Computing Symmetric Functions with *AND/OR* Circuits and A Single *MAJORITY* Gate[*]

Zhi-Li Zhang[†] David A. Mix Barrington [†] Jun Tarui [‡]

Abstract

Fagin *et al.* characterized those symmetric Boolean functions which can be computed by small *AND/OR* circuits of constant depth and unbounded fan-in. Here we provide a similar characterization for *d-perceptrons* — *AND/OR* circuits of constant depth and unbounded fan-in with a single *MAJORITY* gate at the output. We show that a symmetric function has small (quasipolynomial, or $2^{\log^{O(1)} n}$ size) *d*-perceptrons *iff* it has only poly-log many *sign changes* (i.e., it changes value $\log^{O(1)} n$ times as the number of positive inputs varies from zero to n). A consequence of the lower bound is that a recent construction of Beigel is optimal. He showed how to convert a constant-depth unbounded fan-in *AND/OR* circuit with poly-log many *MAJORITY* gates into an equivalent *d*-perceptron — we show that more than poly-log *MAJORITY* gates cannot in general be converted to one.

[*] The first author was supported by grants CCR-8812567 and CCR-9008416. The second author was supported by NSF Computer and Computation Theory grants CCR-8922098 and CCR-9207829. The third author was supported in part by the ESPRIT II BRA Programme of the EC under contract 7141 (ALCOM II).

[†] Computer Science Department, University of Massachusetts at Amherst, Massachusetts 01003, USA. E-mail addresses: zhzhang@cs.umass.edu, and barrington@cs.umass.edu.

[‡] Department of Computer Science, University of Warwick, Coventry, CV4 7AL, United Kingdom. E-mail address: jun@dcs.warwick.ac.uk.

1 Introduction

1.1 The d-Perceptron Model

The power of constant-depth circuits of unbounded fan-in AND and OR gates (*e.g.* the well-known AC^0 circuits) is by now fairly well understood [FSS, Aj, Hå, Ya]. One of the major open problems of complexity theory is to place any non-trivial bounds on the computing power of constant depth circuits of unbounded fan-in threshold or $MAJORITY$ gates. The class TC^0, of languages recognized by polynomial-size families of such circuits, might be equal to NP for all we can prove. A natural approach to bridging the gap between AND/OR circuits and threshold circuits is to consider models which combine the two kinds of gates.

One very old example of such a model is the *perceptron* of Minsky and Papert [MP], which can be viewed as a $MAJORITY$ gate whose inputs are ANDs of the input variables. These original perceptrons are rather limited and their computing power is well understood. But recently, perceptrons have been revived in a new form [BRS]. Along with a probabilistic version, there has emerged what we will call the d-perceptron, a constant-depth unbounded fan-in circuit which has AND and OR gates except for a single $MAJORITY$ gate at the output. It has been shown that such circuits require exponential size (exponentially many gates) to compute the MOD_2 function [Gr], to approximate the MOD_2 function [ABFR], or to compute or approximate the MOD_c function for any constant c [BS]. These d-perceptrons are closely linked to a model of computation which is interesting in its own right, where one evaluates a multilinear polynomial in the input variables, with coefficients in the integers or the reals, and outputs the sign of the result. (This can be extended to polynomials over the complex numbers, using an *ad hoc* notion of "sign" [BS].) Furthermore, the d-perceptron model is robust, in that other circuit models with a limited use of threshold gates can be mapped into it [BRS, ABFR, Be].

In the study of threshold computations, *harmonic analysis* has found many interesting applications (for some recent examples, see [Br, KKL, LMN]). Of particular interest to us is the work by Linial *et al.* [LMN], where they showed that any AC^0 function (any function computable by a poly-size constant-depth AND/OR circuit) can be closely approximated by a low-degree polynomial, a result which has consequences for the learnability of AC^0 functions. This work, together with [ABFR, BS], has been the chief inspiration for our work. However, our proof of the lower bound result is based on the *random restriction* technique [FSS, Hå, Ya].

1.2 Complexity of Symmetric Functions

The *symmetric* boolean functions are those which are invariant under any permutation of the inputs. We can describe a symmetric boolean function by giving its *spectrum*, which is the sequence $\langle f(0), \ldots, f(n) \rangle$, where each $f(i)$ is the value of the function when i of the n inputs are one. All symmetric functions are in TC^0, because they have linear-size depth-2 threshold circuits. In any model the complexity theory of the symmetric functions forms a subtheory of that of all boolean functions, and in some models this theory can be interesting and beautiful.

For example, consider the well-understood model of constant-depth unbounded fan-in AND/OR circuits. A theorem of Fagin *et al.* [FKPS], using the exponential lower bound for $PARITY$ due to Yao [Ya, Hå], gives an elegant characterization of the symmetric functions which have polynomial size in this model (are in the class AC^0). These functions are those whose spectra are constant except for a poly-log section at either end. That is, there is some function $g(n) = \log^{O(1)} n$ such that for each n, $f(i)$ is constant in the range $g(n) \leq i \leq n - g(n)$. Our principal result is that a similar characterization holds in the d-perceptron model we consider.

1.3 The Main Result

In the d-perceptron model we add a single $MAJORITY$ gate to the AND/OR circuit, and thus we immediately allow new symmetric functions, such as $MAJORITY$ itself, to be computed. In previous work in this model [ABFR], two key parameters of a boolean function have proven to be its *strong degree* and *weak degree*. These are based on a space of polynomials over the real numbers, where the boolean domain is taken to be $\{-1,1\}$ rather than $\{0,1\}$. The strong degree is the minimal degree of a polynomial whose sign always agrees with the target boolean function. The weak degree is the minimal degree of a polynomial, not identically zero, whose sign agrees with the target boolean function whenever the polynomial is nonzero. For symmetric functions, these two degrees are equal [ABFR], and furthermore they are equal to the number of *sign changes* of the spectrum (the number of i for which $f(i) \neq f(i+1)$). It turns out that this parameter of symmetric functions give us an exact characterization of the symmetric boolean functions computable in the d-perceptron model as stated below.

Theorem 1 *A symmetric boolean function can be computed by a quasipolynomial size d-perceptron* iff *it has only poly-log many sign changes.*

The organization of the paper is as follows. In section 2 we define some notations and terminologies. In section 3 we give the easier part of the proof of Theorem 1 — the upper bound. In section 4 we give the harder part of the proof of Theorem 1 — the lower bound. Finally in section 5 we conclude our work and present some open problems.

2 Preliminaries

We will consider functions from $\{-1,1\}^n$ to the reals R (with boolean functions being the special case with range $\{-1,1\}$) as multilinear polynomials over R with input variables $\{x_1,\ldots,x_n\}$. The *size* of a polynomial is the number of nonzero coefficients, and the *degree* is the maximum number of variables appearing in any term with nonzero coefficient. We use $[n]$ to denote the set $\{0,1,2,\ldots,n\}$, and by $|\underline{x}|$ we mean the number of -1's in $\underline{x}$ (in general we think of -1 as "true" and 1 as "false").

A symmetric boolean function is a boolean function whose value only depends on $|\underline{x}|$. It can be proved that, over the reals, we can regard a symmetric boolean function as a function of $x = |\underline{x}|$. Hence, in this way, we convert a n-variable symmetric boolean function $f(\underline{x})$ (where $\underline{x} \in \{-1,1\}^n$) into a univariate real function $f'(x)$ (where $x = |\underline{x}|$) such that $deg(f) =$

$deg(f')$. In the sequel, we will use f to denote $f(\underline{x})$ and $f'(x)$ interchangeably. If $i \in [n]$, we say i is a *sign change* of a symmetric function f if $f(i) \neq f(i+1)$. The number of sign changes of f is equal to the cardinality of the set $\{i | f(i) \neq f(i+1)\}$. We will call $\langle f(0), f(1), \ldots, f(n) \rangle$ the *sign change spectrum* of f.

Following [ABFR], we define strong and weak representations of boolean functions as follows.

Definition 1 *We say a polynomial $F(\underline{x})$ over the reals R* strongly represents *a boolean function $f(\underline{x})$ if $sgn(F(\underline{x})) = f(\underline{x})$ for all $\underline{x} \in \{-1, 1\}^n$. Here $sgn(F(\underline{x})) = 1$ if $F(\underline{x}) > 0$ and $sgn(F(\underline{x})) = -1$ if $F(\underline{x}) < 0$.*

We say a polynomial $F(\underline{x})$ over the reals R weakly represents *a boolean function $f(\underline{x})$ if $F(\underline{x})$ is not identically zero and for all $\underline{x} \in \{-1, 1\}^n$ such that $F(\underline{x}) \neq 0$, $sgn(F(\underline{x})) = f(\underline{x})$.*

Definition 2 *Let $f(\underline{x})$ be a boolean function. The* strong degree *of $f(\underline{x})$ (denoted $d_s(f)$) is the minimum degree among all polynomials strongly representing $f(\underline{x})$, and the* weak degree *of $f(\underline{x})$ (denoted $d_s(f)$) is the minimum degree among all polynomials weakly representing $f(\underline{x})$.*

Notice that in general the weak degree of a boolean function may well be smaller than its strong degree. However, the following fact as first observed in [ABFR] says that this cannot happen for symmetric boolean functions.

Lemma 2 ([ABFR]) *Let $f(\underline{x})$ be a symmetric boolean function with k sign changes, then $d_s(f) = d_w(f) = k$.*

We call this quantity the *degree* of a symmetric boolean function.

The above lemma was proved in [ABFR] by exploring the duality relationship of certain function spaces. We note that it has a simpler proof using the symmetrization technique [MP]: Let $F(\underline{x})$ be a strong or weak representation of a symmetric boolean function $f(\underline{x})$. Note that $F(\underline{x})$ is not necessarily symmetric itself; however, we can easily use $F(\underline{x})$ to construct a symmetric function $G(\underline{x})$. Formally, $G(\underline{x}) = \sum_{\sigma \in S_n} F^{\sigma}(\underline{x})$, where S_n is the nth symmetric group and $F^{\sigma}(\underline{x}) = F(\underline{x}^{\sigma}) = F(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$. Obviously $G(\underline{x})$ strongly (weakly) represents $f(\underline{x})$ if $F(\underline{x})$ strongly (weakly) represents

$f(\underline{x})$, and $G(\underline{x})$ has the same degree as $F(\underline{x})$. Since $G(\underline{x})$ is symmetric, $G(\underline{x})$ can be written as a univariate real function of $|\underline{x}|$. Therefore the degree of $G(\underline{x})$ is at least the number of sign changes of $f(\underline{x})$, and thus that of $F(\underline{x})$. But it is easy to construct a real polynomial strongly representing $f(\underline{x})$ such that its degree equals the number of sign changes of $f(\underline{x})$.

3 The Upper Bound

We first define the d-perceptron model as introduced in [ABFR, BRS].

Definition 3 *A* d-perceptron *is a circuit with a $MAJORITY$ gate at the top and depth-d unbounded fan-in AND/OR subcircuits feeding into the $MAJORITY$ gate. The size of a d-perceptron is the number of gates in the circuit.*

We will be interested in d-perceptrons of polynomial and of *quasipolynomial* ($2^{\log^{O(1)} n}$) size. Note that by [FKPS], ordinary constant-depth AND/OR circuits of quasipolynomial size can compute no more symmetric functions than can circuits of polynomial size. For more on quasipolynomial size circuit classes, see [Ba].

Consider a polynomial strongly representing a boolean function on $\{0,1\}^n$ such that the coefficients are positive integers bounded by a quasipolynomial in n. It is easy to see that such a polynomial corresponds to a quasipolynomial size 1-perceptron whose gates on the first level are poly-log fan-in ANDs. In the following lemma we show that any symmetric boolean function with only poly-log degree (*i.e.* sign changes) has such a low degree polynomial representation, and hence can be computed by such 1-perceptrons.

Lemma 3 *Any symmetric boolean function with only poly-log degree can be computed by a quasipolynomial size* 1*-perceptron with ANDs of poly-log fan-in.*

Proof: Let $0 \leq c_1 < c_2 < \ldots < c_k < n$ be the positions of the sign changes of $f(\underline{x})$, where $k = \log^{O(1)} n$. Then the following function

$$F(\underline{x}) = (-1)^{\delta} \prod_{i=1}^{k} (c_i + \frac{1}{2} - \sum_{i=1}^{n} x_i), \text{ where } \delta = \begin{cases} 0 & \text{if } f(0) > 0 \\ 1 & \text{if } f(0) < 0 \end{cases}$$

agrees with $f(\underline{x})$ in sign (note here $\underline{x} \in \{0,1\}^n$).

F($\underline{x}$) is a poly-log degree polynomial with rational coefficients, but we can easily make $F(\underline{x})$ a polynomial with integer coefficients (without changing its sign) by multiplying it by an appropriate positive constant.

The problem left now is to convert the polynomial into one with only positive coefficients. Reversing the procedure used in [MP] to prove the *Positive Normal Form Theorem*, we can eliminate all the negative coefficients without blowing up either the degree or the size of the coefficients. ∎

4 The Lower Bound

In this section, we will prove that any symmetric boolean function with more than poly-log sign changes cannot be computed by any quasipolynomial size d-perceptron. The proof makes use of some key observations by Linial, et al [LMN], and uses the very technique of *random restriction*, first introduced in [FSS] and refined in [Ya, Hå], that gave the first exponential lower bound for AND/OR circuits.

A random restriction is a random mapping of the input variables to 0, 1 and $*$ according to some probability distribution. The function obtained from $f(x_1, \ldots, x_n)$ by applying a random restriction ρ is denoted by f^{ρ}, and its variables are those x_i for which $\rho(x_i) = *$. For our purpose, we will assume that ρ assigns values to each input variable independently and $Pr[0] = Pr[1] = \frac{1-Pr[*]}{2}$.

A simple observation is that any random restriction of a symmetric boolean function is still symmetric; furthermore, its sign change spectrum is a subinterval of that of the original function.

Recall that a *minterm* of a boolean function is a set of variables such that a partial assignment to the variables in the set makes the function identically

1, but no partial assignment to any subset of the set makes the function identically 1. Similarly, a *maxterm* is a set of variables such that a partial assignment to the variables in the set makes the function identically 0, but no partial assignment to a subset of the set makes the function identically 0.

A useful fact, which was independently discovered in [BI,HH, Ta], and explicitly stated in [LMN], states that if all the minterms and maxterms of a boolean function f have size at most s and t respectively, then f can be evaluated by a decision tree of depth at most st. Since each branch of the decision tree corresponds to a monomial over the reals, we see that f can be represented as a real polynomial of degree at most st. This observation will be used in the proof of lemma 5 below.

It is well-known that with high probability, a random restriction of an AC^0 function will have small minterm size and maxterm size. This can be proved by a repeated applications of Håstad's switching lemma [BoS, LMN]. We state this fact formally as follows.

Lemma 4 ([LMN]) *Let f be a boolean function computed by an AND/OR circuit of size M and depth d. Then*

$$Pr[f^\rho \text{ has a minterm or a maxterm of size } > t] \leq M2^{-t}$$

where ρ is a random restriction such that $Pr[*] = 1/(10t)^d$.

Lemma 5 *Let f be a symmetric boolean function on n variables. Suppose f can be computed by a d-perceptron such that the fan-in of the $MAJORITY$ gate is N, the size of each AND/OR subcircuit is at most M, and the depth is at most d, where $N, M \leq 2^{t-1}$. Then for a positive fraction of the random restrictions ρ in a distribution with $Pr[*] = p = \frac{n}{(10t)^d}$, f^ρ is a function of at least $np = \frac{n}{(10t)^d}$ variables, the number of sign changes of f^ρ is at most $O(t^2)$, and the sign change spectrum of f^ρ is a subinterval with its center at most $O(\sqrt{n})$ off the center of the sign change spectrum of f.*

Proof: Let ρ be a random restriction with $Pr[*] = p = \frac{1}{(10t)^d}$. Denote by f_i, $1 \leq i \leq N$, the subfunctions computed by the AND/OR subcircuits, and by

f_i^ρ, $1 \le i \le N$, the functions obtained by the random restriction. Applying Lemma 4 to each f_i, we have

$$Pr[f_i^\rho \text{ has a minterm or maxterm of size } > 2t] \le M2^{-2t}.$$

Hence,

$$\begin{aligned} & Pr[\bigwedge_{i=1}^{N} f_i^\rho \text{ has only minterms and maxterms of size } \le 2t] \\ \ge & \ 1 - \sum_{i=1}^{N} Pr[f_i^\rho \text{ has a minterm or maxterm of size } > 2t] \\ \ge & \ 1 - NM2^{-2t} \ge 1 - 2^{2t-2}2^{-2t} \ge \frac{3}{4} \end{aligned}$$

On the other hand, the expected number of variables assigned $*$ is $np = \frac{n}{(10t)^d}$. By the normal approximation to binomial distribution, for n large, we see that with probability at least $\frac{1}{4}$, ρ will assign $*$'s to at least np variables and an almost equal number of 0 and 1's to the rest of the input variables.

Therefore, there must be a ρ such that $f^\rho = MAJORITY(f_1^\rho, \ldots, f_N^\rho)$ is a function on at least np variables, and each f_i^ρ has both minterms and maxterms of size $\le 2t$. It follows that f_i^ρ can be represented by a $(-1,1)$-valued real function of degree at most $4t^2$, hence $g^\rho = \sum_{i=1}^N f_i^\rho - \frac{1}{2}$ is a strong representation of f^ρ. Since g^ρ has degree at most $4t^2$, f^ρ can have at most $4t^2$ sign changes. ■

Remark: If we choose $t = \log^{O(1)} n$, then f^ρ can have only poly-log many sign changes. Therefore for the original function f, there must exist a subinterval of length at least $\frac{n}{(10)^d} = \frac{n}{\log^{O(1)} n}$, near the center of the sign change spectrum of f, such that f has at most poly-log many sign changes in that interval.

To prove the result that any symmetric function of more than poly-log sign changes cannot be computed by any quasipolynomial size d-perceptron, we need to use a shifting technique to locate an interval in the sign change spectrum of the function such that we can apply the above lemma to obtain a contradiction.

Lemma 6 *If f is a symmetric boolean function of more than poly-log sign changes, then f cannot be computed by a quasipolynomial size d-perceptron for any constant d.*

Proof: Suppose the opposite is true: there exists a quasipolynomial size d-perceptron for some constant d. Let c be such that $N, M \leq 2^{\log^c n - 1}$ where N, M are as in lemma 5. Let $s(n)$ be the sign change function of f, by the hypothesis $s(n) = \log^{\omega(1)} n$.

Consider the interval $[s^{\frac{1}{2}}(n), n - s^{\frac{1}{2}}(n)]$ of the sign change spectrum of f, the number of sign changes in this interval is $\Omega(s(n))$. Without loss of generality, we assume that there are $\Omega(s(n))$ sign changes in $[s(n)^{\frac{1}{2}}, \frac{n}{2}]$. Partition this interval into k intervals of the form $[2^i s^{\frac{1}{2}}(n), 2^{i+1} s^{\frac{1}{2}}(n)]$, where $0 \leq i \leq k-1$ and $k = \log n - \frac{1}{2} \log s(n) - 1 = O(\log n)$.

We further partition each of the intervals $[2^i s^{\frac{1}{2}}(n), 2^{i+1} s^{\frac{1}{2}}(n)]$ into δ subintervals of length $\frac{2^i s^{\frac{1}{2}}(n)}{\delta}$ where $\delta = (10t)^d$ and $t = \log^c n$, *i.e.* $\delta = (10 \log^c n)^d = O(\log^{dc} n)$. We contend that one of the subintervals must have $\omega(t^2)$ sign changes, since otherwise, the total number of sign changes in $[s^{\frac{1}{2}}(n), \frac{n}{2}]$ is at most $\sum_{i=0}^{k-1} O(t^2)\delta = O(t^2 \delta k) = \log^{O(1)} n$, a contradiction. Therefore for some i, $0 \leq i \leq k-1$, a subinterval of length $\frac{2^i s^{\frac{1}{2}}(n)}{\delta} = \log^{\omega(1)} n$ has $\omega(t^2)$ sign changes.

By an appropriate partial assignment to the input variables, we obtain from f a function f', of $2^i s^{\frac{1}{2}}(n)$ variables, whose sign change spectrum is identical to an interval of the sign change spectrum of f which contains the aforementioned subinterval at center. Note that the circuit for f induces a circuit for f' of size at most that for f and thus its N', M' are bounded by $2^{t-1} = 2^{\log^c n - 1}$. Therefore, applying lemma 5 to f', we have that there exists a random restriction ρ such that f'^{ρ} contains the subinterval as its sign change spectrum. However, f'^{ρ} can have only $O(t^2) = \log^{O(1)} n$ many sign changes, hence we arrive at a contradiction. ∎

We have now completed the proof of our main Theorem 1, by combining Lemma 3 and Lemma 6. We conclude with a consequence of Lemma 6 — the optimality of a recent construction of Beigel [Be].

Beigel shows that d-perceptrons serve as a normal form for AND/OR

circuit s augmented by a small number of $MAJORITY$ gates. In particular, a circuit of unbounded fan-in, quasipolynomially many AND, OR, and NOT gates, and poly-log many $MAJORITY$ gates can be converted into a d-perceptron of quasipolynomial size. We can now show there is no general way to eliminate more than poly-log many $MAJORITY$s in this way.

Corollary 7 *Let $m = \log^{\omega(1)} n$. There exists a symmetric boolean function computable by circuits of $\Theta(m)$ $MAJORITY$ gates (with no other gates) but not computable by any d-perceptron family of quasipolynomial size.*

Proof: By Lemma 6, any symmetric function with exactly m sign changes will do. ■

5 Conclusion and Open Problems

In this paper, we proved an *if and only if* condition for a symmetric boolean function to be computable by a quasipolynomial size d-perceptron circuit. This work extends the line of an earlier work by Fagin, *et al* [FKPS] where they gave an *if and only if* condition for a symmetric function to be computable by a polynomial size AND/OR circuit.

In an attempt to capture symmetric functions in other models, in [ZB] we also studied the size complexity of symmetric functions in the parity-threshold model, *i.e.* circuits consisting of a $MAJORITY$ gate whose inputs are $PARITY$ gates. We conjectured that an analogous *if and only if* condition exists, but we were only able to partially resolve the problem under a certain technical condition. One particular case of interest is that for any constant $p > 2$, MOD_p is not computable by any quasipolynomial size parity-threshold circuit.

The analysis of threshold computation by algebra over fields of characteristic zero has proved somewhat fruitful. The computation of circuits of AND, OR, and MOD_p gates has been very well explained using algebra over fields of characteristic p ([Ra], [Sm]). Is it possible to combine the two methods, or otherwise place limits on the power of the following perceptron-like model:

a $MAJORITY$ gate, whose inputs are constant-depth $AND/OR/MOD_p$ circuits?

References

[Aj] M. Ajtai. $\sum_1^1$-formulae on finite structures. *Annals of Pure and Applied Logic*, **24** (1983), 1-48.

[ABFR] J. Aspnes, R. Beigel, M. Furst and S. Rudich. On the expressive power of voting polynomials. *Proceedings of the 23rd Annual Symposium on Theory of Computing* (1991), 402-409.

[Ba] D. A. M. Barrington. Quasipolynomial size circuit classes. *Proceedings: Structure in Complexity Theory, Seventh Annual Conference* (1992), 86-93.

[Be] R. Beigel. Do extra threshold gates help? *Proceedings of the 24th Annual Symposium on Theory of Computing* (1992), 450-454.

[BI] M. Blum and R. Impagliazzo. Generic oracles and oracle classes. *Proceedings of the 28th Annual Symposium on Foundations of Computer Science* (1987), 118-126.

[BoS] R. Boppana and M. Sipser. The complexity of finite functions. *Handbook of Theoretical Computer Science*, Vol. A, ed. by J. van Leeuwen (Elsevier and MIT Press, 1990).

[Br] J. Bruck. Harmonic analysis of polynomial threshold functions. *SIAM J. Disc. Math.* **3:2** (1990), 168-177.

[BRS] R. Beigel, N. Reingold and D. Spielman. The perceptron strikes back. *Proceedings of the 6th Annual Conference on Structrur in Complexity Theory* (1991), 286-291.

[BS] D. A. Mix Barrington and H. Straubing. Complex polynomials and circuit lower bounds for modular counting. *Proceedings of*

LATIN '92 (1st Latin American Symposium on Theoretical Informatics) (1992), 24-31.

[FKPS] R. Fagin, M. M. Klawe, N. J. Pippenger, and L. Stockmeyer. Bounded depth, polynomial size circuits for symmetric functions. *Theoretical Computer Science* **36** (1985), 239-250.

[FSS] M. Furst, J. Saxe, and M. Sipser. Parity, circuits, and the polynomial time hierarchy. *Math. System Theory* **17** (1984), 13-27.

[Gr] F. Green. An oracle separating $\oplus P$ from PP^{PH}, *Proc. 5th Structure in Complexity Theory* (1990), 295-298.

[Hå] J. Håstad. *Computational Limitations of Small-Depth Circuits.* (Cambridge, MA, MIT Press, 1986).

[HH] J. Hartmanis and L. A. Hemachandra. One-way functions, robustness and non-isomorphism of NP-complete sets. Technical Report DCS TR86-796 (1987), Cornell University.

[KKL] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. *Proceedings of 29th Annual ACM Symposium on Theory of Computing* (1988), 68-80.

[LMN] N. Linial, Y. Mansour and N. Nisan. Constant depth circuit, fourier transform and learnability. *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science* (1989), 574-579.

[MP] M. L. Minsky and S. Papert. *Perceptrons* (Cambridge, MA, MIT Press, 1988). Original edition 1968.

[Ra] A. A. Razborov. Lower bounds for the the size of circuits of bounded depth with basis $\wedge$, $\oplus$. *Math. Zametki* **41:4** (1987), 598-607 (in Russian). English translation *Math. Notes Acad. Sci. USSR* **41:4** (1987), 333-338.

[Sm] R. Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. *Proceedings of 19th Annual ACM Symposium on Theory of Computing* (1987), 77-82.

[Ta] G. Tardos. Query complexity, or why is it difficult to separate $NP^A \cap$ co-NP^A from P^A by a random oracle A? Manuscript (1988).

[Ya] A. C.-C. Yao. Separating the polynomial-time hierarchy by oracles. *Proceedings 26th Annual IEEE Symposium on Foundations of Computer Science* (1985), 1-10.

[ZB] Z.-L. Zhang and D. A. Mix Barrington. Lower bounds for symmetric functions in perceptron-like models. COINS Technical Report 91-81 (1991), University of Massachusetts at Amherst.