Query Complexity in Expectation

Jedrzej Kaniewski[1,2], Troy Lee[1,3], and Ronald de Wolf[4,5](✉)

[1] Centre for Quantum Technologies,
National University of Singapore, Singapore, Singapore
[2] QuTech, Delft University of Technology, Delft, The Netherlands
[3] School of Physical and Mathematical Sciences, NTU, Singapore, Singapore
[4] Centrum Wiskunde and Informatica, Amsterdam, The Netherlands
[5] University of Amsterdam, Amsterdam, The Netherlands
j.kaniewski@nus.edu.sg, troyjlee@gmail.com, rdewolf@cwi.nl

Abstract. We study the query complexity of computing a function $f : \{0,1\}^n \to \mathbb{R}_+$ *in expectation*. This requires the algorithm on input x to output a nonnegative random variable whose expectation equals $f(x)$, using as few queries to the input x as possible. We exactly characterize both the randomized and the quantum query complexity by two polynomial degrees, the nonnegative literal degree and the sum-of-squares degree, respectively. We observe that the quantum complexity can be unboundedly smaller than the classical complexity for some functions, but can be at most polynomially smaller for Boolean functions. These query complexities relate to (and are motivated by) the extension complexity of polytopes. The *linear* extension complexity of a polytope is characterized by the randomized *communication* complexity of computing its slack matrix in expectation, and the *semidefinite* (psd) extension complexity is characterized by the analogous quantum model. Since query complexity can be used to upper bound communication complexity of related functions, we can derive some upper bounds on psd extension complexity by constructing efficient quantum query algorithms. As an example we give an exponentially-close entrywise approximation of the slack matrix of the perfect matching polytope with psd-rank only $2^{n^{1/2+\varepsilon}}$. Finally, we show randomized and quantum query complexity in expectation corresponds to the Sherali-Adams and Lasserre hierarchies, respectively.

1 Introduction

We study the complexity of computing a function $f : \{0,1\}^n \to \mathbb{R}_+$ *in expectation*, where our algorithm on input x should output a nonnegative real number whose expectation (over the algorithm's internal randomness) exactly equals $f(x)$. Getting the expectation right is easier than computing the function value $f(x)$ itself, and suffices in some applications. Suppose we want to approximate $F(x) = \sum_{i=1}^{m} f_i(x)$ that depends on $x \in \{0,1\}^n$. Then we can just compute each $f_i(x)$ *in expectation* and output the sum of the results. By linearity of expectation, the output will have expectation $F(x)$, and it will be tightly concentrated around its expectation if the random variables are not too wild (so the

M.M. Halldórsson et al. (Eds.): ICALP 2015, Part I, LNCS 9134, pp. 761–772, 2015.
DOI: 10.1007/978-3-662-47672-7_62

Central Limit Theorem applies). It is not necessary to compute or even approximate any of the values $f_i(x)$ themselves for this. This illustrates that computing functions in expectation is an interesting model in its own right. Additionally, it is motivated by connections with the *extension complexity* of polytopes that are used in combinatorial optimization (roughly: the minimal size of linear or semidefinite programs for optimizing over such a polytope), as described below.

The complexity of computing f can be measured in different ways, and here we will focus on *query* complexity. We measure the complexity of computing a function in expectation by the (worst-case) number of queries to the input $x \in \{0,1\}^n$ that the best algorithm uses. We study both *randomized* and *quantum* versions of this model and show that both of these query complexities can be exactly characterized by natural notions of polynomial degree. In Section 3 we show that the randomized query complexity of computing f in expectation equals the "nonnegative literal degree" of f, which is the minimal d such that f can be written as a nonnegative linear combination of products of up to d variables or negations of variables. In Section 4 we show that the quantum complexity equals the "sum-of-squares degree", which is the minimal d such that there exist polynomials p_i of degree at most d satisfying $f(x) = \sum_i p_i(x)^2$ for all $x \in \{0,1\}^n$.

In Section 5 we observe that quantum and classical query complexities (equivalently: the above two types of polynomial degree) can be arbitrarily far apart. For example, the function $f(x) = (\sum_{i=1}^n x_i - 1)^2$ is the square of a degree-1 polynomial and hence computable in expectation with only 1 quantum query, while randomized algorithms need n queries to get this expectation right. In contrast, we show that for functions with range $\{0,1\}$ the gap can be at most cubic.

Lower bounds on the quantum query complexity can be obtained from lower bounding the sum-of-squares degree of the function at hand, which is often non-trivial. Using techniques from approximation theory, we prove that $f(x) = (\sum_{i=1}^n x_i - 1)(\sum_{i=1}^n x_i - 2)$ has sum-of-squares degree $\Omega(\sqrt{n})$. Hence quantum algorithms require $\Omega(\sqrt{n})$ queries to compute this function in expectation.

Our main motivation for studying query complexity in expectation comes from combinatorial optimization, in particular from linear and semidefinite programs. Many optimization problems can be formulated as maximizing or minimizing a linear function over a polytope. For example, in the Traveling Salesman Problem on n-vertex undirected graphs, one wants to minimize a linear function (the length of the tour) over the polytope $P \subseteq \mathbb{R}^{\binom{n}{2}}$ that is the convex hull of all Hamiltonian cycles in the complete n-vertex graph K_n. Representing this polytope as the feasible region of a small linear or semidefinite program would allow us to efficiently solve the problem using the ellipsoid or interior-point methods.

Informally, the **linear extension complexity** of a polytope $P \subseteq \mathbb{R}^d$ is the minimum number of linear inequalities (over the d variables of P and possibly auxiliary variables) whose feasible region projects down to P. Small linear extension complexity means there is a small linear program to optimize over P.

Motivated by erroneous claims [33] that the TSP polytope had polynomial linear extension complexity (implying P = NP), Yannakakis [36] showed that "symmetric" linear extensions of the Traveling Salesman Polytope need $2^{\Omega(n)}$

linear inequalities. He showed the same for the perfect matching polytope (which is spanned by all perfect matchings in K_n), despite the fact that finding a maximum matching can be done efficiently! For a long time, generalizing these lower bounds to arbitrary (possibly non-symmetric) linear extensions was an open question. However, recently Fiorini et al. [15] proved a $2^{\Omega(n^{1/2})}$ lower bound on the linear extension complexity of the TSP polytope. Subsequently Rothvoß [30] proved a $2^{\Omega(n)}$ lower bound for the perfect matching polytope, which via a reduction implies the same bound for TSP. Chan et al. [10] obtained lower bounds on linear extension complexity for constraint satisfaction problems via a different route: roughly put, they showed that arbitrary linear extensions are not much more powerful than the specific linear extensions produced by the "Sherali-Adams Hierarchy"; hence they could obtain lower bounds on linear extension complexity from known bounds on the Sherali-Adams hierarchy.

The **positive semidefinite (psd) extension complexity** of polytope P, which replaces the linear programs by potentially more powerful semidefinite programs, is the minimal dimension of a semidefinite program whose feasible region projects down to P. In contrast to the case of linear extension complexity, very few lower bounds on psd extension complexity are known. Until recently, there were only a few lower bounds for "symmetric" psd extensions [14,24]. However, in a *very* recent breakthrough, Lee et al. [23] generalized the approach of [10] to show that arbitrary psd extensions are not much more powerful than the specific psd extensions produced by the "Lasserre Hierarchy". In particular they showed that the TSP polytope has psd extension complexity $2^{\Omega(n^{1/13})}$.

Surprisingly, there is a very close connection between these extension complexities and the model of computing functions in expectation, albeit for the *communication complexity* of computing a 2-input function. More precisely, suppose Alice receives input x, Bob receives input y, and they want to compute some function $g(x, y)$ (which may also be viewed as a matrix). In the usual setting of communication complexity [20], one of the parties (let's say Bob) has to output this value $g(x, y)$ exactly, either with probability 1 or with high probability. However, we may also consider how much communication they need to compute $g(x, y)$ *in expectation*, i.e., now Bob needs to output a nonnegative random variable whose expected value equals $g(x, y)$. Faenza et al. [13] showed that the logarithm of the linear extension complexity of a polytope P equals the randomized communication complexity of computing (in expectation) a matrix associated with P, known as the *slack matrix*. Lifting this result to the quantum/psd case, Fiorini et al. [15] showed that the logarithm of the *psd* extension complexity equals the one-way *quantum* communication complexity of computing the slack matrix of P in expectation; in this model Alice sends a single quantum message to Bob. These connections show that studying (linear and psd) extension complexity of a polytope P is *equivalent* to studying (randomized and one-way quantum) communication complexity in expectation, of the slack matrix of P.

How is the *query* complexity of computing a function in expectation related to this *communication* complexity? Many functions of interest in communication complexity are of the form $g(x, y) = f(x \wedge y)$ for some Boolean function

$f : \{0,1\}^n \to \{0,1\}$, where the AND-connective is applied bitwise. Functions of this form also arise as (submatrices of) slack matrices of interesting polytopes, e.g. the correlation polytope. Quite generally across the usual models of worst-case complexity (deterministic, randomized or quantum) upper bounds on the *query complexity* of f imply upper bounds on the *communication complexity* of g. In Section 7 we show that this also holds for the randomized and quantum models of computing a function in expectation. As this leads to multi-round communication protocols, it implies that the one-way and two-way quantum communication complexity of computing a function in expectation are equal.

In Section 7.1 we give an application of the connection between query algorithms and communication complexity (equivalently, *psd rank*), by deriving an exponentially-close entrywise approximation of the slack matrix S of the perfect matching polytope with psd rank $2^{n^{1/2+\varepsilon}}$. This psd rank is surprisingly low in view of the fact that Rothvoß [30] showed that the nonnegative rank of S is $2^{\Omega(n)}$, and Braun and Pokutta [5] showed that any $\tilde{S}$ that is $O(1/n)$-close to S still needs nonnegative rank $2^{\Omega(n)}$. This result about approximating the slack matrix for matching in low psd rank, fits in a recent line of non-quantum results derived using tools and techniques from quantum information theory (see [11]).

Communication protocols derived from query algorithms have a specific structure. In spirit, this is somewhat similar to looking at linear/psd extensions derived from hierarchies of specific linear or semidefinite programs like the Sherali-Adams and Lasserre hierarchies. In Section 2.3 we show these two relaxations actually correspond in a precise sense: just as the linear and psd extension complexities are characterized by models of communication complexity in expectation, the Sherali-Adams and Lasserre hierarchies are characterized by randomized and quantum models of query complexity in expectation, respectively. This follows from known characterizations of these hierarchies in terms of polynomial degrees that exactly correspond to the ones considered here.

Remark: Due to space limitations, many of the proofs have been omitted from this version. These can be found in the longer version at `arXiv:1411.7280`.

2 Preliminaries

2.1 Polytopes and Extension Complexity

While most of this paper is about *query* complexity in expectation, much of it is motivated by (the hope to port our results to) *communication* complexity in expectation and its consequences for linear and semidefinite extension complexity of polytopes. Hence we start with the latter. A polytope $P \subseteq \mathbb{R}^d$ has both an *inner description* as the convex hull of a set $V \subseteq \mathbb{R}^d$ of points, $P = \text{conv}(V)$; and an *outer description* as the intersection of halfspaces, $P = \{x \in \mathbb{R}^d : Ax \leq b\}$. A *slack matrix* integrates information from these two descriptions:

Definition 1. *Let* $P = \text{conv}(V) = \{x : Ax \leq b\}$ *be a polytope. The slack matrix* M *of* P *has columns labeled by* $v \in V$ *and rows labeled by constraints* $A_i x \leq b_i$*, with entries* $M(i,v) = b_i - A_i v$.

Definition 2. *Let M be a nonnegative matrix. A nonnegative factorization of M of size d consists of two sets of d-dimensional nonnegative vectors $\{a_x\}, \{b_y\}$ such that $M(x, y) = a_x^T b_y$ for all x, y. The nonnegative rank of M, denoted $\mathrm{rk}_+(M)$, is the minimal size among all nonnegative factorizations of M. Equivalently, it is the minimum number of nonnegative rank-one matrices whose sum is M.*

Definition 3. *Let M be a nonnegative matrix. A psd factorization of M of size d consists of two sets of d-by-d psd matrices $\{A_x\}, \{B_y\}$ such that $M(x, y) = \mathrm{Tr}(A_x B_y)$ for all x, y. The psd rank of M, denoted $\mathrm{rk}_{\mathrm{psd}}(M)$, is the minimal size among all psd factorizations of M.*

A nonnegative factorization is a psd factorization by diagonal matrices.

The *linear extension complexity* of a polytope P is the minimum number of facets of a (higher-dimensional) polytope which projects to P. The *semidefinite (psd) extension complexity* of P is the minimum d such that an affine slice of the cone of d-by-d positive semidefinite matrices projects to P. These complexity measures can be captured in terms of the above notions of rank of a slack matrix:

Theorem 1 ([16,36]). *The linear extension complexity of a polytope P is the nonnegative rank of a slack matrix of P. The semidefinite (psd) extension complexity of P is the psd rank of a slack matrix of P.*

A polytope may have different slack matrices associated with it, depending on which inner and outer description are used. By Theorem 1 these slack matrices all have the same nonnegative and psd rank.

One of our targets is the correlation polytope: $\mathrm{COR}_n = \{xx^T : x \in \{0, 1\}^n\}$. Fiorini et al. [15] showed that lower bounds on the linear/semidefinite extension complexity of the correlation polytope imply lower bounds on several other polytopes of interest, including the Traveling Salesman Polytope. The next lemma from [28] gives a family of submatrices of the slack matrix of COR_n.

Lemma 1. *Let $p(z) = a + bz + cz^2$ be a single-variate degree-2 polynomial nonnegative on $\{0, 1, \ldots, n\}$. The matrix $M(x, y) = p(|x \wedge y|)$ for $(x, y) \in \{0, 1\}^n$ is a submatrix of a slack matrix for the correlation polytope COR_n.*

In Section 6 we consider the matrix $M(x, y) = (|x \wedge y| - 1)(|x \wedge y| - 2)$ and its associated query problem $f(x) = (|x| - 1)(|x| - 2)$, where $|x|$ is Hamming weight.

2.2 Polynomials

We will study two types of polynomials that are obviously nonnegative on the Boolean cube: nonnegative literal polynomials and sum-of-squares polynomials.

Definition 4 (nonnegative literal degree). *A nonnegative literal polynomial is a nonnegative linear combination of products of variables and negations of variables, i.e., it can be written as*

$$p(x) = \sum_{S \subseteq [n]} \sum_{b \in \{0,1\}^{|S|}} \alpha_{S,b} \prod_{i \in S} ((-1)^{b_i} x_i + b_i) \qquad (1)$$

where each $\alpha_{S,b} \geq 0$*. Its degree is* $\max\{|S| : \alpha_{S,b} \neq 0\}$*. The nonnegative literal degree of* $f : \{0,1\}^n \to \mathbb{R}_+$*, denoted* $\mathrm{ldeg}_+(f)$*, is the minimum degree of a nonnegative literal polynomial* p *that equals* f *on* $\{0,1\}^n$*.*

This measure has also been called the *nonnegative junta certificate degree* [23].

Definition 5 (sum-of-squares degree). *Let* d *be a natural number. A sum-of-squares polynomial of degree* d *is a polynomial* p *that can be written in the form* $p(x) = \sum_{i \in \mathcal{P}} p_i(x)^2$*, where* $\mathcal{P}$ *is a finite index set and the* p_i *are polynomials of degree* $\leq d$*. The sum-of-squares (sos) degree of* $f : \{0,1\}^n \to \mathbb{R}_+$*, denoted* $\deg_{sos}(f)$*, is the minimum* d *for which such a* p *equals* f *on* $\{0,1\}^n$*.*

Note that a sum-of-squares polynomial of degree d is actually a polynomial of degree $2d$; we allow this slight abuse of notation in order to give a clean characterization in Theorem 3 below.

2.3 The Sherali-Adams and Lasserre Hierarchies

Consider the optimization problem

$$\alpha(f) = \max_{x \in \{0,1\}^n} f(x) \tag{2}$$

where f is given by a multilinear polynomial. Many important optimization problems can be cast in this framework, including NP-hard ones. For example finding the maximum cut in a graph $G = (V, E)$ with n vertices corresponds to the quadratic function $f(x) = \sum_{\{i,j\} \in E} x_i(1 - x_j) + x_j(1 - x_i)$.

If $c \geq \alpha(f)$, then $c - f$ is nonnegative on $\{0,1\}^n$. One way we can witness this is by expressing $c - f$ as a polynomial which is obviously nonnegative for all $x \in \{0,1\}^n$. The *Sherali-Adams hierarchy* [31] looks for a witness in the form of a nonnegative literal polynomial. The sum-of-squares or *Lasserre hierarchy* looks for a witness in the form of a sum-of-squares polynomial [21,29,32].

If we can find a nonnegative literal polynomial p of degree d such that $c - f(x) = p(x)$, then this witnesses that the optimal value is upper bounded as $\alpha(f) \leq c$. Moreover, determining if the nonnegative literal polynomial degree of $c - f(x)$ is at most d can be formulated as a linear program of size $n^{O(d)}$. The value of the d-round Sherali-Adams relaxation for (2) is the smallest value of c such that $c - f(x)$ is a degree-d nonnegative literal polynomial. Thus the smallest d for which a Sherali-Adams relaxation certifies an *optimal* upper bound, is exactly the nonnegative literal degree $\mathrm{ldeg}_+(\alpha(f) - f)$ of the function $\alpha(f) - f$.

Similarly, if we can find $p_i : \{0,1\}^n \to \mathbb{R}$ of degree at most d, such that $c - f(x) = \sum_i p_i(x)^2$, then this witnesses that $\alpha(f) \leq c$. Searching for such polynomials p_i can be expressed as a semidefinite program of size $n^{O(d)}$. The smallest value of c such that $c - f$ is degree-d sum-of-squares is known to be equivalent to the relaxation of (2) given by the d^{th} level of the Lasserre hierarchy. The level of the Lasserre hierarchy required to exactly capture (2) is thus $\deg_{sos}(\alpha(f) - f)$.

3 Randomized Query Complexity in Expectation

In this section we define and characterize classical randomized query complexity in expectation, characterize it by the nonnegative literal degree, and relate it to the Sherali-Adams hierarchy. A *randomized decision tree* is a probability distribution μ over deterministic decision trees. We consider deterministic decision trees with leaves labeled by nonnegative real numbers. A randomized decision tree computes a function $f : \{0,1\}^n \to \mathbb{R}_+$ *in expectation* if for every $x \in \{0,1\}^n$ the expected output of the tree on input x is $f(x)$. The *cost* of such a tree is, as usual, the maximum cost, that is the length of a longest path from the root to a leaf, of a deterministic decision tree that has nonzero μ-probability.

Definition 6. *The randomized query complexity of computing f in expectation, denoted* $\mathrm{RE}(f)$*, is the minimum cost among all randomized decision trees that compute f in expectation.*

Theorem 2. *Let $f : \{0,1\}^n \to \mathbb{R}_+$. Then* $\mathrm{RE}(f) = \mathrm{ldeg}_+(f)$*.*

Referring back to Section 2.3, this gives a connection between randomized query complexity in expectation and the Sherali-Adams hierarchy: the smallest d such that the d-round Sherali-Adams relaxation certifies the optimal upper bound $\alpha(f)$ on the maximization problem (2), is exactly $\mathrm{RE}(\alpha(f) - f)$.

4 Quantum Query Complexity in Expectation

Here we study *quantum* query complexity in expectation, characterize it by sum-of-squares degree, and relate it to the Lasserre hierarchy. We assume familiarity with quantum computing [27] and query complexity [9].

We define the quantum query complexity of computing a function $f : \{0,1\}^n \to \mathbb{R}_+$ in expectation. A T-query algorithm is described by unitaries $U_0, \dots, U_T$ and a final POVM measurement $\{E_\theta\}_{\theta \in \Theta}$, where each E_θ is a psd matrix labeled by nonnegative real θ, and $\sum_{\theta \in \Theta} E_\theta = I$. As usual, on input x the query algorithm proceeds from the initial state $|\bar{0}\rangle$ by alternately applying a unitary and the query oracle O_x (which maps $|i, b\rangle \mapsto |i, b \oplus x_i\rangle$), so that the final state of the algorithm after T queries is $|\psi_x^T\rangle = U_T O_x \dots O_x U_1 O_x U_0 |\bar{0}\rangle$. Let $E = \sum_{\theta \in \Theta} \theta E_\theta$. As the probability of output θ upon measuring $|\psi_x^T\rangle$ is $\mathrm{Tr}(E_\theta |\psi_x^T\rangle\langle\psi_x^T|)$, the expected value of the output is $\mathrm{Tr}(E|\psi_x^T\rangle\langle\psi_x^T|)$. The algorithm *computes f in expectation* if $f(x) = \mathrm{Tr}(E|\psi_x^T\rangle\langle\psi_x^T|)$ for every $x \in \{0,1\}^n$.

Definition 7. *The quantum query complexity of computing f in expectation, denoted* $\mathrm{QE}(f)$*, is the minimum T for which there is a T-query quantum algorithm computing f in expectation.*

Theorem 3. *Let $f : \{0,1\}^n \to \mathbb{R}_+$. Then* $\mathrm{QE}(f) = \deg_{sos}(f)$*.*

Proof. $\mathrm{QE}(f) \geq \deg_{sos}(f)$. Say there is a T-query algorithm to compute f in expectation. Let $|\psi_x^T\rangle$ denote its state on input x after T queries. By the polynomial method [2], the amplitude of each basis state in $|\psi_x^T\rangle$ is an n-variate multilinear polynomial in x of degree $\leq T$. We have $f(x) = \sum_\theta \theta\langle\psi_x^T|E_\theta|\psi_x^T\rangle$. Let $E_\theta = \sum_i \lambda_i |e_\theta^i\rangle\langle e_\theta^i|$ be the eigenvalue decomposition of E_θ, where each $\lambda_i \geq 0$. Then $\langle\psi_x^T|E_\theta|\psi_x^T\rangle = \sum_i \lambda_i |\langle\psi_x^T|e_\theta^i\rangle|^2$. Since $\langle\psi_x^T|e_\theta^i\rangle$ is a linear combination of amplitudes of $|\psi_x^T\rangle$, it is a degree $\leq T$ polynomial in x. Since the coefficients θ and λ_i are nonnegative, this gives a representation of $\langle\psi_x^T|E_\theta|\psi_x^T\rangle$ as a sum-of-squares polynomial of degree $\leq T$.

$\mathrm{QE}(f) \leq \deg_{sos}(f)$. Let $d = \deg_{sos}(f)$. We first exhibit a quantum algorithm for the special case where $f = p^2$ for some degree-d polynomial p. This is inspired by the proof of [35, Theorem 2.3]. Let $p = \sum_s \widehat{p}(s)(-1)^{x \cdot s}$ be the Fourier representation of p, where s ranges over $\{0,1\}^n$. Because p has degree d, we have $\widehat{p}(s) \neq 0$ only if $|s| \leq d$. The algorithm is as follows:

1. Prepare n-qubit state $c\sum_s \widehat{p}(s)|s\rangle$, where $c = 1/\sqrt{\sum_s \widehat{p}(s)^2}$ is a constant.
2. Apply a unitary that maps $|s\rangle \mapsto (-1)^{x \cdot s}|s\rangle$ for all s of weight $|s| \leq d$; one can show that this can be implemented using d queries.
3. Apply the n-qubit Hadamard transform to the state.
4. Measure the state and output $2^n/c^2$ if the result was 0^n, otherwise output 0.

Note that the amplitude of the basis state $|0^n\rangle$ after step 3 is $\frac{c}{\sqrt{2^n}}\sum_s \widehat{p}(s)(-1)^{x \cdot s} = \frac{c}{\sqrt{2^n}}p(x)$. Hence the probability that the final measurement results in outcome 0^n is $(\frac{c}{\sqrt{2^n}}p(x))^2$, and the expected value of the output is $(\frac{c}{\sqrt{2^n}}p(x))^2 \cdot 2^n/c^2 = p(x)^2 = f(x)$, as desired. Now consider the general case where $f = \sum_{i \in \mathcal{P}} p_i^2$. The algorithm chooses one $i \in \mathcal{P}$ uniformly at random and runs the above algorithm to produce an output with expected value $p_i(x)^2$. It finally outputs that output multiplied by $|\mathcal{P}|$. Clearly, this uses at most d queries to x, and the expected value of its final output is $\frac{1}{|\mathcal{P}|}\sum_i p_i(x)^2|\mathcal{P}| = \sum_i p_i(x)^2 = f(x)$. □

This connects quantum query complexity in expectation and the Lasserre hierarchy: the smallest level d of the Lasserre hierarchy that certifies the optimal upper bound $\alpha(f)$ on the maximization problem (2), is exactly $\mathrm{QE}(\alpha(f) - f)$.

5 Gaps and Relations between RE(*f*) and QE(*f*)

For some $f : \{0,1\}^n \to \mathbb{R}_+$, the quantum query complexity in expectation $\mathrm{QE}(f)$ can be *much* smaller than its classical counterpart $\mathrm{RE}(f)$. An extreme example is the n-bit function $f(x) = (|x| - 1)^2$, where $\mathrm{QE}(f) = 1$ by Theorem 3, but $\mathrm{RE}(f) = n$. The latter holds because on the all-0 input the algorithm needs to produce a nonzero output with positive probability, but on weight-1 inputs it can never output anything nonzero, hence a classical algorithm needs n queries on the all-0 input. In contrast, if the range of f is Boolean, then we can show that $\mathrm{QE}(f)$ is at most polynomially smaller than $\mathrm{RE}(f)$:

Theorem 4. *For every* $f : \{0,1\}^n \to \{0,1\}$ *we have* $\mathrm{RE}(f) \leq 16\mathrm{QE}(f)^3$.

The main reason this query complexity result is interesting is that the analogous statement for *communication* complexity is equivalent to the longstanding log-rank conjecture! The communication version of Theorem 4 would say that for all *Boolean* matrices M, the quantum and classical communication complexity of computing M in expectation are at most polynomially far apart. As noted by Fiorini et al. [15], this is equivalent to $\log \mathrm{rk}_+(M) \leq \mathrm{polylog}(\mathrm{rk}_{\mathrm{psd}}(M))$, which in turn is equivalent to the log-rank conjecture. Presumably such a communication version will be substantially harder to prove than the above query version. However, in many cases results in query complexity "mirror" (often much harder) results in communication complexity, so our Theorem 4 may be viewed as (weak) evidence for the log-rank conjecture.

6 A Quantum Query Complexity Lower Bound

Here we show that the n-bit function $f(x) = (|x|-1)(|x|-2)$ has $\mathrm{QE}(f) = \Omega(\sqrt{n})$. This result is motivated by the fact that a strong lower bound on the psd rank of the closely related matrix $M(x, y) = (|x \wedge y|-1)(|x \wedge y|-2)$ would have important consequences for the correlation polytope (M is a submatrix of the slack matrix for the correlation polytope, see Lemma 1). We hope that the methods of this section may in the future help lower bound this psd rank as well.

We prove our query complexity lower bound by showing the corresponding lower bound on the sum-of-squares degree of f. As is common in query complexity lower bounds by the polynomial method [2], we will use a symmetrization argument to define a single-variate polynomial $Q : \mathbb{R} \to \mathbb{R}$ that behaves well on $[n]$, and then use Markov's lemma from approximation theory to bound the degree of Q. A new complication in our setting is the following. If $f(x) = \sum_i p_i(x)^2$ then we would like to define a "symmetrized" polynomial $g : [n] \to \mathbb{R}$ where $g(k) = \mathbb{E}_{x:|x|=k} \left[\sum_i p_i(x)^2\right]$. However, we do not know how to prove that g remains a nonnegative polynomial. To get around this, we define symmetrized polynomials $q_i(k) = \mathbb{E}_{x:|x|=k} \left[p_i(x)\right]$ for each p_i individually, then recombine the symmetrized polynomials as $Q(k) = \sum_i q_i(k)^2$. We are then able to bound the sum-of-squares degree of Q.

Theorem 5. *If* $f(x) = (|x| - 1)(|x| - 2)$ *for* $x \in \{0, 1\}^n$, $\deg_{sos}(f) \geq \sqrt{n/48}$.

7 Psd Rank and Query Complexity in Expectation

Fiorini et al. [15] defined a *one-way* model of quantum communication to compute a matrix in expectation, and showed that this complexity is characterized by the logarithm of the psd rank. We show below that this characterization still holds for the more general *two-way* communication model, which allows multiple rounds of communication. Hence one-way and two-way quantum communication complexity are the same for computation in expectation.

We will not formally define the model of two-way quantum communication complexity (see [34] for more technical details), instead just highlighting the differences of the model of computing a function in expectation to the normal model. As usual, Alice and Bob each start with their own input, x and y respectively, and then the protocol specifies whose turn it is to speak and what message they send to the other party. At the end of the protocol Bob must output a *nonnegative* number, which is a random variable z that depends on the inputs x and y as well as on the internal randomness of the protocol.

The major difference with the usual model is the notion of when a protocol is correct. Let M be a matrix with nonnegative real entries whose rows are indexed by Alice's possible inputs, and whose columns are indexed by Bob's inputs. We say a protocol *computes the matrix M in expectation* if, for every (x, y), $M(x, y)$ equals the expected value of the output z on input (x, y). As usual, the *cost* of the protocol is the worst-case number of qubits communicated (over all rounds).

Definition 8. *The quantum communication complexity of computing a matrix M in expectation, denoted $\mathrm{QCE}(M)$, is the minimum q such that there exists a quantum protocol of cost q that computes M in expectation. The minimum q when we restrict to one-way protocols is denoted $\mathrm{QCE}^1(M)$.*

It turns out that two-way quantum communication complexity is not more powerful than its one-way cousin: both correspond to the psd rank.

Theorem 6. $\log \mathrm{rk}_{\mathrm{psd}}(M) \leq \mathrm{QCE}(M) \leq \mathrm{QCE}^1(f) \leq \lceil \log(\mathrm{rk}_{\mathrm{psd}}(M) + 1) \rceil$.

7.1 Upper Bounds on psd Rank from Quantum Algorithms

We can show that efficient quantum query algorithms for computing functions $f : \{0,1\}^n \to \mathbb{R}_+$ in expectation give rise to an efficient quantum communication protocol to compute the matrix $M_f(x, y) = f(x \wedge y)$ in expectation, and hence to a low-rank psd factorization of M_f. We state it more generally:

Theorem 7. *Let Y be a finite set. For every $y \in Y$, let $f_y : \{0,1\}^n \to \mathbb{R}_+$ satisfy $\mathrm{QE}(f_y) \leq T$. Define a $2^n \times |Y|$ matrix M by $M(x, y) = f_y(x)$. Then $\mathrm{QCE}(M) \leq 2T(\log(n) + 1)$, and hence $\mathrm{rk}_{\mathrm{psd}}(M) \leq (2n)^{2T}$.*

Lee et al. [23] independently proved a similar upper bound on psd rank in terms of the sos-degree of f_y rather than quantum query complexity.

As an application we will derive an exponentially-close entrywise approximation of the slack matrix S of the perfect matching polytope, by a matrix with psd rank not much bigger than $2^{\sqrt{n}}$. This shows a big difference to the case of nonnegative rank: Braun and Pokutta [5] show that any $\tilde{S}$ that is $O(1/n)$-close to S needs nonnegative rank $2^{\Omega(n)}$.

Edmonds gave a complete description of the facets of the perfect matching polytope for the complete n-vertex graph K_n [12]. The key are the *odd-set* inequalities: for a perfect matching M, viewed as a vector $M \in \{0,1\}^{\binom{n}{2}}$ of weight $m = n/2$, and an odd-sized set $U \subseteq [n]$, the associated inequality says $|\delta(U) \cap M| \geq 1$, where $\delta(U) \in \{0,1\}^{\binom{n}{2}}$ denotes the cut induced by U. In addition, there are $O(n^2)$ degree and nonnegativity constraints. Thus the corresponding slack matrix S has columns indexed by all perfect matchings M in K_n and rows indexed by odd-sized sets U with entries $S_{UM} = |\delta(U) \cap M| - 1$. There are $O(n^2)$ additional rows for the degree and nonnegativity constraints.

In the full version of this paper we show that the m-bit function $g(z) = |z| - 1$ can be approximated (in expectation) up to exponentially small error with quantum query complexity $O(m^{1/2+\varepsilon} \log m)$. Define $f_M(x) = g(x_M)$, where x_M denotes the restriction of n-bit string x to the m positions in the support of M. Applying Theorem 7 and adding $O(n^2)$ rows for the other constraints gives:

Theorem 8. $\forall \varepsilon > 0$ *there is a matrix* $\tilde{S}$ *of psd rank* $2^{O(n^{1/2+\varepsilon}(\log n)^2)}$ *s.t.*

1. $S_{UM} - 2^{-(n/2)^{2\varepsilon}} \leq \tilde{S}_{UM} \leq S_{UM}$ *for the entries where* $|\delta(U) \cap M| > (n/2)^{2\varepsilon}$*;*
2. $\tilde{S}_{xy} = S_{xy}$ *for all other entries.*

Acknowledgments. We thank Srinivasan Arunachalam, David Steurer, Mario Szegedy and Henry Yuen for useful discussions, Sebastian Pokutta for useful discussions and for pointing us to [5], and James Lee for sending us a version of [23]. Troy Lee is supported in part by the Singapore National Research Foundation under NRF RF Award No. NRF-NRFF2013-13. Ronald de Wolf is partially supported by a Vidi grant from the Netherlands Organization for Scientific Research (NWO) which ended in 2013, ERC Consolidator Grant QPROGRESS, and by the European Commission IST STREP project QALGO 600700.

References

1. Arunachalam, S., Yuen, H., de Wolf, R.: Unpublished manuscript, August 2014
2. Beals, R., Buhrman, H., Cleve, R., Mosca, M., de Wolf, R.: Quantum lower bounds by polynomials. Journal of the ACM **48**(4), 778–797 (2001)
3. Blekherman, G., Gouveia, J., Pfeiffer, J.: Sums of squares on the hypercube, February 18, 2014. arXiv/1402.4199
4. Brassard, G., Høyer, P., Mosca, M., Tapp, A.: Quantum amplitude amplification and estimation. In: Quantum Computation and Quantum Information: A Millennium Volume, AMS Contemporary Mathematics Series, vol. 305, pp. 53–74 (2002)
5. Braun, G., Pokutta, S.: The matching polytope does not admit fully-polynomial size relaxation schemes. In: Proc. of 26th SODA, pp. 837–846 (2015)
6. Buhrman, H., Cleve, R., Wigderson, A.: Quantum vs. classical communication and computation. In: Proc. of 30th ACM STOC, pp. 63–68 (1998)
7. Buhrman, H., Cleve, R., de Wolf, R., Zalka, C.: Bounds for small-error and zero-error quantum algorithms. In: Proc. of 40th IEEE FOCS, pp. 358–368 (1999)
8. Buhrman, H., de Wolf, R.: Communication complexity lower bounds by polynomials. In: Proc. of 16th IEEE Complexity (CCC), pp. 120–130 (2001)
9. Buhrman, H., de Wolf, R.: Complexity measures and decision tree complexity: A survey. Theoretical Computer Science **288**(1), 21–43 (2002)
10. Chan, S.O., Lee, J.R., Raghavendra, P., Steurer, D.: Approximate constraint satisfaction requires large LP relaxations. In: Proc. of 54th IEEE FOCS, pp. 350–359 (2013)
11. Drucker, A., de Wolf, R.: Quantum proofs for classical theorems. Theory of Computing (2011). ToC Library, Graduate Surveys 2

12. Edmonds, J.: Maximum matching and a polyhedron with 0,1-vertices. Journal of research of the National Bureau of Standards-B **69B**(1,2), 125–130 (1965)
13. Faenza, Y., Fiorini, S., Grappe, R., Tiwary, H.R.: Extended formulations, nonnegative factorizations, and randomized communication protocols. In: Mahjoub, A.R., Markakis, V., Milis, I., Paschos, V.T. (eds.) ISCO 2012. LNCS, vol. 7422, pp. 129–140. Springer, Heidelberg (2012)
14. Fawzi, H., Saunderson, J., Parrilo, P.: Equivariant semidefinite lifts and sum-of-squares hierarchies, December 23, 2013. arXiv:1312.6662
15. Fiorini, S., Massar, S., Pokutta, S., Tiwary, H.R., de Wolf, R.: Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds. In: Proc. of 44th ACM STOC, pp. 95–106 (2012)
16. Gouveia, J., Parrilo, P., Thomas, R.: Lifts of convex sets and cone factorizations. Mathematics of Operations Research **38**(2), 248–264 (2013). arXiv:1111.3164
17. Grigoriev, D.: Complexity of Positivstellensatz proofs for the knapsack. Computational Complexity **10**, 139–154 (2001)
18. Grover, L.K.: A fast quantum mechanical algorithm for database search. In: Proc. of 28th ACM STOC, pp. 212–219 (1996). quant-ph/9605043
19. Kremer, I.: Quantum Communication. MSc thesis, Hebrew University (1995)
20. Kushilevitz, E., Nisan, N.: Communication complexity. Cambridge UP (1997)
21. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM Journal on Optimization **11**(3), 796–817 (2001)
22. Laurent, M.: Lower bound for the number of iterations in semidefinite hierarchies for the cut polytope. Mathematics of operations research **28**(4), 871–883 (2003)
23. Lee, J.R., Raghavendra, P., Steurer, D.: Lower bounds on the size of semidefinite programming relaxations, November 24, 2014. To appear in STOC 2015. arXiv:1411.6317
24. Lee, J.R., Raghavendra, P., Steurer, D., Tan, N.: On the power of symmetric LP and SDP relaxations. In: Proc. of 29th IEEE Complexity (CCC), pp. 13–21 (2014)
25. Midrijanis, G.: Exact quantum query complexity for total Boolean functions, March 23, 2004. quant-ph/0403168
26. Minsky, M., Papert, S.: Perceptrons. MIT Press (1987)
27. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press (2000)
28. Padberg, M.: The boolean quadric polytope. Math. prog. **45**, 139–172 (1989)
29. Parrilo, P.: Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. Ph.D. thesis, Caltech (2000)
30. Rothvoß, T.: The matching polytope has exponential extension complexity. In: Proc. of 46th ACM STOC, pp. 263–272 (2014)
31. Sherali, H.D., Adams, W.P.: A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming. SIAM Journal on Discrete Mathematics **3**, 411–430 (1990)
32. Shor, N.Z.: An approach to obtaining global extremums in polynomial mathematical programming problems. Cybernetics **23**, 695–700 (1987)
33. Swart, T.: P = NP. Tech. rep., University of Guelph (1986), revision 1987
34. de Wolf, R.: Quantum communication and complexity. Theoretical Computer Science **287**(1), 337–353 (2002)
35. de Wolf, R.: Nondeterministic quantum query and quantum communication complexities. SIAM Journal on Computing **32**(3), 681–699 (2003)
36. Yannakakis, M.: Expressing combinatorial optimization problems by linear programs. Journal of Computer and System Sciences **43**(3), 441–466 (1991)
37. Yao, A.C.C.: Quantum circuit complexity. In: Proc. of 34th IEEE FOCS, pp. 352–360 (1993)