Optimal Quantum Circuits for General Phase Estimation

Wim van Dam,[1] G. Mauro D'Ariano,[2] Artur Ekert,[3] Chiara Macchiavello,[2] and Michele Mosca[4]

[1]*Departments of Computer Science and Physics, University of California, Santa Barbara, Santa Barbara, California 93106-5110, USA*

[2]*Dipartimento di Fisica "A. Volta" and CNISM, via Bassi 6, 27100 Pavia, Italy*

[3]*Mathematical Institute, University of Oxford, 24-29 St. Giles', Oxford, OX1 3LB, United Kingdom, and Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542*

[4]*Institute for Quantum Computing, University of Waterloo, N2L 3G1, Waterloo, ON, Canada, St. Jerome's University, N2L 3G3, Waterloo, ON, Canada, and Perimeter Institute for Theoretical Physics, N2L 2Y5, Waterloo, ON, Canada*

(Received 24 September 2006; published 1 March 2007)

We address the problem of estimating the phase ϕ given N copies of the phase-rotation gate u_ϕ. We consider, for the first time, the optimization of the general case where the circuit consists of an arbitrary input state, followed by any arrangement of the N phase rotations interspersed with arbitrary quantum operations, and ending with a general measurement. Using the polynomial method, we show that, in all cases where the measure of quality of the estimate $\tilde{\phi}$ for ϕ depends only on the difference $\tilde{\phi} - \phi$, the optimal scheme has a very simple fixed form. This implies that an optimal general phase estimation procedure can be found by just optimizing the amplitudes of the initial state.

DOI: 10.1103/PhysRevLett.98.090501 PACS numbers: 03.67.Lx, 03.65.Ta, 03.65.Vf

The possibility of encoding information into the relative phase of quantum systems is often exploited in quantum information processing tasks and several kinds of applications (e.g., [1,2]). Moreover, information is encoded into phase properties in some quantum cryptographic protocols [3] and in some precision measurements, such as the schemes on which atomic clocks are based [4]. Therefore, the issue of estimating the phase in the most efficient way is of great interest.

We phrase the phase estimation problem as follows. Let u_ϕ be a single qubit gate that in a prescribed "computational" basis $\{|0\rangle, |1\rangle\}$ maps the state $|0\rangle$ to $|0\rangle$ and $|1\rangle$ to $e^{i\phi}|1\rangle$. We assume that we have no prior knowledge about ϕ. The objective is to estimate ϕ using some procedure that will output some guess $\tilde{\phi}$. We characterize the quality of an estimate by a "cost function" $C(\phi, \tilde{\phi})$, which specifies the penalty associated with guessing $\tilde{\phi}$ when the actual phase is ϕ. We are given N identical single qubit quantum gates u_ϕ, and the goal is to use these gates along with any other operations in order to produce an estimate of ϕ. The optimal procedure is the one that has the minimum expected cost.

Most of the previous work on phase estimation assumes some fixed state encoding the phases, and the only thing to be optimized is the final general measurement [i.e., a measurement based on a positive-operator-valued measure (POVM)] [5–7]. More recent work [8] fixes the way the phase gates are applied and optimizes the choice of input state and final POVM or achieves optimal bounds for specific cost functions (e.g., [9] or [10], which is based on the Cramer-Rao bound).

The crucial point is that in this Letter we are not restricted to preparing some input state, then applying all of the phase rotations, and then performing an optimal POVM. We consider the case where one has full freedom over how to use the phase-rotation gates in an experiment designed to optimally estimate the phase. Any realistic experiment of this type can be viewed as computation, completely specified by a quantum circuit acting on some finite number of qubits and involving, apart from the N copies of the u_ϕ gates, some finite number of arbitrary quantum gates of our choice. In fact, many quantum algorithms, including Shor's quantum algorithm for factoring integers, can be phrased in terms of such phase estimations [1,2]. This originally provided the motivation for this work.

We assume that the phase ϕ is chosen uniformly from $[0, 2\pi)$ [11] and that a suitable quantum circuit containing N copies of the u_ϕ gates outputs some value y with probability $\Pr(y|\phi)$. From y we infer, following a prescribed rule, the estimate $\tilde{\phi}_y$. The quality of the whole procedure is quantified by the expected cost $\bar{C}$, given by

$$\bar{C} = \frac{1}{2}\sum_y \int_{\phi=0}^{2\pi} d\phi \, \Pr(y|\phi) C(\phi, \tilde{\phi}_y). \tag{1}$$

The next part of this Letter describes a very simple procedure for estimating ϕ that only requires one to optimize the choice of initial state to an otherwise fixed procedure. The rest of the Letter then reduces the very general case we have described above to this very simple case.

For a given cost function, the quality of an estimation procedure depends on both $\Pr(y|\phi)$ and the inference rule $y \mapsto \tilde{\phi}_y$. The optimal protocol gives the minimum possible average $\bar{C}$. We restrict attention to cost functions C that depend only on $\phi - \tilde{\phi}_y$ and, therefore, adopt the notation

$C(\phi, \tilde{\phi}_y) = C(\phi - \tilde{\phi}_y)$. We will make only the following very weak assumption on the cost function (which corresponds to a more general class of cost functions than the "Holevo" class that is typically considered [12]):

$$\int_{\phi=0}^{2\pi} d\phi |C(\phi)| < \infty. \tag{2}$$

We will deal with specific cost functions later.

Let us start by describing a simple and natural approach for estimating the phase ϕ, illustrated in Fig. 1.

Procedure 1.—(a) Prepare m qubits in state $|x\rangle = \sum_{j=0}^{N} \alpha_j |j\rangle$, with $N \leq 2^m - 1$. The exact values of α_j depend on the cost function to be maximized. (b) Apply the u_ϕ gates to effect U_ϕ

$$U_\phi \sum_{j=0}^{N} \alpha_j |j\rangle = \sum_{j=0}^{N} \alpha_j e^{ij\phi} |j\rangle. \tag{3}$$

(c) Apply the inverse quantum Fourier transform to obtain

$$2^{-m/2} \sum_{y=0}^{2^m-1} \left(\sum_{j=0}^{N} \alpha_j e^{ij(\phi - (2\pi y/2^m))} \right) |y\rangle, \tag{4}$$

measure y, and calculate the estimate $\tilde{\phi}_y = 2\pi y/2^m$.

The surprising claim is the following. Given *any* function C satisfying Eq. (2), the minimum of $\bar{C}$ obtained by optimizing the α_j in procedure 1 is the infimum of all values obtainable by *any* realistic experiment (as we described above and illustrate in Fig. 2). It is important to also note that, apart from the preparation of the initial state, the above procedure can be implemented using the N black boxes u_ϕ and a number of elementary gates polynomial in $\log(N)$. Efficient preparation of states is discussed in Ref. [13]. Exact implementation of quantum Fourier transforms is discussed in Ref. [14], and arbitrarily good approximations are discussed in Refs. [1,15]. How to generalize the circuit in Fig. 1 to work for any positive integer N is shown in Ref. [8].

The remainder of this Letter will prove this claim by a sequence of reductions.

Let us start with a very general circuit (Fig. 2) which uses $m + d$ qubits, where m and d can be arbitrarily large. The first m qubits are measured after the computation, yielding the output $0 \leq y \leq 2^m - 1$, whereas the remaining d qubits are discarded.

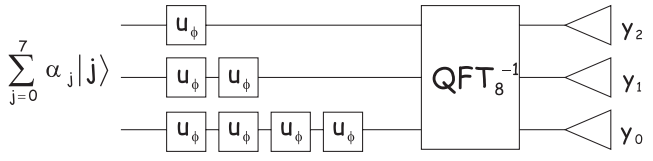

FIG. 1. A simple approach for estimating ϕ in the case that $N = 7$. Optimizing over the input amplitudes α_j produces an optimal estimate of ϕ.

Since we are allowing arbitrarily many extra "ancilla" qubits, and since any classical feedback scheme can, in principle, be implemented by a unitary operation using a sufficiently large ancilla, the family of schemes that can be implemented by a quantum circuit of this form includes any scheme using finite dimensional state spaces.

For convenience, we let the output y correspond to the phase estimate $\bar{\phi}_y = \frac{2\pi y}{M}$, for some $M \leq 2^m$. Without loss of generality [16], we can assume $M \geq N + 1$, which is necessary for this reduction. For a fixed approximation scheme (using finite means) and cost function satisfying Eq. (2), and assuming a uniform prior distribution of the ϕ, this simplifying assumption will give us a scheme with an expected cost that is at most $\bar{C} + \epsilon_M$, where $\epsilon_M \to 0$ as $M \to \infty$, and $\bar{C}$ is the lowest expected cost for any possible scheme. Thus, the infimum of the $\bar{C}$ over all such restricted schemes equals the infimum of the $\bar{C}$ over all possible such schemes.

In fact, we will also show later that assuming the inference rule has this special form does not cost us anything. That is, the infimum of the expected costs of all of the schemes using the inference rule $y \mapsto \frac{2\pi y}{M}$ for any $M \geq N + 1$ is the infimum of the possible expected costs using *any* inference rule $y \mapsto \tilde{\phi}_y$.

Suppose we came up with a general circuit that performs an estimation of ϕ according to some prescribed set of criteria. Let us first show that such a circuit is equivalent, for our purposes, to another one, which has a much simpler structure.

The state at the output of the circuit can be written as

$$\sum_{y=0}^{M-1} \sum_{z=0}^{2^d-1} \alpha(y, z, \phi) |y\rangle |z\rangle, \tag{5}$$

where each amplitude $\alpha(y, z, \phi)$ is a polynomial in $e^{i\phi}$ of degree at most N:

$$\alpha(y, z, \phi) = \sum_{j=0}^{N} \frac{\alpha_j(y, z)}{\sqrt{M}} e^{ij\phi}. \tag{6}$$

This fact follows by an induction proof just as in Ref. [17], where the polynomial method is applied to an oracle revealing one of many Boolean variables. The coefficients can be expressed as polynomials in $e^{i\phi}$, initially of de-

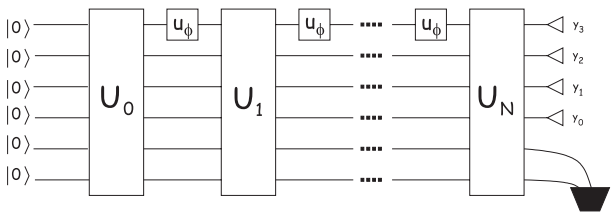

FIG. 2. The most general approach for estimating the unknown phase rotation. This subsumes procedure 1, as well as more complicated procedures with classical feedback.

gree 0. Every application of u_ϕ increases the degree by at most 1. Any intermediate unitaries transform the coefficients linearly and, thus, do not increase the degree.

Since we assume that the cost function is of the form $C(\phi, \tilde{\phi}_y) = C(\phi - \tilde{\phi}_y)$, then by an argument [18] similar to the Hunt-Stein theorem [12] the optimal conditional probability

$$\Pr(y|\phi) = \sum_z |\alpha(y, z, \phi)|^2 \tag{7}$$

can without loss of generality be assumed to depend only on the difference $\phi - \frac{2\pi y}{M}$ and, therefore, equals

$$\Pr(0|\phi - 2\pi y/M) = \sum_z |\alpha(0, z, \phi - 2\pi y/M)|^2. \tag{8}$$

To simplify the notation, we let $\alpha_j(z) = \alpha_j(0, z)$. Therefore, a circuit that produces amplitudes

$$|\alpha(y, z, \phi)| = \left| \sum_{j=0}^{N} \alpha_j(z) e^{ij(\phi - (2\pi y/M))} \right| 1/\sqrt{M} \tag{9}$$

also leads to the optimal estimation of ϕ. Thus, the following simple estimation procedure, whose circuit is illustrated in Fig. 3, performs equally well: (a) Prepare $m + d$ qubits in state $|x\rangle = \sum_{z=0}^{2^d-1} \sum_{j=0}^{N} \alpha_j(z)|j\rangle|z\rangle$. For this preparation to be possible, m has to be chosen such that $N < 2^m$. (b) Apply the u_ϕ gates to effect U_ϕ on the first m qubits

$$U_\phi \sum_{z=0}^{2^d-1} \sum_{j=0}^{N} \alpha_j(z)|j\rangle|z\rangle = \sum_{z=0}^{2^d-1} \sum_{j=0}^{N} \alpha_j(z) e^{ij\phi}|j\rangle|z\rangle. \tag{10}$$

(c) Apply the inverse quantum Fourier transform [19] $|j\rangle \mapsto (1/\sqrt{M}) \sum_{y=0}^{M-1} e^{-i(2\pi yj/M)}|y\rangle$, to obtain

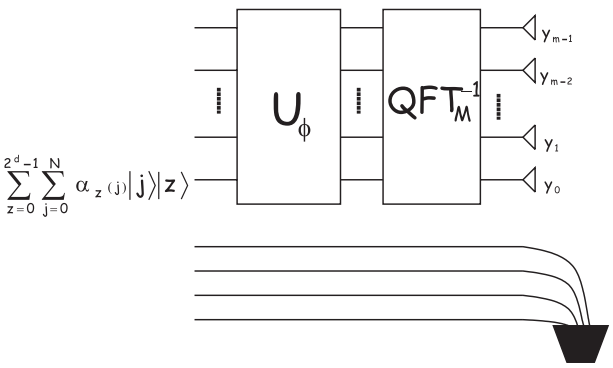

FIG. 3. Without loss of generality, we can assume that our phase estimation procedure has the form illustrated in this figure, where $2^m \geq M \geq N + 1$. The top register contains m qubits, and the bottom d qubits are ancilla qubits that may be entangled with the first m qubits but are discarded.

$$\frac{1}{\sqrt{M}} \sum_{z=0}^{2^d-1} \sum_{y=0}^{M-1} \left(\sum_{j=0}^{N} \alpha_j(z) e^{ij(\phi - (2\pi y/M))} \right) |y\rangle|z\rangle, \tag{11}$$

and measure y.

The following two observations lead to further simplifications.

Let us first notice that in this procedure the role of the d auxiliary qubits is restricted to the initial preparation of the most general state of the first m qubits (all subsequent operations are restricted to these m qubits). The expected cost for a mixture cannot be less than all of the individual expected costs for the contributing pure states; hence, either some of the contributing costs are less or they are all equal. In either case, a judicious choice of a pure state of the m qubits leads to equally good or better phase estimation. This argument implies that, without loss of generality, we can drop the d ancilla qubits, restrict our circuit to only m qubits (plus some ancilla bits that might be used to implement U_ϕ using N copies of u_ϕ), and run the estimation on pure states.

Second, in the description above the quantum Fourier transform is parametrized by M, where $N + 1 \leq M \leq 2^m$, but in fact any $M' \geq N + 1$, for example $M' = 2^m$, will work equally well, while keeping the same $\alpha_j(z)$ values (we can drop the dependence on z, as shown in the previous paragraph) that were defined in Eq. (6) for a specific value of M. To see this, consider a cost function of the form $C(\phi - \tilde{\phi}_y)$. The expected cost, for some M', is

$$\begin{aligned}
\bar{C} &= \frac{1}{2\pi M'} \int_{\phi=0}^{2\pi} d\phi \sum_{y=0}^{M'-1} \left| \sum_{j=0}^{N} \alpha_j e^{ij(\phi - 2\pi y/M')} \right|^2 \\
&\quad \times C\left(\phi - \frac{2\pi y}{M'}\right) \\
&= \frac{1}{2\pi M'} \sum_{y=0}^{M'-1} \int_{\phi'=0}^{2\pi} d\phi' \left| \sum_{j=0}^{N} \alpha_j e^{ij\phi'} \right|^2 C(\phi') \\
&= \frac{1}{2\pi} \int_{\phi'=0}^{2\pi} d\phi' \left| \sum_{j=0}^{N} \alpha_j e^{ij\phi'} \right|^2 C(\phi'),
\end{aligned}$$

where $\phi' = \phi - \frac{2\pi y}{M'}$. The expected cost does not depend on M'. In other words, for any $M, M' \geq N + 1$, any expected cost achievable with inference rule $y \mapsto \frac{2\pi y}{M}$ is also achievable with inference rule $y \mapsto \frac{2\pi y}{M'}$.

Recall that we mentioned in the introduction that as $M \to \infty$ the difference between the optimal $\bar{C}$ assuming the inference rule $y \mapsto \bar{\phi}_y = \frac{2\pi y}{M}$ and the optimal $\bar{C}$ without such an assumption is $\epsilon_M \to 0$. Since we have just shown that for all $M \geq N + 1$, the expected cost $\bar{C}$ is constant, this means that the difference ϵ_M is in fact 0 for all $M \geq N + 1$. In other words, assuming a inference rule of the form $\tilde{\phi}_y = \frac{2\pi y}{M}$ does not cost us anything as long as we use $M \geq N + 1$.

It is clear that the exact value of the expected cost now depends only on the α_j values, that is, on the initial state, which means that, given a specific cost function, all we have to do is to choose an optimal initial state.

We emphasize that the schemes in Figs. 1 and 2 provide an optimal covariant estimation scheme even for general cost functions not necessarily of the Holevo class.

Let us now address the problem of optimal input states for two different cost functions. First, we look at the minimization of the "1−fidelity" cost function $C_F(\phi, \tilde{\phi}_y) = \sin^2[(\phi - \tilde{\phi}_y)/2]$. The minimum cost is achieved with the initial state

$$|x_N^{\text{optimal}}\rangle = \sum_{j=0}^{N} \sqrt{\frac{2}{N+2}} \sin\left(\frac{(j+1)\pi}{N+2}\right)|j\rangle. \qquad (12)$$

The error in fidelity of this protocol goes to zero according to the square of the number of black boxes used $\bar{C}_F = O(1/N^2)$. It is interesting to note that the fidelity of the more conventional approach to phase-rotation estimation with the uniform initial state ($\alpha_j = 1/\sqrt{N+1}$ for all j) only tends to zero *linearly* in N. That is, $\bar{C}_F = \Omega(1/N)$.

Another cost function that is commonly used is the window function that allows any error smaller than δ: $C_W^\delta(\phi, \tilde{\phi}) = 0$ if $|\phi - \tilde{\phi}| < \delta$, but $C_W^\delta(\phi, \tilde{\phi}) = 1$ if $|\phi - \tilde{\phi}| \geq \delta$. The minimization of this cost leads to optimal states with amplitudes $\alpha_j = 1/\sqrt{N+1}$, which corresponds to what is effectively used by Shor's algorithm [1,2,20], and provides an expected cost in $O(\frac{1}{\delta N})$.

In this Letter, we have addressed the general problem of finding the optimal estimating procedure for the real parameter ϕ given N copies of the single qubit phase rotation u_ϕ within a general quantum circuit in finite dimensions. We considered the general case where the circuit consists of an arbitrary input state followed by any arrangement of the N phase rotations interspersed with arbitrary quantum operations. The main result was the proof that in all cases, and for any covariant cost function, the optimal phase estimation procedure is equivalent to a quantum Fourier transform in an appropriate basis.

Our result is very general and gives a recipe for finding the best achievable phase estimation for a given cost function. In practice, once we know the minimum cost possible, we can also search for and use easier-to-implement phase estimation procedures that achieve the same, or similar, expected cost. Because of the generality of our main result, it will surely find many other interesting applications in physical and computational scenarios.

This is an application of the polynomial method to "black boxes" encoding continuous variables, in this case, one real parameter. The method can also be applied to several real parameters, as well as combinations of continuous and discrete parameters.

M. M. is supported by DTO-ARO, NSERC, CFI, ORDCF, CIAR, CRC, ORF, and Ontario-MRI. This work was supported in part by MIUR through PRIN 2005 and by the EC through the project SECOQC.

[1] A. Y. Kitaev, Russ. Math. Surv. **52**, 1191 (1997).
[2] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proc. R. Soc. A **454**, 339 (1998).
[3] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India, 1984* (IEEE, New York, 1984), pp. 175–179.
[4] See, for example, D. J. Wineland *et al.*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **37**, 515 (1990).
[5] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[6] A. S. Holevo, Rep. Math. Phys. **13**, 379 (1978).
[7] R. Derka, V. Buzek, and A. Ekert, Phys. Rev. Lett. **80**, 1571 (1998).
[8] W. van Dam, G. M. D'Ariano, A. Ekert, C. Macchiavello, and M. Mosca (to be published).
[9] E. Knill, G. Ortiz, and R. Somma, Phys. Rev. A **75**, 012328 (2007).
[10] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[11] A uniform prior is also relevant in other scenarios, for example, if we are working in an adversarial scenario where Alice fixes her approximation scheme and then Bob (the adversary) picks the phase in order maximize the expected cost of Alice's estimate. Regardless of Bob's strategy, Alice can "uniformize" it by adding a uniform (or arbitrarily close to uniform) random phase shift to whatever phase shift gate is provided by Bob. This means that any adversary is no more powerful than an adversary that picks a phase uniformly at random.
[12] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
[13] P. Kaye and M. Mosca, in Proceedings of the International Conference on Quantum Information (Optical Society of America, Washington, D.C., 2002), OSA CD-ROM, PB28.
[14] M. Mosca and C. Zalka, Int. J. Quantum. Inform. **2**, 91 (2004).
[15] L. Hales and S. Hallgren, in *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* (IEEE, Los Alamitos, CA, 2000), pp. 515–525.
[16] Any scheme with $0 < M \leq N$ can be turned into an equivalent one with $M \geq N + 1$, e.g., by embedding the output into Hilbert space that is k times larger (for sufficiently large k) and outputting ky instead of y.
[17] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf, J. ACM **48**, 778 (2001).
[18] This involves adding a random phase parameter ϕ_r selected uniformly at random from $\{0, 2\pi\frac{1}{M}, 2\pi\frac{2}{M}, \ldots, 2\pi\frac{M-1}{M}\}$ to the phase to be estimated. This can be done by adding a u_{ϕ_r} gate after each unknown u_ϕ. Then apply the optimal estimation procedure and subtract ϕ_r from the estimate.
[19] This is where our technique requires $M \geq N + 1$.
[20] P. Shor, SIAM J. Comput. **26**, 1484 (1997).