshold circuits, this gives a separation of the corresponding complexity classes. We were not able to prove lower bounds for circuits of larger depth.

Probabilistic threshold circuits of depth d can be simulated by deterministic threshold circuits of depth $d+1$ (with polynomial increase in size). If the problem to be computed satisfies some symmetry conditions then there is a simulation in the other direction too.

We also consider reliability in the context of bounded depth, unbounded fanin circuits. The study of reliable computation by circuits of unreliable gates (i.e., gates which output an incorrect value with a certain probability) goes back to von Neumann [21]. More recent results are due to Dobrushin and Ortyukov [10] and Pippenger [23]. These results show, that in the bounded fanin case every function can be computed reliably over any complete basis without a significant increase in complexity (depth and size).

In the bounded depth, unbounded fanin case the situation is quite different. The class of functions which are computable reliably with bounded depth $\wedge$, $\vee$, $\neg$ circuits of any size is very restricted. These are the functions for which $\{\mathbf{x}: f(\mathbf{x})=1\}$ can be obtained from a bounded number of subcubes by Boolean operations. In contrast, threshold circuits with unreliable gates can be used for reliable computation in bounded depth.

For threshold gates one can consider another notion of unreliability, which may have some practical relevance if it becomes technically feasible to build threshold gates with large fanin. A threshold gate is called imprecise if it behaves unpredictably when the evaluated sum is near the threshold value. We prove a general lower bound which is useful in determining the permissible amount of imprecision.

We also consider a logical characterization of TC° using the generalized quantifier $\#_k$ ("there are at least k elements...") introduced by Immerman and Lander [17]. It is shown that CONNECTIVITY of graphs cannot be defined by sentences of bounded quantifier rank using threshold quantifiers and a successor relation.

There are many open problems about threshold circuits. Some of these are mentioned at the end of the paper.

2. Definitions and Examples

A *threshold function* is of the form

$$T_k^{\boldsymbol{\alpha}}(y_1, ..., y_m) = \begin{cases} 1, & \text{if } \sum_{i=1}^m \alpha_i y_i \geqslant k \\ 0, & \text{otherwise,} \end{cases}$$

where the α_i's are the *weights*, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_m) \in Z^m$ is the *weight vector*, and $k \in Z$ is the *threshold value*. We use the notation

$$\alpha = \sum_{i=1}^m |\alpha_i|.$$

We shall also use the notations $T_k^m(y_1, ..., y_m)$, $T_{\leqslant k}^m(y_1, ..., y_m)$ for the functions "at least k 1's," "at most k 1's."

A *threshold circuit* C is a Boolean circuit where every gate computes a threshold function. We assume that the gates are on levels. Level 0 contains $2n+2$ nodes labeled $x_1, ..., x_n, \bar{x}_1, ..., \bar{x}_n$, 0, and 1. Every edge connects a node on level l to a node on level $l+1$. A circuit is a formula if every gate has outdegree $\leqslant 1$.

The *size* of C is the number of gates, its *depth* is the length of a longest path. The *weight* of C is the maximal absolute value of weights occurring in gates of C.

If v is a gate of C and $\mathbf{x}$ is an input, we write $C_v(\mathbf{x})$ for the value computed at v for input $\mathbf{x}$. $C(\mathbf{x})$ is the output value.

TC_d° is the class of languages $L \subseteq \{0, 1\}^*$ for which there is a polynomial p and a sequence $(C_n)_{n \in \mathbf{N}}$ of depth d threshold circuits such that C_n computes $L \cap \{0, 1\}^n$ and both the size and the weight of C_n are bounded by $p(n)$.

$TC^\circ = \bigcup_{d \in \mathbf{N}} TC_d^\circ$ is the class of languages recognizable by threshold circuits of bounded depth such that the size and the weight are bounded by a polynomial. (As for AC°, one can define uniform variants but we do not pursue this in this paper.)

We give some examples of functions in TC°.

PROPOSITION 2.1. *Symmetric functions can be computed by depth-2 threshold circuits of linear size and weight* 1.

Proof. Let $f(x_1, ..., x_k)$ be a symmetric function described by $S = \{s_1, ..., s_k\} \subseteq \{0, ..., n\}$ (i.e., $f(\mathbf{x}) = 1 \Leftrightarrow \sum_{i=1}^n x_i \in S$). Then the circuit

$$T_{k+1}^{2k}(T_{s_1}^n(\mathbf{x}), T_{\leqslant s_1}^n(\mathbf{x}), ..., T_{s_k}^n(\mathbf{x}), T_{\leqslant s_k}^n(\mathbf{x}))$$

computes f. (Observe that the number of 1's that are input to the final gate is always either k or $k+1$). ▮

Another example is given by the SUM-EQUALITY function. (SUM-EQUALITY$_n(x_1, ..., x_n, y_1, ..., y_n) = 1$ iff $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$.)

PROPOSITION 2.2. SUM-EQUALITY *is in* TC_2°.

INNER PRODUCT MOD 2_n is the $2n$-variable function $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 \oplus \cdots \oplus x_n y_n$. Proposition 2.1 implies that this function is in TC_3°. (It will be shown not to belong TC_2°.)

Other problems in TC_3° are ADDITION, SUBTRACTION, and COMPARISON of two n-bit numbers.†

The reductions of Chandra *et al.* [8] and the results of Beame *et al.* [6] imply that the following problems are also in TC°: MULTIPLICATION of two n-bit numbers, SORTING of n n-bit numbers, MULTIPLICATION of n n-bit numbers, and DIVISION of two n-bit numbers (the last two, following [6], require nonuniformity). This relationship is further elaborated in Reif [26].

We mention some modifications of the circuit model for which the class of functions computable in bounded depth and polynomial size is equal to TC°:

1. Restriction to nonnegative weights: negative weights can be eliminated by replacing $\alpha_i y_i$ by $\alpha_i(1 - \bar{y}_i)$ if $\alpha_i < 0$ everywhere in the circuit, and computing for every gate v not only $C_v(\mathbf{x})$ but $\overline{C_v}(\mathbf{x})$ too. This transformation does not increase the depth and at most doubles the size.

2. Restriction to majority gates: as the weights are polynomial, one can use multiple edges and new edges from the constants to obtain majority gates.

3. Generalization to threshold gates with arbitrary weights: as the number of bits of the weights can be assumed to be polynomial and the addition of several numbers can be done in bounded depth and polynomial size using majority gates [8], these gates can be simulated by subcircuits performing this task.

† *Note added in proof.* Recently Siu and Bruck have shown that ADDITION and COMPARISON are in TC_2°; furthermore Siu and Roychowdhury have shown that MULTIPLICATION and DIVISION are in TC_3°.

4. Gates for every symmetric function: equivalence follows from Proposition 2.1 and modification 2 above.

Proposition 2.1 suggests that depth d circuits with gates computing symmetric functions can be simulated by depth $2d$ threshold circuits. This simulation can be improved.

THEOREM 2.3. *Let $f\colon \{0,1\}^n \to \{0,1\}$ be computed by a depth d, size s circuit C of gates computing symmetric functions. Then f is computed by a depth $d+1$ threshold circuit C' of size $O((s+n)^2)$ and weight* 1.

LEMMA 2.4. *Let $f\colon \{0,1\}^n \to \{0,1\}$ be computed by a depth-2, size-s circuit C of gates computing symmetric functions. Then f is computed by a depth-2 size-$O(sn)$ circuit C' such that C' has threshold gates (with weights* 0 *or* 1) *on level* 1 *and a gate computing a symmetric function on level* 2.

Proof. Let C be of the form

$$S(S_1(y_{11}, ..., y_{1r_1}), ..., S_m(y_{m1}, ..., y_{mr_m})),$$

where y_{ij} $(1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant r_i)$ are variables or their negations, S is given by $\{s_1, ..., s_k\} \subseteq \{0, ..., m\}$ and S_i is given by $\{s_{i,1}, ..., s_{i,k_i}\} \subseteq \{0, ..., r_i\}$.

Let the gates on the first level of C' be $T^{r_i}_{s_{i,j}}(y_{i1}, ..., y_{ir_i})$ and $T^{r_i}_{\leqslant s_{i,j}}(y_{i1}, ..., y_{ir_i})$ for $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant k_i$ and the final gate S' be given by

$$\left\{\sum_{i=1}^{m} k_i + s_1, ..., \sum_{i=1}^{m} k_i + s_k\right\} \subseteq \left\{0, ..., 2\sum_{i=1}^{m} k_i\right\}.$$

Then C' computes f by the observation proving Proposition 2.1. ∎

Proof of Theorem 2.3. C' is constructed by applying the construction of Lemma 2.4 $d-1$ times starting at levels 1 and 2 and, finally, applying the construction of Proposition 2.1. The size bound follows by observing that each gate v in C corresponds to $2k$ gates in C', where k is the size of the set describing the symmetric function computed at v. ∎

Other equivalent models (threshold parallel computer, TRAM, threshold Turing-machine, and a deterministic variant of the Boltzmann-machine) are given by Parberry and Schnitger [22]. The equivalence to circuits with ADDITION and MULTIPLICATION follows from the reductions of [8]. The equivalence to probabilistic threshold circuits is given in Section 4. A logical characterization of TC° is described in Section 7.

Note that the classes TC°_d are not necessarily invariant under these simulations. The definition used here has the advantage that for every d, TC°_d is closed under p-projections [18, 30] and has a natural complete problem MAJ^d: $\mathrm{MAJ}^d(x_1, ..., x_n)$ is computed by a depth d, indegree $n^{1/d}$ tree of MAJORITY gates.

3. Lower Bounds for Depth-2 Circuits

PARITY cannot be computed by a single threshold gate as every threshold function is monotone increasing or decreasing in every variable. Therefore Proposition 2.1 implies $TC_1^\circ \subsetneq TC_2^\circ$. Now we show that the second and third levels of the hierarchy are also different.

THEOREM 3.1. $TC_2^\circ \subsetneq TC_3^\circ$.

This follows from:

LEMMA 3.2. *Fix any $\varepsilon > 0$ and polynomial p. Assume that C is a depth-2 threshold circuit with weight $\leqslant p(n)$ computing* INNER PRODUCT MOD 2 *of two n-bit strings. Then if n is sufficiently large, the size of C is at least $2^{(1/2-\varepsilon)n}$.*

Before proving the lemma we define discriminators.

DEFINITION. Let C be a circuit with n inputs and $A, B \subseteq \{0, 1\}^n$ be disjoint sets. Let P_A (resp. P_B) denote the uniform probability distribution on A (resp. B). Then C is an *ε-discriminator* for A and B if

$$|P_A(C(\mathbf{x})=1) - P_B(C(\mathbf{x})=1)| \geqslant \varepsilon.$$

If f is an n-variable function, C is an *ε-discriminator for f* if it is an ε-discriminator for $A = \{\mathbf{x}: f(\mathbf{x})=1\}$ and $B = \{\mathbf{x}: f(\mathbf{x})=0\}$.

LEMMA 3.3. *Let $T_k^\alpha(C_1, ..., C_m)$ be a circuit with n inputs, $\alpha = \sum_{i=1}^m |\alpha_i|$ and $A, B \subseteq \{0, 1\}^n$ be disjoint sets such that the circuit accepts A and rejects B. Then some subcircuit C_i $(1 \leqslant i \leqslant m)$ is a $(1/\alpha)$-discriminator for A and B.*

Proof. Let the random variable $C_i^A(\mathbf{x})$ (resp. $C_i^B(\mathbf{x})$) be the output of C_i when $\mathbf{x}$ is distributed uniformly on A (resp. B). Then $\sum_{i=1}^m \alpha_i C_i^A(\mathbf{x}) \geqslant k$ and $\sum_{i=1}^m \alpha_i C_i^B(\mathbf{x}) \leqslant k-1$. Taking expectations and rearringing we obtain

$$\begin{aligned} 1 &\leqslant \sum_{i=1}^{m} \alpha_i (E(C_i^A(\mathbf{x})) - E(C_i^B(\mathbf{x})) \\ &\leqslant \alpha \cdot \max_{1 \leqslant i \leqslant m} |P_A(C_i(\mathbf{x})=1) - P_B(C_i(\mathbf{x})=1|. \quad \blacksquare \end{aligned}$$

This lemma reduces the problem of proving a lower bound for depth d circuits to showing that there are no depth $(d-1)$ ε-discriminators for large ε. The lemma holds for arbitrary probability distributions. Later (Theorem 4.6) we shall prove a partial converse of this lemma.

The following lemma is a special case of a lemma due to Lindsey (see [3, 9, 11]).

LEMMA 3.4 (Lindsey). *For every* $X, Y \subseteq \{0, 1\}^n$,

$$||\{(\mathbf{x}, \mathbf{y}) \in X \times Y: \mathbf{x} \cdot \mathbf{y} = 1\}| - |\{(\mathbf{x}, \mathbf{y}) \in X \times Y: \mathbf{x} \cdot \mathbf{y} = 0\}|| \leqslant \sqrt{|X| \cdot |Y| \cdot 2^n}.$$

Proof of Lemma 3.2. Let $T_k^{\alpha}(C_1, ..., C_m)$ be a circuit satisfying the assumption of the lemma. Thus the C_i are threshold gates. Let C_i be a gate of the form

$$\sum_{j=1}^{n} \beta_j x_j + \sum_{j=1}^{n} \gamma_j y_j \geqslant l.$$

Let

$$X_u := \left\{\mathbf{x}: \sum_{j=1}^{n} \beta_j x_j = u\right\},$$

$$Y_u := \left\{\mathbf{y}: \sum_{j=1}^{n} \gamma_j y_j \geqslant l - u\right\}$$

(thus $|u| \leqslant np(n)$). Then

$$\{(\mathbf{x}, \mathbf{y}): C_i(\mathbf{x}, \mathbf{y}) = 1\} = \bigcup_{u=-np(n)}^{np(n)} X_u \times Y_u.$$

Applying Lemma 3.4,

$$||\{(\mathbf{x}, \mathbf{y}): C_i(\mathbf{x}, \mathbf{y}) = 1, \mathbf{x} \cdot \mathbf{y} = 1\}| - |\{(\mathbf{x}, \mathbf{y}): C_i(\mathbf{x}, \mathbf{y}) = 1, \mathbf{x} \cdot \mathbf{y} = 0\}|| \leqslant \sum_{u=-np(n)}^{np(n)} \sqrt{|X_u|\ |Y_u|\ 2^n} \leqslant (2n+1)\ p(n)\ 2^{3n/2}$$

and as $|\{(\mathbf{x}, \mathbf{y}): \mathbf{x} \cdot \mathbf{y} = 1\}| = 2^{2n-1} - 2^{n-1}$, $|\{(\mathbf{x}, \mathbf{y}): \mathbf{x} \cdot \mathbf{y} = 0\}| = 2^{2n-1} + 2^{n-1}$, if C_i is an ε'-discriminator for INNER PRODUCT MOD 2, then $\varepsilon' \leqslant 2^{-(1/2 - \varepsilon/2)n}$ if n is sufficiently large. Hence in the original circuit $\alpha \geqslant 2^{(1/2 - \varepsilon/2)n}$ from Lemma 3.3. But $\alpha \leqslant mp(n)$, thus

$$m \geqslant \frac{1}{p(n)} 2^{(1/2 - \varepsilon/2)n} \geqslant 2^{(1/2 - \varepsilon)n}$$

if n is sufficiently large. ▌

Proof of Theorem 3.1. INNER PRODUCT MOD 2 is not in TC_2° by Lemma 3.2. On the other hand, it is in TC_3° as it was remarked after Proposition 2.2. ▌

Lemma 3.2 implies further lower bounds using reduction by p-projections.

COROLLARY 3.5. *Fix any polynomial p. Let C be a depth-2 threshold circuit with weight* $\leqslant p(n)$. *Then the following holds*:

(a) *if C computes MULTIPLICATION of two n-bit numbers then it has size* $2^{\Omega(n/\log n)}$;

(b) *if C computes CONNECTIVITY of n-vertex graphs then it has size* $2^{\Omega(n)}$;

(c) *if C computes* MAJ^3 *of n bits then it has size* $2^{\Omega(n^{1/3})}$.

Proof. (a) Let $\mathbf{x} = x_1 0 \cdots 0 x_2 0 \cdots 0 x_{n-1} 0 \cdots 0 x_n$, $\mathbf{y} = y_n 0 \cdots 0 \cdots y_2 0 \cdots 0 y_1$ where $\lfloor \log_2 n \rfloor$ 0's are inserted between any two variables. Then $x_1 y_1 \oplus \cdots \oplus x_n y_n$ is the mth digit of the product $\mathbf{x} \times \mathbf{y}$, where m is the length of $\mathbf{x}$.

(b) The fact that INNER PRODUCT MOD 2 is a p-projection of CONNECTIVITY follows from a result of Skyum [29]. A more efficient construction is obtained by taking a cycle of length $4n$ and replacing every second pair of opposite edges by the labeled subgraph of Fig. 1. (Other edges are labeled 0.)

(c) Proposition 2.1 shows that INNER PRODUCT MOD 2 is computed by a depth-3 formula of $O(n)$ variable majority gates which leads to the desired projection. ∎

The proof method of Lemma 3.2 can also be used to prove lower bounds for depth-2 threshold circuits having arbitrary weights on the first level.

THEOREM 3.6. *Let ε, p, C be as in Lemma 3.2, except that there is no bound on the size of weights for gates on level 1. Then if n is sufficiently large, C has size at least* $2^{(1/3-\varepsilon)n}$.

Proof (Outline). Following the proof of Lemma 3.2 it suffices to show that if C_i is a threshold gate of the form

$$\sum_{j=1}^{n} \beta_j x_j + \sum_{j=1}^{n} \gamma_j y_j \geqslant l,$$

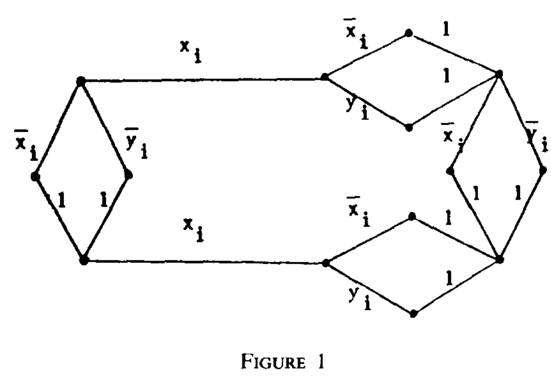

FIGURE 1

then

$$||\{(\mathbf{x}, \mathbf{y}): C_i(\mathbf{x}, \mathbf{y}) = 1, \mathbf{x} \cdot \mathbf{y} = 1\}| - |\{(\mathbf{x}, \mathbf{y}): C_i(\mathbf{x}, \mathbf{y}) = 1, \mathbf{x} \cdot \mathbf{y} = 0\}|| = O(2^{5n/3})$$

if n is sufficiently large.

Form a $2^n \times 2^n$ matrix where the rows (resp. columns) are indexed by $\mathbf{x}$'s in increasing order of $\boldsymbol{\beta}\mathbf{x}$ (resp. $\mathbf{y}$'s in increasing order of $\gamma\mathbf{y}$), and entry $(\mathbf{x}, \mathbf{y})$ is $C_i(\mathbf{x}, \mathbf{y})$. In this matrix every entry either to the right of, or below an entry which is 1, is also equal to 1. Divide the matrix into $2^{2n/3}$ square submatrices of size $2^{2n/3} \times 2^{2n/3}$. There are at most $2 \cdot 2^{n/3}$ squares containing both a 0 and a 1. The 1 entries not in these squares can be covered by $2^{n/3}$ rectangles of height $2^{n/3}$ and width $\leqslant 2^n$. Thus using Lemma 3.4 the difference above is at most

$$2 \cdot 2^{n/3} \cdot 2^{4n/3} + 2^{n/3} \cdot \sqrt{2^{2n/3} \cdot 2^n \cdot 2^n} = 3 \cdot 2^{5n/3}. \quad \blacksquare$$

4. Probabilistic Threshold Circuits

A probabilistic threshold circuit is a threshold circuit with two kinds of inputs, $\mathbf{x} = (x_1, ..., x_n)$ and unbiased random bits $\mathbf{r} = (r_1, ..., r_l)$. $P(C(\mathbf{x}, r) = 1)$ is the probability that C accepts $\mathbf{x}$.

For $A, B \subseteq \{0, 1\}^n$, $\varepsilon > 0$, we say that C gives advantage ε to A over B if

$$P(C(\mathbf{x}, \mathbf{r}) = 1) \geqslant \tfrac{1}{2} + \varepsilon \qquad \text{for every } \mathbf{x} \in A,$$

$$P(C(\mathbf{x}, \mathbf{r}) = 1) \leqslant \tfrac{1}{2} - \varepsilon \qquad \text{for every } \mathbf{x} \in B.$$

RTC_d° is the class of languages $L \subseteq \{0, 1\}^*$ for which there is a polynomial p and a sequence of depth d threshold circuits $(C_n)_{n \in N}$ such that C_n gives advantage ε_n to $L \cap \{0, 1\}^n$ over $\bar{L} \cap \{0, 1\}^n$, $\varepsilon_n^{-1} \leqslant p(n)$, and the size, the weight, and the number of random bits used by C_n are all bounded by $p(n)$:

$$RTC_d^\circ = \bigcup_{d \in \mathbf{N}} RTC_d^\circ.$$

We shall use the standard Chernoff inequality: if S_m is the sum of m independent random variables each taking value 1 (resp. 0) with probability p (resp. $1 - p$) then $P(S_m \geqslant pm + h) \leqslant e^{-(2h^2/m)}$ and $P(S_m \leqslant pm - h) \leqslant e^{-(2h^2/m)}$.

Lemma 4.1. *Let $A, B \subseteq \{0, 1\}^n$ be disjoint sets. Let C be a depth-d probabilistic circuit with size, weight, and number of random bits all bounded by $p(n)$ which gives advantage ε to A over B. Then there is depth-$(d+1)$, deterministic circuit C' which accepts A, rejects B and has size and weight bounded by $O(p(n)\, n\varepsilon^{-2})$.*

Proof. Consider $T_{m/2}^m(C_1(\mathbf{x}, \mathbf{r}_1), ..., C_m(\mathbf{x}, \mathbf{r}_m))$, where the C_i's are disjoint copies of C. For every $\mathbf{x}$, $C_i(\mathbf{x}, \mathbf{r}_i)$ $(1 \leqslant i \leqslant m)$ are independent random variables. For every $\mathbf{x} \in A$, $P(C_i(\mathbf{x}, \mathbf{r}_i) = 1) \geqslant \frac{1}{2} + \varepsilon$, so $P(\sum_{i=1}^m C_i(\mathbf{x}, \mathbf{r}_i) < m/2) < e^{-2\varepsilon^2 m}$ from the Chernoff inequality and similarly for $\mathbf{x} \in B$, $P(\sum_{i=1}^{m} C_i(\mathbf{x}, \mathbf{r}_i) \geqslant m/2) < e^{-2\varepsilon^2 m}$. The probability that the circuit above gives an incorrect output for some $\mathbf{x} \in A \cup B$ is $\leqslant 2^n \cdot e^{-2\varepsilon^2 m}$, which is < 1 for some $m = O(n\varepsilon^{-2})$. For this m, some choice of the random bits $\mathbf{r}_1, ..., \mathbf{r}_m$ works for every $\mathbf{x}$. ▮

PROPOSITION 4.2. *$RTC_d^\circ \subseteq TC_{d+1}^\circ$, hence $RTC^\circ = TC^\circ$.*

Proof. Follows directly from the previous lemma. ▮

Note that the corresponding statement holds for AC° only if $\varepsilon_n^{-1} = (\log n)^{O(1)}$ [2]. Now we show that the two hierarchies differ on the second level.

THEOREM 4.3. *$TC_2^\circ \subsetneq RTC_2^\circ$.*

This follows directly from Lemma 3.2 and the following lemma (giving another proof of Theorem 3.1).

LEMMA 4.4. *INNER PRODUCT MOD 2 is in RTC_2°.*

LEMMA 4.5. *Let $L \subseteq \{0, 1\}^*$, $A_n = L \cap \{0, 1\}^n$, $B_n = \bar{L} \cap \{0, 1\}^n$, and p be a polynomial. Suppose there is a sequence $(C_n)_{n \in N}$ of depth d probabilistic circuits such that C_n has size, weight, and number of random bits all bounded by $p(n)$ and*

(1) *for* $\mathbf{x} \in A_n$, $P(C_n(\mathbf{x}, \mathbf{r}) = 1) \geqslant \alpha_n$

(2) *for* $\mathbf{x} \in B_n$, $P(C_n(\mathbf{x}, \mathbf{r}) = 1) \leqslant \beta_n$

(3) $0 < (\alpha_n - \beta_n)^{-1} \leqslant p(n)$.

Then $L \in RTC^\circ$.

Proof. Assume C_n uses random bits $r_1, ..., r_l$ and is of the form $T_k^{\boldsymbol{\alpha}}(C^1, ..., C^m)$. Assume w.l.o.g. $(\alpha_n + \beta_n)/2 > \frac{1}{2}$.

Let p, q be integers with $q \leqslant 2^p$ and take p new random bits $\mathbf{r}' = (r'_1, ..., r'_p)$. Let $s_1, ..., s_q$ be different elementary conjunctions of the new random bits. Consider the circuit C' given by

$$T_{k+\alpha+1}^{\boldsymbol{\alpha}'}(C^1, ..., C^m, s_1, ..., s_q),$$

where $\alpha = \sum_{i=1}^{m} |\alpha_i|$ and $\boldsymbol{\alpha}'$ is obtained by appending the weight $(\alpha + 1)$ q times to $\boldsymbol{\alpha}$. Then

$$\begin{aligned} P(C'(\mathbf{x}, \mathbf{r}, \mathbf{r}') = 1) &= P(C(\mathbf{x}, \mathbf{r}) = 1) \cdot P(\exists i \leqslant q \colon s_i = 1) \\ & P(C(\mathbf{x}, \mathbf{r}) = 1) \cdot q/2^p \end{aligned}$$

and a simple calculation shows that for some $p = O(-\log(\alpha_n - \beta_n))$ and a suitable q, C' gives advantage $\frac{1}{8}(\alpha_n - \beta_n)$ as required in the definition of RTC°. ▮

Proof of Lemma 4.4. We describe a depth-2 probabilistic circuit for INNER PRODUCT MOD 2. Let $\mathbf{x} = (x_1, ..., x_n)$, $\mathbf{y} = (y_1, ..., y_n)$, and let $\mathbf{r} = (r_1, ..., r_n)$ be random bits. Consider the circuit

$$T^{2n}_{\leqslant n-1}(x_1 \wedge y_1 \wedge r_1, \bar{x}_1 \vee \bar{y}_1 \vee r_1, ..., x_n \wedge y_n \wedge r_n, \bar{x}_n \vee \bar{y}_n \vee r_n).$$

Observe that if $x_i = y_i = 1$ then $x_i \wedge y_i \wedge r_i = \bar{x}_i \vee \bar{y}_i \vee r_i = r_i$, otherwise $x_i \wedge y_i \wedge r_i = 0$, $\bar{x}_i \vee \bar{y}_i \vee r_i = 1$. If for $(\mathbf{x}, \mathbf{y})$ there are k indices i with $x_i = y_i = 1$, this fixes $n-k$ 0's and $n-k$ 1's for $T^{2n}_{\leqslant n-1}$. The remaining bits are generated randomly in pairs of 0's and 1's. Thus if k is odd, we shall never have exactly n 0's and 1's. Hence for such an input the probability of acceptance is $\frac{1}{2}$. If k is even, the probability of having n 0's and n 1's is $\sim \sqrt{2/\pi k} \geqslant \sqrt{2/\pi n}$, and the probability of having $n-2t$ 0's is equal to the probability of having $n+2t$ 0's. Hence for such an input the probability of acceptance is $\leqslant \frac{1}{2} - \sqrt{2/\pi n}$. The proof is completed by applying Lemma 4.5. ∎

In the remainder of this section we present some evidence that the simulation of probabilistic threshold circuits by deterministic ones must increase the depth in general.

Let G be a group of permutations on $\{1, ..., n\}$. Then G induces a group G^* of permutations on $\{0, 1\}^n$.

DEFINITION. $A_n, B_n \subseteq \{0, 1\}^n$ are homogeneous if A_n and B_n are the orbits of G^* for some group of permutations G.

For example, considering CONNECTIVITY of graphs, $A_n = \{\text{Hamilton cycles}\}$ and $B_n = \{\text{disjoint unions of two cycles of length } n/2\}$ are homogeneous sets of inputs, where G is the group of permutations of the edges induced by relabelings of vertices.

THEOREM 4.6. *Let $d \geqslant 3$ and let $A_n, B_n \subseteq \{0, 1\}^n$ be disjoint homogeneous sets. Then the following are equivalent:*

(a) *A_n and B_n can be separated by depth $d+1$ (deterministic) circuits with size and weight bounded by a polynomial,*

(b) *there are depth d ε_n-discriminators for A_n and B_n with size, weight, and ε_n^{-1} bounded by a polynomial,*

(c) *there are depth-d probabilistic circuits giving advantage ε_n to A_n over B_n such that the size, the weight, the number of random bits, and ε_n^{-1} are all bounded by a polynomial.*

Proof. (a) $\Rightarrow$ (b) is Lemma 3.3, (c) $\Rightarrow$ (a) is Lemma 4.1, therefore we only have to prove (b) $\Rightarrow$ (c).

Let C be an ε_n-discriminator for A_n and B_n, with $P_{A_n}(C(\mathbf{x}) = 1) = \alpha_n$, $P_{B_n}(C(\mathbf{x}) = 1) = \beta_n$. Assume w.l.o.g. $\varepsilon_n = \alpha_n - \beta_n > 0$.

For $\gamma \in G$ let C^γ denote the circuit computing $C(\gamma^*(\mathbf{x}))$, where γ^* is the element of G^* induced by γ. Let P_G be the uniform probability distribution on G. Then as G^* is transitive on A_n and B_n, for every $\mathbf{x} \in A_n$ (resp. $\mathbf{x} \in B_n$), $P_G(C^\gamma(\mathbf{x})=1)=\alpha_n$ (resp. $P_G(C^\gamma(\mathbf{x})=1)=\beta_n$).

The Chernoff inequality implies as in Lemma 4.1, that there are permutations $\gamma_1, ..., \gamma_m$ with $m=O(n\varepsilon^{-2})$ such that if P'_G denotes the uniform probability distribution on $\{\gamma_1, ..., \gamma_m\}$then for every $\mathbf{x} \in A_n$ (resp. $\mathbf{x} \in B_n$), $P'_G(C^\gamma(\mathbf{x})=1) \geqslant \alpha_n - \varepsilon_n/3$ (resp. $P'_G(C^\gamma(\mathbf{x})=1) \leqslant \beta_n + \varepsilon_n/3$).

Assume w.l.o.g. that $m=2^l$ and $r_1, ..., r_l$ are random bits. Let C' be a circuit such that

$$C'(x_1, ..., x_n, r_1, ..., r_l) = C(\gamma_i^*(\mathbf{x})),$$

where $(r_1, ..., r_l)$ is i in binary. Then applying Lemma 4.5 to center the probabilities $\alpha_n - \varepsilon_n/3$ and $\beta_n + \varepsilon_n/3$ around $\frac{1}{2}$, we obtain a circuit giving advantage $\varepsilon_n/12$ to A_n over B_n.

What remains is to describe the construction of C'.

LEMMA 4.7. *Let D be a threshold circuit of the form $T_l^{\boldsymbol{\beta}}(D_1, ..., D_m)$ where all weights in $\boldsymbol{\beta}$ are nonnegative and D^* be a new circuit called the* switch. *Form a circuit D' of the form $T_l^{\boldsymbol{\beta}'}(D_1, ..., D_m, D^*)$, where $\boldsymbol{\beta}'$ is $\boldsymbol{\beta}$ appended by a coefficient l. Then if D^* outputs* 0, *D' behaves as D, otherwise it outputs constant* 1.

Proof. Clear. ∎

Continuing the proof of Theorem 4.6, let C be of the form $T_k^{\boldsymbol{\alpha}}(C_1, ..., C_m)$, where we assume w.l.o.g. that all weights are nonnegative. Let $s_1, ..., s_{2^l}$ be all the elementary disjunctions of the random bits $r_1, ..., r_l$. (Thus always exactly one of them is 0.)

Let C_i^j be C_i equipped with switch s_j $(1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant 2^l)$ and let C' be of the form

$$T^{\boldsymbol{\alpha}'}_{\alpha \cdot (2^l-1)+k}(C_1^1, ..., C_m^1, ..., C_1^{2^l}, ..., C_m^{2^l}),$$

where $\boldsymbol{\alpha}'$ is $\boldsymbol{\alpha}$ repeated 2^l times and $\alpha = \sum_{i=1}^m \alpha_i$. (The C_i^j's have disjoint sets of gates but use the same random bits $r_1, ..., r_l$). The correctness of the construction and the size bounds follow from the remarks preceding Lemma 4.7. The switch has depth 1, so the depth of C is not increased if it is at least 3. ∎

We note that to prove (a) directly from (b) it suffices to take $O(n\varepsilon^{-2})$ circuits $C(\gamma^*(\mathbf{x}))$ with γ chosen at random from G and apply a suitable threshold function to the outputs of these circuits.

5. Circuits with Unreliable Threshold Gates

A gate is called ε-unreliable if with probability ε it computes an incorrect output (ε is called the unreliability of the gate). In a circuit of unreliable gates the gate failures are assumed to be independent. The unreliability of different gates in a circuit may be different. C is a circuit of $\geqslant\varepsilon$-unreliable gates if it consists of gates v that are $\varepsilon(v)$-unreliable and $\varepsilon(v)\geqslant\varepsilon$ for every v. C computes a function f with error probability δ if for every $\mathbf{x}$, C computes $f(\mathbf{x})$ with probability $\geqslant 1-\delta$. We assume $\varepsilon\leqslant\frac{1}{2}$.

First we consider bounded depth, unbounded fanin circuits with $\wedge$, $\vee$, $\neg$ gates.

Let $A(f)$ be the smallest number of subcubes from which $\{\mathbf{x}: f(\mathbf{x})=1\}$ can be obtained by union, intersection, and complementation.

PROPOSITION 5.1. *For every r and $\delta>0$ there is an $\varepsilon>0$ such that every function f with $A(f)\leqslant r$ can be computed by a depth $6r+4$ circuit of $\wedge$, $\vee$, $\neg$ gates with error probability δ, for every assignment of unreliabilities $\leqslant\varepsilon$ to its gates.*

Proof. If $A(f)\leqslant r$, f can be written in the form $g(h_1, ..., h_{2r})$, where h_i $(1\leqslant i\leqslant 2r)$ is a conjunction or disjunction of some of the variables of f. Let F' be a formula of depth $\leqslant 2r$ computing $g(y_1, ..., y_{2r})$ with fanin $\leqslant 2$ [19] and let F be obtained from F' by replacing the variables y_i by h_i.

Let G be a subformula of F of depth i computing a function f_G. We claim that there is a circuit C_G of depth $3i-2$ which computes f_G with error probability δ for every assignment of unreliabilities $\leqslant\varepsilon$ to the gates (where ε depends on δ). For $i=1$ this is obvious. For $i>1$ the fanin is $\leqslant 2$, so we can use induction as in Pippenger [23, Theorem 2.4] (compute f_G three times from its subcircuits and take majority, requiring two additional levels). (If F is levelled then the circuit constructed will also be levelled.) ∎

Note that in the construction above the size of the circuit computing f is also bounded by a function of r.

Now we can prove a converse of the proposition above which shows that if a function f can be computed reliably then $A(f)$ is bounded by a function of the depth. (Here we do not need to assume that the circuit is levelled.)

THEOREM 5.2. *For every $\varepsilon>0$ and $\delta>0$, if f is any function computable by a depth d circuit of $\geqslant\varepsilon$-unreliable $\wedge$, $\vee$, $\neg$ gates with error probability $\frac{1}{2}-\delta$ then*

$$A(f)=d^{O(d)}.$$

Proof. Let C be a depth d circuit of $\geqslant\varepsilon$-unreliable gates computing f with error probability $\frac{1}{2}-\delta$. Let s be a number satisfying

$$(1-\varepsilon)^{s+1}\cdot s^{d-2}<\delta.$$

An edge of C is a *red* edge if its tail is a gate (not an input variable). We construct a circuit C' from C as follows: starting from the output gate upwards, replace every $\wedge$-gate (resp. $\vee$-gate) with red indegree $>s$ by an $\wedge$-gate (resp. $\vee$-gate) with a single constant 0 (resp. 1) input (the new gate has the same unreliability as the original one); delete superfluous gates.

LEMMA 5.3. *For every input* $\mathbf{x}$,

$$|P(C(\mathbf{x})=1)-P(C'(\mathbf{x})=1)|<\delta.$$

Proof. Let v be an $\wedge$-gate (the proof for $\vee$ gates follows similarly). Let $v_1, ..., v_t$ $(t>s)$ be the inputs of v and assume that if v_i is input to v_j then $i<j$. Then

$$P\left(\bigwedge_{i=1}^{t} C_{v_i}(\mathbf{x})=1\right)=\prod_{i=1}^{t} P\left(C_{v_i}(\mathbf{x})=1 \,\middle|\, \bigwedge_{j=1}^{i-1} C_{v_j}(\mathbf{x})=1\right)$$
$$\leqslant (1-\varepsilon)^t \leqslant (1-\varepsilon)^{s+1},$$

as for every input, every gate outputs 0 with probability $\geqslant \varepsilon$.

The number of new gates in C' is $\leqslant s^{d-2}$, as new gates occur in depth $\geqslant 2$, and every gate in C' has red indegree $\leqslant s$. Hence, if B denotes the event that all replaced $\wedge$-gates (resp. $\vee$-gates) output a 0 (resp. a 1) in C, then

$$P(B) \geqslant 1-(1-\varepsilon)^{s+1} s^{d-2} > 1-\delta.$$

But

$$P(C(\mathbf{x})=1 \mid B)=P(C'(\mathbf{x})=1 \mid B),$$

so the lemma follows by considering the event $C(\mathbf{x})=1$ under conditions B and $\bar{B}$. ▮

C' contains $\leqslant s^{d-1}$ gates. The variables input to each gate v correspond to a subcube or its complement. Let the subcubes be $Q_1, ..., Q_t$ $(t \leqslant s^{d-1})$.

If $A(f)>s^{d-1}$, then there are $\mathbf{y}$ and $\mathbf{z}$ such that $f(\mathbf{y})=1$, $f(\mathbf{z})=0$, and for every i, $\mathbf{y} \in Q_i$ if and only if $\mathbf{z} \in Q_i$. Then $P(C(\mathbf{y})=1) \geqslant \frac{1}{2}+\delta$, $P(C(\mathbf{z})=1) \leqslant \frac{1}{2}-\delta$, and

$$\begin{aligned} |P(C(\mathbf{y})=1-P(C(\mathbf{z})=1)| &\leqslant |P(C(\mathbf{y})=1)-P(C'(\mathbf{y})=1)| \\ &\quad + |P(C'(\mathbf{y})=1)-P(C'(\mathbf{z})=1)| \\ &\quad + |P(C'(\mathbf{z})=1)-P(C(\mathbf{z})=1)| < 2\delta, \end{aligned}$$

using the lemma above (the middle term is 0), a contradiction. The bound follows by a computation showing $s=O(d \log d)$. ▮

Theorem 5.2 can be used to give lower bounds to the depth required for the reliable computation of several simple functions in AC°. As an example we mention EQUALITY, i.e., the function $\bigwedge_{i=1}^{n} (x_i = y_i)$.

COROLLARY 5.4. *For every* $\varepsilon>0$, *and* $\delta>0$, *every circuit of* $\geqslant \varepsilon$*-unreliable* $\wedge$, $\vee$, $\neg$ *gates computing EQUALITY with error probability* $\frac{1}{2}-\delta$ *has depth* $\Omega(\log n/\log\log n)$.

Proof. By Theorem 5.2 it is sufficient to show $A(\text{EQUALITY}) \geqslant n$. Let $Q_1, ..., Q_m$ be subcubes of the $2n$-dimensional cube generating $\{(\mathbf{x}, \mathbf{x}): \mathbf{x} \in \{0, 1\}^n\}$. Each Q_i can be written as $Q_i' \times Q_i''$, where Q_i' (resp. Q_i'') is a subcube in the first (resp. second) n dimensions. If for some $\mathbf{x}_1 \neq \mathbf{x}_2$ it holds that $\mathbf{x}_1 \in Q_i' \Leftrightarrow \mathbf{x}_2 \in Q_i'$ for every $i = 1, ..., m$, then we get a contradiction, considering $(\mathbf{x}_1, \mathbf{x}_1)$ and $(\mathbf{x}_2, \mathbf{x}_1)$. Hence $m \geqslant n$. ▌

Now we turn to circuits of unreliable threshold gates. We show that in this case bounded depth circuits can be used for reliable computation. In Theorem 5.5 we use the assumption that each gate has the same unreliability.

THEOREM 5.5. *Let f be a function computed by a threshold circuit C of depth d, size s, and weight w. Then for every $\varepsilon < \frac{1}{8}$, f is computed by a depth-d circuit C' of ε-unreliable threshold gates with error probability 2ε, such that C' has size $O(w^{3(d-1)} s^{3(d-1)^2})$, and weight w.*

This follows from the following lemma.

LEMMA 5.6. *Let F be a threshold formula of depth d, size s, and weight w computing a function f. Then for $\varepsilon < \frac{1}{8}$ and for every t, f can be computed by a depth-d threshold formula F' of ε-unreliable gates with error probability δ, where $\delta = \varepsilon + 1/t$ such that F' has size $O(((ws)^3 + (\log(ts))^3)^{d-1})$ and weight w.*

Proof. By induction on d, for $d = 1$ the statement is obvious. For $d > 1$ let F be of the form $T_k^{\boldsymbol{\alpha}}(F_1, ..., F_m)$, where $\alpha = (\alpha_1, ..., \alpha_m)$ and F_i computes f_i $(1 \leqslant i \leqslant m)$.

Let $F_{i,j}'$ be disjoint formulas of ε-unreliable gates computing f_i with error probability δ' for $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant N$, where δ' and N will be specified later. Let

$$F' = T_l^{\boldsymbol{\alpha}'}(F_{1,1}', ..., F_{1,N}', ..., F_{m,1}', ..., F_{m,N}'),$$

where $\boldsymbol{\alpha}' = (\alpha_1, ..., \alpha_1, ..., \alpha_m, ..., \alpha_m)$, i.e., each F_i' is repeated N times. We claim that for some choice of l, F' computes F with error probability δ.

For a fixed input $\mathbf{x}$ let the random variable $y_{i,j}$ be the output of $F_{i,j}'$. The $y_{i,j}$'s are independent as the $F_{i,j}'$'s are disjoint.

Let $I_0 := \{i: f_i(\mathbf{x}) = 0\}$, $I_1 := \{i: f_i(\mathbf{x}) = 1\}$ and $\alpha^* = \sum_{i=1}^m \alpha_i$.

If $f_i(\mathbf{x}) = 0$ (resp. $f_i(\mathbf{x}) = 1$) then $\varepsilon \leqslant P(y_{i,j} = 1) \leqslant \delta'$ (resp. $1 - \delta' \leqslant P(y_{i,j} = 1) \leqslant 1 - \varepsilon$). The final gate in F' evaluates the sum

$$\sum_{i=1}^{m} \sum_{j=1}^{N} \alpha_i y_{i,j} = \sum_{i \in I_0} \alpha_i \sum_{j=1}^{N} y_{i,j} + \sum_{i \in I_1} \alpha_i \sum_{j=1}^{N} y_{i,j}.$$

The Chernoff inequality (see [7, 32]) implies that for $i \in I_0$,

$$P\left(\varepsilon N - N^{2/3} \leqslant \sum_{j=1}^{N} y_{i,j} \leqslant \delta' N + N^{2/3}\right) \geqslant 1 - 2e^{-2N^{1/3}}$$

and for $i \in I_1$,

$$P\left((1-\delta')\,N - N^{2/3} \leqslant \sum_{j=1}^{N} y_{i,j} \leqslant (1-\varepsilon)\,N + N^{2/3}\right) \geqslant 1 - 2e^{-2N^{1/3}};$$

thus

$$P\left(a \leqslant \sum_{i=1}^{m} \sum_{j=1}^{N} \alpha_i y_{i,j} \leqslant b\right) \geqslant 1 - 2me^{-2N^{1/3}},$$

where

$$\begin{aligned} a &= \sum_{i \in I_0} \alpha_i(\varepsilon N - N^{2/3}) + \sum_{i \in I_1} \alpha_i((1-\delta')\,N - N^{2/3}) \\ &= (\varepsilon N - N^{2/3}) \sum_{i=1}^{m} \alpha_i + N \cdot \sum_{i \in I_1} \alpha_i(1-\varepsilon-\delta') \\ &= (\varepsilon N - N^{2/3}) \cdot \alpha^* + (1-\varepsilon-\delta')\,N \cdot \sum_{i=1}^{m} \alpha_i f_i(\mathbf{x}) \end{aligned}$$

and

$$\begin{aligned} b &= \sum_{i \in I_0} \alpha_i(\delta' N + N^{2/3}) + \sum_{i \in I_1} \alpha_i((1-\varepsilon)\,N + N^{2/3}) \\ &= (\delta' N + N^{2/3}) \sum_{i=1}^{m} \alpha_i + N \cdot \sum_{i \in I_1} \alpha_i(1-\varepsilon-\delta') \\ &= (\delta' N + N^{2/3})\,\alpha^* + (1-\varepsilon-\delta')\,N \cdot \sum_{i=1}^{m} \alpha_i f_i(\mathbf{x}). \end{aligned}$$

A simple calculation shows that if $\varepsilon < \frac{1}{8}$, $\delta' < \varepsilon + 1/4ws$ and $N > (8ws)^3$ then there is an integer l between $(\delta' N + N^{2/3})\,\alpha^* + (1-\varepsilon-\delta')\,N(k-1)$ and $(\varepsilon N - N^{2/3})\,\alpha^* + (1-\varepsilon-\delta')\,Nk$. With this integer l, F' computes f with error probability $\varepsilon + 2me^{-2N^{1/3}}$ which is smaller than the required $\varepsilon + 1/t$ if $N > (\frac{1}{2}\ln(2tm))^3$. The bound on the size of F' follows by induction. ∎

Proof of Theorem 5.5. To construct C', first transform C into a formula of size $0(s^{d-1})$ and then apply Lemma 5.6 with $t = \varepsilon^{-1}$. ∎

As a corollary we obtain that functions in TC° can be computed reliably with bounded depth, polynomial size circuits of unreliable threshold gates.

COROLLARY 5.7. *Let $(f_N)_{n \in \mathbf{N}} \in TC^\circ$ and $\varepsilon < \frac{1}{8}$. Then for some d and polynomial p there is a sequence of depth-d circuits $(C_n)_{n \in \mathbf{N}}$ of ε-unreliable gates, such that C_n computes f_n with error probability* 2ε, *and the size and the weight of C_n are at most $p(n)$.*

Proof. Is clear from Theorem 5.5. ∎

Finally we mention one result for the general case of threshold circuits with $\varepsilon(v)$-unreliable gates v, where $\varepsilon(v) \in [0, \varepsilon_{\max}]$ may vary with v. We show that one can simulate a threshold circuit C by a circuit of the same depth with $\varepsilon(v)$-unreliable gates, $\varepsilon(v) \in [0, \varepsilon_{\max}]$, provided that $\varepsilon_{\max}$ is very small. More precisely we require that $\varepsilon_{\max} \leqslant 1/8\alpha$, where α is the maximal sum of weight $\sum_{i=1}^{m} |\alpha_i|$ for gates $T_k^{\boldsymbol{\alpha}}$ in C. Although this condition requires, in general, that $\varepsilon_{\max}$ become smaller when n grows, the result is of some interest in the case of threshold circuits C with $\alpha = o(n)$ (e.g., if all weights in C are constants and the fanin is $o(n)$).

THEOREM 5.8. *Let f be a function computed by a threshold circuit C of depth d, size s, and weight w. Further, assume that $\alpha := \sum_{i=1}^{m} |\alpha_i| \leqslant \alpha_0$ for every gate $T_k^{\boldsymbol{\alpha}}$ in C. Then for $\varepsilon_{\max} := 1/8\alpha_0$ there is a circuit C' of depth d, weight w, and size $O(((8\alpha_0 + \log(2s^{d-1}/\varepsilon_{\max}))^3 \cdot s^{d-1})^{d-1})$ s.t. C' computes f with error probability $\leqslant 2\varepsilon_{\max}$ for every assignment of unreliabilities $\varepsilon(v) \in [0, \varepsilon_{\max}]$ to the gates v in C'.*

Proof. By the observation before Theorem 2.3 we may assume that all weights α_i in circuit C are positive. As in the proof of Theorem 5.5 we first replace C by the corresponding threshold formula $\tilde{C}$ of size $O(s^{d-1})$. Then one proves a variation of Lemma 5.6 for an arbitrary threshold formula $\tilde{C}$ of depth d, size s^{d-1}, weight w, and $\sum_{i=1}^{m} |\alpha_i| \leqslant \alpha_0$ for every gate in $\tilde{C}$. The same construction as in Lemma 5.6 yields for $\varepsilon_{\max} \leqslant 1/8\alpha_0$ and $N \geqslant (8\alpha_0)^3 + (\log 2s^{d-1}/\varepsilon_{\max})^3$ a threshold formula C' with the desired properties (in the proof we now use the more general version of the Chernoff inequality for a sum of random variables with different expected values, see Spencer [32]). ∎

6. CIRCUITS WITH IMPRECISE THRESHOLD GATES

Let $T_k^{\boldsymbol{\alpha}}(y_1, ..., y_m)$, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_m)$, be a threshold gate. When $\sum_{i=1}^{m} \alpha_i y_i$ is near to the threshold k, a small error in the evaluation may change the output. In this section we assume that the gate makes no error if $|\sum_{i=1}^{m} \alpha_i y_i - k|$ is larger than some sensitivity parameter $S(\alpha)$, whereas the output is unpredictable otherwise.

Fix a function $S: \mathbf{N} \to \mathbf{N}$, which is called the *sensitivity function.*

DEFINITION. $T_k^{\boldsymbol{\alpha}, S}(y_1, ..., y_m)$ (where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_m)$ and $\alpha = \sum_{i=1}^{m} |\alpha_i|$) is an S-imprecise threshold gate if

(a) when $\sum_{i=1}^{m} \alpha_i y_i \geqslant k + S(\alpha)$, it outputs 1;

(b) when $\sum_{i=1}^{m} \alpha_i y_i < k - S(\alpha)$, it outputs 0;

(c) otherwise it outputs 0 or 1.

A computation of a circuit of S-imprecise gates is *legal* if the outputs of every gate satisfy (a) and (b). Thus the behavior of such a circuit is not completely determined as it may have several different legal computations. A circuit of S-imprecise gates computes a function f *precisely* if every legal computation gives the correct value.

First we observe that under some conditions on S, S-imprecise gates can be used for precise computation.

PROPOSITION 6.1. *If $S(2\beta\alpha)<\beta$ then the threshold gate $T_k^{\alpha}(y_1, ..., y_m)$ and the S-imprecise threshold gate*

$$T_{\beta(2k-1)}^{2\beta\alpha, S}(y_1, ..., y_m)$$

give the same output for every input and for every legal computation of the imprecise gate.

Proof. If $\sum_{i=1}^m \alpha_i y_i \geqslant k$ (resp. $\sum_{i=1}^m \alpha_i y_i \leqslant k-1$) then $\sum_{i=1}^m 2\beta\alpha_i y_i \geqslant \beta(2k-1)+S(2\beta\alpha)$ (resp. $\sum_{i=1}^m \beta\alpha_i y_i < \beta(2k-1) - S(2\beta\alpha)$). ∎

In the natural special case $S(\alpha)=\varepsilon\alpha$ this shows that, e.g., majority gates with fanin less than $1/2\varepsilon$ can be simulated by imprecise gates. Circuits of imprecise gates built using this simulation have large depth. The bounds below imply that, in general, small depth circuits of imprecise gates cannot be used for precise computation. The lower bounds hold for circuits which are not necessarily levelled.

For a function f, let

$$c(f, \mathbf{x}) := |\{i: f(\mathbf{x}) \neq f(\mathbf{x}^{(i)})\}|$$

(where $\mathbf{x}^{(i)}$ is $\mathbf{x}$ changed in the ith component), and

$$c(f) := \max_{\mathbf{x}} \{c(f, \mathbf{x})\}$$

be the *critical complexity* of f.

For a circuit C of S-imprecise gates let the *sensitivity ratio* of C be

$$I_S(C) := \max \left\{ \frac{\alpha}{S(\alpha)} : T_k^{\alpha, S} \text{ is a gate of } C \right\}.$$

THEOREM 6.2. *If C is a depth-d circuit of S-imprecise gates computing a function f precisely, then*

$$c(f) \leqslant (I_S(C))^d.$$

Proof. If v is a gate of C then $C_v(\mathbf{x})$ denotes the output of v assuming that every gate of C works correctly. For every input $\mathbf{x}$ we define the threshold circuit $C^{\mathbf{x}}$ as follows:

(1) the graph of $C^{\mathbf{x}}$ and the weights in each gate are the same as in C,

(2) if v is a gate of C of the form $T_k^{\alpha, S}$ and $C_v(\mathbf{x})=0$ (resp. $C_v(\mathbf{x})=1$) then the threshold value at v is $k+S(\alpha)$ (resp. $k-S(\alpha)$).

Thus $C^{\mathbf{x}}$ is a "standard" threshold circuit. Note that a computation of $C^{\mathbf{x}}$ on any input is a legal computation of C. $C_v^{\mathbf{x}}(\mathbf{y})$ is the output of v in $C^{\mathbf{x}}$ for input $\mathbf{y}$. If v corresponds to an input variable x_i we write $C_v(\mathbf{y}) = C_v^{\mathbf{x}}(\mathbf{y}) = y_i$.

LEMMA 6.3. *Let v be a gate of C of the form $T_k^{\boldsymbol{\alpha}, S}$ and $v_1, ..., v_m$ be the gates input to v. Then for every inputs $\mathbf{x}$ and $\mathbf{y}$ if $C_v(\mathbf{x}) \neq C_v^{\mathbf{x}}(\mathbf{y})$ then*

$$\sum_{j=1}^{m} |C_{v_j}(\mathbf{x}) - C_{v_j}^{\mathbf{x}}(\mathbf{y})| \cdot |\alpha_j| \geqslant S(\alpha).$$

Proof. Assume $C_v(\mathbf{x}) = 1$, $C_v^{\mathbf{x}}(\mathbf{y}) = 0$ (the other case follows the same way). Then

$$\sum_{j=1}^{m} \alpha_j C_{v_j}(\mathbf{x}) \geqslant k$$

and the definition of $C^{\mathbf{x}}$ implies that

$$\sum_{j=1}^{m} \alpha_j C_{v_j}^{\mathbf{x}}(\mathbf{y}) < k - S(\alpha). \quad \blacksquare$$

Let

$$h_v^{\mathbf{x}} := \sum_{i=1}^{n} |C_v(\mathbf{x}) - C_v^{\mathbf{x}}(\mathbf{x}^{(i)})|.$$

LEMMA 6.4. $h_v^{\mathbf{x}} \leqslant (I_S(C))^{\mathrm{depth}(v)}$.

Proof. By induction on the depth of v. If $\mathrm{depth}(v) = 1$ then $v_1, ..., v_m$ are input variables $x_{i_1}, ..., x_{i_m}$. Then

$$\begin{aligned} h_v^{\mathbf{x}} \cdot S(\alpha) &\leqslant \sum_{i=1}^{n} |C_v(\mathbf{x}) - C_v^{\mathbf{x}}(\mathbf{x}^{(i)})| \\ &\quad \cdot \sum_{j=1}^{m} |C_{v_j}(\mathbf{x}) - C_{v_j}^{\mathbf{x}}(\mathbf{x}^{(i)})|\, |\alpha_j| \\ &\leqslant \sum_{i=1}^{n} \sum_{j=1}^{m} |C_{v_j}(\mathbf{x}) - C_{v_j}^{\mathbf{x}}(\mathbf{x}^{(i)})|\, |\alpha_j| \\ &\leqslant \sum_{i=1}^{n} \sum_{\{j:\, i_j = i\}} |\alpha_j| = \alpha, \end{aligned}$$

where the first inequality follows from the previous lemma and the third inequality follows from the fact that $C_{v_j}(\mathbf{x}) \neq C_{v_j}^{\mathbf{x}}(\mathbf{x}^{(i)})$ only if $i_j = i$.

If $\text{depth}(v) = d > 1$ then

$$\begin{aligned} h_v^{\mathbf{x}} \cdot S(\alpha) &\leqslant \sum_{i=1}^{n} \sum_{j=1}^{m} |C_{v_j}(\mathbf{x}) - C_{v_j}^{\mathbf{x}}(\mathbf{x}^{(i)})| \, |\alpha_j| \\ &= \sum_{j=1}^{m} |\alpha_j| \cdot \sum_{i=1}^{n} |C_{v_j}(\mathbf{x}) - C_{v_j}^{\mathbf{x}}(\mathbf{x}^{(i)})| \\ &\leqslant \sum_{j=1}^{m} |\alpha_j| \cdot (I_S(C))^{d-1} = \alpha \cdot (I_S(C))^{d-1}, \end{aligned}$$

where the first inequality follows as above and the second inequality follows by induction. ∎

Theorem 6.2 now follows from applying Lemma 6.4 to the input $\mathbf{x}$ maximizing $C(f, \mathbf{x})$. ∎

Using results of Simon [28] and Turán [33] this leads to the following lower bounds.

COROLLARY 6.5. *Let $\varepsilon > 0$ and $S(\alpha) = \varepsilon\alpha$. If f is an arbitrary nondegenerate Boolean function (resp. a nontrivial property of r-vertex graphs where $n = \binom{r}{2}$) and C is a depth-d circuit of S-imprecise gates computing f precisely then $d = \Omega(\log \log n)$ (resp. $d = \Omega(\log n)$).*

Proof. If f is a nondegenerate Boolean function (resp. a nontrivial graph property) then $c(f) = \Omega(\log n)$ [28] (resp. $c(f) = \Omega(\sqrt{n})$ [33]). ∎

The remark after Proposition 6.1 implies that the lower bound for graph properties cannot be improved, in general.

For general sensitivity functions $S(\alpha)$ write S in the form $S(\alpha) = \alpha^{\varepsilon(\alpha)}$ $(\varepsilon(\alpha) < 1)$. We assume that $\varepsilon(\alpha)$ is a nondecreasing function (as, e.g., in the case of $S(\alpha) = \alpha^{\varepsilon}$ and $S(\alpha) = \alpha/\log \alpha$). ∎

THEOREM 6.6. (a) *If $\lim_{\alpha \to \infty} \varepsilon(\alpha) < 1$ then for every $(f_n)_{n \in N} \in TC^{\circ}$, f_n can be computed by bounded depth circuits of S-imprecise gates with polynomial size and weight.*

(b) *If $\lim_{\alpha \to \infty} \varepsilon(\alpha) = 1$ then there is no sequence of nondegenerate Boolean functions $(f_n)_{n \in \mathbf{N}}$ with $\log c(f_n) = \Omega(\log n)$ which can be computed by circuits as in* (a).

(c) If $1 - \varepsilon(\alpha) = o(\log \log \alpha / \log \alpha)$ *then there is no sequence of nondegenerate Boolean functions $(f_n)_{n \in N}$ which can be computed by circuits as in* (a).

Proof. (a) follows from Proposition 6.1, (b) and (c) follow from Theorem 6.2 similarly to Corollary 6.5. ∎

7. Threshold Quantifiers

We fix a similarity type $\mathscr{S}$ containing relations only. A_n denotes a model on the ground set $\{1, ..., n\}$ (all models considered are finite).

A threshold quantifier is of the form $\#_k^r(x_1, ..., x_r)$, where r is the arity of the quantifier and k is the threshold value. Formulas are built from atomic formulas using $\wedge$, $\vee$, $\neg$ and threshold quantifiers ($\vee$ and $\wedge$ of any number of formulas is allowed; $\#_k^r(x_1, ..., x_r)$ binds $x_1, ..., x_r$). The interpretation of the quantifiers is the following:

$$\begin{aligned} &A_n \models \#_k^r(x_1, ..., x_r)\, \psi(x_1, ..., x_r) \\ &\quad \Leftrightarrow |\{(a_1, ..., a_r) \colon A_n \models \psi(a_1, ..., a_r)\}| \geqslant k. \end{aligned}$$

(Thus $\exists$ and $\forall$ are special cases.) Let $\mathscr{A}$ be a class of structures and $\mathscr{A}_n$ be the class of structures in $\mathscr{A}$ on the ground set $\{1, ..., n\}$. A sentence ϕ_n defines $\mathscr{A}_n$ if for every A_n it holds that

$$A_n \in \mathscr{A}_n \Leftrightarrow A_n \models \phi_n.$$

A sequence $\Phi = (\phi_n)_{n \in \mathbf{N}}$ defines $\mathscr{A}$ if ϕ_n defines $\mathscr{A}_n$ for every n.

The quantifier rank (QR), depth (D), and size (S) of formulas is defined inductively. For atomic formulas each measure is 0. For $\phi = \neg\, \psi$, $\mathrm{QR}(\phi) = \mathrm{QR}(\psi)$, $\mathrm{D}(\phi) = \mathrm{D}(\psi) + 1$, $\mathrm{S}(\phi) = \mathrm{S}(\psi) + 1$. For $\phi = \phi_1 \vee \cdots \vee \phi_m$, $\mathrm{QR}(\phi) = \max_{1 \leqslant i \leqslant m} \mathrm{QR}(\phi_i)$, $\mathrm{D}(\phi) = 1 + \max_{1 \leqslant i \leqslant m} \mathrm{QR}(\phi_i)$, $\mathrm{S}(\phi) = 1 + \sum_{i=1}^m \mathrm{S}(\phi_i)$ (similarly for $\phi = \phi_1 \wedge \cdots \wedge \phi_m$). For $\phi = \#_k^r \psi$, $\mathrm{QR}(\phi) = \mathrm{QR}(\psi) + r$, $\mathrm{D}(\phi) = \mathrm{D}(\psi) + r$, $\mathrm{S}(\phi) = \mathrm{S}(\psi) + 1$. $\Phi = (\phi_n)_{n \in \mathbf{N}}$ is of bounded quantifier rank (resp. depth) and polynomial size, if for some d and polynomial p, $\mathrm{QR}(\phi_n) \leqslant d$ (resp. $D(\phi_n) \leqslant d$) and $S(\phi_n) \leqslant p(n)$ for every n.

Structures are encoded as usual (an r-ary relation on $\{1, ..., n\}$ by n^r bits). A class $\mathscr{A}$ of structures corresponds to a sequence $(f_n^{\mathscr{A}})_{n \in \mathbf{N}}$ of Boolean functions, where $f_n^{\mathscr{A}}$ is the characteristic function of $\mathscr{A}_n$ (having $p(n)$ variables for some polynomial p).

The relationship between threshold circuits and threshold quantifiers is given by the following proposition.

Proposition 7.1. *The sequence $(f_n^{\mathscr{A}})_{n \in \mathbf{N}}$ is in TC^0 if and only if $\mathscr{A}$ is defined by a Φ of bounded depth and polynomial size, where the language of Φ contains a new relation symbol P and the interpretation of P is fixed for every n.*

Proof (Outline). We consider the "only if" part as the other direction is straightforward. The argument is a modification of the proof of the analogous result for AC^0 and first-order logic in Gurevich and Lewis [14] and Immerman [16]. Let C_n be a bounded depth, polynomialsize threshold circuit computing $f_n^{\mathscr{A}}$. We may assume that all gates are majority gates. The gates, edges, and paths of C_n can be encoded by s-tuples over $\{1, ..., n\}$ for some constant s and thus the description of C_n can be encoded into a single relation of constant arity over $\{1, ..., n\}$. The desired φ_n giving the output of C_n is built recursively, using this relation, from formulas describing the output of gates on previous levels. ∎

Proposition 7.1 shows that undefinability results imply lower bounds for threshold circuits. Undefinability with special relations for P imply lower bounds for "structured" threshold circuits. Below we give an example.

Let the similarity type $\mathscr{S}$ contain $R(x, y)$ (adjacency) only, and the additional relation with a fixed interpretation be $s(x, y)$ always interpreted as the standard successor relation on $\{1, ..., n\}$). Let CONNECTIVITY $:= \{G = (V, E, S)$: $G' = (V, E)$ is a connected undirected graph, S is as above$\}$.

THEOREM 7.2. *CONNECTIVITY is not definable by any Φ (over $\mathscr{S}$ as above) of bounded quantifier rank.*

First we state a general lemma about threshold quantifiers of arity > 1.

LEMMA 7.3. *If $\mathscr{A}_n$ is defined by a sentence ϕ then it is also defined by a sentence ψ of the same quantifier rank such that ψ contains only unary threshold quantifiers.*

Proof. For simplicity we show how to eliminate a binary threshold quantifier $\#_k^2(x_1, x_2)$. Let $d^{(i)} = (d_1^{(i)}, ..., d_n^{(i)})$ $(i = 1, ..., N)$ be all possible degree sequences of directed graphs on n vertices having k edges (without multiple edges but with loops allowed). (I.e., $\sum_{j=1}^n d_j^{(i)} \cdot j = k$, and $d_j^{(i)}$ is the number of vertices with outdegree j.) Then

$$\mathscr{A}_n \models \#_k^2(x_1, x_2)\, \varphi(x_1, x_2)$$
$$\Leftrightarrow \mathscr{A}_n \models \bigvee_{i=1}^{N} \left(\bigwedge_{j=1}^{n} \#_{t_{i,j}}^1(x_1)(\#_j^1(x_2)\, \varphi(x_1, x_2)) \right),$$

where $t_{i,j} = \sum_{l=j}^n d_l^{(i)}$. ∎

(We note that the above transformation may increase the size exponentially.)

The proof of Theorem 7.2 uses a variant of the Fraïssé–Ehrenfeucht games introduced by Immerman and Lander [7]. Given $G_1 = (V_1, E_1, S_1)$ and $G_2 = (V_2, E_2, S_2)$, players I and II play m moves. In move i,

(1) player I selects a structure G_{l_i} $(l_i = 1$ or $2)$ and a set $V_{l_i}^i \subseteq V_{l_i}$,

(2) player II selects a set $V_{3-l_i}^i \subseteq V_{3-l_i}$ in the other structure with $|V_{l_i}^i| = |V_{3-l_i}^i|$,

(3) player I chooses $v_i^{3-l_i} \in V_{3-l_i}^i$,

(4) player II chooses $v_i^{l_i} \in V_{l_i}^i$.

After m moves, II wins if the bijection $v_i^1 \Leftrightarrow v_i^2$ $(1 \leqslant i \leqslant m)$ is an isomorphism between $\{v_1^1, ..., v_m^1\} \subseteq V_1$ and $\{v_1^2, ..., v_m^2\} \subseteq V_2$; otherwise I wins. $G_1 \cong_m G_2$ if II has a winning strategy in the game.

Let G_1 and G_2 be structures of size n and ϕ be a sentence with unary threshold quantifiers such that $\mathrm{QR}(\phi) \leqslant m$.

LEMMA 7.4. *If* $G_1 \cong_m G_2$ *then* $G_1 \models \phi \Leftrightarrow G_2 \models \phi$.

Proof. Omitted. It is analogous to the proofs of similar statements for other versions of the Fraïssé-Ehrenfeucht game. ∎

Proof of Theorem 7.2. From Lemmas 7.3 and 7.4 it suffices to show that for every m there are structures G_1 and G_2 such that $G_1 \in$ CONNECTIVITY, $G_2 \notin$ CONNECTIVITY and $G_1 \cong_m G_2$.

The structures $G_1 = (V_1, E_1, S_1)$, $G_2 = (V_2, E_2, S_2)$ are defined as follows:

$$V_1 = V_2 = \{1, ..., n\},\ n \geqslant 2^{m+5};$$

$S_1 = S_2$ are the standard successor relation;

$$E_1 = \left\{(i, i+1)\colon 1 \leqslant i \leqslant n-1,\ i \neq \left\lfloor \frac{2n}{7} \right\rfloor, \left\lfloor \frac{5n}{7} \right\rfloor\right\}$$
$$\cup \left\{\left(\left\lfloor \frac{n}{7} \right\rfloor, \left\lfloor \frac{6n}{7} \right\rfloor\right), \left(\left\lfloor \frac{3n}{7} \right\rfloor, \left\lfloor \frac{4n}{7} \right\rfloor\right)\right\};$$
$$E_2 = \left\{(i, i+1)\colon 1 \leqslant i \leqslant n-1,\ i \neq \left\lceil \frac{2n}{7} \right\rceil, \left\lceil \frac{5n}{7} \right\rceil\right\}$$
$$\cup \left\{\left(\left\lfloor \frac{n}{7} \right\rfloor, \left\lfloor \frac{4n}{7} \right\rfloor\right), \left(\left\lfloor \frac{3n}{7} \right\rfloor, \left\lfloor \frac{6n}{7} \right\rfloor\right)\right\}.$$

(The two structures are shown in Fig. 2). The winning strategy of player II is given as follows: Let the vertices chosen in the first i moves be $\{v_1^1, ..., v_i^1\}$ in G_1 and $\{v_1^2, ..., v_i^2\}$ in G_2. Assume that

(a) the mapping $w_j^1 \leftrightarrow w_j^2$ $(1 \leqslant j \leqslant 10)$ (see Fig. 2), $v_j^1 \leftrightarrow v_j^2$ $(1 \leqslant j \leqslant i)$ is an isomorphism between

$$A_i^1 = \{w_1^1, ..., w_{10}^1, v_1^1, ..., v_i^1\} \quad \text{and} \quad A_i^2 = \{w_1^2, ..., w_{10}^2, v_1^2, ..., v_i^2\},$$

(b) if a_1, b_1 and a_2, b_2 are corresponding vertices in A_i^1 (resp. A_i^2) then either $b_1 - a_1 = b_2 - a_2$ (i.e., the distances in the successor relation are the same) or $2^{m+1-i} < \min(|b_1 - a_1|, |b_2 - a_2|)$.

Claim. If $i < m$, player II can maintain assumptions (a) and (b) after the $(i+1)$th move. (Clearly the claim implies the existence of a winning strategy (both assumptions hold for $i = 0$).)

To prove the claim, consider interval $I_a = [a - 2^{m-i}, a + 2^{m-i}]$ for every $a \in A_i^1 \cup A_i^2$. The assumptions imply that if for $a_1, b_1 \in A_i^1$, $I_{a_1} \cap I_{b_1} \neq \emptyset$, then $b_1 - a_1 = b_2 - a_2$ (and thus, in particular, $I_{a_2} \cap I_{b_2} \neq \emptyset$), where a_2, b_2 are

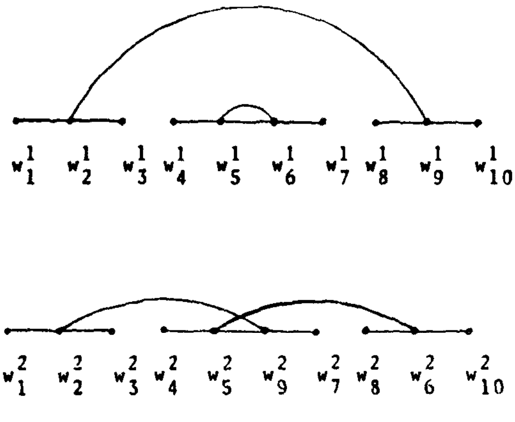

FIGURE 2

the vertices corresponding to a_1, b_1. Therefore $\bigcup_{a \in A_i^1} I_a$ and $\bigcup_{a \in A_i^2} I_a$ form isomorphic sets of intervals (considering lengths) and $|\{1, ..., n\} - \bigcup_{a \in A_i^1} I_a| = |\{1, ..., n\} - \bigcup_{a \in A_i^2} I_a|$.

Now assume that in the $(i+1)$th move I selects V_1^{i+1} in G_1. Then II selects V_2^{i+1} in G_2 as follows: he "imitates" the choice of I in $\bigcup_{a \in A_i^1} I_a$ by choosing corresponding vertices in $\bigcup_{a \in A_i^2} I_a$ and selects the same number of vertices outside these intervals arbitrarily. The assumptions guarantee that any pair of vertices selected outside the intervals defined above will maintain the assumptions for $i+1$. After this choice the choice of II in the last phase of step $i+1$ is clear. ∎

We note that the proof, in fact, implies a $\log_2 n - O(1)$ lower bound to the quantifier rank of any sentence with threshold quantifiers which defines CONNECTIVITY for n vertex graphs. Thus for this property threshold quantifiers are not more powerful than $\exists$ and $\forall$.

8. Some Open Problems

The most important open problems about threshold circuits are to separate the classes TC_d° and to show $TC^\circ \subsetneq NC^1$ (candidates showing this are mentioned in the Introduction).It would also be interesting to show that CONNECTIVITY is not in TC°.

There are open problems concerning circuits of depth 2. There are no lower bounds known for depth-2 circuits of

(a) gates computing symmetric functions (note that by Theorem 2.3 these circuits are simulated by depth-3 threshold circuits of weight 1 with quadratic increase in size),

(b) threshold gates with arbitrary weights (these circuits are called Gamba-perceptrons in [20], where the problem of proving a lower bound is also posed).

There are also no lower bounds known for probabilistic threshold circuits of depth 2.

Does Corollary 5.7 remain true if the circuits C_n are required to compute f_n reliably for every assignment of unreliabilities $\leqslant \varepsilon$ to the gates (where ε does not depend on n as in Theorem 5.8)?

Are there any nondegenerate Boolean functions computable in depth $o(\log n)$ by circuits of imprecise threshold gates with sensitivity function $S(\alpha) = \varepsilon\alpha$? (A candidate would be the "addressing function" [28] which has critical complexity $O(\log n)$.)

It would also be interesting to extend Theorem 7.2 to graphs with a linear order on the vertices.

References

1. M. Ajtai, Σ_1^1 formulae on finite structures, *Ann. Pure Appl. Logic* **24** (1984), 1–48.
2. M. Ajtai and M. Ben-Or, A theorem on probabilistic constant depth computations, *in* "Proceedings, 16th Annu. ACM Sympos. Theory of Comput., 1984," pp. 471–474.
3. L. Babai, P. Frankl, and J. Simon, Complexity classes in communication complexity, *in* "IEEE 27th Annu. Found of Comput. Sci., 1986," pp. 337–347.
4. D. A. Barrington, Bounded-width, polynomial size branching programs recognize exactly those languages in NC^1, *in* "18th Annu. Proceedings, ACM Sympos Theory of Comput., 1986," pp. 1–5.
5. D. A. Barrington and D. Thérien, Finite monoids and the fine structure of NC^1, *in* "19th Annu. Proceedings, ACM Sympos. Theory of Comput., 1987," pp. 101–109.
6. P. W. Beame, S. A. Cook, and H. J. Hoover, Log depth circuits for division and related problems, *SIAM J. Comput.* **15** (1986), 994–1003.
7. B. Bollobás, "Random Graphs," Academic Press, New York, 1985.
8. A. K. Chandra, L. J. Stockmeyer, and U. Vishkin, Constant depth reducibility, *SIAM J. Comput.* **13** (1984), 423–439.
9. B. Chor and O. Goldreich, Unbiased bits from sources of weak randomness and probabilistic communication complexity, *in* "26th Annu. IEEE Found. of Comput. Sci., 1985," pp. 429–442.
10. R. L. Dobrushin and S. J. Ortyukov, Upper bound for the redundancy of self-correcting arrangements of unreliable functional elements, *Problems Inform. Transmission* **13** (1977), 59–65.
11. P. Erdös and J. Spencer, "Probabilistic Methods in Combinatorics," Akad. Kiadó, Budapest, 1974.
12. R. Fagin, M. M. Klawe, N. J. Pippenger, and L. J. Stockmeyer, Bounded depth, polynomial size circuits for symmetric functions, *Theoret. Comput. Sci.* **36** (1985), 239–250.
13. M. Furst, J. B. Saxe, and M. Sipser, Parity, circuits, and the polynomial time hierarchy, *in* "22nd Annu. IEEE Found. of Comput. Sci., 1981," pp. 260–270.
14. Y. Gurevich and H. R. Lewis, A logic for constant depth circuits, *Inform. and Control* **61** (1984), 65–74.
15. J. Hastad, Almost optimal lower bounds for small depth circuits, *in* "18th Annu. Proceedings, ACM Sympos. Theory of Comput., 1986," pp. 6–20.
16. N. Immerman, Languages which capture complexity classes, *in* "15th Annu. Proceedings, ACM Sympos. Theory of Comput., 1983," pp. 347–354.
17. N. Immerman and E. Lander, Telling graphs apart: A first order approach to graph isomorphism, preprint, 1984.

18. O. B. LUPANOV, Implementing the algebra of logic functions in terms of bounded depth formulas in the basis +, *, −, *Dokl. Acad. Nauk. SSSR* **136** (1961), 5, 1041–1042 (Engl. transl. *Sov. Phys. Dokl.* **6** (1961), 107–108).
19. W. F. MCCOLL AND M. S. PATERSON, The depth of all Boolean functions, *SIAM J. Comput.* **6** (1977), 373–380.
20. M. MINSKY AND S. PAPERT, "Perceptrons: An Introduction to Computational Geometry," MIT Press, Cambridge, MA, 1972.
21. J. VON NEUMANN, Probabilistic logics and the synthesis of reliable organisms from unreliable components, *in* "Automata Studies" (C. E. Shannon and J. McCarthy, Eds.), pp. 43–98, Princeton Univ. Press, Princeton, NJ, 1956.
22. I. PARBERRY AND G. SCHNITGER, Parallel computation with threshold functions, *in* "Structure in Complexity Theory" (A. L. Selman, Ed.), Lect. Notes in Comput. Sci., Vol. 223, pp. 272–290, Springer-Verlag, New York/Berlin, 1986.
23. N. PIPPENGER, On networks of noisy gates, *in* "26th Annu. IEEE Found. of Comput. Sci., 1985," pp. 30–38.
24. A. A. RAZBOROV, Lower bounds for the monotone complexity of some Boolean functions, *Dokl. Akad. Nauk.* **281** (1985), 798–801.
25. A. A. RAZBOROV, Lower bounds for the size of circuits of bounded depth with basis $\{\wedge, \oplus\}$, preprint; *Mat. Zametki* **41** (1987), 598–607. English translation in: *Math. Notes of the Academy of Sciences of the USSR* **41**, 333–338.
26. J. H. REIF, On threshold circuits and polynomial computation, *in* "Proceedings, 2nd Structure in Complexity Theory Conf., 1987," pp. 118–125.
27. D. E. RUMELHARDT, J. L. MCCLELLAND, AND THE PDP RESEARCH GROUP, "Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1, MIT Press, Cambridge, MA, 1986.
28. H. U. SIMON, A tight $\Omega(\log \log n)$ bound on the time for parallel RAM's to compute nondegenerate Boolean functions, *in* "Found. in Comput. Theory, 1983," Lecture Notes in Computer Science, Vol. 158, pp. 151–153, Springer-Verlag, New York/Berlin, 1983.
29. S. SKYUM, A measure in which Boolean negation is exponentially powerful, *Inform. Process. Lett.* **17** (1983), 125–128.
30. S. SKYUM AND L. G. VALIANT, A complexity theory based on Boolean algebra, *J. Assoc. Comput. Mach.* **32** (1985), 484–502.
31. R. SMOLENSKY, Algebraic methods in the theory of lower bounds for Boolean circuit complexity, *in* "19th Annu. Proceedings, ACM Sympos. Theory of Comput., 1987," pp. 77–82.
32. J. SPENCER, "Ten Lectures on the Probabilistic Method," SIAM, Philadelphia, 1987.
33. GY. TURÁN, The critical complexity of graph properties, *Inform. Process. Lett.* **18** (1984), 151–153.
34. A. C. YAO, Separating the polynomial-time hierarchy by oracles, *in* "26th Annu IEEE Found. of Comput. Sci., 1985," pp. 1–10.