A new rank technique for formula size lower bounds

Troy Lee[*]
troyjlee@gmail.com

Abstract

We introduce a new technique for proving formula size lower bounds based on matrix rank. A simple form of this technique gives bounds at least as large as those given by the method of Khrapchenko, originally used to prove an n^2 lower bound on the parity function. Applying our method to the parity function, we are able to give an exact expression for the formula size of parity: if $n = 2^\ell + k$, where $0 \leq k < 2^\ell$, then the formula size of parity on n bits is exactly $2^\ell(2^\ell + 3k) = n^2 + k2^\ell - k^2$. Such a bound cannot be proven by any of the lower bound techniques of Khrapchenko, Nečiporuk, Koutsoupias, or the quantum adversary method, which are limited by n^2.

1 Introduction

One of the most important open problems in complexity theory is to prove superlinear lower bounds on the circuit size of an explicit Boolean function. While this seems quite difficult, a modest amount of success has been achieved in the weaker model of formula size, a formula being a circuit where every gate has fan-out exactly one. The current best lower bound on the formula size of an explicit function is $n^{3-o(1)}$ [Hås98].

Besides proving larger lower bounds, many open questions remain about the formula size of basic Boolean functions—functions which are both very important in practice and are the constant companions of complexity theorists. One of the most startling such questions is the gap in our knowledge about the formula size of the majority function: the best lower bound is $\lceil n/2 \rceil^2$ while the best upper bound is $O(n^{4.57})$ [PPZ92]. Even in the monotone case, where a formula consists of only AND and OR gates, the best lower bound is $\lfloor n/2 \rfloor n$ [Rad97], while the best upper bound is $O(n^{5.3})$ by Valiant's beautiful construction [Val84].

One obstacle to proving larger formula size lower bounds seems to be what we call the n^2 barrier—most generic lower bound techniques seem to get stuck around n^2. The technique of Nečiporuk [Neč66] is limited to bounds of size $n^2/\log n$; the methods of Khrapchenko [Khr71], originally used to show a n^2 lower bound on the formula size of parity, Koutsoupias [Kou93], and the recent quantum adversary method [LLS06] all cannot prove lower bounds larger than

[*] LRI, Université Paris-Sud. Supported by a Rubicon grant from the Netherlands Organisation for Scientific Research (NWO). Part of this work conducted while at CWI, Amsterdam.

n^2; Karchmer, Kushilevitz, and Nisan [KKN95] introduce a promising technique based on linear programming but at the same stroke show that it cannot prove lower bounds larger than $4n^2$.

We introduce a new technique for proving formula size lower bounds based on matrix rank. Karchmer and Wigderson [KW88] show that formula size can be phrased as a communication complexity game, specifically as the communication complexity of a relation. Although matrix rank is one of the best tools available for proving lower bounds on the communication complexity of *functions* it has proved difficult to adapt to the relational case. Razborov [Raz90] uses matrix rank to show superpolynomial lower bounds on *monotone* formula size, but also shows [Raz92] that his method is limited to $O(n)$ bounds for general formulas.

While in its full generality our method seems difficult to apply, we give a simplified form which always gives bounds at least as large as the method of Khrapchenko, and even the quantum adversary method, and which *can* break the n^2 barrier: we apply it to the parity function and give an *exact expression* for the formula size of parity. Let $\oplus_n$ denote the parity function on n-bits, and let $L(f)$ denote the the number of leaves in a smallest formula which computes f.

Theorem 1 *If $n = 2^\ell + k$ where $0 \leq k < 2^\ell$, then*

$$L(\oplus_n) = 2^\ell(2^\ell + 3k) = n^2 + k2^\ell - k^2.$$

In Section 3 we present our method and show that it gives bounds at least as large as those of Khrapchenko. In Section 4 we apply the method to the parity function to prove Theorem 1. Finally, in Section 5 we look at the relative strength of different formula size techniques and show that the linear programming method of Karchmer, Kushilevitz, and Nisan [KKN95] is always at least as large as the quantum adversary method [LLS06].

2 Preliminaries

We will make use of Jensen's inequality. We will use the following form:

Lemma 2 (Jensen's Inequality) *Let $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function and a_i a set of positive real numbers for $i = 1, \ldots, n$. Then*

$$\phi\left(\frac{\sum_{i=1}^n a_i x_i}{\sum_{i=1}^n a_i}\right) \leq \frac{\sum_{i=1}^n a_i \phi(x_i)}{\sum_{i=1}^n a_i}.$$

2.1 Linear algebra

We will use some basic concepts from linear algebra. For a matrix A, let A^* be the transpose conjugate of A, that is $A^*[i,j] = \overline{A[j,i]}$. A matrix is Hermitian if $A = A^*$. We will use $\leq$ to refer to entrywise comparision of matrices: that is $A \leq B$ if $A[i,j] \leq B[i,j]$ for all (i,j). The shorthand $A \geq 0$ means that all entries of A are nonnegative. The rank of A, denoted by $\mathrm{rk}(A)$, is the number of linearly independent columns of A. The trace of A, written $\mathrm{Tr}(A)$, is the sum of the diagonal entries of A. For a Hermitian n-by-n matrix A, let $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$ be the eigenvalues of A. Let $\sigma_i(A) = \sqrt{\lambda_i(A^*A)}$ be the i^{th} singular value of A.

We will make use of three matrix norms. The Frobenius norm is the ℓ_2 norm of a matrix thought of as a long vector—that is

$$\|A\|_F = \sqrt{\sum_{i,j} A[i,j]^2}.$$

Notice also that $\|A\|_F^2 = \mathrm{Tr}(A^*A) = \sum_i \sigma_i^2(A)$. We will also use the trace norm, $\|A\|_{tr} = \sum_i \sigma_i(A)$. Finally, the spectral norm $\|A\| = \sigma_1(A)$. A very useful relationship between Frobenius norm, trace norm, and rank is the following:

Lemma 3 *Let A be a n-by-m matrix with $n \leq m$.*

$$\left\lceil \frac{\|A\|_{tr}^2}{\|A\|_F^2} \right\rceil \leq \mathrm{rk}(A).$$

Proof: The rank of A equals the number of nonzero singular values of A. Thus by the Cauchy–Schwarz inequality,

$$\left(\sum_{i=1}^n \sigma_i\right)^2 \leq \mathrm{rk}(A) \cdot \sum_{i=1}^n \sigma_i^2.$$

As rank is an integer, we obtain

$$\left\lceil \frac{\|A\|_{tr}^2}{\|A\|_F^2} \right\rceil \leq \mathrm{rk}(A).$$

□

A useful tool to lower bound the trace norm is the following:

Lemma 4

$$\|A\|_{tr} = \max_B \frac{|\mathrm{Tr}(A^*B)|}{\|B\|}.$$

This lemma expresses the fact that the trace norm and spectral norm are dual. For Theorem 1 we need only the following simple bound on the trace norm: if there are k distinct numbers $i_1, \ldots, i_k$ and k distinct numbers $j_1, \ldots, j_k$ such that $A[i_r, j_r] = 1$ for all $1 \leq r \leq k$, then $\|A\|_{tr} \geq k$.

2.2 Formula size and communication complexity

A formula is a binary tree with nodes labeled by AND and OR gates, and leaves labeled by literals, that is either a variable or its negation. The size of a formula is its number of leaves. The formula size of a Boolean function f, written $L(f)$, is the size of a smallest formula which computes f.

Karchmer and Wigderson [KW88] characterize formula size in terms of a communication game. Since this characterization, nearly all formula size lower bounds have been phrased in the language of communication complexity.

Let X, Y, Z be finite sets and $R \subseteq X \times Y \times Z$ a relation. In the communication problem for R, Alice is given some $x \in X$, Bob some $y \in Y$, and they wish to output some $z \in Z$ such that $(x, y, z) \in R$. A communication protocol is a binary tree with each internal node v labeled either by a function $a_v : X \to \{0, 1\}$ if Alice speaks at this node, or by a function $b_v : Y \to \{0, 1\}$ if Bob speaks. Each leaf is labeled by an element $z \in Z$. We say that a protocol P computes a relation R if for every $(x, y) \in X \times Y$, walking down the tree according to the functions a_v, b_v leads to a leaf labeled with z such that $(x, y, z) \in R$. We let $C^P(R)$ denote the number of leaves in a smallest protocol which computes R.

For a Boolean function $f : \{0, 1\}^n \to \{0, 1\}$, let $X = f^{-1}(0)$ and $Y = f^{-1}(1)$. We associate with f a relation $R_f \subseteq X \times Y \times [n]$, where $R_f = \{(x, y, i) : x \in X, y \in Y, x_i \neq y_i\}$.

Theorem 5 (Karchmer–Wigderson) $L(f) = C^P(R_f)$

An important notion in communication complexity is that of a combinatorial rectangle. A combinatorial rectangle of $X \times Y$ is a set which can be expressed as $X' \times Y'$ for some $X' \subseteq X$ and $Y' \subseteq Y$. A set $S \subseteq X \times Y$ is called monochromatic for the relation R if there is some $z \in Z$ such that $(x, y, z) \in R$ for all $(x, y) \in S$. Let $C^D(R)$ be the number of rectangles in a smallest partition of $X \times Y$ into combinatorial rectangles monochromatic for R. We will often refer to this informally as the rectangle bound. A basic fact, which can be found in [KN97], is that $C^D(R) \leq C^P(R)$. The rectangle bound is also somewhat tight—Karchmer, Kushilevitz, and Nisan [KKN95] show that $C^P(R) \leq C^D(R)^{\log C^D(R)}$.

3 Rank technique

One of the best techniques for showing lower bounds on the communication complexity of a function $f : X \times Y \to \{0, 1\}$ is matrix rank, originally used by [MS82]. If M_f is a matrix with rows labeled from X, columns labeled from Y and where $M_f[x, y] = f(x, y)$, then $\mathrm{rk}(M_f)$ lower bounds the number of leaves in a communication protocol for f.

Let X, Y, Z be finite sets and $R \subseteq X \times Y \times Z$ a relation. In order to apply the rank bound, we first restrict the relation to a (non-Boolean) function by means of what we call a selection function. A selection function $S : X \times Y \to Z$ for the relation R takes input (x, y) and outputs some z such that $(x, y, z) \in R$. That is, it simply selects one of the possible valid outputs of the relation on input (x, y). We let $R|_S = \{(x, y, z) : S(x, y) = z\}$.

Theorem 6 $C^P(R) = \min_S C^P(R|_S)$.

Proof: For any selection function S, we have $C^P(R) \leq C^P(R|_S)$, as a protocol for $R|_S$ is in particular a protocol for R.

To see $C^P(R) \geq \min_S C^P(R_S)$, let P be an optimal protocol for R. We define a selection function based on this protocol, that is, let $S(x, y) = z$ if and only if (x, y) lead to a leaf labeled z

by P. Now the protocol P also solves $R|_S$ and the claim follows. □

With the help of selection functions, we can now use rank as in the functional case.

Theorem 7 *Let $R \subseteq X \times Y \times Z$ be a relation. To a selection function S, we associate a set of matrices $\{S_z\}$ over $X \times Y$ where $S_z[x,y] = 1$ if $S(x,y) = z$ and $S_z[x,y] = 0$ otherwise. Then*

$$C^D(R) \geq \min_S \sum_{z \in Z} \mathrm{rk}(S_z).$$

Proof: Let $\mathcal{R}$ be an optimal rectangle partition of R satisfying $|\mathcal{R}| = \mathcal{C}^{\mathcal{D}}(\mathcal{R})$. We let $\mathcal{R}$ define a selection function in the natural way, setting $S(x,y) = z$ where z is the lexicographically least color of the rectangle in $\mathcal{R}$ which contains (x,y).

We now show for this particular choice

$$C^D(R) \geq \sum_{z \in Z} \mathrm{rk}(S_z),$$

which gives the theorem. Clearly $C^D(R)$ is equal to the sum over all z of the number of rectangles labeled z by the partition $\mathcal{R}$. Thus it suffices to show that $\mathrm{rk}(S_z)$ lower bounds the number of rectangles labeled by z. Consider some z and say that there are k monochromatic rectangles $B_1, \ldots, B_k$ labeled z. As each B_i is a combinatorial rectangle we can write it as $B_i = V_i \times W_i$ for $V_i \subseteq X$ and $W_i \subseteq Y$. Let v_i be the characteristic vector of V_i, that is $v_i[x] = 1$ if $x \in V_i$ and $v_i[x] = 0$ otherwise, and similarly for w_i with W_i. Then we can express S_z as $S_z = \sum_{i=1}^k v_i w_i^*$ and so $\mathrm{rk}(S_z) \leq k$. □

In general, this bound seems quite difficult to apply because of the minimization over all selection functions. We will now look at a simplified form of this method where we get around this difficulty by using Lemma 3 to lower bound the rank.

Corollary 8 *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function, and let $X = f^{-1}(0), Y = f^{-1}(1)$. Let c_i be the number of pairs $(x,y) \in X \times Y$ which differ only in position i, and let $s_1, \ldots, s_n$ be n nonnegative integers which sum to $|X||Y|$. Then*

$$C^D(R_f) \geq \min_{\substack{s_i \\ \sum_i s_i = |X||Y|}} \sum_i \left\lceil \frac{c_i^2}{s_i} \right\rceil.$$

Proof: By Theorem 7 and Lemma 3

$$C^D(R_f) \geq \min_S \sum_i \mathrm{rk}(S_i) \geq \min_S \sum_i \left\lceil \frac{\|S_i\|_{tr}^2}{\|S_i\|_F^2} \right\rceil. \tag{1}$$

For the c_i many (x, y) pairs which differ only in position i, any selection function S must choose i. As the string y differing from x only in position i is unique, this means that we can permute the rows and columns of S_i to obtain a matrix with trace at least c_i, and so $\|S_i\|_{tr} \geq c_i$. The Frobenius norm squared of a zero/one matrix is simply the number of ones, thus $\|S_i\|_F^2$ is simply the number of (x, y) pairs for which the selection function S chooses i. As the selection function is total, $\sum_i \|S_i\|_F^2 = |X||Y|$. The claim follows. □

The simplified version of the rank method given in Corollary 8 is already strong enough to imply Khrapchenko's method, which works as follows. Let f be a Boolean function, and as before let $X = f^{-1}(0), Y = f^{-1}(1)$. Let C be the set of $(x, y) \in X \times Y$ which have Hamming distance one. Khrapchenko's bound is then $|C|^2/|X||Y|$.

Theorem 9 *The bound given in Corollary 8 is at least as large as that of Khrapchenko.*

Proof: Let c_i be the number of $(x, y) \in X \times Y$ which differ only in position i, and let $\{s_i\}$ be such that $\sum_i s_i = |X||Y|$ and which minimize the bound given in Corollary 8. We now apply Jensen's inequality, Lemma 2, with $\phi(x) = 1/x, x_i = s_i/c_i$, and $a_i = c_i$ to obtain

$$\sum_i \frac{c_i^2}{s_i} \geq \frac{(\sum_i c_i)^2}{\sum_i s_i} = \frac{|C|^2}{|X||Y|}.$$

□

4 Application to parity

In this section, we look at an application of the rank technique to the parity function. For both the upper and lower bounds, we will use the communication complexity setting of Karchmer and Wigderson. In this setting, Alice is given some x with even parity, Bob some y with odd parity, and they wish to find some i such that $x_i \neq y_i$. We first show the upper bound.

Proposition 10 *Let $n = 2^\ell + k$, where $0 \leq k < 2^\ell$. Then $L(\oplus_n) \leq 2^\ell(2^\ell + 3k)$.*

Proof: The basic idea is binary search. First imagine that n is a power of two. Bob begins by saying the parity of the left half of y. Alice then says the parity of the left half of x. If these parities differ, then they continue playing on the left half, otherwise they continue playing on the right half. With each round they halve the size of the playing field, and use two bits of communication. Thus after $\log n$ rounds and $2 \log n$ bits of communication they determine an i on which x and y differ. This gives a formula of size n^2.

When n is not a power of two, then at some point Alice and Bob will not be able to split the playing field evenly between left and right halves. To govern how Alice and Bob decompose n, consider a binary tree with the following properties:

- The root is labeled by n.
- The label of a node equals the sum of its sons
- Each leaf is labeled by 1.

Any such tree gives a protocol of the above type in the following way:

- Alice and Bob begin at the root, Alice playing with x and Bob with y. If the left son of the root is n_1, then Alice and Bob exchange the parities of the first n_1 bits of x and y respectively. If these disagree, then they continue playing with the substrings consisting of the first n_1 bits of x and y respectively. If these agree then they continue playing on the last $n - n_1$ bits of x and y respectively.
- Say that Alice and Bob have arrived at node v playing with strings x' and y' respectively, and that the left son of v is labeled by n_1. Alice and Bob exchange the parities of the first n_1 bits of x' and y'. If these agree then they continue playing on the last $n - n_1$ bits of x' and y' respectively.

The following claim gives the number of leaves in such a protocol.

Claim 11 *Let T be a binary decomposition of n as above. Then*

$$L(\oplus_n) \leq \sum_{\ell \in T} 2^{\mathrm{depth}(\ell)},$$

where the sum is taken over the leaves ℓ of T.

Proof: We count the number of transcripts. Consider a path from root to a leaf. At each step in this path, there are two messages that could lead to taking that step. Namely, if the step is a left step, then Alice and Bob disagree in parity at this step and thus the message exchange leading to this is either 01 or 10. Similarly, if the step is a right step then Alice and Bob agreed in parity at this step and the messages which could be exchanged are 00 or 11. Thus the total number of transcripts in the parity protocol from a given tree is $\sum_{\ell \in T} 2^{\mathrm{depth}(\ell)}$. □

We use this claim to prove Proposition 10. Consider a binary decomposition of n where the sons of any node labeled by an even number have the same value and the sons of any node labeled by an odd number differ by one. This decomposition will have $2k$ many leaves at depth $\ell + 1$ and $2^\ell - k$ many leaves at depth ℓ. The claim then gives

$$L(\oplus_n) \leq 2k(2^{\ell+1}) + (2^\ell - k)2^\ell = 2^\ell(2^\ell + 3k)$$

□

Proposition 12 *Let $n = 2^\ell + k$, where $0 \leq k < 2^\ell$. Then $L(\oplus_n) \geq 2^\ell(2^\ell + 3k)$.*

Proof: Let S be any selection function. For every i, there are 2^{n-1} entries of the matrix S_i which *must* be one, namely the entries x, y which differ only on position i. If S only assigns these entries to have the label i, then S_i is a permutation matrix and so has rank 2^{n-1}. Thus to reduce the rank of S_i, the selection function S must therefore assign more (x, y) pairs to also have the label i. The catch is that S must do this for all i simultaneously, and we will bound how well it can do this.

Notice that for parity on n-bits, $|X| = |Y| = 2^{n-1}$. For every i there are 2^{n-1} pairs (x, y) which differ only in position i. Thus applying Corollary 8 with $c_i = 2^{n-1}$ for all i, we obtain

$$\mathrm{C}^D(R) \geq \min_{s_i : \sum_i s_i = 2^{2n-2}} \sum_{i=1}^{n} \left\lceil \frac{2^{2n-2}}{s_i} \right\rceil . \tag{2}$$

Notice that if we were to ignore the ceilings, then we are minimizing over a convex function $\phi(x) = 1/x$ and so Jensen's inequality gives that the minimum is obtained when all s_i are equal. In this case $s_i = 2^{2n-2}/n$ and so $\sum_i 2^{2n-2}/s_i = n^2$.

To get bound larger than n^2 we need to take the ceiling functions into account. If n is not a power of two, then $2^{2n-2}/n$ will not be an integer, whereas each s_i is an integer—this means that it is no longer possible to have all s_i values equal and $\sum_i s_i = 2^{2n-2}$. It is this imbalance that will lead to a larger lower bound.

We transform (Equation (2)) in a series of steps. First, notice that

$$\min_{s_i : \sum_i s_i = 2^{2n-2}} \sum_{i=1}^{n} \left\lceil \frac{2^{2n-2}}{s_i} \right\rceil = \min_{s'_i : \sum_i s'_i \leq 2^{2n-2}} \sum_{i=1}^{n} \left\lceil \frac{2^{2n-2}}{s'_i} \right\rceil . \tag{3}$$

The right hand side is clearly less than the left hand side as the minimization is taken over a larger set. The left hand side is less than the right hand side as given a solution $\{s'_i\}$ to the right hand side, we can obtain a solution to the left hand side which is not larger by setting $s_i = s'_i$ for $i = 1, \ldots, n-1$, and $s_n = 2^{2n-2} - \sum_{i=1}^{n-1} s'_i \geq s'_n$.

Now we observe that there is an optimal solution $\{s_i\}$ to (eqnrefmin2) where each $2^{2n-2}/s_i$ is an integer, and so each s_i is a power of two. If $2^{2n-2}/s_i$ is not an integer, then we can set s'_i to the largest power of two less than s_i and $\lceil 2^{2n-2}/s_i \rceil = 2^{2n-2}/s'_i$, and the sum of s'_i does not increase.

Thus assume that each s_i is a power of two, say $s_i = 2^{a_i}$. We can now rewrite (Equation (3)) as

$$\min_{\substack{a_i \\ \sum_i 2^{a_i} \leq 2^{2n-2}}} \sum_i 2^{2n-2-a_i}$$

The values $\{a_i\}$ which achieve this minimum will maximize

$$\max_{\substack{a_i \\ \sum_i 2^{a_i} \leq 2^{2n-2}}} \sum_i a_i .$$

We now show that there is an optimal solution to this maximization problem where $|a_i - a_j| \leq 1$ for all i, j. If $a_i - a_j > 2$ then we can let $a'_i = a_i - 1$ and $a'_j = a_j + 2$, so that $a'_i + a'_j > a_i + a_j$ and $2^{a'_j} \leq 2^{a_j} + 2^{a_i - 1}$ so $2^{a'_i} + 2^{a'_j} \leq 2^{a_i} + 2^{a_j}$. If $a_i - a_j = 2$ then by setting $a'_i = a_i - 1$ and $a'_j = a_j + 1$ then we still have $a'_i + a'_j = a_i + a_j$, and have saved on weight, $2^{a'_i} + 2^{a'_j} < 2^{a_i} + 2^{a_j}$.

By performing these transformations, we can turn any solution into one where $|a_i - a_j| \leq 1$ and whose value is at least as good. Now if we have $|a_i - a_j| \leq 1$ and $\sum_i 2^{a_i} = 2^{2n-2}$, it follows that $a_i = 2n - \ell - 2$ for $2^\ell - k$ many values of i and $a_i = 2n - \ell - 3$ for $2k$ many values of i. This gives

$$\begin{aligned} \min_{a_i} \sum_{i=1}^{n} 2^{2n-2-a_i} &= (2^\ell - k)2^\ell + 2k2^{\ell+1} \\ &= 2^\ell(2^\ell + 3k). \end{aligned}$$

□

5 Hierarchy of techniques

In this section, we present two results clarifying the hierarchy of available techniques for proving lower bounds on formula size. Laplante, Lee, and Szegedy [LLS06] show that the quantum adversary method gives bounds at least as large as the method of Koutsoupias [Kou93] which is in turn at least as large as the bound of Khrapchenko. Here we show that the linear programming bound of Karchmer, Kushilevitz, and Nisan [KKN95] and a slight variation of our bound, as presented in (Equation (1)), are both always at least as large as the quantum adversary method.

We first describe the methods in question. Karchmer, Kushilevitz, and Nisan notice that for a relation $R \subseteq X \times Y \times Z$ the rectangle bound $C^D(R)$ can be written as an integer program. Indeed, let $\mathcal{R}$ be the set of all rectangles which are monochromatic with respect to the relation R. To represent the relationship between inputs (x, y) and the rectangles of $\mathcal{R}$ we use a $|X| \cdot |Y|$-by-$|\mathcal{R}|$ incidence matrix A, where for $(x, y) \in X \times Y$ and $S \in \mathcal{R}$ we let $A[(x, y), S] = 1$ if $(x, y) \in S$. Now a set of rectangles can be described by a $|\mathcal{R}|$-length vector α, with each entry $\alpha[S] \in \{0, 1\}$. If α represents a partition, then $A\alpha = \vec{1}$, and the number of rectangles in such a partition is simply $\sum_S \alpha[S]$. Karchmer, Kushilevitz, and Nisan relax this integer program to a linear program by replacing the condition $\alpha[S] \in \{0, 1\}$ with $0 \leq \alpha[S] \leq 1$.

Definition 13 (Linear programming bound [KKN95]) *Let $f : \{0,1\}^n \rightarrow \{0,1\}$ be a Boolean function, R_f the relation corresponding to f, and α a vector indexed by rectangles monochromatic with respect to R_f. The linear programming bound, denoted* $\mathrm{LP}(f)$*, is then*

$$\mathrm{LP}(f) = \min_{\substack{\alpha : A\alpha = \vec{1} \\ 0 \leq \alpha[S] \leq 1}} \sum_S \alpha[S].$$

Ambainis [Amb02, Amb03] developed the quantum adversary method to prove lower bounds on bounded-error quantum query complexity. Laplante, Lee, and Szegedy show that the square of the adversary bound is lower bound on formula size. The adversary bound can be phrased as a maximization problem of the spectral norm of a matrix associated with f [BSS03].

Definition 14 (Adversary bound) *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function, and $X = f^{-1}(0)$ and $Y = f^{-1}(1)$. Let Γ be a $|X|$-by-$|Y|$ matrix, and let Γ_i be the matrix such that $\Gamma_i[x,y] = \Gamma[x,y]$ if $x_i \neq y_i$ and $\Gamma_i[x,y] = 0$ otherwise, for $1 \leq i \leq n$. Then*

$$\mathrm{ADV}(f) = \max_{\substack{\Gamma \geq 0 \\ \Gamma \neq 0}} \frac{||\Gamma||}{\max_i ||\Gamma_i||}.$$

First we show that a slightly more sophisticated version of our bound (Equation (1)) is always at least as large as the quantum adversary method. A problem with (Equation (1)) is that it cannot take advantage of the fact that certain inputs to a function might be harder than others. To give a concrete example, the bound given by (Equation (1)) on the function $f : \{0,1\}^{2n} \to \{0,1\}$ on $2n$ bits which is just the parity of the first n bits is worse than the bound for parity on n bits. To remedy this, we let u be a unit vector of length $|X|$ and v be a unit vector of length $|Y|$ and consider the matrix $S_i \circ uv^*$ instead of the matrix S_i. As $\mathrm{rk}(S_i \circ uv^*) \leq \mathrm{rk}(S_i)$, we can again apply Theorem 7 and Lemma 3 to obtain

$$C^D(R_f) \geq \min_S \max_{\substack{u,v \\ ||u||=||v||=1}} \sum_i \frac{||S_i \circ uv^*||_{tr}^2}{||S_i \circ uv^*||_F^2}. \tag{4}$$

Theorem 15 *The bound given by (Equation (4)) is at least as large as $\mathrm{ADV}(f)^2$.*

Proof: Starting from (Equation (4)) we first apply Jensen's inequality with $\phi(x) = 1/x$, $x_i = ||S_i \circ uv^*||_F^2 / ||S_i \circ uv^*||_{tr}$, and $a_i = ||S_i \circ uv^*||_{tr}$ to obtain:

$$\min_S \max_{\substack{u,v \\ ||u||=||v||=1}} \sum_i \frac{||S_i \circ uv^*||_{tr}^2}{||S_i \circ uv^*||_F^2} \geq \min_S \max_{\substack{u,v \\ ||u||=||v||=1}} \frac{\left(\sum_i ||S_i \circ uv^*||_{tr}\right)^2}{\sum_i ||S_i \circ uv^*||_F^2}.$$

As the selection function is total we have $\sum_i ||S_i \circ uv^*||_F^2 = ||uv^*||_F^2 = 1$.

Now we use (Lemma 4) to lower bound $||S_i||_{tr}$. One can think of the weight matrix Γ in the adversary bound as the matrix from (Lemma 4) which witnesses that the trace norm of the S_i's is large:

$$\begin{aligned} \min_S \max_{\substack{u,v \\ ||u||=||v||=1}} \left(\sum_i ||S_i \circ uv^*||_{tr} \right)^2 &\geq \min_S \max_{\substack{\Gamma \geq 0 \\ \Gamma \neq 0}} \max_{\substack{u,v \\ ||u||=||v||=1}} \left(\sum_i \frac{|\mathrm{Tr}((\Gamma \circ S_i)vu^*|}{||\Gamma \circ S_i||} \right)^2 \\ &\geq \min_S \max_{\substack{\Gamma \geq 0 \\ \Gamma \neq 0}} \max_{\substack{u,v \\ ||u||=||v||=1}} \left(\sum_i \frac{|\mathrm{Tr}((\Gamma \circ S_i)vu^*|}{||\Gamma_i||} \right)^2. \end{aligned}$$

This step follows as $0 \leq \Gamma \circ S_i \leq \Gamma_i$ and for matrices A, B if $0 \leq A \leq B$ then $||A|| \leq ||B||$. We can now continue

$$\begin{aligned}
\min_S \max_{\substack{\Gamma\geq 0\\ \Gamma\neq 0}} \max_{\substack{u,v\\ \|u\|=\|v\|=1}} \left(\sum_i \frac{|\mathrm{Tr}((\Gamma\circ S_i)vu^*|}{\|\Gamma_i\|}\right)^2 &\geq \min_S \max_{\substack{\Gamma\geq 0\\ \Gamma\neq 0}} \max_{\substack{u,v\\ \|u\|=\|v\|=1}} \left(\frac{\sum_i \mathrm{Tr}((\Gamma\circ S_i)vu^*)}{\max_i \|\Gamma_i\|}\right)^2\\
&= \max_{\substack{\Gamma\geq 0\\ \Gamma\neq 0}} \max_{\substack{u,v\\ \|u\|=\|v\|=1}} \left(\frac{\mathrm{Tr}(\Gamma vu^*)}{\max_i \|\Gamma_i\|}\right)^2\\
&= \max_{\substack{\Gamma\geq 0\\ \Gamma\neq 0}} \left(\frac{\|\Gamma\|}{\max_i \|\Gamma_i\|}\right)^2.
\end{aligned}$$

□

Now we show that the bound given by the linear programming method is also always at least as large as that given by the adversary method.

Theorem 16 $\mathrm{LP}(f) \geq \mathrm{ADV}^2(f)$.

Proof: Let α be a solution to the linear program associated with f. By definition we have $\sum_{S:(x,y)\in S} \alpha[S] = 1$ for every (x,y). Let u,v be unit vectors such that $|u^*\Gamma v| = \|\Gamma\|$. We will need some notation to label submatrices of Γ and portions of u,v. For a combinatorial rectangle $S = U\times V$, let $\Gamma_S[x,y] = A[x,y]$ if $(x,y)\in S$ and $\Gamma[x,y]=0$ otherwise. Similarly, let $u_S[x] = u[x]$ if $x\in U$ and $u_S[x]=0$ otherwise, and similarly for v_S. Now

$$\begin{aligned}
\|\Gamma\| &= \sum_{x,y} \Gamma[x,y]u[x]v[y]\\
&= \sum_{x,y}\sum_{S:(x,y)\in S} \alpha[S]\Gamma[x,y]u[x]v[y]\\
&= \sum_S \alpha[S] \sum_{(x,y)\in S} \Gamma[x,y]u[x]v[y]\\
&\leq \sum_S \alpha[S]\|\Gamma_S\|\|u_S\|\|v_S\|\\
&\leq \left(\sum_S \alpha[S]\|\Gamma_S\|^2\right)^{1/2} \left(\sum_S \alpha[S]\|u_S\|\|v_S\|\right)^{1/2},
\end{aligned}$$

where the first inequality follows from the definition of spectral norm, and the second uses the Cauchy–Schwarz inequality. Notice that

$$\sum_S \alpha[S]\|\Gamma_S\|^2 = \sum_{x,y} \alpha[S]|u[x]|^2|v[y]|^2 = 1.$$

Thus

$$\|\Gamma\|^2 \leq \sum_S \alpha[S]\|\Gamma_S\|^2 \leq \max_S \|\Gamma_S\|^2 \sum_S \alpha[S],$$

and so

$$\sum_S \alpha[S] \geq \max_{\Gamma} \frac{||\Gamma||^2}{\max_S ||\Gamma_S||^2} \geq \max_{\Gamma} \frac{||\Gamma||^2}{\max_i ||\Gamma_i||^2}.$$

□

Acknowledgments

I would like to thank Anna Gál for helpful discussions on the topics of this paper, and the anonymous referees for many beneficial comments.

References

[Amb02] A. Ambainis. Quantum lower bounds by quantum arguments. *Journal of Computer and System Sciences*, 64:750–767, 2002.

[Amb03] A. Ambainis. Polynomial degree vs. quantum query complexity. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science*, pages 230–239. IEEE, 2003.

[BSS03] H. Barnum, M. Saks, and M. Szegedy. Quantum decision trees and semidefinite programming. In *Proceedings of the 18th IEEE Conference on Computational Complexity*, pages 179–193, 2003.

[Hås98] J. Håstad. The shrinkage exponent is 2. *SIAM Journal on Computing*, 27:48–64, 1998.

[Khr71] V.M. Khrapchenko. Complexity of the realization of a linear function in the case of Π-circuits. *Math. Notes Acad. Sciences*, 9:21–23, 1971.

[KKN95] M. Karchmer, E. Kushilevitz, and N. Nisan. Fractional covers and communication complexity. *SIAM Journal on Discrete Mathematics*, 8(1):76–92, 1995.

[KN97] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[Kou93] E. Koutsoupias. Improvements on Khrapchenko's theorem. *Theoretical Computer Science*, 116(2):399–403, 1993.

[KW88] M. Karchmer and A. Wigderson. Monotone connectivity circuits require superlogarithmic depth. In *Proceedings of the 20th ACM Symposium on the Theory of Computing*, pages 539–550, 1988.

[LLS06] S. Laplante, T. Lee, and M. Szegedy. The quantum adversary method and classical formula size lower bounds. *Computational Complexity*, 15:163–196, 2006.

[MS82] K. Melhorn and E. Schmidt. Las Vegas is better than determinism in VLSI and distributed computing. In *Proceedings of the 14th ACM Symposium on the Theory of Computing*, pages 330–337. ACM, 1982.

[Neč66] E. I. Nečiporuk. A Boolean function. *Soviet Mathematics–Doklady*, 7:999–1000, 1966.

[PPZ92] M. Paterson, N. Pippenger, and U. Zwick. Optimal carry save networks. In *Boolean function complexity*, pages 174–201. London Mathematical Society Lecture Note Series 169, Cambridge University Press, 1992.

[Rad97] J. Radhakrishnan. Better lower bounds for monotone threshold formulas. *Journal of Computer and System Sciences*, 54(2):221–226, 1997.

[Raz90] A. Razborov. Applications of matrix methods to the theory of lower bounds in computational complexity. *Combinatorica*, 10(1):81–93, 1990.

[Raz92] A. Razborov. On submodular complexity measures. In M. Paterson, editor, *Boolean function complexity*, volume 169 of *London Math. Soc. Lecture Notes Series*, pages 76–83. Cambridge University Press, 1992.

[Val84] L.G. Valiant. Short monotone formulae for the majority function. *Journal of Algorithms*, 5:363–366, 1984.