Query Complexity: Worst-Case Quantum Versus Average-Case Classical

Scott Aaronson*

February 25, 2012

1 Introduction

In this note we investigate the relationship between worst-case quantum query complexity and average-case classical query complexity. Specifically, we show that if a quantum computer can evaluate a total Boolean function f with bounded error using T queries in the worst case, then a deterministic classical computer can evaluate f using $O(T^5)$ queries in the average case, under a uniform distribution of inputs. If f is monotone, we show furthermore that only $O(T^3)$ queries are needed.

Previously, Beals et al. [3] showed that if a quantum computer can evaluate f with bounded error using T queries in the worst case, then a deterministic classical computer can evaluate f using $O(T^6)$ queries in the worst case, or $O(T^4)$ if f is monotone. The optimal bound is conjectured to be $O(T^2)$, but improving on $O(T^6)$ remains an open problem. Relating worst-case quantum complexity to average-case classical complexity may suggest new ways to reduce the polynomial gap in the ordinary worst-case versus worst-case setting.

2 Preliminaries

Let $f : \{0,1\}^n \rightarrow \{0,1\}$ be a total Boolean function. Following [3], we let $D(f)$, $R_0(f)$, and $R_2(f)$ respectively denote the deterministic, zero-error, and bounded-error classical query complexities of f, and let $Q_E(f)$, $Q_0(f)$, and $Q_2(f)$ denote the corresponding quantum query complexities. We have:

- $n \geq D(f) \geq Q_E(f) \geq Q_0(f) \geq Q_2(f)$ and $n \geq D(f) \geq R_0(f) \geq R_2(f) \geq Q_2(f)$,
- $D(f) = O(R_2(f)^3)$ and $D(f) = O(R_0(f)^2)$ [5],
- $D(f) = O(Q_2(f)^6)$, or $O(Q_2(f)^4)$ if f is monotone [3], and
- $D(f) = O(Q_0(f)^4)$ [4].

Let μ be the uniform distribution over $\{0,1\}^n$. Following Ambainis and de Wolf [2], we let $D^\mu(f)$ be the average-case deterministic query complexity under the uniform distribution of inputs. (Note that in [2], μ can be non-uniform, whereas here it is always uniform.) The average-case bounded-error analogs of $D^\mu(f)$,

*217 West Avenue, Cornell University, Ithaca NY 14850. Email: sja8@cornell.edu.

$R_2^\mu(f)$ and $Q_2^\mu(f)$ in the classical and quantum settings respectively, can be super-exponentially smaller than $D^\mu(f)$ [2]. On the other hand, we have $D^\mu(f) = R_0^\mu(f)$ by Yao's minimax principle: viewing the questioner's choice of query algorithm and the oracle's choice of response algorithm in matrix-game terms, if the oracle is committed to a fixed randomized strategy (as it is in the average-case setting), then the questioner has nothing to gain by using randomization (assuming the questioner's goal is the same, namely to evaluate f with probability 1). Therefore we need not consider $R_0^\mu(f)$.

Here we show that $D^\mu(f) = O(Q_2(f)^5)$, or $O(Q_2(f)^3)$ if f is monotone. The proof has two components. Theorem 1 gives a deterministic classical algorithm for evaluating f with few queries in the average case, yielding an upper bound on $D^\mu(f)$. The theorem is a refinement of [3, Lemma 5.3], which gives an upper bound on $D(f)$. Theorem 2 gives a lower bound on $Q_2(f)$ in terms of the expected block sensitivity. The bound is obtained via the quantum adversary argument, which was recently introduced by Ambainis [1].

Given $X \in \{0,1\}^n$ and a block B of variables, let $X(B)$ be the input obtained from X by flipping the values of all the variables in B. Following [5, 3]:

- A 1-*certificate* is an assignment $C : B \to \{0,1\}$ of values to a block B of variables, such that $f(X) = 1$ whenever X is consistent with C. The size of C is $|B|$. A 0-*certificate* is defined similarly. The *certificate complexity* $C_X(f)$ of X is the size of the smallest $f(X)$-certificate that agrees with X. The certificate complexity $C(f)$ of f is the maximum of $C_X(f)$ over all X.

- The *block sensitivity* $\mathrm{bs}_X(f)$ of X is the maximum number b of disjoint blocks $B_1, \ldots, B_b$ of variables such that for all $1 \le i \le b$, $f(X) \ne f(X(B_i))$. The block sensitivity $\mathrm{bs}(f)$ of f is the maximum of $\mathrm{bs}_X(f)$ over all X.

Let $C_\mu(f) = E_\mu[C_X(f)]$ be the mean of $C_X(f)$ over all X. Likewise let $\mathrm{bs}_\mu(f) = E_\mu[\mathrm{bs}_X(f)]$ be the mean of $\mathrm{bs}_X(f)$ over all X. Let $\mathrm{bs}_\mu^{(1)}(f) = E_{\mu(1)}[\mathrm{bs}_X(f)]$ be the mean block sensitivity among X such that $f(X) = 1$, and let $\mathrm{bs}_\mu^{(0)}(f) = E_{\mu(0)}[\mathrm{bs}_X(f)]$ be the mean block sensitivity among X such that $f(X) = 0$.

3 Results

First we relate $D^\mu(f)$ to the mean block sensitivity $\mathrm{bs}_\mu(f)$, along the lines of [3, Lemma 5.3].

Theorem 1 $D^\mu(f) \le 2\,\mathrm{bs}_\mu(f) C(f)$.

Proof. Let a *satisfying* certificate be one that agrees with X; let a *consistent* certificate be one that agrees with the X-values queried so far. The following algorithm returns a satisfying 0-certificate in expected number of queries at most $\mathrm{bs}_\mu^{(0)}(f)C(f)$, *assuming* that $f(X) = 0$ (the expectation is over the uniform distribution of all X satisfying this condition).

> Choose a consistent 1-certificate and query those of its variables whose X-values are still unknown.
> Repeat until a satisfying 0-certificate is found.

Call this algorithm A_0. A_0 can be made deterministic by choosing certificates in some fixed lexicographic order. To see that A_0 always returns a satisfying 0-certificate, note that, for the special case $f(X) = 0$ that we're considering, A_0 reduces to Algorithm A of [3, Lemma 5.3]. A always returns a satisfying 0-certificate when $f(X) = 0$, therefore so does A_0.

It remains to show that the expected number of queries used by A_0 is at most $\mathrm{bs}_\mu^{(0)}(f)C(f)$. Suppose that, after A_0 has queried k 1-certificates, $C_1, \ldots, C_k$, no satisfying 0-certificate has yet been found. Then there exists a Y consistent with the bits queried so far such that $f(Y) = 1$. Furthermore, Y contains a satisfying 1-certificate C_{k+1}. We will derive from these C_i disjoint blocks $B_i \subseteq X$ such that f is sensitive to each B_i on X. For each $1 \leq i \leq k+1$, let B_i be the set of variables on which X and C_i disagree. Clearly each B_i is non-empty. Now, $X(B_i)$ agrees with C_i, therefore $f(X(B_i)) = 1$, so that f is sensitive to each B_i on X. Let v be a variable in some B_i; then $X^{(v)} = Y^{(v)} \neq C_i^{(v)}$. For $j > i$, C_j has been chosen consistent with all variables queried so far (including v), so we cannot have $X^{(v)} = Y^{(v)} \neq C_j^{(v)}$, hence $v \notin B_j$. Therefore all B_i and B_j are disjoint. It follows that k (the number of 1-certificates queried) can be at most $\mathrm{bs}(X)$.

Now, since the input is chosen uniformly at random among all X with $f(X) = 0$, the expectation of $\mathrm{bs}(X)$ is $\mathrm{bs}_\mu^{(0)}(f)$. Therefore A_0 returns a satisfying 0-certificate after querying an expected number of certificates at most $\mathrm{bs}_\mu^{(0)}(f)$, or after an expected total number of queries at most $\mathrm{bs}_\mu^{(0)}(f)C(f)$.

An analogous algorithm, A_1, returns a satisfying 1-certificate in expected number of queries at most $\mathrm{bs}_\mu^{(1)}(f)C(f)$, assuming that a 1-certificate exists (i.e. that $f(X) = 1$). Suppose that we interleave A_0 and A_1, alternating between the two until either A_0 or A_1 halts and returns a certificate, and that when X is chosen from μ, $f(X) = 1$ with probability p. Then the expected total number of queries is at most

$2p\,\mathrm{bs}_\mu^{(1)}(f)C(f) + 2(1-p)\,\mathrm{bs}_\mu^{(0)}(f)C(f) = 2\,\mathrm{bs}_\mu(f)C(f)$. ■

One can show, using a similar argument, that $D^\mu(f) \leq 2\,\mathrm{bs}(f)C_\mu(f)$. We conjecture that $D^\mu(f) = O(\mathrm{bs}_\mu(f)C_\mu(f))$, but are unable to show this.

We next give a lower bound on $Q_2(f)$ in terms of the mean block sensitivity. The proof is along the lines of Ambainis [1]; for completeness, we recapitulate some of the material in that manuscript.

Theorem 2 *$Q_2(f) \geq (1/2 - \sqrt{2}/3)\,\mathrm{bs}_\mu(f)$.*

Proof. For each X, choose $\mathrm{bs}_X(f)$ disjoint minimal blocks $B_1^{(X)}, \ldots, B_{\mathrm{bs}_X(f)}^{(X)}$ such that for all i, $f(X) \neq f(X(B_i^{(X)}))$. (By minimal, we mean that for each i, no proper sub-block $B\prime$ of $B_i^{(X)}$ has the property that $f(X) \neq f(X(B\prime))$.) Call $X(B_1^{(X)}), \ldots, X(B_{\mathrm{bs}_X(f)}^{(X)})$ the *block-neighbors* of X. (Note that if Y is a block-neighbor of X, X is not necessarily a block-neighbor of Y.)

Let A be a quantum algorithm to evaluate $f(X)$ with probability of error $\varepsilon = 1/3$. Following [1], instead of running A with a single string as input, we run A with the uniform superposition of all strings in μ as input. Let $\mathcal{H}_I$ be the 'input subspace' spanned by basis vectors $|X\rangle$ corresponding to the possible inputs X. Let ρ_k be the density matrix of $\mathcal{H}_I$ after A has made k queries. Let S_k be the sum of $(\rho_k)_{X,Y}$ for all ordered pairs (X, Y) such that Y is a block-neighbor of X. Suppose that A makes a total number of queries T. Then:

1. $S_0 = \mathrm{bs}_\mu(f)$. This is because there are 2^n input strings, the mean number of block-neighbors of a string is $\mathrm{bs}_\mu(f)$, and every entry of ρ_0 is 2^{-n}.
2. $S_T \leq 2\sqrt{\varepsilon(1-\varepsilon)}\,\mathrm{bs}_\mu(f) = (2\sqrt{2}/3)\,\mathrm{bs}_\mu(f)$, by [1, Lemma 1].
3. $S_{k-1} - S_k \leq 2$.

Together, these statements imply that $T \geq (1/2 - \sqrt{2}/3)\,\mathrm{bs}_\mu(f)$.

We now prove the third statement. Express the state before the k^{th} query as

$|\psi_{k-1}\rangle = \sum_{i,a,z,X} \alpha_{i,a,z,X} |i, a, z\rangle \otimes |X\rangle$

where i is the index of the variable X_i being queried, a is a bit for recording the answer, and z is a collection of extra work bits. Then after the k^{th} query we have

$|\psi_k\rangle = \sum_{i,a,z,X} \alpha_{i,a,z,X} |i, a \oplus X_i, z\rangle \otimes |X\rangle = \sum_{i,a,z,X} \alpha_{i,a\oplus X_i,z,X} |i, a, z\rangle \otimes |X\rangle.$

Define

$|\psi_{i,a,z}\rangle = \sum_X \alpha_{i,a,z,X} |X\rangle$ and $|\psi'_{i,a,z}\rangle = \sum_X \alpha_{i,a\oplus X_i,z,X} |X\rangle.$

Then $\rho_{k-1,i} = \sum_{a,z} |\psi_{i,a,z}\rangle\langle\psi_{i,a,z}|$ and $\rho_{k,i} = \sum_{a,z} |\psi'_{i,a,z}\rangle\langle\psi'_{i,a,z}|$ are the components of ρ_{k-1} and ρ_k respectively corresponding to querying X_i. We can then represent ρ_{k-1} as $\sum_{i=1}^n \rho_{k-1,i}$ and ρ_k as $\sum_{i=1}^n \rho_{k,i}$. Then $S_{k-1} - S_k \leq \sum_{i=1}^n S_{k,i}$ where

$S_{k,i} = \sum_{X,Y:\ Y \text{ a block-neighbor of } X} |(\rho_{k-1,i})_{X,Y} - (\rho_{k,i})_{X,Y}|.$

The only entries that differ in $\rho_{k,i}$ and $\rho_{k-1,i}$ are the ones that correspond to X, Y with $X_i \neq Y_i$. For every X, there is at most one block-neighbor Y having this property. (The fact that we're dealing with block-neighbors, rather than with ordinary neighbors as in [1], doesn't change this.) Therefore

$\sum_{X,Y:\ Y \text{ a block-neighbor of } X} (\rho_{k-1,i})_{X,Y} \leq \sum_{X,Y:\ Y \text{ a block-neighbor of } X} \left[(\rho_{k-1,i})_{X,X} + (\rho_{k-1,i})_{Y,Y}\right]/2 \leq \sum_X (\rho_{k-1,i})_{X,X} = \operatorname{Tr} \rho_{k-1,i}.$

A similar result is true for $\rho_{k,i}$. So we have that $S_{k,i} \leq \operatorname{Tr} \rho_{k-1,i} + \operatorname{Tr} \rho_{k,i}$ and that

$S_{k-1} - S_k \leq \sum_i S_{k,i} \leq \sum_i (\operatorname{Tr} \rho_{k-1,i} + \operatorname{Tr} \rho_{k,i}) = 2.$ ■

Combining Theorem 1 and Theorem 2 with $C(f) \leq \operatorname{bs}(f)^2$ [5] and $\operatorname{bs}(f) \leq 16Q_2(f)^2$ [3] we obtain

$D^\mu(f) \leq 2\operatorname{bs}_\mu(f)C(f) \leq 12(3+2\sqrt{2})\operatorname{bs}(f)^2 Q_2(f) \leq 3072(3+2\sqrt{2})Q_2(f)^5 \approx 17905Q_2(f)^5.$

When f is monotone, $C(f) = \operatorname{bs}_\mu(f)$ [5], so we obtain

$D^\mu(f) \leq 2\operatorname{bs}_\mu(f)\operatorname{bs}(f) \leq 192(3+2\sqrt{2})Q_2(f)^3 \approx 1119Q_2(f)^3.$

4 Some Open Problems

- For the case of zero-error quantum algorithms, Buhrman et al. [4] showed that $D(f) = O(Q_0(f)^4)$. Can we relate $D^\mu(f)$ to $Q_0(f)$?
- What can we say when μ is non-uniform?

References

[1] A. Ambainis. Quantum lower bounds by quantum arguments. Submitted, 1999. http://www.cs.berkeley.edu/~ambainis/ps/dens.ps.gz .

[2] A. Ambainis and R. de Wolf. Average-case quantum query complexity. To appear in *Proceedings of STACS'2000*. http://xxx.lanl.gov/abs/quant-ph/9904079 .

[3] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. In *Proceedings of 39th FOCS*, pages 352–361, 1998. http://xxx.lanl.gov/abs/quant-ph/9802049 .

[4] H. Buhrman, R. Cleve, R. de Wolf, and C. Zalka. Bounds for small-error and zero-error quantum algorithms. In *Proceedings of 40th FOCS*, 1999. http://xxx.lanl.gov/abs/cs.CC/9904019 .

[5] N. Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999-1007, 1991. Earlier version in STOC'89.