Average-Case Verification of the Quantum Fourier Transform Enables Worst-Case Phase Estimation

Noah Linden[1] and Ronald de Wolf[2]

[1]School of Mathematics, University of Bristol. n.linden@bristol.ac.uk
[2]QuSoft, CWI and University of Amsterdam, the Netherlands. rdewolf@cwi.nl

The quantum Fourier transform (QFT) is a key primitive for quantum computing that is typically used as a subroutine within a larger computation, for instance for phase estimation. As such, we may have little control over the state that is input to the QFT. Thus, in implementing a good QFT, we may imagine that it needs to perform well on arbitrary input states. *Verifying* this worst-case correct behaviour of a QFT-implementation would be exponentially hard (in the number of qubits) in general, raising the concern that this verification would be impossible in practice on any useful-sized system. In this paper we show that, in fact, we only need to have good *average-case* performance of the QFT to achieve good *worst-case* performance for key tasks—phase estimation, period finding and amplitude estimation. Further we give a very efficient procedure to verify this required average-case behaviour of the QFT.

1 Introduction

1.1 Verification of quantum circuits

Massive efforts are currently being expended around the world on building large quantum computers, in academia and industry. Because of the fragility of quantum hardware and the quantum states it produces, it is crucial to be able to *test* that the hardware works as advertised. Such a test may involve some quantum hardware itself, but should be more "lightweight" than the procedure that is being tested, in order to avoid circularity.

There is an important issue with testing that is sometimes overlooked in high-level discussions of the topic: if the circuit of interest is a subroutine in a larger computation, then we may have little control of the state that is input to it; thus we would like to verify that it works on the *worse-case* input state. However, we can typically only test its behavior for an *average-case* input state, because there are far too many possible input states to test them all. In general, testing worst-case correctness of a given n-qubit circuit would take resources that scale exponentially in n. This means that efficient verification of worst-case correctness typically requires additional assumptions, ranging from restrictions on the class of circuits one is verifying (for instance Clifford circuits [FL11, dSLCP11, LW21]) to cryptographic assumptions (as in Mahadev's approach [Mah18], which also assumes the computation starts with a fixed initial state). For further discussion and pointers to related work on verification of quantum hardware, we refer to our recent paper [LW21] and to the general survey [EHW+20].[1]

[1]The issue of average-case vs worst-case behavior has also recently received attention in the area of Hamiltonian simulation [ZZS+21, CB21].

In this paper we focus on the situation where we want to apply a quantum Fourier transform (QFT), or its inverse, within the context of a larger quantum computation that we already trust to a sufficient extent. We will give a lightweight procedure for verifying certain average-case behaviour of the QFT circuit, given the ability only to apply it as a black-box. We will then show that this enables us to use the QFT to achieve good *worst-case* performance for key tasks—phase estimation, period finding and amplitude estimation.

1.2 The quantum Fourier transform

The quantum Fourier transform is one of the most important (possibly *the* most important) component of quantum algorithms. It is key in Shor's factoring algorithm [Sho97] and in the standard approach to amplitude estimation [BHMT02], which generalises Grover's search algorithm [Gro96] and which is an important subroutine in many other quantum algorithms.[2] Let $N = 2^n$ and $\omega_N = e^{2\pi i/N}$. The n-qubit QFT is the unitary F_N that maps n-bit basis state $|k\rangle$ as

$$|k\rangle \mapsto |\hat{k}\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \omega_N^{jk} |j\rangle,$$

where the "jk" in the exponent denotes multiplication of two n-bit integers. Interestingly, the complicated-looking Fourier basis state $|\hat{k}\rangle$ is actually a product state of n individual qubits:

$$|\hat{k}\rangle = \bigotimes_{\ell=1}^{n} \frac{1}{\sqrt{2}} \left(|0\rangle + e^{2\pi i k/2^\ell} |1\rangle\right) \tag{1}$$

Leveraging this product structure, there is a well-known circuit of $O(n^2)$ gates that implements F_N exactly. It uses n Hadamard gates, $O(n^2)$ controlled versions of

$$R_s = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^s} \end{pmatrix}$$

for different integers s, and a few SWAP gates at the end [NC00, Section 5.1]. One can also obtain an *approximate* circuit from this with only $O(n \log n)$ gates, by dropping the R_s gates where s is bigger than $c \log n$ for some constant c [Cop94] (R_s gates with large s are very close to the identity, so dropping them incurs very little error). The resulting circuit differs from F_N by only an inverse-polynomially small error in operator norm.

One may also consider the inverse QFT F_N^{-1}, where the phases are ω_N^{-jk} instead of ω_N^{jk}. This has equally efficient exact and approximate quantum circuits, since we can just reverse a circuit for F_N and invert its gates to get a circuit for F_N^{-1}.

[2] It is often possible to avoid doing the full QFT in these applications. For instance, one can do phase estimation in a bit-by-bit manner [Kit95] or by using the block-encoding framework of [GSLW19] as done in [MRTC21, Ral21]; and one can do amplitude estimation by judiciously chosen numbers of Grover iterations [AR20]. However, replacing the QFT by "something else" raises the question of the verification of those "something else" components. In this paper our goal is not avoid the QFT but to show that it can be efficiently tested for average-case correctness, and then used in worst-case applications. We feel our results should actually favor the use of the QFT as a component in quantum algorithms: in this paper we come not to bury the QFT, but to praise it.

1.3 Testing a purported QFT or QFT^{-1}

In this paper we are interested in the situation where we have a channel[3] C that we can run only as a black-box on states $|\psi\rangle$ of our choice; C is supposed to implement F_N, or F_N^{-1} depending on the application. We would like to test to what extent C is correct. It is not practical to test whether $C(|\psi\rangle)$ is approximately the right state for *all* possible $|\psi\rangle$, since C could differ from F_N^{-1} in only one "direction"; in fact, testing whether $C(|\hat{k}\rangle)$ is close to $F_N^{-1}|\hat{k}\rangle = |k\rangle$ for *all* $k \in \{0,1\}^n$ requires $\Omega(\sqrt{2^n})$ runs of C.[4] However, it turns out that we can efficiently test whether $C(|\hat{k}\rangle)$ and $F_N^{-1}|\hat{k}\rangle$ are close on *average* over all $k \in \{0,1\}^n$. Fortunately, good performance on most Fourier basis states $|\hat{k}\rangle$ suffices for the applications we care about in the rest of the paper, which run the inverse QFT on individual Fourier basis states or mixtures thereof, or on superpositions dominated by a small number of Fourier basis states.

A Fourier basis state $|\hat{k}\rangle$ is a product state by Eq. (1), so it is relatively easy ("lightweight") to prepare, at least approximately. Each qubit in the product state $|\hat{k}\rangle$ is of the form $\frac{1}{\sqrt{2}}(|0\rangle + e^{i\phi}|1\rangle)$ for some phase ϕ that depends on k and on the location of the qubit. It suffices to prepare each of those qubits with $O(\log n)$ bits of precision[5] in the phase ϕ in order to prepare $|\hat{k}\rangle$ up to inverse-polynomially small error. One might be worried that this preparation effectively requires us to do an approximate QFT, which would defeat our purpose of testing a purported QFT black-box C; however, it is much easier to prepare known product states such as Fourier basis states than it is to implement a QFT on an arbitrary unknown state. In particular, in regimes starting from a few dozen qubits (which is the current state of the art of quantum hardware), preparing a Fourier basis state to sufficient precision seems doable while tomography on a channel C would already be prohibitively expensive.

Assuming we can prepare Fourier basis states $|\hat{k}\rangle$ sufficiently precisely, in Section 2 we give a simple test that approximates the average error (defined as infidelity, i.e., 1 minus fidelity) which C makes on Fourier basis states, up to additive approximation error ε. Our procedure uses $O(1/\varepsilon^2)$ runs, each of which prepares an n-qubit product state (namely a random Fourier basis state), applies C to it, and measures the resulting n-qubit state in the computational basis.

1.4 Using an average-case-correct QFT^{-1} for worst-case phase estimation

Suppose we have a channel C that has passed our test, so we can be confident that C is close in an average-case sense to the inverse QFT. What can we use C for?

---

[3] A *channel* is a completely positive trace-preserving map on density matrices, in our case taking n-qubit mixed states to n-qubit mixed states. An n-qubit unitary is a special case of this. If channel C is run on pure state $|\psi\rangle$, then we will use the notation $C(|\psi\rangle)$ to abbreviate the resulting state $C(|\psi\rangle\langle\psi|)$.

[4] This follows from the well-known fact that we need $\Omega(\sqrt{2^n})$ queries to a string $x \in \{0,1\}^{2^n}$ to decide whether $x = 0^{2^n}$ [BBBV97], as follows. If we can make queries $O_x : |k\rangle \mapsto (-1)^{x_k}|k\rangle$ and we define C as the n-qubit unitary $O_x F_N^{-1}$, then $C = F_N^{-1}$ if $x = 0^{2^n}$, and otherwise C is far from F_N^{-1} on at least one input state $|\hat{k}\rangle$. Accordingly, if we can distinguish those two cases with T runs of C, then T queries to x can decide whether $x = 0^{2^n}$, which implies T must be $\Omega(\sqrt{2^n})$. This lower bound is optimal, since we can Grover search [Gro96] over all $k \in \{0,1\}^n$ to look for one where $C|\hat{k}\rangle$ differs significantly from $F_N^{-1}|\hat{k}\rangle = |k\rangle$.

[5] Each bit of precision can be rotated in with one single-qubit gate R_s. To see that $2\log n$ bits of precision in each phase suffice, note that this gives a fidelity $\geq 1 - O(1/n^2)$ per qubit, which (because we are dealing with product states) multiplies out to a fidelity $(1 - O(1/n^2))^n \geq 1 - O(1/n)$ for the n-qubit product state as a whole.

Phase estimation, originally due to Kitaev [Kit95], is the following application of the inverse QFT. Suppose we can apply an m-qubit unitary U in a controlled manner, and are given an eigenstate $|\phi\rangle$ of U with eigenvalue $e^{2\pi i\theta}$ for some unknown $\theta \in [0,1)$. The goal is to estimate θ. Standard phase estimation (reviewed in Section 3.1 below) obtains an n-bit approximation to θ by using $O(2^n)$ controlled applications of U in order to (exactly or approximately) prepare the state $F_N|\theta_1 \dots \theta_n\rangle$, where $\theta_1 \dots \theta_n$ are the n most significant bits of the binary expansion of $\theta = 0.\theta_1 \dots \theta_n \dots$ Applying an inverse QFT then gives us θ itself, or at least a good approximation of θ.

Now suppose our channel C for the inverse QFT is only average-case correct. In that case phase estimation will fail if $F_N|\theta_1 \dots \theta_n\rangle$ happens to be one of the Fourier basis states on which C fails significantly. However, in Section 3 we show how an average-case-correct C actually suffices to implement phase estimation *in the worst case* (assuming the other components of phase estimation work sufficiently well). We do this by a simple trick whereby we randomise the phase we are estimating. Then we can use our average-case-correct C to recover a good approximation to that randomised phase with good probability, and afterwards undo the randomisation to obtain a good approximation to θ itself. A moderately small upper bound on the average-case error in the inverse QFT is good enough to make this work: in the case where the eigenphase θ can be written exactly with n bits of precision we can tolerate an average infidelity up to almost $1/2$ (see end of Section 3.2), while for the more general case where θ needs more than n bits of precision we can tolerate an average infidelity up to 0.041 (see end of Section 3.3).

The advantage of this approach is that we can efficiently test whether a given C has small *average-case* error, while we cannot efficiently test whether C has small *worst-case* error. If the average error is a sufficiently small constant, then also a small constant approximation error ε suffices for this test, hence only a constant number of runs of C suffices to achieve high confidence in the approximate correctness of the inverse QFT.

1.5 Applications

As mentioned, two of the most important quantum algorithms known to date are Shor's algorithm for integer factoring [Sho97], whose quantum core is period-finding, and amplitude estimation [BHMT02]. Both rely on an inverse quantum Fourier transform, and both may fail miserably if that inverse Fourier transform happens to fail on the particular state that the algorithm applies it to. In Section 4 we show that both algorithms can still be made to work with high success probability if we only have an average-case-correct inverse QFT at our disposal.

1.6 Pros and cons of our approach

Before going into technical details let us clarify and emphasize several aspects of our approach.

First, our approach is only relevant for implementations of QFT (or QFT^{-1}) that work reasonably well for most (inverse) Fourier basis states. This typically won't be true in a setting where a few of the gates in the circuit can be completely wrong. However, our approach is relevant for a QFT circuit where many (maybe even all) of the gates are *slightly* wrong in various ways, for instance if the per-gate error times the total number of gates ($O(n \log n)$ for the approximate QFT circuit) is at most a constant. If the number n of qubits is a few dozen, then a per-gate error on the order of $1/n \log n$ is not very far from current technology.

Second, apart from small average-case error (infidelity averaged over Fourier basis states) we don't assume much about C. This average-case error could arise in many ways, which could even be picked by our adversary: C could make a small error on most Fourier basis states, or be completely wrong on, say, a small constant fraction of those states, or anything in between. Our approach can deal with all these cases. We are *not* assuming any relatively benign and smooth error model such as depolarizing noise. This also illustrates the difference between worst-case and average-case error: C being completely wrong on some (maybe even a constant fraction of) Fourier basis states means it has terrible worst-case error, and yet our average-case to worst-case reduction shows that it can still be turned into something quite serviceable for worst-case applications. Even in the benign case of random rather than adversarial noise, the action of C is still likely to be worse on some Fourier basis states than on others, and our worst-case-to-average-case approach has the benefit of smoothing this out, reducing the worst-case error probability.

Third, if a channel C passed our test then we can conclude that it works well on an average Fourier basis state and hence (by our average-case-to-worst-case reduction) can be made to work well on *all* Fourier basis states. We can certainly *not* conclude that C will work well on arbitrary superpositions of Fourier basis states, since every state is a superposition of Fourier basis states, and testing that C works well on every possible state requires an exponential number of runs of C (see footnote 4). So a C that passed our test cannot just be used in every application of the inverse QFT. Fortunately, in the case of phase estimation, the inverse QFT will be applied to an individual Fourier basis state, or a mixture of Fourier basis states, or to a superposition state dominated by $O(1)$ Fourier basis states. As we prove in Sections 3.2 and 3.3, in such cases a C that passed our test still works well enough.

Fourth, for the applications related to phase estimation, let us emphasize that we assume the state is measured in the computational basis right after the inverse QFT and all earlier non-QFT parts of the algorithm leading up to that state are essentially perfect. This is a strong assumption. For factoring via period-finding, the preparation of the periodic state involves a circuit for modular exponentiation which uses polynomially more gates (and hence is more prone to error) than the inverse QFT. Similarly, to do amplitude estimation with n bits of precision one can use an n-qubit approximate inverse QFT circuit with $O(n \log n)$ gates, which pales in significance (and error-proneness) compared to the roughly 2^n controlled Grover iterations that are done prior to the inverse QFT. Nevertheless, we feel it is a sensible modular approach to try to isolate parts of important algorithms that can be tested by themselves in a lightweight manner. The fact that the average-case error of the QFT can be efficiently tested in such a lightweight manner, and that small average-case error suffices for many of its applications, should make the QFT a more attractive component to use in larger algorithms.

2 Testing average-case correctness of F_N^{-1} on Fourier basis states

In this section we show how one can efficiently test, in a lightweight manner, that a given quantum channel C is close to the n-qubit inverse Fourier transform F_N^{-1} on an average Fourier basis state. We can test average-case closeness to F_N completely analogously, but for concreteness we focus on F_N^{-1} in this section. We give a procedure to estimate the *infidelity* between C and F_N^{-1}, averaged over the Fourier basis states, which is our measure of average-case error here. The fidelity between mixed states ρ and σ is defined as

$$F(\rho, \sigma) = \mathrm{Tr}\left(\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\right)^2.$$

Fidelity is symmetric. It is 0 if ρ and σ are orthogonal, it is 1 if they are equal, and otherwise it lies in $(0,1)$. If $\sigma = |\psi\rangle\langle\psi|$ is pure, then $F(\rho, |\psi\rangle) = \langle\psi|\rho|\psi\rangle$. The *in*fidelity between ρ and σ is defined as 1 minus fidelity. We now show that average infidelity of a purported inverse QFT is relatively easy to estimate.

Theorem 1. *Let C be a channel from n qubits to n qubits, $|\hat{k}\rangle$ be a uniformly random Fourier basis state, and define the average infidelity between C and F_N^{-1} by*

$$\eta = \mathbb{E}_k[1 - F(C(|\hat{k}\rangle)\,, F_N^{-1}|\hat{k}\rangle)\,].$$

There exists a procedure that estimates η up to additive error ε, with success probability $1-\delta$, using $O(\log(1/\delta)/\varepsilon^2)$ runs, each of which prepares an n-qubit product state, runs C on it, and measures the resulting n-qubit state in the computational basis.

Proof. Choose $k \in \{0,1\}^n$ uniformly at random and prepare n-qubit product state $|\hat{k}\rangle = F_N|k\rangle$. Run C on $|\hat{k}\rangle$ and measure the resulting state in the computational basis. Output 1 if the n-bit measurement outcome is k, and output 0 otherwise. Because $F_N^{-1}|\hat{k}\rangle = |k\rangle$ is a pure state, we have

$$\eta = \mathbb{E}_k[\,1 - F(C(|\hat{k}\rangle)\,, |k\rangle)\,] = \mathbb{E}_k[\,1 - \langle k|\,C(|\hat{k}\rangle)\,|k\rangle\,] = 1 - \Pr[\,\text{output } 1\,] = \Pr[\,\text{output } 0\,].$$

The Chernoff bound implies that if we repeat this procedure $r = O(\log(1/\delta)/\varepsilon^2)$ times, then the frequency of 0s among the r output bits equals η up to $\pm\varepsilon$, except with probability $\leq \delta$.[6] □

Thus we have a procedure to test whether the average infidelity of our purported black-box for the inverse QFT is small. The procedure has modest overhead: $O(\log(1/\delta)/\varepsilon^2)$ runs, each involving a preparation of a Fourier basis state $|\hat{k}\rangle$, one run of C, and one n-qubit measurement in the computational basis. Referring back to the discussion in the penultimate paragraph of Section 1.3, preparing $|\hat{k}\rangle$ can be done with polynomially small error using $O(n)$ single-qubit gates, each with $O(\log n)$ bits of precision in their phase.

The same procedure could be used to test a black-box C for F_N rather than for F_N^{-1}; the only difference is that we would prepare product state $F_N^{-1}|k\rangle$ at the start of each run rather than $F_N|k\rangle$.

While we do not estimate average infidelity averaged over arbitrary states, or over an arbitrary orthonormal basis, averaging over the particular basis of Fourier basis states turns out to be sufficient for our purposes as we'll see next.

3 Using average-case-correct F_N^{-1} for worst-case phase estimation

We saw that it is hard to test a purported QFT or inverse-QFT black-box for *worst-case* correctness, but relatively easy to test it for *average-case* correctness on the set of QFT basis states. Here we will show that average-case correctness actually suffices for phase estimation *even in the worst case*.

[6] The Chernoff bound actually implies something slightly stronger for the relevant case where η is close to 0 (i.e., where C works reasonably well), namely that $r = O(\eta\log(1/\delta)/\varepsilon^2)$ repetitions suffice. In particular, if ε is set to a small constant times η, then $r = O(\log(1/\delta)/\varepsilon)$ repetitions suffice rather than $O(\log(1/\delta)/\varepsilon^2)$.

3.1 Basic phase estimation

As mentioned in the introduction, in the setup for phase estimation we can apply an m-qubit unitary U in a controlled manner, and are given an eigenstate $|\phi\rangle$ of U with eigenvalue $e^{2\pi i\theta}$ for some unknown $\theta \in [0,1)$. The goal is to estimate θ with roughly n bits of precision. We work on $n+m$ qubits that start in state $|0^n\rangle \otimes |\phi\rangle$. We first apply n Hadamard gates to obtain the uniform superposition $\frac{1}{\sqrt{2^n}} \sum_{j\in\{0,1\}^n} |j\rangle$ in the first register. Applying the $(n+m)$-qubit unitary

$$V = \sum_{j\in\{0,1\}^n} |j\rangle\langle j| \otimes U^j \tag{2}$$

gives a phase $e^{2\pi ij\theta}$ to state $|j\rangle|\phi\rangle$, where $j\theta$ is the product of n-bit integer $j \in \{0,\ldots,2^n-1\}$ and $\theta \in [0,1)$. This puts the first register into the state

$$\frac{1}{\sqrt{2^n}} \sum_{j\in\{0,1\}^n} e^{2\pi ij\theta}|j\rangle. \tag{3}$$

The cost of V is $O(2^n)$ controlled applications of U. We assume the above is implemented perfectly, or with very small error. Now we apply (to the first register) a channel C that implements F_N^{-1} and we measure the resulting n-qubit state in the computational basis, hoping that the resulting n bits give us the most significant bits of θ.

The simplest situation arises when the initial state $|\phi\rangle$ in the second register is an eigenstate of U with eigenphase $\theta = 0.\theta_1\ldots\theta_n$ that requires only n bits of precision. In this case the state of Eq. (3) is $F_N|\theta_1\ldots\theta_n\rangle$. The inverse QFT will map this to $|\theta_1\ldots\theta_n\rangle$, and the final measurement will give us the bits $\theta_1\ldots\theta_n$ with certainty. However, two complications can arise.

First, the initial state $|\phi\rangle$ could be a *superposition* of eigenstates of U (each with eigenphases requiring only n bits of precision) rather than one eigenstate. In this case phase estimation still gives useful results; the effect is the same as starting with a *mixture* of eigenstates in the second register. For example, if instead of one eigenstate $|\phi\rangle$ we start (in the second register) with a superposition $\alpha|\phi\rangle+\beta|\phi'\rangle$ of two normalised eigenstates, with distinct associated n-bit phases θ and θ', respectively, then before the final measurement the state is $\alpha|\theta_1\ldots\theta_n\rangle|\phi\rangle+\beta|\theta'_1\ldots\theta'_n\rangle|\phi'\rangle$; measuring the first n qubits gives θ with probability $|\alpha|^2$ and gives θ' with probability $|\beta|^2$.

A second complication that can arise is when θ requires more than n bits of precision. In that case the state of Eq. (3) to which we apply C will be a superposition of n-qubit Fourier basis states: something of the form $\sum_k \alpha_k|\hat{k}\rangle$ rather than one Fourier basis state. If C is a perfect F_N^{-1} then this doesn't matter: the resulting state after applying C will be $\sum_k \alpha_k|k\rangle$. However, if channel C is an imperfect implementation of the inverse QFT, then interference between different terms could cause trouble, and we have to be careful about this in the next subsections. For now, let us record the useful fact that the state of Eq. (3) is always dominated by a few Fourier basis states that correspond to good approximations of θ. Variants of this fact are already known (e.g. [CEMM98, Appendix C]) but for completeness we give a proof in the appendix.

Proposition 2. *Let $N = 2^n$, $\theta \in [0,1)$, and let coefficients $\alpha_k \in \mathbb{C}$ be such that*

$$\frac{1}{\sqrt{2^n}} \sum_{j\in\{0,1\}^n} e^{2\pi ij\theta}|j\rangle = \sum_{k=0}^{N-1} \alpha_k|\hat{k}\rangle.$$

Let $k^* = \lfloor 2^n\theta \rfloor \in \{0, \ldots, N-1\}$ be the n-bit integer corresponding to the first n bits in the binary expansion of θ, and $S = \{k^* - K + 1, \ldots, k^* - 1, k^*, k^* + 1, \ldots, k^* + K\}$ be the $2K$ integers "around" $2^n\theta$ (these integers should be taken mod N). Then $\sum_{k \notin S} |\alpha_k|^2$ can be made as small as we want by choosing K sufficiently large (independent of N):

$$\sum_{k \notin S} |\alpha_k|^2 \leq \frac{1}{4}\left(\frac{1}{K} + \frac{1}{K-1}\right).$$

Note that every k in the above set S provides an approximation of θ with small additive error, because $|\theta - k/N| < K/N \pmod 1$. It will suffice to take $K = O(1)$ below. The above upper bound can probably be improved somewhat; in the appendix we also calculate numerical upper bounds on $\sum_{k \notin S} |\alpha_k|^2$ for small values of K.

3.2 Worst-case phase estimation via average-case-correct F_N^{-1}: n-bit case

Now we switch to the scenario where we do not have a perfect inverse QFT available, but instead have a channel C which has small average infidelity w.r.t. F_N^{-1} in the sense of Theorem 1. Note that if C is usually close to F_N^{-1} but not on the particular Fourier basis state $F_N|\theta_1 \ldots \theta_n\rangle$ that we (approximately) prepared using V, then recovering the particular phase θ that we are interested in may fail miserably, even if θ can be represented exactly with n bits of precision and all the other components of phase estimation work perfectly. In other words, an average-case-correct F_N^{-1} does not guarantee that phase estimation works in the worst case, i.e., for each possible θ. However, we can do a relatively simple worst-case-to-average-case reduction to deal with the situation that C is not the perfect F_N^{-1}.

Our idea is to choose a uniformly random offset $\lambda \in [0, 1)$ that can be described with n bits of precision, change the phase θ to $\theta' = \theta + \lambda \bmod 1$, and then apply our purported F_N^{-1} on the state and measure in the hope of obtaining an approximation of θ', from which we acn then subtract λ to obtain an approximation of θ itself. In this section we first describe what happens in the case where the unknown θ can be described exactly with n bits; in the next section we deal with the more subtle general case where θ needs more than n bits. Note that if θ can be described exactly by n bits, then $\theta' = 0.\theta'_1 \ldots \theta'_n$ can be as well. In particular, $\theta'_1 \ldots \theta'_n$ is now a uniformly random n-bit string and $F_N|\theta'_1 \ldots \theta'_n\rangle$ is a *uniformly random Fourier basis state* (on which C is likely to work well if its average infidelity w.r.t. F_N^{-1} is small).

Let us first consider how to change the phase. One way to do this is to change U to $U' = e^{2\pi i\lambda}U$. This has the effect that the unitary V of Eq. (2), with U replaced by U', induces an extra phase of $e^{2\pi ij\lambda}$ on basis state $|j\rangle$, resulting in

$$\frac{1}{\sqrt{2^n}} \sum_{j \in \{0,1\}^n} e^{2\pi ij(\theta+\lambda)}|j\rangle = \frac{1}{\sqrt{2^n}} \sum_{j \in \{0,1\}^n} e^{2\pi ij\theta'}|j\rangle = F_N|\theta'_1 \ldots \theta'_n\rangle.$$

However, a probably more efficient way to achieve the same is to leave U as it is, and instead modify the n Hadamard gates at the start of the phase estimation procedure. If we change the ℓth Hadamard to a single-qubit gate that maps

$$|0\rangle \mapsto \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i 2^{n-\ell}\lambda}|1\rangle),$$

then the n-bit basis state $|j\rangle = |j_1 \ldots j_n\rangle$ gets a phase

$$\prod_{\ell=1}^{n} e^{2\pi i j_\ell 2^{n-\ell}\lambda} = e^{2\pi i(\sum_{\ell=1}^{n} j_\ell 2^{n-\ell})\lambda} = e^{2\pi ij\lambda}.$$

Thus the uniform superposition $\frac{1}{\sqrt{2^n}}\sum_{j\in\{0,1\}^n}|j\rangle$ that we prepared in the original phase estimation procedure now becomes

$$\frac{1}{\sqrt{2^n}}\sum_{j\in\{0,1\}^n} e^{2\pi ij\lambda}|j\rangle.$$

Now we apply V of Eq. (2) to multiply in the phases $e^{2\pi ij\theta}$, and the n-qubit register becomes

$$\frac{1}{\sqrt{2^n}}\sum_{j\in\{0,1\}^n} e^{2\pi ij(\theta+\lambda)}|j\rangle = F_N|\theta'_1\dots\theta'_n\rangle.$$

We have thus changed the phase from θ to $\theta' = \theta+\lambda$ as desired. Applying a perfect F_N^{-1} would give us $|\theta'_1\dots\theta'_n\rangle$ with certainty, from which we learn $\theta = \theta'-\lambda \bmod 1$.

Now suppose we have a channel C available that is not the perfect F_N^{-1} but that has small average infidelity w.r.t. the perfect F_N^{-1}, averaged uniformly over the Fourier basis states $|\hat{k}\rangle$:

$$\mathbb{E}_k[1-\langle k|\,C(|\hat{k}\rangle)\,|k\rangle\,] \leq \eta, \tag{4}$$

for some small constant η. This could be tested by running the procedure of Theorem 1.

Suppose we run C on a specific Fourier basis state $|\hat{k}\rangle$, for some $k\in\{0,1\}^n$ which would be the binary representation of the number θ. The intended n-bit outcome of a measurement in the computational basis on $C(|\hat{k}\rangle)$ would be k. Define $\eta_k = 1-\langle k|\,C(|\hat{k}\rangle)\,|k\rangle$, which is the probability of not getting the intended outcome. It could be the case that we happen to run C on a $|\hat{k}\rangle$ where η_k is particularly large; the error probability η_k could even be 1 for some k. In that case basic phase estimation would fail. However, we have $\mathbb{E}_{k'}[\eta_{k'}]\leq\eta$ by Eq. (4). So if we do our worst-case-to-average-case reduction, shifting θ by a random λ (equivalently, changing k to a uniformly random k' by adding a uniformly random n-bit integer $N\lambda$ to it, mod N), we obtain the following theorem:

Theorem 3 (case where θ is an n-bit number)**.** *Let C be a channel from n qubits to n qubits with average infidelity $\leq\eta$ w.r.t. F_N^{-1} (in the sense of Theorem 1). Let $\theta = 0.\theta_1\dots\theta_n \in \{0,1/2^n,\dots,(2^n-1)/2^n\}$ be fixed, and $\lambda\in\{0,1/2^n,\dots,(2^n-1)/2^n\}$ be uniformly random. If we apply C to state $\frac{1}{\sqrt{2^n}}\sum_{j\in\{0,1\}^n} e^{2\pi ij(\theta+\lambda)}|j\rangle$, measure in the computational basis, and subtract λ from the measurement outcome, then we get θ except with probability $\leq\eta$.*

As long as $\eta<1/2$, we can reduce the error probability to an arbitrarily small δ by $O(\log(1/\delta))$ repetitions.

The above theorem is for the basic case where the second register contains one eigenstate $|\phi\rangle$ of U, and the corresponding eigenphase can be described exactly with n bits of precision. Let us consider again the two complications mentioned near the end of Section 3.1. The first complication is where the state $|\phi\rangle$ in the second register is a *superposition* of multiple eigenstates of U rather than one eigenstate, but still assuming the eigenphases can all be decribed with at most n bits. In this case we may treat the first register as containing a mixture of different states, and still use Theorem 3. We will typically be using the estimated eigenphase to approximate some quantity of interest. If the obtained eigenphase would give a good approximation to that quantity with probability at least p in the case of a perfect F_N^{-1}, then it will still give a good approximation with probability at least $p(1-\eta)$ in the case of our imperfect channel C; that is the situation we will be in for the application to period-finding in Section 4.1. The second complication mentioned in Section 3.1 (when the eigenphases need more than n bits) is more subtle, and we deal with it next.

3.3 Worst-case phase estimation via average-case-correct F_N^{-1}: general case

If we run channel C on a *superposition* of Fourier basis states in the *first* register, then interference effects could occur between the different parts of the superposition when C is applied, and we have to be more careful. This is the situation when we do phase estimation for the case where the phase θ needs more than n bits of precision, as mentioned at the end of Section 3.1. We will now analyse what happens in this case in more detail.

Fortunately, by Proposition 2 the state $|\psi\rangle$ on which we apply C will still be dominated by $O(1)$ Fourier basis states, each of which corresponds to a good approximation of θ. Choose integer K and let S be the set of $2K$ elements of Proposition 2. We can write $|\psi\rangle$ as

$$|\psi\rangle = \sum_{k\in S} \alpha_k |\hat{k}\rangle + \alpha_\rho |\rho\rangle,$$

where $|\rho\rangle$ is the normalised "rest" of the state (which consists of the non-S Fourier basis states and hence is orthogonal to the $|\hat{k}\rangle$ with $k \in S$), and $|\alpha_\rho|^2 = 1 - \sum_{k\in S} |\alpha_k|^2$ is small, depending on our choice of K.

The phase estimation procedure ends by measuring $C(|\psi\rangle)$ in the computational basis. Every $k \in S$ will be a good measurement outcome, in the sense of corresponding to a good approximation of θ. If C were perfect then $C(|\psi\rangle)$ would be $\sum_k \alpha_k |k\rangle$, and measuring in the computational basis would give a good outcome k with probability $\sum_{k\in S} |\alpha_k|^2$, which is close to 1 by Proposition 2. However, if C is not perfect then we have to worry about the interaction between the errors that C makes on the different parts of the state. The following easy lemma implies that measuring a state $|\psi\rangle$ in a basis in which $|\psi\rangle$ has small support, cannot increase probabilities by too much.

Lemma 4. *If* $|\psi\rangle = \sum_{j\in T} \beta_j |j\rangle$ *(not necessarily normalised), then* $|\psi\rangle\langle\psi| \preceq |T| \sum_{j\in T} |\beta_j|^2 |j\rangle\langle j|$.

Proof. Let M be the matrix on the right-hand side, and consider an arbitrary state $|\phi\rangle = \sum_j \gamma_j |j\rangle$. Using Cauchy-Schwarz, we have

$$\langle\phi|(|\psi\rangle\langle\psi|)|\phi\rangle = \left|\sum_{j\in T} \beta_j^* \gamma_j\right|^2 \leq |T| \sum_{j\in T} |\beta_j|^2 |\gamma_j|^2 = \langle\phi|M|\phi\rangle,$$

which implies the lemma. □

The prefactor $|T|$ on the right-hand side is optimal whenever all β_j's are non-zero, as follows. Define $B = \sum_{j\in T} 1/|\beta_j|^2$ and $|\phi\rangle = \frac{1}{\sqrt{B}} \sum_{j\in T} \frac{1}{\beta_j^*} |j\rangle$. Then $\langle\phi| \left(\sum_{j\in T} |\beta_j|^2 |j\rangle\langle j|\right) |\phi\rangle = |T|/B$ while $\langle\phi|(|\psi\rangle\langle\psi|)|\phi\rangle = |\langle\phi|\psi\rangle|^2 = |T|^2/B$. Hence the factor-$|T|$ on the right-hand side is necessary.

In order to upper bound the probability of obtaining a bad outcome when measuring $C(|\psi\rangle\langle\psi|)$ in the computational basis, let $P_{\text{bad}} = \sum_{k\notin S} |k\rangle\langle k|$ be the projector on the bad outcomes. Define $\eta_k = 1 - \langle k| C(|\hat{k}\rangle) |k\rangle$, which is the probability of getting a measurement outcome other than k when measuring $C(|\hat{k}\rangle)$. We have $\mathbb{E}_k[\eta_k] \leq \eta$ by Eq. (4). For $k \in S$, we have $\text{Tr}(|k\rangle\langle k| C(|\hat{k}\rangle)) + \text{Tr}(P_{\text{bad}} C(|\hat{k}\rangle)) \leq \text{Tr}(C(|\hat{k}\rangle)) = 1$, hence $\text{Tr}(P_{\text{bad}} C(|\hat{k}\rangle)) \leq \eta_k$.

First consider the case that we do not shift θ by a random λ. Below we will use Lemma 4 twice, and also use the fact that the channel C is linear and preserves the positive semidefinite (psd) ordering[7], to upper bound the probability that the final measurement

[7] This follows because C is linear and positivity-preserving: if $\sigma \preceq \rho$, then $0 \preceq C(\rho - \sigma) = C(\rho) - C(\sigma)$ and hence $C(\sigma) \preceq C(\rho)$. The latter in turn implies (in fact is equivalent to) the property that $\text{Tr}(M\, C(\sigma)) \leq \text{Tr}(M\, C(\rho))$ for all psd operators M.

outcome lies outside of S.

$$\Pr[\text{bad outcome}] = \mathrm{Tr}\left(P_{\text{bad}}\, C(|\psi\rangle\langle\psi|)\right)$$

First apply Lemma 4 with T containing 2 states: $|\rho\rangle$ and the normalised version of $|\psi_S\rangle = \sum_{k\in S}\alpha_k|\hat{k}\rangle$, so that $|\psi\rangle = \| \, |\psi_S\rangle \, \| \dfrac{|\psi_S\rangle}{\| \, |\psi_S\rangle \, \|} + \alpha_\rho|\rho\rangle$. We obtain:

$$\begin{aligned}\Pr[\text{bad outcome}] &\leq \mathrm{Tr}\left(P_{\text{bad}}\, C\left(2\| \, |\psi_S\rangle \, \|^2 \frac{|\psi_S\rangle\langle\psi_S|}{\| \, |\psi_S\rangle \, \|^2} + 2|\alpha_\rho|^2|\rho\rangle\langle\rho|\right)\right)\\ &= 2\mathrm{Tr}\left(P_{\text{bad}}\, C\left(|\psi_S\rangle\langle\psi_S|\right)\right) \,+\, 2|\alpha_\rho|^2\mathrm{Tr}\left(P_{\text{bad}}\, C\left(|\rho\rangle\langle\rho|\right)\right)\end{aligned}$$

Apply Lemma 4 to $|\psi_S\rangle$ with $T = S$, and for the second term use $\mathrm{Tr}\left(P_{\text{bad}}\, C\left(|\rho\rangle\langle\rho|\right)\right) \leq 1$:

$$\begin{aligned}&\leq 2\mathrm{Tr}\left(P_{\text{bad}}\, C\left(|S|\sum_{k\in S}|\alpha_k|^2|\hat{k}\rangle\langle\hat{k}|\right)\right) + 2|\alpha_\rho|^2\\ &= 2|S|\left(\sum_{k\in S}|\alpha_k|^2\,\mathrm{Tr}\left(P_{\text{bad}}\, C(|\hat{k}\rangle)\right)\right) + 2|\alpha_\rho|^2\\ &\leq 2|S|\left(\sum_{k\in S}|\alpha_k|^2\eta_k\right) + 2|\alpha_\rho|^2.\end{aligned}$$

The latter upper bound on the probability of a bad outcome could be large if the $k \in S$ happen to be among the few k's that have large η_k-values. However, the average over all N η_k-values is small: at most η. Now consider the case where we do use our worst-case-to-average-case reduction, shifting all k by the same uniformly random n-bit integer $N\lambda$, and similarly shift the notion of a "bad outcome". Then we can upper bound the overall probability of a bad outcome by using linearity of expectation, as follows:

$$\begin{aligned}\mathbb{E}_\lambda\left[\Pr[\text{bad outcome}]\right] &\leq \mathbb{E}_\lambda\left[2|S|\left(\sum_{k\in S}|\alpha_k|^2\eta_{k+N\lambda}\right) + 2|\alpha_\rho|^2\right]\\ &= 2|S|\left(\sum_{k\in S}|\alpha_k|^2\,\mathbb{E}_\lambda\left[\eta_{k+N\lambda}\right]\right) + 2|\alpha_\rho|^2\\ &\leq 2|S|\left(\sum_{k\in S}|\alpha_k|^2\eta\right) + 2|\alpha_\rho|^2\\ &= 2|S|(1-|\alpha_\rho|^2)\eta + 2|\alpha_\rho|^2\\ &= 2|S|\eta + 2(1-|S|\eta)|\alpha_\rho|^2. \qquad (5)\end{aligned}$$

Note the tradeoff here: $|S| = 2K$ increases with our choice of K, while $|\alpha_\rho|^2$ decreases.

We summarize the above calculations in a theorem, using Proposition 2 to bound $|\alpha_\rho|^2$:

Theorem 5 (case where θ may need more than n bits of precision)**.** *Let C be a channel from n qubits to n qubits with average infidelity $\leq \eta$ w.r.t. F_N^{-1} (in the sense of Theorem 1). Let $\theta \in [0,1)$ be fixed, and $\lambda \in \{0, 1/2^n, \ldots, (2^n-1)/2^n\}$ be uniformly random. If we apply C to state $\frac{1}{\sqrt{2^n}}\sum_{j\in\{0,1\}^n} e^{2\pi i j(\theta+\lambda)}|j\rangle$, measure in the computational basis, and subtract λ from the measurement outcome, then for every integer $K \geq 1$, the probability to get an outcome $k \in \{0,\ldots,N-1\}$ such that $|\theta - k/N| > K/N$ (mod 1) is at most*

$$4K\eta + (1/2 - K\eta)\left(\frac{1}{K} + \frac{1}{K-1}\right).$$

For example, if we use $K = 4$ (considering the $|S| = 2K = 8$ k's that are closest to $2^n\theta$ to be the good measurement outcomes) and we set $\eta \leq 0.015$, then the probability of a bad outcome will be < 0.497. Thus the probability of obtaining a good approximation of the phase θ is > 0.503, and we can amplify this success probability to be close to 1 by taking the median of multiple runs of this procedure.

We can improve these bounds on the tolerable average infidelity η a bit by plugging the numerical upper bounds on $|\alpha_\rho|^2$ from the end of the appendix into Eq. (5), taking $N = 2^{10}$ for concreteness (for other values of N the numbers are very similar). With $K = 4$ and using $|\alpha_\rho|^2 \approx 0.05$ from the appendix, the probability of a bad outcome is $\leq 16\eta + 2(1 - 8\eta)0.05$, which is < 0.5 as long as the average infidelity η is ≤ 0.026. With $K = 3$, using $|\alpha_\rho|^2 \approx 0.067$ from the appendix, the probability of a bad outcome is $\leq 12\eta + 2(1 - 6\eta)0.067$, which is < 0.5 as long as the average infidelity η is ≤ 0.032. With $K = 2$, using $|\alpha_\rho|^2 \approx 0.099$ from the appendix, we can tolerate $\eta \leq 0.041$. Even with these numerical improvements, the tolerable η is still much smaller here than in the case where θ is an n-bit number (Theorem 3): there we could tolerate any $\eta < 1/2$.

4 Applications: period-finding and amplitude estimation

4.1 Period-finding

Fix $N = 2^n$.[8] For fixed *period* $r < N$ and variable *offset* $s \in \{0, \ldots, r-1\}$, define periodic state

$$|\pi_s\rangle = \frac{1}{\sqrt{p}} \sum_{z=0}^{p-1} |s + zr\rangle,$$

where $p = |\{z : 0 \leq s + zr < N\}|$ is the number of basis states occurring in the superposition. This p will be N/r rounded up; N has $\lfloor N/r \rfloor$ "complete" sequences of r indices followed by one "incomplete" sequence of $N - r\lfloor N/r \rfloor$ indices. Shor [Sho97] showed via classical number theory that the ability to find the period r given such a state suffices for factoring integers, and gave an efficient quantum algorithm for period-finding. Subsequently Kitaev [Kit95] showed how to do period-finding using phase estimation, and then Cleve, Ekert, Mosca, and Macchiavello [CEMM98] showed that Shor's and Kitaev's approaches to period-finding are basically the same.

In order to do period-finding via phase estimation starting from a periodic state, we let U be the "+1 mod N" operator:

$$U|x\rangle = |x + 1 \bmod N\rangle.$$

It is easily verified that the eigenstates of U are the states $F_N^{-1}|j\rangle$ with corresponding eigenvalue $\omega_N^j = e^{2\pi ij/N}$. Note that n bits of precision suffice for each of these eigenphases.

Because these eigenstates form a basis, any state $|\phi\rangle$ can be written as a superposition $\sum_{j=0}^{N-1} \alpha_j F_N^{-1}|j\rangle$ of eigenstates of U with some coefficients α_j. We already saw how phase estimation using U acts when we start with such a superposition in the second register: with probability $|\alpha_j|^2$ it returns the n-bit number j/N exactly. We now determine these α_j coefficients for the case where our starting state is the periodic state $|\pi_s\rangle = \sum_j \alpha_j F_N^{-1}|j\rangle$, by multiplying $|\pi_s\rangle$ with F_N (we do this only as a calculational device; we are not implementing

[8] To avoid confusion with Shor's algorithm: our N here is the dimension of the QFT, not an integer to be factored.

the QFT physically):

$$\sum_{j=0}^{N-1} \alpha_j |j\rangle = F_N|\pi_s\rangle = \frac{1}{\sqrt{p}} \sum_{z=0}^{p-1} F_N|s+zr\rangle$$
$$= \frac{1}{\sqrt{p}} \sum_{z=0}^{p-1} \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \omega_N^{j(s+zr)} |j\rangle = \sum_{j=0}^{N-1} \underbrace{\frac{\omega_N^{js}}{\sqrt{pN}} \sum_{z=0}^{p-1} \omega_N^{jzr}}_{\alpha_j} |j\rangle.$$

We now want to show that the amplitude is concentrated around integer multiples of N/r. First consider the special case where N/r happens to be an integer. If $j = cN/r$ for some integer $c \in \{0, \ldots, r-1\}$, then we have $\omega_N^{jzr} = 1$ for all z and hence $|\alpha_j|^2 = p/N = 1/r$; there are r such j's, each with squared amplitude $1/r$, so the j's that are not integer multiples of N/r will have amplitude 0 in this case.

In the general case where N/r is not an integer, let j be the closest integer to cN/r for some $c \in \{0, \ldots, r-1\}$ (i.e., $j = cN/r + \delta$ for $\delta \in (-1/2, 1/2]$). Then

$$|\alpha_j|^2 = \frac{1}{pN} \left| \sum_{z=0}^{p-1} \omega_N^{jzr} \right|^2 = \frac{1}{pN} \frac{|1-\omega_N^{pjr}|^2}{|1-\omega_N^{jr}|^2} = \frac{1}{pN} \frac{|1-e^{2\pi i p\delta r/N}|^2}{|1-e^{2\pi i \delta r/N}|^2} = \frac{1}{pN} \frac{\sin(\pi p\delta r/N)^2}{\sin(\pi\delta r/N)^2}$$
$$\geq \frac{1}{pN} \frac{(\frac{2}{\pi}\pi p\delta r/N)^2}{(\pi\delta r/N)^2} = \frac{4p}{\pi^2 N} \geq \frac{4}{\pi^2 r},$$

using $\frac{2}{\pi}x \leq \sin(x) \leq x$ for $x \in [0, \pi/2]$, and assuming $\delta \neq 0$. If indeed $\delta \neq 0$, then the probability that the measurement outcome j is one of the two integers in the interval $(cN/r - 1, cN/r + 1)$ is $\geq 8/(\pi^2 r)$.[9] If $\delta = 0$, then $|\alpha_j|^2 = p/N \geq 1/r$. Accordingly, with probability $\geq 8/\pi^2$ our measurement outcome j is cN/r (rounded up or down) for some random integer $c \in \{0, \ldots, r-1\}$. Note that in that case we have $|j/N - c/r| < 1/N$, so the known ratio j/N is a very good approximation to the unknown ratio c/r. Shor showed that if c and r are coprime (which, by classical number theory, happens with largish probability $\Omega(1/\log\log r)$), then continued-fraction expansion on the known ratio j/N yields the period r (as mentioned, this suffices for factoring).

Using our worst-case to average-case reduction, this method for period-finding still works when we only have a good-on-average inverse QFT for our phase estimation, provided the initial periodic state $|\pi_s\rangle$ is prepared sufficiently well and the other components of phase estimation (the unitary V from Eq. (2) and the n modified Hadamard gates) also work sufficiently well. Note that the relevant eigenphases can all be described by n bits here, so we only need to refer to the result in Section 3.2 and not to the more complicated result in Section 3.3, where the tolerable η is worse. By the result in Section 3.2, since the probability to find a good j is $8/\pi^2$ using a perfect inverse QFT, this probability is still $\geq (1-\eta)8/\pi^2$ if we instead use a channel C that has average infidelity $\leq \eta$ w.r.t. the perfect inverse QFT.

[9] It need not be the case that each of the two outcomes $\lceil cN/r \rceil$ and $\lfloor cN/r \rfloor$ has probability $\geq 4/(\pi^2 r)$, because one of them will correspond to a $\delta \notin (-1/2, 1/2]$. However, the sum of these two probabilities is at least $8/(\pi^2 r)$, which may be verified by noting that the function $f(x) = \sin(\pi x)^2/\sin(\pi x/p)^2 + \sin(\pi(1-x))^2/\sin(\pi(1-x)/p)^2$ is at least $8p^2/\pi^2$ on the interval $x \in [0, 1/2]$. We note that this calculation is similar to that in the appendix; here, however, we have a superposition of eigenstates each with n-bit eigenphase; in the appendix similar calculations are needed to deal with the situation that the eigenphase needs more than n bits of precision.

4.2 Amplitude estimation

Suppose we have a unitary quantum algorithm A that maps

$$|0^m\rangle \mapsto A|0^m\rangle = \sin(\mu)|\phi_1\rangle|1\rangle + \cos(\mu)|\phi_0\rangle|0\rangle,$$

for some arbitrary normalised states $|\phi_1\rangle$ and $|\phi_0\rangle$, and some angle $\mu \in [0, \pi/2]$. We would like to estimate the angle μ or (which comes to the same thing) the amplitude $\sin(\mu)$. Such "amplitude estimation" can be used for instance for optimal quantum approximate counting [BHMT02] and is a key component of many other quantum algorithms that involve estimating various quantities. For completeness we here sketch the elegant method of Brassard et al. [BHMT02] that reduces amplitude estimation to phase estimation. Consider the unitary

$$U = AR_0A^{-1}(I \otimes Z),$$

where $R_0 = 2|0^m\rangle\langle 0^m| - I$ reflects about $|0^m\rangle$. This R_0 can be implemented using a circuit with $O(m)$ elementary gates. Consider how U acts in the 2-dimensional space $\mathcal{S}$ spanned by $|\phi_1\rangle|1\rangle$ and $|\phi_0\rangle|0\rangle$.[10] U is the product of two reflections: first $I \otimes Z$ reflects about $|\phi_0\rangle|0\rangle$, and then AR_0A^{-1} reflects about $A|0^m\rangle$. This product of two reflections corresponds (in the space $\mathcal{S}$) to a rotation over twice the angle μ that exists between $|\phi_0\rangle|0\rangle$ and $A|0^m\rangle = \sin(\mu)|\phi_1\rangle|1\rangle + \cos(\mu)|\phi_0\rangle|0\rangle$. Such a rotation over angle 2μ has two eigenvectors in $\mathcal{S}$, with respective eigenvalues $e^{i2\mu}$ and $e^{-i2\mu}$.

How do we use phase estimation to estimate μ? We start in state $A|0^m\rangle$, which lies in $\mathcal{S}$ and hence is some linear combination of the two eigenvectors (with unknown coefficients, but that doesn't matter). If we run phase estimation, then the output will either be an estimate of μ/π or of $-\mu/\pi$ (or rather, $1 - \mu/\pi$). Since we assumed $\mu \in [0, \pi/2]$, the phase μ/π that we're estimating lies in $[0, 1/2]$. If our estimate is in $[0, 1/2]$ then we'll assume it's μ/π, and if our estimate is in $[1/2, 1)$ then we'll assume it's $1 - \mu/\pi$. Either way we obtain a good estimate of μ.

This method still works with an only-good-on-average channel for F_N^{-1} if we do our worst-case to average-case reduction. If we want n bits of precision in our estimate of μ, then the cost will be $O(2^n)$ applications of U. Note that here we need the more complicated result of Section 3.3, since the eigenphases $\pm\mu/\pi$ may require more than n bits of precision.

5 Summary and future work

In this paper we did two things. First, we showed that one can efficiently test whether a given n-qubit channel C implements the inverse QFT well, on average over all Fourier basis states. Our procedure estimates (with success probability $\geq 1 - \delta$) the average (in)fidelity up to $\pm\varepsilon$ using $O(\log(1/\delta)/\varepsilon^2)$ runs, each of which uses $O(n)$ single-qubit gates to prepare a product state, one run of C, and a measurement in the computational basis.

Second, we showed that such an average-case-correct inverse QFT suffices to implement phase estimation *in the worst case*. This implies that an average-case-correct inverse QFT also suffices for period-finding (as in Shor's algorithm) and for amplitude estimation, provided the other components of those procedures work sufficiently well.

Practical methods of verification for large numbers of qubits will be vital for future quantum computers. Here we have shown that such verification is possible for several

[10] It is very helpful to picture this similarly to the usual analysis of Grover's algorithm, in a 2d plane where the vertical axis corresponds to the state $|\phi_1\rangle|1\rangle$ and the horizontal axis corresponds to $|\phi_0\rangle|0\rangle$.

key algorithmic primitives. We feel it would be very interesting to find more examples of worst-case-to-average-case reductions for quantum computing.

One more example would be the quantum algorithm of Jordan [Jor05] for computing the gradient of a differentiable function $f : \mathbb{R}^d \to \mathbb{R}$, i.e., the vector of the d partial derivatives of f evaluated at a given point. Let's say we want to approximate the gradient with n bits of precision for each of its d entries. Jordan's algorithm sets up a uniform superposition over some grid of inputs $(x_1, \ldots, x_d)$ in d n-qubit registers, then computes f once in the phase, and then applies an inverse QFT to each of the d registers. If f is affine-linear, i.e., there are coefficients $a, b_1, \ldots, b_d \in \mathbb{R}$ such that $f(x) = a + b_1 x_1 + \cdots + b_d x_d$ for all $x \in \mathbb{R}^n$, and each b_i can be represented exactly with n bits of precision, then the final state gives $b_1, \ldots, b_d$ which is exactly the gradient of f. If some of the b_i's need more than n bits of precision then at the end of the algorithm the ith register contains b_i with high probability. This algorithm "costs" only one f-evaluation, d runs of the n-qubit inverse QFT, and $O(dn)$ other elementary gates. In contrast, classical algorithms need d f-evaluations to compute the gradient. In the more general case where f is fairly smooth but only close to an affine-linear function on the grid of points that we feed into it, a much more complicated analysis due to Gilyén, Arunachalam, and Wiebe [GAW19] determines the complexity of approximating the gradient in various norms, in terms of how many times f needs to be evaluated.

What if we only have a good-on-average inverse QFT available? If f is affine-linear then one may confirm that the analysis of our Section 3.3 implies that Jordan's algorithm can still be made to work via our worst-case-to-average-case reduction. If f is only close to affine-linear then it is not clear what happens, and whether the more subtle analysis of [GAW19] can also be made to work with an imperfect inverse QFT. We leave this question to further work.

Acknowledgements.

We thank Timothy Browning for helpful discussions regarding some number theory for Shor's algorithm, and the anonymous Quantum referees for very helpful comments that substantially improved the presentation and rigour of the paper.

NL was partially supported by the UK Engineering and Physical Sciences Research Council through grants EP/R043957/1, EP/S005021/1, EP/T001062/1. RdW was partially supported by the Dutch Research Council (NWO) through Gravitation-grant Quantum Software Consortium, 024.003.037, and through QuantERA ERA-NET Cofund project QuantAlgo 680-91-034.

This paper did not involve any underlying data.

References

[AR20] Scott Aaronson and Patrick Rall. Quantum approximate counting, simplified. In *Proceedings of 3rd Symposium on Simplicity in Algorithms (SOSA)*, pages 24–32, 2020. arXiv:1908.10846.

[BBBV97] Charles H. Bennett, Ethan Bernstein, Gilles Brassard, and Umesh Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, 1997. quant-ph/9701001.

[BHMT02] Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. In *Quantum Computation and Quantum*

Information: A Millennium Volume, volume 305 of *AMS Contemporary Mathematics Series*, pages 53–74. 2002. quant-ph/0005055.

[CB21] Chi-Fang Chen and Fernando G. S. L. Brandão. Concentration for Trotter error. arXiv:2111.05324, 9 Nov 2021.

[CEMM98] Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca. Quantum algorithms revisited. In *Proceedings of the Royal Society of London*, volume A454, pages 339–354, 1998. quant-ph/9708016.

[Cop94] Don Coppersmith. An approximate Fourier transform useful in quantum factoring. IBM Research Report No. RC19642, quant-ph/0201067, 1994.

[dSLCP11] Marcus da Silva, Oliver Landon-Cardinal, and David Poulin. Practical characterization of quantum devices without tomography. *Physical Review Letters*, 107:210404, 2011. arXiv:1104.3835.

[EHW+20] Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud, and Elham Kashefi. Quantum certification and benchmarking. *Nature Reviews Physics*, 2:382–390, 2020. arXiv:1910.06343.

[FL11] Steven T. Flammia and Yi-Kai Liu. Direct fidelity estimation from few Pauli measurements. *Physical Review Letters*, 106:230501, 2011. arXiv:1104.4695.

[GAW19] András Gilyén, Srinivasan Arunachalam, and Nathan Wiebe. Optimizing quantum optimization algorithms via faster quantum gradient computation. In *Proceedings of 30th ACM-SIAM SODA*, pages 1425–1444, 2019. arXiv:1711.00465.

[Gro96] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of 28th ACM STOC*, pages 212–219, 1996. quant-ph/9605043.

[GSLW19] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of 51st ACM STOC*, pages 193–204, 2019. arXiv:1806.01838.

[Jor05] Stephen P. Jordan. Fast quantum algorithm for numerical gradient estimation. *Physical Review Letters*, 95:050501, 2005. quant-ph/0405146.

[Kit95] Alexey Yu. Kitaev. Quantum measurements and the Abelian stabilizer problem. quant-ph/9511026, 12 Nov 1995.

[LW21] Noah Linden and Ronald de Wolf. Lightweight detection of a small number of large errors in a quantum circuit. *Quantum*, 5(436), 2021. arXiv:2009.08840.

[Mah18] Urmila Mahadev. Classical verification of quantum computations. In *Proceedings of 59th IEEE FOCS*, pages 259–267, 2018. arXiv:1804.01082.

[MRTC21] John M. Martyn, Zane M. Rossi, Andrew K. Tan, and Isaac L. Chuang. A grand unification of quantum algorithms. *PRX Quantum*, 2:040203, 2021. arXiv.2105.02859.

[NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Ral21] Patrick Rall. Faster coherent quantum algorithms for phase, energy, and amplitude estimation. *Quantum*, 5(566), 2021. arXiv:2103.09717.

[Sho97] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997. Earlier version in FOCS'94. quant-ph/9508027.

[ZZS+21] Qi Zhao, You Zhou, Alexander F. Shaw, Tongyang Li, and Andrew M. Childs. Hamiltonian simulation with random inputs. arXiv:2111.04773, 8 Nov 2021.

A Proof of Proposition 2

Let $N = 2^n$, $\theta \in [0,1)$, and let coefficients $\alpha_k \in \mathbb{C}$ be such that

$$\frac{1}{\sqrt{N}} \sum_{j\in\{0,1\}^n} e^{2\pi i j\theta}|j\rangle = \sum_{k=0}^{N-1} \alpha_k |\hat{k}\rangle. \tag{6}$$

Let $k^* = \lfloor N\theta \rfloor \in \{0,\dots,N-1\}$ be the n-bit integer corresponding to the first n bits in the binary expansion of θ, $x = N\theta - k^* \in [0,1)$ be the fractional part of $N\theta$. Let $S = \{k^*-K+1,\dots,k^*-1,k^*,k^*+1,\dots,k^*+K\}$ be the set of $2K$ integers "around" $2^n\theta$ (these integers should be taken mod N). These $2K$ integers come in K pairs: $\{k^*,k^*+1\},\{k^*-1,k^*+2\},\{k^*-2,k^*+3\},\dots,\{k^*-K+1,k^*+K\}$.

Our goal is to show that the total probability $P = \sum_{k\in S}|\alpha_k|^2$ of terms in the set S is close to 1. By applying inverse QFT to both sides of Eq. (6) and rewriting the geometric series, we have

$$\alpha_k = \frac{1}{N}\frac{e^{2\pi i(N\theta-k)}-1}{e^{2\pi i(N\theta-k)/N}-1}, \text{ hence } |\alpha_k|^2 = \frac{1}{N^2}\frac{\sin^2(\pi(N\theta-k))}{\sin^2(\pi(N\theta-k)/N)} = \frac{1}{N^2}\frac{\sin^2(\pi x)}{\sin^2(\pi(N\theta-k)/N)}.$$

So

$$P = \frac{1}{N^2}\sum_{k=k^*-K+1}^{k^*+K}\frac{\sin^2(\pi x)}{\sin^2(\pi(k^*-k+x)/N)} = \frac{1}{N^2}\sum_{k=-K+1}^{K}\frac{\sin^2(\pi x)}{\sin^2(\pi(x-k)/N)} \tag{7}$$

We will now prove an upper bound on the probability of error (the analysis is similar to that in [NC00] section 5.2.1).

$$\begin{aligned}
1-P &= \frac{1}{N^2}\sum_{k=-N/2+1}^{-K}\frac{\sin^2(\pi x)}{\sin^2(\pi(x-k)/N)} + \frac{1}{N^2}\sum_{k=K+1}^{N/2}\frac{\sin^2(\pi x)}{\sin^2(\pi(x-k)/N)} \\
&\le \frac{\sin^2(\pi x)}{4}\Big(\sum_{k=-N/2+1}^{-K}\frac{1}{(x-k)^2} + \sum_{k=K+1}^{N/2}\frac{1}{(x-k)^2}\Big) \\
&= \frac{\sin^2(\pi x)}{4}\sum_{k=K+1}^{N/2}\Big(\frac{1}{(k-x)^2}+\frac{1}{(k-1+x)^2}\Big) \\
&\le \frac{\sin^2(\pi x)}{4}\int_K^{N/2} dt\Big(\frac{1}{(t-x)^2}+\frac{1}{(t-1+x)^2}\Big) \\
&\le \frac{\sin^2(\pi x)}{4}\int_K^{\infty} dt\Big(\frac{1}{(t-x)^2}+\frac{1}{(t-1+x)^2}\Big) \\
&= \frac{\sin^2(\pi x)}{4}\Big(\frac{1}{(K-x)}+\frac{1}{(K-1+x)}\Big) \\
&\le \frac{1}{4}\Big(\frac{1}{(K-x)}+\frac{1}{(K-1+x)}\Big) \\
&\le \frac{1}{4}\Big(\frac{1}{K}+\frac{1}{K-1}\Big).
\end{aligned} \tag{8}$$

The final inequality is easy to see because the function is maximised on $[0,1]$ at the two endpoints of that interval ($x = 0$ or $x = 1$). We use this bound in Proposition 2 in the body of the paper.

In fact, numerical evidence suggests that the above rigorous bound is rather loose. The probability of Eq. (7) is symmetric around $x = 1/2$, because replacing x by $1 - x$ leaves P invariant (replacing x by $1 - x$ interchanges the two squared amplitudes $|\alpha_{k^*+k}|^2$ and $|\alpha_{k^*-k+1}|^2$, leaving the sum of the pair invariant). Hence P has a local optimum at $x = 1/2$. Plotting the graphs for small K strongly suggests that P is actually minimised at $x = 1/2$. If this is indeed the case, then (using the symbol $\tilde{P}$ to denote this conjectured bound)

$$\begin{aligned}
\tilde{P} &\geq \frac{1}{N^2} \sum_{k=-K+1}^{K} \frac{1}{\sin^2(\pi(\frac{1}{2}-k)/N)} \\
&= \frac{1}{N^2} \sum_{k=-K}^{K-1} \frac{1}{\sin^2(\pi(\frac{1}{2}+k)/N)} \\
&\geq \frac{4}{\pi^2} \sum_{k=-K}^{K-1} \frac{1}{(2k+1)^2} \\
&= \frac{8}{\pi^2} \sum_{k=0}^{K-1} \frac{1}{(2k+1)^2} \\
&= \frac{8}{\pi^2} \Big(\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} - \sum_{k=K}^{\infty} \frac{1}{(2k+1)^2} \Big) \\
&= \frac{8}{\pi^2} \Big(\frac{\pi^2}{8} - \sum_{k=K}^{\infty} \frac{1}{(2k+1)^2} \Big) \\
&\geq \frac{8}{\pi^2} \Big(\frac{\pi^2}{8} - \int_{K-1}^{\infty} \frac{dy}{(2y+1)^2} \Big) \\
&= 1 - \frac{4}{\pi^2(2K-1)}
\end{aligned} \tag{9}$$

Numerical evidence shows that this asymptotic bound is reasonably close to the sum for moderate-sized and large K, but still somewhat off for small K:

- For $K = 2$, the proven upper bound for the error of Eq. (8) is 0.375, whereas the upper bound for the error using (9) is 0.135, and the numerical evaluation of the exact formula (7) for $1 - P$ for $K = 2$, e.g. $N = 2^{10}$ and $x = 1/2$, is 0.099.
- For $K = 3$, the proven upper bound for the error of Eq. (8) is 0.208, whereas the upper bound for the error using (9) is 0.081, and the numerical evaluation of the exact formula (7) for $1 - P$ for $K = 3$, e.g. $N = 2^{10}$ and $x = 1/2$, is 0.067.
- For $K = 4$, the proven upper bound for the error of Eq. (8) is 0.146, whereas the upper bound for the error using (9) is 0.058, and the numerical evaluation of the exact formula (7) for $1 - P$ for $K = 4$, e.g. $N = 2^{10}$ and $x = 1/2$, is 0.050.