An Improved Quantum Fourier Transform Algorithm and Applications[*]

Lisa Hales
Group in Logic and the Methodology of Science
University of California at Berkeley
hales@cs.berkeley.edu

Sean Hallgren[†]
Computer Science Division
University of California at Berkeley
hallgren@cs.berkeley.edu

Abstract

We give an algorithm for approximating the quantum Fourier transform over an arbitrary Z_p which requires only $O(n \log n)$ steps where $n = \log p$ to achieve an approximation to within an arbitrary inverse polynomial in n. This improves the method of Kitaev [11] which requires time quadratic in n. This algorithm also leads to a general and efficient Fourier sampling technique which improves upon the quantum Fourier sampling lemma of [8]. As an application of this technique we give a quantum algorithm which finds the period of an arbitrary periodic function, i.e. a function which may be many-to-one within each period. We show that this algorithm is efficient (polylogarithmic in the period of the function) for a large class of periodic functions. Moreover, using standard quantum lower-bound techniques we show that this characterization is tight. That is, this is the maximal class of periodic functions with an efficient quantum period-finding algorithm.

1 Introduction

All known quantum algorithms which give exponential advantage over classical computation are based on the ability of a quantum computer to efficiently solve a hidden subgroup problem[1] over a group with exponentially many elements. A particularly famous example is Shor's algorithm for factoring [13]. Its key step is finding the period of a given periodic function. But this is just a twist on the hidden subgroup problem for cyclic groups, the added twist being the fact that, in this case, the group itself is unknown. In fact, the methodology behind much of the research in quantum algorithms can be summarized as follows: first reduce the question at hand to something resembling a hidden subgroup problem, then extend the existing hidden subgroup machinery to give a quantum algorithm for this particular variant. Some examples of this methodology include [13, 11, 14, 3, 6, 7, 9, 12].

In this paper we focus on the second of these goals: extending the existing hidden subgroup machinery. The key step in solving a hidden subgroup problem is the ability to compute the quantum Fourier transform over the group in question. Our first result is a new and more efficient way of approximating this transform over an arbitrary cyclic group. Our algorithm combines ideas from the eigenvalue estimation procedures of [11] and [4] with methods from [8] and [10] which are based on the relationship between the Fourier transforms over different cyclic groups of an arbitrary fixed superposition. Our methods also yield an extremely general Fourier sampling technique which can be used to solve Shor's variant of the hidden subgroup problem in a particularly clean manner.

We also use this Fourier sampling technique to further extend the existing hidden subgroup machinery. In particular, we address the variant of the standard hidden subgroup problem in which the given function is not necessarily distinct on distinct cosets of the subgroup in question. We give a quantum algorithm for the one-dimensional version of this problem, namely, finding the period of a periodic function which may be many-to-one within each period. While previous results in this direction have been extremely restricted[2], we give a complete characterization of the class of periodic functions whose periods can be efficiently recovered by a quantum machine, giving an efficient algorithm for this class and proving a corresponding lower bound which shows this class to be maximal. Our lower bound result is a slight variant of the standard technique of [1] for establishing a lower bound on the number of ora-

[*]To appear in the the 41st Annual Symposium on Foundations of Computer Science (FOCS), November 2000, Redondo Beach, California.

[†]Supported by an NDSEG Fellowship, a GAANN Fellowship, NSF Grant CCR-9800024, and Air Force Grant F30602-00-2-0601.

[1]The hidden subgroup problem can be described as follows: given access to the values of a function defined on a group and constant on the cosets of some fixed subgroup, find the subgroup.

[2][3] gives a quantum algorithm finding the period of a function which is m-to-one on each period, but m must be both polylogarithmic in the period r and less than the smallest divisor of r. See also [12] for a similar algorithm which takes time quadratic in m.

cle queries in a quantum computation. It appears that our results also extend to the multidimensional version of the same hidden subgroup variant.

Section 2 presents our Fourier transform algorithm and Section 3 the proof of the main lemma establishing the correctness of the algorithm. Section 4 discusses our Fourier sampling technique, and Section 5 gives our period-finding application and the corresponding lower bound.

2 An Approximate Fourier Transform over an Arbitrary Z_p

Let $|\alpha\rangle = \sum_{i<p} \alpha_i |i\rangle$ be a superposition with p an arbitrary positive integer, and let $n = \log p$. Suppose that we wish to perform the quantum Fourier transform of $|\alpha\rangle$ over Z_p. We show how to approximate this transform with high probability by performing just two transforms over powers of 2 together with a classical rounding procedure and a measurement. Moreover, to produce a superposition $|\gamma\rangle$ such that

$$\| |\gamma\rangle - F_p|\alpha\rangle \| < \epsilon$$

the largest transform can be taken to be a power of 2 with just $O\left(n + \log(1/\epsilon)\right)$ bits. Combining this with the method of [5] for ϵ-approximating the Fourier tranform over 2^n in just $O\left(n \log(\frac{n}{\epsilon})\right)$ steps we establish the following:

Theorem 1 *There is a quantum algorithm which ϵ-approximates the quantum Fourier transform over Z_p for an arbitrary n-bit p and any ϵ and which runs in time $O\left(n \log \frac{n}{\epsilon} + \log^2 \frac{1}{\epsilon}\right)$*

This algorithm compares favorably to [11] which requires time quadratic in n to achieve approximations to within an arbitrary inverse polynomial. Our algorithm continues to work for arbitrarily small ϵ but no longer gives a substantial speedup over earlier methods. We maintain, however, that in practice a polynomial approximation suffices for any efficient quantum algorithm due to the unitary evolution of the quantum computation (see [1]). Moreover our algorithm is extremely simple and gives insight into the structure of the Fourier transform itself.

2.1 The Algorithm

We now describe our algorithm given $|\alpha\rangle = \sum_{i<p} \alpha_i |i\rangle$ and parameters s and $q > ps$. We will also require a supply of approximately $\log(q/p) + 1$ clean bits in the second register.

Algorithm 1 *Input:* $|\alpha\rangle$

1. *Compute the Fourier transform over Z_s in the second register:*
$$|\alpha\rangle|0\rangle \longrightarrow |\alpha\rangle \sum_{i<s} \frac{1}{\sqrt{s}} |i\rangle.$$

2. *Repeat $|\alpha\rangle$ s-times as follows:*
$$\begin{aligned} |\alpha\rangle \sum_{i<s} \frac{1}{\sqrt{s}} |i\rangle &= \sum_{j<p} \alpha_j |j\rangle \sum_{i<s} \frac{1}{\sqrt{s}} |i\rangle \\ &\longrightarrow \sum_{j<p, i<s} \frac{1}{\sqrt{s}} \alpha_j |j + ip\rangle \\ &\stackrel{def}{=} |\tilde{\alpha}\rangle. \end{aligned}$$

3. *Transform $|\tilde{\alpha}\rangle$ over Z_q,*
$$|\tilde{\alpha}\rangle \longrightarrow F_q|\tilde{\alpha}\rangle$$

4. *Map*
$$|j\rangle|0\rangle \longrightarrow |i\rangle|t\rangle$$
where $j = \lfloor \frac{q}{p} i \rceil + t$ and $|t| \leq \frac{q}{2p}$[3].

5. *Measure the second register.*

To prove the correctness of the algorithm and consequently Theorem 1, we will need the following lemma, about the superposition $F_q(|\tilde{\alpha}\rangle)$ output by Step 3 of the algorithm.

Let $|\beta\rangle = \sum_{i<p} \beta_i |i\rangle = F_p|\alpha\rangle$ be the superposition we wish to approximate. Then

Lemma 1 *There exists a fixed vector $|\eta\rangle$ supported on the integers in the interval $(-\frac{q}{2p}, \frac{q}{2p})$ such that*

$$\|F_q|\tilde{\alpha}\rangle - \sum_{i<p} \beta_i |\eta^i\rangle\| \leq \frac{4ps}{q} + \frac{8 \ln p}{\sqrt{s}}.$$

where $|\eta^i\rangle$ is the vector $|\eta\rangle$ with indices shifted by $\lfloor \frac{q}{p} i \rceil$.

This lemma says that we can choose s and q so that the output of the algorithm after Step 3, $F_q|\tilde{\alpha}\rangle$, is very close to the vector $\sum_{i<p} \beta_i |\eta^i\rangle$ which has the following structure:

$$\sum_{i<p} \beta_i |\eta^i\rangle = \sum_{i<p} \sum_{|t|<q/2p} \beta_i \eta_t |\lfloor \frac{q}{p} i \rceil + t\rangle$$

Applying Step 4 to this vector would yield

$$\sum_{i<p} \sum_{|t|<q/2p} \beta_i \eta_t |i\rangle|t\rangle = \left(\sum_{i<p} \beta_i |i\rangle \right) \left(\sum_{|t|<q/2p} \eta_t |t\rangle \right),$$

and clearly a measurement of the last register would yield the exact superposition

$$\sum_{i<p} \beta_i |i\rangle$$

[3] Ties can be broken arbitrarily

which we desire to approximate.

Since $F_q|\tilde{\alpha}\rangle$ is close to the superposition $\sum_{i<p}\beta_i|\eta^i\rangle$ the output of the algorithm should be close to this desired superposition with high probability. More formally we can prove the following:

Corollary 1 *Let $|\gamma_t\rangle$ be the output of the algorithm above. For any $\epsilon > 0$, if $s = \Omega\left(\frac{\ln^2 p}{\epsilon^4}\right)$ and $q = \Omega\left(\frac{ps}{\epsilon^2}\right)$, then with probability at least $1-\epsilon^2$,*

$$\||\gamma_t\rangle - F_p|\alpha\rangle\| < \epsilon.$$

Notice that we are free to choose s and q to be powers of 2 and that q can be chosen to have $O(n + \log\frac{1}{\epsilon})$ bits as claimed previously. Furthermore, notice that Step 2 consists of multiplying an n-bit number by an $m = O\left(\log n + \log\frac{1}{\epsilon}\right)$-bit number, requiring less than $O(nm)$ steps. A similar argument can be made for the division in Step 4. Thus the Fourier transform over Z_q is the computational bottleneck in the algorithm. Thus Theorem 1 follows immediately from this corollary together with the results of [5]. The proof of the corollary uses a standard averaging argument and is in Appendix A.

3 Proof of Lemma 1

Recall that $|\beta\rangle = \sum_{i<p}\beta_i|i\rangle = F_p|\alpha\rangle$. Then $F_{ps}|\tilde{\alpha}\rangle = \sum_{i<p}\beta_i|is\rangle$ and $F_q|\tilde{\alpha}\rangle = \sum_{i<p}\beta_i F_q(F_{ps}^{-1}(|is\rangle))$.

For convenience we shall denote $F_q(F_{ps}^{-1}(|is\rangle))$ by $|\delta_i\rangle$, and thus

$$F_q|\tilde{\alpha}\rangle = \sum_{i<p}\beta_i|\delta_i\rangle.$$

Recall that we are trying to show that there exists some superposition $|\eta\rangle$ with support contained in the integers in the interval $(-\frac{q}{2p}, \frac{q}{2p})$ such that $F_q|\tilde{\alpha}\rangle$ is close to $\sum_{i<p}\beta_i|\eta^i\rangle$ where $|\eta^i\rangle$ is $|\eta\rangle$ shifted by $\lfloor\frac{q}{p}i\rceil$. To this end we shall show that $|\delta_i\rangle$ is very close to $|\delta_0\rangle$ shifted by $\lfloor\frac{q}{p}i\rceil$ and that $|\delta_0\rangle$ has almost all of its amplitude on the integers between $-\frac{q}{2p}$ and $\frac{q}{2p}$.

Let the vector $|B_i\rangle$, for "bump", be the superposition $|\delta_i\rangle$ restricted to the integers in the interval $(\frac{q}{p}i - \frac{q}{2p}, \frac{q}{p}i + \frac{q}{2p})$, and $|T_i\rangle = |\delta_i\rangle - |B_i\rangle$ be the tails of $|\delta_i\rangle$ outside this set.

Note that

$$F_q|\tilde{\alpha}\rangle = \sum_{i<p}\beta_i|B_i\rangle + \sum_{i<p}\beta_i|T_i\rangle.$$

We begin by bounding $\|\sum_{i<p}\beta_i|T_i\rangle\|$, thus showing that $\|F_q|\tilde{\alpha}\rangle - \sum_{i<p}\beta_i|B_i\rangle\|$ is very small. We will then finish off the proof by showing that each $|B_i\rangle$ is very close to $|B_0\rangle$ shifted by $\lfloor\frac{q}{p}i\rceil$.

More formally, we will prove the following two claims:

Claim 1

$$\|F_q|\tilde{\alpha}\rangle - \sum_{i<p}\beta_i|B_i\rangle\| = \|\sum_{i<p}\beta_i|T_i\rangle\| \le \frac{8\ln p}{\sqrt{s}}$$

The claim states that making s large (i.e. increasing the number of repetitions of $|\alpha\rangle$) reduces the effect of the tails.

Claim 2 *Let $|B_0^i\rangle$ be the superposition $|B_0\rangle$ shifted by $\lfloor\frac{q}{p}i\rceil$. Then*

$$\||B_0^i\rangle - |B_i\rangle\| < \frac{4ps}{q}.$$

From Claim 2 and the fact that the $|B_i\rangle$ have disjoint supports,

$$\|\sum_{i<p}\beta_i|B_0^i\rangle - \sum_{i<p}\beta_i|B_i\rangle\| \le \frac{4ps}{q}.$$

Combining this with claim 1 via the triangle inequality we have,

$$\|F_q|\tilde{\alpha}\rangle - \sum_{i<p}\beta_i|B_0^i\rangle\| \le \frac{4ps}{q} + \frac{8\ln p}{\sqrt{s}},$$

as desired.

3.1 Proof of Claim 1

Proof: In order to bound $\|\sum_{i<p}\beta_i|T_i\rangle\|$ we will use the following observation which establishes that the amplitudes in $|\delta_i\rangle$ fall off quickly away from $\lfloor\frac{q}{p}i\rceil$. The proof of Observation 1 is in Appendix B

Observation 1

$$|(|\delta_i\rangle)_j| = \left|\frac{1}{q}\frac{1}{ps}\sum_{k<ps}\omega_q^{k(j-\frac{q}{p}i)}\right| \le \frac{\sqrt{q}}{\sqrt{ps}}\frac{2}{\left|j-\frac{q}{p}i\right|_q}$$

where $|x|_q = \begin{cases} x \bmod q & \textit{if } 0 \le x \bmod q \le q/2 \\ -x \bmod q & \textit{otherwise} \end{cases}$

We now use this to bound $\|\sum_{i<p}\beta_i|T_i\rangle\|$. The first equality below is by the definition of $|T_i\rangle$ and the second is by the above observation:

$$\begin{aligned}\|\sum_{i<p}\beta_i|T_i\rangle\|^2 &= \sum_{j<q}\Big|\sum_{i\neq\lfloor\frac{p}{q}j\rceil}\beta_i(|T_i\rangle)_j\Big|^2 \\ &\le \sum_{j<q}\frac{4q}{ps}\left(\sum_{i\neq\lfloor\frac{p}{q}j\rceil}|\beta_i|\frac{1}{\left|j-\frac{q}{p}i\right|_q}\right)^2.\end{aligned}$$

This expression is almost maximized by taking the $\beta_i = 1/\sqrt{p}$ for all i. In particular, the expression can be bounded by four times its value at this vector. A proof of this is in Appendix C. This bound is not actually necessary to prove the time bounds in Theorem 1–the bound found by taking $\beta_i = 1$ suffices–but it does improve the constants.

$$\|\sum_{i<p} \beta_i |T_i\rangle\|^2 \leq \sum_{j<q} \frac{16q}{p^2 s} \left(\sum_{i \neq \lfloor \frac{p}{q} j \rceil} \frac{1}{\left| j - \frac{q}{p} i \right|_q} \right)^2 .$$

Using the fact that the smallest denominator $\left| j - \frac{q}{p} i \right|_q$ is at least $\frac{q}{2p}$ and the rest are spaced out by $\frac{q}{p}$ we have

$$\sum_{i \neq \lfloor \frac{p}{q} j \rceil} \frac{1}{\left| j - \frac{q}{p} i \right|_q} \leq \frac{4p \ln p}{q}.$$

Therefore

$$\|F_q|\tilde{\alpha}\rangle - \sum_{i<p} \beta_i |B_i\rangle\| = \|\sum_{i<p} \beta_i |T_i\rangle\| \leq \frac{8 \ln p}{\sqrt{s}}, \qquad (1)$$

as desired. ■

3.2 Proof of Claim 2

Proof: We first note that to show that $\||B_0^i\rangle - |B_i\rangle\| < \frac{4ps}{q}$ it suffices to show that $\||\delta_0^i\rangle - |\delta_i\rangle\| < \frac{4ps}{q}$, where $|\delta_0^i\rangle$ is $|\delta_0\rangle$ shifted by $\lfloor \frac{q}{p} i \rceil$.

But

$$\begin{aligned} \||\delta_0^i\rangle - |\delta_i\rangle\|^2 &= \|F_q^{-1}|\delta_0^i\rangle - F_q^{-1}|\delta_i\rangle\|^2 \\ &= \sum_{j<ps} \left| \frac{1}{\sqrt{ps}} \omega_p^{-\lfloor \frac{q}{p} i \rceil j} - \frac{1}{\sqrt{ps}} \omega_q^{-ij} \right|^2 \\ &\leq \sum_{j<ps} \left| \frac{1}{\sqrt{ps}} \omega_p^{-ij} \left(\omega_q^{\epsilon_i j} - 1 \right) \right|^2 \end{aligned}$$

where $\epsilon_i = \lfloor \frac{q}{p} i \rceil - \frac{q}{p} i$. But since j only goes up to ps,

$$\left| \omega_q^{\epsilon_i j} - 1 \right| < \frac{4ps}{q}$$

and thus

$$\begin{aligned} &\sum_{j<ps} \left| \frac{1}{\sqrt{ps}} \omega_p^{-ij} \left(\omega_q^{\epsilon_i j} - 1 \right) \right|^2 \\ &\quad \leq \sum_{j<ps} \left| \frac{1}{\sqrt{ps}} \omega_p^{-ij} \right|^2 \left| \omega_q^{\epsilon_i j} - 1 \right|^2 \\ &\quad \leq \left(\frac{4ps}{q} \right)^2, \end{aligned}$$

as desired. ■

4 Fourier Sampling

In many quantum algorithms (see [2, 13, 3, 14]) the Fourier transform occurs as the final quantum step and a measurement of the superposition immediately follows. We refer to this procedure as Fourier sampling [2]. Suppose that we wish to sample $F_p|\alpha\rangle$ for some given p and $|\alpha\rangle$. In this situation we need only insure that the distribution we sample from is ϵ-close to the distribution induced by $F_p|\alpha\rangle$, i.e. we need not worry about the phases of the amplitudes in the final superposition. Because of this we need merely perform the first 3 steps of Algorithm 1, then measure the resulting superposition. We can then perform the rounding from Step 4 classically and output the resulting i. More precisely,

Algorithm 2 *Input:* $|\alpha\rangle$

1. *Compute the Fourier transform over Z_s in the second register:*
$$|\alpha\rangle|0\rangle \longrightarrow |\alpha\rangle \sum_{i<s} \frac{1}{\sqrt{s}} |i\rangle.$$

2. *Repeat $|\alpha\rangle$ s-times as follows:*
$$\begin{aligned} |\alpha\rangle \sum_{i<s} \frac{1}{\sqrt{s}} |i\rangle &= \sum_{j<p} \alpha_j |j\rangle \sum_{i<s} \frac{1}{\sqrt{s}} |i\rangle \\ &\longrightarrow \sum_{j<p, i<s} \frac{1}{\sqrt{s}} \alpha_j |j + ip\rangle \\ &\stackrel{def}{=} |\tilde{\alpha}\rangle. \end{aligned}$$

3. *Transform $|\tilde{\alpha}\rangle$ over Z_q:*
$$|\tilde{\alpha}\rangle \longrightarrow F_q|\tilde{\alpha}\rangle$$

4. *Measure.*

We then compute and ouput the integer i which minimizes[4] *the quantity*

$$\left| j - \frac{q}{p} i \right|,$$

where j was the measured value.

Let $\mathcal{D}_{F_p|\alpha\rangle}$ be the distribution on $\{0, ..., (p-1)\}$ induced by measuring $F_p|\alpha\rangle$ and let $\mathcal{D}$ be the distribution induced by Algorithm 2.

Using the following lemma from [2]:

[4] Ties can be broken arbitrarily

Lemma 2 *[[2],Lemma 3.2.6] Let $|\phi\rangle$ and $|\psi\rangle$ be two unit length vectors with $\||\phi\rangle - |\psi\rangle\| \leq \epsilon$. Then the total variation distance (denoted by $\| * \|_1$) between the probability distributions resulting from measurements of $|\phi\rangle$ and $|\psi\rangle$ is at most 4ϵ.*

and the results of the previous section it is easy to see that

Corollary 2 *Suppose that $s = \Omega\left(\frac{\ln^2 p}{\epsilon^2}\right)$ and $q = \Omega\left(\frac{ps}{\epsilon}\right)$, then*

$$\|\mathcal{D}_{F_p|\alpha\rangle} - \mathcal{D}\|_1 < \epsilon.$$

4.1 Fourier Sampling When Z_p is Unknown

In some of the most interesting quantum algorithms which use Fourier sampling neither p nor $|\alpha\rangle$ is known. In particular, it is often possible, even without knowledge of p or $|\alpha\rangle$, to generate a superposition of arbitrary length which consists of repetitions of $|\alpha\rangle$. In this case Fourier sampling can yield useful information about p and $|\alpha\rangle$. The goal in such cases is to sample from the set of proper fractions with denominator p and with numerators distributed according to $\mathcal{D}_{F_p|\alpha\rangle}$.[5]

Our initial superposition in this case is $|\tilde{\alpha}\rangle$, that is, $|\alpha\rangle$ repeated some (possibly non-integral) number of times. After computing the Fourier transform over Z_q of this superposition we make a measurement. We are then left with the task of computing i such that

$$|j - \frac{q}{p}i| < \frac{q}{2p}.$$

More precisely,

Algorithm 3 *We begin with the input $|\tilde{\alpha}\rangle$ where*

$$|\tilde{\alpha}\rangle = \sum_{k<ps} \alpha_{(k \bmod p)}|k\rangle$$

1. *Transform $|\tilde{\alpha}\rangle$ over Z_q:*

$$|\tilde{\alpha}\rangle \longrightarrow F_q|\tilde{\alpha}\rangle$$

2. *Measure.*

Let j be the measured value. We wish to compute and output the fraction i/p where i minimizes

$$\left|j - \frac{q}{p}i\right|.$$

Unfortunately, the task of minimizing $\left|j - \frac{q}{p}i\right|$ is now non-trivial since p is not known. Suppose that t is some known upper bound on p. We can use the continued fraction expansion of $\frac{j}{q}$ to get the fraction with denominator less than t which is very close to $\frac{j}{q}$. But we this will only be the desired fraction $\frac{i}{p}$ if we can guarantee that

$$\left|j - \frac{q}{p}i\right| < \frac{q}{2t^2}.$$

Fortunately, we can in fact tighten the results of the previous section in order to guarantee this with high probability. We first make the following claim which follows by bounding the appropriate sum using Observation 1. The proof is left to the reader.

Let $|B_0\rangle$ be as defined previously and let $|B_0\rangle_{\frac{q}{2t^2}}$ be $|B_0\rangle$ restricted to the set of integers between $-\frac{q}{2t^2}$ and $+\frac{q}{2t^2}$. Then

Claim 3

$$\||B_0\rangle - |B_0\rangle_{\frac{q}{2t^2}}\| \leq \frac{4t}{\sqrt{ps}}$$

In the case that s is integral we can use Lemma 1 which, together with Claim 3, yields

$$\|F_q|\tilde{\alpha}\rangle - \sum_{i<p} \beta_i |B_0^i\rangle_{\frac{q}{2t^2}}\| \leq \frac{4ps}{q} + \frac{8\ln p}{\sqrt{s}} + \frac{4t}{\sqrt{ps}}. \quad (2)$$

When s is non-integral we can just use the fact that $|\alpha\rangle$ repeated s times and $|\alpha\rangle$ repeated $\lfloor s\rceil$ times have l_2-distance equal to $O(1/s)$ to get the above bound plus an extra term of size $O(1/s)$.

We denote by $\hat{\mathcal{D}}_{F_p|\alpha\rangle}$ the distribution on proper fractions with denominator p and numerators distributed according to $\mathcal{D}_{F_p|\alpha\rangle}$, and we let $\hat{\mathcal{D}}$ be the distribution over the larger set of fractions with denominator less that t output by Algorithm 3 followed by a continued fraction expansion on parameters t, s, and p. Then the following corollary follows from Equation 2 together with Lemma 2:

Corollary 3 *Suppose that $s = \Omega(\frac{t^2}{\epsilon^2 p})$ and $q = \Omega(\frac{ps}{\epsilon})$, then*

$$\|\hat{\mathcal{D}}_{F_p|\alpha\rangle} - \hat{\mathcal{D}}\|_1 < \epsilon.$$

We will apply this corollary in Section 4.2 to give a new proof of the correctness of Shor's factoring algorithm and in Section 5 to prove the correctness of our new period-finding algorithm.

4.2 Fourier Sampling 1-1 Periodic Functions: Shor's Algorithm Reproved

We give a new proof of the correctness of Shor's algorithm for finding the period of a one-to-one periodic function in terms of our Fourier sampling technique. This is a

[5]Recall that $\mathcal{D}_{F_p|\alpha\rangle)}$ is the distribution on $\{0, ..., (p-1)\}$ induced by measuring $F_p(|\alpha\rangle)$

nice application of the previous section and it also is useful as a reference for understanding the more complicated many-to-one case discussed in Section 5. Let f be a one-to-one periodic function with period $p < 2^n$.

Algorithm 4 *Let* $s = \Omega\left(\frac{2^{2n}}{p\epsilon^2}\right)$ *and* $q = \Omega\left(\frac{ps}{\epsilon}\right)$.

1. *We prepare the input to the Fourier transform as follows:*

$$|0\rangle|0\rangle \longrightarrow \sum_{i<sp} |i\rangle|f(i)\rangle \longrightarrow \sum_{j<s} |i_0 + jp\rangle|a\rangle$$

where the last transformation comes from measuring the value of f.

2. *Fourier transform over* Z_q*:*

$$\sum_{j<s} |i_0 + jp\rangle|a\rangle \longrightarrow F_q(\sum_{j<s} |i_0 + jp\rangle)|a\rangle$$

3. *Measure the first register.*

Divide the output by q*, and then use continued fractions to round to the nearest fraction with denominator less than* 2^n.

Notice that the input to the Fourier transform is of the form $|\tilde{\alpha}\rangle$ for $|\alpha\rangle = |i_0\rangle$. Since for any i_0 $F_p(|i_0\rangle)$ has amplitudes with norms identically equal to $\frac{1}{\sqrt{p}}$, measuring $F_p|\alpha\rangle$ gives the uniform distribution on the set $\{0, \ldots, (p-1)\}$. By Corollary 3 we can conclude that the distribution of outputs from our procedure is ϵ-close to the uniform distribution over proper fractions with denominator p.

We know that for some constant c a $\frac{c}{\log\log p}$ fraction of the integers less than p are relatively prime to p, thus as long as our distribution is $\frac{c/2}{\log\log p}$-close to the uniform distribution over proper fractions with denominator p we will see the correct period with significant probability after just $O(\log\log p)$ repetitions of the algorithm.

As in [13] we can test the denominator of the fraction output by our procedure to see if it is the period by evaluating the function at just two values. Since the distribution of outputs from our procedure is actually over the set of fractions with denominators less than 2^n (as opposed to denominator p) there is a small (less than ϵ at each measurement) chance that will see a denominator which is a multiple of the period and such a denominator will pass our test. But by repeating the algorithm a sufficient number of times (as opposed to terminating as soon as the test is passed) and taking the greatest common divisor of all the denominators which pass our test we will retrieve the true period with very high probability. Another alternative is to make ϵ so small that even after $O(\log\log p)$ repetitions of the algorithm we are unlikely to have seen any fractions outside the desired set.

5 Fourier Sampling Arbitrary Periodic Functions

We now give the period-finding algorithm which is an application of the Fourier sampling technique from the previous section. Throughout our discussion f and g will denote periodic functions with domain the positive integers and range $\{0, ..., 2^n - 1\}$. We assume that their periods, denoted p_f and p_g respectively, satisfy $p_f, p_g < 2^n$. For convenience we shall drop the subscript, referring to the period of f as just p, whenever possible without confusion.

We emphasize that we allow our functions to be many-to-one within each period, a characteristic which we shall refer to as many-to-one periodic. Note that for any periodic function f there are many-to-one periodic functions g which agree with f on all but an exponentially small fraction of values but have strictly larger periods. From the lower bound results of [1], which are based on the unitary evolution of quantum computation, we know that quantum algorithms which can distinguish between exponentially close functions typically require exponentially many queries to the function's values. Thus we should expect that no efficient, i.e. polynomial in n, quantum algorithm can correctly determine the period of all such many-to-one periodic functions. Our lower bound result found in Section 5.2 formalizes this intuition.

Our first task, then, is to delineate the collection of classes of periodic functions which **do** have efficient quantum period-finding algorithms. To this end we make the following definitions:

We first define a metric on periodic functions. Intuitively, f and g are ϵ-close if they disagree on at most an ϵ fraction of their values.

Definition 1 *Let* f *and* g *be as above. Then* $\boldsymbol{D}(f, g)$ *is the fraction of points in the set* $\{0, ..., p_f p_g - 1)\}$ *satisfying* $f(x) \neq g(x)$.

Using this metric we define the following classes of functions.

Definition 2 *For any function* $d(n)$ *let*

$$C_{1/d(n)} = \{f | \forall g \text{ with } p_g < p_f,\ D(f, g) > 1/d(n)\}.$$

Informally, then, f is in $C_{1/d(n)}$ if in order to reduce the period of f one must change at least a $1/d(n)$ fraction of its values. Such an f can be viewed as being $1/d(n)$-robust with respect to its period. Now we can state the theorem.

Theorem 2 *Given any polynomial* $d(n)$ *there is an efficient quantum algorithm* $\boldsymbol{A}$[6] *which computes the period of any* $f \in C_{1/d(n)}$ *with probability at least* $3/4$.

[6]Throughout the paper we will assume that $\boldsymbol{A}$ has a blackbox subroutine for computing values of f

We note that by testing values of the function we can rule out the incorrect outputs of any such algorithm unless they happen to be multiples of the period. But by repeating the procedure and taking the greatest common divisor of all outputs which pass our test we can amplify the success probability of the algorithm to $1 - 2^{-e(n)}$ for any polynomial $e(n)$.

We also show in Section 5.2 that the restriction to functions in $C_{1/d(n)}$ for $d(n)$ a polynomial is necessary:

Theorem 3 *Let $d(n) = o(2^n)$ be given. Suppose that $\boldsymbol{A}$ is a quantum algorithm which correctly computes the period of any $f \in C_{1/d(n)}$ with probability at least $3/4$. Then $\boldsymbol{A}$ has worst-case run-time $\Omega(\sqrt[4]{d(n)})$.*

5.1 Fourier Sampling Many-to-One Periodic Functions

We now give our algorithm for finding the period of an arbitrary periodic function. Let $f \in C_{1/d(n)}$ with period $p < 2^n$ be given.

Algorithm 5 *Let $s = \Omega\left(\frac{2^{2n}}{p\epsilon^2}\right)$ and $q = \Omega\left(\frac{ps}{\epsilon}\right)$*

1. *We prepare the input to the Fourier transform as follows:*

$$|0\rangle|0\rangle \longrightarrow \sum_{i<sp} |i\rangle|f(i)\rangle \longrightarrow \sum_{i<sp, f(i)=a} |i\rangle|a\rangle$$

where the last transformation comes from measuring the value of f.

2. *Fourier transform over Z_q:*

$$\sum_{i<sp, f(i)=a} |i\rangle|a\rangle \to F_q\left(\sum_{i<sp, f(i)=a} |i\rangle\right)|a\rangle$$

3. *Measure the first register, divide the output by q, and then use continued fractions to round to the nearest fraction with denominator less than 2^n.*

Repeat the above procedure t-times obtaining the fractions $j_1/r_1, \dots, j_t/r_t$. Output the least common multiple of $r_1, \dots, r_t$.

We will show that if t is chosen to be a sufficiently large polynomial then with very high probability Algorithm 5 outputs the correct period p. Our analysis proceeds just as in the previous section. We first note that, since f is periodic with period p the input to the Fourier transform is again of the form $|\tilde{\alpha}\rangle$, namely it is some superposition $|\alpha\rangle$ over the integers $\{0, ..., (p-1)\}$ repeated s-times. Thus we can use Corollary 3 to say that, for a fixed $|\tilde{\alpha}\rangle$, the distribution of outputs of the continued fraction procedure in Step 3 is ϵ-close to the distribution $\hat{\mathcal{D}}_{F_p|\alpha\rangle}$ [7].

There are two new difficulties which arise in the many-to-one case. The first, which is merely technical, is that while in the one-to-one periodic function case $\mathcal{D}_{F_p|\alpha\rangle}$ did not depend on $|\alpha\rangle$, since all possible $|\alpha\rangle$ gave rise to the uniform distribution, this no longer holds for an arbitrary periodic function. Thus we now need to analyze the distribution $\mathcal{D}_{F_p|\alpha\rangle}$ where $|\alpha\rangle$ itself is a random variable distributed according to the measurement at the end of Step 1. We shall denote this distribution by $\mathcal{D}_{F_p(f)}$ since it is the distribution on the set $\{0, ..., (p-1)\}$ induced by Fourier sampling over Z_p the output of the procedure

$$\sum_{i<p} |i\rangle|f(i)\rangle \quad \xrightarrow{\text{measure f}} \quad \sum_{i<p, f(i)=a} |i\rangle|a\rangle.$$

It is easy to see that Corollary 3 implies that the distribution output by Algorithm 5 is ϵ-close to $\hat{\mathcal{D}}_{F_p(f)}$.

By choosing ϵ sufficiently small, then, we can effectively assume that the outputs of Step 3 of our algorithm are contained in the set of proper fractions with denominator p and distributed according to $\hat{\mathcal{D}}_{F_p(f)}$. We are then left with a deeper problem. Since, in contrast to the one-to-one case, $|\alpha\rangle = \sum_{i<p, f(i)=a} |i\rangle$ is not necessarily a simple point-mass, $\mathcal{D}_{F_p(f)}$ is not necessarily uniform. Recall that in the previous algorithm we used the fact our outputs were uniform over the proper fractions with denominator p to conclude that our quantum procedure had a significant chance of outputting a fraction $\frac{i}{p}$ such that $gcd(i,p) = 1$, enabling us to recover the period. Unfortunately, in the many-to-one case this is not true. In fact, one can construct periodic functions satisfying our hypotheses for which the outputs $\frac{i}{p}$ of the sampling procedure **never** satisfy $gcd(i,p) = 1$.

Recall, however, that we are taking the least common multiple of all the denominators output by our algorithm. The only way this can fail to be p is if there is some fixed divisor a of p such that **every** fraction output by our algorithm is of the form ia/p. Fortunately, we can show that by repeating the algorithm a sufficient number of times the likelihood of this will be very small. In particular, we can prove the following Lemma:

Lemma 3 *Suppose that $f \in C_{1/d(n)}$ and j is chosen according to $\mathcal{D}_{F_p(f)}$. Then for every prime a dividing p,*

$$\Pr(a|j) < 1 - 1/8d^2(n).$$

The proof of the Lemma can be found in Appendix D. The idea is that if there is a a dividing p such that a divides

[7] Recall that $\mathcal{D}_{F_p|\alpha\rangle}$ is the distribution on $\{0, ..., (p-1)\}$ induced by measuring $F_p|\alpha\rangle$ and $\hat{\mathcal{D}}_{F_p|\alpha\rangle}$ is the distribution on proper fractions with denominator p and numerators distributed according to $\mathcal{D}_{F_p|\alpha\rangle}$.

a sample of $\mathcal{D}_{F_p(f)}$ with extremely high probability then the function f must be very close to a function with period p/a, violating the assumption that $f \in C_{1/d(n)}$.

We can now formalize the discussion in the previous paragraphs and prove the correctness of the above algorithm for $\epsilon = 1/d^3(n)$ and $t = 20d^2(n)\log n$.

Proof: Let $\hat{\mathcal{D}}$ be the distribution on fractions with denominators less than 2^n output by the algorithm. By Corollary 3 and our choice of ϵ, with very high probability (at least $1 - t\epsilon = 1 - 20/d(n)$) all the fractions we see are from the set of proper fractions with denominator p. Let us assume this is the case and let us denote them by $i_1/p, ..., i_t/p$.

Again by the corollary and choice of ϵ together with Lemma 3 we know that if i is chosen according to $\hat{\mathcal{D}}$ then for every prime a dividing p,

$$\Pr(a|i) < 1 - 1/8d^2(n) + \epsilon < 1 - 1/9d^2(n).$$

Since there are at most n primes dividing p, the probability that there exists a prime a dividing p which also divides all t of the numerators is less than

$$n\left(1 - 1/9d^2(n)\right)^t < 1/7.$$

Thus when we take the least common multiple of the denominators of our fractions we will see p with probability at least $3/4$, as desired. ■

5.2 Proof of Theorem 3

We need the following definition and theorem from [1]

Definition 3 *Let $|\phi_i\rangle$ be the superposition of A^f on input x at time i. We denote by $q_y(|\phi_i\rangle)$ the sum of squared magnitudes in $|\phi_i\rangle$ of configurations of M which are querying the oracle on string y.*

Theorem 4 *Let $|\phi_i\rangle$ be the superposition of A^f on input x at time i. Let $\epsilon > 0$. Let $S \subseteq [0, T-1] \times \Sigma_*$ be a set of time-strings pairs such that $\sum_{(i,y)\in S} q_y(|\phi_i\rangle) \leq \frac{\epsilon^2}{T}$. Now suppose the answer to each query $(i, y) \in S$ is modified to some arbitrary fixed $a_{i,y}$ (these answers need not be consistent with an oracle). Let $|\phi_i'\rangle$ be the time i superposition of A on input x with oracle answers modified as stated above. Then $\||\phi_i\rangle - |\phi_i'\rangle\| \leq \epsilon$.*

In our case we wish to use Theorem 4 to show that if a quantum algorithm correctly computes the period of any $f \in C_{1/d(n)}$ with constant probability then it must make at least $\Omega(\sqrt[4]{d(n)})$ queries to the function's values. To this end we first look at the algorithm's behavior when $f(x) = 0$ for all x (Note that the all-zeroes function is in every class $C_{1/d(n)}$).

We wish to use this behavior to generate a function $g \in C_{1/d(n)}$ which has period greater than 1 and which the algorithm cannot distinguish from the all-zeroes function without making lots of queries. This is similar to earlier applications of Theorem 4 but with the added complication that g must be periodic and at least $1/d(n)$ away from any function of smaller period. We ensure periodicity by first deciding on the period p of g and then changing the value of the function simultaneously on all points of the form $x + kp$. We show that the latter complication can be resolved by choosing g to have prime period and to be sufficiently different from the all-zeroes function.

Proof: (Proof of Theorem 3) Given A computing the period of any function in $C_{1/d(n)}$ in time T, we initially examine A^o where o is the all-zeroes function.

Fix a prime p such that $\sqrt{d(n)} < p < 2^n$. For $0 \leq x < p$ let

$$S_x = [0, T-1] \times \{y | y = x + kp\}.$$

The average value of $\sum_{(i,y)\in S_x} q_y(|\phi_i\rangle)$ is $\frac{T}{p}$ and thus at least $1/2$ of the sets S_x satisfy

$$\sum_{(i,y)\in S_x} q_y(|\phi_i\rangle) \leq 2\frac{T}{p}. \tag{3}$$

Let T be any set of $3p/\sqrt{d(n)}$ x which satisfy (3). We let our new function g satisfy $g(x + kp) = 1$ for $x \in T$ and $g(x) = o(x) = 0$ otherwise. Note that $g(x)$ has period our chosen prime p and that $D(o, g) \geq 3/\sqrt{d(n)}$.

Furthermore, let $S_T = \bigcup_{x\in T} S_x \subseteq [0, T-1] \times \Sigma_*$. Then

$$\sum_{(i,y)\in S_T} q_y(|\phi_i\rangle) \leq \frac{6T}{\sqrt{d(n)}},$$

and we can take the ϵ of Theorem 4 to be $\frac{\sqrt{6}T}{\sqrt[4]{d(n)}}$. Thus in order for our algorithm A to distinguish between the all-zeros function o and our new period-p function g with constant probability, A must have worst-case run-time $\Omega(\sqrt[4]{d(n)})$.

To prove our theorem, however, we need to verify that our function g is actually in $C_{1/d(n)}$. We need the following claim whose simple proof is in Appendix E.

Claim 4 *For any periodic functions f and g, if $D(f, g) < \epsilon^2 < 1/16$ then there is a function h with $p_h = gcd(p_f, p_g)$ and $D(h, g) < 3\epsilon$.*

Think of the g in the claim as being our g constructed above. We need to argue that there are no functions of smaller period within $1/d(n)$ of g. By our claim if such a function f existed then there would be a function h with period $p_h = gcd(p_f, p_g) = 1$ (since the period of g is a prime) and $D(h, g) < 3/\sqrt{d(n)}$. But g and the all-zeroes function, which is the only plausible candidate for h, have distance at least $3/\sqrt{d(n)}$, a contradiction. ■

References

[1] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, Oct. 1997.

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

[3] D. Boneh and R. J. Lipton. Quantum cryptanalysis of hidden linear functions (extended abstract). In D. Coppersmith, editor, *Advances in Cryptology—CRYPTO '95*, volume 963, pages 424–437. Springer-Verlag, 27–31 Aug. 1995.

[4] R. Cleve, E. Ekert, C. Macchiavello, and M. Mosca. Quantum algorithms revisited. *Proc. Roy. Soc. Lond. A*, 454:339–354, 1998.

[5] D. Coppersmith. An approximate fourier transform useful in quantum factoring. Technical Report RC19642, IBM, 1994.

[6] M. Ettinger and P. Høyer. On quantum algorithms for noncommutative hidden subgroups. In *Symposium on Theoretical Aspects in Computer Science*, University of Trier, 4–6 Mar. 1999.

[7] M. Grigni, L. Schulman, and U. Vazirani. Quantum mechanical algorithms for the non-abelian hidden subgroup problem. Manuscript, 1997.

[8] L. Hales and S. Hallgren. Quantum fourier sampling simplified. In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, pages 330–338, Atlanta, Georgia, 1–4 May 1999.

[9] S. Hallgren, A. Russell, and A. Ta-Shma. Normal subgroup reconstruction and quantum computation using group representations. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, pages 627–635, Portland, Oregon, 21–23 May 2000.

[10] P. Høyer. Simplified proof of the fourier sampling theorem. *Information Processing Letters*, 75:139–143, 2000.

[11] A. Y. Kitaev. Quantum measurements and the abelian stabilizer problem. Technical report, quant-ph/9511026, 1995.

[12] M. Mosca and A. Ekert. The hidden subgroup problem and eigenvalue estimation on a quantum computer. In *QCQS: NASA International Conference on Quantum Computing and Quantum Communications, QCQS*. LNCS, 1998.

[13] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, Oct. 1997.

[14] D. R. Simon. On the power of quantum computation. *SIAM J. Comput.*, 26(5):1474–1483, Oct. 1997.

A Proof of Corollary 1 from Lemma 1

Proof: We apply a standard averaging argument to conclude the bound. In particular, suppose that for some superposition $|\psi\rangle$ and vector $|\phi\rangle$,

$$\||\psi\rangle - |\phi\rangle\| < \epsilon$$

and let $|\psi\rangle = \sum_t |\gamma_t\rangle$, where $|\gamma_t'\rangle = |\gamma_t\rangle/\sqrt{p_t}$ is the result of measuring t in $|\psi\rangle$, and p_t is the probability of measuring t. Let $|\phi\rangle = \sum_t |\phi_t\rangle$, where $|\phi_t\rangle/||\phi_t\rangle|$ is the result of measuring t in $|\phi\rangle$ and let $|\phi_t'\rangle = |\phi_t\rangle/\sqrt{p_t}$. Finally let $B = \{t : \||\gamma_t'\rangle - |\phi_t'\rangle\|^2 > \epsilon\}$. Suppose that the probability of sampling a $t \in B$ is greater than ϵ. Then $\sum_{t\in B} p_t \||\gamma_t'\rangle - |\phi_t'\rangle\|^2 > \epsilon \sum_{t\in B} p_j > \epsilon^2$. But $\sum_t p_t \||\gamma_t'\rangle - |\phi_t'\rangle\|^2 = \sum_t \||\gamma_t\rangle - |\phi_t\rangle\|^2 \leq \epsilon^2$, a contradiction. Thus with probability at least $1-\epsilon$ we have

$$\||\gamma_t'\rangle - |\phi_t'\rangle\| < \sqrt{\epsilon}.$$

We wish to compare $|\gamma_t'\rangle$ to the unit vector in the direction of $|\phi_t'\rangle$, it is easy to see that the difference between these superpositions is at most $2\sqrt{\epsilon}$. ■

B Proof of Observation 1

Proof: Recall that

$$|\delta_i\rangle = F_q F_{ps}^{-1} |is\rangle = \frac{1}{\sqrt{qps}} \sum_{j=0}^{q-1} \sum_{k=0}^{ps-1} \omega^{k(\frac{j}{q} - \frac{i}{p})} |j\rangle.$$

We have

$$\begin{aligned} |(|\delta_i\rangle)_j| &= \left| \frac{1}{\sqrt{qps}} \sum_{k=0}^{ps-1} \omega^{k(\frac{j}{q}-\frac{i}{p})} \right| \leq \frac{1}{\sqrt{qps}} \left| \frac{1-\omega^{psj/q}}{1-\omega^{(\frac{j}{q}-\frac{i}{p})}} \right| \\ &\leq \sqrt{\frac{q}{ps}} \frac{2}{|j - i\frac{q}{p}|_q}, \end{aligned}$$

since $|1-\omega^{(\frac{j}{q}-\frac{i}{p})}| = |1-\omega_q^{|j-\frac{q}{p}i|_q}| \geq \frac{1}{q}|j - i\frac{q}{p}|_q$. ■

C Proof of Bound in Claim 1

Claim 5 *For any unit vector $x \in R^p_{\geq 0}$,*

$$\begin{aligned} &\sum_{j=0}^{q-1} \left(\sum_{\substack{i=0 \\ i \neq \lfloor jp/q \rceil}}^{p-1} \frac{x_i}{|j - \frac{q}{p} i|_q} \right)^2 \\ &\qquad \leq \frac{4}{p} \sum_{j=0}^{q-1} \left(\sum_{\substack{i=0 \\ i \neq \lfloor jp/q \rceil}}^{p-1} \frac{1}{|j - \frac{q}{p} i|_q} \right)^2 \end{aligned}$$

Proof: The bound of the claim is equivalent to bounding the operator norm of the p by q matrix M with entries

$$M_{ij} = \frac{1}{|j - \frac{q}{p} i|_q}$$

unless $i = \lfloor jp/q \rceil$ in which case $M_{ij} = 0$.

Note that our matrix has the property that each row is very close to being the shift by $\lfloor \frac{q}{p} \rceil$ of the previous row and all its entries are nonnegative reals. Given a p by p matrix with the property that each row is the shift by one of the previous row, i.e. $A_{ij} = A_{i+1,j+1}$, and all its entries are nonnegative reals it is easy to show that its operator norm is exactly the sum of a row. This is because its eigenvalues are all of the form $\sum_i \omega_p^{jk} A_{ij}$ and since the entries A_{ij} are nonnegative reals this is maximized when $j = 0$ and the corresponding eigenvector has all equal entries.

Now, while our matrix is not of this form, by reindexing and changing the denominators of the entries only slightly the expression $||M(x)||^2$ becomes

$$||\hat{M}x||^2 = \sum_{t=0}^{\lceil q/p \rceil -1} \sum_{k=0}^{p-1} \left(\sum_{\substack{i=0 \\ i \neq \lfloor dp/q \rceil}}^{p-1} \frac{x_i}{|\frac{q}{p}k + t - \frac{q}{p}i|_q} \right)^2$$

Notice that the matrix giving rise to each of the inner sums in this new expression is p by p and has the properties discussed previously. Thus each individual sum, and therefore the entire sum is maximized by choosing the entries of x_i to be equal.

Finally, we can relate this new sum to our original expression as follows: we changed the denominators of our original matrix entries by at most $1/2$, since these denominators were all larger than four this at most doubled/halved the squares of the entries themselves, thus we have:

$$\frac{1}{2} \leq \frac{||\hat{M}x||^2}{||Mx||^2} \leq 2.$$

Finally, we use this to bound the operator norm of M. Let x_0 maximize $||\hat{M}x||^2$. Then for any x we have $||Mx||^2 \leq 2||\hat{M}x||^2 \leq 2||\hat{M}x_0||^2 \leq 4||Mx_0||^2$. Thus our expression is bounded by four times it's value at the unit vector with entries uniformly equal to $\frac{1}{\sqrt{p}}$, as claimed. ■

D Proof of Lemma 3

Let $f \in C_{1/d(n)}$ with period p be given. Suppose that there is a prime a dividing p, such that

$$\Pr(a|j) > 1 - 1/8d^2(n).$$

where the probability is over j chosen according to $\mathcal{D}_{F_p(f)}$. We will show that f must in fact be close to some periodic function g with period p_g dividing p/a, contradicting our assumption that $f \in C_{1/d(n)}$.

Proof: We begin by specifying g. Recall that g is to be a function close to f but with period p/a. We begin by examining the values of f on flights of the form $(x, x+p/a, ..., x+(a-1)p/a)$. For convenience we shall denote this flight by $[x + kp/a]$ and the corresponding a-tuple of values of f by $[f(x + kp/a)]$. Since we must choose g to be constant on each flight and we need to agree with f on as many values as possible we will define $g(x_0)$ to be the majority value of $[f(x_0 + kp/a)]$ if one exists and 0 otherwise. More formally,

Definition 4 *For any k-tuple $s = (s_1, ...s_k)$ let*

$$maj(s) = \begin{cases} t & \textit{if at least half of the } s_i \textit{ are } t \\ 0 & \textit{otherwise.} \end{cases}$$

For $0 \leq x < p/a$ let

$$\hat{g}(x) = \text{maj}\left([f(x + kp/a]\right).$$

Then $g(x)$ is the unique function with period p/a extending $\hat{g}$.

We now prove that g and f are close:

Claim 6 $D(f, g) < 1/d(n)$.

Proof: For $0 \leq x < p$ let N be the size of the set

$$\{y | 0 \leq y < p \text{ and } f(y) = f(x)\}$$

and $|\alpha(x)\rangle$ be the superposition

$$\frac{1}{\sqrt{N}} \sum_{y, f(y) = f(x)} |y\rangle.$$

Also, for $0 \leq x < p$ let $|\hat{\alpha}(x)\rangle$ be the vector

$$\frac{1}{\sqrt{N}} \sum_{y, g(y) = f(x)} |y\rangle.$$

We shall suppress the dependence of these vectors on x, referring to them as $|\alpha\rangle$ and $|\hat{\alpha}\rangle$, when x is clear from the context. Our goal is to show that, for most x, these vectors are very close. Notice first that we can derive a formula for the distance between these two vectors as follows:

Let c_z be the fraction of the flight $[f(z + kp/a)]$ equal to $f(x)$. In other words c_z is the fraction of values in $[z + kp/a]$ with non-zero amplitude in $|\alpha\rangle$. Then

$$|||\alpha\rangle - |\hat{\alpha}\rangle||^2 = \frac{Na}{p} \sum_{0 \leq z < p/a} \min\{\frac{c_z}{N}, \frac{1 - c_z}{N}\}. \tag{4}$$

We wish to use the assumption of our lemma about the distribution $\mathcal{D}_{F_p(f)}$ to get a bound on the above expression. We first note that the distribution $\mathcal{D}_{F_p|\alpha(x)\rangle}$ where x is chosen uniformly from the set $\{0, ..., (p-1)\}$ and the distribution $\mathcal{D}_{F_p(f)}$ are identical. Thus by our lemma if x is chosen uniformly at random in $\{0, ..., (p-1)\}$ then with probability at least $1 - 1/2d(n)$, measuring $|\beta\rangle = F_p|\alpha(x)\rangle$ yields a j such that $a|j$ with probability at least $1 - 1/4d(n)$.

For such x let $|\hat{\beta}\rangle$ be the restriction of $|\beta\rangle$ to indices which are multiples of a. Then $|\hat{\beta}\rangle$ (which may no longer have unit length) and $|\beta\rangle$ satisfy

$$|||\beta\rangle - |\hat{\beta}\rangle||^2 < 1/4d(n).$$

Furthermore, $F_p^{-1}|\hat{\beta}\rangle$ is periodic with period p/a, that is $F_p^{-1}|\hat{\beta}\rangle = \sum_y \gamma_y|y\rangle$ satisfies

$$\gamma_y = \gamma_{y+p/a}$$

for all y.

Since F_p is unitary we have

$$\||\alpha\rangle - F_p^{-1}(|\hat{\beta}\rangle)\|^2 < 1/4d(n),$$

and thus $|\alpha\rangle$ is close to this p/a-periodic superposition. Recall that c_z is the fraction of the flight $[f(z+kp/a)]$ equal to $f(x)$, in other words c_z is the fraction of values in $[z+kp/a]$ with non-zero amplitude in $|\alpha\rangle$. Using the fact that $|\alpha\rangle$ is uniformly equal to $\frac{1}{\sqrt{N}}$ on its support and $F_p^{-1}(|\hat{\beta}\rangle)$ is p/a-periodic we have the following formula for $\||\alpha\rangle - F_p^{-1}(|\hat{\beta}\rangle)\|^2$:

$$\frac{Na}{p}\sum_{0\le z<p/a} c_z\left|\frac{1}{\sqrt{N}} - \gamma_z\right|^2 + (1-c_z)|\gamma_z|^2. \quad (5)$$

and thus

$$\frac{Na}{p}\sum_{0\le z<p/a} c_z\left|\frac{1}{\sqrt{N}} - \gamma_z\right|^2 + (1-c_z)|\gamma_z|^2 < 1/4d(n)$$

For a fixed set of c_z's, the formula in (5) is minimized by taking $\gamma_z = \frac{c_z}{\sqrt{N}}$. Thus we will only increase our upper bound on (5) if we instead use the condition

$$\begin{aligned}\frac{Na}{p}\sum_{0\le z<p/a} c_z\left|\frac{1}{\sqrt{N}} - \frac{c_z}{\sqrt{N}}\right|^2 + (1-c_z)\left|\frac{c_z}{\sqrt{N}}\right|^2 \\ = \frac{d}{r}\sum_{0\le z<p/a} c_z(1-c_z) \\ < 1/4d(n).\end{aligned}$$

But this clearly yields

$$\||\alpha\rangle - |\hat{\alpha}\rangle\|^2 = \frac{Na}{p}\sum_{0\le z<p/a}\min\{\frac{c_z}{N}, \frac{1-c_z}{N}\} < 1/2d(n).$$

Going back to the definitions of $|\alpha\rangle$ and $|\hat{\alpha}\rangle$, this implies that at most a $1/2d(n)$ fraction of $\{0 \le y < p | f(y) = f(x)\}$ satisfy $g(y) \ne f(y)$. Since this was true for at least a $1 - 1/2d(n)$ fraction of x's, the fraction of x's such that $f(x) \ne g(x)$ is at most $1/d(n)$, as desired. ■ ■

E Proof of Claim 4

Proof: Let $p_h = gcd(p_f, p_g)$. Fix k and l such that $lp_f - kp_g = p_h$. We will define a function h which is constant on flights of the form $[x + kp_h] = (x, x + p_h, x + 2p_h, \ldots)$ and within 3ϵ of g. Since $D(f, g) < \epsilon^2$, with probability at least $1-\epsilon$ when we choose a random flight $[x + kp_h]$ at least a $1-\epsilon$ fraction of points y in that flight will satisfy $f(y) = g(y)$. For such a "good" flight, choose y and z independently at random in the flight and let j satisfy $jp_h = y - z$. Then the point $w = y + jkp_g = z + jlp_f$ is uniformly distributed over the flight. Thus $f(z) = f(w) = g(w) = g(y)$ with probability at least $1-\epsilon$. Putting these two facts together we get that when y and z are chosen at random in a "good" flight, $g(y) = g(z)$ with probability at least $1-2\epsilon$. Using the fact that $\epsilon < 1/4$, this implies that at least a $1-2\epsilon$ fraction of points in the flight share the same g value. We let the value of h on all points in the flight be this overwhelming g value, and for "bad" flights we define h to be uniformly 0. Then it follows that $D(g, h) < 2\epsilon + \epsilon = 3\epsilon$, as claimed. ■