ICCM 2007 · Vol. II · 1–4

# Approximate polynomial degree of Boolean functions and its applications

Yaoyun Shi*

**Abstract**

The approximate polynomial degree of a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is the smallest $k$ such that there is a degree-$k$ polynomial approximating $f$ on 0/1 inputs. We survey recent developments on this subject.

**2000 Mathematics Subject Classification:** 20C30, 20J05.
**Keywords and Phrases:** polynomial approximation, Boolean functions complexity, decision tree complexity, communication complexity

## 1. Introduction

Let $n \geq 1$ be an integer. Denote by $[n] := \{1, 2, \cdots, n\}$. The set of $n$-bit binary strings is denoted by $\{0,1\}^n$. We also identify $\{0,1\}^n$ with the powerset of $[n]$ such that $x = x_1 x_2 \cdots x_n \in \{0,1\}^n$ identifies with $\{i : x_i = 1\}$. Denote by $\mathcal{F}_n := \{f : \{0,1\}^n \to \mathbb{R}\}$, and $\mathcal{B}_n := \{f : \{0,1\}^n \to \{0,1\}\} \subseteq \mathcal{F}_n$.

A function $f \in \mathcal{B}_n$ can be represented as a polynomial

$$f(x_1, x_2, \cdots, x_n) = \sum_{s \subseteq [n]} \alpha_s \Pi_{i \in s} x_i,$$

where each variable $x_i$ has at most degree 1. Such a multi-linear representation is unique. Denote the total degree of the polynomial by $\deg(f)$. For example, the ANDfunction $(x_1 \wedge x_2 \wedge \cdots \wedge x_n) = \Pi_{i=1}^n x_i$, and $\deg(\mathrm{AND}_n) = n$.

We are interested in approximating $f \in \mathcal{B}_n$ by a $\tilde{f} \in \mathcal{F}_n$ with the smallest possible degree.

*Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan 48109-2121, USA. E-mail: shiyy@eecs.umich.edu

**Definition 1..1.** *Let $\epsilon \in (0, 1/2)$. The* $\epsilon$-approximate polynomial degree *of $f \in \mathcal{B}_n$, denoted by $\widetilde{deg}_\epsilon(f)$, is the samllest integer $k$ such that there exists $\tilde{f} \in \mathcal{F}_n$ with $\deg(\tilde{f}) = k$, and $|\tilde{f}(x) - f(x)| \leq \epsilon$, for all $x \in \{0,1\}^n$.*

We are interested in estimating the asymptotic growth of $\widetilde{\deg}(f_n)$ for families $\{f_n \in \mathcal{F}_n : n \to \infty\}$. Since an $\epsilon$-approximation can be used to construct a $\epsilon'$ approximation with a factor of $O(\frac{1}{(1-2\epsilon)^2} \log \frac{1}{\epsilon'})$ increase in the degree (see Appendix for a proof), the choice of the constant $\epsilon$ is not important for our purpose. We shall set $\epsilon = 1/3$ when it is omitted.

While the use of polynomials for the study of Boolean functions has a long history, the study of the above notion of polyomial approximation of Boolean functions is more recent and its power is yet to be fully explored.

## 2. Discrete versions of Markov and Bernstein Inequalities

For a univariate polynomial $P(x)$, denote by $\|P\| := \max_{x \in [-1,1]} |P(x)|$.

**Theorem 2..1 (A. A. Markov).** *For a polynomial $P(x)$ of degree $d$,*

$$\|P'\| \leq d^2 \|P\|.$$

**Theorem 2..2 (Bernstein).** *For a polynomial $P(x)$ of degree $d$,*

$$|P'(x)| \leq \frac{d\|P\|}{\sqrt{1-x^2}}, \quad \forall x \in (-1,1).$$

**Theorem 2..3 (Paturi[15]).** *Let $P(x)$ be a polynomial of degree $d$, $a, b \in \mathbb{Z}$, $a < b$, and $\xi \in [a,b]$. Suppose that*

*(1) $|P(i)| \leq 1$ for all integers $i \in [a,b]$, and*
*(2) $|P(\lceil xi \rceil) - P(\xi)| \geq c$ for some constant $c > 0$.*

*Then*

$$d = \Omega(\sqrt{(\xi - a + 1)(b - \xi + 1)}.$$

*In particular,*

$$d = \Omega(\sqrt{b-a}).$$

This Theorem may appear rather technical, yet it yields a complete characterization of the approximate degree of symmetric Boolean functions. For $x \in \{0,1\}^n$, denote by $|x|$ the Hamming weight of $x$, i.e., the number of 1's in $x$. A function $f \in \mathcal{B}_n$ is *symmetric* if for a function $D : \{0, 1, \cdots, n\} \to \{0,1\}$ such that $f(x) = D(|x|)$, for all $x \in \{0,1\}^n$. Denote by $\ell_0(f) := \max\{i : 0 \leq i \leq n/2,\ D(i-1) \neq D(i)\} \cup \{0\}$, and $\ell_1(f) := \max\{n - i : n/2 \leq i \leq n-1,\ D(i) \neq D(i+1)\} \cup \{0\}$.

**Theorem 2..4 (Paturi).** *For any symmetric $f \in \mathcal{B}_n$,*

$$\widetilde{deg}(f) = \Theta(\sqrt{n(\ell_0(f) + \ell_1(f))}).$$

Another direct consequence of Theorem 2..3 states that if a function is highly sensitive to the changes of its input bits, then it has a high approximation degree. For an $x \in \{0,1\}^n$ and $i \in [n]$, denote by $x^i \in \{0,1\}^n$ the $n$-bit string identical to $x$ except on the $i$'th position. Define the *sensitivity* of $f \in \mathcal{B}_n$, denoted by $\mathrm{S}(f)$, to be the maximum size of a set $S$ such that there exists $x \in \{0,1\}^n$ with $f(x) \neq f(x^i)$, for all $i \in S$.

**Theorem 2..5 (Nisan and Szegedy).** *For any $f \in \mathcal{B}_n$, $\widetilde{deg}(f) = \Omega(\sqrt{S(f)})$.*

The proof goes by *symmetrizing* $f$. More specificially, define $\bar{f} : \{0, \cdots, k\} \to [0,1]$ as follows.

$$\bar{f}(i) := \frac{1}{\binom{k}{i}} \sum_{z \subseteq [k], |z|=i} f(x \oplus z).$$

Then $\bar{f}(1) \neq \bar{f}(0)$, and $\bar{f}(i) \in [0,1]$, for all $1 \leq i \leq k$. Therefore, $\widetilde{\deg}(f) = \Omega(\sqrt{k})$.

Essentially the same proof gives a better bound in terms of the *block-sensitivity* of $f$, denoted by $\mathrm{bs}(f)$, which is the maximum number $k$ such that there exist $x \in \{0,1\}^n$, and $k$ pairwise non-intersecting subsets $S_1$, $S_2$, $\cdots$, $S_k \subseteq [n]$ flipping all the Boolean varaibles in any of which will change the function value.

**Theorem 2..6 (Nisan and Szegedy).** *For any $f \in \mathcal{B}_n$, $\widetilde{deg}(f) = \Omega(\sqrt{bs(f)})$.*

Both bounds are tight in general — the OR function is an example on which both inequalities become equal (asymptotically). It is a long-standing open problem if $\mathrm{bs}(f) = O(\mathrm{S}(f)^{O(1)})$, for all $f \in \mathcal{B}_n$. It is known that $\widetilde{\deg}(f) = O(\mathrm{bs}(f)^3)$, thus the question is equivalent to if $\widetilde{\deg}(f) = O(\mathrm{S}(f)^{O(1)})$.

# 3. Relations with other complexity measures

It turns out that $\widetilde{\deg}(f)$ and $\deg(f)$ are polynomially related for any $f \in \mathcal{B}$.

**Theorem 3..1 (Nisan and Segedy; Beal et al.).** *For any $f \in \mathcal{B}$,* $\deg(f) = O(\widetilde{deg}(f)^6)$.

The proof was done by relating both quantities to the *deterministic decision tree complexity* of $f$. In the decision tree model of computation, an algorithm accesses an input $x = x_1 \cdots x_n \in \{0,1\}^n$ by making queries of the type "$x_i =?$". The complexity of the algorithm is the maximum number of such queries for the worst case input, and the minimum complexity of a correct algorithm is the *deterministic decision tree complexity* of $f$, denoted by $D(f)$.

Equivalently, a deterministic decision tree can be represented as a rooted tree such that each internval vertex is labeled with a variable and each edge, as well as each leaf, is labeled with either 0 or 1. On an input $x = x_1 x_2 \cdots x_n \in \{0,1\}^n$, the output of the algorithm is the label of the leaf reached from the root following the edge labels equal to the values of the internal variables. Thus, a decision tree of depth $k$ gives a polynomial of degree $\leq k$ for $f$:

$$f_T(x_1, x_2, \cdots, x_n) = \sum_{\ell} \Pi_i r_i,$$

where $\ell = v_1 v_2 \cdots v_i \cdots$ is a path from the root to a leaf labeled with 1, and $r_i$ is either the variable label $x_i$ of the internal vertex $v_i$ or its negation, depending on if the edge $v_i v_{i+1}$ is labeled 1 or 0. Therefore, $\mathrm{D}(f) \geq \deg(f) \geq \widetilde{\deg}(f)$. It turns out that $D(f) \leq \mathrm{bs}(f)^3$ [14]. Thus $\widetilde{\deg}(f) = \Omega(D(f)^{1/6}) = \Omega(\deg(f)^{1/6})$.

The approximate degree $\widetilde{\deg}(f)$ is also a lower bound for the *randomized* and even *quantum* decision tree complexity of $f$. In both the randomized and the quantum models, we allow the algorithm to make error with a probablity no more than 1/3 (or any constant bounded from above by 1/2). A randomized decision tree algorithm is a probablistic distribution on deterministic decision tree. Thus a randomized decision tree algorithm that makes $k$ queries (on the worst case input) gives a polynomial $\tilde{f}(x_1, x_2, \cdots, x_n)$ which is the convext combination of the polynomials of each underlying deterministic decision tree. The degree $\deg(\tilde{f}) \leq k$. Thus the randomized decision tree complexity $\mathrm{R}(f) \geq \widetilde{\deg}(f)$.

The quantum decision tree model is slightly more involved; but it follows from the definition that the acceptance probablity of a quantum decision tree algorithm that makes $T$ queries is a polynomial of degree $\leq 2T$. Thus the quantum decision tree cmplexity $\mathrm{Q}(f) \geq \widetilde{\deg}(f)/2$.

Since $\widetilde{\deg}(f)$ is polynomially related to $\mathrm{R}(f)$, and by definition, $\mathrm{Q}(f) \leq \mathrm{R}(f) \leq D(f)$, those results together implies that the following complexity measures are polynomially related: $\widetilde{\deg}(f)$, $\deg(f)$, $\mathrm{Q}(f)$, $\mathrm{R}(f)$, and $D(f)$. Asymptotic gaps between each neiboring pairs in this list are also known (see, e.g. the survey [5]).

# 4. Approximate degree lower bounds for partial functions

While the various decision tree complexities are polynomially related for all *total* functions $f \in \mathcal{B}$, exponential gaps are possible for *partial* functions. For example, at the heart of the celebrated quantum algorithm by Peter Shor for Interger Factorization is a super-fast decision tree algorithm solving the Period Finding Problem: given a function $f : \mathbb{Z}_N to \mathbb{Z}_N$ as an oracle, find the minimum $r$ such that $f(x) = f(x+r)$, for all $x \in \mathbb{Z}_N$. The function is promised to be periodic and the complexity measure is the number of evaluations of $f(x)$ on an $x$ chosen by the algorithm. Shor's algorithm finds $r$ in $O(\log^3 N \log\log N)$ number of function evaluations while any classical algorithm requires $N^{\Omega(1)}$ evalutions [7].

The polynomial method remains a powerful tool for proving quantum decision tree lower bounds for partial functions. A problem of particular interest is the Collision-Finding Problem: Given a two-to-one function $f : [N] \to [M]$ as an oracle, find $i$ and $j$, such that $i \neq j$ and $f(i) = f(j)$. A super-efficient quantum algorithm for Collision would in fact have a disppointing consequence: it'd be likely that no digital signature scheme could be resilient to quantum attack. A closely related problem is Element Distinctness Problem, which is to decide if $f : [N] \to [M]$ is one-to-one. The complexity of this elementary computational problem is surprisingly difficult to analyze, and has been the subject of many studies. An algorithm for solving Collision-Finding can be used to decide if an input function $f : [N] \to [M]$ is one-to-one or two-to-one. This latter task, referred to as the Collision Problem can in turn be solved by applying an algorithm for Element Distinctness on a random set of $\Theta(\sqrt{N})$ inputs to $f$. Thus a $t(N)$ lower bound for Collision Problem implies a $t(\Theta(N^2))$ lower bound for Element Distinctness.

Consider now lower-bouding the quantum complexity of the Collision Problem [1, 13]. Let $\delta_{ij}$, $1 \leq i \leq N$, $1 \leq j \leq M$, be a Boolean variable indicating if $f(i) = j$. It follows from the definition of the quantum decision tree model that any $T$-query quantum algorithm for the Collision Problem gives rise to a polynomial $P(\delta_{11}, \cdots, \delta_{NM})$ of degree $\leq 2T$ approximating a Boolean function for the Collision Problem. Thus a lower bound on $\deg(P)$ translates to a quantum lower bound. Note that $P$ is partial function, as not all assignments to the $\delta_{ij}$ variables correspond to a valid Collision Problem instance. While $P$ is not a totally symmetric function, it has a high degree of symmetry: for any permutation $\sigma$ on $[N]$ and any permutation $\tau$ on $[M]$, $P(\delta_{\sigma(i),\tau(j)}) = P(\delta_{ij})$.

The idea to make use of this symmetry is to averaging $P$ over a family of distributions parameterized by a triple $(m, a, b)$, where $0 \leq m \leq N$, $a \mid m$, and $b \mid N - m$. A triple satisfying those conditions is said to be *valid*. For valid triple, the corresponding ditribution is a random permutation of a fixed function which is $a$-to-1 on $m$ inputs and $b$-to-one on the remaining $N - m$ inputs. A remarkable property of the resulting averaged function $Q(m, a, b)$ is that it is a polynomial of degree $\leq \deg(P)$ in $m$, $a$, $b$. Furthermore, we have $Q(N, 1, 1) \approx 0$, $Q(N, 2, 2) \approx 1$, and there are plenty of valid triples in the grid $[N] \times [N] \times [N]$. By restricting $Q$ to various lines of valid triples, one can apply Theorem 2..3 to derive the $\Omega(N^{1/3})$ lower bound.

This lower bound implies a $\Omega(N^{2/3})$ degree lower bound for Element Distinctness Problem when $M = \Omega(N^2)$ through the reduction described above. This restriction on $M$ was removed by Ambainis[2], who proved that for any problem on an oracle function $f : [N] \to [M]$, $M \geq N$, that is symmetric under permutations of the domain and the range, $M$ does not affect the polynomial degree. Thus Element Distinctness Problem for $M = N$ also requires $\Omega(N^{2/3})$ degree to approxiate. An interesting consequence of this lower bound is that the two level AND-OR tree

$$T_{\wedge_N \vee_N}(x_{11}, x_{12}, \cdots, x_{NN}) := \bigwedge_{i=1}^{N} \bigvee_{j=1}^{N} x_{ij}$$

requires $\Omega(N^{2/3})$ degree to approximate. It remains open to close the gap between this bound and the best known upper bound of $O(N)$ degree.

## 5. Application in communication complexity

We now turn to a surprising application of the approximate degree method in resolving an important problem in communication complexity.

Let $f$ be a Boolean function whose domain $\mathrm{dom}(f) \subseteq X \times Y$, for some finite set $X$ and $Y$. Consider the computation of $f$ by two parties, Alice and Bob, who has $x \in X$ and $y \in Y$, respectively. Unless $f$ trivially depends on one of its arguments, they need to communicate in order to compute $f(x, y)$. The minimum amount of information they need to exchange for the worst case input is the *communication complexity* of $f$, which has a few variants including deterministic ($\mathrm{DC}(f)$), randomized ($\mathrm{RC}(f)$), or quantum ($\mathrm{QC}(f)$) complexities. In the randomized and quantum cases, a protocol is allowed to start with shared random coins (or quantum entanglement in the quantum case), and to err with a probability $\leq 1/3$. Since its introduction by Yao in 1979 [20], communication complexity has developed into a major branch of complexity

theory with a wide range of applications in VLSI design, circuit complexity, data streaming computation, etc. The monograph [12] suveys results up to 1997.

The major problems on communication complexities are often to prove lower bound on specfic problems, either in an application context, or to compare the power of different communication models. In particular, the Log-Rank Conjecture is one of the best known open problems in communication complexity. We identify a function $F : X \times Y \to \{0,1\}$ with the matrix $[F(x,y)]_{x\in X, y\in Y}$, and denote by $\widetilde{\text{rank}}(F) = \min\{\text{rank}(\tilde{F}) : \|\tilde{F} - F\|_\infty \le 1/3\}$ the *approximate rank* of $F$.

**Conjecture 5..1 (Log-Rank Conjecture).** *For any $F : X \times Y \to \{0,1\}$, both $R(F) = \log(\widetilde{rank}(F))^{O(1)}$.*

The following is perhaps the most important open problem concerning quantum communication complexity.

**Conjecture 5..2 (Log-Equivalence Conjecture).** *For any $F : X \times Y \to \{0,1\}$, $R(F) = O(Q(F)^{O(1)})$.*

Since it is known that $Q(F) = \Omega(\log \widetilde{\text{rank}}(F))$, the Log-Rank Conjecture implies the Log-Equivalence Conjecture.

A simple example to illustrate the favor of such problems is the Equality Problem: deciding if two input strings $x,\ y \in \{0,1\}^n$ are identical. It is not hard to show that $D(\text{Equality Problem}) = n$ while $R(\text{Equality Problem}) = O(\log n)$.

An important and highly nontrivial problem is $\mathsf{DISJ}$, where $X,\ Y \subseteq 2^{[n]}$, and

$$\mathsf{DISJ}_n(x,y) = \begin{cases} 1 & x \cap y \neq \oslash, \\ 0 & \text{otherwise.} \end{cases}$$

It is now known that $\text{DC}(\mathsf{DISJ}_n) = \Theta(n)$, where $\text{QC}(\mathsf{DISJ}_n) = \Theta(\sqrt{n})$. We shall focus on the proof for $\text{QC}(\mathsf{DISJ}_n) = \Omega(\sqrt{n})$ [16].

The starting point of Razoborov's proof is the following lemma.

**Lemma 5..3 (Razborov-Yao).** *Let $F \in \{0,1\}^{N\times M}$. If there is a quantum protocol computing $F$ (as a function of its row and column indices) by exchanging $q$ qubits with error probablity $\le \epsilon$, then its acceptance probability matrix $\tilde{F}$ satisfies*

$$\|\tilde{F} - F\|_\infty \le \epsilon, \quad \textit{and,}$$

$$\|\tilde{F}\|_{\mathsf{tr}} \le \sqrt{NM}2^{2q}.$$

For an integer $k$, $0 \le k \le n$, denote by $\binom{[n]}{k}$ the set of $k$-element subsets of $[n]$. Let $N := \binom{n}{k}$. To prove $Q(\mathsf{DISJ}_n) = \Omega(\sqrt{n})$, it suffices

to prove the same lower bound on $\mathsf{DISJ}_n\mid_{\binom{[n]}{k}\times\binom{[n]}{k}}$, the restriction of $\mathsf{DISJ}_n$ on $\binom{[n]}{k}\times\binom{[n]}{k}$ for some $k$. For $s\in\{0,1,\cdots,k\}$, let $J_s$ be the $\binom{n}{k}\times\binom{n}{k}$ Boolean matrix whose $(x,y)$ entry, $x,\ y\in\binom{[n]}{k}$, is 1 if and only if $|x\cap y|=s$. The matrices $\{J_s:0\le s\le k\}$ form a *Johnson scheme* (c.f. [8]). As a consequence, those matrices simulaneously diagonize and the eigenvalues have explicit formulae described by the *Hahn polynomials*. Let $\mu_s:=\frac{1}{N\binom{k}{s}\binom{n-k}{k-s}}J_s$ be the normalized $J_s$.

**Lemma 5..4.** *The matrices $\mu_s$, $0\le s\le k$, share the same eigenspaces $E_0$, $E_1$, $\cdots$, $E_k$. The eigenvalues $p_t(s)$ of $\mu_s$ corresponding to $E_t$ is a polynomial of degree $t$, and $|p_t(s)|=\frac{1}{N}\exp(-\Omega(t))$ whenever $k\le n/4$ and $s\le k/2$.*

Fix a quantum protocol for $\mathsf{DISJ}_n$ with $q$ qubits and error probability $\le 1/6$, and let $\tilde{F}\in\mathbb{R}^{N\times N}$ its acceptance probability matrix on $\binom{[n]}{k}\times\binom{[n]}{k}$. Then

$$g(s):=\mathrm{trace}\tilde{F}^T\mu_s\approx\begin{cases}0 & s=0,\\ 1 & s\ge 1.\end{cases}\tag{5.1}$$

With a slight abuse of notation, we use $E_t$ to denote the projection to $E_t$. By Lemma 5..4,

$$g(s)=\sum_t p_t(s)\mathrm{trace}(\tilde{F}^TE_t).$$

Thus when $k\le n/4$ and $s\le k/2$, for some constant $c>0$,

$$|\sum_{t>cq}p_t(s)\mathrm{trace}(\tilde{F}^TE_t)|\le\max_{t>cq}|p_t(s)|\cdot\|\tilde{F}^T\|_{\mathrm{tr}}\le 1/6.$$

Thus

$$\tilde{g}(s):=\sum_{0\le t\le cq-1}p_t(s)\mathrm{trace}(\tilde{F}^TE_t)$$

is a polynomial of degree $\le cq$ and approximates $\mathsf{OR}_{k/2}$. Since $\widetilde{\deg}(\mathsf{OR}_n)=\Omega(\sqrt{n})$, this means $q=\Omega(\sqrt{n})$ with $k=n/4$.

This lower bound method applies to all functions $F:\{0,1\}^n\times\{0,1\}^n\to\{0,1\}$ such that for some *symmetric* $f\in\mathcal{B}_n$,

$$F(x,y)=f(x_1\wedge y_1,x_2\wedge y_2,\cdots,x_n\wedge y_n),$$

where $x_i$ and $y_i$ are the $i$'th bit of $x$ and $y$, respectively. Razborov called those functions *symmetric predicate*, which should not be confused with those $F$ with $F(x,y)=F(y,x)$ for all $x$ and $y$.

**Theorem 5..5.** *For all symmetric predicate $F$ based on a symmetric $f \in \mathcal{B}_n$, $Q(F) = \Theta^*(\sqrt{n\ell_0(f)} + \ell_1(f))$, where $\Theta^*$ suppresses log factors. Furthermore, $R(F) = O(Q(F)^2)$.*

Recently, the author and Zhu [18] broadened the connection of approximate degree and communication complexity and and proved the Log-Equivalence Conjecture for a larger class of functions than symmetric predicates. We consider what we called the *block-composition* of a building block $g_k : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}$ with an arbitrary $f_n \in \mathcal{B}_n$, and seek conditions under which the composition $f_n \square g_k$ has polynomially related quantum and randomized communication complexity. We show that this is the case when $g_k$ itself is of a sufficiently high complexity measured by a combinatorial parameter.

To illustrate the main idea, consider $g_k$ being the Inner Product function $\mathrm{IP}_k$:

$$\mathrm{IP}_k(u,v) := \sum_{i=1}^{k} u_i v_k \mod 2.$$

**Theorem 5..6 (Shi and Zhu [18]).** *Let $k \geq 2\log_2 n + 5$. Then for any $f_n \in \mathcal{B}_n$, $D(f_n \square IP_k) = O((f_n \square IP_k)^7)$.*

The conclusion is in fact true with probability 1 for a random $g_k$.

The classical upper bound is obtained by having Bob simulate a decision tree algorithm for $f_n$, having Alice send him all the bits in a block $x_i$ when the $g_k(x_i, y_i)$ is needed. This takes $k\mathrm{D}(f_n) = O(k\widetilde{\deg}(f)^6)$ bits. It is not hard to show that $Q(\mathrm{IP}_k) = \Omega(k)$ thus it remains to prove $Q(f_n \square \mathrm{IP}_k) = \Omega(\widetilde{\deg}(f_n))$.

For $s \subseteq [n]$, denote by $\chi_s : \{0,1\}^n \to \{\pm 1\}$, $\chi_s(x) = (-1)^{s \cdot x}$. The set $\{\chi_s : s \subseteq [n]\}$ forms an orthonormal basis for $\mathcal{F}_n$ (the Walsh-Hadamard basis) with respect to the inner product $\langle f, g \rangle := \frac{1}{2^n} \sum_{x \subseteq [n]} f(x)g(x)$. Furthermore, $\deg(\chi_s) = |s|$.

Suppose $d = \widetilde{\deg}_\epsilon(f_n)$. Then there is no solution to the linear program whose unknowns are $\alpha_s$, $s \subseteq [n]$, $|s| \leq d-1$:

$$\Big| \sum_{\substack{s \in \subseteq [n] \\ |s| \leq d-1}} \alpha_s \chi_s(x) - f_n(x) \Big| \leq \epsilon.$$

By the duality of linear programming, there exists a polnomial $q$ such that

1. $q$ is orthogonal to all polynomials of degree $\leq d-1$. Thus, for some $\hat{q}_s \in \mathbb{R}$,
$$q = \sum_{\substack{s \in \subseteq [n] \\ |s| \geq d}} \hat{q} \chi_s.$$

2. $q^T f_n = 1$. This means that $q$ overlaps with $f_n$ substantially.
3. $\|q\|_1 \geq 1/\epsilon$. Thus $q$ is almost a distribution with signs.
4. $|\hat{q}| \leq \frac{1}{\epsilon 2^n}$.

Those remarkable properties are then used to turn $q$ into a witness that the truetable $F$ of the block-composition requires high trace norm to approximate.

# 6. Open Problems

As mentioned above, the question of whether $\mathrm{bs}(f)$ and $\mathrm{S}(f)$ are polynomially related is still open. Shi [17] showed that $\mathrm{S}(f)$ is polynomially related to another notion of approximation degree. Let $\ell = (a, b) \in \{0,1\}^n \times \{0,1\}^n$ be a line segment inside the the $n$-dimensional hypercube. A function $f \in \mathcal{B}_n$ can be extended to $[0,1]^n$ by convexity. The restriction of this extended $f$ on $\ell$, denoted by $f\mid_\ell$ is a univariate polynomial. Define the linear approximate polynomial degree, denoted by $\widetilde{\deg}(f)$, to be the smallest degree of a polynomial approximating $f\mid_\ell$. Then $\widetilde{\deg}(f) = \Omega(\sqrt{\mathrm{S}(f)})$ and $\widetilde{\deg}(f) \leq \mathrm{S}(f)$. Thus the block-sensitivity v.s. sensitivity question becomes an approximation theory problem: are $\widetilde{\deg}(f)$ and $\widetilde{\deg}(f)$ polynomially related.

Extending the results of Razborov and Shi and Zhu to functions of the form $f_n \square \wedge$ for an arbitrary $f_n$ would be a significant progress toward resolving the Log-Equivalence Conjecture. It appears more critical to $Q(f_n \square \wedge)$ is $\tilde{m}(f)$, the minimum number of monomials of an approximating polynomial of $f_n$, than its approximate degree, as suggested by the work of Burhman and de Wolf [4]. Developing techniques to lower bound $\tilde{m}(f)$ is of great interest.

A long-standing open problem in theoretical computer science and combinatorics is the randomized decision tree complexity of nontrivial monotone graph properties. A graph on $n$ vertices can be represented by the $\binom{n}{2}$-bit characteristic string of its edges. A Boolean function $\mathcal{P} \in \mathcal{B}_{\binom{n}{2}}$ is called a graph property if it is invariant on isomorphic graphs, and is called monotone if adding an edge to a graph would not change the function value from 1 to 0. The *Evasiveness Conjecture* states that for any nontrivial (i.e. non-constant) monotone graph property $\mathcal{P}$, $D(\mathcal{P}) = \binom{n}{2}$. The Conjecture was proved only for $n$ being a power of prime [10], and other restricted classes of graphs (e.g.[6]). In the prime power result, and likely all other known cases, is actually stronger: $\deg(\mathcal{P}) = \binom{n}{2}$. No example is known that $\deg(\mathcal{P}) < \binom{n}{2}$.

Parallel to Evasiveness Conjecture are the conjectures that $\mathrm{R}(\mathcal{P}) = \Theta(n^2)$ and $\mathrm{Q}(\mathcal{P}) = \Omega(n)$. The best known lower bounds to date are $\mathrm{R}(\mathcal{P}) = \Omega(n^{4/3})$ [9] and $\mathrm{Q}(\mathcal{P}) = \Omega(n^{2/3})$ [19]. Applying the approximate

degree method appears to be a promising direction.

The approximate degree method has also been very successful in two other areas: direct product theorems and time-space tradeoffs. Intuitively, solving $k$ instances of a problem simulatenously requires $k$ times of the resouce for solving each instance. It is, however, rather difficult to prove in general. Strong lower bounds in the following form, so called "strong product theorem", were proved by the approximate degree method [11, 3] for some functions (such as $\mathsf{OR}_n$) in the decision tree model: the probability of computing $k$ instances of a function $f$ simultaneously would be expoenentially small in $k$ if the algorithm asks substantially less queries than $k$ times of those required for computing one instance of $f$. Such strong result in turn implies strong tradeoff relations between the number of queries and the size of the storage required [11, 3]. Proving similar results for a broader class of functions is an important open problem.

All the approximate degree lower bounds we know were proved by reducing the function considered to a univariat function. That $\deg(f)$ and $\widetilde{\deg}(f)$ are polynomially related for $f \in \mathcal{B}$ means that this symmetrization approach is satisfactory if one is not concerned with the constants in the exponent. However, to precisely determine the approximate degree of less symmetric functions appears to require fundamentally different approach. A specific case of interset is that of two level AND-OR tree. univariat functions.

# References


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

[2] A. Ambainis. Quantum lower bounds for collision and element distinctness with small range. Pre-print: quant-ph/0305179, 2003.

[3] A. Ambainis, R. Špalek, and R. de Wolf. A new quantum lower bound method,: with applications to direct product theorems and time-space tradeoffs. In ACM, editor, *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing 2006, Seattle, WA, USA, May 21–23, 2006*, pages 618–633, pub-ACM:adr, 2006. ACM Press.

[4] H. Buhrman and R. de Wolf. Communication complexity lower bounds by polynomials. In *IEEE Conference on Computational Complexity*, pages 120–130, 2001.

[5] H. Buhrman and R. de Wolf. Complexity measures and decision tree complexity: a survey. *Theoretical Computer Science*, 288(1):21–43, Sept. 2002.

[6] A. Chakrabarti, S. Khot, and Y. Shi. Evasiveness of subgraph containment and related properties. *sigcom*, 31(3):866–875, 2002.
[7] R. Cleve. The query complexity of order-finding. *Inf. Comput.*, 192(2):162–171, 2004.
[8] C. D. Godsil. *Algebraic combinatorics*. Chapman and Hall Mathematics Series. Chapman & Hall, New York, 1993.
[9] P. Hajnal. An $\omega(n^{\frac{4}{3}})$ lower bound on the randomized complexity of graph properties. *Combinatorica*, 11(2):131–143, 1991.
[10] J. Kahn, M. Saks, and D. Sturtevant. A topological approach to evasiveness. In *24th Annual Symposium on Foundations of Computer Science*, pages 31–33, Los Alamitos, Ca., USA, Nov. 1982. IEEE Computer Society Press.
[11] H. Klauck, R. Špalek, and R. de Wolf. Quantum and classical strong direct product theorems and optimal time-space tradeoffs. *SIAM Journal on Computing*, 36(5):1472–1493, 2007.
[12] E. Kushilevitz and N. Nisan. *Communication complexity*. Cambridge University Press, 1997.
[13] S. Kutin. A quantum lower bound for the collision problem. Preprint: quant-ph/0304162, 2003.
[14] N. Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, Dec. 1991.
[15] R. Paturi. On the degree of polynomials that approximate symmetric Boolean functions (preliminary version). In *Proceedings of the Twenty-Fourth Annual ACM Symposium on the Theory of Computing*, pages 468–474, Victoria, British Columbia, Canada, May 1992.
[16] A. A. Razborov. Quantum communication complexity of symmetric predicates (Russian). *Izvestiya:Mathematics*, 67(1):145–159, 2003. English translation available at http://genesis.mi.ras.ru/ razborov/qcc_eng.ps.
[17] Y. Shi. Approximating linear restrictions of Boolean functions. Manuscript.
[18] Y. Shi and Y. Zhu. The quantum communication complexity of block-composed functions. Manuscript, 2007.
[19] A. C. Yao, 2001. Presentation at Institute for Advance Stuties.
[20] A. C.-C. Yao. Some complexity questions related to distributive computing. In *Eleventh Annual ACM Symposium on Theory of Computing (STOC '79)*, pages 209–213, New York, Apr. 1979. ACM.