Electronic Colloquium on Computational Complexity, Report No. 134 (2012)

Real Advantage

Alexander Razborov* Emanuele Viola†

October 3, 2012

Abstract

We highlight the challenge of proving correlation bounds between boolean functions and integer-valued polynomials, where any non-boolean output counts against correlation.

We prove that integer-valued polynomials of degree $\frac{1}{2}\lg_2 \lg_2 n$ have zero correlation with parity. Such a result is false for modular and threshold polynomials. Its proof is based on a strengthening of an anti-concentration result by Costello, Tao, and Vu (Duke Math. J. 2006).

1 Introduction

The polynomial method has been one of the most successful tools in theoretical computer science. It has had many applications, for example, in complexity and learning theory. The surveys [Bei93, Vio09, She08] provide starting points to learn more about this method.

As is well-known, however, there are several problems about polynomials which have resisted decades of attacks. The purpose of this work is to highlight certain new basic problems about polynomials which appear to stand in the way of further progress, making some initial technical contributions along the way.

A challenge that we are interested in is that of proving correlation bounds. Specifically, for two functions $f, g : \{0,1\}^n \to R$, where R represents the set of real numbers, define their "correlation" to be the quantity:

$$\mathrm{Cor}_n(f,g) \stackrel{\text{def}}{=} \mathbf{P}_x[f(x) = g(x)] - 1/2,$$

where x is uniformly distributed over $\{0,1\}^n$.

Most of the research has so far concentrated on the case in which both f and g are boolean, and in order to incorporate into this framework arbitrary multivariate polynomials,

*University of Chicago, razborov@cs.uchicago.edu. Part of this work was done while the author was at Steklov Mathematical Institute, supported by the Russian Foundation for Basic Research, and at Toyota Technological Institute, Chicago

†Supported by NSF grant CCF-0845003. Email: viola@ccs.neu.edu

ISSN 1433-8092

one clearly has to convert them to boolean functions $b : \{0,1\}^n \to \{0,1\}$. Historically, there have been two prominent methods of doing this. The first applies to polynomials with integer coefficients and consists in declaring $b(x) = 1$ if $p(x)$ is divisible by a prescribed integer m, and 0 otherwise. We refer to these functions b as *modular polynomials*. The second declares $b(x) = 1$ if $p(x) > t$ for some prescribed threshold t, and 0 otherwise. (Here the involved polynomials may be also assumed to have integer coefficients without loss of generality [MTT61, Mur71].) We refer to these functions b as *threshold polynomials*.

It is an open problem to exhibit an explicit boolean function $f : \{0,1\}^n \to \{0,1\}$ such that $\mathrm{Cor}_n(b, f) = o(1/\sqrt{n})$ for any modular polynomial b whose underlying polynomial has degree $\lg_2 n$, cf. [Vio09]. Interestingly, the same problem is open even for threshold polynomials.

These parallel developments make it very natural (and hopefully instructive) to consider the "universal", "umbrella" model of *real polynomials* where we just view the polynomial itself as computing a boolean function, and any output of the polynomial outside of $\{0,1\}$ is counted as an error. As it appears, this generic setting was considered before, but in a somewhat ad hoc way and mostly as a building block for other constructions (see e.g. [ABFR94, Section 5]). We are not aware of any previous attempts to study it independently, and, as a consequence, we are not aware of any lower bound techniques for real polynomials that are not derivative from modular or threshold case. In particular we highlight the following challenge:

Challenge 1.1. Exhibit an explicit boolean function $f : \{0,1\}^n \to \{0,1\}$ such that $\mathrm{Cor}_n(p, f) = o(1/\sqrt{n})$ for any real polynomial $p : \{0,1\}^n \to R$ of degree $\lg_2 n$.

Another motivation for considering real correlation in its unprocessed form comes from connections to matrix analysis. It is well-known that analogies between polynomial approximations and matrix approximations are extremely important and influential in Theory and other areas like Machine Learning (see e.g. [She08]). Viewed under this angle, our model is a straightforward analogy of matrix rigidity [Val77] that still remains one of the greatest unresolved mysteries in the modern Complexity Theory.

Note that the Parity function is a candidate for f in Challenge 1.1. We also note that solving Challenge 1.1 is a pre-requisite for solving the corresponding problem for threshold polynomials. Similarly, the special case of Challenge 1.1 when the polynomials have integer coefficients is a pre-requisite for solving the corresponding problem for modular polynomials.

We do not know how to address Challenge 1.1. However, we can prove that the correlation with parity is zero for low degrees.

Theorem 1.2. There exists an $\epsilon > 0$ such that $\mathrm{Cor}_n(p, \mathrm{parity}) = 0$ for every n and every real polynomial $p : \{0,1\}^n \to R$ of degree $\leq \frac{1}{2} \lg_2 \lg_2 n$.

Theorem 1.2 follows easily from an anti-concentration result which is presented in §3 and is our main technical contribution. The proof of Theorem 1.2 is at the end of §3.

We note that the "zero behavior" in Theorem 1.2 does not hold for the other models mentioned earlier. Specifically, for threshold polynomials it follows from the results [ABFR94] of Aspnes, Beigel, Furst, and Rudich that increasing degree always increases correlation

with parity. For modular polynomials when the modulus is $m = 3$, it follows from Green's result [Gre04] that quadratic polynomials have better correlation with parity than linear polynomials. Thus, Theorem 1.2 appears to be the first authentic lower bound in our model.

Theorem 1.2 also raises the problem of determining the smallest degree for which the correlation with parity becomes strictly positive. Using standard techniques we note an $O(\sqrt{n})$ upper bound on this degree. In fact, the bound applies to any symmetric function, i.e., a function whose value only depends on the Hamming weight of the input. The same result obviously holds for threshold polynomials. The proof does not immediately apply to polynomials modulo m, because the coefficients may have denominators divisible by m. But a similar result for modular polynomials is obtained in Theorem 7 in [Vio09].

Fact 1.3. Let $f : \{0,1\}^n \to R$ be a real-valued, symmetric function. There exists a real polynomial $p : \{0,1\}^n \to R$ of degree $O(\sqrt{n})$ such that $\mathrm{Cor}_n(p, f) \geq 0.99$.

Proof. We construct a polynomial that computes f exactly on inputs of Hamming weight $n/2 - c\sqrt{n}, \ldots, n/2 + c\sqrt{n}$. For a suitable constant c, this yields the result by, say, Chebyshev's bound.

Denote by $h(w)$ the value of f on inputs of Hamming weight w. By interpolation, we can compute h on inputs in the above range by a univariate polynomial $p(w)$ of degree $2c\sqrt{n}$. The desired multivariate polynomial is then $p(\sum_i x_i)$.■

2 Notation

We let $[n] \stackrel{\text{def}}{=} \{1, 2, \ldots, n\}$, and $[n]^{\leq k} \stackrel{\text{def}}{=} \{\, S \in [n] \mid |S| \leq k \,\}$.

In this paper, we are interested in real-valued functions $f : \{0,1\}^n \longrightarrow R$. Every such function has a unique representation as a *multi-linear* polynomial $p(x_1, \ldots, x_n) = \sum_{S \subseteq [n]} c_S x_S$, where $c_S \in R$ and we have introduced the natural abbreviation $x_S \stackrel{\text{def}}{=} \prod_{i \in S} x_i$. $\mathrm{Supp}(p) \stackrel{\text{def}}{=} \{\, S \subseteq [n] \mid c_S \neq 0 \,\}$ corresponds to the set of non-zero coefficients.

Unless specifically noted otherwise, all probabilities and expectations in this paper are calculated w.r.t. the uniform distribution on $\{0,1\}^n$. For a real vector $\vec{c} = (c_1, \ldots, c_n)$, $\ell_1(\vec{c}) \stackrel{\text{def}}{=} \sum_i |c_i|$ is its ℓ_1-norm.

3 An anti-concentration result

In [CTV06], Costello, Tao and Vu proved the following remarkable result.

Proposition 3.1. Let k be a fixed positive integer, and $p(x_1, \ldots, x_n) = \sum_{S \in [n]^{\leq k}} c_S x_S$ be a multi-linear polynomial of degree $\leq k$. Then for any real interval I of length 1 we have $\mathbf{P}[p \in I] \leq O(m^{-a_k})$, where

$$m \stackrel{\text{def}}{=} \frac{1}{n^{k-1}} \left| \left\{\, S \in [n]^{\leq k} \mid |c_S| \geq 1 \,\right\} \right| \tag{1}$$

and

$$a_k \stackrel{\text{def}}{=} 2^{-(k^2+k)/2}. \tag{2}$$

In this paper, we are mostly interested in the following corollary easily obtainable by an indefinite scaling of the original polynomial.

Corollary 3.2. Under the same assumptions as in Proposition 3.1, for every fixed $b \in R$ we have the bound

$$\mathbf{P}[p = b] \leq O(s^{-a_k}),$$

where a_k is again given by (2) and

$$s \stackrel{\text{def}}{=} \frac{1}{n^{k-1}}|\text{Supp}(p)|.$$

Unfortunately, however, the size of the support does not behave well with respect to restrictions and thus can be hardly used for our purposes.

We improve Proposition 3.1 by relaxing the density restriction to something more combinatorial. Let us call the term x_S *significant* if $|S| = k$ (i.e., it has the highest possible degree) and $|c_S| \geq 1$. Choose a maximal possible set $\{x_{S_1}, \ldots, x_{S_r}\}$ of significant terms with *mutually disjoint sets of variables*. Then every term x_S with $|c_S| \geq 1$ either has degree $< k$ or contains at least one variable in common with one of the terms $x_{S_1}, \ldots, x_{S_r}$ which implies that their overall number is $O(rn^{k-1})$. Thus, $r \geq \Omega(m)$ (as given by (1)) and we are going to prove that the existence of a large set of mutually disjoint significant terms is the *only* property necessarily to derive the conclusion of Proposition 3.1. As a small by-product, we slightly improve their numerical bound.

Theorem 3.3. Let $p(x_1, \ldots, x_n) = \sum_{S \in [n]^{\leq k}} c_S x_S$ be a multi-linear polynomial of degree k, and assume that there exist r terms $x_{S_1}, \ldots, x_{S_r}$ of degree k each and with mutually disjoint sets of variables such that $|c_{S_i}| \geq 1$ $(1 \leq i \leq r)$. Then for any real interval I of length 1 we have $\mathbf{P}[p \in I] \leq O(r^{-b_k})$, where

$$b_k \stackrel{\text{def}}{=} (2k2^k)^{-1}. \tag{3}$$

Proof. We first condition (in the least advantageous way) on the values of all variables not appearing in the terms $x_{S_1}, \ldots, x_{S_r}$. Since these terms have maximal possible degree k, substitutions of other variables do not produce any effect on the coefficients $c_{S_1}, \ldots, c_{S_r}$. Thus, we may assume from the very beginning that our polynomial has precisely kr variables indexed by $S_1 \dot\cup \ldots \dot\cup S_r$. Renaming the variables, let, say

$$x_{S_j} = x_{1j}x_{2j}\ldots x_{kj},$$

where all variables x_{ij} $(i \in [k], j \in [r])$ are pairwise distinct.

Let us call a term t *cross-term* if it has degree k and, moreover, is of the form $x_{1j_1}\ldots x_{kj_k}$ for some function $j : [k] \longrightarrow [r]$. Cross-terms x_{S_j} themselves (i.e., those for which the function j is a constant) will be called *principal*. Our first task is to get rid of all terms

that are not cross-terms (and this is where we suffer the enormous loss in the exponent in (3)); the proof is virtually identical to [CTV06] but we include it nonetheless for the sake of completeness.

Write $p = p(X_1, \ldots, X_k)$, where

$$X_i \stackrel{\text{def}}{=} \{\, x_{ij} \mid j \in [r] \,\},$$

introduce an isomorphic set of variables $Y_1, \ldots, Y_k$ and form the alternating sum

$$\widehat{p}(X_1, \ldots, X_k, Y_1, \ldots, Y_k) \stackrel{\text{def}}{=} \sum_{a \in \{0,1\}^k} (-1)^{\sum_i a_i} p(W_1^{a_1}, \ldots, W_k^{a_k}), \tag{4}$$

where $W_i^a \stackrel{\text{def}}{=} \begin{cases} X_i & \text{if } a = 0 \\ Y_i & \text{if } a = 1. \end{cases}$

The Decoupling lemma [CTV06, Lemma 6.3] implies that

$$\mathbf{P}[p(X_1, \ldots, X_k) \in I] \leq \mathbf{P}\left[\bigwedge_{a \in \{0,1\}^k} p(W_1^{a_1}, \ldots, W_k^{a_k}) \in I\right]^{1/2^k} \leq \mathbf{P}[\widehat{p}(X_1, \ldots, X_k, Y_1, \ldots, Y_k) \in J]^{1/2^k}, \tag{5}$$

where $J = [-2^k, 2^k]$ and we assume that $Y_1, \ldots, Y_k$ are picked from $\{0,1\}^r$ uniformly and independently of $X_1, \ldots, X_k$.

On the other hand, (4) defines a linear mapping from the space of multi-linear polynomials in (kr) variables, and on every term $t = t_1(X_1) \ldots t_k(X_k)$ it produces $(t_1(X_1) - t_1(Y_1))(t_2(X_2) - t_2(Y_2)) \ldots (t_k(X_k) - t_x(Y_k))$. Which implies that this mapping vanishes on all terms that are not cross-terms (as for them there exists an i such that $t_i = 1$), and on every cross-term the mapping acts simply as the substitution $x_{ij} \mapsto z_{ij} \stackrel{\text{def}}{=} x_{ij} - y_{ij}$. Thus,

$$\widehat{p}(X_1, \ldots, X_k, Y_1, \ldots, Y_k) = q(Z_1, \ldots, Z_k),$$

where q is the polynomial p from which we have erased all terms that are not cross-terms. Random variables z_{ij} are i.i.d, and for every individual variable z_{ij}, $\mathbf{P}[z_{ij} = 1] = \mathbf{P}[z_{ij} = -1] = 1/4$, and $\mathbf{P}[z_{ij} = 0] = 1/2$.

Definition 3.4. A polynomial $q(Z_1, \ldots, Z_k)$ of degree k (where $Z_j = \{z_{j1}, \ldots, z_{jr}\}$) is a *$(k, r)$-polynomial* if it is *multi-linear in $Z_1, \ldots, Z_k$*[1], and, moreover, the coefficient in front of all r principal terms is at least one in the absolute value.

Note that q is a (k, r)-polynomial.

Our main technical contribution is the following anti-concentration inequality:

[1] meaning that every monomial appearing in q may contain at most one variable from each of these groups

Claim 3.5. For any (k,r)-polynomial q and any interval I of length 1,

$$\mathbf{P}[q(Z_1,\ldots,Z_k)\in I]\le C_1 3^k r^{-1/(2k)},$$

where $C_1>0$ is an absolute constant.

To complete the proof of Theorem 3.3, employ (5), divide the interval J in $O(2^k)$ intervals of length 1, and apply Claim 3.5 to our alternating sum $\widehat{p}(X_1,\ldots,X_k,Y_1,\ldots,Y_k)=q(Z_1,\ldots,Z_k)$.■

Proof of Claim 3.5.

The proof is by induction on k.

Base case $k=1$ immediately follows from Kolmogorov-Rogozin inequality [Rog61, Theorem 1] generalizing Littlewood-Offord lemma [LO43, Erd45] to a wider class of distributions. We state it here at the level of generality appropriate to our purposes.

Proposition 3.6 ([Rog61])**.** Let z be a discrete random variable that takes on at least two different values. Then for any real coefficient vector $\vec{c}=(c_1,\ldots,c_n)$ and any interval I of length 1,

$$\mathbf{P}[c_1z_1+\cdots+c_nz_n\in I]\le C_0\cdot\ell_1(\vec{c})^{-1/2},$$

where $z_1,\ldots,z_n$ are i.i.d copies of z, and the constant C_0 depends only on the distribution of z.

In particular, $\mathbf{P}[q(Z_1)\in I]\le C_0r^{-1/2}$ which is what we need, as long as $C_1\ge C_0$.

Inductive step. Assume now that $k\ge 2$, and that the Claim is already established for $(k-1,r')$-polynomials, for arbitrary r'. Expand our (k,r)-polynomial q as

$$q(Z_1,\ldots,Z_k)=q_0(Z_1,\ldots,Z_k)+\sum_{j=1}^{r}z_{1j}z_{2j}\ldots z_{(k-1)j}R_j(Z_k),$$

where terms in $q_0(Z_1,\ldots,Z_k)$ are not divisible by any of $z_{1j}z_{2j}\ldots z_{(k-1)j}$, and

$$R_j(Z_k)=c_{j1}z_{k1}+\cdots+c_{jr}z_{kr}+c_j$$

is an affine form in the variables Z_k such that $|c_{jj}|\ge 1$.

What we need to prove for the inductive step is that after we assign $Z_k=\{z_{k1},\ldots,z_{kr}\}$ at random (according to their distribution), then with high probability $|R_j(Z_k)|$ will be large for many j. Indeed, note that after we assign A_k to Z_k we obtain a multi-variate polynomial of degree $k-1$, and the coefficient of the j-th principal term is $|R_j(A_k)|$, because terms in q_0 cannot contribute to principal terms. Assigning further (in an arbitrary way) all those variables z_{ij} $(i\in[k-1])$ for which $|R_j(A_k)|$ is small, we arrive at $(k-1,r')$-polynomials, and we can apply the inductive assumption.

Set

$$s\stackrel{\text{def}}{=}\frac{1}{500}r^{1/k},$$

and let us call the form R_j *large* if $\sum_{j'=1}^{r} |c_{jj'}| \geq s$ and *small* otherwise.

Case 1. *At least $r/2$ forms (say, $R_1, \ldots, R_{r/2}$) are large.* This case is analyzed similarly to [CTV06]. For every large form R_j, Proposition 3.6 implies that $\mathbf{P}[|R_j(Z_k)| < 1] \leq 2C_0 s^{-1/2}$. Picking at random $j \in_U [r/2]$, and applying Markov's inequality, we obtain

$$\mathbf{P}_{Z_k}\left[\mathbf{P}_{j\in_U[r/2]}[|R_j(Z_k)| < 1] \geq 1/2\right] \leq 4C_0 s^{-1/2}.$$

For every individual assignment A_k to the variables Z_k such that $\mathbf{P}_{j\in_U[r/2]}[|R_j(A_k)| < 1] < 1/2$ (followed by additionally assigning the variables z_{ij} with $i \in [k-1]$ and $|R_j(A_k)| < 1$) we are left with a $(k-1, r')$-polynomial for $r' \geq r/4$. Hence we have, by the inductive assumption:

$$\mathbf{P}[q(Z_1, \ldots, Z_k) \in I] \leq 4C_0 s^{-1/2} + C_1 3^{k-1}(r/4)^{-1/(2k-2)} \leq C_1 3^k (r/4)^{-1/(2k)}$$

as long as the constant C_1 is large enough relative to C_0.

Case 2. *At least $r/2$ forms (again, $R_1, \ldots, R_{r/2}$) are small.* Arrange coefficients $c_{jj'}$ ($j, j' \in [r/2]$) in the form of a square matrix C: every diagonal element of this matrix satisfies $|c_{jj}| \geq 1$, and the ℓ_1-norm of off-diagonal elements $\sum_{j\neq j'\in[r/2]} |c_{jj'}|$ is bounded by $\sum_j \left(\sum_{j'=1}^{r} |c_{jj'}|\right) \leq \frac{rs}{2}$.

Pick now at random a principal submatrix C^* of size $\frac{r}{10s} \times \frac{r}{10s}$ corresponding to $J \in [r/2]^{\frac{r}{10s}}$. (That is, keep the rows and columns in J.) Every individual off-diagonal element $c_{jj'}$ remains in this matrix with probability $\leq \frac{1}{25s^2}$, therefore the expectation of $\sum_{j\neq j'\in J} |c_{jj'}|$ does not exceed $\frac{r}{50s}$, and we choose any particular J for which this is the case. Applying Markov's inequality again, we find that

$$\left|\left\{ j \in J \;\middle|\; \sum_{\substack{j'\in J\\ j'\neq j}} |c_{jj'}| \geq 1/4 \right\}\right| \leq \frac{2r}{25s},$$

and hence there exist $\geq \frac{r}{10s} - \frac{2r}{25s} = \frac{r}{50s}$ indices $J' \subseteq J$ such that

$$\sum_{\substack{j'\in J\\ j'\neq j}} |c_{jj'}| \leq 1/4 \tag{6}$$

for any $j \in J'$.

Now we first set all variables z_{kj} with $j \notin J'$ at random. This converts affine forms R_j ($j \in J'$) to the form

$$c_{jj} z_{kj} + \sum_{\substack{j'\in J'\\ j'\neq j}} c_{jj'} z_{kj'} + x_j,$$

where x_j are certain fixed numbers. Even if we do not have any way to control these x_j, from (6) we still know that $|R_j(z_k)| \leq 1/2$ implies that the term $c_{jj} z_{kj}$ belongs to a fixed

interval of length $\leq 1 + \frac{1}{4} + \frac{1}{4} < 2$, and thus there exists a fixed value $a_j \in \{\pm 1\}$ such that $z_{kj} = a_j$ implies $|R_j(Z_k)| \geq 1/2$. Applying Chernoff's bound, we conclude that

$$\mathbf{P}_{Z_k}\left[|\{j \in J' \mid \ |R_j(Z_k)| \geq 1/2\}| \leq \frac{r}{500s}\right] \leq 2^{-\epsilon r/s}$$

for a fixed constant $\epsilon > 0$, and in the good (i.e., when $|\{j \in J' \,|\, |R_j(Z_k)| \geq 1/2\}| \geq \frac{r}{500s}$) case we, as before, are left with an $(k-1, \frac{r}{500s})$-polynomial, scaled by a factor of 2. Noting that $\frac{r}{500s} = r^{1-1/k}$, by inductive assumption we get

$$\mathbf{P}[q(Z_1, \ldots, Z_k) \in I] \leq 2^{-\epsilon r/s} + 2C_1 3^{k-1} \left(\frac{r}{500s}\right)^{-1/(2k-2)} \leq C_1 3^k r^{-1/(2k)},$$

again, provided the constant C_1 is large enough. (The factor 2 in front of C_1 is the scaling factor.)

This completes the proof of Claim 3.5.■

Proof. [Of Theorem 1.2] Suppose the hypothesis of Theorem 3.3 is satisfied with $r = \sqrt{n}$. Then the probability that the polynomial outputs a boolean value is

$$O\left(1/\sqrt{n}^{1/(2k2^k)}\right) \leq 1/2$$

when $k \leq \frac{1}{2}\lg_2 \lg_2 n$. This proves the theorem.

Otherwise, we can cover all the degree-k terms by $\leq \sqrt{n}k$ variables. Fix these variables in the most disadvantageous way (this reduces degree of the polynomial), and iterate.

After $\leq k$ iterations either the hypothesis of Theorem 3.3 is satisfied with $r = \sqrt{n}$, in which case we reason as above, or else we end up with a polynomial of degree 0 with $n - \sqrt{n}k^2 \geq 1$ variables, in which case the theorem is again true.■

References

[ABFR94] James Aspnes, Richard Beigel, Merrick Furst, and Steven Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994.

[Bei93] Richard Beigel. The polynomial method in circuit complexity. In *8th Structure in Complexity Theory Conference*, pages 82–95. IEEE, 1993.

[CTV06] Kevin P. Costello, Terence Tao, and Van Vu. Random symmetric matrices are almost surely nonsingular. *Duke Math. J.*, 135(2):395–413, 2006.

[Erd45] Paul Erdős. On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc.*, 51:898–902, 1945.

[Gre04] Frederic Green. The correlation between parity and quadratic polynomials mod 3. *J. of Computer and System Sciences*, 69(1):28–44, 2004.

[LO43] John Littlewood and Albert Offord. On the number of real roots of a random algebraic equation. *III. Rec. Math. [Mat. Sbornik] N.S.*, 12:277–286, 1943.

[MTT61] Saburo Muroga, Iwao Toda, and Satoru Takasu. Theory of majority decision elements. *J. Franklin Inst.*, 271:376–418, 1961.

[Mur71] Saburo Muroga. *Threshold logic and its applications.* Wiley-Interscience, New York, 1971.

[Rog61] B. A. Rogozin. An estimate for concentration functions. *Theory of Probability and its Applications*, 6:94–97, 1961.

[She08] Alexander A. Sherstov. Communication lower bounds using dual polynomials. *Bulletin of the EATCS*, 95:59–93, 2008.

[Val77] L. G. Valiant. Graph-theoretic arguments in low-level complexity. In *Proceedings of the 6th MFCS, Lecture Notes in Computer Science,* 53, pages 162–176, New York/Berlin, 1977. Springer-Verlag.

[Vio09] Emanuele Viola. On the power of small-depth computation. *Foundations and Trends in Theoretical Computer Science*, 5(1):1–72, 2009.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de