Optimal phase estimation and square root measurement

2001 J. Phys. A: Math. Gen. 34 7017

(http://iopscience.iop.org/0305-4470/34/35/327)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **34** (2001) 7017–7027 PII: S0305-4470(01)19131-6

Optimal phase estimation and square root measurement

Masahide Sasaki[1], **Alberto Carlini**[1] **and Anthony Chefles**[2]

[1] Communications Research Laboratory, Ministry of Posts and Telecommunications Koganei, Tokyo 184-8795, Japan
[2] Department of Physical Sciences, University of Hertfordshire, Hatfield AL10 9AB, Herts, UK

E-mail: psasaki@crl.go.jp

Received 15 November 2000
Published 24 August 2001
Online at stacks.iop.org/JPhysA/34/7017

Abstract
We present an optimal strategy having finite outcomes for estimating a single parameter of the displacement operator on an arbitrary finite-dimensional system using a finite number of identical samples. Assuming the uniform *a priori* distribution for the displacement parameter, an optimal strategy can be constructed by making the *square root measurement* based on uniformly distributed sample points. This type of measurement automatically ensures the global maximality of the figure of merit, that is, the so-called average score or fidelity. Quantum circuit implementations for the optimal strategies are provided in the case of a two-dimensional system.

PACS numbers: 03.67.−a, 03.65.Bz, 89.70.+c

1. Introduction

Quantum state estimation deals with how to estimate unknown parameters of a quantum state as precisely as possible. This problem was studied extensively in the 1970s in the context of the formalism of the probability operator measure (POM) and of quantum Bayesian inference. Basic formulations were already established and optimal strategies were found in various cases. The relevant works are reviewed in [1, 2].

In the 1990s, the same problem was revisited in a new context where one is allowed to use finite samples of a quantum system to be estimated, while most of earlier works were concerned with the estimation using a single sample. Massar and Popescu [3] obtained the optimal strategy for estimating a quantum pure state of a spin-1/2 system, say, $|\rho\rangle$, from N identically prepared samples $|\rho\rangle^{\otimes N}$. Their strategy is based on the use of an infinite continuous set of projectors in the Hilbert space of $|\rho\rangle^{\otimes N}$. For a given unknown quantum system, the optimal estimation strategy is not unique. As shown in [2], one can always find an optimal strategy consisting of an infinite continuous set of POMs. On the other hand, for

0305-4470/01/357017+11$30.00

a finite-dimensional system there must exist a discrete and finite POM achieving the same optimal bound as shown by Derka *et al* [4]. From the point of view of the physical realization the latter is preferable, while the former might be a more useful mathematical tool to derive the maximum attainable average fidelity (which is a commonly used figure of merit for state estimation). In [4] an algorithm is described for constructing such optimal and finite POMs for an arbitrary finite-dimensional system in a pure state. Latorre *et al* [5] then studied the optimal strategy with a minimum number of outcomes for a spin-$1/2$ system, and showed explicit forms of optimal minimal measurements for $N = 1$–5. Their analysis was extended to the cases of a mixed state of a spin-$1/2$ system [6] and of an arbitrary spin system in a pure state [7]. In [6], the closed form expressions for the maximum average score, the optimal minimal POM and its number of outcomes were derived by using the symmetric properties of the totally symmetric subspace supported by a tensor product of N identical samples. To construct the optimal minimal strategy explicitly, however, some parameters are to be determined and remain unsolved for larger N ($\geqslant 6$). For higher dimensional systems it becomes more difficult to find concrete forms for the optimal minimal strategy [7]. As for the maximum average fidelity the explicit expression for an arbitrary N was obtained in [8]. This bound was derived by using the fact that quantum optimal state estimation using N samples can be viewed as the limiting case $M \to \infty$ of universal optimal cloning generating M copies from N inputs for which the maximum average fidelity was given by Werner [9] [3].

Although substantial progress has been obtained in the quantum state estimation, it is still a difficult and open problem how to find explicit and physically realizable solutions for optimal strategies in an algorithmic way, especially in the case of higher dimensional systems and larger numbers of samples. Moreover, discussions given so far in the literature for ensuring the optimality of discrete and finite POMs were focused only on the condition for extremality and not on the full conditions for the existence of a global maximum, which are reviewed in [1, 2]. In general, seeking all extrema and picking up the point corresponding to the global maximum is not necessarily a trivial task for complex systems.

In this paper, we focus on a single-parameter estimation of an arbitrary finite-dimensional system in a pure state and give finite element optimal strategies that can be constructed in a straightforward way and ensure the global maximality conditions for the POM.

2. Optimal phase estimation

Consider a finite-dimensional system described in a Hilbert space $\mathcal{H}$ and let $\{|0\rangle, |1\rangle, \ldots, |K\rangle\}$ be its basis built from the eigenstates of the observable $\hat{O}$ on $\mathcal{H}$, $\hat{O}|k\rangle = k|k\rangle$. Such a system may be, e.g., an optical field produced by quantum scissors [10] (with $\hat{O}$ the photon number operator) or a spin j system described by a superposition of the eigenstates $\{|j, m\rangle, m = -j, \ldots, j\}$ (with $\hat{O}$ the spin operator). The problem we consider is the estimation of a unitary evolution $\hat{u}$ specified by a displacement parameter θ, that is, $\hat{u}(\theta) = \mathrm{e}^{-\mathrm{i}\theta\hat{O}}$. We suppose that the initial state of the system is known *a priori* and reads

$$|\psi(0)\rangle = \sum_{k=0}^{K} c_k\,|k\rangle \tag{2.1}$$

[3] The optimal cloning map of [9] may have a connection with the infinite continuous version of the optimal state estimation strategy.

(with c_k non-zero arbitrary complex coefficients), but we do not have any *a priori* knowledge about θ. After the evolution, the system will be in a state

$$|\psi(\theta)\rangle = \hat{u}(\theta)\,|\psi(0)\rangle = \sum_{k=0}^{K} c_k \mathrm{e}^{-\mathrm{i}\theta k}\,|k\rangle\,. \tag{2.2}$$

It is assumed that N identical samples of the system are available. The combined system is then described on the *totally symmetric bosonic subspace* of $\mathcal{H}^{\otimes N}$ [4, 9] as

$$|\Psi(\theta)\rangle = |\psi(\theta)\rangle^{\otimes N} = {\sum_{\vec{n}}}' C(\vec{n})\mathrm{e}^{-\mathrm{i}\theta\sum_k k n_k}\,|\vec{n}\rangle \tag{2.3}$$

where

$$C(\vec{n}) \equiv \sqrt{N!}\prod_{k=0}^{K}\frac{c_k^{n_k}}{\sqrt{n_k!}} \tag{2.4}$$

$\sum_{\vec{n}}'$ means the summation over $(K+1)$-tuples $\vec{n} \equiv (n_0, \ldots, n_K)$ with $\sum_{k=0}^{K} n_k = N$, and $|\vec{n}\rangle$ $(\equiv|n_0, \ldots, n_K\rangle)$ is the occupation number basis. The dimensionality D_{B} of this space is $D_{\mathrm{B}} = \binom{N+K}{K}$. For our present purpose, however, it is enough to consider the smaller subspace spanned by the eigenstates of the compound operator $\hat{O}_{\mathrm{T}} \equiv \sum_{i=1}^{N}\hat{O}(i)$, where $\hat{O}(i)$ is the observable for the ith sample. Let $\{\vec{n}_i^{(J)}\}$ be the set of $(K+1)$-tuples that satisfy $\sum_{k=0}^{K} k n_k = J$ and define $A_J \equiv \sqrt{\sum_i |C(\vec{n}_i^{(J)})|^2}$. Then the state in equation (2.3) can be rewritten as

$$|\Psi(\theta)\rangle = \sum_{J=0}^{KN} A_J \mathrm{e}^{-\mathrm{i}\theta J}\,|J\rangle \tag{2.5}$$

where

$$|J\rangle = A_J^{-1}\sum_i C\left(\vec{n}_i^{(J)}\right)\left|n_i^{(J)}\right\rangle \qquad \hat{O}_{\mathrm{T}}\,|J\rangle = J\,|J\rangle\,. \tag{2.6}$$

The POM describing the optimal estimation strategy is constructed in the $D_{\mathrm{T}} = (KN+1)$-dimensional subspace $\mathcal{H}_{\mathrm{T}}$ spanned by the set $\{|J\rangle\}$.

Such a POM $\{\hat{\mu}_m\}$ should maximize the following score:

$$\bar{S}(N) = \sum_m \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\,\theta\,\mathrm{Tr}\,(\hat{\mu}_m\hat{\Psi}(\theta))|\langle\psi_m|\psi(\theta)\rangle|^2 \tag{2.7}$$

where $\hat{\Psi}(\theta) = |\Psi(\theta)\rangle\,\langle\psi(\theta)|$ and $|\psi_m\rangle$ is a guessed state according to the mth outcome of the measurement. Optimality can be discussed along with the conditions for quantum Bayesian optimization [1, 2]. Let us introduce the *score operators*

$$\hat{W}_m \equiv \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\,\theta\,\hat{\Psi}(\theta)|\langle\psi_m|\psi(\theta)\rangle|^2. \tag{2.8}$$

They include all the *a priori* information. Then the necessary and sufficient conditions such that a certain set $\{\hat{\mu}_m\}$ globally maximizes the average score for a fixed set of $\{|\psi_m\rangle\}$ are expressed as [1, 2]

$$\begin{aligned} &\text{(i)} \quad \hat{\Gamma} \equiv \textstyle\sum_m \hat{W}_m\hat{\mu}_m \text{ is Hermitian and } (\hat{\Gamma} - \hat{W}_m)\hat{\mu}_m = 0 \quad \forall m \\ &\text{(ii)} \quad \hat{\Gamma} - \hat{W}_m \geqslant 0 \quad \forall m \end{aligned} \tag{2.9}$$

where $\hat{\Gamma}$ is called the Lagrange operator, and the average score (2.7) can then be rewritten as

$$\bar{S}(N) = \mathrm{Tr}\,\hat{\Gamma}. \tag{2.10}$$

The optimal estimation strategy can be constructed in the following way. First take M states corresponding to uniformly distributed sample points in $\theta \in [0, 2\pi)$; that is,

$$|\psi_m\rangle = \sum_{k=0}^{K} c_k \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{M}k} |k\rangle \qquad (m = 0, \ldots, M-1) \tag{2.11}$$

and let us denote its N tensor product states as $|\Psi_m\rangle = |\psi_m\rangle^{\otimes N}$. Then consider the state vector

$$|\mu_m\rangle \equiv \hat{\Psi}^{-\frac{1}{2}} |\Psi_m\rangle \qquad \hat{\Psi} \equiv \sum_m |\Psi_m\rangle \langle\Psi_m|. \tag{2.12}$$

The set $\{\hat{\mu}_m \equiv |\mu_m\rangle \langle\mu_m|\}$ is easily seen to be a set of non-negative Hermitian operators satisfying the resolution of the identity on $\mathcal{H}_T$, and thus a POM. This POM is often called the *square root measurement* [11–13]. If we take $M \geqslant KN + 1$ sample points, $\{\hat{\mu}_m\}$ also works as the optimal estimation strategy. Under the condition $M \geqslant KN + 1$ we have, in fact, that (cf. equation (2.5) with $\theta = 2\pi m/M$)

$$\hat{\Psi} = M \sum_{J=0}^{KN} A_J^2 |J\rangle \langle J| \tag{2.13}$$

because

$$\sum_{m=0}^{M-1} \mathrm{e}^{\mathrm{i}\frac{2\pi m}{M}n} = M\delta_{n,0} \qquad \text{for} \quad -KN \leqslant n \leqslant KN. \tag{2.14}$$

Therefore, from equation (2.12) we get

$$|\mu_m\rangle = \frac{1}{\sqrt{M}} \sum_{J=0}^{KN} \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{M}J} |J\rangle . \tag{2.15}$$

To prove optimality, let us first rewrite the score function as (see equations (2.2) and (2.11))

$$|\langle\psi_m|\psi(\theta)\rangle|^2 = d_0 + \sum_{L=1}^{K} d_L \left(\mathrm{e}^{\mathrm{i}(\frac{2\pi m}{M}-\theta)L} + \mathrm{e}^{-\mathrm{i}(\frac{2\pi m}{M}-\theta)L}\right) \tag{2.16}$$

where

$$d_L = \sum_{k=0}^{K-L} |c_{k+L}c_k|^2. \tag{2.17}$$

By substituting equation (2.16) into equation (2.8), we obtain

$$\begin{aligned}\hat{W}_m = d_0 \sum_{J=0}^{KN} A_J^2 |J\rangle\langle J| + \sum_{L=1}^{K} d_L \sum_{J=0}^{KN-L} A_J A_{J+L} \\ \times \left(\mathrm{e}^{\mathrm{i}\frac{2\pi m}{M}L} |J\rangle\langle J+L| + \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{M}L} |J+L\rangle\langle J|\right).\end{aligned} \tag{2.18}$$

We then have

$$\hat{\Gamma} = \sum_{m=0}^{M-1} \hat{W}_m \hat{\mu}_m = d_0 \sum_{J=0}^{KN} A_J^2 |J\rangle\langle J| + \sum_{L=1}^{K} d_L \sum_{J=0}^{KN-L} A_J A_{J+L} (|J\rangle\langle J| + |J+L\rangle\langle J+L|) \tag{2.19}$$

where the orthogonality relation equation (2.14) was used under the condition $M \geqslant KN + 1$. The operator $\hat{\Gamma} - \hat{W}_m$ is now

$$\hat{\Gamma} - \hat{W}_m = \sum_{L=1}^{K} d_L \sum_{J=0}^{KN-L} A_J A_{J+L} \Big[|J\rangle\langle J| + |J+L\rangle\langle J+L| - \left(\mathrm{e}^{\mathrm{i}\frac{2\pi m}{M}L} |J\rangle\langle J+L| + \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{M}L} |J+L\rangle\langle J| \right) \Big]. \tag{2.20}$$

Now look at each operator enclosed by [...] in equation (2.20). This 2×2 matrix has the eigenvalues 0 and *2* and is non-negative. So is the sum of them. Since the coefficients d_L and A_J in equation (2.20) are also positive, the second condition (ii) of equation (2.9) is proved. The first condition (i) can be easily checked by direct calculation. The maximum average score is then given by

$$\bar{S}_{\mathrm{MAX}}(N) = d_0 \sum_{J=0}^{KN} A_J^2 + 2 \sum_{L=1}^{K} d_L \sum_{J=0}^{KN-L} A_J A_{J+L}. \tag{2.21}$$

This maximum is independent of the number of sample points *M*; that is, for $M \geqslant KN + 1$ the attainable average score is saturated. In fact, in the limit of $M \to \infty$ we can construct the infinite continuous POM

$$\mathrm{d}\hat{\Pi}(\varphi) \equiv |\mu(\varphi)\rangle\langle\mu(\varphi)| \mathrm{d}\varphi/2\pi \qquad \int_0^{2\pi} \mathrm{d}\hat{\Pi}(\varphi) = \hat{I} \tag{2.22}$$

where

$$|\mu(\varphi)\rangle \equiv \left(\frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\phi \, \hat{\Psi}(\phi) \right)^{-\frac{1}{2}} |\Psi(\varphi)\rangle = \sum_{J=0}^{KN} \mathrm{e}^{-\mathrm{i}\varphi J} \, |J\rangle \tag{2.23}$$

and this attains the same maximum as equation (2.21). Thus it is proved that the measurement state vector (2.12) (also (2.15)) provides the optimal estimation strategy. For the minimum number of sample points $M = KN + 1$, $\{|\mu_m\rangle\}$ is an orthonormal set, that is, a von Neumann measurement.

Here we mention other strategies $\{\hat{\mu}_m^{\perp}\}$ which extremize the average score (2.7) (for the same $\hat{W}_m$ as in equation (2.8)), that is, satisfy the first condition (i) of equation (2.9), but not necessarily the second condition (ii). Consider, for example, the states orthogonal to the *N*-tensor-product sample states $|\Psi_m\rangle$, that is, the states

$$|\Psi_m^{\perp}\rangle = \sum_{J=0}^{KN} \binom{KN}{J} A_J^{-1} (-1)^J \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{M}J} \, |J\rangle \,. \tag{2.24}$$

The square root measurement for discriminating $\{|\Psi_m^{\perp}\rangle\}$ is given as

$$|\mu_m^{\perp}\rangle \equiv \left(\sum_{m=0}^{M-1} |\Psi_m^{\perp}\rangle\langle\Psi_m^{\perp}| \right)^{-\frac{1}{2}} |\Psi_m^{\perp}\rangle = \frac{1}{\sqrt{M}} \sum_{J=0}^{KN} (-1)^J \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{M}J} \, |J\rangle \,. \tag{2.25}$$

The Lagrange operator for this measurement is

$$\hat{\Gamma}^{\perp} = \sum_{m=0}^{M-1} \hat{W}_m \hat{\mu}_m^{\perp} = d_0 \sum_{J=0}^{KN} A_J^2 |J\rangle\langle J| + \sum_{L=1}^{K} d_L \sum_{J=0}^{KN-L} A_J A_{J+L} (-1)^L (|J\rangle\langle J| + |J+L\rangle\langle J+L|) \tag{2.26}$$

and we have that

$$\hat{\Gamma}^{\perp} - \hat{W}_m = \sum_{L=1}^{K} d_L \sum_{J=0}^{KN-L} A_J A_{J+L} \Big[(-1)^L (|J\rangle\langle J| + |J+L\rangle\langle J+L|) - \Big(\mathrm{e}^{\mathrm{i}\frac{2\pi m}{M}L} |J\rangle\langle J+L| + \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{M}L} |J+L\rangle\langle J| \Big) \Big] \tag{2.27}$$

which is easily seen to satisfy

$$\left(\hat{\Gamma}^{\perp} - \hat{W}_m \right) \left| \mu_m^{\perp} \right\rangle = 0. \tag{2.28}$$

Thus the POM $\{\hat{\mu}_m^{\perp} \equiv |\mu_m^{\perp}\rangle\langle\mu_m^{\perp}|\}$ is an extremal solution for the average score (2.7). However, for example, in the case of a two-dimensional system ($K = 1$ in equation (2.1)), equation (2.27) reduces to

$$\hat{\Gamma}^{\perp} - \hat{W}_m = -d_1 \sum_{J=0}^{N-1} A_J A_{J+1} \Big[|J\rangle\langle J| + |J+1\rangle\langle J+1| + \mathrm{e}^{\mathrm{i}\frac{2\pi m}{M}} |J\rangle\langle J+1| + \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{M}} |J+1\rangle\langle J| \Big] \tag{2.29}$$

which is clearly a non-positive definite operator; that is, $\hat{\Gamma}^{\perp} - \hat{W}_m \leqslant 0$. In contrast to equation (2.9), this means that $\{\hat{\mu}_m^{\perp}\}$ attains a *global minimum* of the average score. In the general case of $K > 1$, it is not necessarily the case that $\hat{\Gamma}^{\perp} - \hat{W}_m$ is positive or negative definite. Thus, in general, the set $\{\hat{\mu}_m^{\perp}\}$ may represent strategies that attain either *local* maxima or minima.

3. Quantum circuit for the optimal strategy

Let us consider physical implementations of the optimal estimation strategy represented by equation (2.15). This is a collective measurement on a finite sample system. This sort of measurement can, in principle, be realized by a quantum circuit acting on the combined system and a separable measurement on each sample system. From this point of view, the main problem is a synthesis of an appropriate quantum circuit. When we take the minimum number of outputs $M = KN + 1$, the measurement basis $\{|\mu_m\rangle\}$ is orthonormal, and equation (2.15) is the discrete Fourier transform in the subspace $\mathcal{H}_{\mathrm{T}}$. The discrete Fourier transform is a fundamental tool in quantum computation. The corresponding circuit working on qubit systems is already known (see, e.g., [14]). Therefore in the case of a two-dimensional system (that is, $K = 1$ in equation (2.2)), we may apply this result to synthesizing the optimal estimation strategy.

Let us start with the simplest two-dimensional system with $N = 2$. The state to be measured is

$$|\psi(\theta)\rangle^{\otimes 2} = A_0\,|0\rangle_{\mathrm{T}} + A_1 \mathrm{e}^{-\mathrm{i}\theta}\,|1\rangle_{\mathrm{T}} + A_2 \mathrm{e}^{-\mathrm{i}2\theta}\,|2\rangle_{\mathrm{T}} \tag{3.1}$$

where $A_0 = c_0^2$, $A_1 = \sqrt{2} c_0 c_1$, $A_2 = c_1^2$, and

$$|0\rangle_{\mathrm{T}} = |00\rangle \tag{3.2a}$$

$$|1\rangle_{\mathrm{T}} = \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) \tag{3.2b}$$

$$|2\rangle_{\mathrm{T}} = |11\rangle \tag{3.2c}$$

assuming that the coefficients c_i are real for eliminating inessential phase factors. Here we have used the subscript T for denoting the basis $|J\rangle$ of the $(KN + 1)$-dimensional subspace $\mathcal{H}_{\mathrm{T}}$, and the basis states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$ span the space $\mathcal{H}^{\otimes 2}$. The optimal strategy with the minimal outputs is represented as

$$|\mu_m\rangle = \frac{1}{\sqrt{3}} \left(|0\rangle_{\mathrm{T}} + \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{3}} |1\rangle_{\mathrm{T}} + \mathrm{e}^{-\mathrm{i}\frac{4\pi m}{3}} |2\rangle_{\mathrm{T}} \right) \quad (m = 0, 1, 2). \tag{3.3}$$

This expression is, however, rather inconvenient for directly applying the discrete Fourier transform quantum network, which is usually defined in a 2^n-dimensional space corresponding to n qubit systems. Therefore, it seems better to consider the measurement circuit in a four-dimensional space $\mathcal{H}^{\otimes 2}$ spanned by the basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. In fact we also have the four-output optimal, nonminimal ($M = 4$) strategy in $\mathcal{H}_{\mathrm{T}}$ as

$$|\mu_m\rangle = \frac{1}{\sqrt{4}}\left(|0\rangle_{\mathrm{T}} + \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{4}}|1\rangle_{\mathrm{T}} + \mathrm{e}^{-\mathrm{i}\frac{4\pi m}{4}}|2\rangle_{\mathrm{T}}\right) \qquad (m = 0, 1, 2, 3) \tag{3.4}$$

and this can be extended to a von Neumann measurement in $\mathcal{H}^{\otimes 2}$. It is thus helpful to transform the state of the input samples by the unitary operator $\hat{T}^{(2)}$ defined by the circuit shown in figure 1 such that

$$\hat{T}^{(2)}|\psi(\theta)\rangle^{\otimes 2} = A_0|00\rangle + A_1\mathrm{e}^{-\mathrm{i}\theta}|01\rangle + A_2\mathrm{e}^{-\mathrm{i}2\theta}|10\rangle\,. \tag{3.5}$$

We may then apply the measurement

$$|\mu_m\rangle = \frac{1}{\sqrt{4}}\left(|00\rangle + \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{4}}|01\rangle + \mathrm{e}^{-\mathrm{i}\frac{4\pi m}{4}}|10\rangle\right) \qquad (m = 0, 1, 2, 3) \tag{3.6}$$

or its Naimark extension in $\mathcal{H}^{\otimes 2}$

$$|\Pi_m\rangle = \frac{1}{\sqrt{4}}\left(|00\rangle + \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{4}}|01\rangle + \mathrm{e}^{-\mathrm{i}\frac{4\pi m}{4}}|10\rangle + \mathrm{e}^{-\mathrm{i}\frac{6\pi m}{4}}|11\rangle\right) \qquad (m = 0, 1, 2, 3) \tag{3.7}$$

which, using the discrete Fourier transform $\hat{U}_{\mathrm{DFT}}$ shown by the circuit of figure 2, can be explicitly written as

$$|\Pi_0\rangle = \hat{U}^{\dagger}_{\mathrm{DFT}}|00\rangle \tag{3.8a}$$
$$|\Pi_1\rangle = \hat{U}^{\dagger}_{\mathrm{DFT}}|10\rangle \tag{3.8b}$$
$$|\Pi_2\rangle = \hat{U}^{\dagger}_{\mathrm{DFT}}|01\rangle \tag{3.8c}$$
$$|\Pi_3\rangle = \hat{U}^{\dagger}_{\mathrm{DFT}}|11\rangle\,. \tag{3.8d}$$

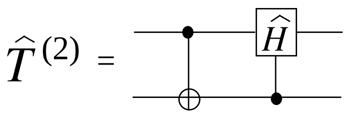

Figure 1. The circuit for the basis transformation from $\{|0\rangle_{\mathrm{T}}, |1\rangle_{\mathrm{T}}, |2\rangle_{\mathrm{T}}\}$ to $\{|00\rangle, |01\rangle, |10\rangle\}$. $\hat{H}$ is the Hadamard transformation.

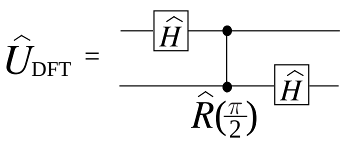

Figure 2. The circuit for the discrete Fourier transform on two qubit systems. The two bit gate $\hat{R}(\phi)$ performs the transformation $|x\rangle|y\rangle \mapsto \mathrm{e}^{\mathrm{i}xy\phi}|x\rangle|y\rangle$.

Thus the optimal phase estimation can be realized by first performing the unitary transformation $\hat{U}_{\mathrm{DFT}}\hat{T}^{(2)}$ on the input state $|\psi(\theta)\rangle^{\otimes 2}$, then measuring the transformed state in the basis $\{|00\rangle, |10\rangle, |01\rangle, |11\rangle\}$ (which is a separable measurement), and finally deciding the phase as $\theta = 0, \frac{\pi}{2}, \pi$ or $\frac{3\pi}{2}$, according to whether the outcome is $|00\rangle, |10\rangle, |01\rangle$ or $|11\rangle$, respectively. This is summarized in figure 3.

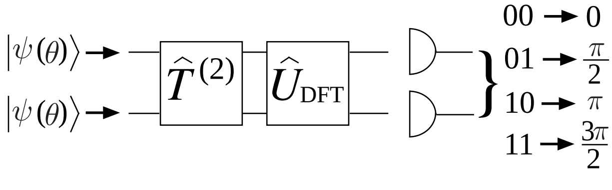

Figure 3. The circuit structure for the optimal estimation strategy in the case of N=2.

Similarly, in the case of $N = 3$, the input state is represented as

$$|\psi(\theta)\rangle^{\otimes 3} = A_0|\bar{0}\rangle_{\mathrm{T}} + A_1\mathrm{e}^{-\mathrm{i}\theta}|\bar{1}\rangle_{\mathrm{T}} + A_2\mathrm{e}^{-\mathrm{i}2\theta}|\bar{2}\rangle_{\mathrm{T}} + A_3\mathrm{e}^{-\mathrm{i}3\theta}|\bar{3}\rangle_{\mathrm{T}} \tag{3.9}$$

where

$$|\bar{0}\rangle_{\mathrm{T}} = |000\rangle \tag{3.10a}$$

$$|\bar{1}\rangle_{\mathrm{T}} = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle) \tag{3.10b}$$

$$|\bar{2}\rangle_{\mathrm{T}} = \frac{1}{\sqrt{3}}(|110\rangle + |101\rangle + |011\rangle) \tag{3.10c}$$

$$|\bar{3}\rangle_{\mathrm{T}} = |111\rangle. \tag{3.10d}$$

Let $\hat{T}^{(3)}$ be the unitary operator which converts the basis states $\{|0\rangle_{\mathrm{T}}, |1\rangle_{\mathrm{T}}, |2\rangle_{\mathrm{T}}, |3\rangle_{\mathrm{T}}\}$ into $|0\rangle \otimes \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, respectively; that is,

$$\hat{T}^{(3)}\,|\psi(\theta)\rangle^{\otimes 3} = |0\rangle \otimes \left(A_0\,|00\rangle + A_1\mathrm{e}^{-\mathrm{i}\theta}\,|01\rangle + A_2\mathrm{e}^{-\mathrm{i}2\theta}\,|10\rangle + A_3\mathrm{e}^{-\mathrm{i}3\theta}\,|11\rangle\right). \tag{3.11}$$

The estimation strategy can then be constructed again in the four-dimensional space $\mathcal{H}^{\otimes 2}$, where the minimal optimal measurement is actually given by equation (3.7) and can be realized just as in the previous case. The unitary operator $\hat{T}^{(3)}$ can be effected by the circuit shown in figure 4 which consists of three main blocks. In the first block, the operator $\hat{T}^{(2)}$ acts on the first two qubits of the $\mathcal{H}_T$ basis states, which gives

$$|\bar{0}\rangle_{\mathrm{T}} \mapsto |000\rangle \tag{3.12a}$$

$$|\bar{1}\rangle_{\mathrm{T}} \mapsto \frac{1}{\sqrt{3}}|001\rangle + \sqrt{\frac{2}{3}}|010\rangle \tag{3.12b}$$

$$|\bar{2}\rangle_{\mathrm{T}} \mapsto \frac{1}{\sqrt{3}}|100\rangle + \sqrt{\frac{2}{3}}|011\rangle \tag{3.12c}$$

$$|\bar{3}\rangle_{\mathrm{T}} \mapsto |101\rangle. \tag{3.12d}$$

The second block transforms the state $|100\rangle$ into the state $|110\rangle$ and the last block $\hat{S}^{(3)}(\hat{v}_1)$, which includes the conditional rotation on one qubit

$$\hat{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} \\ -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \tag{3.13}$$

leads to the final basis states $|0\rangle \otimes \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ as required.

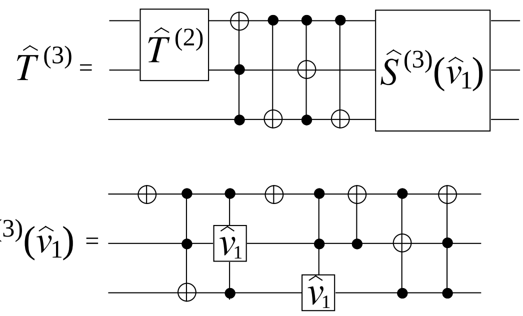

Figure 4. The circuit which converts the basis states $\{|\bar{0}\rangle_{\mathrm{T}}, |\bar{1}\rangle_{\mathrm{T}}, |\bar{2}\rangle_{\mathrm{T}}, |\bar{3}\rangle_{\mathrm{T}}\}$ into $|0\rangle \otimes \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

Finally we mention the case of $N = 4$. Let us observe that the $\mathcal{H}_T$ basis can be written as

$$|\bar{\bar{0}}\rangle_{\mathrm{T}} = |000\rangle \otimes |0\rangle \tag{3.14a}$$

$$|\bar{\bar{1}}\rangle_{\mathrm{T}} = \frac{1}{2}[(|001\rangle + |010\rangle + |100\rangle) \otimes |0\rangle + |000\rangle \otimes |1\rangle] \tag{3.14b}$$

$$|\bar{\bar{2}}\rangle_{\mathrm{T}} = \frac{1}{\sqrt{6}}[(|001\rangle + |010\rangle + |100\rangle) \otimes |1\rangle + (|110\rangle + |101\rangle + |011\rangle) \otimes |0\rangle] \tag{3.14c}$$

$$|\bar{\bar{3}}\rangle_{\mathrm{T}} = \frac{1}{2}[(|110\rangle + |101\rangle + |011\rangle) \otimes |1\rangle + |111\rangle \otimes |0\rangle] \tag{3.14d}$$

$$|\bar{\bar{4}}\rangle_{\mathrm{T}} = |111\rangle \otimes |1\rangle. \tag{3.14e}$$

The first three qubits can be transformed into $|000\rangle$, $|001\rangle$, $|010\rangle$ or $|011\rangle$ by applying $\hat{T}^{(3)}$. Thus the first qubit of all the basis states becomes $|0\rangle$, and can be factorized out. By further applying on the remaining three qubits a CC-NOT gate, the operator $\hat{S}^{(3)}(\hat{v}_2)$ with

$$\hat{v}_2 = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \tag{3.15}$$

and the operator $\hat{W}^{(3)}$ as shown in figure 5, the basis states $\{|\bar{\bar{0}}\rangle_{\mathrm{T}}, \ldots, |\bar{\bar{4}}\rangle_{\mathrm{T}}\}$ are finally transformed into $|0\rangle \otimes \{|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle\}$, respectively.

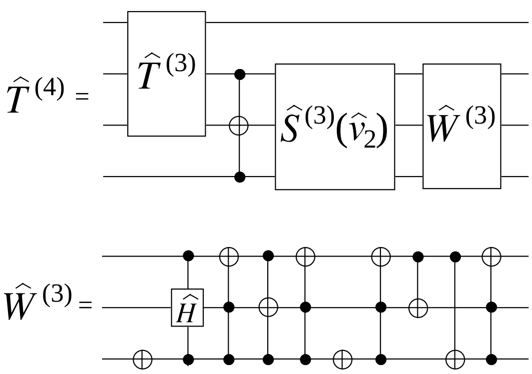

Figure 5. The circuit which transforms the basis states $\{|\bar{\bar{0}}\rangle_{\mathrm{T}}, \ldots, |\bar{\bar{4}}\rangle_{\mathrm{T}}\}$ into $|0\rangle \otimes \{|000\rangle, |001\rangle, |010\rangle, |011\rangle, |100\rangle\}$, respectively.

Let $\hat{T}^{(4)}$ be this basis transformation represented by the circuit of figure 5. After transforming the input state by $\hat{T}^{(4)}$, it is then sufficient to perform the measurement on the last three qubits. Let $\{|L\rangle_3\,; L = 0, 1, \ldots, 7\}$ be the three-qubit basis $\{|000\rangle, |001\rangle, \ldots, |111\rangle\}$. Then the minimal optimal measurement is given as

$$|\mu_m\rangle = \frac{1}{\sqrt{5}} \sum_{L=0}^{4} \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{5}L} |L\rangle_3 \qquad (m = 0, 1, 2, 3, 4). \tag{3.16}$$

However, for applying the discrete Fourier transform network, it is convenient to take the other optimal strategy consisting of the overcomplete states

$$|\mu_m\rangle = \frac{1}{\sqrt{8}} \sum_{L=0}^{4} \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{8}L} |L\rangle_3 \qquad (m = 0, 1, \ldots, 7). \tag{3.17}$$

These can be orthogonalized in the eight-dimensional space $\mathcal{H}^{\otimes 3}$ as

$$|\Pi_m\rangle = \frac{1}{\sqrt{8}} \sum_{L=0}^{7} \mathrm{e}^{-\mathrm{i}\frac{2\pi m}{8} L} |L\rangle_3 \qquad (m = 0, 1, \ldots, 7). \tag{3.18}$$

This is just the discrete Fourier transform in $\mathcal{H}^{\otimes 3}$, and can be written in the form $|\Pi_m\rangle = \hat{U}^{\dagger}_{\mathrm{DFT}} |m\rangle_3$. The circuit corresponding to $\hat{U}_{\mathrm{DFT}}$ is found in [14]. Therefore the optimal estimation is realized by applying the unitary transform $(\hat{I} \otimes \hat{U}_{\mathrm{DFT}}) \cdot \hat{T}^{(4)}$ on the input state $|\psi(\theta)\rangle^{\otimes 4}$ and then by measuring the last three qubits of the transformed state in the basis $\{|L\rangle_3\}$. According to the outcome, we decide the phase to be $\exp(-\mathrm{i}2\pi L/8)$.

The case of larger N can be treated in a similar way by applying the circuits used in the case of lower N inductively. In the general case of higher dimensional systems, one should first develop basic tools for constructing quantum circuits (for some recent progress see, e.g., [15]), and at present practical circuit synthesizations remain an open problem.

4. Concluding remarks

We have shown how to construct the optimal strategies for estimating a displacement parameter of an arbitrary finite-dimensional system in a pure state. These are based on the square root measurement for discriminating the states corresponding to the uniformly distributed sample points of the parameter. We have assumed that the *a priori* probability distribution of the parameter is uniform. When the *a priori* distribution is not uniform, or the system to be estimated is in a mixed state, the strategy based on the square root measurement is not in general optimal.

Within the assumption of a uniform *a priori* distribution and the purity of the system to be estimated, it is a remaining problem whether our method applies to the estimation of two or more parameters. The simplest case would be the estimation of a two-state system using finite identical samples $|\psi(\theta, \phi)\rangle^{\otimes N}$. As for the infinite continuous POM, we can show that the square root measurement

$$\mathrm{d}\hat{\Pi}(\theta, \phi) \equiv |\mu(\theta, \phi)\rangle\langle\mu(\theta, \phi)| \frac{\mathrm{d}\phi\, \mathrm{d}\theta \sin\theta}{2\pi} \qquad \int_0^{2\pi} \int_0^{\pi} \mathrm{d}\hat{\Pi}(\theta, \phi) = \hat{I} \tag{4.1}$$

where

$$|\mu(\theta, \phi)\rangle \equiv \left[\frac{1}{4\pi} \int_0^{2\pi} \mathrm{d}\phi' \int_0^{\pi} \mathrm{d}\theta' \sin\theta' \left(|\psi(\theta, \phi)\rangle\, \langle\psi(\theta, \phi)| \right)^{\otimes N} \right]^{-\frac{1}{2}} |\psi(\theta, \phi)\rangle^{\otimes N} \tag{4.2}$$

provides the optimal strategy. Whether its discrete and finite version should exist and be built from the uniformly distributed sample points of (θ, ϕ) is still an open question at present.

Acknowledgments

The authors would like to thank O Hirota, S M Barnett, A S Holevo and C A Fuchs for valuable discussions. A Cerlini's research is supported by JISTEC under grant no 199016. A Chefles acknowledges the British Council for financial support.

References

[1] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[2] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)

[3] Massar S and Popescu S 1995 *Phys. Rev. Lett.* **74** 1259
[4] Derka R, Buzek V and Ekert A K 1998 *Phys. Rev. Lett.* **80** 1571
[5] Latorre J I, Pascual P and Tarrach R 1998 *Phys. Rev. Lett.* **81** 1351
[6] Vidal G, Latorre J I, Pascual P and Tarrach R 1999 *Phys. Rev.* A **60** 126
[7] Acín A, Latorre J I and Pascual P 2000 *Phys. Rev.* A **61** 022113
[8] Bruß D and Macchiavello C 1999 *Phys. Lett.* A **253** 249
[9] Werner R F 1998 *Phys. Rev.* A **58** 1827
[10] Barnett S M, Chefles A, Pegg D T and Phillip L S 2000 *Quantum Communication, Computing, and Measurement* vol 2, ed Kumar, D'Ariano and O Hirota (Dordrecht: Kluwer Academic/New York: Plenum) pp 125–36
[11] Holevo A S 1978 *Theor. Prob. Appl.* **23** 411
[12] Hausladen P and Wootters W K 1994 *J. Mod. Opt.* **41** 2385
[13] Ban M, Kurokawa K, Momose R and Hirota O 1997 *Int. J. Theor. Phys.* **36** 1269
[14] Cleve R, Ekert A, Macchiavello C and Mosca M 1998 *Proc. R. Soc. Lond.* A **454** 339
[15] Alber G, Delgado A, Gisin N and Jex I 2000 *Preprint* arXiv quant-ph/0008022