A QUANTUM FOURIER TRANSFORM ALGORITHM

CHRIS LOMONT

Abstract. Algorithms to compute the quantum Fourier transform over a cyclic group are fundamental to many quantum algorithms. This paper describes such an algorithm and gives a proof of its correctness, tightening some claimed performance bounds given earlier. Exact bounds are given for the number of qubits needed to achieve a desired tolerance, allowing simulation of the algorithm.

1. Introduction

Most quantum algorithms giving an exponential speedup over classical algorithms rely on efficiently computing Fourier transforms over some finite group [2, 4, 6, 10, 11, 12]. The Abelian group case depends on fast quantum algorithms for doing Fourier transforms over cyclic groups [7, 13, 14]. The thesis [7] and the paper [8] describe such a quantum algorithm, but the proofs are incorrect. This note attempts to correct those proofs, and in the process obtains stronger bounds for many of their results, and a few weaker ones. The end result is a proof of the correctness of their algorithm, with concrete bounds suitable for quantum simulation instead of the asymptotic bounds listed in their papers. Our final result is theorem 16.

2. Preliminaries

Efficient algorithms for the quantum Fourier transform over finite Abelian groups are constructed from the algorithms for the transform over cyclic groups, which in turn reduce to computing the transform efficiently over prime order groups [7, 13]. Efficient algorithms for computing the quantum Fourier transform over a cyclic group of order 2^m for a positive integer m are well known, and are used in Shor's factoring algorithm: see Coppersmith [5] and Shor [15, 16] for example. We will show an algorithm for computing the quantum Fourier transform over an odd order cyclic group. The algorithm (containing minor errors) is given in [7] and [8], but their proofs of the correctness of the algorithm and subsequent performance bounds are incorrect. This paper corrects those proofs and obtains new bounds. For applications of this algorithm, see [7] and [8]. A proof of similar ideas using a different method is in [9].

Date: March 2004.

2000 *Mathematics Subject Classification.* 03D15, 42A85, 68Q05, 68W40, 81P68.

Key words and phrases. algorithms, quantum computers, Hidden Subgroup Problem, quantum Fourier transform, cyclic quantum Fourier transform.

Research supported by AFRL grant F30602-03-C-0064.

2.1. **Notation and basic facts.** We fix three integers: an odd integer $N \geq 3$, $L \geq 2$ a power of 2, and $M \geq LN$ a power of 2. This gives $(M, N) = 1$, which we need later.

Some notation and facts to clarify the presentation:

- $\sqrt{-1}$ will be written explicitly, as i will always denote an index.
- For an integer $n > 1$, let $\omega_n = e^{2\pi\sqrt{-1}/n}$ denote a primitive n^{th} root of unity.
- Fact: $\left|1 - e^{\theta\sqrt{-1}}\right| \leq |\theta|$ as can be seen from arc length on the unit circle. If $-\pi \leq \theta \leq \pi$ we also[1] have $\left|\frac{\theta}{2}\right| \leq \left|1 - e^{\theta\sqrt{-1}}\right|$. Thus for real values α we have $|1 - \omega_M^\alpha| \leq \left|\frac{2\pi\alpha}{M}\right|$, etc.
- $\log n$ denotes log base 2, while $\ln n$ is the natural log. Since M and L are powers of two, $\lceil \log M \rceil = \lfloor \log M \rfloor = \lfloor \log M \rceil = \log M$, and similarly for L, but we often leave the symbols to emphasize expressions are integral.
- For a real number x, $\lceil x \rceil$ is the smallest integer greater than or equal to x, $\lfloor x \rfloor$ is the largest integer less than or equal to x, and $\lfloor x \rceil$ is the nearest integer, with ties rounding up[2]. We often use the three relations:
$$\begin{aligned} x - \frac{1}{2} \leq \lfloor x \rceil &\leq x + \frac{1}{2} \\ x - 1 < \lfloor x \rfloor &\leq x \\ x \leq \lceil x \rceil &< x + 1 \end{aligned}$$
- Indices: i and s will be indices from $0, 1, \ldots, N-1$. j will index from $0, 1, \ldots, L-1$. k will index from $0, 1, \ldots, M-1$. a and b will be arbitrary indices. t will index from a set C_s, defined in definition 2 below.
- Given $i \in \{0, 1, \ldots, N-1\}$, let $i' = \left\lfloor \frac{M}{N} i \right\rceil$ denote the nearest integer to $\frac{M}{N} i$ with ties broken as above. Similarly for s and s'. Note $0 \leq i' \leq M-1$.
- For a real number x and positive real number n, let $x \bmod n$ denote the real number y such that $0 \leq y < n$ and $y = x + mn$ for an integer m. Note that we do not think of $x \bmod n$ as an equivalence class, but as a real number in $[0, n)$.
- $|u\rangle$ and $|v\rangle$ are vectors in spaces defined later, and given a vector $|u\rangle$ denote its coefficients relative to the standard (orthonormal) basis $\{|0\rangle, |1\rangle, \ldots, |n-1\rangle\}$ by $u_0, u_1, \ldots, u_{n-1}$, etc.
- For a real number x, let
$$|x|_M = \begin{cases} x \bmod M & \text{if} \quad 0 \leq (x \bmod M) \leq \frac{M}{2} \\ -x \bmod M & \text{otherwise} \end{cases}$$
Thus $0 \leq |x|_M \leq \frac{M}{2}$. Properties of this function are easiest to see by noting it is a sawtooth function, with period M, and height $M/2$.
- For an integer s set $\delta_s = \left\lfloor \frac{M}{N} s \right\rceil - \frac{M}{N} s$. Then $|\delta_s| \leq \frac{1}{2}$.
- The (unitary) Fourier transform over a cyclic group of order N is denoted F_N. Thus if $|u\rangle = \sum_{i=0}^{N-1} u_i |i\rangle$, then $F_N|u\rangle = \frac{1}{\sqrt{N}} \sum_{i,s=0}^{N-1} u_i \omega_N^{is} |s\rangle$. We write $|\hat{u}\rangle = F_N|u\rangle$, with coefficients $\hat{u}_i$.
- $\sum_{i=0}^{N-1} |u_i|^2 = 1$ implies $\sum_i |u_i| \leq \sqrt{N}$.

[1] This range can be extended slightly.

[2] We could break ties arbitrarily with the same results.

The majority of the errors in [7] and [8] resulted from misunderstanding the consequences of their versions of the following two definitions. The first defines sets of integers which will play an important role:

Definition 1. *For $i = 0, 1, \ldots, N-1$, let (i) denote the set of integers in the open interval $\left(i' - \frac{M}{2N} + \frac{1}{2}, i' + \frac{M}{2N} - \frac{1}{2}\right)$ taken* mod M. *Recall $i' = \left\lfloor \frac{M}{N} i \right\rceil$.*

The second definition we make precise is a division and remainder operation:

Definition 2. *Given M, N as above. Set $\alpha = \left\lfloor \frac{M}{2N} + \frac{1}{2} \right\rfloor$, and $\beta = \left\lceil \frac{M}{2N} - \frac{3}{2} \right\rceil$. We define the map $\Delta : \{0, 1, \ldots, M-1\} \to \{0, 1, \ldots, N-1\} \times \{-\alpha, -\alpha+1, \ldots, \alpha\}$, as follows: for any $k \in \{0, 1, \ldots, M-1\}$, let $k \xrightarrow{\Delta} (s, t)$, via*

$$\begin{aligned} k' &= \left\lfloor k \frac{N}{M} \right\rceil \\ t &= k - \left\lfloor k' \frac{M}{N} \right\rceil \\ s &= k' \bmod N \end{aligned}$$

We extend this definition to a transform of basis elements $|k\rangle$ via

$$\Delta |k\rangle = |s\rangle |t + \alpha\rangle$$

and extend to all vectors by linearity.

Finally, from the image of Δ, define $C_s = \{\ t\ \mid\ (s,t) \in \text{Image } \Delta\}$ to be those values of t appearing for a fixed s. Thus $\sum_{k=0}^{M-1} |k\rangle \xrightarrow{\Delta} \sum_{s=0}^{N-1} \sum_{t \in C_s} |s\rangle |t + \alpha\rangle$.

We will show the integers $\{-\beta, \ldots, \beta\} \subseteq C_s \subseteq \{-\alpha, \ldots, \alpha\}$ for all s, which is why we defined β with the Δ definition. α and β remain fixed throughout the paper.

For the proofs to work, we need that the sets (i) are disjoint and have the same cardinality. Neither [7] nor [8] define these sets mod M, although perhaps it is implied. [7] makes a similar definition without the $\frac{1}{2}$ terms, but the resulting sets are *not* then disjoint. [8] makes a similar definition, but uses $\frac{M}{N} i$ instead of i' and drops the $\frac{1}{2}$ terms, which results in sets of varying cardinality. To see the differences, check the three definitions using $M = 32$ and $N = 5$. Both [7] and [8] implicitely assume their resulting sets are disjoint and of constant cardinality in numerous places, invalidating many proofs. Note also that the mod M condition gives $M - 1, 0 \in (0)$ when $M > 3N$. We now show that the sets defined here have the required properties:

Lemma 3. *For $i_1 \neq i_2 \in \{0, 1, \ldots, N-1\}$,*

$$|(i_1)| = |(i_2)| \tag{1}$$

$$(i_1) \bigcap (i_2) = \varnothing \tag{2}$$

Proof. Each set is defined using an interval of constant width, centered at an integer, so the sets will have the same cardinality. To show disjointness, for any integer a, take the rightmost bound $R_a = \left\lfloor \frac{M}{N} a \right\rceil + \frac{M}{2N} - \frac{1}{2}$ of an interval and compare it to

the leftmost bound $L_{a+1} = \left\lfloor \frac{M}{N}(a+1) \right\rceil - \frac{M}{2N} + \frac{1}{2}$ of the next interval:

$$L_{a+1} - R_a = \left\lfloor \frac{M}{N}(a+1) \right\rceil - \left\lfloor \frac{M}{N}a \right\rceil - \frac{M}{N} + 1 \tag{3}$$

$$\geq \left(\frac{M}{N}(a+1) - \frac{1}{2}\right) - \left(\frac{M}{N}a + \frac{1}{2}\right) - \frac{M}{N} + 1 \tag{4}$$

$$= 0 \tag{5}$$

giving that the open intervals are disjoint. Thus taking the integers in the intervals mod M remains disjoint (which requires $i_1, i_2 \leq N-1$). □

The second error which propagates throughout the proofs in [7] and [8] stems from misconceptions about the division operation Δ. Both papers treat the image of the Δ map as a cartesian product, that is, the range on t is the same for all values s ($M = 32$ and $N = 5$ illustrates how this fails to give a bijection with their definitions). However, the image is not a cartesian product; the values t assumes depend on s, otherwise we would have that M is a multiple of N. In other words, the cardinality of C_s depends on s, with bounds given in the following lemma, where we show that our definition works and list some properties:

Lemma 4. *Using the notation from definition* 2*,*

1) the map Δ is well defined, and a bijection with its image,

2) $\alpha = \beta + 1$,

3) the sets of integers satisfy $\{-\beta, \ldots, \beta\} \subseteq C_s \subseteq \{-\alpha, \ldots, \alpha\}$ for all $s \in \{0, 1, \ldots, N-1\}$.

Proof. Given a k in $\{0, 1, \ldots, M-1\}$, let $\Delta(k) = (s, t)$. Clearly $0 \leq s \leq N-1$. Set $\alpha = \left\lfloor \frac{M}{2N} + \frac{1}{2} \right\rfloor$. To check that $-\alpha \leq t \leq \alpha$, note

$$\frac{N}{M}k - \frac{1}{2} \leq k' \leq \frac{N}{M}k + \frac{1}{2} \tag{6}$$

giving

$$\frac{M}{2N} + \frac{1}{2} \geq t = k - \left\lfloor \frac{M}{N}k' \right\rceil \geq -\left(\frac{M}{2N} + \frac{1}{2}\right) \tag{7}$$

and t integral allows the rounding operation. Thus the definition makes sense.

Next we check that both forms of Δ in the definition are bijections. Suppose $k_1 \neq k_2$ are both in $\{0, 1, \ldots, M-1\}$, with images $\Delta(k_r) = (s_r, t_r), r = 1, 2$. Let $k'_r = \left\lfloor \frac{N}{M}k_r \right\rceil, r = 1, 2$. Note $0 \leq k'_r \leq N$.

Assume $(s_1, t_1) = (s_2, t_2)$. If $k'_1 = k'_2$, then

$$t_1 = k_1 - \left\lfloor \frac{M}{N}k'_1 \right\rceil = k_1 - \left\lfloor \frac{M}{N}k'_2 \right\rceil \tag{8}$$

$$\neq k_2 - \left\lfloor \frac{M}{N}k'_2 \right\rceil = t_2 \tag{9}$$

a contradiction. So we are left with the case $k'_1 \neq k'_2$. In order for $s_1 = s_2$ we have (without loss of generality) $k'_1 = 0, k'_2 = N$. But then $t_1 = k_1 \geq 0$ and $t_2 = k_2 - M \leq M - 1 - M = -1$, a contradiction. Thus Δ in the first sense is a bijection.

The second interpretation follows easily, since $-\alpha \leq t \leq \alpha$ gives $0 \leq t + \alpha \leq 2\alpha$. So the second register needs to have a basis with at least $2\alpha + 1$ elements, which

causes the number of qubits needed[3] to implement the algorithm to be $\lceil \log M \rceil + 2$ instead of $\lceil \log M \rceil$.

To see $\alpha = \beta + 1$, bound $\alpha - \beta$ using the methods above, and[4] one obtains $2 > \alpha - \beta > 0$.

All integers between $\left\lfloor \frac{M}{N}(s+1) \right\rceil$ and $\left\lfloor \frac{M}{N} s \right\rceil$ inclusive must be of the form $t_1 + \left\lfloor \frac{M}{N} s \right\rceil$ for $t_1 \in C_s$ or of the form $t_2 + \left\lfloor \frac{M}{N}(s+1) \right\rceil$ for $t_2 \in C_{s+1}$. This range contains $\left\lfloor \frac{M}{N}(s+1) \right\rceil - \left\lfloor \frac{M}{N} s \right\rceil + 1 \geq \frac{M}{N}$ integers, and at most $\alpha + 1$ of these are of the form $t_2 + \left\lfloor \frac{M}{N}(s+1) \right\rceil$ with $t_2 \in C_{s+1}$. This leaves at least $\left\lceil \frac{M}{N} \right\rceil - \alpha \geq \frac{M}{2N} - \frac{3}{2}$ that have to be of the form $t_1 + \left\lfloor \frac{M}{N} s \right\rceil$ with $t_1 \in C_s$, implying $\beta \in C_s$. Similar arguments give $\pm\beta \in C_s$, thus $\{-\beta, \ldots, \beta\} \subseteq C_s \subseteq \{-\alpha, \ldots, \alpha\}$ for all s. □

Δ is efficient to implement as a quantum operation, since it is efficient classically [3, Chapter 4]. Finally we note that Δ, being a bijection, can be extended to a permutation of basis vectors $|k\rangle$, thus can be considered an efficiently implementable unitary operation.

We define some vectors we will need. For $i \in \{0, 1, \ldots, N-1\}$ define

$$\begin{aligned}
|A^i\rangle &= F_M F_{LN}^{-1} |Li\rangle \\
&= \frac{1}{\sqrt{LMN}} \sum_{k=0}^{M-1} \sum_{a=0}^{LN-1} \omega_N^{-ai} \omega_M^{ak} |k\rangle \\
|B^i\rangle &= |A^i\rangle \text{ restricted to integers in the set } (i) \\
&= \sum_{b \in (i)} A_b^i |b\rangle \\
&= \frac{1}{\sqrt{LMN}} \sum_{b \in (i)} \sum_{a=0}^{LN-1} \omega_N^{-ai} \omega_M^{ab} |b\rangle \\
|T^i\rangle &= |A^i\rangle \text{ restricted to integers outside the set } (i) \\
&= \sum_{b \notin (i)} A_b^i |b\rangle \\
&= |A^i\rangle - |B^i\rangle \\
&= \frac{1}{\sqrt{LMN}} \sum_{b \notin (i)} \sum_{a=0}^{LN-1} \omega_N^{-ai} \omega_M^{ab} |b\rangle
\end{aligned}$$

Think A^i for actual values, B^i for bump functions, and T^i for tail functions. Note that the coefficients B_b^i and T_b^i are just A_b^i for b in the proper ranges.

We also define three equivalent shifted versions of $|B^0\rangle$. Note that to make these definitions equivalent we require the sets (i) to have the same cardinality. Let $|S^i\rangle = \sum_{b \in (0)} B_b^0 |b + i'\rangle = \sum_{b \in (0)} A_b^0 |b + i'\rangle = \sum_{b \in (i)} A_{b-i'}^0 |b\rangle$, where each $b \pm i'$ expression is taken mod M. The $|S^i\rangle$ have *disjoint support*, which follows from lemma 3, and will be important for proving theorem 13.

[3] This is proven in theorem 16.

[4] $(M, N) = 1$ is used to get the strict inequalities.

3. The Algorithm

The algorithm takes a unit vector (quantum state) $|u\rangle$ on $\lceil \log N \rceil$ qubits[5], does a Fourier transform F_L, L a power of two, on another register containing $|0\rangle$ with $\lceil \log M \rceil - \lceil \log N \rceil + 2$ qubits, to create[6] a superposition, and then reindexes the basis to create L (normalized) copies of the coefficients of $|u\rangle$, resulting in $|u_L\rangle$. Then another power of two Fourier transform F_M is applied. The division Δ results in a vector very close to the desired output $F_N|u\rangle$ in the first register, with garbage in the second register (with some slight entanglement). The point of this paper is to show how close the output is to this tensor product. We use $\lceil \log M \rceil + 2$ qubits, viewed in two ways: as a single register $|k\rangle$, or as a $\lceil \log N \rceil$ qubit first register, with the remaining qubits in the second register, written $|s\rangle|t\rangle$. We note that merely $\lceil \log M \rceil$ qubits may not be enough qubits to hold some of the intermediate results. The algorithm is:

3.1. **The odd cyclic quantum Fourier transform algorithm.**

$$|u\rangle|0\rangle \xrightarrow{F_L} \frac{1}{\sqrt{L}} \sum_{i=0}^{N-1} \sum_{j=0}^{L-1} u_i |i\rangle|j\rangle \tag{10}$$

$$\xrightarrow{\text{multiply}} \frac{1}{\sqrt{L}} \sum_{i,j} u_i |i + jN\rangle \tag{11}$$

$$= |u_L\rangle \tag{12}$$

$$\xrightarrow{F_M} \frac{1}{\sqrt{LM}} \sum_{i,j} \sum_{k=0}^{M-1} u_i \omega_M^{(i+jN)k} |k\rangle \tag{13}$$

$$\xrightarrow{\Delta} \frac{1}{\sqrt{LM}} \sum_{i,j} u_i \sum_{s=0}^{N-1} \sum_{t \in C_s} \omega_M^{(i+jN)\left(t + \left\lfloor \frac{M}{N} s \right\rceil\right)} |s\rangle|t + \alpha\rangle \tag{14}$$

$$= \frac{1}{\sqrt{N}} \sum_{i,s=0}^{N-1} u_i \omega_N^{is} |s\rangle \sqrt{\frac{N}{LM}} \sum_{t \in C_s} \sum_{j=0}^{L-1} \omega_M^{(i+jN)(t+\delta_s)} |t + \alpha\rangle \tag{15}$$

$$= |v\rangle \tag{16}$$

$|u_L\rangle$ is the vector that is L copies of the coefficients from $|u\rangle$, normalized. $|v\rangle$ is the algorithm output.

Notice that $F_N|u\rangle$ appears in the output in line 15, but the rest is unfortunately dependent on s and i. However the dependence is small: if C_s were the same for all s, if the δ_s, which are bounded in magnitude by $\frac{1}{2}$, were actually zero, and if the i dependence were dropped, then the output would leave $F_N|u\rangle$ in the first register. The paper shows this is approximately true, and quantifies the error.

4. Initial bounds

We need many bounds to reach the final theorem, which we now begin proving. [7] makes the mistake of missing the -1 in the following lemma[7]; [8], using a

[5]Recall logs are base 2.

[6]Note it may be more efficient to apply the Hadamard operator H to each qubit in $|0\rangle$.

[7]Using the definitions in [7], a $-\frac{1}{2}$ instead of -1 is sufficient. Even then, however, $M = 128$, $N = 37$, $i = 12$, and $k = 40$ shows the error. Compare our lemma 5 to the proof of claim 4, section 9.2.3, in [7].

different definition for the (i), is correct in dropping the -1. To avoid these subtle errors we thus prove

Lemma 5. *For integers $N > 2$, $M \geq 2N$, and any $i \in \{0, 1, \ldots, N-1\}$, $k \in \{0, 1, \ldots, M-1\}$, with $k \notin (i)$, we have*

$$\left| k - \frac{M}{N} i \right|_M \quad \geq \quad \frac{M}{2N} - 1 \tag{17}$$

Proof. The sets (i) are disjoint, so we do two cases. If $i = 0$, then $k \notin (0)$ implies

$$\frac{M}{2N} - \frac{1}{2} \leq k \leq M - \frac{M}{2N} + \frac{1}{2} \tag{18}$$

from which it follows that

$$\left| k - \frac{M}{N} 0 \right|_M \geq \frac{M}{2N} - \frac{1}{2} > \frac{M}{2N} - 1 \tag{19}$$

If $i \neq 0$, then either k is less than the integers in (i) or greater than the integers in (i), giving two subcases. Subcase 1:

$$0 \leq k \leq \left\lfloor \frac{M}{N} i \right\rceil - \frac{M}{2N} + \frac{1}{2} \leq \frac{M}{N} i - \frac{M}{2N} + 1 \tag{20}$$

implying

$$\frac{M}{2N} - 1 \leq \frac{M}{N} i - k \leq \frac{M}{N} i \leq M - \frac{M}{N} \tag{21}$$

which gives the bound. Subcase 2 is then

$$\frac{M}{N} i + \frac{M}{2N} - 1 \leq \left\lfloor \frac{M}{N} i \right\rceil + \frac{M}{2N} - \frac{1}{2} \leq k \leq M - 1 \tag{22}$$

which implies

$$\frac{M}{2N} - 1 \leq k - \frac{M}{N} i \leq M - 1 - \frac{M}{N} i \tag{23}$$

giving the bound and the proof. □

We now bound many of the $|A^i\rangle$ coefficients. Our bound has a factor of π not in [7] and [8], making it somewhat tighter, and we avoid special cases[8] where the statement would not be true.

Lemma 6. *For $k \in \{0, 1, \ldots, M-1\}$ and $i \in \{0, 1, \ldots, N-1\}$, with $\frac{k}{M} - \frac{i}{N}$ not an integer, then*

$$\left| A_k^i \right| \leq \sqrt{\frac{M}{LN}} \, \frac{2}{\pi \left| k - \frac{M}{N} i \right|_M} \tag{24}$$

Proof. We rewrite from the definition

$$A_k^i \quad = \quad \frac{1}{\sqrt{LMN}} \sum_{a=0}^{LN-1} \omega_M^{a\left(k - \frac{M}{N} i\right)} \tag{25}$$

(26)

[8] [7] and [8] missed these cases by not placing any restriction such as our hypothesis that $\frac{k}{M} - \frac{i}{N}$ is non-integral. Compare our lemma 6 with Observation 2, section 9.2.3 in [7] and with Observation 1, section 3.1, in [8].

which is a geometric series. By hypothesis, $\omega_M^{\left(k-\frac{M}{N}i\right)} \neq 1$, so we can sum as[9]

$$\left|A_k^i\right| = \frac{1}{\sqrt{LMN}} \left| \frac{1-\omega_M^{LN\left(k-\frac{M}{N}i\right)}}{1-\omega_M^{\left(k-\frac{M}{N}i\right)}} \right| \tag{27}$$

The numerator is bounded above by 2, and the denominator satisfies

$$\left|1-\omega_M^{\left(k-\frac{M}{N}i\right)}\right| = \left|1-\omega_M^{\left|k-\frac{M}{N}i\right|_M}\right| \tag{28}$$

$$\geq \frac{\pi\left|k-\frac{M}{N}i\right|_M}{M} \tag{29}$$

These together give

$$\left|A_k^i\right| \leq \sqrt{\frac{M}{LN}}\,\frac{2}{\pi\left|k-\frac{M}{N}i\right|_M} \tag{30}$$

□

Note our initial requirement that $(M,N)=1$ is strong enough to satisfy the non-integral hypothesis in lemma 6, except for the case $i=k=0$, which we will avoid.

Next we bound a sum of these terms. We fix $\gamma = \frac{1}{2} - \frac{N}{M}$ for the rest of this paper.

Lemma 7. *Given integers $N>2$ and $M>2N$, with N odd. Let $\gamma = \frac{1}{2}-\frac{N}{M}$. For a fixed integer $k \in \{0,1,\ldots,M-1\}$,*

$$\sum_{\substack{i=0\\ k\notin(i)}}^{N-1} \frac{1}{\left|k-\frac{M}{N}i\right|_M} \leq \frac{2N}{M}\left(\frac{1}{\gamma} + \ln\left|\frac{N-1}{2\gamma}+1\right|\right) \tag{31}$$

Proof. The minimum value of the denominator is at least $\frac{M}{2N}-1$ by lemma 5, and the rest are spaced out by $\frac{M}{N}$, but can occur twice[10] since the denominator is a sawtooth function going over one period, giving that

$$\sum_{\substack{i=0\\ k\notin(i)}}^{N-1} \frac{1}{\left|k-\frac{M}{N}i\right|_M} \leq 2\sum_{a=0}^{\frac{N-1}{2}} \frac{1}{\frac{M}{2N}-1+\frac{M}{N}a} \tag{32}$$

$$= \frac{2N}{M}\left(\frac{1}{\gamma} + \sum_{a=1}^{\frac{N-1}{2}} \frac{1}{\gamma+a}\right) \tag{33}$$

$$\leq \frac{2N}{M}\left(\frac{1}{\gamma} + \int_0^{(N-1)/2} \frac{1}{x+\gamma}dx\right) \tag{34}$$

$$= \frac{2N}{M}\left(\frac{1}{\gamma} + \ln\left|\frac{N-1}{2\gamma}+1\right|\right) \tag{35}$$

□

[9]Without this requirement, the sum would be LN, much different than the claimed sum. The hypotheses avoid the resulting divide by zero.

[10]Both [7] and [8] appear to overlook this fact.

The generality of the above lemma would be useful where physically adding more qubits than necessary would be costly, since the lemma lets the bound tighten as $\frac{N}{M}$ decreases. However the following corollary is what we will use in the final theorem.

Corollary 8. *Given integers $N \geq 13$ and $M \geq 16N$, with N odd. For a fixed value $k \in \{0, 1, \ldots, M-1\}$,*

$$\sum_{\substack{i=0 \\ k \notin (i)}}^{N-1} \frac{1}{\left|k - \frac{M}{N} i\right|_M} \leq \frac{4N \ln N}{M} \tag{36}$$

Proof. Using lemma 7, $M \geq 16N$ gives $\frac{1}{\gamma} \leq \frac{16}{7}$ and

$$\begin{aligned} \frac{1}{\gamma} + \ln\left|\frac{N-1}{2\gamma} + 1\right| &\leq \frac{16}{7} + \ln\left|\frac{8(N-1)}{7} + 1\right| && (37) \\ &= \ln\left(e^{\frac{16}{7}}\left(\frac{8(N-1)}{7} + 1\right)\right) && (38) \\ &\leq \ln\left(\frac{8}{7}\, e^{\frac{16}{7}} N\right) && (39) \\ &\leq 2 \ln N && (40) \end{aligned}$$

where the last step required $N \geq \left(\frac{8}{7}\, e^{\frac{16}{7}}\right) > 11.2$. The corollary follows. □

[7] claimed an incorrect bound[11] of $\frac{2N \ln N}{M}$ in section 9.2.3, and [8] obtained the correct $\frac{4N \ln N}{M}$ in section 3.1, but both made the errors listed above.

Next we prove a bound on a sum of the above terms, weighted with a real unit vector. This will lead to a bound on the tails $\left\|\sum_i \hat{u}_i |T^i\rangle\right\|$. Our bound has an extra term compared to the claimed bounds in [7] and [8], but corrects an error in their proofs.

Lemma 9. *Given integers $N \geq 13$ and $M \geq 16N$, with N odd. For any unit vector $x \in \mathbb{R}^N$*

$$\sum_{k=0}^{M-1} \left| \sum_{\substack{i=0 \\ k \notin (i)}}^{N-1} \frac{x_i}{\left|k - \frac{M}{N} i\right|_M} \right|^2 \leq \frac{22N \ln^2 N}{M} + \frac{32N^3}{M^2} \tag{41}$$

Proof. We split the expression into three parts, the first of which we can bound using methods from [7] and [8], and the other two terms we bound separately.

Using the Δ operator from definition 2, along with the values α and β defined there, and using lemma 4, we can rewrite each k with $k = t + \left\lfloor \frac{M}{N} k' \right\rceil = t + \frac{M}{N} k' + \delta_s$. Since s differs from k' by a multiple of N, and the $|x|_M$ function has period M, in $\left|\frac{M}{N}(k' - i) + t + \delta_s\right|_M$ we can replace k' with s. Rewrite the left hand side of inequality 41 as

$$\sum_{k=0}^{M-1} \left| \sum_{\substack{i=0 \\ k \notin (i)}}^{N-1} \frac{x_i}{\left|k - \frac{M}{N} i\right|_M} \right|^2 = \sum_{s=0}^{N-1} \sum_{t \in C_s} \left| \sum_{\substack{i=0 \\ s \neq i}}^{N-1} \frac{x_i}{\left|\frac{M}{N}(s - i) + t + \delta_s\right|_M} \right|^2 \tag{42}$$

[11] This fails, for example, at $M = 256$, $N = 13$, $k = 26$, using either their definitions or our definitions.

Letting $\Delta k = (s,t)$, note that $k \notin (i)$ if and only if $s \neq i$, which can be shown from the definitions and the rounding rules used earlier. To simplify notation, write $q_{i,s}^t = \frac{M}{N}(s-i) + t + \delta_s$. We have not changed the values of the denominators, so $|q_{i,s}^t|_M \geq \frac{M}{2N} - 1$ by lemma 5 for all $i, (s,t)$ in this proof.

We want to swap the s and t sums, but we need to remove the t dependence on s. Again using lemma 4, we can split the expression into the three terms[12]:

$$\sum_{t=-\beta}^{\beta} \sum_{s=0}^{N-1} \left| \sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{x_i}{\left|q_{i,s}^t\right|_M} \right|^2 \tag{43}$$

$$+ \sum_{s \text{ with } \alpha \in C_s} \left| \sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{x_i}{\left|q_{i,s}^\alpha\right|_M} \right|^2 \tag{44}$$

$$+ \sum_{s \text{ with } -\alpha \in C_s} \left| \sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{x_i}{\left|q_{i,s}^{-\alpha}\right|_M} \right|^2 \tag{45}$$

Next we bound the first term 43. For a unit vector x and fixed t we rewrite the s,i sum as the norm of a square matrix P_t acting on x, so that the sum over s and i becomes

$$\|P_t x\|^2 = \sum_{s=0}^{N-1} \left| \sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{x_i}{\left|q_{i,s}^t\right|_M} \right|^2 \tag{46}$$

We also define similarly to each P_t a matrix Q_t which is the same except for minor modifications to the denominator:

$$\|Q_t x\|^2 = \sum_{s=0}^{N-1} \left| \sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{x_i}{\left|q_{i,s}^t - \delta_s\right|_M} \right|^2 \tag{47}$$

Note this matrix is circulant[13], since each entry in the matrix only depends on $s-i$. Also each entry is nonnegative[14]. Thus the expression is maximized by the vector $y = \frac{1}{\sqrt{N}}(1,1,\ldots,1)$ as shown in each of [7], [8], and [9]. Now we relate these matrix expressions. Recall $|q_{i,s}^t|_M \geq \frac{M}{2N} - 1$ and $|\delta_s| \leq \frac{1}{2}$. Set $\lambda = \frac{N}{M-2N}$. Then we find lower and upper bounds

$$1 - \lambda = 1 - \frac{1}{2(\frac{M}{2N} - 1)} \leq \frac{\left|q_{i,s}^t\right|_M - \frac{1}{2}}{\left|q_{i,s}^t\right|_M} \leq \frac{\left|q_{i,s}^t - \delta_s\right|_M}{\left|q_{i,s}^t\right|_M} \tag{48}$$

and

$$\frac{\left|q_{i,s}^t - \delta_s\right|_M}{\left|q_{i,s}^t\right|_M} \leq \frac{\left|q_{i,s}^t\right|_M + \frac{1}{2}}{\left|q_{i,s}^t\right|_M} \leq 1 + \frac{1}{2(\frac{M}{2N} + 1)} = 1 + \lambda \tag{49}$$

[12]The second two terms are missed in [7] and [8].

[13]That is, each row after the first is the cyclic shift by one from the previous row.

[14]$|q_{i,s}^t - \delta_s|_M \geq |q_{i,s}^t|_M - \frac{1}{2} \geq \frac{M}{2N} - \frac{3}{2} > 0$ since $M > 3N$

Rewriting

$$\|P_t x\|^2 = \sum_{s=0}^{N-1} \left| \sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{x_i}{\left|q_{i,s}^t - \delta_s\right|_M} \frac{\left|q_{i,s}^t - \delta_s\right|_M}{\left|q_{i,s}^t\right|_M} \right|^2 \tag{50}$$

and using the bounds gives

$$(1-\lambda)^2 \|Q_t x\|^2 \leq \|P_t x\|^2 \leq (1+\lambda)^2 \|Q_t x\|^2 \tag{51}$$

Then since y maximizes $\|Q_t x\|^2$,

$$\|P_t x\|^2 \leq (1+\lambda)^2 \|Q_t x\|^2 \leq (1+\lambda)^2 \|Q_t y\|^2 \leq \left(\frac{1+\lambda}{1-\lambda}\right)^2 \|P_t y\|^2 \tag{52}$$

giving that we can bound the leftmost term by $\left(\frac{1+\lambda}{1-\lambda}\right)^2$ times the norm at y. $\left(\frac{1+\lambda}{1-\lambda}\right)^2$ takes on values between 1 and $\frac{225}{169} \approx 1.33$ for $M \geq 16N$, better than the constant 4 in [7] and [8].

Combined with corollary 8 this allows us to bound term 43:

$$\sum_{t=-\beta}^{\beta} \sum_{s=0}^{N-1} \left| \sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{x_i}{\left|q_{i,s}^t\right|_M} \right|^2 \leq \sum_t \frac{225}{169} \sum_{s=0}^{N-1} \left| \sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{\frac{1}{\sqrt{N}}}{\left|q_{i,s}^t\right|_M} \right|^2 \tag{53}$$

$$\leq (2\beta+1) \frac{225}{169} \frac{N}{N} \left(\frac{4N \ln N}{M}\right)^2 \tag{54}$$

$$\leq \frac{M}{N} \frac{225}{169} \left(\frac{4N \ln N}{M}\right)^2 \tag{55}$$

$$\leq \frac{22N \ln^2 N}{M} \tag{56}$$

Now we bound the other two terms, 44 and 45. We need the following fact, which can be shown with calculus: the expression $\left|\sum_{i=0}^{N-1} a_i x_i\right|$ subject to the condition $\sum_{i=0}^{N-1} x_i^2 = 1$, has maximum value $\sqrt{\sum_{i=0}^{N-1} a_i^2}$. Then term 44 can be bounded using a similar technique as in the proof of lemma 8. Again we take $\gamma = \frac{1}{2} - \frac{N}{M}$.

$$\sum_{s \text{ with } \alpha \in C_s} \left| \sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{x_i}{\left|q_{i,s}^\alpha\right|_M} \right|^2 \leq \sum_s \left| \sqrt{\sum_{\substack{i=0 \\ s\neq i}}^{N-1} \frac{1}{\left|q_{i,s}^\alpha\right|_M^2}} \right|^2 \tag{57}$$

$$\leq N \frac{2N^2}{M^2} \left(\frac{1}{\gamma^2} + \sum_{a=1}^{\frac{N-1}{2}} \frac{1}{\left(\frac{1}{2} - \frac{N}{M} + a\right)^2} \right) \tag{58}$$

$$\leq \frac{2N^3}{M^2} \left(\frac{1}{\gamma^2} + \frac{1}{\gamma} - \frac{1}{\frac{N-1}{2} + \gamma} \right) \tag{59}$$

$$\leq \frac{16N^3}{M^2} \tag{60}$$

Term 45 is bound with the same method and result, and adding these three bounds gives the desired inequality 41. □

Similar to the proof of the previous lemma, Both [7] and [8] claim the following bound

$$\sum_{k=0}^{M-1}\left|\sum_{\substack{i=0\\k\notin(i)}}^{N-1}\frac{x_i}{\left|k-\frac{M}{N}i\right|_M}\right|^2 \le \frac{4}{N}\sum_{k=0}^{M-1}\left|\sum_{\substack{i=0\\k\notin(i)}}^{N-1}\frac{1}{\left|k-\frac{M}{N}i\right|_M}\right|^2 \tag{61}$$

leading to (in their papers)

$$\sum_{k=0}^{M-1}\left|\sum_{\substack{i=0\\k\notin(i)}}^{N-1}\frac{x_i}{\left|k-\frac{M}{N}i\right|_M}\right|^2 \le \frac{64N\ln^2 N}{M} \tag{62}$$

So our bound tightens their 64 to a 22, but has a new term accounting for the extra pieces in the proof. However, both [7] (section 9.2.4) and [8] (appendix C) had the following flaws their proofs. Both proofs rearranged the left expression to be bounded by a matrix norm, and then "rearranged" the matrix to be square. This fails due to the subtle nature of the Δ operation they implicitly used. They claimed the resulting matrix differed only slightly from their previous one, which is false, since many terms may have to be changed from 0 to a large value. They relied on the resulting matrix being circulant and being close to the to their initial expression, which it is not due to these extra terms. Our proof above is based on their methods, but avoids the errors they made by pulling out the incorrect terms and bounding them separately, resulting in the extra term in our bound.

We now use these lemmata to bound the tails $\left\|\sum_i \hat{u}_i|T^i\rangle\right\|$. The bound $\frac{8\ln N}{\sqrt{L}}$ was claimed in [7], section 9.2.3, and [8], section 3.1, but our new terms from lemma 9 give us a more complicated bound:

Lemma 10. *Given three integers: an odd integer $N \ge 13$, $L \ge 2$ a power of two, and $M \ge 16N$ a power of two, then*

$$\left\|\sum_{i=0}^{N-1}\hat{u}_i|T^i\rangle\right\| \le \frac{2}{\pi}\sqrt{\frac{22\ln^2 N}{L}+\frac{32N^2}{LM}} \tag{63}$$

Proof.

$$\left\|\sum_{i=0}^{N-1}\hat{u}_i|T^i\rangle\right\|^2 = \sum_{k=0}^{M-1}\left|\sum_{\substack{i=0\\k\notin(i)}}^{N-1}\hat{u}_i T_k^i\right|^2 \tag{64}$$

$$\le \sum_k \frac{4M}{\pi^2 LN}\left(\sum_{\substack{i=0\\k\notin(i)}}^{N-1}\frac{|\hat{u}_i|}{\left|k-\frac{M}{N}i\right|_M}\right)^2 \tag{65}$$

$$\le \frac{4M}{\pi^2 LN}\left(\frac{22N\ln^2 N}{M}+\frac{32N^3}{M^2}\right) \tag{66}$$

Taking square roots gives the result. Note that the requirements of lemma 6 are satisfied when obtaining line 65, since we avoid the $k = i = 0$ case, and $(M, N) = 1$. □

Next we show that the shifted $|S^i\rangle$ are close to the $|B^i\rangle$, which will allow us to show the algorithm output is close to a tensor product. This mirrors [7] claim 5, section 9.2.1, and [8] claim 2, section 3. In both cases their constant was 4, where we obtain the better bound $\frac{\pi}{\sqrt{3}}$.

Lemma 11.

$$\left\| |S^i\rangle - |B^i\rangle \right\| \leq \frac{\pi LN}{M\sqrt{3}} \tag{67}$$

Proof. Recall $|S^i\rangle = \sum_{b\in(i)} A^0_{b-i' \bmod M}|b\rangle$ and $|B^i\rangle = \sum_{b\in(i)} A^i_b|b\rangle$. It is important these are supported on the same indices! Also recall that $|A^i\rangle = F_M F_{LN}^{-1}|Li\rangle$ and that F_M is unitary. Then (dropping mod M throughout for brevity)

$$\left\| |S^i\rangle - |B^i\rangle \right\|^2 = \left\| \sum_{b\in(i)} A^0_{b-i'}|b\rangle - \sum_{b\in(i)} A^i_b|b\rangle \right\|^2 \tag{68}$$

$$\leq \left\| \sum_{k=0}^{M-1} A^0_{k-i'}|k\rangle - \sum_{k=0}^{M-1} A^i_k|k\rangle \right\|^2 \tag{69}$$

$$= \left\| F_M^{-1} \left(\sum_{k=0}^{M-1} A^0_k|k+i'\rangle - |A^i\rangle \right) \right\|^2 \tag{70}$$

$$= \sum_{a=0}^{LN-1} \left| \frac{1}{\sqrt{LN}} \omega_M^{-ai'} - \frac{1}{\sqrt{LN}} \omega_N^{-ai} \right|^2 \tag{71}$$

$$= \frac{1}{LN} \sum_{a=0}^{LN-1} \left| \omega_M^{-ai'} \left(1 - \omega_M^{a\delta_i} \right) \right|^2 \tag{72}$$

and this can be bounded by

$$\frac{1}{LN} \sum_{a=0}^{LN-1} \left| \frac{2\pi a \delta_i}{M} \right|^2 \leq \frac{\pi^2}{LNM^2} \sum_{a=0}^{LN-1} a^2 \leq \frac{\pi^2}{LNM^2} \frac{(LN)^3}{3} \tag{73}$$

Taking square roots gives the bound. □

In the above proof, to obtain line 69 we needed that $|S^i\rangle$ and $|B^i\rangle$ have the same support, but $|S^i\rangle$ is a shifted version of $|B^0\rangle$, so we implicitly needed all the sets (i) to have the same cardinality. This is not satisfied in [8] (although it is needed) but is met in [7].

For the rest of he paper we need a set which is (0) without mod M applied: let Λ be those integers in the closed interval $\left[-\left\lfloor \frac{M}{2N} - \frac{1}{2} \right\rfloor, \left\lfloor \frac{M}{2N} - \frac{1}{2} \right\rfloor \right]$. Then

Lemma 12.

$$\Delta|S^i\rangle = |i\rangle \sum_{t\in\Lambda} A^0_t|t+\alpha\rangle \tag{74}$$

Proof. By definition, $|S^i\rangle = \sum_{b\in(0)} A_b^0|b + \lfloor \frac{M}{N} i \rceil \bmod M\rangle$. $\Delta\left(b + \lfloor \frac{M}{N} i \rceil\right) = (i, b)$ (the proof uses $(M, N) = 1$), and Δ a bijection implies $\Delta|b + \lfloor \frac{M}{N} i \rceil \bmod M\rangle = |i\rangle|b + \alpha\rangle$. The rest follows[15]. □

5. Main results

Now we are ready to use the above lemmata to prove the main theorem.

Theorem 13. *Given three integers: an odd integer $N \geq 13$, $L \geq 16$ a power of two, and $M \geq LN$ a power of two. Then the output $|v\rangle$ of the algorithm in section* 3.1 *satisfies*

$$\left\| |v\rangle - F_N|u\rangle \otimes \sum_{t\in\Lambda} A_t^0|t + \alpha\rangle \right\| \leq \frac{2}{\pi}\sqrt{\frac{22\ln^2 N}{L} + \frac{32N^2}{LM}} + \frac{\pi LN}{M\sqrt{3}} \tag{75}$$

Proof. Note

$$|\hat{u}\rangle := F_N|u\rangle = \sum_{i=0}^{N-1} \hat{u}_i|i\rangle \qquad\qquad F_M|u_L\rangle = \sum_{i=0}^{N-1} \hat{u}_i|A^i\rangle \tag{76}$$

Using lemma 12 and that Δ is unitary allows us to rewrite the left hand side as

$$\left\| |v\rangle - \sum_{\substack{s=0 \\ t\in C_s}}^{N-1} \hat{u}_s A_t^0|s\rangle|t + \alpha\rangle \right\| = \left\| \Delta F_M|u_L\rangle - \sum_{s=0}^{N-1} \hat{u_s}\Delta|S^s\rangle \right\| \tag{77}$$

$$= \left\| \sum_{s=0}^{N-1} \hat{u_s}|A^s\rangle - \sum_{s=0}^{N-1} \hat{u_s}|S^s\rangle \right\| \tag{78}$$

$$= \left\| \sum_{s=0}^{N-1} \hat{u_s}(|B^s\rangle + |T^s\rangle) - \sum_{s=0}^{N-1} \hat{u_s}|S^s\rangle \right\| \tag{79}$$

By the triangle inequality this is bounded by

$$\left\| \sum_{s=0}^{N-1} \hat{u_s}|T^s\rangle) \right\| + \left\| \sum_{s=0}^{N-1} \hat{u_s}|B^s\rangle) - \sum_{s=0}^{N-1} \hat{u_s}|S^s\rangle \right\| \tag{80}$$

which in turn by lemmata 10 and 11 is bounded by

$$\frac{2}{\pi}\sqrt{\frac{22\ln^2 N}{L} + \frac{32N^2}{LM}} + \frac{\pi LN}{M\sqrt{3}}\sqrt{\sum_s |\hat{u_s}|^2} \tag{81}$$

The last expression has $\||\hat{u}\rangle\| = 1$, which gives the result. Note that to obtain line 81 we needed the supports of the $|B^s\rangle$ disjoint, and that the $|S^i\rangle$ and $|B^i\rangle$ have the same support[16]. □

This shows that the output of the algorithm in section 3.1 is close to a tensor product of the desired output $F_N|u\rangle$ and another vector (which is not in general a unit vector). Since a quantum state is a unit vector, we compare the output to a unit vector in the direction of our approximation via:

[15] It is tempting to use C_0 instead of Λ, but this is not correct in all cases.

[16] This is not satisfied in [7], and the overlapping portions make that proof invalid.

Lemma 14. *Let $\vec{a}$ be a unit vector in a finite dimensional vector space, and $\vec{b}$ any vector in that space. For any $0 \leq \epsilon \leq 1$, if $\left\|\vec{a} - \vec{b}\right\| \leq \epsilon$ then the unit vector $\vec{b'}$ in the direction of $\vec{b}$ satisfies $\left\|\vec{a} - \vec{b'}\right\| \leq \epsilon\sqrt{2}$.*

Proof. Simple geometry shows the distance is bounded by $\sqrt{2(1-\sqrt{1-\epsilon^2})}$, and this expression divided by ϵ has maximum value $\sqrt{2}$ on $(0,1]$. The $\epsilon = 0$ case is direct. □

So we only need a $\sqrt{2}$ factor to compare the algorithm output with a unit vector which is $F_N|u\rangle$ tensor another unit vector. We let $|\psi\rangle$ denote the unit length vector in the direction of $\sum_{t\in\Lambda} A_t^0|t+\alpha\rangle$ for the rest of this paper[17].

For completeness, we repeat arguments from [8, 9] to obtain the operation complexity and probability distribution, and we show concrete choices for M and L achieving a desired error bound.

To show that measuring the first register gives measurement statistics which are very close to the desired distribution, we need some notation. Given two probability distributions $\mathcal{D}$ and $\mathcal{D}'$ over $\{0,1,\ldots,M-1\}$, let $|\mathcal{D}-\mathcal{D}'| = \sum_{k=0}^{M-1}|\mathcal{D}(k)-\mathcal{D}'(k)|$ denote the total variation distance. Then a result[18] of Bernstein and Vazirani [1] states that if the distance between any two states is small, then so are the induced[19] probability distributions:

Lemma 15 ([1], Lemma 3.6)**.** *Let $|\alpha\rangle$ and $|\beta\rangle$ be two normalized states, inducing probability distributions $\mathcal{D}_\alpha$ and $\mathcal{D}_\beta$. Then for any $\epsilon > 0$*

$$\||\alpha\rangle - |\beta\rangle\| \leq \epsilon \Rightarrow |\mathcal{D}_\alpha - \mathcal{D}_\beta| \leq 2\epsilon + \epsilon^2 \tag{82}$$

independent of what basis is used for measurement.

Combining this with theorem 13 and lemmata 14 and 15 gives the final result

Theorem 16.

1) *Given an odd integer $N \geq 13$, and any $\sqrt{2} \geq \epsilon > 0$. Choose $L \geq 16$ and $M \geq LN$ both integral powers of 2 satisfying*

$$\frac{2}{\pi}\sqrt{\frac{22\ln^2 N}{L} + \frac{32N^2}{LM}} + \frac{\pi LN}{M\sqrt{3}} \leq \frac{\epsilon}{\sqrt{2}} \tag{83}$$

Then there is a unit vector $|\psi\rangle$ such that the output $|v\rangle$ of the algorithm in section 3.1 satisfies

$$||v\rangle - F_N|u\rangle \otimes |\psi\rangle\| \leq \epsilon \tag{84}$$

2) *We can always find such an L and M by choosing*

$$L = c_1\frac{\sqrt{N}}{\epsilon^2} \tag{85}$$

$$M = c_2\frac{N^{\frac{3}{2}}}{\epsilon^3} \tag{86}$$

[17] The subscripts in A_t^0 are taken mod M.

[18] Their statement is a bound of 4ϵ, but their proof gives the stronger result listed above. We choose the stronger form to help minimize the number of qubits needed for simulations.

[19] The induced distribution from a state $|\phi\rangle$ is $\mathcal{D}(k) = |\langle k|\phi\rangle|^2$.

for some constants c_1, c_2 satisfying

$$65 \leq c_1 \leq 2 \times 65 \tag{87}$$

$$735 \leq c_2 \leq 2 \times 735 \tag{88}$$

3) *The algorithm requires $\lceil \log M \rceil + 2$ qubits. By claim 2 a sufficient number of qubits is then $\left\lceil 12.53 + 3 \log \frac{\sqrt{N}}{\epsilon} \right\rceil$. The algorithm has operation complexity $O(\log M(\log\log M + \log 1/\epsilon))$. Again using claim 2 yields an operation complexity of*

$$O\left(\log \frac{\sqrt{N}}{\epsilon}\left(\log\log \frac{\sqrt{N}}{\epsilon} + \log 1/\epsilon\right)\right) \tag{89}$$

4) *The induced probability distributions $\mathcal{D}_v$ from the output and $\mathcal{D}$ from $F_N|u\rangle \otimes |\psi\rangle$ satisfy*

$$|\mathcal{D}_v - \mathcal{D}| \leq 2\epsilon + \epsilon^2 \tag{90}$$

Proof. Claim 1 follows directly from theorem 13 and lemma 14. Claim 1 and lemma 15 give claim 4.

To get claim 2, note that for the bound to be met, we must have $\frac{\ln^2 N}{L} < \epsilon^2$, $\frac{N^2}{LM} < \epsilon^2$, and $\frac{LN}{M} < \epsilon$. Trying to keep M small as N and ϵ vary leads to the forms for L and M chosen. If we substitute lines 85 and 86 into 83 and simplify, we get

$$\frac{4}{\pi}\sqrt{\frac{11\ln^2 N}{c_1\sqrt{N}} + \frac{16\epsilon^3}{c_1 c_2}} + \frac{\pi\sqrt{2}}{\sqrt{3}}\frac{c_1}{c_2} \leq 1 \tag{91}$$

The left hand side is largest when $\epsilon = \sqrt{2}$ and $N = 55$, so it is enough to find constants c_1 and c_2 such that

$$\frac{4}{\pi}\sqrt{\frac{11\ln^2 55}{c_1\sqrt{55}} + \frac{32\sqrt{2}}{c_1 c_2}} + \frac{\pi\sqrt{2}}{\sqrt{3}}\frac{c_1}{c_2} \leq 1 \tag{92}$$

Ultimately we want L and M to be powers of two, so we find a range for each of c_1 and c_2 such that the upper bound is at least twice the lower bound, and such that all pairs of values (c_1, c_2) in these ranges satisfy inequality 92. To check that the claimed ranges work, note that for a fixed c_1, the expression increases as c_2 decreases, so it is enough to check the bound for $c_2 = 735$. After replacing c_2 in the expression with 735, the resulting expression has first and second derivatives with respect to c_1 over the claimed range, and the second derivative is positive, giving that the maximum value is assumed at an endpoint. So we only need to check inequality 92 at two points: $(c_1, c_2) = (65, 735)$ and $(2 \times 65, 735)$, both of which work. Thus the bound is met for all (c_1, c_2) in the ranges claimed. With these choices for M and L, note that $L \geq 16$ and $M \geq LN \Leftrightarrow c_2 \geq \epsilon c_1$, which is met over the claimed range, so all the hypothesis for claim 1 are satisfied.

Finally, to prove claim 3, algorithm 3.1 and the proof of lemma 4 give that we need $\lceil \log N \rceil$ qubits in the first register and $\max\{\lceil \log L \rceil, \lceil \log(2\alpha + 1) \rceil\}$ qubits in the second register. $L \leq \frac{M}{N} < 2\alpha + 1$ gives that it is enough to have $\lceil \log(2\alpha + 1) \rceil$ qubits in the second register. Then $2\alpha + 1 \leq \frac{M}{2N} + 2$ gives

$$\lceil \log(2\alpha + 1) \rceil \leq \lceil 1 + \log M - \log N \rceil = 2 + \lceil \log M \rceil - \lceil \log N \rceil \tag{93}$$

Thus $\lceil \log M \rceil + 2$ is enough qubits[20] for the algorithm. By claim 2, we can take $M \leq 2 \times 735 \frac{N^{3/2}}{\epsilon^3}$ giving $\lceil \log M \rceil + 2 \leq \left\lceil 12.53 + 3 \log \frac{\sqrt{N}}{\epsilon} \right\rceil$.

As noted in [7] and [8], the most time consuming step in algorithm 3.1 is the F_M Fourier computation. Coppersmith [5] shows how to ϵ approximate the quantum Fourier transform[21] for order $M = 2^m$ with operation complexity of $O(\log M(\log\log M + \log 1/\epsilon))$. Using this to approximate our approximation within error ϵ gives the time complexities in claim 3, finishing the proof. □

6. Conclusion

These bounds allow simulation for many choices of N and ϵ. However the choices for M and L given in theorem 16 can usually be improved, and were merely given to show such values can be found. For example, the following table shows, for different N and ϵ combinations, a triple (g, m, l) of integers, with the choice from line 86 being $M = 2^g$; yet in each case $M = 2^m$ and $L = 2^l$ is the pair with minimal m satisfying the hypotheses for theorem 16. Thus choosing M and L carefully may allow lower qubit counts, such as the $N = 13$, $\epsilon = 0.10$ case.

ϵ	N=13	N=25	N=51	N=101	N=251	N=501
.001	45,45,28	47,47,28	48,48,29	50,50,29	52,52,30	53,53,30
.01	36,35,21	37,37,22	38,38,23	40,40,23	42,42,23	43,43,24
.05	29,28,17	30,30,17	31,31,18	33,33,18	35,35,19	36,36,19
.10	26,25,15	27,27,15	28,28,16	30,30,16	32,32,17	33,33,17
.20	23,22,13	24,24,13	25,25,14	27,27,14	29,29,15	30,30,15
.30	21,20,12	22,22,12	24,24,12	25,25,13	27,27,13	29,28,14
.40	20,19,11	21,21,11	22,22,12	24,24,12	26,26,13	27,27,13

Table 1. Values

We also simulated this algorithm for the combinations above requiring 22 or fewer qubits. The first test computed the algorithm error on random input vectors (states)[22]. The middle set of columns in table 2, where $(M, L) = (2^m, 2^l)$, shows the maximal error observed in the column labelled "observed ϵ" over 100 random vectors. Note that the observed error is *much* smaller than the required bound; for example, with $N = 25$, $\epsilon = 0.3$ the max observed error is actually 0.0182. This led to the second set of experiments, with results in the third set of columns in table 2, where we tried all legal M, L combinations until we found the one with smallest M value that met the desired error bound, when tested over 5000 random vectors. This seemed to show that the qubit cost could almost be cut in half in practice. As a final test, for $N = 501$ and $\epsilon = 0.2$, the theorem requires 30 qubits, but empirical testing showed 15 suffices for our 5000 test vectors, which had a maximal error of 0.18. These results show that it is likely that significant tightening of the bounds presented here is possible, resulting in qubit savings.

[20] An example requiring $\lceil \log M \rceil + 2$ qubits is $M = 1024$, $N = 65$, so the bound is tight.

[21] Many authors give a simple quantum circuit doing the quantum Fourier transform over a power 2^m with time complexity $O(m^2)$; see for example [3, Chapter 5 and endnotes]. However, this requires m elementary operations, which seems a little like cheating. Requiring a finite fixed number of elementary operations would give a time complexity of $O(m^3)$.

[22] The left hand side of line 84 is the error computed.

N	ϵ	m	l	observed ϵ	best m	best l	ϵ_2
13	0.4	19	11	0.0362329	9	4	0.353615
13	0.3	20	12	0.0409662	10	4	0.212023
13	0.2	22	13	0.0187127	11	4	0.158535
25	0.4	21	11	0.0193478	10	4	0.309438
25	0.3	22	12	0.0181997	11	4	0.193214
51	0.4	22	12	0.0332493	11	4	0.294778

TABLE 2. Simulation results

REFERENCES

1. Ethan Bernstein and Umesh Vazirani, *Quantum complexity theory*, SIAM Journal on Computing **26** (1997), no. 5, 1411–1473.
2. Thomas Beth, Markus Püschel, and Martin Rötteler, *Fast quantum Fourier transforms for a class of non-abelian groups*, Proc. of Applied Algebra Algebraic Algorithms, and Error-Correction Codes (AAECC-13, Springer-Verlag, 1999, volume 1719 in Lecture Notes in Computer Science, pp. 148–159.
3. I. L. Chuang and M. A. Nielsen, *Quantum computation and quantum information*, Cambridge University Press, Cambridge, 2000.
4. R. Cleve, E. Ekert, C. Macchiavello, and M. Mosca, *Quantum algorithms revisited*, Proc. Roy. Soc. Lond. A **454** (1998), 339–354.
5. D. Coppersmith, *An approximate Fourier transform useful in quantum computing*, IBM Technical Report RC 19642 (1994), `quant-ph/0201067`.
6. M. Grigni, L. J. Schulman, M. Vazirani, and U. V. Vazirani, *Quantum mechanical algorithms for the nonabelian hidden subgroup problem*, Proc. 33rd ACM Symp. on Theory of Computing, 2001, pp. 68–74.
7. Lisa Hales, *The quantum Fourier transform and extensions of the abelian subgroup problem*, Ph.D. thesis, University of California at Berkeley, Berkeley, CA, 2002.
8. Lisa Hales and Sean Hallgren, *An improved quantum Fourier transform algorithm and applications*, Proc. 41st Ann. Symp. on Foundations of Computer Science, 2000, Redonda Beach, California, 12-14 November, pp. 515–525.
9. Peter Høyer, *Simplified proof of the Fourier sampling theorem*, Information Processing Letters **75** (2000), no. 4, 139–143.
10. Alexi Yu. Kitaev, *Quantum measurements and the Abelian stabilizer problem*, `quant-ph/9511026`, 1995, Nov 20.
11. S. J. Lomonaco and L.H. Kauffman, *Quantum hidden subgroup problems: A mathematical perspective*, 2002, `quant-ph/0201095`.
12. Christopher Moore, Daniel Rockmore, Alexander Russell, and Leonard Schulman, *The hidden subgroup problem in affine groups: Basis selection in Fourier sampling*, `quant-ph/0211124`, 2002.
13. Michele Mosca, *Quantum computer algorithms*, Ph.D. thesis, Wolfson College, University of Oxford, Oxford, United Kingdom, 1999.
14. Michele Mosca and Christof Zalka, *Exact quantum Fourier transforms and discrete logarithm algorithms*, `quant-ph/0301093`, 2003.
15. P. W. Shor, *Algorithms for quantum computation: discrete logarithms and factoring*, Proceedings, 35th Annual Symposium on Fundamentals of Comp. Science (FOCS), 1994, pp. 124–134.
16. ———, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Computing **26** (1997), no. 5, 1484–1509.

E-mail address: `clomont@cybernet.com, clomont@math.purdue.edu`

URL: `www.math.purdue.edu/~clomont`

URL: `www.cybernet.com`

Current address: Cybernet Systems Corporation, 727 Airport Blvd., Ann Arbor, MI, 48108-1639 USA.