Quantum Lower Bound for Approximate Counting via Laurent Polynomials

Scott Aaronson[*]

Abstract

We consider the following problem: estimate the size of a nonempty set $S \subseteq [N]$, given both quantum queries to a membership oracle for S, and a device that generates equal superpositions $|S\rangle$ over S elements. We show that, if $|S|$ is neither too large nor too small, then approximate counting with these resources is still quantumly hard. More precisely, any quantum algorithm needs either $\Omega\left(\sqrt{N/|S|}\right)$ queries or else $\Omega\left(\min\left\{|S|^{1/4}, \sqrt{N/|S|}\right\}\right)$ copies of $|S\rangle$. This means that, in the black-box setting, quantum sampling does *not* imply approximate counting. The proof uses a novel generalization of the polynomial method of Beals et al. to Laurent polynomials, which can have negative exponents.

1 Introduction

The *quantum query complexity of approximate counting* was one of the first topics studied in quantum algorithms. Given a nonempty finite set $S \subseteq [N]$ (here and throughout, $[N] = \{1, \ldots, N\}$), suppose we want to estimate its cardinality, $|S|$, to within some multiplicative accuracy ε. This is a fundamental task in theoretical computer science, used as a subroutine for countless other tasks.

As is standard in quantum algorithms, we work in the so-called *black-box model* (see [10]), where we assume only that we're given a membership oracle for S: an oracle that, for any $i \in [N]$, tells us whether $i \in S$. We can, however, query the oracle in quantum superposition. How many queries must a quantum computer make, as a function of both N and $|S|$, to solve this problem with high probability?

For classical randomized algorithms, one can show that $\Theta\left(\frac{N}{|S|}\right)$ membership queries are necessary and sufficient, for approximate counting to within some constant accuracy $\varepsilon > 0$. Moreover, *any* accuracy ε is achievable at the cost of a $O\left(\frac{1}{\varepsilon^2}\right)$ multiplicative overhead. Intuitively, in the worst case, we might need $\Theta\left(\frac{N}{|S|}\right)$ queries just to find *any* S elements, but once we do, estimating their frequency is just a standard statistics problem. Furthermore, for the $O\left(\frac{N}{|S|}\right)$ estimation strategy to work, we don't need to suppose (circularly) that $|S|$ is approximately known in advance, but can decide when to halt dynamically, depending on when the first S elements are found.

In the quantum case, Brassard, Høyer, and Tapp [9] gave an algorithm for approximate counting that makes only $O\left(\sqrt{\frac{N}{|S|}}\right)$ queries, for any constant $\varepsilon > 0$. Moreover, they showed how to achieve

[*] University of Texas at Austin. Email: aaronson@cs.utexas.edu. Supported by a Vannevar Bush Fellowship from the US Department of Defense, a Simons Investigator Award, and the Simons "It from Qubit" collaboration.

any accuracy ε with $O\left(\frac{1}{\varepsilon}\right)$ multiplicative overhead. To do so, one uses amplitude amplification, the basic primitive of Grover's search algorithm [14]. The original algorithm of Brassard et al. [9] also used quantum phase estimation, in effect *combining* Grover's algorithm with Shor's period-finding algorithm. However, it's a folklore fact that one can remove the phase estimation, and adapt Grover search with an unknown number of marked items, to get an approximate count of the number of marked items as well.

On the lower bound side, it follows immediately from the optimality of Grover's algorithm (i.e., the BBBV Theorem [7]) that even with a quantum computer, in the black-box setting, $\Omega\left(\sqrt{\frac{N}{|S|}}\right)$ queries are *needed* for approximate counting to any constant accuracy.

In practice, when trying to estimate the size of a set $S \subseteq [N]$, often we can do more than make membership queries to S. At the least, often we can efficiently generate nearly-uniform *samples* from S, for instance by using Markov Chain Monte Carlo techniques. To give two examples, if S is the set of perfect matchings in a bipartite graph, or the set of grid points in a high-dimensional convex body, then we can efficiently sample S using the seminal algorithms of Jerrum, Sinclair, and Vigoda [15] or of Dyer, Frieze, and Kannan [11], respectively.

Sometimes we can even "QSample" S—a term coined in 2003 by Aharonov and Ta-Shma [5], and which simply means that we can approximately prepare the uniform superposition

$$|S\rangle := \frac{1}{\sqrt{|S|}} \sum_{i \in S} |i\rangle$$

via a polynomial-time quantum algorithm (where "polynomial" here means $(\log N)^{O(1)}$). Because we need to uncompute any garbage, the ability to prepare $|S\rangle$ as a coherent superposition is a more stringent requirement than the ability to sample S. Indeed, as Aharonov and Ta-Shma [5] pointed out, the quantum lower bound for finding collisions [1, 4] has the corollary that, in the black-box setting, there are classes of sets S that can be efficiently sampled but *not* efficiently QSampled.

On the other hand, Aharonov and Ta-Shma [5], and Grover and Rudolph [13], observed that many interesting sets S can be QSampled as well. In particular, this holds for all sets S such that we can approximately count not only S itself, but also the restrictions of S obtained by fixing bits of its elements. Or, what's known to be equivalent [21], it holds for all sets S such that we can efficiently sample not only the uniform distribution over S elements, but also the conditional distributions obtained by fixing bits. So in particular, the set of perfect matchings in a bipartite graph, and the set of grid points in a convex body, can both be efficiently QSampled. There are other sets that can be QSampled but not because of this reduction. A simple example would be a set S such that $|S| \geq \frac{N}{\operatorname{polylog} N}$: in that case we can efficiently prepare $|S\rangle$ using postselection, but approximately counting S's restrictions might be hard.

Recently Paul Burchard (personal communication) posed the following question to us: are there any sets that can be QSampled even though they *can't* be approximately counted? Or, better: do QSampling and membership testing, *together*, imply approximate counting? I.e., if we have polynomial-time quantum algorithms to prepare the state $|S\rangle$, and also to decide membership in S, is that enough to let us approximately count S?

Our main result is that, in the black-box setting, the answer to this question is no. More precisely, we show that any quantum algorithm to decide whether $|S| = w$ or $|S| = 2w$, promised that one of those is the case, must either make $\Omega\left(\sqrt{\frac{N}{w}}\right)$ membership queries to S, or else use $\Omega\left(\min\left\{w^{1/4},\sqrt{\frac{N}{w}}\right\}\right)$ copies of $|S\rangle$. So if (for example) we set $w := N^{2/3}$, then any quantum algorithm must either query S or prepare the state $|S\rangle$ at least $\Omega\left(N^{1/6}\right)$ times. This means that there's at most a quadratic speedup compared to classical approximate counting.

In Section 3, we discuss the prospects for improving this lower bound. We conjecture that the bound could be improved, at least to involve $\Omega\left(w^{1/3}\right)$ rather than $\Omega\left(w^{1/4}\right)$, by using stronger results from approximation theory; indeed, a user on MathOverflow already proved one of the requisite results about polynomial degree. However, we also observe that our lower bound *cannot* be improved to involve $\omega\left(w^{1/3}\right)$ without going beyond the polynomial method. While we do not do that in this paper, we give a viable approach: using a hybrid argument, inspired by recent work of Zhandry [23], we show that better lower bounds for our problem would follow from better lower bounds purely on the number of copies of $|S\rangle$ (ignoring the number of queries).

Our lower bounds are within a polynomial factor of the best known quantum upper bounds for approximate counting. As mentioned before, Brassard et al. [9] gave a quantum algorithm to solve the problem using $O\left(\sqrt{\frac{N}{w}}\right)$ queries (and no copies of $|S\rangle$). At the opposite extreme, it's easy to solve the problem using $O\left(\sqrt{w}\right)$ copies of $|S\rangle$ (and no queries), by simply measuring each copy of $|S\rangle$ in the computational basis and then searching for birthday collisions. Alternatively, one can solve the problem using $O\left(\frac{N}{w}\right)$ copies of $|S\rangle$ (and again, no queries), by projecting each copy onto the state $\frac{1}{\sqrt{N}}\left(|1\rangle+\cdots+|N\rangle\right)$ and then counting how many of the projections succeed. We're not aware of any quantum algorithm for approximate counting that combines membership queries with QSampling in an interesting way, though neither can we rule such an algorithm out. Of course, our main result limits the advantage that any such algorithm could achieve.

In our view, at least as interesting as our main result is the technique used to achieve it. In 1998, Beals et al. [6] famously observed that, if a quantum algorithm Q makes T queries to an input X, then Q's acceptance probability can be written as a real multilinear polynomial in the bits of X, of degree at most $2T$. And thus, crucially, if we want to *rule out* a fast quantum algorithm to compute some function $f\left(X\right)$, then it suffices to show that any real polynomial p that approximates f pointwise must have high degree. This general transformation, from questions about quantum algorithms to questions about polynomials, has been used to prove many results that were not known otherwise at the time, including the quantum lower bound for the collision problem [1, 4] and the first direct product theorems for quantum search [2, 16].

In our case, the difficulty is that the quantum algorithm starts with many copies of the state $|S\rangle$. As a consequence of this—and specifically, of the $\frac{1}{\sqrt{|S|}}$ normalizing factor in $|S\rangle$—when we write the average acceptance probability of our algorithm as a function of $|S|$, we find that we get a *Laurent polynomial*: a polynomial that can contain both positive and negative integer powers of $|S|$. The degree of this polynomial (the highest power of $|S|$) encodes the sum of the number of queries and the number of copies of $|S\rangle$, while the "anti-degree" (the highest power of $\frac{1}{|S|}$) encodes the number of copies of $|S\rangle$. We're thus faced with the task of lower-bounding the degree and the anti-degree of a Laurent polynomial that's bounded at integer points and that encodes the approximate counting problem.

We address this using a switching argument that, as far as we know, is new in quantum query complexity. Writing our Laurent polynomial as $q\left(|S|\right)=u\left(|S|\right)+v(\frac{1}{|S|})$, where u and v are ordinary polynomials, we show that, if u and v both have low enough degree (namely, $\deg(u) = o\left(\sqrt{\frac{N}{w}}\right)$ and $\deg(v) = o\left(w^{1/4}\right)$), then we get "unbounded growth" in their values. That is: for approximation theory reasons, either u or v must attain large values, far outside of $[0,1]$, at some integer values of $|S|$. But that means that, for q itself to be bounded in $[0,1]$ (and thus represent a probability), the other polynomial must *also* attain large values. And that, in turn, will force the first polynomial to attain even larger values, and so on forever—thereby proving that these polynomials could not have existed.

We observe that, if we considered the broader class of rational functions, then there *are* rational functions of low degree that represent approximate counting. This follows, for example, from the connection between rational functions and *postselected* quantum algorithms [17], together with Aaronson's $\mathsf{PostBQP} = \mathsf{PP}$ theorem [3]. Thus, our proof relies on the fact that Laurent polynomials are an extremely special kind of rational function.

2 Result

Define $x \in \{0,1\}^N$, the "characteristic string" of the set $S \subseteq [N]$, by $x_i = 1$ if $i \in S$ and $x_i = 0$ otherwise.

Our starting point is the well-known *symmetrization lemma* of Minsky and Papert [18] (see also Beals et al. [6] for its application to quantum query complexity), by which we can often reduce questions about multivariate polynomials to questions about univariate ones.

Lemma 1 (Symmetrization Lemma, Minsky and Papert [18]) *Let $p : \{0,1\}^N \to \mathbb{R}$ be a real multilinear polynomial of degree d, and let*

$$q(k) := \mathrm{E}_{|X|=k}\left[p(X)\right].$$

Then q can be written as a real polynomial in k of degree at most d.

By using Lemma 1, we now prove the key fact that relates quantum algorithms, of the type we're considering, to real Laurent polynomials in one variable. The following lemma generalizes the connection between quantum algorithms and real polynomials established by Beals et al. [6].

Lemma 2 *Let Q be a quantum algorithm that receives R copies of $|S\rangle$ and makes T queries to $\mathcal{O}_S$. Let*

$$q(k) := \mathrm{E}_{|S|=k}\left[\Pr\left[Q^{\mathcal{O}_S}\left(|S\rangle^{\otimes R}\right) \text{ accepts}\right]\right].$$

Then $q : \mathbb{R} \to \mathbb{R}$ is a univariate Laurent polynomial, with maximum exponent at most $2T + R$ and minimum exponent at least $-R$.

Proof. Let $|\psi_t\rangle$ be Q's state immediately after the t^{th} query. Then we can write Q's initial state as

$$|\psi_0\rangle = |S\rangle^{\otimes R} = \frac{1}{|S|^{R/2}} \sum_{i_1,\ldots,i_R \in [N]} x_{i_1} \cdots x_{i_R} \left|i_1,\ldots,i_R\right\rangle.$$

Thus, each amplitude is a complex multilinear polynomial in $X = (x_1,\ldots,x_N)$ of degree R, divided by $|S|^{R/2}$.

Like Beals et al. [6], we now consider how amplitudes change as Q progresses. Each query, to an index $i \in [N]$, multiplies the amplitude of the associated basis state by $1 - 2x_i$, increasing the amplitude's degree as a polynomial by 1. Meanwhile, between the t^{th} and $(t+1)^{st}$ queries, Q can apply an arbitrary unitary transformation U_t, which does not depend on X and hence does not increase degree. Since $x_i^2 = x_i$ for all i, we can also maintain multilinearity without loss of generality.

It follows that Q's final state has the form

$$U_T \left|\psi_T\right\rangle = \sum \alpha_z\left(X\right)\left|z\right\rangle,$$

where each $\alpha_z\left(X\right)$ is a complex multilinear polynomial in X of degree at most $R+T$, again divided by $\left|S\right|^{R/2}$. Since X itself is real-valued, it follows that the real and imaginary parts of $\alpha_z\left(X\right)$, considered individually, are real multilinear polynomials in X of degree at most $R+T$ divided by $\left|S\right|^{R/2}$.

Hence, if we let

$$p\left(X\right) := \Pr\left[Q^{\mathcal{O}_S}\left(\left|S\right\rangle^{\otimes R}\right) \text{ accepts}\right],$$

then

$$p\left(X\right) = \sum_{\text{accepting } z}\left|\alpha_z\left(X\right)\right|^2 = \sum_{\text{accepting } z}\left(\operatorname{Re}^2\alpha_z\left(X\right) + \operatorname{Im}^2\alpha_z\left(X\right)\right)$$

is a real multilinear polynomial in X of degree at most $2\left(R+T\right)$, divided through (in every monomial) by $\left|S\right|^R = \left|X\right|^R$.

Now consider

$$q\left(k\right) := \operatorname{E}_{\left|X\right|=k}\left[p\left(X\right)\right].$$

By Lemma 1, this is a real univariate polynomial in $\left|X\right|$ of degree at most $2\left(R+T\right)$, divided through (in every monomial) by $\left|S\right|^R = \left|X\right|^R$. Or said another way, it's a real Laurent polynomial in $\left|X\right|$, with maximum exponent at most $R+2T$ and minimum exponent at least $-R$. ■

Besides , We'll need several results from approximation theory, each of which has previously been used (in some form) in other applications of the polynomial method to quantum lower bounds. We start with the basic inequality of Markov.

Lemma 3 (Markov) *Let p be a real polynomial, and suppose that*

$$\max_{x,y\in[a,b]}\left|p\left(x\right) - p\left(y\right)\right| \leq H.$$

Then

$$\left|p'\left(x\right)\right| \leq \frac{H}{b-a}\deg\left(p\right)^2$$

for all $x \in [a,b]$.

We'll also need a bound that was explicitly stated by Paturi [19], and which amounts to the folklore fact that, among all degree-d polynomials that are bounded within a given range, the Chebyshev polynomials have the fastest growth outside that range.

Lemma 4 (Paturi) *Let p be a real polynomial, and suppose that $\left|p\left(x\right)\right| \leq 1$ for all $\left|x\right| \leq 1$. Then for all $x \leq 1+\mu$, we have*

$$\left|p\left(x\right)\right| \leq \exp\left(2\deg\left(p\right)\sqrt{2\mu+\mu^2}\right).$$

We now state a useful corollary of Lemma 4, which says (in effect) that slightly shrinking the domain of a low-degree real polynomial can only modestly shrink its range.

Corollary 5 *Let p be a real polynomial of degree d, and suppose that*

$$\max_{x,y\in[a,b]} |p(x) - p(y)| \geq H.$$

Let $\varepsilon \leq \frac{1}{100d^2}$ and $a' := a + \varepsilon(b-a)$. Then

$$\max_{x,y\in[a',b]} |p(x) - p(y)| \geq \frac{H}{2}.$$

Proof. Suppose by contradiction that

$$|p(x) - p(y)| < \frac{H}{2}$$

for all $x, y \in [a', b]$. By affine shifts, we can assume without loss of generality that $|p(x)| < \frac{H}{4}$ for all $x \in [a', b]$. Then by Lemma 4, for all $x \in [a, b]$ we have

$$|p(x)| < \frac{H}{4} \cdot \exp\left(2d\sqrt{2\left(\frac{1}{1-\varepsilon} - 1\right) + \left(\frac{1}{1-\varepsilon} - 1\right)^2}\right) \leq \frac{H}{2}.$$

But this violates the hypothesis. ■

Finally, we'll need a bound that relates the range of a low-degree polynomial on a discrete set of points to its range on a continuous interval. The following lemma generalizes a result due to Ehlich and Zeller [12] and Rivlin and Cheney [20], who were interested only in the case where the discrete points are evenly spaced.

Lemma 6 *Let p be a real polynomial of degree at most $\sqrt{N}$, and let $0 = z_1 < \cdots < z_M = N$ be a list of points such that $z_{i+1} - z_i \leq 1$ for all i (the simplest example being the integers $0, \ldots, N$). Suppose that*

$$\max_{x,y\in[0,N]} |p(x) - p(y)| \geq H.$$

Then

$$\max_{i,j} |p(z_i) - p(z_j)| \geq \frac{H}{2}.$$

Proof. Suppose by contradiction that

$$|p(z_i) - p(z_j)| < \frac{H}{2}$$

for all i, j. By affine shifts, we can assume without loss of generality that $|p(z_i)| < \frac{H}{4}$ for all i. Let

$$c := \max_{x\in[0,N]} \frac{|p(x)|}{H/4}.$$

If $c \leq 1$, then the hypothesis clearly fails, so assume $c > 1$. Suppose that the maximum, $|p(x)| = \frac{cH}{4}$, is achieved between z_i and z_{i+1}. Then by basic calculus, there exists an $x^* \in [z_i, z_{i+1}]$ such that

$$\left|p'(x^*)\right| > \frac{2(c-1)}{z_{i+1} - z_i} \cdot \frac{H}{4} \geq \frac{(c-1)H}{2}.$$

So by Lemma 3,

$$\frac{(c-1)H}{2} < \frac{cH/4}{N} \deg(p)^2.$$

Solving for c, we find

$$c < \frac{2N}{2N - \deg(p)^2} \leq 2.$$

But if $c < 2$, then $\max_{x \in [0,N]} |p(x)| < \frac{H}{2}$, which violates the hypothesis. ■

We're now ready to prove the main result of this paper.

Theorem 7 *Let Q be a quantum algorithm that receives R copies of $|S\rangle$, makes T queries to $\mathcal{O}_S$, and decides whether $|S| = w$ or $|S| = 2w$ with success probability at least $2/3$, promised that one of those is the case. Then either $T = \Omega\left(\sqrt{\frac{N}{w}}\right)$ or*

$$R = \Omega\left(\min\left\{w^{1/4}, \sqrt{\frac{N}{w}}\right\}\right).$$

Proof. Let

$$q(k) := \mathrm{E}_{|S|=k}\left[\Pr\left[Q^{\mathcal{O}_S}\left(|S\rangle^{\otimes R}\right) \text{ accepts}\right]\right].$$

Then by Lemma 2, we can write q as a Laurent polynomial, like so:

$$q(k) = u(k) + v(1/k),$$

where u is a real polynomial in k with $\deg(u) \leq 2T + R$, and v is a real polynomial in $1/k$ with $\deg(v) \leq R$. So to prove the theorem, it suffices to show that either $\deg(u) = \Omega\left(\sqrt{\frac{N}{w}}\right)$, or else $\deg(v) = \Omega\left(w^{1/4}\right)$. To do so, we'll assume that $\deg(u) = o\left(\sqrt{\frac{N}{w}}\right)$ and $\deg(v) = o\left(w^{1/4}\right)$, and derive a contradiction.

Our high-level strategy is as follows: we'll observe that, if approximate counting is successfully being solved, then either u or v must attain a large first derivative somewhere in its domain. By the approximation theory lemmas that we proved earlier, this will force that polynomial to have a large range—even on a subset of integer (or inverse-integer) points. But the sum, $u(k) + v(1/k)$, is bounded in $[0,1]$ for all $k \in [N]$. So if one polynomial has a large range, then the other does too. But this forces the *other* polynomial to have a large derivative somewhere in its domain, and therefore (by approximation theory) to have an even larger range, forcing the first polynomial to have an even larger range to compensate, and so on. As long as $\deg(u)$ and $\deg(v)$ are both small enough, this endless switching will force both u and v to attain *unboundedly* large values—with the fact that one polynomial is in k, and the other is in $1/k$, crucial to achieving the desired "explosion." Since u and v are polynomials on compact sets, such unbounded growth is an obvious absurdity, and this will give us the desired contradiction.

In more detail, we will study the following quantities.

$$\begin{array}{ll}
G_u := \max_{x,y\in\left[\sqrt{w},2w\right]} \left|u\left(x\right)-u\left(y\right)\right| & G_v := \max_{x,y\in\left[\frac{1}{N},\frac{1}{w}\right]} \left|v\left(x\right)-v\left(y\right)\right| \\
\Delta_u := \max_{x\in\left[\sqrt{w},2w\right]} \left|u'\left(x\right)\right| & \Delta_v := \max_{x\in\left[\frac{1}{N},\frac{1}{w}\right]} \left|v'\left(x\right)\right| \\
H_u := \max_{x,y\in\left[\sqrt{w},N\right]} \left|u\left(x\right)-u\left(y\right)\right| & H_v := \max_{x,y\in\left[\frac{1}{N},\frac{1}{\sqrt{w}}\right]} \left|v\left(x\right)-v\left(y\right)\right| \\
I_u := \max_{x,y\in\left[w,N\right]} \left|u\left(x\right)-u\left(y\right)\right| & I_v := \max_{x,y\in\left[\frac{1}{2w},\frac{1}{\sqrt{w}}\right]} \left|v\left(x\right)-v\left(y\right)\right| \\
L_u := \max_{x,y\in\left\{w,\ldots,N\right\}} \left|u\left(x\right)-u\left(y\right)\right| & L_v := \max_{x,y\in\left\{\sqrt{w},\ldots,2w\right\}} \left|v\left(\frac{1}{x}\right)-v\left(\frac{1}{y}\right)\right|
\end{array}$$

We have $0 \leq q\left(k\right) \leq 1$ for all $k \in [N]$, since in those cases $q\left(k\right)$ represents a probability. Since Q solves approximate counting, we also have $q\left(w\right) \leq \frac{1}{3}$ and $q\left(2w\right) \geq \frac{2}{3}$. This means in particular that either

(i) $u\left(2w\right) - u\left(w\right) \geq \frac{1}{6}$, and hence $G_u \geq \frac{1}{6}$, or else

(ii) $v\left(\frac{1}{2w}\right) - v\left(\frac{1}{w}\right) \geq \frac{1}{6}$, and hence $G_v \geq \frac{1}{6}$.

We will show that either case leads to a contradiction.

We have the following inequalities regarding u:

$$\begin{array}{ll}
G_u \geq L_v - 1 & \text{by the boundedness of } q \\
\Delta_u \geq \frac{G_u}{2w} & \text{by basic calculus} \\
H_u \geq \frac{\Delta_u\left(N-\sqrt{w}\right)}{\deg(u)^2} & \text{by Lemma 3} \\
I_u \geq \frac{H_u}{2} & \text{by Corollary 5} \\
L_u \geq \frac{I_u}{2} & \text{by Lemma 6}
\end{array}$$

Here the fourth inequality uses the fact that, setting $\varepsilon := \frac{\sqrt{w}}{N}$, we have $\deg\left(u\right) = o\left(\frac{1}{\sqrt{\varepsilon}}\right)$ (thereby satisfying the hypothesis of Corollary 5), while the fifth inequality uses the fact that $\deg\left(u\right) = o\left(\sqrt{N}\right)$.

Meanwhile, we have the following inequalities regarding v:

$$\begin{array}{ll}
G_v \geq L_u - 1 & \text{by the boundedness of } q \\
\Delta_v \geq G_v w & \text{by basic calculus} \\
H_v \geq \frac{\Delta_v\left(\frac{1}{\sqrt{w}}-\frac{1}{N}\right)}{\deg(v)^2} & \text{by Lemma 3} \\
I_v \geq \frac{H_v}{2} & \text{by Corollary 5} \\
L_v \geq \frac{I_v}{2} & \text{by Lemma 6}
\end{array}$$

Here the fourth inequality uses the fact that, setting $\varepsilon := \frac{1/2w}{1/\sqrt{w}} = \frac{1}{2\sqrt{w}}$, we have $\deg\left(v\right) = o\left(\frac{1}{\sqrt{\varepsilon}}\right)$ (thereby satisfying the hypothesis of Corollary 5). The fifth inequality uses the fact that, if we set $V\left(x\right) := v\left(x/w\right)$, then the situation satisfies the hypothesis of Lemma 6: we are interested in the range of V on the interval $\left[\frac{1}{2},\sqrt{w}\right]$, compared to its range on discrete points $\frac{w}{\sqrt{w}}, \frac{w}{\sqrt{w}+1}, \ldots, \frac{w}{2w}$ that are spaced at most 1 apart from each other; and we also have $\deg\left(V\right) = \deg\left(v\right) = o\left(w^{1/4}\right)$.

All that remains is to show that, if we insert either $G_u \geq \frac{1}{6}$ or $G_v \geq \frac{1}{6}$ into the coupled system of inequalities above, then we get unbounded growth and the inequalities have no solution. Let us collapse the two sets of inequalities to

$$\begin{aligned} L_u &\geq \frac{1}{4} \frac{N - \sqrt{w}}{\deg(u)^2} \frac{G_u}{2w} = \Omega\left(\frac{N}{w \deg(u)^2} G_u\right), \\ L_v &\geq \frac{1}{4} \frac{\frac{1}{\sqrt{w}} - \frac{1}{N}}{\deg(v)^2} G_v w = \Omega\left(\frac{\sqrt{w}}{\deg(v)^2} G_v\right). \end{aligned}$$

Hence

$$\begin{aligned} G_u &\geq L_v - 1 = \Omega\left(\frac{\sqrt{w}}{\deg(v)^2} G_v\right) - 1, \\ G_v &\geq L_u - 1 = \Omega\left(\frac{N}{w \deg(u)^2} G_u\right) - 1. \end{aligned}$$

By the assumption that $\deg(v) = o\left(w^{1/4}\right)$ and $\deg(u) = o\left(\sqrt{\frac{N}{w}}\right)$, we have $\frac{\sqrt{w}}{\deg(v)^2} \gg 1$ and $\frac{N}{w \deg(u)^2} \gg 1$. Plugging in $G_u \geq \frac{1}{6}$ or $G_v \geq \frac{1}{6}$, this is enough to give us unbounded growth. ■

3 Improvements

At our request, user "fedja" on MathOverflow kindly proved the following lemma in approximation theory (see the link[1] for the proof):

Lemma 8 (fedja) *Let p be a real polynomial, and suppose that $|p(1/k)| \leq 1$ for all $k \in [2w]$, and that $p\left(\frac{1}{w}\right) \leq \frac{1}{3}$ while $p\left(\frac{1}{2w}\right) \geq \frac{2}{3}$. Then $\deg(p) = \Omega\left(w^{1/3}\right)$.*

Interestingly, Lemma 8 turns out to be tight. We give the construction for completeness:

Lemma 9 (fedja) *For all w, there is a real polynomial p such that $|p(1/k)| \leq 1$ for all $k \in [2w]$, and $p\left(\frac{1}{w}\right) \leq \frac{1}{3}$ while $p\left(\frac{1}{2w}\right) \geq \frac{2}{3}$, and $\deg(p) = O\left(w^{1/3}\right)$.*

Proof. Assuming for simplicity that w is a perfect cube, consider

$$u(x) := (1 - x)(1 - 2x) \cdots \left(1 - w^{1/3} x\right).$$

Notice that $\deg(u) = w^{1/3}$ and $u\left(\frac{1}{k}\right) = 0$ for all $k \in \left[w^{1/3}\right]$. Furthermore, we have $|u(x)| \leq 1$ for all $x \in \left[0, \frac{1}{w^{1/3}}\right]$, and also $u(x) \in \left[1 - O\left(\frac{1}{w^{1/3}}\right), 1\right]$ for all $x \in \left[0, \frac{1}{w}\right]$. Now, let v be the Chebyshev polynomial of degree $w^{1/3}$, affinely adjusted so that $|v(x)| \leq 1$ for all $x \in \left[0, \frac{1}{w^{1/3}}\right]$ rather than all $|x| \leq 1$, and with a large jump between $\frac{1}{2w}$ and $\frac{1}{w}$. Then the product, $p(x) := u(x) v(x)$, has degree $2w^{1/3}$ and satisfies all the requirements. ■

It seems plausible that, by using Lemma 8, we could give a modest improvement to Theorem 7, which would involve $\Omega\left(w^{1/3}\right)$ rather than $\Omega\left(w^{1/4}\right)$. Unfortunately, there are technical difficulties

[1] See https://mathoverflow.net/questions/302113/real-polynomial-bounded-at-inverse-integer-points

in doing so, since relaxing the assumption $\deg(v) = o\left(w^{1/4}\right)$ to $\deg(v) = o\left(w^{1/3}\right)$ breaks several steps in the proof simultaneously. We leave the details to future work.

In any case, Lemma 9 presumably means that, to prove a lower bound involving $\Omega(\sqrt{w})$, one would need to go beyond the polynomial method. (We say "presumably" because we can't rule out the possibility of using the polynomial method in some way completely different from how it was used in Theorem 7.)

In the remainder of this section, we give what we think is a viable path to going beyond the polynomial method. Specifically, we observe that our problem—of lower-bounding the number of copies of $|S\rangle$ *and* the number of queries to $\mathcal{O}_S$ needed for approximate counting of S—can be reduced to a pure problem of lower-bounding the number of copies of $|S\rangle$. To do so, we use a hybrid argument, closely analogous to an argument recently given by Zhandry [23] in the context of quantum money.

Given a subset $S \subseteq [L]$, let $|S\rangle$ be a uniform superposition over S elements. Then let

$$\rho_{L,w,k} := \mathrm{E}_{S\subseteq[L]\;:\;|S|=w}\left[\left(|S\rangle\langle S|\right)^{\otimes k}\right]$$

be the mixed state obtained by first choosing S uniformly at random subject to $|S| = w$, then taking k copies of $|S\rangle$. Given two mixed states ρ and σ, recall also that the *trace distance*, $\|\rho - \sigma\|_{\mathrm{tr}}$, is the maximum bias with which ρ can be distinguished from σ by a single-shot measurement.

Theorem 10 *Let $2w \leq L \leq N$. Suppose $\|\rho_{L,w,k} - \rho_{L,2w,k}\|_{\mathrm{tr}} \leq \frac{1}{10}$. Then any quantum algorithm Q requires either $\Omega\left(\sqrt{\frac{N}{L}}\right)$ queries to $\mathcal{O}_S$ or else $\Omega(k)$ copies of $|S\rangle$ to decide whether $|S| = w$ or $|S| = 2w$ with success probability at least $2/3$, promised that one of those is the case.*

Proof. Choose a subset $S \subseteq [N]$ uniformly at random, subject to $|S| = w$ or $|S| = 2w$, and consider S to be fixed. Then suppose we choose $U \subseteq [N]$ uniformly at random, subject to both $|U| = L$ and $S \subseteq U$. Consider the hybrid in which Q is still given R copies of the state $|S\rangle$, but now gets oracle access to $\mathcal{O}_U$ rather than $\mathcal{O}_S$. Then so long as Q makes $o\left(\sqrt{\frac{N}{L}}\right)$ queries to its oracle, we claim that Q cannot distinguish this hybrid from the "true" situation (i.e., the one where Q queries $\mathcal{O}_S$) with $\Omega(1)$ bias. This claim follows almost immediately from the BBBV Theorem [7]. In effect, Q is searching the set $[N] \setminus S$ for any elements of $U \setminus S$ (the "marked items," in this context), of which there are $L - |S|$ scattered uniformly at random. In such a case, we know that $\Omega\left(\sqrt{\frac{N-|S|}{L-|S|}}\right) = \Omega\left(\sqrt{\frac{N}{L}}\right)$ quantum queries are needed to detect the marked items with constant bias.

Next suppose we first choose $U \subseteq [N]$ uniformly at random, subject to $|U| = L$, and consider U to be fixed. We then choose $S \subseteq U$ uniformly at random, subject to $|S| = w$ or $|S| = 2w$. Note that this produces a distribution over (S, U) pairs identical to the distribution that we had above. In this case, however, since U is fixed, queries to $\mathcal{O}_U$ are no longer relevant. The only way to decide whether $|S| = w$ or $|S| = 2w$ is by using our copies of $|S\rangle$—of which, by assumption, we need $\Omega(k)$ to succeed with constant bias, even after having fixed U. ■

One might think that Theorem 10 would lead to immediate improvements to our lower bound. In practice, however, the best lower bounds that we currently have, even purely on the number of copies of $|S\rangle$, come from the Laurent polynomial method (Theorem 7)! Having said that, we are optimistic that one could obtain a lower bound that beat Theorem 7 at least when w is small, by combining Theorem 10 with a brute-force computation of trace distance.

4 Discussion and Open Problems

In Theorem 7, can the bound $\min\left\{w^{1/4}, \sqrt{\frac{N}{w}}\right\}$ be tightened to $\min\left\{\sqrt{w}, \frac{N}{w}\right\}$, matching the upper bounds that come from the birthday paradox and projective measurements? If so, then how far can one go toward proving this using the (Laurent) polynomial method, and where does one start to need new techniques?

Also, suppose our task was to distinguish the case $|S| = w$ from the case $|S| = (1+\varepsilon)\,w$, rather than merely w from $2w$. Then what is the optimal dependence on ε? As we said in Section 1, it's known that $O\left(\frac{1}{\varepsilon}\sqrt{\frac{N}{w}}\right)$ quantum queries to $\mathcal{O}_S$ suffice to solve this problem. One can also show without too much difficulty that

$$O\left(\min\left\{\frac{\sqrt{w}}{\varepsilon}, \frac{N}{\varepsilon^2 w}\right\}\right)$$

copies of $|S\rangle$ suffice. On the lower bound side, what generalizations of Theorem 7 can we prove that incorporate ε? We note that our current argument doesn't automatically generalize; one would need to modify something to continue getting growth in the polynomials u and v after the first iteration.

Is there any interesting real-world example of a class of sets for which QSampling and membership testing are both efficient, but approximate counting is not? (I.e., the behavior that this paper showed can occur in the black-box setting?)

Finally, our favorite open problem in this area: can we show that there's no black-box QMA protocol for approximate counting? In other words: that there's no $(\log N)^{O(1)}$-qubit quantum state that Merlin can send to Arthur, so that Arthur becomes convinced after $(\log N)^{O(1)}$ queries that $|S|$ is $2w$ rather than w (promised that one of those is the case)? Arthur's task is "easier" than the task considered in this paper, in that Merlin can send him an arbitrary witness state $|\psi_S\rangle$, rather than just the specific state $|S\rangle$; but also "harder," in that Merlin can cheat and send the wrong state. We thus obtain a problem that's formally incomparable to the one solved here, yet which seems very closely related.

Ruling out a black-box QMA protocol for approximate counting is equivalent to asking for an oracle relative to which $\mathsf{SBP} \not\subset \mathsf{QMA}$, where SBP (Small Bounded-Error Polynomial-Time), defined by Böhler et al. [8], is the complexity class that captures the power of approximate counting. We note that $\mathsf{MA} \subseteq \mathsf{SBP} \subseteq \mathsf{AM}$, and that an oracle relative to which $\mathsf{AM} \not\subset \mathsf{QMA}$ already follows from the work of Vereshchagin [22]. We also note that, under strong derandomization assumptions, we'd have $\mathsf{NP} = \mathsf{MA} = \mathsf{SBP} = \mathsf{AM}$, and hence $\mathsf{SBP} \subseteq \mathsf{QMA}$ in the unrelativized world.

5 Acknowledgments

I'm grateful to Paul Burchard for suggesting the problem to me, and to "fedja" for letting me include Lemmas 8 and 9.

References

[1] S. Aaronson. Quantum lower bound for the collision problem. In *Proc. ACM STOC*, pages 635–642, 2002. quant-ph/0111102.

[2] S. Aaronson. Limitations of quantum advice and one-way communication. *Theory of Computing*, 1:1–28, 2005. Earlier version in CCC'2004. quant-ph/0402095.

[3] S. Aaronson. Quantum computing, postselection, and probabilistic polynomial-time. *Proc. Roy. Soc. London*, A461(2063):3473–3482, 2005. quant-ph/0412187.

[4] S. Aaronson and Y. Shi. Quantum lower bounds for the collision and the element distinctness problems. *J. of the ACM*, 51(4):595–605, 2004.

[5] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proc. ACM STOC*, pages 20–29, 2003. quant-ph/0301023.

[6] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *J. of the ACM*, 48(4):778–797, 2001. Earlier version in FOCS'1998, pp. 352-361. quant-ph/9802049.

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

[8] E. Böhler, C. Glaßer, and D. Meister. Error-bounded probabilistic computations between MA and AM. *J. Comput. Sys. Sci.*, 72(6):1043–1076, 2006.

[9] G. Brassard, P. Høyer, and A. Tapp. Quantum counting. In *Proc. Intl. Colloquium on Automata, Languages, and Programming (ICALP)*, pages 820–831, 1998. arXiv:quant-ph/9805082.

[10] H. Buhrman and R. de Wolf. Complexity measures and decision tree complexity: a survey. *Theoretical Comput. Sci.*, 288:21–43, 2002.

[11] M. E. Dyer, A. M. Frieze, and R. Kannan. A random polynomial time algorithm for approximating the volume of convex bodies. *J. of the ACM*, 38(1):1–17, 1991. Earlier version in STOC'1989.

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

[13] L. Grover and T. Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. arXiv:quant-ph/0208112, 2002.

[14] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proc. ACM STOC*, pages 212–219, 1996. quant-ph/9605043.

[15] M. Jerrum, A. Sinclair, and E. Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries. *J. of the ACM*, 51(4):671–697, 2004. Earlier version in STOC'2001.

[16] H. Klauck, R. Špalek, and R. de Wolf. Quantum and classical strong direct product theorems and optimal time-space tradeoffs. *SIAM J. Comput.*, 36(5):1472–1493, 2007. Earlier version in FOCS'2004. quant-ph/0402123.

[17] U. Mahadev and R. de Wolf. Rational approximations and quantum algorithms with postselection. *Quantum Information and Computation*, 15(3-4):295–307, 2015. arXiv:1401.0912.

[18] M. Minsky and S. Papert. *Perceptrons (2nd edition).* MIT Press, 1988. First appeared in 1968.

[19] R. Paturi. On the degree of polynomials that approximate symmetric Boolean functions. In *Proc. ACM STOC*, pages 468–474, 1992.

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

[21] A. Sinclair and M. Jerrum. Approximate counting, uniform generation and rapidly mixing Markov chains. *Inf. Comput.*, 82(1):93–133, 1989.

[22] N. Vereshchagin. On the power of PP. In *Proc. Conference on Computational Complexity*, pages 138–143, 1992.

[23] M. Zhandry. Quantum lightning never strikes the same state twice. arXiv:1711.02276, 2017.