Electronic Colloquium on Computational Complexity, Report No. 23 (2013)

APPROXIMATING THE AND-OR TREE

ALEXANDER A. SHERSTOV

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.

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], [11], [32], [1] provide a fairly comprehensive overview of this body of work. Several kinds of representation have been studied [25, 24, 5, 7, 26], 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

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

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 $1/3$ 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 [24], 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 algorithm design and complexity theory. The earliest use of the approximate degree was to prove circuit lower bounds and oracle separations of complexity classes [28, 40, 5, 21, 22, 35]. For 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, 19]. The approximate degree has enabled remarkable progress in communication complexity, with complete resolutions of problems that were once considered hopelessly hard, e.g., [10, 29, 12, 36, 30, 32]. 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 was used to obtain the fastest known algorithms for PAC-learning DNF formulas [41, 20] and read-once formulas [4] and the fastest known algorithm for agnostically learning disjunctions [18]. Another well-known

* Computer Science Department, UCLA, Los Angeles, California 90095. ✉ sherstov@cs.ucla.edu
Supported by NSF CAREER award CCF-1149018.

ISSN 1433-8092

Bound	**Reference**
$O(n)$	Høyer, Mosca, and de Wolf [16]
$\Omega(\sqrt{n})$	Nisan and Szegedy [26]
$\Omega(\sqrt{n \log n})$	Shi [39]
$\Omega(n^{0.66\dots})$	Ambainis [3]
$\Omega(n^{0.75})$	Sherstov [34]
$\Omega(n)$	This paper

TABLE 1. Approximate degree of the AND-OR tree.

use of the approximate degree is an algorithm for approximating the inclusion-exclusion formula based on its initial terms [23, 17, 33, 14].

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, 26, 39, 2, 3, 33, 38]. In addition, quantum query algorithms have been used to prove *upper* bounds on the approximate degree [16, 14, 4], and duality-based methods have yielded *lower* bounds [27, 34, 37]. Nevertheless, our understanding of this complexity measure remains fragmented, with few general results available [26, 38].

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 [26, 39, 16, 3, 34] and was 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 (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; see Remark 3.3.

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 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) $Lf \approx \mathrm{AND}_m$ 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 [26] 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 $\mathrm{OR}_n^{-1}(0)$ 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 [15], where it was used to obtain lower bounds for nondeterministic and Merlin-Arthur communication protocols.

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 based on the same idea—discovered and used for the first time in [15]—that OR_n has a one-sided dual polynomial. The two papers differ in how they use this idea 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 $\|f - p\|_\infty \leqslant \varepsilon$. 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:

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

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, 37] of the approximate degree is as follows.

FACT 2.1. *Let $f : X \to \mathbb{R}$ be given, for a finite set $X \subset \mathbb{R}^n$. Then* $\deg_\varepsilon(f) \geqslant d$ *if and only if there exists a function $\psi : X \to \mathbb{R}$ 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 $\mathrm{AND}_n, \mathrm{OR}_n : \{0,1\}^n \to \{0,1\}$. Their approximate degree was determined by Nisan and Szegedy [26].

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 [15, 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 $\psi : \{0,1\}^n \to \mathbb{R}$ such that*

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

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

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 [15]). Recall from Theorem 2.2 that $\deg_{1/3}(\mathrm{OR}_n) = \Omega(\sqrt{n})$. Thus, (2.1) shows that $\deg_{\frac{1-\varepsilon}{2}}(\mathrm{OR}_n) \geqslant \delta\sqrt{n}$ for a sufficiently small constant $\delta = \delta(\varepsilon) > 0$. Now the dual characterization of the approximate degree (Fact 2.1) provides a function $\psi : \{0,1\}^n \to \mathbb{R}$ 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):

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

∎

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. *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

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

Fix a real polynomial p with

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

CLAIM 3.1. $\|\mathrm{AND}_m - Lf\|_\infty < \varepsilon.$

CLAIM 3.2. $\deg p \geqslant \delta\sqrt{n}\deg Lp.$

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

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

where we have used (3.4) and Claim 3.1 in bounding the marked quantities. For $\varepsilon = 1/6$, we arrive at $\|\mathrm{AND}_m - Lp\|_\infty \leqslant 1/3$ and therefore $\deg Lp = \Omega(\sqrt{m})$ by Theorem 2.2. Now Claim 3.2 implies that $\deg p = \Omega(\sqrt{mn})$. ∎

Proof of Claim 3.1. 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

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

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

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

whence

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

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

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

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

Proof of Claim 3.2. By the linearity of L, it suffices to consider factored polynomials p of the form $p(x) = \prod_{i=1}^{m} p_i(x_{i,1}, x_{i,2}, \dots, x_{i,n})$. 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,

$$\deg Lp \leqslant |\{i : \deg p_i \geqslant \delta\sqrt{n}\}| \leqslant \frac{\deg p}{\delta\sqrt{n}}.$$ ∎

REMARK 3.3. It is shown in [38] that for any Boolean functions f and g, the composition $f(g, g, \dots, g)$ has approximate degree $O(\deg_{1/3}(f)\deg_{1/3}(g))$. Since AND_m and OR_n have approximate degree $\Theta(\sqrt{m})$ and $\Theta(\sqrt{n})$, respectively, the lower bound (3.1) is tight for all m and n.

ACKNOWLEDGMENTS

The author is thankful to Mark Bun and Justin Thaler for their useful feedback on an earlier version of this manuscript.

REFERENCES

[1] S. Aaronson. The polynomial method in quantum and classical computing. In *Proceedings of the Forty-Ninth Annual IEEE Symposium on Foundations of Computer Science* (FOCS), page 3, 2008.

[2] S. Aaronson and Y. Shi. Quantum lower bounds for the collision and the element distinctness problems. *J. ACM*, 51(4):595–605, 2004.

[3] A. Ambainis. Polynomial degree and lower bounds in quantum complexity: Collision and element distinctness with small range. *Theory of Computing*, 1(1):37–46, 2005.

[4] A. Ambainis, A. M. Childs, B. Reichardt, R. Špalek, and S. Zhang. Any AND-OR formula of size N can be evaluated in time $N^{1/2+o(1)}$ on a quantum computer. In *Proceedings of the Forty-Eighth Annual IEEE Symposium on Foundations of Computer Science* (FOCS), pages 363–372, 2007.

[5] J. Aspnes, R. Beigel, M. L. Furst, and S. Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994.

[6] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001.

[7] R. Beigel. The polynomial method in circuit complexity. In *Proc. of the Eigth Annual Conference on Structure in Complexity Theory*, pages 82–95, 1993.

[8] R. Beigel. Perceptrons, PP, and the polynomial hierarchy. *Computational Complexity*, 4:339–349, 1994.

[9] H. Buhrman, R. Cleve, R. de Wolf, and C. Zalka. Bounds for small-error and zero-error quantum algorithms. In *Proceedings of the Fortieth Annual IEEE Symposium on Foundations of Computer Science* (FOCS), pages 358–368, 1999.

[10] H. Buhrman and R. de Wolf. Communication complexity lower bounds by polynomials. In *Proceedings of the Sixteenth Annual IEEE Conference on Computational Complexity* (CCC), pages 120–130, 2001.

[11] H. Buhrman and R. de Wolf. Complexity measures and decision tree complexity: A survey. *Theor. Comput. Sci.*, 288(1):21–43, 2002.

[12] H. Buhrman, N. K. Vereshchagin, and R. de Wolf. On computation and communication with small bias. In *Proceedings of the Twenty-Second Annual IEEE Conference on Computational Complexity* (CCC), pages 24–32, 2007.

[13] M. Bun and J. Thaler. Dual lower bounds for approximate degree and Markov-Bernstein inequalities. Manuscript, 2013.

[14] R. de Wolf. A note on quantum algorithms and the minimal degree of ε-error polynomials for symmetric functions. *Quantum Information and Computation*, 8(10):943–950, 2008.

[15] D. Gavinsky and A. A. Sherstov. A separation of NP and coNP in multiparty communication complexity. *Theory of Computing*, 6(10):227–245, 2010.

[16] P. Høyer, M. Mosca, and R. de Wolf. Quantum search on bounded-error inputs. In *Proc. of the 30th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 291–299, 2003.

[17] J. Kahn, N. Linial, and A. Samorodnitsky. Inclusion-exclusion: Exact and approximate. *Combinatorica*, 16(4):465–477, 1996.

[18] A. T. Kalai, A. R. Klivans, Y. Mansour, and R. A. Servedio. Agnostically learning halfspaces. *SIAM J. Comput.*, 37(6):1777–1805, 2008.

[19] H. Klauck, R. Špalek, and R. de Wolf. Quantum and classical strong direct product theorems and optimal time-space tradeoffs. *SIAM J. Comput.*, 36(5):1472–1493, 2007.

[20] A. R. Klivans and R. A. Servedio. Learning DNF in time $2^{\tilde{O}(n^{1/3})}$. *J. Comput. Syst. Sci.*, 68(2):303–318, 2004.

[21] M. Krause and P. Pudlák. On the computational power of depth-2 circuits with threshold and modulo gates. *Theor. Comput. Sci.*, 174(1–2):137–156, 1997.

[22] M. Krause and P. Pudlák. Computing Boolean functions by polynomials and threshold circuits. *Comput. Complex.*, 7(4):346–370, 1998.

[23] N. Linial and N. Nisan. Approximate inclusion-exclusion. *Combinatorica*, 10(4):349–365, 1990.

[24] M. L. Minsky and S. A. Papert. *Perceptrons: An Introduction to Computational Geometry*. MIT Press, Cambridge, Mass., 1969.

[25] S. Muroga. *Threshold Logic and Its Applications*. John Wiley & Sons, New York, 1971.

[26] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4:301–313, 1994.

[27] R. O'Donnell and R. A. Servedio. New degree bounds for polynomial threshold functions. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing* (STOC), pages 325–334, 2003.

[28] R. Paturi and M. E. Saks. Approximating threshold circuits by rational functions. *Inf. Comput.*, 112(2):257–272, 1994.

[29] A. A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya of the Russian Academy of Sciences, Mathematics*, 67:145–159, 2002.

[30] A. A. Razborov and A. A. Sherstov. The sign-rank of AC^0. *SIAM J. Comput.*, 39(5):1833–1855, 2010. Preliminary version in *Proceedings of the Forty-Ninth Annual IEEE Symposium on Foundations of Computer Science* (FOCS), 2008.

[31] M. E. Saks. Slicing the hypercube. *Surveys in Combinatorics*, pages 211–255, 1993.

[32] A. A. Sherstov. Communication lower bounds using dual polynomials. *Bulletin of the EATCS*, 95:59–93, 2008.

[33] A. A. Sherstov. Approximate inclusion-exclusion for arbitrary symmetric functions. *Computational Complexity*, 18(2):219–247, 2009. Preliminary version in *Proceedings of the Twenty-Third Annual IEEE Conference on Computational Complexity* (CCC), 2008.

[34] A. A. Sherstov. The intersection of two halfspaces has high threshold degree. In *Proceedings of the Fiftieth Annual IEEE Symposium on Foundations of Computer Science* (FOCS), pages 343–362, 2009.

[35] A. A. Sherstov. Separating AC^0 from depth-2 majority circuits. *SIAM J. Comput.*, 38(6):2113–2129, 2009. Preliminary version in *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing* (STOC), 2007.

[36] A. A. Sherstov. The pattern matrix method. *SIAM J. Comput.*, 40(6):1969–2000, 2011. Preliminary version in *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing* (STOC), 2008.

[37] A. A. Sherstov. Strong direct product theorems for quantum communication and query complexity. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing* (STOC), pages 41–50, 2011. Full version available as ECCC Report TR11-040.

[38] A. A. Sherstov. Making polynomials robust to noise. In *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing* (STOC), pages 747–757, 2012.

[39] Y. Shi. Approximating linear restrictions of Boolean functions. Manuscript, 2002.

[40] K.-Y. Siu, V. P. Roychowdhury, and T. Kailath. Rational approximation techniques for analysis of neural networks. *IEEE Transactions on Information Theory*, 40(2):455–466, 1994.

[41] J. Tarui and T. Tsukiji. Learning DNF by approximating inclusion-exclusion formulae. In *Proceedings of the Fourteenth Annual IEEE Conference on Computational Complexity* (CCC), pages 215–221, 1999.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de