The quantum adversary method and classical formula size lower bounds

Sophie Laplante
LRI, Université Paris-Sud
laplante@lri.fr

Troy Lee
CWI and University of Amsterdam
Troy.Lee@cwi.nl

Mario Szegedy
Rutgers University
szegedy@cs.rutgers.edu

September 9, 2005

Abstract

We introduce two new complexity measures for Boolean functions, which we name sumPI and maxPI. The quantity sumPI has been emerging through a line of research on quantum query complexity lower bounds via the so-called quantum adversary method [Amb02, Amb03, BSS03, Zha05, LM04], culminating in [ŠS05] with the realization that these many different formulations are in fact equivalent. Given that sumPI turns out to be such a robust invariant of a function, we begin to investigate this quantity in its own right and see that it also has applications to classical complexity theory.

As a surprising application we show that $\mathsf{sumPI}^2(f)$ is a lower bound on the formula size, and even, up to a constant multiplicative factor, the probabilistic formula size of f. We show that several formula size lower bounds in the literature, specifically Khrapchenko and its extensions [Khr71, Kou93], including a key lemma of [Hås98], are in fact special cases of our method. The second quantity we introduce, $\mathsf{maxPI}(f)$, is always at least as large as $\mathsf{sumPI}(f)$, and is derived from sumPI in such a way that $\mathsf{maxPI}^2(f)$ remains a lower bound on formula size.

Our main result is proven via a combinatorial lemma which relates the square of the spectral norm of a matrix to the squares of the spectral norms of its submatrices. The generality of this lemma implies that our methods can also be used to lower bound the communication complexity of relations, and a related combinatorial quantity, the rectangle partition number.

To exhibit the strengths and weaknesses of our methods, we look at the sumPI and maxPI complexity of a few examples, including the recursive majority of three function, a function defined by Ambainis [Amb03], and the collision problem.

1 Introduction

A central and longstanding open problem in complexity theory is to prove superlinear lower bounds for the circuit size of an explicit Boolean function. While this seems quite difficult, a modest amount of success has been achieved in the slightly weaker model of formula size, a formula being simply a circuit where every gate has fan-out at most one. The current best formula size lower bound for an explicit function is $n^{3-o(1)}$ by Håstad [Hås98].

In this paper we show that part of the rich theory developed around proving lower bounds on quantum query complexity, namely the so-called quantum adversary argument, can be brought to bear on formula size lower bounds. This adds to the growing list of examples of how studying quantum computing has led to new results in classical complexity, including [SV01, KW03, Aar04, LM04], to cite a few.

The roots of the quantum adversary argument can be traced to the hybrid argument of [BBBV97], who use it to show a $\Omega(\sqrt{n})$ lower bound on quantum search. Ambainis developed a more sophisticated adversary argument [Amb02] and later improved this method to the full-strength quantum adversary argument [Amb03]. Further generalizations include Barnum, Saks, and Szegedy [BSS03] with their spectral method and Zhang [Zha05] with his strong adversary method. Laplante and Magniez [LM04] use Kolmogorov complexity to capture the adversary argument in terms of a minimization problem. This line of research culminates in recent work of Špalek and Szegedy [ŠS05] who show that in fact all the methods of [Amb03, BSS03, Zha05, LM04] are equivalent.

The fact that the quantum adversary argument has so many equivalent definitions indicates that it is a natural combinatorial property of Boolean functions which is worthwhile to investigate on its own. We give this quantity its own name, sumPI, and adopt the following primal formulation of the method, from [ŠS05, LM04]. Letting $S \subseteq \{0,1\}^n$ and $f : S \to \{0,1\}$, be a Boolean function we say

$$\mathsf{sumPI}(f) = \min_{p} \max_{\substack{x,y \\ f(x)\neq f(y)}} \frac{1}{\sum_{\substack{i \\ x_i \neq y_i}} \sqrt{p_x(i)p_y(i)}}, \tag{1}$$

where $p = \{p_x : x \in S\}$ is a family of probability distributions on the indices $[n]$. If $Q_\epsilon(f)$ is the two sided error quantum query complexity of f then $Q_\epsilon(f) = \Omega(\mathsf{sumPI}(f))$. We show further that $\mathsf{sumPI}^2(f)$ is a lower bound on the formula size of f. Moreover, $\mathsf{sumPI}^2(f)$ generalizes several formula size lower bounds in the literature, specifically Khrapchenko and its extensions [Khr71, Kou93], and a key lemma of [Hås98] used on the way to proving the current best formula size lower bounds for an explicit function.

We also introduce

$$\mathsf{KI}(f) = \min_{\alpha \in \Sigma^*} \max_{\substack{x,y \\ f(x)\neq f(y)}} \min_{i : x_i \neq y_i} K(i|x,\alpha) + K(i|y,\alpha),$$

where K is prefix-free Kolmogorov complexity. This formulation arises from the quantum and randomized lower bounds of [LM04]. This formulation is especially interesting because of the intuition that it provides. For example, it allows for a very simple proof that circuit depth $\mathsf{d}(f) \geq \mathsf{KI}(f)$, using the Karchmer-Wigderson characterization of circuit depth [KW88].

We define a quantity closely related to 2^{KI}, which we call maxPI.

$$\mathsf{maxPI}(f) = \min_{p} \max_{\substack{x,y \\ f(x)\neq f(y)}} \frac{1}{\max_{i : x_i \neq y_i} \sqrt{p_x(i)p_y(i)}}. \tag{2}$$

Notice that this is like sumPI but where the sum is replaced by a maximum. By definition, maxPI is larger than sumPI, but its square is still a lower bound on formula size.

We prove our main results by transforming in two steps the problem of proving formula size lower bounds into a problem with a more combinatorial flavor which is easier to work with. First, we use the elegant characterization given by Karchmer and Wigderson [KW88] of formula size in terms of the communication complexity of a relation. We then use the well-known property that a

successful communication protocol partitions a relation into rectangles of a certain form. We then lower bound the size of the minimal such rectangle partition. A sufficient condition for a measure to lower bound the size of such a partition is that it is subadditive on disjoint rectangles. Our main lemma shows that the spectral norm squared of a matrix A is less than the sum of the squared spectral norms of matrices A_R which partition A.

We look at several concrete problems to illustrate the strengths and weaknesses of our methods. We study the height h recursive majority of three problem, $\mathsf{R{-}MAJ}_3^h$, and show that $Q_\epsilon(\mathsf{R{-}MAJ}_3^h) = \Omega(2^h)$ and a lower bound of 4^h for the formula size. We also look at a function defined by Ambainis [Amb03] to separate the quantum query complexity of a function from the bound given by the polynomial method [BBC+01]. This function gives an example where sumPI^2 can give something much better than Khraphchenko's bound. For total functions, maxPI and sumPI are polynomially related; however, we give an example of a partial function f, namely the collision problem, where $\mathsf{sumPI}(f) = 2$ and $\mathsf{maxPI}(f) = \Theta(\sqrt{n})$. This example shows that in general maxPI is not a lower bound on quantum query complexity as for the collision problem $\mathsf{maxPI}(f) \gg Q_\epsilon(f) = \Theta(n^{1/3})$ [AS04, BHT97].

1.1 Organization

In Section 2, we give the definitions, results, and notation that we use throughout the paper, and introduce the quantities sumPI, maxPI, and KI. In Section 3 we prove some properties of sumPI and maxPI. In Section 4, we show how sumPI and maxPI give rise to formula size lower bounds, for deterministic and probabilistic formula size. In Section 5, we compare our new methods with previous methods in formula size complexity. In Section 6, we investigate the limits of our and other formula lower bound methods. Finally, in Section 7 we apply our techniques to some concrete problems.

2 Preliminaries

We use standard notation such as $[n] = \{1, \ldots, n\}$, $|S|$ for the cardinality of a set S, and all logarithms are base 2. Hamming distance is written d_H.

2.1 Complexity measures of Boolean functions

We use standard measures of Boolean functions, such as sensitivity and certificate complexity. We briefly recall these here, see [BW02] for more details. For a set $S \subseteq \{0,1\}^n$ and Boolean function $f : S \to \{0,1\}$, the sensitivity of f on input x is the number of positions $i \in [n]$ such that changing the value of x in position i changes the function value. The zero-sensitivity, written $s_0(f)$ is the maximum over $x \in f^{-1}(0)$ of the sensitivity of f on x. The one-sensitivity, $s_1(f)$ is defined analogously. The maximum of $s_0(f), s_1(f)$ is the sensitivity of f, written $s(f)$.

A certificate for f on input $x \in S$ is a subset $I \subseteq [n]$ such that for any y satisfying $y_i = x_i$ for all $i \in I$ it must be the case that $f(y) = f(x)$. The zero-certificate complexity of f, written $C_0(f)$ is the maximum over all $x \in f^{-1}(0)$ of the minimum size certificate of x. Similarly, the one-certificate complexity of f, written $C_1(f)$ is the maximum over all $x \in f^{-1}(1)$ of the minimum size certificate of x.

2.2 Linear algebra

For a matrix A (respectively, vector v) we write A^T (resp. v^T) for the transpose of A, and A^* (resp. v^*) for the conjugate transpose of A. For two matrices A, B we let $A \circ B$ be the Hadamard product of A and B, that is $(A \circ B)[x, y] = A[x, y]B[x, y]$. We write $A \geq B$ if A is entrywise greater than B, and $A \succeq B$ when $A - B$ is positive semidefinite, that is $\forall v : v^T(A - B)v \geq 0$. We let $\mathrm{rk}(A)$ denote the rank of the matrix A. We will use the notation $\mathrm{Entrysum}(A)$ for $\sum_{i,j} A[i, j]$.

We will make extensive use of the spectral norm, denoted $\|A\|_2$. For a matrix A,

$$\|A\|_2 = \{\sqrt{\lambda} : \lambda \text{ is the largest eigenvalue of } A^*A\}.$$

For a vector v, we let $|v|$ be the ℓ_2 norm of v.

We will also make use of some other matrix norms. The maximum absolute column sum norm, written $\|A\|_1$ is defined as $\|A\|_1 = \max_j \sum_i |A[i, j]|$, and the maximum absolute row sum norm, written $\|A\|_\infty$ is $\|A\|_\infty = \max_i \sum_j |A[i, j]|$. The Fróbenius norm $\|A\|_F = \sqrt{\sum_{i,j} A[i, j]^2}$ is the ℓ_2 norm of A thought of as a long vector.

We collect a few facts about the spectral norm. These can be found in for example [HJ99].

Proposition 1 *Let A be an arbitrary m by n matrix. Then*

1. $\|A\|_2 = \max_{u,v} \frac{|u^*Av|}{|u||v|}$
2. $\|A\|_2^2 \leq \|A\|_1 \|A\|_\infty$
3. *For nonnegative matrices A, B, if $A \leq B$ then $\|A\|_2 \leq \|B\|_2$*

2.3 Deterministic and probabilistic formulae

A Boolean formula over the standard basis $\{\vee, \wedge, \neg\}$ is a binary tree where each internal node is labeled with $\vee$ or $\wedge$, and each leaf is labeled with a literal, that is, a Boolean variable or its negation. The size of a formula is its number of leaves. We naturally identify a formula with the function it computes.

Definition 2 *Let $f : \{0, 1\}^n \to \{0, 1\}$ be a Boolean function. The formula size of f, denoted $\mathsf{L}(f)$, is the size of the smallest formula which computes f. The formula depth of f, denoted $\mathsf{d}(f)$ is the minimum depth of a formula computing f.*

It is clear that $\mathsf{L}(f) \leq 2^{\mathsf{d}(f)}$; that in fact the opposite inequality $\mathsf{d}(f) \leq O(\log \mathsf{L}(f))$ also holds is a nontrivial result due to Spira [Spi71].

We will also consider probabilistic formulae, that is, a probability distribution over deterministic formulae. We take a worst-case notion of the size of a probabilistic formula. Probabilistic formula size has been studied before, for example in [Val84, Bop89, DZ97, Kla04].

Definition 3 *Let $\{f_j\}_{j \in J}$ be a set of functions with $f_j : S \to \{0, 1\}$ for each $j \in J$. For a function $f : S \to \{0, 1\}$, we say that f is ϵ-approximated by $\{f_j\}_{j \in J}$ if there is a probability distribution $\alpha = \{\alpha_j\}_{j \in J}$ over J such that for every $x \in S$,*

$$\Pr_\alpha[f(x) = f_j(x)] \geq 1 - \epsilon.$$

In particular, if $\max_j \mathsf{L}(f_j) \leq s$, then we say that f is ϵ-approximated by formulas of size s, denoted $\mathsf{L}^\epsilon(f) \leq s$.

Note that even if a function depends on all its variables, it is possible that the probabilistic formula size is less than the number of variables.

2.4 Communication complexity of relations

Karchmer and Wigderson [KW88] give an elegant characterization of formula size in terms of a communication game. We will use this formulation in our proofs. This has the advantage of letting us work in the more general setting of communication complexity of relations and enabling us to use the combinatorial tools of communication complexity. We now describe the setting.

Let X, Y, Z be finite sets, and $R \subseteq X \times Y \times Z$. In the communication game for R, Alice is given some $x \in X$, Bob is given some $y \in Y$ and their goal is to find some $z \in Z$ such that $(x, y, z) \in R$, if such a z exists. A communication protocol is a binary tree where each internal node v is labelled by a either a function $a_v : X \to \{0, 1\}$ or $b_v : Y \to \{0, 1\}$ describing either Alice's or Bob's message at that node, and where each leaf is labelled with an element $z \in Z$. A communication protocol computes R if for all $(x, y) \in X \times Y$ walking down the tree according to a_v, b_v leads to a leaf labelled with z such that $(x, y, z) \in R$, provided such a z exists. The communication cost $\mathsf{D}(R)$ of R is the height of the smallest communication protocol computing R. The protocol partition number $C^P(R)$ is the number of leaves in the smallest communication protocol computing R.

Definition 4 *For any Boolean function f we associate a relation $R_f = \{(x, y, i) : f(x) = 0, f(y) = 1, x_i \neq y_i\}$.*

Theorem 5 (Karchmer-Wigderson) *For any Boolean function f, $\mathsf{L}(f) = C^P(R_f)$ and $\mathsf{d}(f) = \mathsf{D}(R_f)$.*

An advantage of the communication complexity approach to formula size is that we can use the powerful combinatorial tools available for communication complexity lower bounds. At the heart of this approach lies the idea of combinatorial rectangles. A combinatorial rectangle is simply a set $S \subseteq X \times Y$ which can be expressed as $S = X' \times Y'$ for some $X' \subseteq X, Y' \subseteq Y$. We say that a set $S \subseteq X \times Y$ is monochromatic with respect to the relation R if there is a $z \in Z$ such that $(x, y, z) \in R$ for all $(x, y) \in S$. It can be shown that the leaves of a successful communication protocol for R form a disjoint covering of $X \times Y$ by rectangles monochromatic with respect to R. We let $C^D(R)$ be the size of the smallest disjoint covering of $X \times Y$ by monochromatic rectangles. It follows that $C^D(R) \leq C^P(R)$. For more information on communication complexity and proofs of the above results, we suggest [KN97].

2.5 sumPI and the quantum adversary method

Knowledge of quantum computing is not needed for reading this paper; for completeness, however, we briefly sketch the quantum query model. More background on quantum query complexity and quantum computing in general can be found in [BW02, NC00].

As with the classical counterpart, in the quantum query model we wish to compute some function $f : S \to \{0, 1\}$, where $S \subseteq \Sigma^n$, and we access the input through queries. The complexity of f is the number of queries needed to compute f. Unlike the classical case, however, we can now make queries in superposition. Formally, a query O corresponds to the unitary transformation

$$O : |i, b, z\rangle \mapsto |i, b \oplus x_i, z\rangle$$

where $i \in [n], b \in \{0, 1\}$, and z represents the workspace. A t-query quantum algorithm A has the form $A = U_t O U_{t-1} O \cdots O U_1 O U_0$, where the U_k are fixed unitary transformations independent of

the input x. The computation begins in the state $|0\rangle$, and the result of the computation A is the observation of the rightmost bit of $A|0\rangle$. We say that A ϵ-approximates f if the observation of the rightmost bit of $A|O\rangle$ is equal to $f(x)$ with probability at least $1-\epsilon$, for every x. We denote by $Q_\epsilon(f)$ the minimum query complexity of a quantum query algorithm which ϵ-approximates f.

Along with the polynomial method [BBC+01], one of the main techniques for showing lower bounds in quantum query complexity is the quantum adversary method [Amb02, Amb03, BSS03, Zha05, LM04]. Recently, Špalek and Szegedy [ŠS05] have shown that all the strong versions of the quantum adversary method are equivalent, and further that these methods can be nicely characterized as primal and dual.

We give the primal characterization as our principal definition of sumPI.

Definition 6 (sumPI) *Let $S \subseteq \{0,1\}^n$ and $f : S \to \{0,1\}$ be a Boolean function. For every $x \in S$ let $p_x : [n] \to \mathbb{R}$ be a probability distribution, that is, $p_x(i) \geq 0$ and $\sum_i p_x(i) = 1$. Let $p = \{p_x : x \in S\}$. We define the sum probability of indices to be*

$$\mathsf{sumPI}(f) = \min_p \max_{\substack{x,y \\ f(x)\neq f(y)}} \frac{1}{\sum_{\substack{i \\ x_i \neq y_i}} \sqrt{p_x(i)p_y(i)}}$$

We will also use two versions of the dual method, both a weight scheme and the spectral formulation. The most convenient weight scheme for us is the "probability scheme", given in Lemma 4 of [LM04].

Definition 7 (Probability Scheme) *Let $S \subseteq \{0,1\}^n$ and $f : S \to \{0,1\}$ be a Boolean function, and $X = f^{-1}(0), Y = f^{-1}(1)$. Let q be a probability distribution on $X \times Y$, and p_A, p_B be probability distributions on X, Y respectively. Finally let $\{p'_{x,i} : x \in X, i \in [n]\}$ and $\{p'_{y,i} : y \in Y, i \in [n]\}$ be families of probability distributions on X, Y respectively. Assume that $q(x,y) = 0$ when $f(x) = f(y)$. Let P range over all possible tuples $(q, p_A, p_B, \{p'_{x,i}\}_{x,i})$ of distributions as above. Then*

$$\mathsf{PA}(f) = \max_P \min_{\substack{x,y,i \\ f(x)\neq f(y), x_i \neq y_i}} \frac{\sqrt{p_A(x)p_B(y)p'_{x,i}(y)p'_{y,i}(x)}}{q(x,y)}$$

We will also use the spectral adversary method.

Definition 8 (Spectral Adversary) *Let $S \subseteq \{0,1\}^n$ and $f : S \to \{0,1\}$ be a Boolean function. Let $X = f^{-1}(0), Y = f^{-1}(1)$. Let $\Gamma \neq 0$ be an arbitrary $|X| \times |Y|$ nonnegative matrix. For $i \in [n]$, let Γ_i be the matrix:*

$$\Gamma_i[x,y] = \begin{cases} 0 & \text{if } x_i = y_i \\ \Gamma[x,y] & \text{if } x_i \neq y_i \end{cases}$$

Then

$$\mathsf{SA}(f) = \max_\Gamma \frac{\|\Gamma\|_2}{\max_i \|\Gamma_i\|_2}$$

Note that the spectral adversary method was initially defined [BSS03] for symmetric matrices over $X \cup Y$. The above definition is equivalent: if A is a symmetric matrix over $X \cup Y$ satisfying the constraint $A[x,y] = 0$ when $f(x) = f(y)$, then A is of the form $A = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix}$, for some matrix

B over $X \times Y$. Then the spectral norm of A is equal to that of B. Similarly, for any $X \times Y$ matrix A we can form a symmetrized version of A as above preserving the spectral norm.

We will often use the following theorem implicitly in taking the method most convenient for the particular bound we wish to demonstrate.

Theorem 9 (Špalek-Szegedy) *Let $n \geq 1$ be an integer, $S \subseteq \{0,1\}^n$ and $f : S \to \{0,1\}$. Then*

$$\mathsf{sumPI}(f) = \mathsf{SA}(f) = \mathsf{PA}(f)$$

2.6 The KI and maxPI complexity measures

The definition of KI arises from the Kolmogorov complexity adversary method [LM04]. The Kolmogorov complexity $C_U(x)$ of a string x, with respect to a universal Turing machine U is the length of the shortest program p such that $U(p) = x$. The complexity of x given y, denoted $C(x|y)$ is the length of the shortest program p such that $U(\langle p, y \rangle) = x$. When U is such that the set of outputs is prefix-free (no string in the set is prefix of another in the set), we write $K_U(x|y)$. From this point onwards, we fix U and simply write $K(x|y)$. For more background on Kolmogorov complexity consult [LV97].

Definition 10 *Let $S \subseteq \{0,1\}^n$ and $f : S \to \{0,1\}$, let*

$$\mathsf{KI}(f) = \min_{\alpha \in \{0,1\}^*} \max_{\substack{x,y \\ f(x) \neq f(y)}} \min_{i : x_i \neq y_i} K(i|x,\alpha) + K(i|y,\alpha).$$

The advantage of using concepts based on Kolmogorov complexity is that they often naturally capture the information theoretic content of lower bounds. As an example of this, we give a simple proof that KI is a lower bound on circuit depth.

Theorem 11 *For any Boolean function f, $\mathsf{KI}(f) \leq \mathsf{d}(f)$.*

Proof: Let P be a protocol for R_f. Fix x, y with different values under f, and let T_A be a transcript of the messages sent from A to B, on input x, y. Similarly, let T_B be a transcript of the messages sent from B to A. Let i be the output of the protocol, with $x_i \neq y_i$. To print i given x, simulate P using x and T_B. To print i given y, simulate P using y and T_A. This shows that $\forall x, y : f(x) \neq f(y), \exists i : x_i \neq y_i, K(i|x,\alpha) + K(i|y,\alpha) \leq |T_A| + |T_B| \leq \mathsf{D}(R_f)$, where α is a description of A's and B's algorithms. □

Remark A similar proof in fact shows that $\mathsf{KI}(f) \leq 2\mathsf{N}(R_f)$, where N is the nondeterministic communication complexity. Since the bound does not take advantage of interaction between the two players, in many cases we cannot hope to get optimal lower bounds using these techniques.

An argument similar to that in [ŠS05] shows that

$$2^{\mathsf{KI}(f)} = \Theta\left(\min_{p} \max_{\substack{x,y \\ f(x) \neq f(y)}} \frac{1}{\max_i \sqrt{p_x(i) p_y(i)}} \right)$$

Notice that the right hand side of the equation is identical to the definition of sumPI, except that the sum in the denominator is replaced by a maximum. This led us to define the complexity measure maxPI, in order to get stronger formula size lower bounds.

Definition 12 (maxPI) *Let $S \subseteq \{0,1\}^n$ and $f : S \to \{0,1\}$. For every $x \in S$ let $p_x : [n] \to \mathbb{R}$ be a probability distribution. Let $p = \{p_x : x \in S\}$. We define the maximum probability of indices to be*

$$\mathsf{maxPI}(f) = \min_{p} \max_{\substack{x,y \\ f(x)\neq f(y)}} \frac{1}{\max_i \sqrt{p_x(i)p_y(i)}}$$

It can be easily seen from the definitions that $\mathsf{sumPI}(f) \leq \mathsf{maxPI}(f)$ for any f. The following lemma is also straightforward from the definitions:

Lemma 13 *If $S' \subseteq S$ and $f' : S' \to \{0,1\}$ is a domain restriction of $f : S \to \{0,1\}$ to S', then $\mathsf{sumPI}(f') \leq \mathsf{sumPI}(f)$ and $\mathsf{maxPI}(f') \leq \mathsf{maxPI}(f)$.*

3 Properties of sumPI and maxPI

3.1 Properties of sumPI

Although in general, as we shall see, sumPI gives weaker formula size lower bounds than maxPI, the measure sumPI has several nice properties which make it more convenient to use in practice.

The next lemma shows that sumPI behaves like most other complexity measures with respect to composition of functions:

Lemma 14 *Let $g_1, \ldots, g_n$ be Boolean functions, and h be a function, $h : \{0,1\}^n \to \{0,1\}$. If $\mathsf{sumPI}(g_j) \leq a$ for $1 \leq j \leq n$ and $\mathsf{sumPI}(h) \leq b$, then for $f = h(g_1, \ldots, g_n)$, $\mathsf{sumPI}(f) \leq ab$.*

Proof: Let p be an optimal family of distribution functions associated with h and p_j be optimal families of distribution functions associated with g_j. Define the distribution function

$$q_x(i) = \sum_{j\in[n]} p_{g(x)}(j)p_{j,x}(i).$$

Assume that for $x, y \in S$ we have $f(x) \neq f(y)$. It is enough to show that

$$\sum_{i:\, x_i\neq y_i} \sqrt{\sum_{j\in[n]} p_{g(x)}(j)p_{j,x}(i)} \sqrt{\sum_{j\in[n]} p_{g(y)}(j)p_{j,y}(i)} \geq \frac{1}{ab}. \tag{3}$$

By Cauchy–Schwarz, the left hand side of Eq. 3 is greater than or equal to

$$\sum_{i:x_i\neq y_i} \sum_{j\in[n]} \sqrt{p_{g(x)}(j)p_{j,x}(i)} \sqrt{p_{g(y)}(j)p_{j,y}(i)}$$

$$= \sum_{j\in[n]} \left(\sqrt{p_{g(x)}(j)p_{g(y)}(j)} \sum_{i:x_i\neq y_i} \sqrt{p_{j,x}(i)p_{j,y}(i)} \right). \tag{4}$$

As long as $g_j(x) \neq g_j(y)$, by the definition of p_j, we have $\sum_{i:x_i\neq y_i} \sqrt{p_{j,x}(i)}\sqrt{p_{j,y}(i)} \geq 1/a$. Thus we can estimate the expression in Eq. 4 from below by:

$$\frac{1}{a} \sum_{j:g_j(x)\neq g_j(y)} \sqrt{p_{g(x)}(j)p_{g(y)}(j)}.$$

By the definition of p we can estimate the sum (without the $1/a$ coefficient) in the above expression from below by $1/b$, which finishes the proof. □

Another advantage of working with sumPI complexity is the following very powerful lemma of Ambainis [Amb03] which makes it easy to lower bound the sumPI complexity of iterated functions.

Definition 15 *Let $f : \{0,1\}^n \to \{0,1\}$ be any Boolean function. We define the dth iteration of f, written $f^d : \{0,1\}^{n^d} \to \{0,1\}$, inductively as $f^1(x) = f(x)$ and*

$$f^{d+1}(x) = f(f^d(x_1, \ldots, x_{n^d}), f^d(x_{n^d+1}, \ldots, x_{2n^d}), \ldots,$$

$$f^d(x_{(n-1)n^d+1}, \ldots, x_{n^{d+1}}))$$

Lemma 16 (Ambainis) *Let f be any Boolean function and f^d the dth iteration of f. Then* $\mathsf{sumPI}(f^d) \geq (\mathsf{sumPI}(f))^d$.

Combining this with Lemma 14, we get:

Corollary 17 *Let f be any Boolean function and f^d the dth iteration of f. Then* $\mathsf{sumPI}(f^d) = (\mathsf{sumPI}(f))^d$.

Ambainis shows that for total Boolean functions the square root of block sensitivity is a lower bound on the sumPI complexity [Amb02]. This, together with Lemmas 13 and 14 and the results of [NS94, BBC+01] imply the following:

Lemma 18 (Ambainis) *For total Boolean functions the* sumPI *complexity is in polynomial relation with the various (deterministic, randomized, quantum) decision tree complexities and the Fourier degree of the function.*

3.2 Properties of maxPI

One thing that makes sumPI so convenient to use is that it dualizes [ŠS05]. In this section we partially dualize the expression maxPI. The final expression remains a minimization problem, but we minimize over discrete index selection functions, instead of families of probability distributions, which makes it much more tractable. Still, we remark that maxPI can take exponential time (in the size of the truth table of f) whereas, sumPI takes polynomial time in the size of the truth table of f to compute by reduction to semidefinite programming.

Definition 19 (Index selection functions) *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function, $X{=}f^{-1}(0)$, and $Y{=}f^{-1}(1)$. For $i \in [n]$ let D_i be $|X|\times|Y|$ be defined by $D_i[x,y] = 1 - \delta_{x_i,y_i}$. We call the set of n Boolean $(0-1)$ matrices $\{P_i\}_{i\in n}$* index selection functions *if*

1. *$\sum_i P_i = E$, where $E[x,y] = 1$ for every $x \in X$, $y \in Y$. (informally: for every $x \in X$, $y \in Y$ we select a unique index)*

2. $P_i \leq D_i$ (informally: for every $x \in X$, $y \in Y$ the index we select is an i such that $x_i \neq y_i$).

Notice that index selection functions correspond to partitioning $X \times Y$, in such a way that if x, y are in the ith part, then $x_i \neq y_i$.

Theorem 20 (Spectral adversary version of maxPI) *Let f, X, Y be as in the previous definition. Let A be an arbitrary $|X| \times |Y|$ nonnegative matrix satisfying $A[x,y] = 0$ whenever $f(x) = f(y)$. Then*

$$\mathsf{maxPI}(f) = \min_{\{P_i\}_i} \max_{A} \frac{\|A\|_2}{\max_i \|A \circ P_i\|_2},$$

where $\{P_i\}_i$ runs through all index selection functions.

Proof: For a fixed family of probability distributions $p = \{p_x\}$, and for the expression

$$\max_{\substack{x,y \\ f(x)\neq f(y)}} \frac{1}{\max_{i:x_i \neq y_i} \sqrt{p_x(i) p_y(i)}}, \tag{5}$$

let us define the index selection function $P_i[x,y] = 1$ if $i = \mathrm{argmax}_{i:x_i \neq y_i} \sqrt{p_x(i) p_y(i)}$ and 0 otherwise. (Argmax is the smallest argument for which the expression attains its maximal value.) Then the denominator in Eq. 5 becomes equal to $\sum_{i:x_i \neq y_i} \sqrt{p_x(i) p_y(i)} P_i[x,y]$. If we replace the above system of P_is with any other choice of index selection function the value of $\sum_{i:x_i \neq y_i} \sqrt{p_x(i) p_y(i)} P_i[x,y]$ will not increase. Thus we can rewrite Eq. 5 as

$$\max_{\substack{x,y \\ f(x)\neq f(y)}} \frac{1}{\max_{\{P_i\}_i} \sum_{i:x_i \neq y_i} \sqrt{p_x(i) p_y(i)} P_i[x,y]},$$

where here $P_i[x,y]$ runs through all index selection functions. Thus:

$$\mathsf{maxPI}(f) =$$

$$1 \Big/ \max_{p} \min_{\substack{x,y \\ f(x)\neq f(y)}} \max_{\{P_i\}_i} \sum_{i:x_i \neq y_i} \sqrt{p_x(i) p_y(i)} P_i[x,y]. \tag{6}$$

Notice that in Eq. 6 the minimum is interchangeable with the second maximum. The reason for this is that for a fixed p there is a fixed $\{P_i[x,y]\}_i$ system that maximizes $\sum_{i:x_i \neq y_i} \sqrt{p_x(i) p_y(i)} P_i[x,y]$ for all $x, y:\ f(x) \neq f(y)$. Thus:

$$\mathsf{maxPI}(f) =$$

$$1 \Big/ \max_{\{P_i\}_i} \max_{p} \min_{\substack{x,y \\ f(x)\neq f(y)}} \sum_{i:x_i \neq y_i} \sqrt{p_x(i) p_y(i)} P_i[x,y].$$

Following the proof of the main theorem of Špalek and Szegedy we can create the semidefinite version of the above expression. The difference here, however, is that we have to treat $\{P_i\}_i$ (the index selection functions) as a "parameter" of the semidefinite system over which we have to maximize. Unfortunately it also appears in the final expression.

Semidefinite version of maxPI: For fixed $\{P_i\}_i$ let $\mu'_{\max}$ be the solution of the following semidefinite program:

$$\begin{aligned} &\text{maximize } \mu' \\ &\text{subject to } \quad (\forall i) \quad R_i \succeq 0, \\ &\qquad\qquad \textstyle\sum_i R_i \circ I = I, \\ &\qquad\qquad \textstyle\sum_i R_i \circ P_i \geq \mu' F. \end{aligned}$$

Define $\mu_{\max}$ as the maximum of $\mu'_{\max}$, where P_i $(1 \leq i \leq n)$ run through all index selection functions. Then $\mathsf{maxPI} = 1/\mu_{\max}$.

We can dualize the above program and simplify it in same way as was done in Špalek and Szegedy for the case of sumPI with the only change that D_i needs to be replaced with P_i, and that we have to minimize over all choices of $\{P_i\}_i$. □

4 Formula size lower bounds

We transform in two steps the problem of proving lower bounds on formula size steps into a combinatorial problem which is easier to work with. First we apply the theorem of Karchmer and Wigderson [KW88], Theorem 5 which gives an exact characterization of the formula size of f in terms of the communication complexity of a relation associated with f. We then use the well-known fact that the size of the smallest partition of a relation into monochromatic rectangles is a lower bound on the smallest number of leaves in a communication protocol for the relation. We then lower bound the size of such a partition.

A natural way to lower bound the size of the smallest partition is to find a measure which is subadditive on rectangles. Then the measure of the whole space divided by the size of the largest rectangle in the partition will lower bound the number of rectangles in the partition. In the next section we show our key lemma that the spectral norm squared of a matrix is one such measure.

4.1 Key combinatorial lemma

We first prove a combinatorial lemma which is the key to our main result. This lemma relates the spectral norm squared of a matrix to the spectral norm squared of its submatrices, and may also be of independent interest.

Let X and Y be finite sets. A set system $\mathcal{S}$ (over $X \times Y$) will be called a *covering* if $\cup_{S \in \mathcal{S}} S = X \times Y$. Further, $\mathcal{S}$ will be called a *partition* if $\mathcal{S}$ is a covering and the intersection of any two distinct sets from $\mathcal{S}$ is empty. A *rectangle* (over $X \times Y$) is an arbitrary subset of $X \times Y$ of the form $X_0 \times Y_0$ for some $X_0 \subseteq X$ and $Y_0 \subseteq Y$. A set system $\mathcal{R}$ will be called a *rectangle partition* if $\mathcal{R}$ is a partition and each $R \in \mathcal{R}$ is a rectangle. For a subset $S \subseteq X \times Y$ we define:

$$A_S[x, y] = A[x, y], \text{ if } (x, y) \in S \text{ and } 0 \text{ otherwise.} \tag{7}$$

We are now ready to state the lemma:

Lemma 21 *Let A be an arbitrary $|X| \times |Y|$ matrix (possibly with complex entries), and $\mathcal{R}$ a partition of $X \times Y$. Then $\|A\|_2^2 \leq \sum_{R \in \mathcal{R}} \|A_R\|_2^2$*

Proof: By Proposition 1, $\|A\|_2 = \max_{u,v} |u^* A v|$, where the maximum is taken over all unit vectors u, v. Let u, v be the unit vectors realizing this maximimum. Then we have

$$\|A\|_2 = |u^* A v| = \left| u^* \left(\sum_{R \in \mathcal{R}} A_R \right) v \right| = \left| \sum_{R \in \mathcal{R}} u^* A_R v \right|.$$

As each $R \in \mathcal{R}$ is a rectangle, it can be expressed as $R = X_0 \times Y_0$ for some $X_0 \subseteq X$ and $Y_0 \subseteq Y$. Let $u_R[x] = u_[x]$ if $x \in X_0$ and 0 otherwise, and similarly $v_R[y] = v[y]$ if $y \in Y_0$ and 0 otherwise. Notice that $\{u_R\}_{R \in \mathcal{R}}$ do not in general form a partition of u. We now have

$$\begin{aligned} \|A\|_2 &= \left| \sum_{R \in \mathcal{R}} u_R^* A_R v_R \right| \leq \sum_{R \in \mathcal{R}} |u_R^* A_R v_R| \\ &\leq \sum_{R \in \mathcal{R}} \|A_R\|_2 |u_R| |v_R| \end{aligned}$$

by Proposition 1. Applying the Cauchy–Schwarz inequality, we obtain

$$\|A\|_2 \leq \left(\sum_{R \in \mathcal{R}} \|A_R\|_2^2 \right)^{1/2} \left(\sum_{R \in \mathcal{R}} |u_R|^2 |v_R|^2 \right)^{1/2}.$$

Now it simply remains to observe that

$$\sum_{R \in \mathcal{R}} |u_R|^2 |v_R|^2 = \sum_{R \in \mathcal{R}} \sum_{(x,y) \in R} u[x]^2 v[y]^2 = |u|^2 |v|^2 = 1,$$

as $\mathcal{R}$ is a partition of $X \times Y$. □

4.2 Deterministic formulae

In this section, we prove our main result that maxPI is a lower bound on formula size. We first identify two natural properties which are sufficient for a function to be a formula size lower bound.

Definition 22 *A function $\mu : 2^{X \times Y} \to \mathbb{R}^+$ is called a* rectangle measure *if the following properties hold.*

1. *(Subadditivity) For any rectangle partition $\mathcal{R}$ of $X \times Y$, $\mu(X \times Y) \leq \sum_{R \in \mathcal{R}} \mu(R)$.*
2. *(Monotonicity) For any rectangle $R \subseteq X \times Y$, and subset $S \subseteq X \times Y$, if $R \subseteq S$ then $\mu(R) \leq \mu(S)$.*

Theorem 21 and item 3 of Proposition 1 imply that for any $|X| \times |Y|$ matrix A with non-negative entries $S \to \|A_S\|_2^2$ of is a rectangle measure. Other examples include the rank of A_S for any matrix A over any field (see Section 5.4), and the μ-rectangle size bounds of [KKN95] (see Section 5.5).

Let $\mathcal{S}_1, \mathcal{S}_2$ be two families of sets over the same universe. We say that $\mathcal{S}_1$ is *embedded* in $\mathcal{S}_2$ ($\mathcal{S}_1 \prec \mathcal{S}_2$) if for every $S \in \mathcal{S}_1$ there is a $S' \in \mathcal{S}_2$ such that $S \subseteq S'$.

Proposition 23 *Let μ be a rectangle measure over $2^{X\times Y}$, $\mathcal{S}$ be a covering of $X\times Y$ and $\mathcal{R}$ a rectangle partition of $X\times Y$ such that $\mathcal{R} \prec \mathcal{S}$. Then $|\mathcal{R}| \geq \frac{\mu(X\times Y)}{\max_{S\in\mathcal{S}} \mu(S)}$.*

The proof follows by subadditivity and monotonicity of μ.

Theorem 24 (Main Theorem)

$$\mathsf{sumPI}^2(f) \leq \mathsf{maxPI}^2(f) \leq C^D(R_f) \leq \mathsf{L}(f)$$

Proof: We have seen that $\mathsf{sumPI}^2(f) \leq \mathsf{maxPI}^2(f)$, and $C^D(R_f) \leq \mathsf{L}(f)$ follows from the Karchmer–Wigderson communication game characterization of formula size, thus we focus on the inequality $\mathsf{maxPI}^2(f) \leq C^D(R_f)$.

Let $\mathcal{R}$ be a monochromatic rectangle partition of R_f such that $|\mathcal{R}| = C^D(R_f)$, and let A be an arbitrary $|X|\times|Y|$ matrix with nonnegative real entries. For $R \in \mathcal{R}$ let $\mathrm{color}(R)$ be the least index c such that $x_c \neq y_c$ holds for all $(x, y) \in R$. By assumption each R is monochromatic, thus such a color exists. Define

$$S_c = \cup_{\mathrm{color}(R)=c} R.$$

Then $\mathcal{R}$ is naturally embedded in the covering $\{S_c\}_{c\in[n]}$. For any $S \subseteq X\times Y$, let $\mu_A(S) = \|A_S\|_2^2$. By Lemma 21, and item 3 of Proposition 1, μ_A is a rectangle measure. Hence by Proposition 23,

$$\max_A \frac{\|A\|_2^2}{\max_c \|A_{S_c}\|_2^2} \leq C^D(R_f).$$

We have exhibited a particular index selection function, the $\{S_c\}_c$, for which this inequality holds, thus it also holds for $\mathsf{maxPI}^2(f)$ which is the minimum over all index selection functions. □

4.3 Probabilistic Formulae

The properties of sumPI allow us to show that it can be used to lower bound the probabilistic formula size.

Lemma 25 *Let $\epsilon < 1/2$. If $f : S \to \{0, 1\}$ is ϵ-approximated by functions $\{f_j\}_{j\in J}$ with $\mathsf{sumPI}(f_j) \leq s$ for every $j \in J$, then $\mathsf{sumPI}(f) \leq s/(1 - 2\epsilon)$.*

Proof: By assumption there is a probability distribution $\alpha = \{\alpha_j\}_{j\in J}$ such that $\Pr[f(x) = f_j(x)] \geq 1 - \epsilon$. Thus for a fixed $x \in S$, letting $J_x = \{j \in J : f(x) = f_j(x)\}$, we have $\sum_{j\in J_x} \alpha_j \geq 1 - \epsilon$. Hence for any $x, y \in S$ we have $\sum_{j\in J_x\cap J_y} \alpha_j \geq 1 - 2\epsilon$. For convenience, we write $J_{x,y}$ for $J_x \cap J_y$. As $\mathsf{sumPI}(f_j) \leq s$ there is a family of probability distributions p_j such that whenever $f_j(x) \neq f_j(y)$

$$\sum_{\substack{i \\ x_i\neq y_i}} \sqrt{p_{j,x}(i)p_{j,y}(i)} \geq 1/s.$$

Define $p_x(i) = \sum_{j \in J} \alpha_j p_{j,x}(i)$. Let x, y be such that $f(x) \neq f(y)$.

$$\begin{aligned}
&\sum_{\substack{i \\ x_i \neq y_i}} \sqrt{p_x(i) p_y(i)} \\
&= \sum_{\substack{i \\ x_i \neq y_i}} \sqrt{\sum_{j \in J} \alpha_j p_{j,x}(i)} \sqrt{\sum_{j \in J} \alpha_j p_{j,y}(i))} \\
&\geq \sum_{\substack{i \\ x_i \neq y_i}} \sqrt{\sum_{j \in J_{x,y}} \alpha_j p_{j,x}(i)} \sqrt{\sum_{j \in J_{x,y}} \alpha_j p_{j,y}(i)} \\
&\geq \sum_{\substack{i \\ x_i \neq y_i}} \sum_{j \in J_{x,y}} \sqrt{\alpha_j p_{j,x}(i)} \sqrt{\alpha_j p_{j,y}(i)} \\
&= \sum_{j \in J_{x,y}} \left(\alpha_j \sum_{\substack{i \\ x_i \neq y_i}} \sqrt{p_{j,x}(i) p_{j,y}(i)} \right) \\
&\geq \frac{1 - 2\epsilon}{s},
\end{aligned}$$

where for the third step we have used the Cauchy–Schwarz Inequality. □

This lemma immediately shows that the sumPI method can give lower bounds on probabilistic formula size.

Theorem 26 *Let $S \subseteq \{0,1\}^n$ and $f : S \to \{0,1\}$. Then $\mathsf{L}^\epsilon(f) \geq ((1 - 2\epsilon)\mathsf{sumPI}(f))^2$ for any $\epsilon < 1/2$.*

Proof: Suppose that $\{f_j\}_{j \in J}$ gives an ϵ-approximation to f. Using Lemma 25 in the contrapositive implies that there exists some $j \in J$ with $\mathsf{sumPI}(f_j) \geq (1 - 2\epsilon)\mathsf{sumPI}(f)$. Theorem 24 then implies $\mathsf{L}(f_j) \geq ((1 - 2\epsilon)\mathsf{sumPI}(f))^2$ which gives the statement of the theorem. □

5 Comparison among methods

In this section we look at several formula size lower bound techniques in the literature and see how they compare with our methods. A bottleneck in formula size lower bounds seems to have been to go beyond methods which only consider pairs (x, y) with $f(x) \neq f(y)$ which have Hamming distance 1. In fact, the methods of Khrapchenko, Koutsoupias, and a lemma of Håstad can all be seen as special cases of the sumPI method where only pairs of Hamming distance 1 are considered.

5.1 Khrapchenko's method

One of the oldest and most general techniques available for showing formula size lower bounds is Khrapchenko's method [Khr71], originally used to give a tight $\Omega(n^2)$ lower bound for the parity function. This method considers a bipartite graph whose left vertices are the 0-inputs to f and

whose right vertices are the 1-inputs. The bound given is the product of the average degree of the right and left hand sides.

Theorem 27 (Khrapchenko) *Let $S \subseteq \{0,1\}^n$ and $f : S \to \{0,1\}$. Let $A \subseteq f^{-1}(0)$ and $B \subseteq f^{-1}(1)$. Let C be the set of pairs $(x,y) \in A \times B$ with Hamming distance 1, that is $C = \{(x,y) \in A \times B : d_H(x,y) = 1\}$. Then $\mathsf{L}(f) \geq \mathsf{sumPI}(f)^2 \geq \frac{|C|^2}{|A||B|}$.*

Khrapchenko's method can easily be seen as a special case of the probability scheme. Letting A, B, C be as in the statement of the theorem, we set up our probability distributions as follows:

- $p_A(x)=1/|A|$ for all $x \in A$, $p_A(x)=0$ otherwise
- $p_B(x)=1/|B|$ for all $x \in B$, $p_B(x)=0$ otherwise
- $q(x,y)=1/|C|$ for all $(x,y) \in C$, $q(x,y)=0$ otherwise
- $p_{x,i}(y)=1$ if $(x,y) \in C$ and $x_i \neq y_i$, 0 otherwise. Note that this is a probability distribution as for every x there is only one y such that $(x,y) \in C$ and $x_i \neq y_i$.

By Theorem 9 and Theorem 24,

$$\mathsf{L}(f) \geq \min_{\substack{x,y,i \\ f(x) \neq f(y), \\ x_i \neq y_i}} \frac{p_A(x) p_B(y) p'_{x,i}(y) p'_{y,i}(x)}{q(x,y)} = \frac{|C|^2}{|A||B|},$$

where the expression in the middle is a lower bound on $\mathsf{sumPI}(f)^2$.

5.2 The Koutsoupias bound

Koutsoupias [Kou93] extends Khrapchenko's method with a spectral version. The weights are always 1 for pairs of inputs with different function values that have Hamming distance 1, and 0 everywhere else.

Theorem 28 (Koutsoupias) *Let $f : \{0,1\}^n \to \{0,1\}$, and let $A \subseteq f^{-1}(0)$, and $B \subseteq f^{-1}(1)$. Let $C = \{(x,y) \in A \times B : d_H(x,y) = 1\}$. Let Q be a $|B| \times |A|$ matrix $Q[x,y] = C(x,y)$ where C is identified with its characteristic function. Then $\mathsf{L}(f) \geq \mathsf{sumPI}(f)^2 \geq \|Q\|_2^2$.*

Proof: The bound follows easily from the the spectral version of sumPI. Let Q be as in the statement of the theorem. Notice that since we only consider pairs with Hamming distance 1, for every row and column of Q_i there is at most one nonzero entry, which is at most 1. Thus by Proposition 1 we have $\|Q_i\|_2^2 \leq \|Q\|_1 \|Q\|_\infty \leq 1$. The theorem now follows from Theorem 24. □

5.3 Håstad's method

The shrinkage exponent of Boolean formulae is the least upper bound γ such that subject to a random restriction where each variable is left free with probability p, Boolean formulae shrink from size L to expected size $p^\gamma L$. Determining the shrinkage exponent is important as Andreev [And87] defined a function f whose formula size is $\mathsf{L}(f) = n^{1+\gamma}$. Håstad [Hås98] shows the shrinkage exponent of Boolean formulae is 2 and thereby obtains an $n^{3-o(1)}$ formula size lower bound, the largest

bound known for an explicit function. On the way to this result, Håstad proves an intermediate lemma which gives a lower bound on formula size that depends on the probability that restrictions of a certain form occur. He proves that this lemma is a generalization of Khrapchenko's method; we prove that Håstad's lemma is in turn a special case of sumPI. Since Håstad's method uses random restrictions, which at first glance seems completely different from adversary methods, it comes as a surprise that it is in fact a special case of our techniques.

Definition 29 *For any function $f : \{0,1\}^n \to \{0,1\}$,*

1. *A* restriction *is a string in $\{0,1,\star\}^n$ where $\star$ means the variable is left free, and 0 or 1 mean the variable is set to the constant 0 or 1, respectively.*
2. *The restricted function $f|_\rho$ is the function that remains after the non-$\star$ variables in ρ are fixed.*
3. *R_p is the distribution on random restrictions to the variables of f obtained by setting each variable, independently, to $\star$ with probability p, and to 0 or 1 each with probability $\frac{(1-p)}{2}$.*
4. *A* filter *Δ is a set of restrictions which has the property that if $\rho \in \Delta$, then every ρ' obtained by fixing one of the $\star$s to a constant is also in Δ.*
5. *When p is known from the context, and for any event E, and any filter Δ, we write $\Pr[E|\Delta]$ to mean $\Pr_{\rho \in R_p}[E|\rho \in \Delta]$.*

Theorem 30 (Håstad, Lemma 4.1) *Let $f : \{0,1\}^n \to \{0,1\}$. Let A be the event that a random restriction in R_p reduces f to the constant 0, B be the event that a random restriction in R_p reduces f to the constant 1, and let C be the event that a random restriction $\rho \in R_p$ is such that $f|_\rho$ is a single literal. Then*

$$\mathsf{L}(f) \geq \frac{Pr[C|\Delta]^2}{Pr[A|\Delta]Pr[B|\Delta]} \left(\frac{1-p}{2p}\right)^2$$

Proof: We show that the theorem follows from the probability scheme (Definition 7). In this proof we only consider restrictions obtained from R_p that are in the filter Δ. We also abuse notation and use A and B to mean the sets of restrictions in Δ which contribute with non-zero probability to the events A and B respectively.

Implicit in Håstad's proof is the following relation between restrictions in A and B. For every $\rho \in C$, $f|_\rho$ reduces to a single literal, that is, for every $\rho \in C$, there is an i such that $f|_\rho = x_i$ (or $\neg x_i$ if the variable is negated). Define ρ^b to be ρ where x_i is set to b, for $b \in \{0,1\}$ (set x_i to $1-b$ if the variable is negated). To fit into the framework of the probability scheme, let $\overline{\rho^b}$ be ρ^b where all remaining $\star$s are set to 1. This doesn't change the value of the function, because it is already constant on ρ^b. Then we say that $\overline{\rho^0}, \overline{\rho^1}$ are in the relation.

We set $p_A(\sigma) = \frac{\Pr[\sigma]}{\Pr[A|\Delta]}$ for any $\sigma \in A$, and $p_B(\tau) = \frac{\Pr[\tau]}{\Pr[B|\Delta]}$ for any $\tau \in B$, and for every pair $\overline{\rho^0}, \overline{\rho^1}$ in the relation, where $\rho \in C$, $f|_\rho = x_i$ or $\neg x_i$, let

$$\begin{aligned} p'_{\overline{\rho^0},i}(\overline{\rho^1}) &= 1 \\ p'_{\overline{\rho^1},i}(\overline{\rho^0}) &= 1 \\ q(\overline{\rho^0}, \overline{\rho^1}) &= \frac{Pr[\rho]}{Pr[C|\Delta]} \end{aligned}$$

The probabilities are 0 on all other inputs. We can easily verify that the probabilities sum to 1. For p', notice that the Hamming distance between $\overline{\rho^0}$ and $\overline{\rho^1}$ is 1, so when $\overline{\rho^b}$ and i are fixed, there is only a single $\overline{\rho^{1-b}}$ with probability 1.

By Theorem 9 and Theorem 24,

$$\begin{aligned} \mathsf{L}(f) &\geq \frac{p_A(x)p_B(y)p'_{y,i}(x)p'_{x,i}(y)}{q(x,y)^2} \\ &= \frac{Pr[\rho^0]}{Pr[A|\Delta]}\frac{Pr[\rho^1]}{Pr[B|\Delta]}\left(\frac{Pr[C|\Delta]}{Pr[\rho]}\right)^2 \end{aligned}$$

Finally, notice that $Pr[\rho] = \frac{2p}{1-p}Pr[\rho^b]$. □

Remark Håstad actually defines $f|_\rho$ to be the result of reducing the formula for f (not the function) by applying a sequence of reduction rules, for each restricted variable. So there is a subtlety here about whether $f|_\rho$ denotes the reduced formula, or the reduced function, and the probabilities might be different if we are in one setting or the other. However both in his proof and ours, the only thing that is used about the reduction is that if the formula or function reduces to a single literal, then fixing this literal to 0 or to 1 reduces the function to a constant. Therefore, both proofs go through for both settings.

5.4 Razborov's method

Razborov [Raz90] proposes a formula size lower bound technique using matrix rank:

Theorem 31 (Razborov) *Let $R \subseteq X \times Y \times Z$ be a relation and let $\mathcal{R}$ be a partition of $X \times Y$ into monochromatic rectangles satisfying $|\mathcal{R}| = C^D(R)$. Let $\mathcal{S}$ be a covering of $X \times Y$ such that $\mathcal{R} \prec \mathcal{S}$. Then*

$$\max_{A \neq 0} \frac{\mathrm{rk}(A)}{\max_{S \in \mathcal{S}} \mathrm{rk}(A_S)} \leq C^D(R).$$

It can be easily verified that the function $S \to \mathrm{rk}(A_S)$ is a rectangle measure, thus this theorem follows from Proposition 23. Razborov uses Theorem 31 to show superpolynomial monotone formula size lower bounds, but also shows that the method becomes trivial (limited to $O(n)$ bounds) for regular formula size [Raz92]. An interesting difference between matrix rank and and spectral norm is that $\mathrm{rk}(A+B) \leq \mathrm{rk}(A) + \mathrm{rk}(B)$ holds for any two matrices A, B, while a necessary condition for subadditivity of the spectral norm squared is that A, B be disjoint rectangles.

5.5 Karchmer, Kushilevitz, and Nisan

In this section we discuss two methods proposed by Karchmer, Kushilevitz, and Nisan [KKN95] for proving lower bounds on the communication complexity of relations. Our presentation here differs from the original in order to highlight similarities with the present discussion.

Both of the techniques of [KKN95] arise from linear program relaxations of integer program formulations of communication complexity bounds. First they look at nondeterministic complexity, which corresponds to the cover number of a relation $C^N(R)$, that is, the minimum number of monochramatic relations needed to cover the relation R. Writing the linear program relaxation of the cover number, they obtain the following bound:

Theorem 32 *Let $R \subseteq X \times Y \times Z$ be a relation and let $\mathcal{R}$ be a partition of $X \times Y$ into monochromatic rectangles satisfying $|\mathcal{R}| = C^D(R)$. Let $\mathcal{S}$ be a covering of $X \times Y$ such that $\mathcal{R} \prec \mathcal{S}$. Then*

$$C^D(R_f) \geq \max_{A \neq 0} \frac{\|A\|_F^2}{\max_R \|A_R\|_F^2}$$

Notice that this bound looks the same as ours with the spectral norm replaced by the Frobenius norm. It is easy to see that the Frobenius norm squared is both subbadditive and monotone and thus a rectangle measure in the sense of Definition 22. They show some other interesting properties of this measure, such as its logarithm characterizes (up to a $\log n$ factor) nondeterministic communication complexity, and this measure satisfies a direct sum property.

Karchmer, Kushilevitz, and Nisan then turn to formulate the rectangle partition bound as a integer programming problem, and investigate its relaxation as a linear program. They show that, when dualized, this bound has the following form:

Theorem 33 (Karchmer-Kushilevitz-Nisan) *Let $R \subseteq X \times Y \times Z$ be a relation and let $\mathcal{R}$ be a partition of $X \times Y$ into monochromatic rectangles satisfying $|\mathcal{R}| = C^D(R)$.*

$$C^D(R_f) \geq \max_{A \neq 0} \frac{\text{Entrysum}(A)}{\max_{S \in \mathcal{R}} \text{Entrysum}(A_S)}$$

Notice that $S \to \text{Entrysum}(A_S)$ for a matrix A is again a subadditive measure. The essential difference between these two methods is that in the latter one one can use negative weights in the matrix A. This allows one to prove larger formula size lower bounds using the second theorem, but also means that this measure does not satisfy the monotonicity property, and so one must be careful in checking the weights of all monochromatic rectangles. They show that this bound is larger than Khrapchenko's method, but cannot prove lower bounds larger than n^2.

6 Limitations

6.1 Hamming distance 1 techniques

We show that the bounds for a function f given by Khrapchenko's and Koutsoupias' method, and by Håstad's lemma are upper bounded by the product of the zero sensitivity and the one sensitivity of f. We will later use this bound to show a function on n bits for which the best lower bound given by these methods is n and for which an $\approx n^{1.32}$ bound is provable by sumPI^2.

Lemma 34 *The bound given by the Khrapchenko method (Theorem 27), Koutsoupias' method (Theorem 28), and Håstad's Lemma (Theorem 30) for a function f are at most $s_0(f)s_1(f) \leq s^2(f)$.*

Proof: Let A be a nonnegative matrix, with nonzero entries only in positions (x, y) where $f(x) = 0, f(y) = 1$ and the Hamming distance between x, y is one. We first show that

$$\max_A \frac{\|A\|_2^2}{\max_i \|A_i\|_2^2} \leq s_0(f)s_1(f). \tag{8}$$

Let a_{max} be the largest entry in A. A can have at most $s_0(f)$ many nonzero entries in any row, and at most $s_1(f)$ many nonzero entries in any column, thus by item 2 of Propostion 1,

$$\|A\|_2^2 \leq \|A\|_1 \|A\|_\infty \leq a_{max}^2 s_0(f) s_1(f).$$

On the other hand, for some i, the entry a_{max} appears in A_i, and so by item 1 of Proposition 1, $\|A_i\|_2^2 \geq a_{max}^2$. Equation 8 follows.

Now we see that the left hand side of Equation 8 is larger than the three methods in the statement of the theorem. That it is more general than Koutsoupias method is clear. To see that it is more general than the probability schemes method where $q(x,y)$ is only positive if the Hamming distance between x, y is one: given the probability distributions q, p_A, p_B, define the matrix $A[x,y] = q(x,y)/\sqrt{p_A(x)p_B(y)}$. By item 1 of Proposition 1, $\|A\|_2 \geq 1$, witnessed by the unit vectors $u[x] = \sqrt{p_A(x)}$ and $v[y] = \sqrt{p_B(y)}$. As each reduced matrix A_i has at most one nonzero entry in each row and column, by item 2 of Proposition 1 we have

$$\max_i \|A_i\|_2^2 \leq \max_{x,y} \frac{q^2(x,y)}{p_A(x)p_B(y)}.$$

Thus we have shown

$$\max_A \frac{\|A\|_2^2}{\max_i \|A_i\|_2^2} \geq \max_{p_A, p_B, q} \min_{x,y} \frac{p_A(x)p_B(y)}{q^2(x,y)}.$$

□

The only reference to the limitations of these methods we are aware of is Schürfeld [Sch83], who shows that Khrapchenko's method cannot prove bounds greater than $C_0(f)C_1(f)$.

6.2 Limitations of sumPI and maxPI

The limitations of the adversary method are well known [Amb02, LM04, Sze03, Zha05, ŠS05]. Špalek and Szegedy, in unifying the adversary methods, also give the most elegant proof of their collective limitation. The same proof also shows the same limitations hold for the maxPI measure.

Lemma 35 *Let $f : \{0,1\}^n \to \{0,1\}$ be any partial or total Boolean function. If f is total (respectively, partial) then* $\mathsf{maxPI}(f) \leq \sqrt{C_0(f)C_1(f)}$ *(respectively,* $\min\{\sqrt{nC_0(f)}, \sqrt{nC_1(f)}\}$*).*

Proof: Assume that f is total. Take x, y such that $f(x) = 0$ and $f(y) = 1$. We choose any 0-certificate B_0 for x and any 1-certificate B_1 for y and let $p_x(i) = 1/C_0(f)$ for all $i \in B_0$ and $p_y(i) = 1/C_1(f)$ for all $i \in B_1$. As f is total, we have $B_0 \cap B_1 \neq \emptyset$, thus let $j \in B_0 \cap B_1$. For this j we have $p_x(j)p_y(j) \geq 1/\left(C_0(f)C_1(f)\right)$, thus $\min_i 1/p_x(i)p_y(i) \geq C_0(f)C_1(f)$.

The case where f is partial follows similarly. As we no longer know that $B_0 \cap B_1 \neq \emptyset$, we put a uniform distribution over a 0-certificate of x and the uniform distribution over $[n]$ on y or vice versa. □

This lemma implies that sumPI and maxPI are polynomially related for total f.

Corollary 36 *Let f be a total Boolean function. Then* $\mathsf{maxPI}(f) \leq \mathsf{sumPI}^4(f)$.

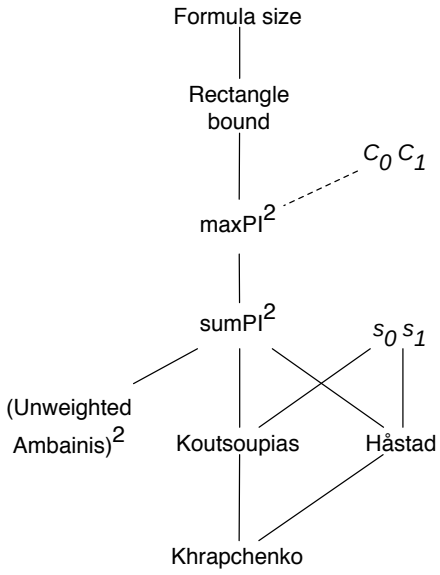

Figure 1: Summary of the methods and their limitations. The containments denoted by solid lines hold for total as well as partial functions. All containments are strict.

Proof: By [Amb02, Thm. 5.2] we know that $\sqrt{bs(f)} \leq \mathsf{sumPI}(f)$. As f is total, by the above lemma we know that $\mathsf{maxPI}(f) \leq \sqrt{C_0(f)C_1(f)}$. This in turn is smaller than $bs(f)^2$ as $C(f) \leq s(f)bs(f)$ [Nis91]. The statement follows. □

Besides the certificate complexity barrier, another serious limitation of the sumPI method occurs for partial functions where every positive input is far in Hamming distance from every negative input. Thus for example, if for any pair x, y where $f(x) = 1$ and $f(y) = 0$ we have $d_H(x, y) \geq \epsilon n$, then by putting the uniform distribution over all input bits it follows that $\mathsf{sumPI}(f) \leq 1/\epsilon$. The measure maxPI does not face this limitation as there we still only have one term in the denominator.

Following this line of thinking, we can give an example of a partial function f where $\mathsf{maxPI}(f) \gg \mathsf{sumPI}(f)$. Such an example is the Collision problem (see Section 7.3), as here any positive and negative input must differ on at least $n/2$ positions. Another family of examples comes from property testing, where the promise is that the input either has some property, or that it is ϵ-far from having the property.

7 Concrete lower bounds

The quantum adversary argument has been used to prove lower bounds for a variety of problems. Naturally, all of these lower bounds carry over to formula size lower bounds. In this section we present some new lower bounds, in order to highlight the strengths and weaknesses of maxPI and sumPI.

7.1 Recursive majorities

As an example of applying sumPI, we look at the recursive majority of three function. We let $\mathsf{R{-}MAJ}_3^h : \{0,1\}^{3^h} \to \{0,1\}$ be the function computed by a complete ternary tree of depth h where every internal node is labeled by a majority gate and the input is given at the leaves.

Recursive majority of three has been studied before in various contexts. It is a monotone function which is very sensitive to noise [MO03], making it useful for hardness amplification in NP [O'D02]. Jayram, Kumar, and Sivakumar [JKS03] give nontrivial lower and upper bounds on the randomized decision tree complexity of recursive majority of three. They show a lower bound of $(7/3)^h$ on the randomized decision tree complexity. As far as we know, the quantum query complexity of recursive majority of three has not yet been investigated. We show a lower bound of 2^h on the quantum query complexity.

Lemma 37 $\mathsf{sumPI}(\mathsf{R{-}MAJ}_3^h) = \mathsf{maxPI}(\mathsf{R{-}MAJ}_3^h) = 2^h$

Proof: To see that $\mathsf{maxPI}(\mathsf{R{-}MAJ}_3^h) \leq 2^h$, observe that $C_0(\mathsf{R{-}MAJ}_3^h) = C_1(\mathsf{R{-}MAJ}_3^h) = 2^h$. The result then follows from Lemma 35.

We now turn to the lower bound. We will first show a lower bound for $\mathsf{R{-}MAJ}_3^1$, the majority of three function, and then apply Lemma 16. Consider the following table, where the rows are indexed by negative instances x, the columns by positive instances y, and 1's indicate when $d_H(x,y) = 1$.

	110	101	011
001	0	1	1
010	1	0	1
100	1	1	0

Interpreting this table as the adjacency matrix of a graph, it is clear that every vertex has degree 2. Thus Khrapchenko's method gives a bound of 4 for the base function. By Theorem 27 we have $\mathsf{sumPI}(\mathsf{R{-}MAJ}_3^1) \geq 2$. Now applying Lemma 16 gives the lemma. □

From Lemma 37 we immediately obtain quantum query complexity and formula size lower bounds:

Theorem 38 *Let* $\mathsf{R{-}MAJ}_3^h$ *be the recursive majority of three function of height* h*. Then* $Q_\epsilon(\mathsf{R{-}MAJ}_3^h) \geq (1 - 2\sqrt{\epsilon(1-\epsilon)})2^h$ *and* $\mathsf{L}^\epsilon(\mathsf{R{-}MAJ}_3^h) \geq (1-2\epsilon)4^h$.

The best upper bound on the formula size of $\mathsf{R{-}MAJ}_3^h$ is 5^h. For this bound, we will use the following simple proposition about the formula size of iterated functions.

Proposition 39 *Let* $S \subseteq \{0,1\}^n$ *and* $f : S \to \{0,1\}$*. If* $\mathsf{L}(f) \leq s$ *then* $\mathsf{L}(f^d) \leq s^d$*, where* f^d *is the* d*th iteration of* f*.*

Proposition 40 $\mathsf{L}(\mathsf{R{-}MAJ}_3^h) \leq 5^h$.

Proof: The formula $(x_1 \wedge x_2) \vee ((x_1 \vee x_2) \wedge x_3)$ computes $\mathsf{R{-}MAJ}_3^1$ and has 5 leaves. Using Proposition 39 gives $\mathsf{L}(\mathsf{R{-}MAJ}_3^h) \leq 5^h$. □

7.2 Ambainis' function

We define a function $\mathrm{f_A} : \{0,1\}^4 \to \{0,1\}$ after Ambainis [Amb03]. This function evaluates to 1 on the following values: 0000, 0001, 0011, 0111, 1111, 1110, 1100, 1000. That is, $f(x) = 1$ when $x_1 \leq x_2 \leq x_3 \leq x_4$ or $x_1 \geq x_2 \geq x_3 \geq x_4$. To obtain this formulation from Ambainis' original definition, exchange x_1 and x_3, and take the negation of the resulting function. There are a few things to notice about this function. The sensitivity of $\mathrm{f_A}$ is 2 on every input. Also on an input $x = x_1x_2x_3x_4$ the value of $\mathrm{f_A}(x)$ changes if both bits sensitive to x are flipped simultaneously, and if both bits insensitive for x are flipped simultaneously.

We will be looking at iterations of the base function $\mathrm{f_A}$ as in Definition 15. Notice that the sensitivity of ${\mathrm{f_A}}^d$ is 2^d on every input $x \in \{0,1\}^{4^d}$.

Lemma 41 $\mathsf{sumPI}({\mathrm{f_A}}^d) = 2.5^d$.

Proof: Ambainis has already shown that $\mathsf{sumPI}({\mathrm{f_A}}^d) \geq 2.5^d$ [Amb03].

We now show the upper bound. We will show an upper bound for the base function $\mathrm{f_A}$ and then use the composition Lemma 14. Every input $x_1x_2x_3x_4$ has two sensitive variables and two insensitive variables. For any $x \in \{0,1\}^4$ we set $p_x(i) = 2/5$ if i is sensitive for x and $p_x(i) = 1/10$ if i is insensitive for x. The claim follows from the following observation: for any $x, y \in \{0,1\}^4$ such that $f(x) \neq f(y)$ at least one of the following holds

- x and y differ on a position i which is sensitive for both x and y. Thus $\sum_i \sqrt{p_x(i)p_y(i)} \geq 2/5$
- x and y differ on at least 2 positions, each of these positions being sensitive for at least one of x, y. Thus $\sum_i \sqrt{p_x(i)p_y(i)} \geq 2\sqrt{1/25} = 2/5$

□

This lemma gives us a bound of $6.25^d \approx N^{1.32}$ on the formula size of ${\mathrm{f_A}}^d$. Since the sensitivity of ${\mathrm{f_A}}^d$ is 2^d, by Lemma 34, the best bound provable by Khrapchenko's method, Koutsoupias' method, and Håstad's lemma is $4^d = N$.

It is natural to ask how tight this formula size bound is. The best upper bound we can show on the formula size of ${\mathrm{f_A}}^d$ is 10^d.

Proposition 42 $\mathsf{L}({\mathrm{f_A}}^d) \leq 10^d$

Proof: It can be easily verified that the following formula of size 10 computes the base function $\mathrm{f_A}$:

$$(\neg x_1 \vee x_3 \vee \neg x_4) \wedge \\ ((\neg x_1 \wedge x_3 \wedge x_4) \vee ((x_1 \vee \neg x_2) \wedge (x_2 \vee \neg x_3))) .$$

This formula was found by computer search. The claim now follows from Proposition 39. □

7.3 Collision problem

In this section we look at the collision problem. This is a promise problem, where for an alphabet Σ the inputs $x = x_1 x_2 \dots x_n \in \Sigma^n$ satisfy one of the following conditions:

- All x_i are different
- For each i there exists exactly one $j \neq i$ such that $x_i = x_j$.

Those inputs satisfying the first condition are positive inputs and those satisfying the second condition are negative. An optimal lower bound for the quantum query complexity of $\Omega(n^{1/3})$ has been given by Aaronson and Shi [AS04]. We now show that the quantum adversary method cannot give better than a constant bound for this problem.

Lemma 43 $\mathsf{sumPI}(\mathrm{f_C}) \leq 2$

Proof: We demonstrate a set of probability distributions p_x such that for any positive instance x and negative instance y we have

$$\sum_{\substack{i \\ x_i \neq y_i}} \sqrt{p_x(i) p_y(i)} \geq 1/2.$$

The upper bound then follows.

Our probability distribution is very simple: for every x, let $p_x(i)$ be the uniform distribution over $[n]$. Any positive and negative instance must disagree in at least $n/2$ positions, thus

$$\sum_{\substack{i \\ x_i \neq y_i}} \sqrt{p_x(i) p_y(i)} \geq \frac{n}{2} \sqrt{\frac{1}{n} \frac{1}{n}} = \frac{1}{2}.$$

□

On the other hand, $\mathsf{maxPI}(\mathrm{f_C}) \geq \sqrt{n/2}$. As there is an upper bound for the collision problem of $O(n^{1/3})$ by Brassard, Høyer, Tapp [BHT97], this also shows that in general $\mathsf{maxPI}(f)$ is not a lower bound on the quantum query complexity of f.

Lemma 44 $\mathsf{maxPI}(\mathrm{f_C}) = \Theta(\sqrt{n})$

Proof: For the upper bound: On every positive instance x, where all x_i are different, we put the uniform distribution over $i \in [n]$; for a negative instance y we put probability $1/2$ on the first position, and probability $1/2$ on the position j such that $y_1 = y_j$. As $y_1 = y_j$, any positive instance x must differ from y on position 1 or position j (or both). Thus $\max_{i, x_i \neq y_i} p_x(i) p_y(i) \geq 1/2n$ and $\mathsf{maxPI}(\mathrm{f_C}) \leq \sqrt{2n}$.

Now for the lower bound. Fix a set of probability distributions p_x. Let x be any positive instance. There must be at least $n/2$ positions i satisfying $p_x(i) \leq 2/n$. Call this set of positions I. Now consider a negative instance y of where $y_j = x_j$ for all $j \notin I$, and y is assigned values in I in an arbitrary way so as to make it a negative instance. For this pair x, y we have $\max_i \sqrt{p_x(i) p_y(i)} \leq \sqrt{2/n}$, thus $\mathsf{maxPI}(\mathrm{f_C}) \geq \sqrt{n/2}$. □

The following table summarizes the bounds from this section.

Function	Input size	sum PI	Q_ϵ	max PI	L	$s_0 s_1$
$\mathsf{R{-}MAJ}_3^h$	$N = 3^h$	$2^h \approx N^{0.63}$	$\Omega(N^{0.63})$	$N^{0.63}$	$\Omega(N^{1.26})$, $O(N^{1.46})$	$N^{1.26}$
$\mathrm{f_A}^h$	$N = 4^h$	$2.5^h \approx N^{0.66}$	$\Omega(N^{0.66})$ [Amb03]	$\leq 3^h \approx N^{0.69}$	$\Omega(N^{1.32})$, $O(N^{1.66})$	N
$\mathrm{f_C}$	N	2	$\Theta(N^{1/3})$	$\Theta(\sqrt{N})$	$\perp$	$\perp$

8 Conclusions and open problems

An outstanding open problem is whether the square of the quantum query complexity lower bounds the formula size. We have given some support to this conjecture by showing it is true for one of the two main techniques of proving lower bounds on quantum query complexity. A simpler problem than the above might be to show the same is true of approximate polynomial degree, the other main lower bound technique for quantum query complexity.

We have seen that many formula size techniques in the literature can be viewed as clever ways of defining a subadditive measure on rectangles. In the search for better formula size lower bounds, it would be interesting to find other such measures; perhaps of particular interest are measures which rely on the disjointness condition for subadditivity, as the spectral norm squared does. Another example of a matrix norm which is subsquare additive on disjoint rectangles is the Frobenius norm, which has also been applied towards communication complexity theoretic ends as in Theorem 32. Let $\sigma_1(A) \geq \cdots \geq \sigma_n(A)$ denote the singular values of A. Noticing that $\|A\|_2^2 = \sigma_1(A)^2$ and $\|A\|_F^2 = \sigma_1(A)^2 + \ldots + \sigma_n(A)^2$, entices us to make the following conjecture:

Conjecture 45 *Let A be a matrix over $X \times Y$ with $n = \min\{|X|, |Y|\}$ and let $\mathcal{R}$ be a rectangle partition of $X \times Y$. Then for any $1 \leq k \leq n$*

$$\sum_{i=1}^{k} \sigma_i^2(A) \leq \sum_{R \in \mathcal{R}} \sum_{i=1}^{k} \sigma_i^2(A_R) \tag{9}$$

Recently, Troy Lee [Lee05] has shown that the conjecture is true for "tree-like" rectangle decompositions $\mathcal{R}$, that is for rectangle decompositions arising from communication protocols. Thus, in particular, in the spectral formulation of sumPI^2, one can replace the spectral norm squared with $\sum_{i=1}^{k} \sigma_i^2(A)$ for any k, and the resulting quantity also lower bounds formula size.

We have seen that the quantum adversary method breaks through the "Hamming distance 1" barrier and subsumes several previous formula size methods, in some cases giving provably stronger lower bounds on formula size. One question remaining is the relationship between sumPI^2 and the technique of Karchmer, Kushilevitz, and Nisan described in Theorem 33. In all the examples we know of Theorem 33 gives lower bounds at least as large as sumPI^2.

Acknowledgments

We would like to thank Frédéric Magniez, Robert Špalek, and Ronald de Wolf for helpful discussions. We also wish to thank Ryan O'Donnell for suggesting to look at the recursive majority of three function, and Xiaomin Chen for help in programming. Finally, we thank the anonymous referees for many exposition improving suggestions.

References

[Aar04] S. Aaronson. Lower bounds for local search by quantum arguments. In *Proceedings of the 36th annual ACM Symposium on Theory of Computing*, pages 465–474, 2004.

[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 44th IEEE Symposium on Foundations of Computer Science*, pages 230–239, 2003.

[And87] A. E. Andreev. On a method for obtaining more than quadratic effective lower bounds for the complexity of Π-schemes. *Moscow Univ. Math. Bull.*, 42(1):63–66, 1987.

[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.

[BBBV97] C.H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journl on Computing*, 26:1510–1523, 1997.

[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.

[BHT97] G. Brassard, P. Høyer, and A. Tapp. Quantum algorithm for the collision problem. *ACM SIGACT News (Cryptology column)*, 28:14–19, 1997.

[Bop89] R. Boppana. Amplification of probabilistic boolean formulas. *Advances in Computing Research*, 5(4):27–45, 1989.

[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.

[BW02] H. Buhrman and R. de Wolf. Complexity measures and decision tree complexity: A survey. *Theoretical Computer Science*, 288:21–43, 2002.

[DZ97] M. Dubiner and U. Zwick. Amplification by read-once formulas. *SIAM J. Comput.*, 26(1):15–38, 1997.

[Hås98] J. Håstad. he shrinkage exponent of de morgan formulae is 2. *SIAM Journal on Computing*, 2719(1):48–64, 1998.

[HJ99] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1999.

[JKS03] T. Jayram, R. Kumar, and D. Sivakumar. Two applications of information complexity. In *Proceedings of the 35th ACM Symposium on the Theory of Computing*, pages 673–682, 2003.

[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.

[Kla04] H. Klauck. One-way communication complexity and the Nečiporuk lower bound on formula size. Technical Report 0111062, cs.CC arXiv, 2004.

[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 STOC*, pages 539–550, 1988.

[KW03] I. Kerenidis and R. de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing of the thirty-fifth annual ACM symposium on Theory of computing*, pages 106—115, 2003.

[LM04] S. Laplante and F. Magniez. Lower bounds for randomized and quantum query complexity using kolmogorov arguments. In *Proceedings of the Nineteenth Annual IEEE Conference on Computational Complexity*, pages 294–304, 2004.

[Lee05] T. Lee. Kolmogorov complexity and formula size lower bounds. PhD thesis, University of Amsterdam. To appear.

[LV97] M. Li and P. Vitányi. An introduction to Kolmogorov complexity and its applications. In *Graduate Texts in Computer Science*. Springer, 1997. Second edition.

[MO03] E. Mossell and R. O'Donnell. On the noise sensitivity of monotone functions. *Random Structures and Algorithms*, 23(3):333–350, 2003.

[NC00] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Nis91] N. Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, 1991.

[NS94] N. Nisan and M. Szegedy. On the degree of boolean functions as real polynomials. *Computational Complexity*, 4:301–313, 1994.

[O'D02] R. O'Donnell. Hardness amplification within NP. In *Proceedings of the 34th ACM Symposium on the Theory of Computing*, pages 751–760. ACM, 2002.

[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.

[Sch83] U. Schürfeld. New lower bounds on the formula size of Boolean functions. *Acta Informatica*, 19(2):183–194, 1983.

[Spi71] P. Spira. On time-hardware complexity tradeoffs for Boolean functions. In *Proceedings of the 4th Hawaii Symposium on System Sciences*, pages 525–527. Western Periodicals Company, North Hollywood, 1971.

[ŠS05] R. Špalek and M. Szegedy. All quantum adversary methods are equivalent. In *Proceedings of the 32th International Colloquium On Automata, Languages and Programming*, volume 3580 of *Lecture Notes in Computer Science*, pages 1299–1311. Springer-Verlag, 2005. quant-ph/0409116.

[SV01] P. Sen and S. Venkatesh. Lower bounds in the quantum cell probe model lower bounds in the quantum cell probe model lower bounds in the quantum cell probe model. In *Proceedings of Intenational Colloquium on Automata, Languages and Programming (ICALP)*, pages 358–369, 2001.

[Sze03] M. Szegedy. An $O(n^{1.3})$ quantum algorithm for the triangle finding problem. Technical report, 2003. quant-ph/0310134.

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

[Zha05] S. Zhang. On the power of Ambainis's lower bounds. *Theoretical Computer Science*, 339(2–3):241–256, 2005.