A Dual Polynomial for OR

Robert Špalek
Google, Inc.*
spalek@google.com

Abstract

We reprove that the approximate degree of the OR function on n bits is $\Omega(\sqrt{n})$. We consider a linear program which is feasible if and only if there is an approximate polynomial for a given function, and apply the duality theory. The duality theory says that the primal program has no solution if and only if its dual has a solution. Therefore one can prove the nonexistence of an approximate polynomial by exhibiting a dual solution, coined the *dual polynomial*. We construct such a polynomial.

1 Introduction

We study the approximation of Boolean functions by real-valued polynomials. This line of research was initiated by Minsky and Papert [MP68]. An n-bit Boolean function f is represented by a multivariate polynomial $p(x_1, \ldots, x_n)$. Nisan and Szegedy [NS94] defined the approximate degree of a function f under the ℓ_∞-norm, denoted $\widetilde{\deg}(f)$, as the smallest degree for which there exists a polynomial that is close to the function pointwise. Several complexity measures have been since shown to be lower-bounded in terms of $\widetilde{\deg}(f)$: circuit size [Bei93], or quantum query complexity [BBC$^+$01]. Consider the OR function on n bits. Nisan and Szegedy [NS94] showed that $\widetilde{\deg}(\mathrm{OR}_n) = \Theta(\sqrt{n})$, and Paturi [Pat92] extended their bound to all symmetric functions.

The existence of an approximate polynomial can be described by a linear program; let us coin it the primal program. Using the duality theory of linear programming, one can show the non-existence of an approximate polynomial for a function f by exhibiting a solution to its dual program, a so-called *dual polynomial* for f. Recently, several papers have appeared that use dual polynomials to prove good communication complexity lower bounds: Sherstov [She07] and Shi and Zhu [SZ07] show two-party quantum communication lower bounds, and Lee and Shraibman [LS08] and Chattopadhyay and Ada [CA08] show multi-party randomized communication lower bounds in the number-on-the-forehead model. The basic idea of these papers is as follows. One defines a special *pattern matrix* (or tensor in the multi-party case) whose entries are values of a certain polynomial. The structure of the pattern matrix allows one to relate properties of the polynomial to properties of the matrix, such as its trace norm. The pattern matrix formed from the dual polynomial forms a witness to the large trace norm of the matrix. The communication complexity is then lower-bounded in terms of the trace norm. None of these papers actually presents an explicit dual polynomial for any function; they only use its existence and some inequalities guaranteed by the duality principle from the known bounds on the approximate degree.

It is natural to ask what a dual polynomial looks like for the simplest functions. In this short note, we address this question and present an asymptotically optimal dual polynomial for the OR function. Our proof extends the ideas of Buhrman and Szegedy [BS03].

*Most of the work conducted while at CWI, Amsterdam, in February 2003.

2 Preliminaries

2.1 Symmetric polynomials

We represent Boolean functions by polynomials in the Fourier basis, where $+1$ corresponds to the logical value 0 (false) and -1 to the logical value 1 (true). In this basis, multiplication corresponds to the exclusive OR. We say that $f : \{\pm 1\}^n \to \{\pm 1\}$ is a *symmetric* function, if $f(x) = f(x_\sigma)$ for every permutation $\sigma \in S_n$ and $x \in \{\pm 1\}^n$, where x_σ denotes a σ-permuted version of x, with $(x_\sigma)_i = x_{\sigma(i)}$.

Let $p : \{\pm 1\}^n \to \Re$ be a polynomial in variables $x_1, \ldots, x_n$. Since $x_i^2 = 1$, we can restrict ourselves to *multilinear* polynomials, where each variable appears with degree at most 1. We say that p has *degree* d and *pure high degree* d', if each term in p is a product of at most d and at least d' variables. We say that p is an *ε-approximation* for a function f, if $|p(x) - f(x)| \leq \varepsilon$ for every $x \in \{\pm 1\}^n$. If p is an ε-approximation of a symmetric function f, then there exists a symmetric polynomial p' with the same degree, pure high degree, and approximation factor: $p'(x) = \frac{1}{n!} \sum_{\sigma \in S_n} p(x_\sigma)$.

Let $[n] = \{0, 1, \ldots, n\}$. Given a symmetric function $f : \{\pm 1\}^n \to \{\pm 1\}$, one can define a single-variate function $F : [n] \to \{\pm 1\}$ such that $f(x) = F(|x|)$, where $|x| = \frac{n - (x_1 + \cdots + x_n)}{2}$ is the Hamming weight of x, i.e., the number of minuses in x. Analogously, following [MP68], given a symmetric multilinear polynomial $p : \{\pm 1\}^n \to \Re$, one can define a *single-variate* polynomial $P : [n] \to \Re$ of the same degree such that

$$P(k) = p(\underbrace{-1, \ldots, -1}_{k}, \underbrace{+1, \ldots, +1}_{n-k}) \qquad \text{for all } k \in [n],$$

$$p = P\left(\frac{n - (x_1 + \cdots + x_n)}{2}\right) \qquad \mod (x_1^2 - 1), \quad \ldots, \quad \mod (x_n^2 - 1).$$

Note that the pure high degree of p does not correspond to the smallest degree of a k-term in $P(k)$. When we talk about the pure high degree of a single-variate polynomial, we mean the pure high degree of its corresponding multilinear polynomial.

Let $p, q : \{\pm 1\}^n \to \Re$. Define a *scalar product* as $p \cdot q = \sum_{x \in \{\pm 1\}^n} p(x)q(x)$. This induces a scalar product $P \cdot Q = \sum_{i=0}^n \binom{n}{i} P(i)Q(i)$ on the space of symmetric polynomials. Similarly, the *ℓ_1-norm* $\|p\|_1 = \sum_{x \in \{\pm 1\}^n} |p(x)|$ induces an ℓ_1-norm $\|P\|_1 = \sum_{i=0}^n \binom{n}{i} |P(i)|$.

Let $p : \{\pm 1\}^n \to \Re$ be a multilinear polynomial of degree d and pure high degree d', and consider $q(x) = p(x) \cdot (x_1 \cdots x_n) \mod (x_i^2 - 1)$. In the functional interpretation, $q(x)$ equals $p(x)$ *multiplied by the parity of* x. Thanks to the term cancellation $x_i^2 = 1$, each term in q corresponds to the complement of a term in p, and therefore q has degree $n - d'$ and pure high degree $n - d$. Now, assume that p (and thus also q) are symmetric, and consider their corresponding single-variate polynomials P, Q. Then $Q(k) = P(k) \cdot (-1)^k$, and the degree of P corresponds to n minus the pure high degree of Q and vice versa.

2.2 Linear program for polynomial approximation

Theorem 1. *A total Boolean function $f : \{\pm 1\}^n \to \{\pm 1\}$ has ε-approximate degree at least d if and only if there exists a polynomial $b : \{\pm 1\}^n \to \Re$ with pure high degree d such that $\frac{\|b\|_1}{b \cdot f} < \frac{1}{\varepsilon}$.*

Proof. f can be ε-approximated by a polynomial of degree $d-1$ is equivalent to the feasibility of the following primal linear program. Consider the *Fourier basis* on the space of multilinear polynomials: $\{\chi_S\}_{S \subseteq \{1,\ldots,n\}}$, where $\chi_S(x) = \prod_{i \in S} x_i$. Let $F = \{\chi_S(x)\}_{x,S}$ denote the Fourier transform over $\mathbb{Z}_2^n$, indexed by $\{\pm 1\}^n$ and $S \subseteq \{1, \ldots, n\}$, and let a denote a vector of Fourier coefficients.

$$\begin{aligned} Fa &\geq f - \varepsilon \\ Fa &\leq f + \varepsilon \\ a_S &= 0 \text{ for } |S| \geq d \end{aligned}$$

The primal program is unfeasible if and only if its dual is feasible. The dual program is as follows.

$$\begin{array}{r} (b^+ - b^-) \cdot f > (b^+ + b^-) \cdot \varepsilon \\ (b^+ - b^-)F = c \\ b^+, b^- \geq 0 \\ c_S = 0 \text{ for } |S| < d \end{array} \quad \Longleftrightarrow \quad \begin{array}{r} b \cdot f > |b| \cdot \varepsilon \\ bF = c \\ c_S = 0 \text{ for } |S| < d \end{array}$$

We can assume that b^+ and b^- of the optimal solution are disjoint, i.e., $b^+(x)b^-(x) = 0$ for each x, otherwise we could lower the right-hand side of the first inequality by subtracting the same constant $\min(b^+(x), b^-(x)) > 0$ from both $b^+(x)$ and $b^-(x)$, and the remaining expressions would stay unchanged. Let $b = b^+ - b^-$ and $|b| = b^+ + b^-$. The constraints $bF = c$ and $c_S = 0$ for $|S| < d$ say that b has pure high degree d. The dual is feasible if and only if there exists such a b with $b \cdot f > |b| \cdot \varepsilon = \varepsilon \|b\|_1$. □

Note that if f is symmetric, then it suffices to look for a dual polynomial b in the space of symmetric polynomials. Let us reformulate the condition in the language of single-variate polynomials.

Corollary 2. *A total symmetric Boolean function $F : [n] \to \{\pm 1\}$ has ε-approximate degree at least d if and only if there exists a polynomial $B : [n] \to \Re$ with pure high degree d such that $\frac{\|B\|_1}{B \cdot F} < \frac{1}{\varepsilon}$.*

3 Dual polynomial for OR

First, we define a certain low-degree polynomial P and show that its norm $\|P\|_1$ is not too large compared to its value $P(0)$. This polynomial will be crucial for defining the dual polynomial for OR. The design of our polynomial comes from extending the ideas of Buhrman and Szegedy [BS03].

Definition 3. *Let $m = \lfloor \sqrt{n} \rfloor$ and let $\mathcal{S} = \{i^2 : i \in [m]\} \cup \{2\}$ denote the set containing the integer squares up to n and the number 2. Define a polynomial*

$$P(x) = 2(-1)^{n-m-1} \frac{m!^2}{n!} \cdot \prod_{i \in [n] - \mathcal{S}} (x - i) \ .$$

The multiplicative factor of P is chosen such that $P(0) = 1$. The degree of P is $n - m - 1$.

Lemma 4. *For every pair of integers k, m with $k \leq m$, $\frac{m!^2}{(m+k)!(m-k)!} \leq 1$.*

Proof. The term is a product of numbers that are all smaller than 1:

$$\frac{m!^2}{(m+k)!(m-k)!} = \frac{m(m-1)\ldots(m-k+1)}{(m+k)(m+k-1)\ldots(m+1)} = \prod_{i=1}^{k} \left(1 - \frac{k}{m+i}\right) \leq 1 \qquad \square$$

Lemma 5. $\binom{n}{2}|P(2)| \leq 12$ *and* $\binom{n}{k^2}|P(k^2)| \leq \frac{8}{k^2}$ *for every $k = 1, 2, \ldots, m$.*

Proof. First, we substitute $x = 2$ into $|P(x)|$ and rewrite the product over $i \in [n] - \mathcal{S}$ as the ratio of two products, one over $i \in [n] - \{0, 1, 2\}$ and one over $i \in \mathcal{S} - \{0, 1, 2\}$. We then pull the $j = 2$ term out of the product in the denominator, use $|j^2 - 2| < j^2 - 4$, and apply Lemma 4.

$$\begin{aligned} |P(2)| &= 2\frac{m!^2}{n!} \frac{(n-2)!}{2\prod_{j=3}^{m} |2 - j^2|} < \frac{m!^2}{n!} \frac{(n-2)!}{\prod_{j=3}^{m} (j^2 - 4)} \\ &= \frac{m!^2}{n!} \frac{(n-2)!}{\prod_{j=3}^{m} (j+2)(j-2)} = \frac{1}{n(n-1)} \frac{m!^2}{\frac{(m+2)!}{4!}(m-2)!} \leq \frac{4!}{n(n-1)} = \frac{12}{\binom{n}{2}}. \end{aligned}$$

Second, we substitute $x = k^2$ to $|P(x)|$ and rewrite the product over $i \in [n] - \mathcal{S}$ as the ratio of two products, one over $i \in [n] - \{k^2\}$ and one over $i \in \mathcal{S} - \{k^2\}$. The term $i = k^2$ does not appear in any of products, because it is 0.

$$\begin{aligned}
|P(k^2)| &= 2\frac{m!^2}{n!} \cdot \frac{\prod_{\substack{i \in [n] \\ i \neq k^2}} |k^2 - i|}{|k^2 - 2| \cdot \prod_{\substack{j \in [m] \\ j \neq k}} (k + j)|k - j|} \\
&= 2\frac{m!^2}{n!} \cdot \frac{k^2!(n - k^2)!}{\frac{(k+m)!}{2k \cdot (k-1)!} \cdot k!(m-k)!} \cdot \frac{1}{|k^2 - 2|} \\
&= 4 \cdot \frac{k^2!(n - k^2)!}{n!} \cdot \frac{m!^2}{(m+k)!(m-k)!} \cdot \frac{1}{|k^2 - 2|}
\end{aligned}$$

Apply Lemma 4 and $|k^2 - 2| \geq k^2/2$, which holds for all integers $k \geq 1$.

$$\leq \frac{4}{\binom{n}{k^2}} \cdot \frac{1}{|k^2 - 2|} \leq \frac{4}{\binom{n}{k^2}} \cdot \frac{1}{k^2/2} \leq \frac{8}{\binom{n}{k^2} k^2}.$$

Note that if we did not include the number 2 into $\mathcal{S}$, in Definition 3, then the upper bound on $|P(k^2)|$ would be much weaker, without the factor of $1/k^2$. □

Now we show that a constant fraction of the norm of P comes from the term $P(0) = 1$.

Theorem 6. $\|P\|_1 < 27$.

Proof. First, use the fact that $P(i) = 0$ for $i \in [n] - \mathcal{S}$, non-square integers i other than 2.

$$\begin{aligned}
\|P\|_1 &= \sum_{i=0}^{n} \binom{n}{i} |P(i)| = \sum_{i \in \mathcal{S}} \binom{n}{i} |P(i)| \\
&= P(0) + \binom{n}{2} P(2) + \sum_{k=1}^{m} \binom{n}{k^2} |P(k^2)|
\end{aligned}$$

Now, use $P(0) = 1$, Lemma 5, and $\sum_k \frac{1}{k^2} = \frac{\pi^2}{6}$.

$$\leq 13 + 8 \sum_{k=1}^{m} \frac{1}{k^2} < 13 + 8\frac{\pi^2}{6} < 27.$$ □

Finally, we are ready to present the dual polynomial for OR.

Theorem 7. *The $\frac{1}{14}$-approximate degree of OR on n bits is at least $\sqrt{n}$.*

Proof. Consider the polynomial

$$Q(k) = (-1)^k P(k) \ ,$$

that is P from Definition 3 multiplied by parity. We show that Q is a good dual polynomial for OR. First, the pure high degree of Q is $n - (n - m - 1) = m + 1 > \sqrt{n}$. Second, we compute the ratio from Corollary 2. Since $\mathrm{OR}(0) = 1$ and $\mathrm{OR}(k) = -1$ for $k \geq 1$, $Q \cdot \mathrm{OR} = 2Q(0) - Q \cdot 1 = 2Q(0)$, because Q has no constant coefficient. Now, we use Theorem 6 to upper-bound the numerator and conclude

$$\frac{\|Q\|_1}{Q \cdot \mathrm{OR}} = \frac{\|P\|_1}{2P(0)} < \frac{27}{2} < 14 \ .$$ □

4 Open problems

The approximate degree of the t-threshold function on n bits is $\Theta(\sqrt{t(n-t)})$ [Pat92]. It would be interesting to find an explicit dual polynomial for the threshold function. A good candidate may be $Q(k) = (-1)^k P(k)$ with

$$P(x) = \prod_{i \in [n]-T} (x - t - i) \ ,$$

where T is a set of integers that can be written as $k^2 - \ell^2$, where $k \in [\lfloor\sqrt{n-t}\rfloor]$ and $\ell \in [\lfloor\sqrt{t}\rfloor]$. Note that $|T| = \Theta(\sqrt{t(n-t)})$.

The approximate degree of the two-level AND-OR tree on n bits (with all gates of fan-in $\sqrt{n}$) is only known to lie between $O(\sqrt{n})$ and $\Omega(\sqrt[3]{n})$. Both bounds have been obtained through quantum algorithms, as follows. Consider a T-query quantum algorithm. Its acceptance probability on input x can be expressed as a $2T$-degree polynomial p in the variables $x_1, \ldots, x_n$ [BBC$^+$01]. If the algorithm computes a function f with bounded error, then p approximates f. Therefore quantum algorithms give approximate polynomials, and approximate degree lower bounds give quantum query lower bounds. For the two-level AND-OR tree, the upper bound is via a quantum search algorithm on noisy inputs [HMW03] and the lower bound is via a reduction from the element distinctness problem [AS04]. Can one compute the approximate degree of the AND-OR tree by showing a good dual polynomial?

Acknowledgments

We thank Harry Buhrman and Mario Szegedy for starting the project, coming up with the crucial ideas, and many fruitful discussions. We also thank Ronald de Wolf for fruitful discussions, and Troy Lee for proofreading.

References

[AS04] S. Aaronson and Y. Shi. Quantum lower bounds for the collision and the element distinctness problems. *Journal of the ACM*, 51(4):595–605, 2004.

[BBC$^+$01] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *Journal of the ACM*, 48(4):778–797, 2001. Earlier version in FOCS'98.

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

[BS03] H. Buhrman and M. Szegedy, 2003. Personal Communication.

[CA08] A. Chattopadhyay and A. Ada. Multiparty communication complexity of disjointness. Technical report, ECCC TR08-002, 2008.

[HMW03] P. Høyer, M. Mosca, and R. de Wolf. Quantum search on bounded-error inputs. In *Proc. of 30th ICALP*, LNCS 2719, pages 291–299, 2003.

[LS08] T. Lee and A. Shraibman. Disjointness is hard in the multi-party number-on-the-forehead model. In *Proc. of 23rd IEEE Complexity*, 2008. To appear.

[MP68] M. Minsky and S. Papert. *Perceptrons*. MIT Press, 1968.

[NS94] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4(4):301–313, 1994. Earlier version in STOC'92.

[Pat92] R. Paturi. On the degree of polynomials that approximate symmetric Boolean functions. In *Proc. of 24th ACM STOC*, pages 468–474, 1992.

[She07] A. Sherstov. The pattern matrix method for lower bounds on quantum communication. Technical report, ECCC TR07-100, 2007.

[SZ07] Y. Shi and Y. Zhu. The quantum communication complexity of block-composed functions. arXiv:0710.0095v3 [quant-ph], 2007.