Quantum Lower Bound for the Collision Problem

Scott Aaronson*

February 1, 2008

Abstract

The collision problem is to decide whether a function $X : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ is one-to-one or two-to-one, given that one of these is the case. We show a lower bound of $\Omega\left(n^{1/5}\right)$ on the number of queries needed by a quantum computer to solve this problem with bounded error probability. The best known upper bound is $O\left(n^{1/3}\right)$, but obtaining any lower bound better than $\Omega(1)$ was an open problem since 1997. Our proof uses the polynomial method augmented by some new ideas. We also give a lower bound of $\Omega\left(n^{1/7}\right)$ for the problem of deciding whether two sets are equal or disjoint on a constant fraction of elements. Finally we give implications of these results for quantum complexity theory.

1 Introduction

The power of quantum computing has been intensively studied for a decade [4, 3, 13, 21, 2, 1, 22]. Apart from possible applications—such as speeding up combinatorial search [13] and breaking public-key cryptography [21]—a major motivation for this work has been to better understand quantum theory itself. Thus, researchers have tried to discover not just the capabilities of quantum computing but also the limitations. This task is difficult, though; proving (for example) that quantum computers cannot solve NP-complete problems in polynomial time would imply $P \neq NP$.

A popular alternative is to study restricted models of computation, and particularly the *query model*, in which one counts only the number of queries to the input, not the number of computational steps. An early result of Bennett, Bernstein, Brassard, and Vazirani [3] showed that a quantum computer needs $\Omega(\sqrt{n})$ queries to search a list of n items for one marked item. (This bound is tight, as evidenced by Grover's algorithm [13].) Subsequently, Beals et al. [2], Ambainis [1], and others obtained lower bounds for many other problems.

But one problem, the collision problem, resisted attempts to prove a lower bound [6, 1]. Because of its simplicity, the problem was widely considered a benchmark for our understanding of quantum query complexity. The collision problem of size n, or Col_n, is defined as follows. Let $X = x_1 \ldots x_n$ be a sequence of n integers drawn from $\{1, \ldots, n\}$, with n even. We are guaranteed that either

(1) X is one-to-one (that is, a permutation of $\{1, \ldots, n\}$), or

*Computer Science Department, University of California, Berkeley, CA 94720-1776. Email: aaronson@cs.berkeley.edu. Supported in part by a National Science Foundation Graduate Fellowship and by the Institute for Quantum Information at the California Institute of Technology.

(2) X is two-to-one (that is, each element of $\{1, \ldots, n\}$ appears in X twice or not at all).

The problem is to decide whether (1) or (2) holds.

We show that $Q_2(\mathrm{Col}_n) = \Omega(n^{1/5})$, where Q_2 is bounded-error quantum query complexity as defined by Beals et al. [2]. Details of the oracle model are given in Section 3. The best known upper bound, due to Brassard, Høyer, and Tapp [5], is $O(n^{1/3})$; thus, our bound is probably not tight. Previously, though, no lower bound better than the trivial $\Omega(1)$ bound was known. How great a speedup quantum computers yield for the problem was apparently first asked by Rains [18].

Previous lower bound techniques failed for the problem because they depended on a function's being sensitive to many disjoint changes to the input. For example, Beals et al. [2] showed that for all total Boolean functions f, $Q_2(f) = \Omega\left(\sqrt{\mathrm{bs}(f)}\right)$, where $\mathrm{bs}(f)$ is the block sensitivity, defined by Nisan [16] to be, informally, the maximum number of disjoint changes (to any particular input X) to which f is sensitive. In the case of the collision problem, though, every one-to-one input differs from every two-to-one input in at least $n/2$ places, so the block sensitivity is $O(1)$. Ambainis' adversary method [1], as currently formulated, faces a related obstacle. In that method we consider the algorithm and input as a bipartite quantum state, and upper-bound how much the *entanglement* of the state can increase via a single query. Yet under the simplest measures of entanglement, the algorithm and input can become highly entangled after $O(1)$ queries, again because every one-to-one input is far from every two-to-one input.

Our proof is an adaptation of the polynomial method, introduced to quantum computing by Beals et al. [2]. Their idea was to reduce questions about quantum algorithms to easier questions about multivariate polynomials. In particular, if a quantum algorithm makes T queries, then its acceptance probability is a polynomial over the input bits of degree at most $2T$. So by showing that any polynomial approximating the desired output has high degree, one obtains a lower bound on T.

To lower-bound the degree of a multivariate polynomial, a key technical trick is to construct a related *univariate* polynomial. Beals et al. [2], using a lemma due to Minsky and Papert [15], replace a polynomial $p(X)$ (where X is a bit string) by $q(|X|)$ (where $|X|$ denotes the Hamming weight of X), satisfying

$$q(k) = \underset{|X|=k}{\mathrm{EX}}\, p(X)$$

and $\deg(q) \leq \deg(p)$.

We construct the univariate polynomial in a different way. We consider a uniform distribution over k-to-one inputs, where k might be greater than 2. Even though the problem is to distinguish $k = 1$ from $k = 2$, the acceptance probability must lie between 0 and 1 for all k, and that is a surprisingly strong constraint. We show that the acceptance probability is *close* to a univariate polynomial in k of degree at most $2T$. We then obtain a lower bound by generalizing a classical approximation theory result of Ehlich and Zeller [11] and Rivlin and Cheney [19]. Much of the proof deals with the complication that k does not divide n in general.

Shi [20] has recently improved our method to obtain a lower bound of $\Omega(n^{1/4})$ for the collision problem.

The paper is organized as follows. Section 2 motivates the collision lower bound within quantum computing, pointing out connections to collision-resistant hash functions, the nonabelian hidden subgroup problem, and information erasure. Section 3 gives technical preliminaries, Section 4 proves the crucial fact that the acceptance probability is "almost" a univariate polynomial, and Section 5 completes the lower bound argument. In Appendix 7 we show a lower bound of $\Omega(n^{1/7})$ for the *set comparison problem*, a variant of the collision problem that is needed for the application to information erasure.

2 Motivation

The most immediate implication of the collision lower bound is that certain problems, notably breaking cryptographic hash functions, are not in *BQP* relative to an oracle. A second implication is that a nonstandard quantum oracle model proposed by Kashefi et al. [14] is exponentially more powerful than the usual oracle model. A third implication, in our view the most interesting one, concerns the computational power of so-called *dynamical quantum theories*. That implication will be discussed in detail in another paper.

2.1 Oracle Hardness Results

The original motivation for the collision problem was to model *(strongly) collision-resistant hash functions* in cryptography. There is a large literature on collision-resistant hashing; see [9, 7] for example. When building secure digital signature schemes, it is useful to have a family of hash functions $\{H_i\}$, such that finding a distinct (x, y) pair with $H_i(x) = H_i(y)$ is computationally intractable. A quantum algorithm for finding collisions using $O(\text{polylog}(n))$ queries would render *all* hash functions insecure against quantum attack in this sense. (Shor's algorithm [21] already renders hash functions based on modular arithmetic insecure.) Our result indicates that collision-resistant hashing might still be possible in a quantum setting.

The collision problem also models the *nonabelian hidden subgroup problem*, of which graph isomorphism is a special case. Given a group G and subgroup $H \leq G$, suppose we have oracle access to a function $f : G \rightarrow \mathbb{N}$ such that for all $g_1, g_2 \in G$, $f(g_1) = f(g_2)$ if and only if g_1 and g_2 belong to the same coset of H. Is there then an efficient quantum algorithm to determine H? If G is abelian, the work of Simon [22] and Shor [21] implies an affirmative answer. If G is nonabelian, though, efficient quantum algorithms are known only for special cases [10, 12]. An $O(\text{polylog}(n))$-query algorithm for the collision problem would yield a polynomial-time algorithm to distinguish $|H| = 1$ from $|H| = 2$, which does not exploit the group structure at all. Our result implies that no such algorithm exists.

2.2 Information Erasure

Let $f : \{0,1\}^n \rightarrow \{0,1\}^m$ with $m \geq n$ be a one-to-one function. Then we can consider two kinds of quantum oracle for f:

(A) a *standard oracle*, one that maps $|x\rangle |z\rangle$ to $|x\rangle |z \oplus f(x)\rangle$, or

(B) an *erasing oracle* (as recently proposed by Kashefi et al. [14]), which maps $|x\rangle$ to $|f(x)\rangle$, in effect "erasing" $|x\rangle$.

Intuitively erasing oracles seem at least as strong as standard ones, though it is not clear how to simulate the latter with the former without also having access to an oracle that maps $|y\rangle$ to $\left|f^{-1}(y)\right\rangle$. The question that concerns us here is whether erasing oracles are *more* useful than standard ones for some problems. One-way functions provide a clue: if f is one-way, then (by assumption) $|x\rangle |f(x)\rangle$ can be computed efficiently, but if $|f(x)\rangle$ could be computed efficiently given $|x\rangle$ then so could $|x\rangle$ given $|f(x)\rangle$, and hence f could be inverted. But can we find, for some problem, an exponential gap between query complexity given a standard oracle and query complexity given an erasing oracle?

In Appendix 7 we extend the collision lower bound to show an affirmative answer. Define the *set comparison problem* of size n, or SetComp_n, as follows. We are given as input two sequences, $X = x_1 \ldots x_n$ and $Y = y_1 \ldots y_n$, such that for each i, $x_i, y_i \in \{1, \ldots, 2n\}$. A query has the form (b, i), where $b \in \{0,1\}$ and $i \in \{1, \ldots, n\}$, and produces as output $(0, x_i)$ if $b = 0$ and $(1, y_i)$ if $b = 1$. Sequences X and Y are both one-to-one; that is, $x_i \neq x_j$ and $y_i \neq y_j$ for all $i \neq j$. We are furthermore guaranteed that either

(1) X and Y are equal as sets (that is, $\{x_1, \dots, x_n\} = \{y_1, \dots, y_n\}$) or

(2) X and Y are far as sets (that is, $|\{x_1, \dots, x_n\} \cup \{y_1, \dots, y_n\}| \geq 1.1n$).

As before the problem is to decide whether (1) or (2) holds.

This problem can be solved with high probability in a constant number of queries using an erasing oracle, by using a trick similar to that of Watrous [23] for verifying group non-membership. First, using the oracle, we prepare the uniform superposition

$$\frac{1}{\sqrt{2n}} \sum_{i \in \{1, \dots, n\}} \left(|0\rangle |x_i\rangle + |1\rangle |y_i\rangle \right).$$

We then apply a Hadamard gate to the first register, and finally we measure the first register. If X and Y are equal as sets, then interference occurs between every $(|0\rangle |z\rangle, |1\rangle |z\rangle)$ pair and we observe $|0\rangle$ with certainty. But if X and Y are far as sets, then basis states $|b\rangle |z\rangle$ with no matching $|1-b\rangle |z\rangle$ have probability weight at least $1/10$, and hence we observe $|1\rangle$ with probability at least $1/20$.

In Appendix 7 we show that $Q_2(\text{SetComp}_n) = \Omega\left(n^{1/7}\right)$; that is, no efficient quantum algorithm using a standard oracle exists for this problem.

3 Preliminaries

Let A be a quantum query algorithm. A basis state of A is written $|\Psi, i, z\rangle$. Then a query replaces each $|\Psi, i, z\rangle$ by $|\Psi \oplus x_i, i, z\rangle$, where x_i is exclusive-OR'ed into some specified location of Ψ (which we cannot assume to be all 0's). We assume without loss of generality that every basis state queries at every step. Between queries, the algorithm can perform any unitary operation that does not depend on the input. At the end z is measured in the standard basis; if $z = 1$ the algorithm returns 'one-to-one' and if $z = 2$ it returns 'two-to-one.' The total number of queries is denoted T. Also, we assume for simplicity that all amplitudes are real; this restriction is without loss of generality [4].

Let $\alpha^{(t)}_{X,\Psi,i,z}$ be the amplitude of basis state $|\Psi, i, z\rangle$ after t queries when the input is X. Also, let $\Delta(x_i, h) = 1$ if $x_i = h$, and $\Delta(x_i, h) = 0$ if $x_i \neq h$. Let $P(X)$ be the probability that A returns $z = 2$ when the input is X. Then we obtain a simple variant of the main lemma of Beals et al. [2].

Lemma 1 *$P(X)$ is a multilinear polynomial of degree at most $2T$ over the $\Delta(x_i, h)$.*

Proof. We show, by induction on t, that for all basis states $|\Psi, i, z\rangle$, $\alpha^{(t)}_{X,\Psi,i,z}$ is a multilinear polynomial of degree at most t over the $\Delta(x_i, h)$. Since $P(X)$ is a sum of squares of $\alpha^{(t)}_{X,\Psi,i,z}$, the lemma follows.

The base case ($t = 0$) holds since, before making any queries, each $\alpha^{(0)}_{X,\Psi,i,z}$ is a degree-0 polynomial over the $\Delta(x_i, h)$. A unitary transformation on the algorithm part replaces each $\alpha^{(t)}_{X,\Psi,i,z}$ by a linear combination of $\alpha^{(t)}_{X,\Psi,i,z}$, and hence cannot increase the degree. Suppose the lemma holds prior to the t^{th} query. Then

$$\alpha^{(t+1)}_{X,\Psi,i,z} = \sum_{1 \leq h \leq n} \alpha^{(t)}_{X,\Psi \oplus h,i,z} \Delta(x_i, h),$$

and we are done. ■

A remark on notation: we sometimes use brackets ($a_{b[c]}$) rather than nested subscripts (a_{b_c}).

4 Reduction to Bivariate Polynomial

Call the point $(g, N) \in \Re^2$ an (n, T)-*quasilattice point* if and only if

(1) g and N are integers, with g dividing N,

(2) $1 \leq g \leq \sqrt{n}$,

(3) $n \leq N \leq n + n/(10T)$, and

(4) if $g = 1$ then $N = n$.

For quasilattice point (g, N), define $\mathcal{D}_n(g, N)$ to be the uniform distribution over all size-n subfunctions of g-1 functions having domain $\{1, \ldots, N\}$ and range a subset of $\{1, \ldots, n\}$. More precisely: to draw an X from $\mathcal{D}_n(g, N)$, we first choose a set $S \subseteq \{1, \ldots, n\}$ with $|S| = N/g \leq n$ uniformly at random. We then choose a g-1 function $\widehat{X} = \widehat{x}_1 \ldots \widehat{x}_N$ from $\{1, \ldots, N\}$ to S uniformly at random. Finally we let $x_i = \widehat{x}_i$ for each $1 \leq i \leq n$.

Let $P(g, N)$ be the probability that algorithm A returns $z = 2$ when the input is chosen from $\mathcal{D}_n(g, N)$:

$$P(g, N) = \underset{X \in \mathcal{D}[n](g,N)}{\mathrm{EX}} P(X).$$

We then have the following surprising characterization:

Lemma 2 *For all sufficiently large n and if $T \leq \sqrt{n}/3$, there exists a bivariate polynomial $q(g, N)$ of degree at most $2T$ such that if (g, N) is a quasilattice point, then*

$$|P(g, N) - q(g, N)| < 0.182$$

(where the constant 0.182 can be made arbitrarily small by adjusting parameters).

Proof. Let I be a product of $\Delta(x_i, h)$ variables, with degree $r(I)$, and let $I(X) \in \{0, 1\}$ be I evaluated on input X. Then define

$$\gamma(I, g, N) = \underset{X \in \mathcal{D}[n](g,N)}{\mathrm{EX}} I(X)$$

to be the probability that monomial I evaluates to 1 when the input is drawn from $\mathcal{D}_n(g, N)$. Then by Lemma 1, $P(X)$ is a polynomial of degree at most $2T$ over X, so

$$P(g, N) = \underset{X \in \mathcal{D}[n](g,N)}{\mathrm{EX}} P(X) = \underset{X \in \mathcal{D}[n](g,N)}{\mathrm{EX}} \sum_{I : r(I) \leq 2t} \beta_I I(X) = \sum_{I : r(I) \leq 2T} \beta_I \gamma(I, g, N)$$

for some coefficients β_I.

We now calculate $\gamma(I, g, N)$. Assume without loss of generality that for all $\Delta(x_i, h_1), \Delta(x_j, h_2) \in I$, either $i \neq j$ or $h_1 = h_2$, since otherwise $\gamma(I, g, N) = 0$.

Define the "range" $Z(I)$ of I to be the set of all h such that $\Delta(x_i, h) \in I$. Let $w(I) = |Z(I)|$; then we write $Z(I) = \{z_1, \ldots, z_{w(I)}\}$. Clearly $\gamma(I, g, N) = 0$ unless $Z(I) \in S$, where S is the range of $\widehat{X}$. By assumption,

$$\frac{N}{g} \geq \frac{n}{\sqrt{n}} \geq 2T \geq r(I)$$

so the number of possible S is $\binom{n}{N/g}$ and, of these, the number that contain Z is $\binom{n-w(I)}{N/g-w(I)}$.

Then, conditioned on $Z \in S$, what is the probability that $\gamma(I,g,N) = 1$? The total number of g-1 functions with domain size N is $N!/(g!)^{N/g}$, since we can permute the N function values arbitrarily, but must not count permutations that act only within the N/g constant-value blocks of size g.

Among these functions, how many satisfy $\gamma(I,g,N) = 1$? Suppose that, for each $1 \le j \le w(I)$, there are $r_j(I)$ distinct i such that $\Delta(x_i, z_j) \in I$. Clearly

$$r_1(I) + \cdots + r_{w(I)}(I) = r(I).$$

Then we can permute the $(N - r(I))!$ function values outside of I arbitrarily, but must not count permutations that act only within the N/g constant-value blocks, which have size either g or $g - r_i(I)$ for some i. So the number of functions for which $\gamma(I,g,N) = 1$ is

$$\frac{(N-r(I))!}{(g!)^{N/g-w(I)} \prod_{i=1}^{w(I)} (g - r_i(I))!}.$$

Putting it all together,

$$\begin{aligned}
\gamma(I,g,N) &= \frac{\binom{n-w(I)}{N/g-w(I)}}{\binom{n}{N/g}} \cdot \frac{(N-r(I))!\,(g!)^{N/g}}{(g!)^{N/g-w(I)}\,N! \prod_{i=1}^{w(I)} (g-r_i(I))!} \\
&= \frac{(N-r(I))!\,(n-w(I))!\,(N/g)!}{N!n!\,(N/g-w(I))!} \cdot \frac{(g!)^{w(I)}}{\prod_{i=1}^{w(I)} (g-r_i(I))!} \\
&= \frac{(N-r(I))!}{N!} \frac{(n-w(I))!}{n!} \cdot \prod_{i=0}^{w(I)-1} \left(\frac{N}{g} - i\right) \prod_{i=1}^{w(I)} \left[g \prod_{j=1}^{r[i](I)-1} (g-j) \right] \\
&= \frac{(N-2T)!}{N!} \frac{(n-w(I))!}{n!} \cdot \prod_{i=r(I)}^{2T-1} (N-i) \prod_{i=0}^{w(I)-1} (N-gi) \prod_{i=1}^{w(I)} \prod_{j=1}^{r[i](I)-1} (g-j) \\
&= \frac{(N-2T)!n!}{N!\,(n-2T)!} \widetilde{q}_{n,T,I}(g,N)
\end{aligned}$$

where

$$\widetilde{q}_{n,T,I}(g,N) = \frac{(n-w(I))!\,(n-2T)!}{(n!)^2} \cdot \prod_{i=r(I)}^{2T-1} (N-i) \prod_{i=0}^{w(I)-1} (N-gi) \prod_{i=1}^{w(I)} \prod_{j=1}^{r[i](I)-1} (g-j)$$

is a bivariate polynomial of total degree at most

$$(2T - r(I)) + w(I) + (r(I) - w(I)) = 2T.$$

(Note that in the case $r_i(I) > g$ for some i, this polynomial evaluates to 0, which is what it ought to do.) Hence

$$P(g,N) = \sum_{I : r(I) \le 2T} \beta_I \gamma(I,g,N) = \frac{(N-2T)!n!}{N!\,(n-2T)!} q(g,N)$$

where

$$q\left(g,N\right)=\sum_{I:r(I)\leq 2T}\beta_{I}\widetilde{q}_{n,T,I}\left(g,N\right).$$

Clearly

$$\frac{(N-2T)!n!}{N!\left(n-2T\right)!}\leq 1.$$

Since $N\leq n+n/\left(10T\right)$ and $T\leq\sqrt{n}/3$, we also have

$$\begin{aligned}\frac{(N-2T)!n!}{N!\left(n-2T\right)!}&\geq\left(\frac{n-2T+1}{N-2T+1}\right)^{2T}\\&\geq\left(1-\frac{n/\left(10T\right)}{N-2T+1}\right)^{2T}\\&\geq\left(1-\frac{n}{10\left[n-\left(2T+1\right)/n\right]}\frac{1}{T}\right)^{2T}\\&\geq\exp\left\{-\frac{1}{5}\frac{n}{n-\left(2T+1\right)/n}\right\}\\&\geq 0.818\end{aligned}$$

for all sufficiently large n.

Thus, since $0\leq P\left(g,N\right)\leq 1$,

$$\left|P\left(g,N\right)-q\left(g,N\right)\right|<0.182$$

and we are done. ■

5 Lower Bound

We are now ready to prove a lower bound for the collision problem. To do so, we generalize an approximation theory result due to Rivlin and Cheney [19] and (independently) Ehlich and Zeller [11]. That result was applied to query complexity by Nisan and Szegedy [17] and later by Beals et al. [2].

Theorem 3 $Q_2\left(\mathrm{Col}_n\right)=\Omega\left(n^{1/5}\right).$

Proof. Let g have range $1\leq g\leq G$. Then the quasilattice points $\left(g,N\right)$ all lie in the rectangular region $R=\left[1,G\right]\times\left[n,n+n/\left(10T\right)\right]$. Recalling the polynomial $q\left(g,N\right)$ from Lemma 2, define

$$d\left(q\right)=\max_{(g,N)\in R}\left(\max\left\{\left|\frac{\partial q}{\partial g}\right|,\frac{n}{10T\left(G-1\right)}\cdot\left|\frac{\partial q}{\partial N}\right|\right\}\right).$$

Suppose without loss of generality that we require $P\left(1,n\right)\leq 1/10$ and $P\left(2,n\right)\geq 9/10$ (that is, algorithm A distinguishes 1-1 from 2-1 functions with error probability at most 1/10). Then, since

$$\left|P\left(g,N\right)-q\left(g,N\right)\right|<0.182$$

by the Intermediate Value Theorem we have

$$d(q) \geq \max_{1 \leq g \leq 2} \frac{\partial q}{\partial g} > 0.8 - 2(0.182) = 0.436.$$

An inequality due to Markov (see [8, 17]) states that, for a univariate polynomial p, if $b_1 \leq p(x) \leq b_2$ for all $a_1 \leq x \leq a_2$, then

$$\max_{a[1] \leq x \leq a[2]} \left| \frac{dp(x)}{dx} \right| \leq \frac{b_2 - b_1}{a_2 - a_1} \deg(p)^2 .$$

Clearly for every point $\left(\widehat{g}, \widehat{N}\right) \in R$, there exists a quasilattice point (g, N) for which $|g - \widehat{g}| \leq 1$ and $\left|N - \widehat{N}\right| \leq G$. For take $g = \lceil \widehat{g} \rceil$—or, in the special case $\widehat{g} = 1$, take $g = 2$, since there is only one quasilattice point with $g = 1$.

Furthermore, since $P(g, N)$ represents an acceptance probability at such a point, we have

$$-0.182 < q(g, N) < 1.182.$$

Observe that for all $\left(\widehat{g}, \widehat{N}\right) \in R$,

$$-0.182 - \left(\frac{10TG(G-1)}{n} + 1\right) d(q) < q\left(\widehat{g}, \widehat{N}\right) < 1.182 + \left(\frac{10TG(G-1)}{n} + 1\right) d(q).$$

For consider a quasilattice point close to $\left(\widehat{g}, \widehat{N}\right)$, and note that the maximum-magnitude derivative is at most $d(q)$ in the g direction and $10T(G-1)\, d(q)/n$ in the N direction.

Let (g^*, N^*) be a point in R at which the weighted maximum-magnitude derivative $d(q)$ is attained. Suppose first that the maximum is attained in the g direction. Then $q(g, N^*)$ (with N^* constant) is a univariate polynomial with

$$\left| \frac{dq(g, N^*)}{dg} \right| > 0.436$$

for some $1 \leq g \leq G$. So

$$\begin{aligned} 2T &\geq \deg(q(g, N)) \\ &\geq \deg(q(g, N^*)) \\ &\geq \sqrt{\frac{d(q)(G-1)}{1.364 + 2d(q)(1 + 10TG(G-1)/n)}} \\ &\geq \sqrt{\frac{0.436(G-1)n}{2.236n + 8.720TG(G-1)}} \\ &= \Omega\left(\min\left\{\sqrt{G}, \sqrt{\frac{n}{TG}}\right\}\right). \end{aligned}$$

Similarly, suppose the maximum $d(q)$ is attained in the N direction. Then $q(g^*, N)$ (with g^* constant) is a univariate polynomial with

$$\left| \frac{dq(g^*, N)}{dN} \right| > \frac{0.436T(G-1)}{n}$$

for some $n \leq N \leq n + n/(10T)$. So

$$2T \geq \sqrt{\frac{(10T(G-1)/n)\, d(q)\, n/(10T)}{1.364 + 2d(q)(1 + 10TG(G-1)/n)}} \geq \Omega\left(\min\left\{\sqrt{G}, \sqrt{\frac{n}{TG}}\right\}\right).$$

One can show that the lower bound on T is optimized when we take $G = n^{2/5} \leq \sqrt{n}$. Then

$$T = \Omega\left(\min\left\{n^{1/5}, \frac{\sqrt{n}}{\sqrt{T}n^{1/5}}\right\}\right),$$
$$T = \Omega\left(n^{1/5}\right)$$

and we are done. ■

6 Acknowledgments

I am grateful to Yaoyun Shi, Ronald de Wolf, Umesh Vazirani, Ashwin Nayak, and Andris Ambainis for helpful comments; to Leonard Schulman, Lawrence Ip, Jordan Kerenidis, and John Preskill for discussions during earlier stages of this work; and to James Lee, Alex Halderman, and Elham Kashefi for discussions and references regarding Section 2.

References

[1] A. Ambainis. Quantum lower bounds by quantum arguments. *Proceedings of STOC'2000*, pages 636–643, 2000. Journal version to appear in *Journal of Computer and System Sciences*. quant-ph/0002066[1].

[2] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *Proceedings of FOCS'98*, pages 352–361, 1998. quant-ph/9802049.

[3] C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, 1997. quant-ph/9701001.

[4] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997.

[5] G. Brassard, P. Høyer, and A. Tapp. Quantum algorithm for the collision problem. *ACM SIGACT News (Cryptology Column)*, 28:14–19, 1997. quant-ph/9705002.

[6] H. Buhrman, C. Dürr, M. Heiligman, P. Høyer, F. Magniez, M. Santha, and R. de Wolf. Quantum algorithms for element distinctness. *Proceedings of IEEE Conference on Computational Complexity (CCC'2001)*, pages 131–137, 2001. quant-ph/0007016.

[1] Available at `www.arxiv.org`.

[7] S. Bakhtiari, R. Safavi-Naini, and J. Pieprzyk. Cryptographic hash functions: a survey. Technical Report 95-09, Department of Computer Science, University of Wollongong, July 1995. Available at ftp://ftp.cs.uow.edu.au/pub/papers/1995/tr-95-09.ps.Z.

[8] E. W. Cheney. *Introduction to approximation theory*, McGraw-Hill, 1966.

[9] I. B. Damgård. Collision free hash functions and public key signature schemes. *Proceedings of Eurocrypt'87*, Volume 304 of *Lecture Notes in Computer Science* (Springer-Verlag), 1988.

[10] M. Ettinger and P. Høyer. On quantum algorithms for noncommutative hidden subgroups. *Advances in Applied Mathematics*, 25(3):239–251, 2000.

[11] H. Ehlich and K. Zeller. Schwankung von Polynomen zwischen Gitterpunkten. *Mathematische Zeitschrift*, 86:41–44, 1964.

[12] M. Grigni, L. Schulman, M. Vazirani, and U. Vazirani. Quantum mechanical algorithms for the nonabelian hidden subgroup problem. *Proceedings of STOC'2001*, pages 68–74, 2001.

[13] L. K. Grover. A fast quantum mechanical algorithm for database search. *Proceedings of STOC'96*, pages 212–219, 1996. quant-ph/9605043.

[14] E. Kashefi, A. Kent, V. Vedral, and K. Banaszek. On the power of quantum oracles, 2001. quant-ph/0109104.

[15] M. Minsky and S. Papert. *Perceptrons*, MIT Press, 1988. First appeared in 1968.

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

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

[18] E. Rains. Talk given at AT&T, Murray Hill, New Jersey, on March 12, 1997.

[19] T. J. Rivlin and E. W. Cheney. A comparison of Uniform Approximations on an interval and a finite subset thereof. *SIAM Journal on Numerical Analysis*, 3(2):311–320, 1966.

[20] Y. Shi. Improving the lower bound on the collision problem to $n^{1/4}$. Manuscript, 2001.

[21] P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997. quant-ph/9508027.

[22] D. Simon. On the power of quantum computation. *Proceedings of FOCS'94*, pages 116–123, 1994.

[23] J. Watrous. Succinct quantum proofs for properties of finite groups. *Proceedings of FOCS'2000*, pages 537–546, 2000. cs.CC/0009002.

7 Appendix: Set Comparison

Here we show that $Q_2\left(\mathrm{SetComp}_n\right) = \Omega\left(n^{1/7}\right)$, where $\mathrm{SetComp}_n$ is the set comparison problem of size n as defined in Section 2.2. We need only redo the proof of Lemma 2; then Theorem 3 goes through largely unchanged.

The idea is the following. We need a distribution of inputs with a parameter g, such that the inputs are one-to-one when $g = 1$ or $g = 2$—since otherwise the problem of distinguishing $g = 1$ from $g = 2$ would be ill-defined for erasing oracles. On the other hand, the inputs must *not* be one-to-one for all $g > 2$—since otherwise the lower bound for standard oracles would apply also to erasing oracles, and we could not obtain a separation between the two. Finally, the algorithm's acceptance probability must be close to a polynomial in g.

Our solution is to consider $\kappa\left(g\right)$-to-one inputs, where

$$\kappa\left(g\right) = 4g^2 - 12g + 9.$$

is a quadratic with $\kappa\left(1\right) = \kappa\left(2\right) = 1$. The total range of the inputs (on sequences X and Y combined) has size roughly n/g; thus, we can tell the $g = 1$ inputs apart from the $g = 2$ inputs using an erasing oracle, even though $\kappa\left(g\right)$ is the same for both. The disadvantage is that, because $\kappa\left(g\right)$ increases quadratically rather than linearly in g, the quasilattice points become sparse more quickly. That is what weakens the lower bound from $\Omega\left(n^{1/5}\right)$ to $\Omega\left(n^{1/7}\right)$. We note that, using the ideas of Shi [20], one can improve our lower bound on $Q_2\left(\mathrm{SetComp}_n\right)$ to $\Omega\left(n^{1/6}\right)$.

Call $(g, N, M) \in \Re^3$ an (n,T)-*super-quasilattice point* if and only if

(1) g is an integer in $\left[1, n^{1/3}\right]$,

(2) N and M are integers in $\left[n, n\left(1 + 1/\left(100T\right)\right)\right]$,

(3) g divides N,

(4) if $g = 1$ then $N = n$,

(5) $\kappa\left(g\right)$ divides M, and

(6) if $g = 2$ then $M = n$.

For super-quasilattice point (g, N, M), we draw input $(X, Y) = (x_1 \ldots x_n, y_1 \ldots y_n)$ from distribution $\mathcal{L}_n\left(g, N, M\right)$ as follows. We first choose a set $S \subseteq \{1, \ldots, 2n\}$ with $|S| = 2N/g \leq 2n$ uniformly at random. We then choose two sets $S_X, S_Y \subseteq S$ with $|S_X| = |S_X| = M/\kappa\left(g\right) \leq |S|$, uniformly at random and independently. Next we choose $\kappa\left(g\right)$-1 functions $\widehat{X} = \widehat{x}_1 \ldots \widehat{x}_N : \{1, \ldots, M\} \rightarrow S_X$ and $\widehat{Y} = \widehat{y}_1 \ldots \widehat{y}_N : \{1, \ldots, M\} \rightarrow S_Y$ uniformly at random and independently. Finally we let $x_i = \widehat{x}_i$ and $y_i = \widehat{y}_i$ for each $1 \leq i \leq n$.

Define sets $X_S = \{x_1, \ldots, x_n\}$ and $Y_S = \{y_1, \ldots, y_n\}$. Suppose $g = 1$ and $N = M = n$; then by Chernoff bounds,

$$\Pr_{(X,Y) \in \mathcal{L}[n](1,n,n)}\left[|X_S \cup Y_S| < 1.1n\right] \leq 2e^{-n/10}.$$

Thus, if algorithm A can distinguish $|X_S \cup Y_S| = n$ from $|X_S \cup Y_S| \geq 1.1n$ with probability at least $9/10$, then it can distinguish $(X, Y) \in \mathcal{L}_n\left(1, n, n\right)$ from $(X, Y) \in \mathcal{L}_n\left(2, n, n\right)$ with probability at least $9/10 - 2e^{-n/10}$. So a lower bound for the latter problem implies an equivalent lower bound for the former.

Define $P(X,Y)$ to be the probability that the algorithm returns that X and Y are far on input (X,Y), and let

$$P(g,N,M)=\underset{(X,Y)\in\mathcal{L}[n](g,N,M)}{\mathrm{EX}} P(X,Y).$$

We then have

Lemma 4 *For all sufficiently large n and if $T\leq n^{1/3}/8$, there exists a trivariate polynomial $q(g,N,M)$ of degree at most $8T$ such that if (g,N,M) is a super-quasilattice point, then*

$$|P(g,N,M)-q(g,N,M)|<\varepsilon$$

for some constant $0<\varepsilon<1/2$.

Proof. By analogy to Lemma 1, $P(X,Y)$ is a multilinear polynomial of degree at most $2T$ over variables of the form $\Delta(x_i,h)$ and $\Delta(y_i,h)$. Let $I(X,Y)=I_X(X)I_Y(Y)$ where I_X is a product of $r_X(I)$ distinct $\Delta(x_i,h)$ variables and I_Y is a product of $r_Y(I)$ distinct $\Delta(y_i,h)$ variables. Let $r(I)=r_X(I)+r_Y(I)$. Define

$$\gamma(I,g,N,M)=\underset{(X,Y)\in\mathcal{L}[n](g,N,M)}{\mathrm{EX}} I(X,Y);$$

then

$$P(g,N,M)=\sum_{I:r(I)\leq 2T}\beta_I\gamma(I,g,N,M)$$

for some coefficients β_I.

We now calculate $\gamma(I,g,N,M)$. As before we assume there are no pairs of variables $\Delta(x_i,h_1),\Delta(x_i,h_2)\in I$ with $h_1\neq h_2$.

Let $Z_X(I)$ be the range of I_X and let $Z_Y(I)$ be the range of I_Y. Then let $Z(I)=Z_X(I)\cup Z_Y(I)$. Let $w_X(I)=|Z_X(I)|$, $w_Y(I)=|Z_Y(I)|$, and $w(I)=|Z(I)|$. By assumption

$$\frac{N}{g}\geq\frac{M}{\kappa(g)}\geq\frac{1}{4}n^{1/3}\geq 2T$$

so

$$\Pr[Z(I)\subseteq S]=\frac{\binom{2n-w(I)}{2N/g-w(I)}}{\binom{2n}{2N/g}}.$$

Then the probability that $Z_X(I)\subseteq S_X$ given $Z(I)\subseteq S$ is

$$\frac{\binom{2N/g-w_X(I)}{M/\kappa(g)-w_X(I)}}{\binom{2N/g}{M/\kappa(g)}}$$

and similarly for the probability that $Z_Y(I) \subseteq S_Y$ given $Z(I) \subseteq S$.

Let $r_{X,1}(I), \ldots, r_{X,w[X](I)}(I)$ be the multiplicities of the range elements in $Z_X(I)$, so that $r_{X,1}(I) + \cdots + r_{X,w[X](I)}(I) = r_X(I)$. Then

$$\Pr\left[I_X(X) \mid Z_X(I) \subseteq S_X\right] = \frac{(M - r_X(I))!}{M!} \prod_{i=1}^{w[X](I)} \prod_{j=0}^{r[X,i](I)-1} (\kappa(g) - j)$$

and similarly for $\Pr\left[I_Y(Y) \mid Z_Y(I) \subseteq S_Y\right]$.

Putting it all together,

$$\begin{aligned}\gamma(I, g, N, M) =& \frac{\binom{2n - w(I)}{2N/g - w(I)}}{\binom{2n}{2N/g}} \frac{(M - r_X(I))!}{M!} \frac{(M - r_Y(I))!}{M!} \frac{\binom{2N/g - w_X(I)}{M/\kappa(g) - w_X(I)}}{\binom{2N/g}{M/\kappa(g)}} \times \\ & \prod_{i=1}^{w[X](I)} \prod_{j=0}^{r[X,i](I)-1} (\kappa(g) - j) \frac{\binom{2N/g - w_Y(I)}{M/\kappa(g) - w_Y(I)}}{\binom{2N/g}{M/\kappa(g)}} \prod_{i=1}^{w[Y](I)} \prod_{j=0}^{r[Y,i](I)-1} (\kappa(g) - j) \\ =& \frac{(2n - w(I))!}{(2n)!} \frac{(M - r_X(I))!}{M!} \frac{(M - r_Y(I))!}{M!} \frac{(2N/g - w_X(I))!}{(2N/g - w(I))!} \frac{(2N/g - w_Y(I))!}{(2N/g)!} \theta_I(g, M)\end{aligned}$$

where

$$\begin{aligned}&\theta_I(g, M) = \\ &\prod_{i=0}^{w[X](I)-1} (M - i\kappa(g)) \prod_{i=1}^{w[X](I)} \prod_{j=1}^{r[X,i](I)-1} (\kappa(g) - j) \prod_{i=0}^{w[Y](I)-1} (M - i\kappa(g)) \prod_{i=1}^{w[Y](I)} \prod_{j=1}^{r[Y,i](I)-1} (\kappa(g) - j)\end{aligned}$$

is a bivariate polynomial in (g, M) of total degree at most $2r(I)$.

Thus

$$\begin{aligned}\gamma(I, g, N, M) =& \frac{(2n - w(I))!}{(2n)!} \left[\frac{(M - 2T)!n!}{M!(n - 2T)!}\right]^2 \left(\frac{(n - 2T)!}{n!}\right)^2 \prod_{i=r[X](I)}^{2T-1} (M - i) \prod_{i=r[Y](I)}^{2T-1} (M - i) \times \\ & \frac{(2N/g - w_X(I)) \cdots (2N/g - (w(I) - 1))}{(2N/g)(2N/g - 1) \cdots (2N/g - (w_Y(I) - 1))} \theta_I(g, M) \\ =& \frac{(2n)^{2T}}{(2N)(2N - g) \cdots (2N - (2T - 1)g)} \left[\frac{(M - 2T)!n!}{M!(n - 2T)!}\right]^2 \widetilde{q}_{n,T,I}(g, N, M)\end{aligned}$$

where

$$\begin{aligned}\widetilde{q}_{n,T,I}(g, N, M) =& \frac{(2n - w(I))!}{(2n)!(2n)^{2T}} \left(\frac{(n - 2T)!}{n!}\right)^2 g^{w[X](I) + w[Y](I) - w(I)} \theta_I(g, M) \prod_{i=r[X](I)}^{2T-1} (M - i) \times \\ & \prod_{i=r[Y](I)}^{2T-1} (M - i) \prod_{i=w[X](I)}^{w(I)-1} (2N - ig) \prod_{i=w[Y](I)}^{2T-1} (2N - ig)\end{aligned}$$

is a trivariate polynomial in (g, N, M) of total degree at most

$$(4T - r(I)) + 2r(I) + (w_X(I) + w_Y(I) - w(I)) + (w(I) - w_X(I)) + (2T - w_Y(I)) \leq 8T.$$

Thus

$$P(g, N, M) = \frac{(2n)^{2T}}{\prod_{i=0}^{2T-1} (2N - gi)} \left[\frac{(M - 2T)!n!}{M!(n - 2T)!} \right]^2 q(g, N, M)$$

where $q(g, N, M)$ is a polynomial of total degree at most $8T$. The argument that q approximates P to within a constant is analogous to that of Lemma 2; note that

$$\left| \frac{(2n)^{2T}}{\prod_{i=0}^{2T-1} (2N - gi)} - 1 \right| = O\left[\left(1 + \frac{1}{T} + \frac{gT}{n} \right)^{2T} \right] = O(1)$$

since $g \leq n^{1/3}$ and $T \leq n^{1/3}/8$. ■

Theorem 5 $Q_2(\mathrm{SetComp}_n) = \Omega(n^{1/7})$.

Proof sketch. The proof is analogous to that of Theorem 3. Let $g \in [1, G]$ for some $G \leq n^{1/3}$. Then the super-quasilattice points (g, N, M) all lie in $R = [1, G] \times [n, n + n/(100T)]^2$. Define $d(q)$ to be

$$\max_{(g,N,M) \in R} \left(\max \left\{ \left| \frac{\partial q}{\partial g} \right|, \frac{n/100T}{(G-1)} \left| \frac{\partial q}{\partial N} \right|, \frac{n/100T}{(G-1)} \left| \frac{\partial q}{\partial M} \right| \right\} \right).$$

Then $d(q) \geq \delta$ for some constant $\delta > 0$, by Lemma 4.

For every point $\left(\widehat{g}, \widehat{N}, \widehat{M}\right) \in R$, there exists a super-quasilattice point (g, N, M) such that $|g - \widehat{g}| \leq 1$, $\left|N - \widehat{N}\right| \leq G$, and $\left|M - \widehat{M}\right| \leq \kappa(G)$. Hence, $q\left(\widehat{g}, \widehat{N}, \widehat{M}\right)$ can deviate from $[0, 1]$ by at most

$$O\left(\left(\frac{TG^3}{n} + 1 \right) d(q) \right).$$

Let (g^*, N^*, M^*) be a point in R at which $d(q)$ is attained. Suppose $d(q)$ is attained in the g direction; the cases of the N and M directions are analogous. Then $q(g, N^*, M^*)$ is a univariate polynomial in g, and

$$\begin{aligned} 8T &\geq \deg(q(g, N^*, M^*)) \\ &= \Omega\left(\sqrt{\frac{d(q)\, G}{1 + d(q) + d(q)\, TG^3/n}} \right) \\ &= \Omega\left(\min\left\{ \sqrt{G}, \sqrt{\frac{n}{TG^2}} \right\} \right). \end{aligned}$$

One can show that the bound is optimized when we take $G = n^{2/7} \leq n^{1/3}$. Then

$$\begin{aligned} T &= \Omega\left(\min\left\{n^{1/7}, \frac{\sqrt{n}}{\sqrt{T}n^{2/7}}\right\}\right), \\ T &= \Omega\left(n^{1/7}\right). \end{aligned}$$

■