THEORY OF COMPUTING, Volume 9 (20), 2013, pp. 653–663
www.theoryofcomputing.org

Approximating the AND-OR Tree

Alexander A. Sherstov*

Received February 12, 2013; Revised May 25, 2013; Published June 19, 2013

Abstract: The *approximate degree* of a Boolean function f is the least degree of a real polynomial that approximates f within $1/3$ at every point. We prove that the function $\bigwedge_{i=1}^{n} \bigvee_{j=1}^{n} x_{ij}$, known as the *AND-OR tree*, has approximate degree $\Omega(n)$. This lower bound is tight and closes a line of research on the problem, the best previous bound being $\Omega(n^{0.75})$. More generally, we prove that the function $\bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} x_{ij}$ has approximate degree $\Omega(\sqrt{mn})$, which is tight. The same lower bound was obtained independently by Bun and Thaler (2013) using related techniques.

ACM Classification: F.0, F.1.3

AMS Classification: 68Q05, 68Q87

Key words and phrases: AND-OR tree, polynomial approximation, polynomial representations of Boolean functions, approximate degree

1 Introduction

Over the past two decades, representations of Boolean functions by real polynomials have played an important role in theoretical computer science. The surveys [7, 31, 10, 32, 1] provide a fairly comprehensive overview of this body of work. Several kinds of representation [24, 23, 5, 7, 25] have been studied, depending on the intended application. For our purposes, a real polynomial p represents a Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ if

$$|f(x) - p(x)| \leq \frac{1}{3}$$

for every $x \in \{0,1\}^n$, In other words, we are interested in the pointwise approximation of Boolean functions by real polynomials. The least degree of a real polynomial that approximates f pointwise within

*Supported by NSF CAREER award CCF-1149018.

 DOI: 10.4086/toc.2013.v009a020

$1/3$ is called the *approximate degree* of f, denoted $\widetilde{\deg}(f)$. The constant $1/3$ is chosen for aesthetic reasons and can be replaced by any other in $(0,1/2)$ without affecting the theory in any way.

The formal study of the approximate degree began in 1969 with the seminal work of Minsky and Papert [23], who famously proved that the parity function in n variables cannot be approximated by a polynomial of degree less than n. Since then, the approximate degree has been used to solve a vast array of problems in complexity theory and algorithm design. The earliest use of the approximate degree was to prove circuit lower bounds and oracle separations of complexity classes [27, 41, 5, 20, 21, 35]. Over the past decade, the approximate degree has been used many times to prove tight lower bounds on quantum query complexity, e. g., [6, 9, 2, 18]. The approximate degree has enabled remarkable progress [12, 28, 11, 36, 29, 32] in communication complexity, with complete resolutions of difficult open problems. The results listed up to this point are of *negative* character, i. e., they are lower bounds in relevant computational models. More recently, the approximate degree has found important algorithmic applications. In computational learning theory, the approximate degree has been used to obtain the fastest known algorithms for PAC-learning DNF formulas [42, 19] and read-once formulas [4] and the fastest known algorithm for agnostically learning disjunctions [17]. Another well-known use of the approximate degree is an algorithm for approximating the inclusion-exclusion formula based on its initial terms [22, 16, 33, 43].

These applications motivate the study of the approximate degree as a complexity measure in its own right. As one would expect, methods of approximation theory have been instrumental in determining the approximate degree for specific Boolean functions of interest [8, 25, 40, 2, 3, 33, 39]. In addition, quantum query algorithms have been used to prove *upper* bounds on the approximate degree [15, 43, 4, 30], and duality-based methods have yielded *lower* bounds [26, 34, 38]. Nevertheless, our understanding of this complexity measure remains fragmented, with few general results available [25, 39].

The limitations of known techniques are nicely illustrated by the so-called AND-OR tree,

$$f(x) = \bigwedge_{i=1}^{n} \bigvee_{j=1}^{n} x_{ij}\,.$$

Despite its seeming simplicity, it has been a frustrating function to analyze. Its approximate degree has been studied for the past 19 years [25, 40, 15, 3, 34] and has been recently re-posed as an open problem by Aaronson [1]. Table 1 gives a quantitative summary of this line of research. The best lower and upper bounds prior to this paper were $\Omega(n^{0.75})$ and $O(n)$, respectively. Our contribution is to close this gap by improving the lower bound to $\Omega(n)$. We obtain the following more general result.

Theorem 1.1 (Main result)**.** *The function $f(x) = \bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} x_{ij}$ has approximate degree*

$$\widetilde{\deg}(f) = \Omega(\sqrt{mn})\,.$$

This lower bound is tight for all m and n, by the results of Høyer et al. [15].

1.1 Proof overview

The problem of approximating a given function f pointwise to within error ε by polynomials of degree at most d can be viewed as a search for a point in the intersection of two convex sets, namely, the

Bound	Reference
$O(n)$	Høyer, Mosca, and de Wolf [15]
$\Omega(\sqrt{n})$	Nisan and Szegedy [25]
$\Omega(\sqrt{n \log n})$	Shi [40]
$\Omega(n^{0.66\ldots})$	Ambainis [3]
$\Omega(n^{0.75})$	Sherstov [34]
$\Omega(n)$	This paper

Table 1: Approximate degree of the AND-OR tree.

ε-neighborhood of f and the set of polynomials of degree at most d. As a result, the *nonexistence* of an approximating polynomial for f is equivalent to the *existence* of a so-called dual polynomial for f, whose defining properties are orthogonality to degree-d polynomials and large inner product with f. Geometrically, the dual polynomial is a separating hyperplane for the two convex sets in question.

Our proof is quite short (barely longer than a page). We view $f(x) = \bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} x_{ij}$ as the componentwise composition of the functions AND_m and OR_n. We use the dual polynomial for OR_n to prove the existence of an operator L with the following properties:

(i) L linearly maps functions $\{0,1\}^{m \times n} \to [-1,1]$ to functions $\{0,1\}^m \to [-1,1]$;

(ii) L decreases the degree of the function to which it is applied by a factor of $\Omega(\sqrt{n})$;

(iii) $Lf \approx \text{AND}_m$ pointwise.

The existence of L directly implies our main result. Indeed, for any polynomial p that approximates f pointwise, the polynomial Lp has degree $\Omega(\sqrt{n})$ times smaller and approximates AND_m pointwise; since the latter approximation task is known [25] to require degree $\Omega(\sqrt{m})$, the claimed lower bound of $\Omega(\sqrt{mn})$ on the degree of p follows.

What makes the construction of L possible is the following very special property of any dual polynomial for OR_n: it maintains the same sign on $\text{OR}_n^{-1}(0)$ and has almost half of its ℓ_1 norm there. We call such dual polynomials *one-sided*. This property was proved several years ago by Gavinsky and the author in [14], where it was used to obtain lower bounds for nondeterministic and Merlin-Arthur communication protocols.

1.2 Independent work by Bun and Thaler

In an upcoming paper, Bun and Thaler [13] independently prove an $\Omega(\sqrt{mn})$ lower bound on the approximate degree of $f(x) = \bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} x_{ij}$. The proof in [13] and ours are both based on the fact that OR_n has a one-sided dual polynomial. The two papers differ in how they use this fact to prove an $\Omega(\sqrt{mn})$ lower bound on the approximate degree. The treatment in this paper is a combination of the dual view (one-sided dual polynomial for OR_n) and the primal view (construction of an approximating polynomial for AND_m). The treatment in [13] is a refinement of [34] and uses exclusively the dual view (construction of a dual polynomial for f using dual polynomials for AND_m and OR_n). In our opinion, the proof in this paper has the advantage of being shorter and simpler. On the other hand, the approach in [13] has the advantage of giving an explicit dual polynomial for f, which is of interest because explicit dual polynomials have found several uses in communication complexity [32].

2 Preliminaries

For a function $f\colon X\to\mathbb{R}$ on a finite set X, we let $\|f\|_\infty=\max_{x\in X}|f(x)|$. The total degree of a multivariate real polynomial $p\colon\mathbb{R}^n\to\mathbb{R}$ is denoted $\deg p$. We use the terms *degree* and *total degree* interchangeably in this paper. For a function $f\colon X\to\mathbb{R}$ on a finite set $X\subset\mathbb{R}^n$, the ε*-approximate degree* $\deg_\varepsilon(f)$ of f is defined as the least degree of a real polynomial p with $\|f-p\|_\infty\leq\varepsilon$. Throughout this paper, we will work with the ε-approximate degree for a small constant $\varepsilon>0$. For Boolean functions $f\colon X\to\{0,1\}$, the choice of constant $0<\varepsilon<1/2$ affects the quantity $\deg_\varepsilon(f)$ by at most a constant factor:

$$c\deg_{1/3}(f)\leq\deg_\varepsilon(f)\leq C\deg_{1/3}(f)\,, \tag{2.1}$$

where $c=c(\varepsilon)$ and $C=C(\varepsilon)$ are positive constants. By convention, one studies $\varepsilon=1/3$ as the canonical case and reserves for it the special symbol $\widetilde{\deg}(f)=\deg_{1/3}(f)$. A dual characterization [36, 38] of the approximate degree is as follows.

Fact 2.1. *Let $f\colon X\to\mathbb{R}$ be given, for a finite set $X\subset\mathbb{R}^n$. Then $\deg_\varepsilon(f)\geq d$ if and only if there exists a function $\psi\colon X\to\mathbb{R}$ such that*

$$\sum_{x\in X}|\psi(x)|=1\,,$$
$$\sum_{x\in X}\psi(x)f(x)>\varepsilon\,,$$

and

$$\sum_{x\in X}\psi(x)p(x)=0$$

for every polynomial p of degree less than d.

We adopt the usual definitions of the Boolean functions $\mathrm{AND}_n,\mathrm{OR}_n\colon\{0,1\}^n\to\{0,1\}$. Their approximate degree was determined by Nisan and Szegedy [25].

Theorem 2.2 (Nisan and Szegedy)**.** *The functions AND_n and OR_n obey*

$$\deg_{1/3}(\mathrm{AND}_n)=\deg_{1/3}(\mathrm{OR}_n)=\Theta(\sqrt{n})\,.$$

By combining the above two theorems, Gavinsky and the author [14, Thm. 5.1] obtained the following result, which plays a key role in this paper.

Theorem 2.3 (Gavinsky and Sherstov)**.** *Fix any constant* $0 < \varepsilon < 1$. *Then there exists a constant* $\delta = \delta(\varepsilon) > 0$ *and a real function* $\psi\colon \{0,1\}^n \to \mathbb{R}$ *such that*

$$\sum_{x\in\{0,1\}^n} |\psi(x)| = 1\,, \tag{2.2}$$

$$\psi(0,0,\dots,0) < -\frac{1-\varepsilon}{2}\,, \tag{2.3}$$

and

$$\sum_{x\in\{0,1\}^n} \psi(x)p(x) = 0 \tag{2.4}$$

for every polynomial p of degree less than $\delta\sqrt{n}$.

For the sake of completeness, we include the proof.

Proof of Theorem 2.3 (adapted from [14]). Recall from Theorem 2.2 that $\deg_{1/3}(\mathrm{OR}_n) = \Omega(\sqrt{n})$. Therefore, (2.1) shows that $\deg_{\frac{1-\varepsilon}{2}}(\mathrm{OR}_n) \geq \delta\sqrt{n}$ for a sufficiently small constant $\delta = \delta(\varepsilon) > 0$. Now the dual characterization of the approximate degree (Fact 2.1) provides a function $\psi\colon \{0,1\}^n \to \mathbb{R}$ that obeys (2.2), (2.4), and

$$\sum_{x\in\{0,1\}^n} \psi(x)\mathrm{OR}_n(x) > \frac{1-\varepsilon}{2}\,. \tag{2.5}$$

It remains to verify (2.3):

$$\begin{aligned}
\psi(0,0,\dots,0) &= \sum_{x\in\{0,1\}^n} \psi(x)(1-\mathrm{OR}_n(x)) \\
&= -\sum_{x\in\{0,1\}^n} \psi(x)\mathrm{OR}_n(x) && \text{by (2.4)} \\
&< -\frac{1-\varepsilon}{2} && \text{by (2.5).}
\end{aligned}$$ □

For probability distributions μ and λ on finite sets X and Y, respectively, we let $\mu \times \lambda$ denote the probability distribution on $X \times Y$ given by $(\mu\times\lambda)(x,y) = \mu(x)\lambda(y)$. The *support* of a probability distribution μ is defined to be $\operatorname{supp}\mu = \{x : \mu(x) > 0\}$.

3 Main Result

We are now in a position to prove our main result.

Theorem 3.1. *The Boolean function* $f(x) = \bigwedge_{i=1}^m \bigvee_{j=1}^n x_{ij}$ *obeys*

$$\deg_{1/3}(f) = \Omega(\sqrt{mn})\,. \tag{3.1}$$

Proof. Let ε be an absolute constant to be named later, $0 < \varepsilon < 1$. Then by Theorem 2.3, there exists a constant $\delta = \delta(\varepsilon) > 0$ and a function $\psi\colon \{0,1\}^n \to \mathbb{R}$ that obeys (2.2)–(2.4). Let μ be the probability distribution on $\{0,1\}^n$ given by $\mu(x) = |\psi(x)|$. Let μ_0 and μ_1 be the probability distributions induced by μ on the sets $\{x : \psi(x) < 0\}$ and $\{x : \psi(x) > 0\}$, respectively. Since $\sum_{x\in\{0,1\}^n} \psi(x) = 0$, the sets $\{x : \psi(x) < 0\}$ and $\{x : \psi(x) > 0\}$ are weighted equally by μ. As a consequence,

$$\mu = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_0\,, \tag{3.2}$$

$$\psi = \frac{1}{2}\mu_1 - \frac{1}{2}\mu_0\,. \tag{3.3}$$

Consider the linear operator L that maps functions $\phi\colon (\{0,1\}^n)^m \to \mathbb{R}$ to functions $L\phi\colon \{0,1\}^m \to \mathbb{R}$ according to

$$(L\phi)(z) = \mathop{\mathbf{E}}_{x_1\sim\mu_{z_1}} \cdots \mathop{\mathbf{E}}_{x_m\sim\mu_{z_m}} \phi(x_1,\ldots,x_m)\,.$$

Fix a real polynomial p with

$$\|f - p\|_\infty \le \varepsilon\,. \tag{3.4}$$

Claim 3.2. $\|\mathrm{AND}_m - Lf\|_\infty < \varepsilon$.

Claim 3.3. $\deg p \ge \delta\sqrt{n}\,\deg Lp$.

Before settling the claims, we finish the proof of the theorem. The linearity of L yields

$$\|\mathrm{AND}_m - Lp\|_\infty \le \underbrace{\|\mathrm{AND}_m - Lf\|_\infty}_{<\varepsilon} + \underbrace{\|L(f-p)\|_\infty}_{\le\varepsilon} < 2\varepsilon\,,$$

where we have used (3.4) and Claim 3.2 in bounding the marked quantities. For $\varepsilon = 1/6$, we arrive at $\|\mathrm{AND}_m - Lp\|_\infty \le 1/3$ and therefore $\deg Lp = \Omega(\sqrt{m})$ by Theorem 2.2. Now Claim 3.3 implies that $\deg p = \Omega(\sqrt{mn})$. □

Proof of Claim 3.2. By (2.3), we have $\psi(x) > 0$ only when $\mathrm{OR}_n(x) = 1$. Hence $\operatorname{supp}\mu_1 \subseteq \mathrm{OR}_n^{-1}(1)$ and

$$(Lf)(1,1,\ldots,1) = \mathop{\mathbf{E}}_{\mu_1\times\cdots\times\mu_1}[f] = \prod_{i=1}^{m} \mathop{\mathbf{E}}_{\mu_1}[\mathrm{OR}_n] = 1\,.$$

It remains to prove that $|(Lf)(z)| < \varepsilon$ for every $z \ne (1,1,\ldots,1)$. We have

$$(Lf)(z) = \mathop{\mathbf{E}}_{\mu_{z_1}\times\cdots\times\mu_{z_m}}[f] = \prod_{i=1}^{m} \mathop{\mathbf{E}}_{\mu_{z_i}}[\mathrm{OR}_n] = \prod_{i=1}^{m}(1 - \mu_{z_i}(0,0,\ldots,0))\,,$$

whence

$$0 \le (Lf)(z) \le 1 - \mu_0(0,0,\ldots,0)\,. \tag{3.5}$$

We know from (2.3) that $\psi(0,0,\ldots,0) < -(1-\varepsilon)/2$, which means in particular that $(0,0,\ldots,0) \in \operatorname{supp}\mu_0$. Therefore

$$\mu_0(0,0,\ldots,0) = 2\mu(0,0,\ldots,0) = 2|\psi(0,0,\ldots,0)| > 1-\varepsilon\,,$$

where the first step uses (3.2). By (3.5), we conclude that $0 \le (Lf)(z) < \varepsilon$. □

Proof of Claim 3.3. By the linearity of L, it suffices to consider factored polynomials p of the form $p(x) = \prod_{i=1}^{m} p_i(x_{i,1}, x_{i,2}, \ldots, x_{i,n})$. In this case we have the convenient formula

$$(Lp)(z) = \prod_{i=1}^{m} \mathop{\mathbf{E}}_{\mu_{z_i}}[p_i]\,.$$

By (2.4) and (3.3), polynomials p_i of degree less than $\delta\sqrt{n}$ obey $\mathbf{E}_{\mu_0}[p_i] = \mathbf{E}_{\mu_1}[p_i]$ and therefore do not contribute to the degree of Lp. As a result,

$$\deg Lp \le |\{i : \deg p_i \ge \delta\sqrt{n}\}| \le \frac{\deg p}{\delta\sqrt{n}}\,. \qquad \square$$

Using the pattern matrix method [36], one can immediately translate the main result of this paper into lower bounds on communication complexity. For example, it follows that the two-party communication problem $f(x,y) = \bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} (x_{ij} \wedge y_{ij})$ has bounded-error quantum complexity $\Omega(\sqrt{mn})$, regardless of prior entanglement. We refer the interested reader to [36, 38, 37] for further details and applications.

Acknowledgments

The author is thankful to Mark Bun, Justin Thaler, and the anonymous reviewers for their valuable feedback on an earlier version of this manuscript.

References

[1] SCOTT AARONSON: The polynomial method in quantum and classical computing. In *Proc. 49th FOCS*, p. 3. IEEE Comp. Soc. Press, 2008. [doi:10.1109/FOCS.2008.91] 653, 654

[2] SCOTT AARONSON AND YAOYUN SHI: Quantum lower bounds for the collision and the element distinctness problems. *J. ACM*, 51(4):595–605, 2004. Preliminary version in STOC'02. [doi:10.1145/1008731.1008735] 654

[3] ANDRIS AMBAINIS: Polynomial degree and lower bounds in quantum complexity: Collision and element distinctness with small range. *Theory of Computing*, 1(3):37–46, 2005. [doi:10.4086/toc.2005.v001a003] 654, 655

[4] ANDRIS AMBAINIS, ANDREW M. CHILDS, BEN W. REICHARDT, ROBERT ŠPALEK, AND SHENGYU ZHANG: Any AND-OR formula of size N can be evaluated in time $N^{1/2+o(1)}$ on a quantum computer. *SIAM J. Comput.*, 39(6):2513–2530, 2010. Preliminary version in FOCS'07. [doi:10.1137/080712167] 654

[5] JAMES ASPNES, RICHARD BEIGEL, MERRICK L. FURST, AND STEVEN RUDICH: The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994. Preliminary version in STOC'91. [doi:10.1007/BF01215346] 653, 654

[6] ROBERT BEALS, HARRY BUHRMAN, RICHARD CLEVE, MICHELE MOSCA, AND RONALD DE WOLF: Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001. Preliminary version in FOCS'98. [doi:10.1145/502090.502097] 654

[7] RICHARD BEIGEL: The polynomial method in circuit complexity. In *Proc. 8th IEEE Conf. on Structure in Complexity Theory (SCT'93)*, pp. 82–95. IEEE Comp. Soc. Press, 1993. [doi:10.1109/SCT.1993.336538] 653

[8] RICHARD BEIGEL: Perceptrons, PP, and the polynomial hierarchy. *Comput. Complexity*, 4(4):339–349, 1994. Preliminary version in SCT'92. [doi:10.1007/BF01263422] 654

[9] HARRY BUHRMAN, RICHARD CLEVE, RONALD DE WOLF, AND CHRISTOF ZALKA: Bounds for small-error and zero-error quantum algorithms. In *Proc. 40th FOCS*, pp. 358–368. IEEE Comp. Soc. Press, 1999. [doi:10.1109/SFFCS.1999.814607] 654

[10] HARRY BUHRMAN AND RONALD DE WOLF: Complexity measures and decision tree complexity: A survey. *Theoret. Comput. Sci.*, 288(1):21–43, 2002. [doi:10.1016/S0304-3975(01)00144-X] 653

[11] HARRY BUHRMAN, NIKOLAI K. VERESHCHAGIN, AND RONALD DE WOLF: On computation and communication with small bias. In *Proc. 22nd IEEE Conf. on Computational Complexity (CCC'07)*, pp. 24–32. IEEE Comp. Soc. Press, 2007. [doi:10.1109/CCC.2007.18] 654

[12] HARRY BUHRMAN AND RONALD DE WOLF: Communication complexity lower bounds by polynomials. In *Proc. 16th IEEE Conf. on Computational Complexity (CCC'01)*, pp. 120–130. IEEE Comp. Soc. Press, 2001. [doi:10.1109/CCC.2001.933879] 654

[13] MARK BUN AND JUSTIN THALER: Dual lower bounds for approximate degree and Markov-Bernstein inequalities. In *Proc. 40th Internat. Colloq. on Automata, Languages and Programming (ICALP'13)*. Springer, 2013. To appear. Preprint available at arXiv. 655, 656

[14] DMITRY GAVINSKY AND ALEXANDER A. SHERSTOV: A separation of NP and coNP in multiparty communication complexity. *Theory of Computing*, 6(1):227–245, 2010. [doi:10.4086/toc.2010.v006a010] 655, 656, 657

[15] PETER HØYER, MICHELE MOSCA, AND RONALD DE WOLF: Quantum search on bounded-error inputs. In *Proc. 30th Internat. Colloq. on Automata, Languages and Programming (ICALP'03)*, pp. 291–299. Springer, 2003. [doi:10.1007/3-540-45061-0_25] 654, 655

[16] JEFF KAHN, NATHAN LINIAL, AND ALEX SAMORODNITSKY: Inclusion-exclusion: Exact and approximate. *Combinatorica*, 16(4):465–477, 1996. [doi:10.1007/BF01271266] 654

[17] ADAM TAUMAN KALAI, ADAM R. KLIVANS, YISHAY MANSOUR, AND ROCCO A. SERVEDIO: Agnostically learning halfspaces. *SIAM J. Comput.*, 37(6):1777–1805, 2008. Preliminary version in FOCS'05. [doi:10.1137/060649057] 654

[18] HARTMUT KLAUCK, ROBERT ŠPALEK, AND RONALD DE WOLF: Quantum and classical strong direct product theorems and optimal time-space tradeoffs. *SIAM J. Comput.*, 36(5):1472–1493, 2007. Preliminary version in FOCS'04. [doi:10.1137/05063235X] 654

[19] ADAM R. KLIVANS AND ROCCO A. SERVEDIO: Learning DNF in time $2^{\tilde{O}(n^{1/3})}$. *J. Comput. System Sci.*, 68(2):303–318, 2004. Preliminary version in STOC'01. [doi:10.1016/j.jcss.2003.07.007] 654

[20] MATTHIAS KRAUSE AND PAVEL PUDLÁK: On the computational power of depth-2 circuits with threshold and modulo gates. *Theoret. Comput. Sci.*, 174(1-2):137–156, 1997. Preliminary version in STOC'94. [doi:10.1016/S0304-3975(96)00019-9] 654

[21] MATTHIAS KRAUSE AND PAVEL PUDLÁK: Computing Boolean functions by polynomials and threshold circuits. *Comput. Complexity*, 7(4):346–370, 1998. Preliminary version in FOCS'95. [doi:10.1007/s000370050015] 654

[22] NATHAN LINIAL AND NOAM NISAN: Approximate inclusion-exclusion. *Combinatorica*, 10(4):349–365, 1990. Preliminary version in STOC'90. [doi:10.1007/BF02128670] 654

[23] MARVIN L. MINSKY AND SEYMOUR A. PAPERT: *Perceptrons: An Introduction to Computational Geometry*. MIT Press, Cambridge, Mass., 1969. 653, 654

[24] SABURO MUROGA: *Threshold Logic and Its Applications*. John Wiley & Sons, New York, 1971. 653

[25] NOAM NISAN AND MARIO SZEGEDY: On the degree of Boolean functions as real polynomials. *Comput. Complexity*, 4(4):301–313, 1994. Preliminary version in STOC'92. [doi:10.1007/BF01263419] 653, 654, 655, 656

[26] RYAN O'DONNELL AND ROCCO A. SERVEDIO: New degree bounds for polynomial threshold functions. *Combinatorica*, 30(3):327–358, 2010. Preliminary version in STOC'03. [doi:10.1007/s00493-010-2173-3] 654

[27] RAMAMOHAN PATURI AND MICHAEL E. SAKS: Approximating threshold circuits by rational functions. *Inform. and Comput.*, 112(2):257–272, 1994. Preliminary version in FOCS'90. [doi:10.1006/inco.1994.1059] 654

[28] ALEXANDER A. RAZBOROV: Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145–159, 2003. [doi:10.1070/IM2003v067n01ABEH000422] 654

[29] ALEXANDER A. RAZBOROV AND ALEXANDER A. SHERSTOV: The sign-rank of AC^0. *SIAM J. Comput.*, 39(5):1833–1855, 2010. Preliminary version in FOCS'08. [doi:10.1137/080744037] 654

[30] BEN W. REICHARDT: Reflections for quantum query algorithms. In *Proc. 22nd Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'11)*, pp. 560–569. ACM Press, 2011. SIAM. [ACM:2133080] 654

[31] MICHAEL E. SAKS: Slicing the hypercube. In KEITH WALKER, editor, *Surveys in Combinatorics, 1993*, volume 187 of *London Mathematical Society Lecture Note Series*, pp. 211–256. Cambridge Univ. Press, 1993. [doi:10.1017/CBO9780511662089.009] 653

[32] ALEXANDER A. SHERSTOV: Communication lower bounds using dual polynomials. *Bulletin of the EATCS*, 95:59–93, 2008. EATCS. 653, 654, 656

[33] ALEXANDER A. SHERSTOV: Approximate inclusion-exclusion for arbitrary symmetric functions. *Comput. Complexity*, 18(2):219–247, 2009. Preliminary version in CCC'08. [doi:10.1007/s00037-009-0274-4] 654

[34] ALEXANDER A. SHERSTOV: The intersection of two halfspaces has high threshold degree. In *Proc. 50th FOCS*, pp. 343–362. IEEE Comp. Soc. Press, 2009. [doi:10.1109/FOCS.2009.18] 654, 655, 656

[35] ALEXANDER A. SHERSTOV: Separating AC^0 from depth-2 majority circuits. *SIAM J. Comput.*, 38(6):2113–2129, 2009. Preliminary version in STOC'07. [doi:10.1137/08071421X] 654

[36] ALEXANDER A. SHERSTOV: The pattern matrix method. *SIAM J. Comput.*, 40(6):1969–2000, 2011. Preliminary version in FOCS'08. [doi:10.1137/080733644] 654, 656, 659

[37] ALEXANDER A. SHERSTOV: The multiparty communication complexity of set disjointness. In *Proc. 44th STOC*, pp. 525–548. ACM Press, 2012. [doi:10.1145/2213977.2214026] 659

[38] ALEXANDER A. SHERSTOV: Strong direct product theorems for quantum communication and query complexity. *SIAM J. Comput.*, 41(5):1122–1165, 2012. Preliminary version in STOC'11. [doi:10.1137/110842661] 654, 656, 659

[39] ALEXANDER A. SHERSTOV: Making polynomials robust to noise. *Theory of Computing*, 9(18):593–615, 2013. Preliminary version in STOC'12. [doi:10.4086/toc.2013.v009a018] 654

[40] YAOYUN SHI: Approximating linear restrictions of Boolean functions. Manuscript. Available at author's home page, 2002. 654, 655

[41] KAI-YEUNG SIU, VWANI P. ROYCHOWDHURY, AND THOMAS KAILATH: Rational approximation techniques for analysis of neural networks. *IEEE Trans. Inform. Theory*, 40(2):455–466, 1994. Preliminary version in NIPS'92. [doi:10.1109/18.312168] 654

[42] JUN TARUI AND TATSUIE TSUKIJI: Learning DNF by approximating inclusion-exclusion formulae. In *Proc. 14th IEEE Conf. on Computational Complexity (CCC'99)*, pp. 215–221. IEEE Comp. Soc. Press, 1999. [doi:10.1109/CCC.1999.766279] 654

[43] RONALD DE WOLF: A note on quantum algorithms and the minimal degree of ε-error polynomials for symmetric functions. *Quantum Inf. Comput.*, 8(10):943–950, 2008. QIC. [ACM:2016989] 654

AUTHOR

Alexander A. Sherstov
Assistant Professor
University of California, Los Angeles
sherstov@cs.ucla.edu
http://www.cs.ucla.edu/~sherstov

ABOUT THE AUTHOR

ALEXANDER A. SHERSTOV completed his Ph. D. in 2009 at the University of Texas at Austin. After a postdoc at Microsoft Research, Alexander joined UCLA as an assistant professor of computer science. His research interests include computational complexity theory, computational learning theory, and quantum computing.