

# Optimal phase estimation in quantum networks

Wim van Dam<sup>1</sup>, G Mauro D'Ariano<sup>2</sup>, Artur Ekert<sup>3,4</sup>,  
Chiara Macchiavello<sup>2</sup> and Michele Mosca<sup>5,6,7</sup>

<sup>1</sup> Departments of Computer Science and Physics, University of California, Santa Barbara, Santa Barbara, CA 93106-5110, USA

<sup>2</sup> Dipartimento di Fisica 'A. Volta' and CNISM, via Bassi 6, 27100 Pavia, Italy

<sup>3</sup> Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford, OX1 3LB, UK

<sup>4</sup> Department of Physics, National University of Singapore, 2 Science Drive 3, 117542, Singapore

<sup>5</sup> Institute for Quantum Computing, University of Waterloo, N2L 3G1, Waterloo, ON, Canada

<sup>6</sup> St Jerome's University, N2L 3G3, Waterloo, ON, Canada

<sup>7</sup> Perimeter Institute for Theoretical Physics, N2L 2Y5, Waterloo, ON, Canada

Received 30 January 2007, in final form 11 April 2007

Published 27 June 2007

Online at [stacks.iop.org/JPhysA/40/7971](http://stacks.iop.org/JPhysA/40/7971)

## Abstract

We address the problem of estimating the phase  $\phi$  given  $N$  copies of the phase rotation  $u_\phi$  within an array of quantum operations in finite dimensions. We first consider the special case where the array consists of an arbitrary input state followed by any arrangement of the  $N$  phase rotations, and ending with a POVM. We optimize the POVM for a given input state and fixed arrangement. Then we also optimize the input state for some specific cost functions. In all cases, the optimal POVM is equivalent to a quantum Fourier transform in an appropriate basis. Examples and applications are given.

PACS numbers: 03.67.-a, 03.65.-w

## 1. Introduction

Extracting information that is encoded in the relative phase of quantum systems is exploited in several quantum information processing tasks and applications. For example, most existing quantum algorithms with super-polynomial speed-up involve phase estimation [3] and more recently it was shown that one natural formulation of the phase estimation problem is BQP-complete [6]. Moreover, information is encoded into phase properties in some quantum cryptographic protocols [4], and in schemes on which atomic clocks are based [5]. Therefore, the issue of estimating the phase in the most efficient way is of fundamental importance.

In this paper we address the problem of optimally estimating the phase  $\phi$  in a quantum network of qubits given  $N$  copies of the phase rotation  $u_\phi = \exp(i|1\rangle\langle 1|\phi)$ , where  $|0\rangle$  and  $|1\rangle$  denote the computational basis of a single qubit.

There is a variety of relevant scenarios that can be considered. For example, the most general scenario, where one can prepare any initial state, and apply an arbitrary quantum circuit

including the  $N$  phase rotation gates, is considered in [1] for arbitrary cost functions, and in [2] for a specific cost function. In [1] it is shown that for any cost function, the general case can be reduced to optimizing the initial state of a procedure with a very specific form. More specific scenarios are relevant, for example, when performing state estimation (e.g. [10]) or phase estimations in other special circumstances, such as in the final step of a dihedral hidden subgroup algorithm [15].

The paper is organized as follows. In section 2 we introduce these specific phase estimation scenarios, where the phase shift gates are applied in a special way. In section 3 we optimize the phase estimation procedure in these cases, by first deriving the optimal POVM for a fixed initial state and cost function, and then by optimizing the average cost also on the form of the initial state. In sections 4, 5 and 6 we give some examples of interest where our results can be applied, namely the case of phase estimation in a system of  $N$  identically prepared qubits, the Shor algorithm and the dihedral hidden subgroup problem. Finally, in section 7 we summarize the main results and discuss their implications. The paper ends with some appendices where some details of the derivations presented in the text are reported.

## 2. Special case

In this section, we will study a special case of the general problem of optimal phase estimation, where we restrict attention to one particular way of applying the phase rotation operator  $u_\phi$  some finite number of times, and we also restrict attention to a special (but widely used) class of ‘cost functions’.

In particular, we consider the task of estimating the unknown phase shift parameter  $\phi \in [0, 2\pi]$  of a unitary transformation  $U_\phi$  acting on  $L$  qubits (with Hilbert space  $\mathcal{H}^{\otimes L}$ ,  $\mathcal{H} \simeq \mathbb{C}^2$ ), with  $U_\phi$  of the form

$$U_\phi = \otimes_{l=1}^L u_\phi^{n_l}, \quad (1)$$

where  $u_\phi$  is the elementary single qubit phase-shift gate defined in the introduction. More explicitly the  $l$ th qubit undergoes the unitary phase shift

$$|0\rangle_l \mapsto |0\rangle_l, \quad (2)$$

$$|1\rangle_l \mapsto e^{in_l\phi}|1\rangle_l, \quad (3)$$

where  $\{|0\rangle_l, |1\rangle_l\}$  is a basis for the  $l$ th qubit and  $n_l$  is an integer number, with the constraint  $\sum_l n_l = N$ .

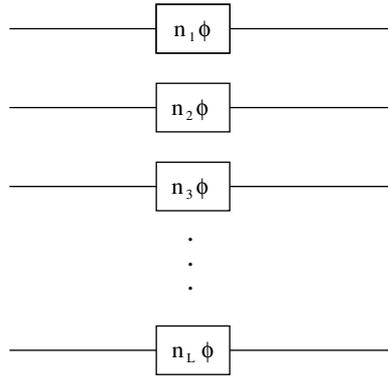
The operator  $U_\phi$ ,  $2\pi$ -periodic with respect to  $\phi$ , operates locally on each of the  $L$  qubits, and in the Schrödinger picture it acts on a known initial state  $|\Psi_0\rangle \in \mathcal{H}^{\otimes L}$ . The general scenario is illustrated in figure 1.

The problem we want to address can be phrased as follows:

*Given the set of numbers  $\{n_l, l = 1, \dots, L\}$ , with  $\sum_l n_l = N$ , we want to find the best measurement procedure to estimate the phase  $\phi$ .*

This estimation problem is posed in the general framework of *quantum estimation theory* [7]. One first defines a *cost function*  $C(\phi_*, \phi)$ , which assesses ‘the cost one has to pay’ for errors in the estimated value  $\phi_*$  when the true one is  $\phi$  (such a function is typically a non-decreasing function of  $|\phi - \phi_*|$  so that we pay more for larger errors in the estimate). For any given cost function  $C(\phi_*, \phi)$  and *a priori* probability distribution  $p_0(\phi)$  for the true value  $\phi$  one then evaluates the average cost in the estimate

$$\bar{C} = \int_0^{2\pi} d\phi p_0(\phi) \int_0^{2\pi} d\phi_* C(\phi_*, \phi) p(\phi_*|\phi), \quad (4)$$



**Figure 1.** Network representation of the action of the unitary operator acting on  $L$  qubits  $U_\phi = \otimes_{l=1}^L u_\phi^{n_l}$ , with  $n_l$  integers. Each box represents a single qubit unitary phase shift operation.

where  $p(\phi_*|\phi)$  is the conditional probability of estimating  $\phi_*$  given the true value  $\phi$ . The conditional probability is given by the Born rule

$$p(\phi_*|\phi) = \text{Tr}[D_{\phi_*} U_\phi |\Psi_0\rangle\langle\Psi_0| U_\phi^\dagger], \tag{5}$$

where  $D_{\phi_*}$  denotes the POVM density [7] of the estimation, with  $D_{\phi_*}$  a positive operator, normalized as

$$\int_0^{2\pi} d\phi_* D_{\phi_*} = I, \tag{6}$$

$I$  denoting the identity operator on  $\mathcal{H}^{\otimes L}$ . The estimation problem consists in finding the optimal POVM which minimizes the average cost  $\bar{C}$  in equation (4).

We assume that  $\phi$  is *a priori* uniformly distributed in  $[0, 2\pi]$  with probability density  $p_0(\phi) = \frac{1}{2\pi}$  (see appendix C for an explanation). Moreover, we want to have no privileged phase value in the estimation, hence the error is paid only as a function of the distance between the true value and the estimated one, so that the cost function is actually an even function of a single variable  $C(\phi_*, \phi) \equiv C(\phi_* - \phi)$ . In the following we will consider some examples of cost functions corresponding to relevant optimization criteria.

For a cost function that depends only on  $\phi_* - \phi$ , one can prove [8] that there exists an optimal POVM which is covariant, namely

$$d\phi_* D_{\phi_*} = U_{\phi_*} \xi U_{\phi_*}^\dagger \frac{d\phi_*}{2\pi}, \tag{7}$$

where  $\xi \equiv 2\pi D_0$  is a positive operator, corresponding to a conditional probability  $p(\phi_*|\phi)$  that will also depend only on  $\phi_* - \phi$ . Therefore, using equation (5) and the invariance of the trace under cyclic permutations one has

$$\begin{aligned} p(\phi_*|\phi) &\equiv p(\phi_* - \phi) = \text{Tr}[U_\phi^\dagger D_{\phi_*} U_\phi |\Psi_0\rangle\langle\Psi_0|] \\ &= \text{Tr}[U_{\phi-\phi_*}^\dagger D_0 U_{\phi-\phi_*} |\Psi_0\rangle\langle\Psi_0|]. \end{aligned} \tag{8}$$

In other words, the POVM is generated from a positive operator  $\xi$  under the action of the unitary Abelian group of the operators  $U_{\phi_*}$ , with  $\phi_* \in [0, 2\pi]$ . Hence the problem is reduced to finding the positive operator  $\xi$  that minimizes the cost  $\bar{C}$  in equation (4).

It is convenient to seek the solution of the optimization problem using the representation of the *shift operator*  $K$  defined as

$$U_\phi = \exp(iK\phi). \tag{9}$$

The operator  $K$  has integer (possibly degenerate) spectrum that is contained in  $\mathbb{Z}_q$  for  $q = N + 1$  (the  $2\pi$ -periodicity of  $U_\phi$  in equation (1) implies integer values for the spectrum of  $K$ ). In the following we will denote by  $P_k$  the projector over the eigenspace of  $K$  corresponding to the eigenvalue  $k \in \mathbb{Z}_q$ . We introduce a (generally incomplete) set of orthonormal states  $\{|k\rangle\}$  uniquely defined as follows. For  $k \in \text{spec } K$ , define  $p_k = \langle \Psi_0 | P_k | \Psi_0 \rangle$  and for those  $k$  such that  $p_k \neq 0$  let

$$|k\rangle = \frac{P_k |\Psi_0\rangle}{\sqrt{p_k}} \quad (10)$$

(recall that the initial state of the  $L$  qubits  $|\Psi_0\rangle$  is known by hypothesis). In the following we will denote by  $\mathbb{S} \subseteq \mathbb{Z}_q$  the set of eigenvalues  $k$  such that  $P_k |\Psi_0\rangle \neq 0$ .

The problem of state estimation is now restricted to the Hilbert space  $\mathcal{K} = \text{Span}\{|k\rangle, k \in \mathbb{S}\}$ , with  $\dim(\mathcal{K}) \leq q$ . Therefore, without loss of generality, the POVM density can be completed in the block diagonal form

$$D_\phi = D_\phi^{\mathcal{K}} \oplus D_\phi^{\mathcal{K}^\perp} \quad (11)$$

on  $\mathcal{H}^{\otimes N} = \mathcal{K} \oplus \mathcal{K}^\perp$ , with  $D_\phi^{\mathcal{K}^\perp}$  any arbitrary POVM density on  $\mathcal{K}^\perp$  (the component of the POVM density  $D_\phi^{\mathcal{K}^\perp}$  acting on the subspace  $\mathcal{K}^\perp$  can be arbitrary because the state  $|\Psi_0\rangle$  has vanishing projection on this subspace and therefore this component does not give any contribution to the average cost).

The above procedure has been designed to deal with the (possible) degeneracy of the operator  $K$ , reducing the problem to the ‘canonical’ phase estimation problem for non-degenerate shift operator  $H$  on the space  $\mathcal{K}$

$$|\Psi_0\rangle \mapsto \exp(iH\phi) |\Psi_0\rangle, \quad H = \sum_{k \in \mathbb{S}} k |k\rangle \langle k| \quad (12)$$

with initial state

$$|\Psi_0\rangle = \sum_{k \in \mathbb{S}} x_k |k\rangle. \quad (13)$$

Note that due to the arbitrariness of the POVM density complement  $D_\phi^{\mathcal{K}^\perp}$  on  $\mathcal{K}^\perp$ , it is always possible to complete  $\mathbb{S}$  to  $\mathbb{Z}_q$ , without changing the optimality of the POVM.

Note that this implies that the optimal POVM depends only on the set  $\mathbb{S}$ , which is contained in the set of integers which can be expressed as  $\sum_i b_i n_i$ , for some  $b_i \in \{0, 1\}$  (also referred to as ‘subset sums’ of the set  $\{n_1, n_2, \dots, n_L\}$ ). The choice of the initial state  $|\Psi_0\rangle$  determines the weight  $x_k$  of each possible integer  $k$  of this form, and we define  $\mathbb{S}$  to be the  $k$  with  $x_k \neq 0$ .

This means that if we were given the freedom to choose the initial state  $|\Psi_0\rangle$ , then all the partitions of  $N$  into parts  $n_1, n_2, \dots, n_L$ , with the same subset sums, lead to the same optimal expected cost for estimating  $\phi$ .

Given full freedom to partition  $N$ , one can always achieve all possible subset sums between 0 and  $N$ . One obvious way to achieve this is to have  $L = N$  and  $n_1 = n_2 = \dots = n_N = 1$ . However, this requires  $N$  qubits. One can also achieve all subset sums with  $O(\log_2 N)$  qubits by letting  $n_i = 2^i$ , for  $0 \leq i < L = \lfloor \log_2 N \rfloor$ , and  $n_L = N - 2^L + 1$ . A simple basis change allows one to assume that  $|k\rangle$  is a tensor product state representing the integer  $k$  expressed in binary (described in appendix A and illustrated in figure A1).

### 3. Optimal phase estimation in the special case

In this section we will review the derivation of the optimal procedure to estimate the phase  $\phi$  in equation (12) [8]. We will first derive in subsection 3.1 the optimal POVM for a fixed

initial state  $|\Psi_0\rangle$  and for a given cost function. As mentioned above, the optimal POVM is the one that minimizes the average cost in equation (4). We will show that a general solution can be found, which depends on the form of the initial state  $|\Psi_0\rangle$  only through the orthonormal set  $\{|k\rangle, k \in \mathbb{S}\}$  and which holds for a large class of cost functions, including, for example, the fidelity and the window function as special cases. Later, in subsection 3.2, we will further optimize the solution by deriving the form of the initial state  $|\Psi_0\rangle$  that leads to the minimum average cost for a chosen cost function and using the optimal POVM. In this way we give a prescription to prepare the initial state of the  $L$  qubits in order to have the most efficient phase estimation for a given network (i.e. for a given set of integer numbers  $\{n_l, l = 1, \dots, L\}$ ) and for a given cost function.

### 3.1. Optimal POVM

In order to find the optimal phase estimation procedure we have to derive the positive operator  $\chi = 2\pi D_0^{\mathcal{K}}$  on the subspace  $\mathcal{K}$ , which minimizes the cost  $\bar{C}$  in equation (4). We will conveniently use the representation of  $\chi$  in the  $|k\rangle$  basis, namely

$$\chi = \sum_{h,k \in \mathbb{S}} |h\rangle\langle k| \chi_{hk}. \tag{14}$$

As a cost function we consider a generic  $2\pi$ -periodic even function, which can be expressed as a Fourier series in the form

$$C(\phi) = - \sum_{l=0}^{\infty} c_l \cos l\phi. \tag{15}$$

The average cost can then be written as

$$\bar{C} = -c_0 - \sum_{l=1}^{\infty} c_l \int \frac{d\phi}{2\pi} \cos l\phi \sum_{hk \in \mathbb{S}} e^{i(h-k)\phi} \chi_{hk} \langle k|\Psi_0\rangle \langle \Psi_0|h\rangle. \tag{16}$$

By calculating the integrals the above expression can be reduced to

$$\bar{C} = -c_0 - \frac{1}{2} \sum_{l=1}^{q-1} c_l \sum_{h,k \in \mathbb{S}, |h-k|=l} \langle \Psi_0|h\rangle \chi_{hk} \langle k|\Psi_0\rangle. \tag{17}$$

Note that the sum over  $l$  is truncated by the vanishing of the  $|\Psi_0\rangle$  components. Positivity of the operator  $\chi$  implies that [8]

$$|\chi_{hk}| \leq \sqrt{\chi_{hh} \chi_{kk}} = 1, \tag{18}$$

where the last equality  $\chi_{hh} = 1 \forall h$  is a consequence of the POVM completeness condition  $\int d\phi D_\phi^{\mathcal{K}} = I_{\mathcal{K}}$  ( $I_{\mathcal{K}}$  is the identity operator on the Hilbert space  $\mathcal{K}$ ). By exploiting equation (18) we can now write the following inequality:

$$\text{sign}(c_l) \sum_{h,k \in \mathbb{S}, |h-k|=l} \langle \Psi_0|h\rangle \chi_{hk} \langle k|\Psi_0\rangle \leq \sum_{h,k \in \mathbb{S}, |h-k|=l} |\langle \Psi_0|h\rangle| |\langle k|\Psi_0\rangle|. \tag{19}$$

The equality in the above equation is achieved only for  $\chi_{hk} = \text{sign}(c_{|h-k|})$  (we recall that  $\langle \Psi_0|k\rangle > 0 \forall k \in \mathbb{S}$  by construction). The average cost is minimized when the equality in equation (19) is satisfied and the minimum value is given by

$$\bar{C} = -c_0 - \frac{1}{2} \sum_{l=1}^{q-1} |c_l| \sum_{h,k \in \mathbb{S}, |h-k|=l} |\langle \Psi_0|h\rangle| |\langle k|\Psi_0\rangle| \tag{20}$$

where we can set  $\text{sign}(0) = 1$ , since the cost  $\bar{C}$  is independent of  $\chi_{hk}$  for  $c_{|h-k|} = 0$ . Note, however, that positivity of the operator  $\chi$  in the form  $\chi_{hk} = \text{sign}(c_{|h-k|})$  is not guaranteed for an arbitrary choice of the coefficients  $c_l$ . For the rest of section 3, we will restrict ourselves to the particular form of coefficients  $c_l \geq 0 \forall l \geq 1$  considered by Holevo [8]. In this case  $\chi$  has all unit elements  $\chi_{hk} = 1$  and is positive. The optimal POVM corresponding to equation (7) can be written as follows:

$$d\phi D_\phi^{\mathcal{K}} \equiv \frac{d\phi}{2\pi} |e^{i\phi}\rangle_{\mathcal{K}\mathcal{K}} \langle e^{i\phi}| \quad (21)$$

where

$$|e^{i\phi}\rangle_{\mathcal{K}} = \sum_{k \in \mathbb{S}} e^{ik\phi} |k\rangle. \quad (22)$$

We want to stress that the Holevo condition on the Fourier coefficients is not very restrictive. In fact, it corresponds to a large class of optimization criteria. It includes, for example, the likelihood criterion for  $C_L(\phi) = -\delta_{2\pi}(\phi)$ , the periodized variance for  $C_V(\phi) = 4 \sin^2(\phi/2)$  and the fidelity optimization  $F(\phi) = |\langle \Psi_0 | U_\phi | \Psi_0 \rangle|^2$  with  $C_F(\phi) = 1 - F(\phi)$  and  $c_l = 2 \sum_{h,k \in \mathbb{S}, |h-k|=l} x_h^2 x_k^2$  where  $x_k = \langle k | \Psi_0 \rangle$ . Note that also a cost function often considered implicitly in computer science, namely the ‘window function’  $C_W(\phi) = 0$  for  $|\phi| \leq \epsilon$ ,  $C_W(\phi) = 1$  for  $|\phi| > \epsilon$ , defined for  $\phi \in [-\pi, +\pi]$ , can be included in this class under some restrictions. Actually, the Fourier coefficients for such a function take the form  $c_0 = \frac{\epsilon}{\pi} - 1$  and  $c_l = \sin(l\epsilon)/l\pi$  for  $l > 0$ , and therefore they are all positive for  $0 < l < q$  when the width of the window function satisfies the condition  $\epsilon \leq \pi/q$ . Since the Fourier coefficients that contribute to the average cost in equation (17) correspond to  $l < q$  and are all positive (apart from  $c_0$  but this does not affect the optimization of the POVM), the optimal POVM for the window function with  $\epsilon \leq \pi/q$  is still given by equation (21). We want to point out that the same POVM (21) would optimize the average cost for Holevo-type cost functions even for a certain class of mixed states<sup>8</sup>.

The optimal POVM in the form (21) does not correspond to a straightforward physical measurement scheme (that is, one with finite resources, including finite resolution) because it gives a continuous value of the phase as the estimated value. We will now find a more convenient discrete description.

We can first complete the set  $\mathbb{S}$  to  $\mathbb{Z}_q$  by appropriate choice of the arbitrary density  $D_\phi^{\mathcal{K}^\perp}$ , extending  $\mathcal{K}$  to the span of those vectors  $|k\rangle$  corresponding to  $P_k |\Psi_0\rangle = 0$ , and restricting  $\mathcal{K}^\perp$  accordingly, without changing the optimality of the POVM. The optimal relevant POVM density is then given by equation (21), with

$$|e^{i\phi}\rangle_{\mathcal{K}} = \sum_{k=0}^{q-1} e^{ik\phi} |k\rangle. \quad (23)$$

Due to the covariance of the optimal POVM and the uniform prior distribution  $p_0(\phi)$ , the average cost  $\bar{C}$  in equation (16) takes exactly the same value if the continuous POVM (21) is restricted to only a set of  $q$  equally spaced values  $\phi_s = \frac{2\pi}{q}s$ ,  $s = 0, 1, \dots, q-1$ . This can be easily proved using the identity

$$\delta_{n0} = \int_0^{2\pi} \frac{d\phi}{2\pi} e^{in\phi} = \frac{1}{q} \sum_{s=0}^{q-1} e^{in\phi_s} \quad (24)$$

<sup>8</sup> This class of mixed states  $\theta$  is characterized by the restriction that the elements of  $\theta$  have fixed phase along the diagonals, i.e.  $\theta_{hk} = |\theta_{hk}| e^{i\phi_{|h-k|}}$ . In this case the optimal  $\chi$  is given by  $\chi_{hk} = e^{-i\phi_{|h-k|}}$ .

with  $n \in \mathbb{Z}_q$ . Therefore, for every covariant POVM there exists always a discrete POVM giving the same average cost. In the present case of Holevo cost the optimal POVM becomes the orthogonal projector-valued measure

$$E_s = |\phi_s\rangle\langle\phi_s| \tag{25}$$

where

$$|\phi_s\rangle = \frac{1}{\sqrt{q}} \sum_{k=0}^{q-1} e^{ik\phi_s} |k\rangle \tag{26}$$

are orthogonal states for  $s = 0, \dots, q - 1$ . Note that from the form of states (26) one can see that this estimation procedure can be implemented by the customary discrete quantum Fourier transform (QFT) network. Exact implementation of the QFT for any positive integer  $q$  was detailed in [12].

### 3.2. Optimal states

We will now perform a further optimization, namely we derive the form of the initial state  $|\Psi_0\rangle$  that leads to the minimum average cost (how to generate such initial states is described in [11]). As we will see soon, this step depends crucially on the form of the cost function, which has to be specified in advance. We first notice that for a covariant POVM the average cost can be recast in the following form:

$$\bar{C} = \text{Tr}[\hat{C}|\Psi_0\rangle\langle\Psi_0|] \tag{27}$$

where the cost operator is defined as

$$\hat{C} = \int_0^{2\pi} d\phi D_\phi^{\mathcal{K}} C(\phi). \tag{28}$$

Without loss of generality we will take  $\mathbb{S} \equiv \mathbb{Z}_q$  in the following. (Note that one can define the operator  $\hat{C}$  with the complete POVM density  $D_\phi$  over  $\mathcal{H} = \mathcal{K} \oplus \mathcal{K}^\perp$ , since the term  $D_\phi^{\mathcal{K}^\perp}$  will not contribute to the trace (27).)

The average cost  $\bar{C}$  can be minimized over the coefficients of  $|\Psi_0\rangle$  by using the Lagrange multipliers method to account for the normalization constraint. One has to minimize the bilinear function

$$\mathcal{L}[\Psi_0] = \langle\Psi_0|\hat{C}|\Psi_0\rangle - \lambda\langle\Psi_0|\Psi_0\rangle \tag{29}$$

which gives the eigenvalue equation

$$\hat{C}|\Psi_0\rangle = \lambda|\Psi_0\rangle \tag{30}$$

where now the Lagrange parameter  $\lambda$  plays the role of an eigenvalue. As mentioned above, we have now to specify the form of the cost function. We will consider two particular cases which correspond to the optimal POVM (21). We first consider the cost function  $C_V(\phi) = 4 \sin^2(\phi/2)$  (more details are given in appendix B). In this case the cost operator in equation (28) takes the form

$$\hat{C}_V = e_+ + e_- - 2 \tag{31}$$

where

$$e_+ = \sum_{k=0}^{q-2} |k+1\rangle\langle k|, \quad e_- = e_+^\dagger. \tag{32}$$

In terms of the coefficients  $x_k = \langle \Psi_0 | k \rangle > 0$  the eigenvalue equation becomes the recurrence relation

$$x_{k+1}(\lambda) + x_{k-1}(\lambda) - (2 + \lambda)x_k(\lambda) = 0 \quad (33)$$

with boundary conditions  $x_{-1}(\lambda) = x_q(\lambda) = 0$ . The solution can be written in terms of the Chebyshev polynomials of the second kind [9], leading to

$$x_j = \sqrt{\frac{2}{q+1}} \sin\left(\frac{j+1}{q+1}\pi\right). \quad (34)$$

The minimum cost in this case is given by

$$\bar{C}_V = \sin^2\left(\frac{\pi}{2(q+1)}\right). \quad (35)$$

As a second example we consider the window function cost  $C_W(\phi)$  defined in the previous subsection with the condition  $\epsilon \leq \pi/q$ , so that the optimal POVM is still given by equation (21). In this case the cost operator takes the form

$$\hat{C}_W = \left(1 - \frac{\epsilon}{\pi}\right) - \frac{1}{2\pi} \sum_{l=1}^{q-1} \frac{\sin l\epsilon}{l} (e_+^l + e_-^l), \quad (36)$$

where

$$e_+^l = \sum_{k=0}^{q-l-1} k + lk, \quad e_-^l = (e_+^l)^\dagger. \quad (37)$$

The eigenvalue equation has the form

$$x_k \left(\lambda + \frac{\epsilon}{\pi}\right) + \frac{1}{2\pi} \sum_{m(\neq k)=0}^{q-1} \frac{\sin(k-m)\epsilon}{k-m} x_m = 0. \quad (38)$$

This has an easy solution for very narrow window functions, such that  $\epsilon \ll 1/q$ . In this case the coefficients  $\sin(m\epsilon)/m$  can be approximated by a constant independent of  $m$  and the solution for the optimal state  $|\Psi_0\rangle$  is simply given by the equally weighted state, i.e.

$$x_k = 1/\sqrt{q}, \quad k = 0, \dots, q-1. \quad (39)$$

The minimum cost in this case takes the form

$$\bar{C}_W = X \left[1 - \frac{\epsilon}{2\pi}(q+1)\right]. \quad (40)$$

#### 4. Example 1: phase estimation of $N$ identically prepared qubits

We specify here the phase estimation problem to a particular case: estimation of the phase shift undergone by  $N$  qubits initially in the same state, namely

$$U_\phi(a|0\rangle + b|1\rangle)^{\otimes N} = (a|0\rangle + b e^{i\phi}|1\rangle)^{\otimes N}, \quad (41)$$

i.e. the unitary operator  $U_\phi$  is given by

$$U_\phi = \otimes_{l=1}^N \exp\left[\frac{i}{2}(\sigma_l^z - 1)\phi\right], \quad (42)$$

where  $\sigma_l^z$  represents the Pauli operator for the  $l$ th qubit. In the network language of figure 1 this case corresponds to  $n_l = 1 \forall l$ . One can see that for both  $a \neq 0$  and  $b \neq 0$  the  $|k\rangle$  basis introduced in subsection 3.1 is given by

$$\begin{aligned} |k\rangle &= e^{i\phi k} |k\rangle_{\text{sym}} \\ &\equiv e^{i\phi k} \binom{N}{k}^{-1/2} \sum_{\{s_i=0,1\}} \delta\left(\sum_i s_i - k\right) \otimes_{i=1}^N |s_i\rangle, \end{aligned} \quad (43)$$

namely it is the basis of the symmetric multiplet with the choice of phases  $\phi_k = (N - k)\arg(a) + k\arg(b)$  (in this case  $q = N + 1$  because the subspace  $\mathcal{K}$  is the symmetric subspace of the  $N$  qubits). The optimal POVM restricted to the symmetrical tensor subspace  $\mathcal{K} = (\mathcal{H}^{\otimes N})_+$  is given in equation (25) for cost functions of the Holevo class.

The optimized state among the ones of the form (41) for the cost function  $C_V(\phi) = 4 \sin^2 \phi/2$ , and also for the fidelity error  $1 - F(\phi)$  with  $F(\phi) = |a|^4 + |b|^4 + 2|a|^2|b|^2 \cos(\phi)$ , is given by  $a = b$  and leads to the minimum cost

$$C_V = \frac{1}{2} - \frac{1}{2^{N+1}} \sum_{k=0}^{N-1} \sqrt{\binom{N}{k} \binom{N}{k+1}}. \tag{44}$$

This result has also been reported in [10].

Note that we can also phrase the problem in a wider sense and ask for the optimal strategy to estimate the phase shift generated by the unitary operator (42) on a generic symmetric state of the  $N$  qubits and optimize the average cost with respect to the initial state, as in subsection 3.2. It is interesting to note that in this case the optimal states for the cost function  $C_V$  and the window cost function, given in equations (34) and (39) respectively, are entangled states of the  $N$  qubits.

Note also that for  $U_\phi$  of the form (42) we can solve an even more general problem, where the initial state  $|\Psi_0\rangle$  does not belong to the symmetric subspace. In this case, in fact, the best POVM is of the form  $Z^\dagger E_s Z$ , where  $|\Psi_0\rangle = Z|\Lambda_0\rangle$  and  $|\Lambda_0\rangle \in (\mathcal{H}^{\otimes N})_+$ , the symmetric subspace of the  $N$  qubits (one can always find a suitable  $Z$  commuting with a  $U_\phi$  of the form in equation (42)).

### 5. Example 2: the Shor algorithm

In this section we will consider another particular case, corresponding to the phase estimation as the final step of the Shor algorithm in the formulation given in [3]. This step was also optimized in [18]. The phase operator  $U_\phi$  in this case has the form

$$U_\phi = \otimes_{l=1}^L u_\phi^{2^{l-1}}. \tag{45}$$

This corresponds to the network representation of figure 1 with  $n_l = 2^{l-1}$ . In this case the problem does not have degeneracy and the subspace  $\mathcal{K}$  is the whole Hilbert space  $\mathcal{H}^{\otimes L}$  of the  $L$  qubits, namely  $q = 2^L$ . The eigenvectors of the operator  $K$  are simply given by the computational basis of the  $L$  qubits

$$|k\rangle = \otimes_{l=1}^L |s_l\rangle \quad k = \sum_{l=1}^L s_l 2^{l-1}. \tag{46}$$

For cost functions of the Holevo class, the optimal POVM in equation (25) here corresponds to the quantum Fourier transform measurement discussed in [3], where a network realization is also given. Here, our general method proves that this measurement procedure is optimal.

Regarding the optimized states in equation (34) for the cost function  $C_V$  one should note that the state is partially entangled. To be useful in practice, we would also need to consider the computational complexity of creating this state for use as a possible initial state for the Shor algorithm.

If the window cost function is considered, in the limit of very small width analysed in the previous section, the optimal state is given by

$$|\Psi_0\rangle = \left[ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \right]^{\otimes L}, \tag{47}$$

namely it is a factorized state of the  $L$  qubits, as the one considered in the scheme of [3].

### 6. Example 3: dihedral hidden subgroup problem

The hidden subgroup problem (HSP) is the problem of finding generators for a subgroup  $K$  of a group  $G$  given a black-box that implements a function  $f : G \rightarrow X$  satisfying  $f(x) = f(y) \Leftrightarrow x - y \in K$ . In other words,  $f$  is constant on cosets of  $K$  and distinct on different cosets.

Assuming a reasonable presentation of the group  $G$ , there is an efficient quantum algorithm for solving the HSP in Abelian groups. This algorithm is a natural generalization of Shor's algorithm [13] and can be cast as a phase estimation problem as outlined in [3] (based on the approach of [14]). There has been limited success in solving the HSP for non-Abelian groups. See [19] for recent results and references. Ettinger and Høyer [15] reduced the HSP for the dihedral group to the following phase estimation problem:

Given a polynomial (in  $n$ ) number of qubits of the form  $|0\rangle + e^{ik_j\phi}|1\rangle$  and the integers  $k_j$ , where the  $k_j$  are selected uniformly at random from  $\{0, 1, 2, \dots, 2^n - 1\}$ , estimate  $\phi$  with error at most  $\frac{1}{2^n}$ .

Ettinger and Høyer showed that an optimal measurement would solve the HSP with high probability.

Our results from section 2 tell us the optimal POVM. Let  $S = \sum_{j=1}^n k_j$ . Let  $S_j$  equal the set of solutions  $(b_1, b_2, \dots, b_n) \in \{0, 1\}^n$  to the equation  $\sum_i b_i k_i = j$ , and let  $n_j = |S_j|$  equal the number of such solutions. For  $j$  with  $n_j \neq 0$ , let  $|S_j\rangle = \frac{1}{\sqrt{n_j}} \sum_{b_1 b_2 \dots b_n | \sum_i b_i k_i = j} |b_1 b_2 \dots b_n\rangle$ .

Note that

$$(|0\rangle + e^{ik_1\phi}|1\rangle) \otimes (|0\rangle + e^{ik_2\phi}|1\rangle) \otimes \dots \otimes (|0\rangle + e^{ik_n\phi}|1\rangle) = \sum_{j=0}^S \frac{\sqrt{n_j}}{2^n} |S_j\rangle.$$

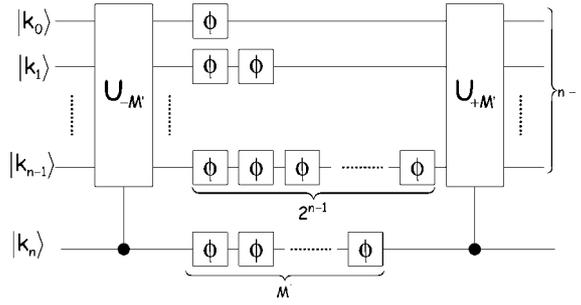
Thus the optimal POVM could be achieved by first performing a unitary basis change  $U$  that maps  $|S_k\rangle \mapsto |k\rangle$ , where the integer  $k$  is represented as a binary string in the computational basis, followed by a quantum Fourier transform in the computational basis, and a measurement in the computational basis.

However, implementing  $U$  would solve the subset sum problem, which is known to be NP-complete (since  $U^{-1}$  maps  $|k\rangle$  to a uniform superposition of strings representing subsets with sum  $k$ ). However, such a measurement is optimal, and a sufficiently good approximation of this measurement would suffice. Regev [16] showed that it suffices to be able to solve the subset sum problem on average in order to find a sufficiently precise estimate of  $\phi$  that allows one to solve the hidden subgroup problem for the dihedral group.

### 7. Discussion

In this paper we have addressed the problem of finding the optimal estimating procedure for the real parameter  $\phi$  given  $N$  copies of the single qubit phase rotation  $u_\phi$  within a quantum network in finite dimensions. We have derived the optimal measurement procedure in the special case where the network consists of an arbitrary input state followed by any arrangement of the  $N$  phase rotations, and found also the optimal states corresponding to some cost functions of interest.

This result is general and can be applied to many cases of interest. In particular, we have considered the phase estimation problem as the final step of the Shor algorithm in the formulation given in [3] and we have proved that the quantum Fourier transform performed in that case is indeed the optimal phase estimation procedure. As another example, we have also shown how our result can be applied to the dihedral hidden subgroup problem.



**Figure A1.** A quantum network for implementing the phase shift operator  $|k\rangle \mapsto e^{i\phi k}|k\rangle$  for  $k \in \{0, 1, \dots, M\}$  represented in binary (that is  $k = k_0 + 2k_1 + \dots + k_n 2^n$ ) and  $M < 2^{n+1}$ . We let  $M' = 2^n - 1 - M$ , and we let  $U_x$  denote the operator that maps  $|j\rangle \mapsto |j + x \bmod 2^n\rangle$ .

**Acknowledgments**

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

**Appendix A. Phase shift circuit**

The network illustrated in figure A1 will achieve the phase shift operator  $|k\rangle \mapsto e^{i\phi k}|k\rangle$  where  $k$  is represented in binary as the string of bits  $k_0 k_1 \dots k_{n-1} k_n$ , where  $k = k_0 + 2k_1 + \dots + 2^{n-1} k_{n-1} + 2^n k_n$ .

**Appendix B. Optimimization of fidelity error**

In this appendix we derive the minimum obtainable value of the fidelity error, given by

$$\bar{C}_F = \frac{1}{2\pi} \sum_{y=0}^{M-1} \int_{\phi=0}^{2\pi} \Pr(y|\phi) \cdot C_F(\phi, \tilde{\phi}_y) d\phi \tag{B.1}$$

$$= \frac{1}{2} - \frac{1}{2} \operatorname{Re} \left( \sum_{j=0}^{N-1} \alpha_j \alpha_{j+1}^* \right). \tag{B.2}$$

We start by decomposing the complex coefficients  $\alpha_j$  into the two real parameters  $p, q$  according to  $\alpha_j = p_j + q_j \sqrt{-1}$  for every  $0 \leq j \leq N$ . This means that we have to minimize the expression

$$1 - 2\bar{C}_F = \sum_{j=0}^{N-1} p_j p_{j+1} + \sum_{j=0}^{N-1} q_j q_{j+1}, \tag{B.3}$$

under the normalization condition for the reals  $p_j$  and  $q_j$  that

$$\sum_{j=0}^N p_j^2 + \sum_{j=0}^N q_j^2 = 1. \tag{B.4}$$

First we will optimize the  $p$  coefficients for the general case  $0 \leq \sum p_j^2 = \mu \leq 1$ . We do this by the introduction of a *Lagrange multiplier*  $\lambda$  and two additional values  $p_{-1} = p_{N+1} = 0$ . The method of Lagrange multipliers requires that the partial derivatives  $\partial/\partial p_t$  of the function  $\sum p_j p_{j+1} + \lambda \sum p_j^2$  have to be zero. This leads to the set of equations for all  $0 \leq t \leq N$ ,

$$\frac{\partial}{\partial p_t} \left( \sum_{j=-1}^{N+1} p_j p_{j+1} + \lambda \sum_{j=-1}^{N+1} p_j^2 \right) = 0, \quad (\text{B.5})$$

and hence the recurrence relation,

$$p_{t+1} = -2\lambda p_t - p_{t-1}. \quad (\text{B.6})$$

This relation can be solved with the help of the Chebyshev polynomials of the second kind (see [9] chapter 8.94, p 1032), using the identification  $U_t(-\lambda) = p_t$ . When taking into account the restrictions  $p_{-1} = p_{N+1} = 0$  and  $\sum p_j^2 = \mu$ , the possible solutions are of the form

$$p_t = \pm \sqrt{\frac{2\mu}{N+2}} \sin\left(\frac{(t+1)k\pi}{N+2}\right) \quad (\text{B.7})$$

for  $k \in \{1, 2, \dots, N+1\}$ . For each of the possible values of  $k$ , the summation that we want to maximize equals (using [9] chapter 1.361, p 32):

$$\sum_{t=0}^{N-1} p_t p_{t+1} = \mu \cos\left(\frac{k\pi}{N+2}\right). \quad (\text{B.8})$$

The absolute maximum hence occurs when  $k = 1$  and equals  $\mu \cos(\pi/(N+2))$ . By the normalization condition of Equation (B.4) we have for the  $q$  values that  $\sum q_j^2 = 1 - \mu$ . Hence, the absolute maximum of the  $\sum q_j q_{j+1}$  summation equals  $(1 - \mu) \cos(\pi/(N+2))$ . The overall result is therefore the following.

The maximum  $1 - 2\bar{C}_F$  as in equation (B.3) is obtained for the  $\alpha$  values

$$\alpha_j = e^{i\psi} \sqrt{\frac{2}{N+2}} \sin\left(\frac{(j+1)\pi}{N+2}\right) \quad (\text{B.9})$$

for any (non-relevant) general phase factor  $e^{i\psi}$ . The investigated minimum then equals

$$\bar{C}_F = \sin^2\left(\frac{\pi}{2N+4}\right), \quad (\text{B.10})$$

which gives  $\bar{C}_F = O(1/N^2)$  as  $N$  goes to infinity.

### Appendix C. Why a uniform prior?

In this paper, we assign a uniform *a priori* probability for the value of  $\phi$ . This is important for the results in this section. There are very natural situations in which one should assume a uniform prior distribution, such as the case when there is no prior knowledge of the value of  $\phi$ . Another scenario (which is typical in computer science) where one is naturally led to consider the uniform prior is in an *adversarial* scenario, where one party, Alice, is doing the phase estimation, and the adversary, Bob, picks the phase. Alice first describes the estimation procedure, and then Bob picks the phase  $\phi$  according to any probability distribution he wishes, with the intention of maximizing the expected cost for Alice.

Alice must optimize her procedure to work well for any probability distribution for  $\phi$ , since Bob will naturally pick the worst case distribution for  $\phi$ . Alice can foil Bob's intentions by 'uniformizing' the distribution of the  $\phi$ . That is, she can guarantee that regardless of the

distribution for  $\phi$ , her procedure will work as well as it does for the uniform distribution for  $\phi$ . That is, she can guarantee that the expected cost for any distribution will equal the expected cost for the uniform prior distribution.

Alice first takes any procedure  $A$  that is optimal given a uniform prior distribution for  $\phi$ . She then augments this procedure in the following way. She pick a random phase  $\phi_r$  with uniform probability over  $[0, 2\pi)$ . She then runs the optimal procedure for estimating  $\phi$  but replaces each instance of  $u_\phi$  with  $u_{\phi+\phi_r}$ . This is easily done by adding a  $u_{\phi_r}$  gate after every instance of a  $u_\phi$  gate. The procedure  $A$  will output an estimate  $\widetilde{u_{\phi+\phi_r}}$  for  $u_{\phi+\phi_r}$ . Alice outputs  $\widetilde{u_\phi} = \widetilde{u_{\phi+\phi_r}} - \phi_r$  as her estimate of  $\phi$ .

This uniformization guarantees that regardless of the distribution for  $\phi$  that Bob chooses, Alice's new procedure performs with the same expected cost as the optimal procedure would with a uniform prior distribution.

However, if Alice is restricted to using finite means, she may not be able to actually sample a uniform prior on the continuous set  $[0, 2\pi)$ . However, she can sample an arbitrarily fine discrete subset of these points, which should be enough for dealing with non-pathological cost functions  $C$ . Note that regardless of how fine a mesh of points Alice samples from to ultimately produce estimates  $\phi^*$ , the values of  $\phi^*$  will come from a finite set. Thus it is possible for an adversary to restrict to choices of  $\phi$  so that although the sets of numbers  $\{\phi - \phi^*\}$  are arbitrarily close for different  $\phi$ , they are still disjoint for different  $\phi$ . This allows for the construction of pathological cost functions for which even the slightest round-off in the outputs  $\phi^*$  can drastically change the expected cost. Thus we need to add further practical assumptions for  $C$  in the case that we do not wish to assume an *a priori* uniform distribution for  $\phi$ , but wish to justify a uniform prior distribution using an adversarial scenario. Essentially, it suffices that most of the time the function does not change very fast. More precisely, it suffices for example that  $|C(\phi)| < B < \infty$ , and that for any  $\delta > 0$  there exists an integer  $N_\delta$  such that cost function does not vary by more than  $\delta$  over intervals of width less than  $1/N_\delta$ , except for a set of points of measure at most  $\delta/B$ . For example, any uniformly continuous cost function satisfies this. Any bounded function that increases monotonically as  $\phi$  tends away from 0 also satisfies this.

Knowing the details of the specific procedure she wishes to uniformize, and knowing the cost function  $C$ , which we assume is reasonably well-behaved as discussed above, and given any  $\epsilon > 0$ , she can pick a large integer  $D_\epsilon$  so that uniformly sampling the phases  $\{2\pi x/D_\epsilon | x = 0, 1, \dots, D_\epsilon - 1\}$  will yield a procedure with expected cost within  $\epsilon$  of the optimal expected cost in the case of a perfectly uniform prior distribution.

The above discussion is only meant to justify that a uniform prior distribution is a meaningful assumption to make. In this paper, we will simply assume that the prior distribution for the  $\phi$  is perfectly uniform in the interval  $[0, 2\pi)$ .

## References

- [1] van Dam W, D'Ariano G M, Ekert A, Macchiavello C and Mosca M 2007 Optimal quantum circuits for general phase estimation *Phys. Rev. Lett.* **98** (9) 090501
- [2] Giovannetti V, Lloyd S and Maccone L 2006 Quantum metrology *Phys. Rev. Lett.* **96** 010401
- [3] Cleve R, Ekert A, Macchiavello C and Mosca M 1998 Quantum algorithms revisited *Proc. R. Soc. London, Ser. A* **454** 339
- [4] Bennett C H and Brassard G 1984 *Proc. IEEE Int. Conf. on Computers, Systems, and Signal Processing (Bangalore, India)* pp 175–9
- [5] See for example Wineland D J *et al* 1990 Progress at NIST toward absolute frequency standards using stored ions *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **A 37** 515
- [6] Wocjan P and Zhang S 2006 *Preprint* [quant-ph/0606179](https://arxiv.org/abs/quant-ph/0606179)
- [7] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)

- [8] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
- [9] Gradshteyn I S and Ryzhik I M 1965 *Table of Integrals, Series and Products* 4th edn (San Diego: Academic)
- [10] Derka R, Buzek V and Ekert A 1998 Universal algorithm for optimal estimation of quantum states from finite ensembles via realizable generalized measurement *Phys. Rev. Lett.* **80** 1571
- [11] Kaye P and Mosca M 2002 Quantum networks for generating arbitrary quantum states *Proc. Int. Conf. on Quantum Information, OSA CD-ROM (Optical Society of America, Washington, DC)* PB28
- [12] Mosca M and Zalka C 2004 Exact quantum Fourier transforms and discrete logarithm algorithms *Int. J. Quantum Inform.* **2** 91–100
- [13] Shor P 1997 Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer *SIAM J. Comput.* **26** 1484–509
- [14] Kitaev A Y 1998 Quantum computations: algorithms and error correction *Russ. Math. Surv.* **52** 1191–249
- [15] Ettinger M and Høyer P 2000 On quantum algorithms for noncommutative hidden subgroups *Adv. Appl. Math.* **25** 239–51
- [16] Regev O 2004 Quantum computation and lattice problems *SIAM J. Comput.* **33** 738–60
- [17] Holevo A S 1978 Estimation of shift parameters of a quantum state *Rep. Math. Phys.* **13**
- [18] Bessen A 2005 Lower bound for quantum phase estimation *Phys. Rev. A* **71** 042313
- [19] Friedl K, Ivanyos G, Magniez F, Santha M and Sen P 2003 Hidden translation and orbit coset in quantum computing *Proc. 35th Annual ACM Symp. on Theory of Computing: STOC'03* pp 1–9