PHYSICAL REVIEW A, **66**, 022308 (2002)

Optimal parameter estimation of a depolarizing channel

Masahide Sasaki,[1,2,*] Masashi Ban,[3] and Stephen M. Barnett[4]
[1]*Communications Research Laboratory, Koganei, Tokyo 184-8795, Japan*
[2]*CREST, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan*
[3]*Advanced Research Laboratory, Hitachi Ltd., 1-280, Higashi-Koigakubo, Kokubunnji, Tokyo 185-8601, Japan*
[4]*Department of Physics and Applied Physics, University of Strathclyde, Glasgow G4 0NG, Scotland*
(Received 22 March 2002; published 14 August 2002)

We investigate strategies for estimating a depolarizing channel for a finite dimensional system. Our analysis addresses the double optimization problem of selecting the best input probe state and the measurement strategy that minimizes the Bayes cost of a quadratic function. In the qubit case, we derive the Bayes optimal strategy for any finite number of input probe particles when bipartite entanglement can be formed in the probe particles.

DOI: 10.1103/PhysRevA.66.022308 PACS number(s): 03.67.Hk, 03.65.Ta

I. INTRODUCTION

In order to design a reliable communication system one requires *a priori* knowledge of the property of a channel. Precise knowledge of the channel allows us to devise appropriate coding, modulation, and filtering schemes. In general, the channel property is not stationary, so one should first acquire and then track the optimal operating point of each device by monitoring the condition of the channel. It is important, therefore, to know how to estimate the channel property in an efficient way, that is, as precisely as possible with minimum resources.

A reasonable assumption is that we know that the channel belongs to a certain parametrized family, and only the values of the parameters are not known. To know them one may input a probe system in an appropriate state into the channel and make a measurement on the output state. Only when an infinite amount of input resource is available, one can determine the channel parameters with perfect accuracy. In the quantum domain, however, the resource is often restricted for various reasons. For example, when one is to monitor a fast quantum dymanics at cryogenic temperatures, the input probe power should be kept as low as possible so as to prevent the system from heating up while obtaining meaningful data in a short time. This restricts the available amount of probe particles. Furthermore, preparing the probe in an appropriate quantum state is usually an elaborate process. Thus to find the efficient estimation strategy relying only on a restricted amount of input resource is of practical importance.

In estimating a quantum channel parameter, given a finite amount of input resource, both the input probe state and the measurement of the output state need to be optimized. This double maximization problem has been studied in the context of estimation of SU(d) unitary operation [1]. Estimating a noisy quantum channel has been discussed in the literature [2–4]. In Ref. [2], the locally optimal strategy, which achieves the Cramér-Rao bound at a local point of the parameter space, was derived for the depolarizing channel for a qubit system. Such a measurement is usually used to get the final estimate around the most likely value of the parameter obtained by some preliminary estimation using a part of the probe particles. By this two stage estimation, one can attain the optimal asymptotic rate at which the estimation accuracy grows with the number of probe particles [5–7]. References [3,4] focus on several noisy qubit channels. They study some reasonable, although not optimal, strategies based on the maximum likelihood estimator, and derive the asymptotic behavior of the cost as a function of the number of input probe qubits.

In contrast, we are concerned here with the Bayes optimal strategy which minimizes the *average* cost. The scenario we have in mind is that one has no particular knowledge about the *a priori* parameter distribution, and the available number of probe particles is strictly limited. We then take into account the possibility of rather large errors. We seek the strategy that works equally well for all possible values of the parameter on average, that is, the strategy which is more universal for various possible situations.

It seems difficult for us to study this problem for the most general probe state. In this paper we deal with the depolarizing channel by assuming that we dispose of M pairs of probe particles and only bipartite entanglement can be formed in each pair. This might be a practically sensible assumption from the view point of optical implementation given current technology. Our problem is to find the best estimation strategy to minimize the average cost. We consider the quadratic of a cost function.

The paper is organized as follows. In Sec. II we analyze the qubit case with a pair of probe particles and derive the optimal estimation strategy. In Sec. III we extend our analysis to the d-dimensional case restricting the probe state to be the maximally entangled pairs instead of arbitrary bipartite entanglement pairs or the completely separable state. Finally in Sec. IV we summarize the results and discuss some physical interpretations of them.

II. QUBIT CASE

Let $\hat{\rho}$ be a density operator in the two-dimensional Hilbert space $\mathcal{H}_2$. The depolarizing channel $\mathcal{L}_\theta$ maps a density operator $\hat{\rho}$ to a density operator which is a mixture of $\hat{\rho}$ and the maximally mixed state,

*Electronic address: psasaki@crl.go.jp

$$\mathcal{L}_\theta \hat{\rho} = \theta\hat{\rho} + \frac{1-\theta}{2}\hat{I}. \tag{1}$$

The parameter θ represents the degree of randomization of polarization. For the map $\mathcal{L}_\theta$ to be completely positive, the parameter θ must lie in the interval $-\frac{1}{3} \leqslant \theta \leqslant 1$.

Let us start with two qubit systems as the input probe. For simplicity we only consider a pure state family of the probe $\hat{\Psi} = |\Psi\rangle\langle\Psi|$. This may be represented in the Schmidt decomposition

$$|\Psi\rangle = \sqrt{x}|0\rangle \otimes |e_0\rangle + \sqrt{1-x}|1\rangle \otimes |e_1\rangle, \tag{2}$$

where $\{|0\rangle,|1\rangle\}$ and $\{|e_0\rangle,|e_1\rangle\}$ are orthonormal basis sets for the first and second probe particle, respectively. What is the best way to use this state? There are two possibilities to consider:

(a) input one qubit of the pair into the channel keeping the other untouched leading to the output state

$$\hat{\Psi}_1(\theta) \equiv (\mathcal{L}_\theta \otimes \hat{I})|\Psi\rangle\langle\Psi|, \tag{3}$$

(b) input both qubits into the channel and have the output state

$$\hat{\Psi}_2(\theta) \equiv (\mathcal{L}_\theta \otimes \mathcal{L}_\theta)|\Psi\rangle\langle\Psi|. \tag{4}$$

A measurement is described by a probability operator measure (POM) $\hat{\Pi}(\theta)$ [8,9], also referred to as a positive operator valued measure (POVM) [10]. The average cost for the quadratic cost function is given by

$$\overline{C}_i(x) = \int_{-(1/3)}^{1} d\tilde{\theta} \int_{-(1/3)}^{1} d\theta (\tilde{\theta}-\theta)^2 z(\theta) \mathrm{Tr}[\hat{\Pi}(\tilde{\theta})\hat{\Psi}_i(\theta)], \tag{5}$$

where $z(\theta)$ is the *a priori* probability distribution of θ, and $\int_{-(1/3)}^{1} d\tilde{\theta}\hat{\Pi}(\tilde{\theta}) = \hat{I}$. It is assumed that we have no *a priori* knowledge about θ, that is, $z(\theta) = \frac{3}{4}$. Given the channel $\mathcal{L}_\theta$, we are to find the optimal probe $|\Psi\rangle$ and the POM $\hat{\Pi}(\theta)$ minimizing the average cost $\overline{C}(x)$.

It is convenient to introduce the *risk* operator

$$\hat{W}(\theta) = \frac{3}{4}\int_{-(1/3)}^{1} d\theta' (\theta-\theta')^2 \hat{\Psi}_i(\theta') \tag{6}$$

$$= \hat{W}^{(2)} - 2\theta\hat{W}^{(1)} + \theta^2\hat{W}^{(0)}, \tag{7}$$

where $\hat{W}^{(k)} \equiv \frac{3}{4}\int_{-(1/3)}^{1} d\theta\theta^k \hat{\Psi}_i(\theta)$. The average cost is then

$$\overline{C}(x) = \mathrm{Tr}\,\hat{\Gamma}, \quad \hat{\Gamma} \equiv \int_{-(1/3)}^{1} d\theta \hat{\Pi}(\theta)\hat{W}(\theta). \tag{8}$$

For a fixed probe state $|\Psi\rangle$, the optimal POM $\hat{\Pi}(\theta)$ is derived from the necessary and sufficient conditions to minimize the average cost [11,12]:

(i) $\hat{\Gamma} = \hat{\Gamma}^\dagger$, and $[\hat{W}(\theta) - \hat{\Gamma}]\hat{\Pi}(\theta) = 0$ for all θ,

(ii) $\hat{W}(\theta) - \hat{\Gamma} \geqslant 0$ for all θ.

The optimal solution for a single parameter estimation with a quadratic cost is well known [13,8]. The optimal POM is constructed by finding the eigenstate $|\theta\rangle$ of the *minimizing* operator $\hat{\Theta}$ which is defined by

$$\hat{\Theta}\hat{W}^{(0)} + \hat{W}^{(0)}\hat{\Theta} = 2\hat{W}^{(1)}, \tag{9}$$

that is, $\hat{\Pi}(\theta) = |\theta\rangle\langle\theta|$ so that $\hat{\Theta}|\theta\rangle = \theta|\theta\rangle$. We then have $\hat{\Gamma} = \hat{W}^{(2)} - \hat{\Theta}\hat{W}^{(0)}\hat{\Theta}$ from which the conditions (i) and (ii) are easily verified.

For a discrete system, one can find the optimal POM with finite elements. Let the spectral decomposition of $\hat{W}^{(0)}$ for our two-qubit system be

$$\hat{W}^{(0)} = \sum_{i=1}^{4} \omega_i |\omega_i\rangle\langle\omega_i|. \tag{10}$$

Then the minimizing operator is

$$\hat{\Theta} = \sum_{i,j=1}^{4} \frac{2}{\omega_i + \omega_j} |\omega_i\rangle\langle\omega_i|\hat{W}^{(1)}|\omega_j\rangle\langle\omega_j|. \tag{11}$$

Let the spectral decomposition of $\hat{\Theta}$ be

$$\hat{\Theta} = \sum_{i=1}^{4} \theta_i |\theta_i\rangle\langle\theta_i|. \tag{12}$$

The optimal POM is then given by

$$\hat{\Pi}(\theta) = \sum_{i=1}^{4} \delta(\theta - \theta_i)|\theta_i\rangle\langle\theta_i|. \tag{13}$$

This implies that the measurement has four outputs at most and we then estimate the channel parameter as one of four θ_i's. Before going on to derive the optimal strategies, let us define some notations. As seen below the output states $\hat{\Psi}_i(\theta)$'s can be written as a direct sum

$$\hat{\Psi}_i(\theta) = \hat{\psi}_i(\theta) \oplus \hat{\phi}_i(\theta), \tag{14}$$

where $\hat{\psi}_i(\theta)$ is in the subspace $\mathcal{H}_\psi$ spanned by $|\mu_1\rangle \equiv |0\rangle \otimes |e_0\rangle$ and $|\mu_2\rangle \equiv |1\rangle \otimes |e_1\rangle$, and $\hat{\phi}_i(\theta)$ in the subspace $\mathcal{H}_\phi$ spanned by $|\nu_1\rangle \equiv |0\rangle \otimes |e_1\rangle$ and $|\nu_2\rangle \equiv |1\rangle \otimes |e_0\rangle$. In the following all 2×2 matrices represent density operators in $\mathcal{H}_\psi$ with $|\mu_1\rangle = \binom{1}{0}$ and $|\mu_2\rangle = \binom{0}{1}$.

A. Case (a)

The output state $\hat{\Psi}_1(\theta)$ is given by

$$\hat{\psi}_1(\theta) = \frac{1}{2}\begin{bmatrix} (1+\theta)x & 2\theta\sqrt{x(1-x)} \\ 2\theta\sqrt{x(1-x)} & (1+\theta)(1-x) \end{bmatrix}, \tag{15}$$

$$\hat{\phi}_1(\theta) = \frac{1-\theta}{2}[(1-x)|\nu_1\rangle\langle\nu_1| + x|\nu_2\rangle\langle\nu_2|]. \tag{16}$$

The elements of the risk operator are

$$\hat{W}^{(0)}=\frac{1}{3}\left(\begin{bmatrix}2x & \sqrt{x(1-x)}\\ \sqrt{x(1-x)} & 2(1-x)\end{bmatrix}\oplus\hat{\varphi}_1\right),\qquad(17)$$

$$\hat{W}^{(1)}=\frac{1}{27}\left(\begin{bmatrix}8x & 7\sqrt{x(1-x)}\\ 7\sqrt{x(1-x)} & 8(1-x)\end{bmatrix}\oplus\hat{\varphi}_1\right),\qquad(18)$$

$$\hat{W}^{(2)}=\frac{1}{27}\left(\begin{bmatrix}6x & 5\sqrt{x(1-x)}\\ 5\sqrt{x(1-x)} & 6(1-x)\end{bmatrix}\oplus\hat{\varphi}_1\right),\qquad(19)$$

where $\hat{\varphi}_1=(1-x)|\nu_1\rangle\langle\nu_1|+x|\nu_2\rangle\langle\nu_2|$. After a lengthy but straightforward calculation (see Appendix A) we have

$$\hat{\Theta}=\frac{2}{9}\begin{bmatrix}1+x & 2\sqrt{x(1-x)}\\ 2\sqrt{x(1-x)} & 2-x\end{bmatrix}\oplus\frac{1}{9}\hat{I}_\phi.\qquad(20)$$

To diagonalize Θ we introduce $r=\sqrt{1+12x(1-x)}$ and

$$\cos\gamma=\sqrt{\frac{r-1+2x}{2r}},\quad \sin\gamma=\sqrt{\frac{r+1-2x}{2r}}.\qquad(21)$$

The eigenstates and eigenvalues are then

$$\begin{aligned}|\theta_1\rangle&=\cos\gamma|\mu_1\rangle+\sin\gamma|\mu_2\rangle,\quad \theta_1=(3+r)/9,\\ |\theta_2\rangle&=-\sin\gamma|\mu_1\rangle+\cos\gamma|\mu_2\rangle,\quad \theta_2=(3-r)/9,\qquad(22)\\ |\theta_3\rangle&=|\nu_1\rangle,\quad \theta_3=1/9,\\ |\theta_4\rangle&=|\nu_2\rangle,\quad \theta_4=1/9.\end{aligned}$$

The average cost finally reads

$$\overline{C}_1(x)=\mathrm{Tr}(\hat{W}^{(2)}-\Theta\hat{W}^{(0)}\Theta)=\frac{8}{81}\left[1+\left(x-\frac{1}{2}\right)^2\right].\qquad(23)$$

This is minimized by the maximally entangled state input

$$|\Psi\rangle=\frac{1}{\sqrt{2}}(|0\rangle\otimes|e_0\rangle+|1\rangle\otimes|e_1\rangle),\qquad(24)$$

for which $\theta_1=\frac{5}{9}$ and $\theta_2=\theta_3=\theta_4=\frac{1}{9}$. Therefore the optimal measurement is actually constructed by the two projectors:

$$\hat{\Pi}_1=|\Psi\rangle\langle\Psi|,\quad \hat{\Pi}_2=\hat{I}-|\Psi\rangle\langle\Psi|,\qquad(25)$$

with the associated guesses $\theta_1=\frac{5}{9}$ and $\theta_2=\frac{1}{9}$, respectively. The minimum average cost is $\overline{C}_{1\,\mathrm{min}}=\frac{8}{81}$.

B. Case (b)

The output state $\hat{\Psi}_2(\theta)=\hat{\psi}_2(\theta)\oplus\hat{\phi}_2(\theta)$ is given by

$$\hat{\psi}_2(\theta)=\begin{bmatrix}\frac{1}{4}-\left(\frac{1}{2}-x\right)\theta+\theta^2 & \theta^2\sqrt{x(1-x)}\\ \theta^2\sqrt{x(1-x)} & \frac{1}{4}+\left(\frac{1}{2}-x\right)\theta+\theta^2\end{bmatrix},\qquad(26)$$

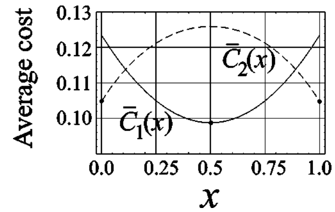

FIG. 1. The average costs as a function of x.

$$\hat{\phi}_2(\theta)=\frac{1-\theta^2}{4}\hat{I}_\phi.\qquad(27)$$

The elements of the risk operator are

$$\hat{W}^{(0)}=\frac{1}{27}\begin{bmatrix}4+9x & 7\sqrt{x(1-x)}\\ 7\sqrt{x(1-x)} & 13-9x\end{bmatrix}\oplus\frac{5}{27}\hat{I}_\phi,\qquad(28)$$

$$\hat{W}^{(1)}=\frac{1}{27}\begin{bmatrix}7x & 5\sqrt{x(1-x)}\\ 5\sqrt{x(1-x)} & 7(1-x)\end{bmatrix}\oplus\frac{1}{27}\hat{I}_\phi,\qquad(29)$$

$$\hat{W}^{(2)}=\frac{1}{405}\begin{bmatrix}4+75x & 61\sqrt{x(1-x)}\\ 61\sqrt{x(1-x)} & 79-75x\end{bmatrix}\oplus\frac{11}{405}\hat{I}_\phi.\qquad(30)$$

The minimizing operator is (see Appendix B)

$$\hat{\Theta}=\frac{1}{17[13+8x(1-x)]}\begin{bmatrix}a & c\\ c & b\end{bmatrix}\oplus\frac{1}{5}\hat{I}_\phi,\qquad(31)$$

where

$$\begin{aligned}a&=7x(35-20x+2x^2),\\ b&=7x(17+16x+2x^2),\qquad(32)\\ c&=9[9-2x(1-x)]\sqrt{x(1-x)}.\end{aligned}$$

To diagonalize it we use $r=\sqrt{(a-b)^2+4c^2}$ and

$$\cos\gamma=\sqrt{\frac{r+a-b}{2r}},\quad \sin\gamma=\sqrt{\frac{r-a+b}{2r}}.\qquad(33)$$

We then have the similar eigenstates to Eq. (22) and the eigenvalues $\theta_1=\theta_+$, $\theta_2=\theta_-$, and $\theta_3=\theta_4=\frac{1}{5}$ with

$$\theta_\pm=\frac{119[1+2x(1-x)]\pm r}{34[13+8x(1-x)]}.\qquad(34)$$

The average cost is then

$$\overline{C}_2(x)=\frac{8[391+606x(1-x)-10x^2(1-x)^2]}{2295[13+8x(1-x)]}.\qquad(35)$$

This is an upward convex function, symmetric with respect to $x=\frac{1}{2}$. The minimum is attained at $x=0,1$, that is, by separable input states. This reads $\overline{C}_{2\mathrm{min}}=\frac{184}{1755}$.

The average costs for cases (a) and (b) are shown in Fig. 1: $\overline{C}_1(x)$ (solid line) and $\overline{C}_2(x)$ (dashed line). We see that

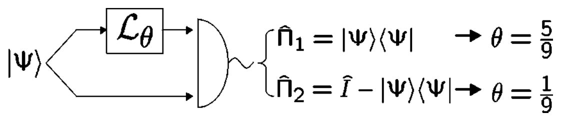

FIG. 2. The optimal estimation strategy using two probe qubits. $|\Psi\rangle$ is the maximally entangled state. The output state is projected onto $\{\hat{\Pi}_1, \hat{\Pi}_2\}$. We guess the channel parameter as $\theta=\frac{5}{9}$ for the outcome $\hat{\Pi}_1$ and $\theta=\frac{1}{9}$ otherwise.

$\overline{C}_{1\min}<\overline{C}_{2\min}$ so that the optimal estimation strategy, using two probe qubits, is to prepare them as a maximally entangled pair and to input one qubit of the pair into the channel keeping the other untouched. The estimation is then obtained by applying the two element POM, Eq. (25), as described in case (a). This strategy is represented schematically in Fig. 2.

When M maximally entangled pairs $|\Psi\rangle^{\otimes M}$ are available, it is best to use them so as to have the output $[(\mathcal{L}_\theta \otimes \hat{I})|\Psi\rangle\langle\Psi|]^{\otimes M}$. The optimal measurement for this can be derived straightforwardly. This is discussed in the next section as a part of an arbitrary finite dimensional case.

III. *d*-DIMENSIONAL CASE

The action of the depolarizing channel on a d-dimensional system is described by

$$\mathcal{L}_\theta \hat{\rho} = \theta \hat{\rho} + \frac{1-\theta}{d}\hat{I}. \tag{36}$$

Complete positivity then implies $-1/(d^2-1) \leqslant \theta \leqslant 1$. For $d \geqslant 3$, we have not succeeded in finding the optimal probe state, even when we restrict ourselves to a pure state. In this section we focus on the most plausible input state, that is, the maximally entangled state, and consider the estimation using M entangled pairs. Only for $d=2$, is the optimality ensured.

It might be interesting to compare the three cases specified by the three different outputs.

(a) M product states of the pair

$$\hat{\Psi}_1(\theta) = (\mathcal{L}_\theta \otimes \hat{I})|\Psi\rangle\langle\Psi| = \theta|\Psi\rangle\langle\Psi| + \frac{1-\theta}{d^2}\hat{I}\otimes\hat{I}, \tag{37}$$

where $|\Psi\rangle$ is the maximally entangled state.

(b) M product states of the pair

$$\hat{\Psi}_2(\theta) = (\mathcal{L}_\theta \otimes \mathcal{L}_\theta)|\Psi\rangle\langle\Psi| = \theta^2|\Psi\rangle\langle\Psi| + \frac{1-\theta^2}{d^2}\hat{I}\otimes\hat{I}. \tag{38}$$

(c) $2M$ product states of

$$\hat{\psi}(\theta) = \mathcal{L}_\theta|0\rangle\langle 0| = \theta|0\rangle\langle 0| + \frac{1-\theta}{d}\hat{I}. \tag{39}$$

[The input state in case (c) can be any pure state in the d-dimensional space.] Let us first consider the case (a). We denote Eq. (37) as

$$\hat{\Psi}(\theta) = f_0(\theta)\hat{a}_0 + f_1(\theta)\hat{a}_1, \tag{40}$$

where

$$\hat{a}_0 \equiv |\Psi\rangle\langle\Psi|, \quad \hat{a}_1 \equiv \hat{I} - |\Psi\rangle\langle\Psi|, \tag{41}$$

and

$$f_0(\theta) = \theta + \frac{1-\theta}{d^2}, \quad f_1(\theta) = \frac{1-\theta}{d^2}. \tag{42}$$

The output state can then be represented as

$$\hat{\Psi}(\theta)^{\otimes M} = \sum_{m=0}^{M} f_0(\theta)^{M-m} f_1(\theta)^m \hat{A}_m, \tag{43}$$

where

$$\hat{A}_m = \sum_{(i_1+\cdots i_M=m)} \hat{a}_{i_1} \otimes \cdots \otimes \hat{a}_{i_M} \tag{44}$$

is the projector onto the symmetric subspace. The risk operator is

$$\hat{W}(\theta) = \sum_{m=0}^{M} [\omega_m^{(2)} - 2\theta\omega_m^{(1)} + \theta^2\omega_m^{(0)}]\hat{A}_m, \tag{45}$$

where $\omega_m^{(k)} \equiv \int_{-(1/3)}^{1} d\theta\, \theta^k f_0(\theta)^{M-m} f_1(\theta)^m$. The optimal POM is

$$\hat{\Pi}(\theta) = \sum_{m=0}^{M} \delta(\theta-\theta_m)\hat{A}_m, \tag{46}$$

where

$$\theta_m \equiv \frac{\omega_m^{(1)}}{\omega_m^{(0)}}. \tag{47}$$

We then note that

$$\hat{W}(\theta) - \hat{\Gamma} = \sum_{m=0}^{M} (\theta-\theta_m)^2 \omega_m^{(0)} \hat{A}_m \geqslant 0, \tag{48}$$

from which it can easily be seen that the conditions (i) and (ii) hold. The minimum average cost is

$$\overline{C}_1(M) = \sum_{m=0}^{M} \left[\omega_m^{(2)} - \frac{(\omega_m^{(1)})^2}{\omega_m^{(0)}} \right] \binom{M}{m} (d^2-1)^m. \tag{49}$$

The other cases can be dealt with in a similar manner. In case (b), we just put

$$f_0(\theta) = \theta^2 + \frac{1-\theta^2}{d^2}, \quad f_1(\theta) = \frac{1-\theta^2}{d^2}. \tag{50}$$

The minimum average cost $\overline{C}_2(M)$ is then given by the same expression as Eq. (49) with $\omega_m^{(k)}$'s defined by $f_0(\theta)$ and $f_1(\theta)$ of Eq. (50).

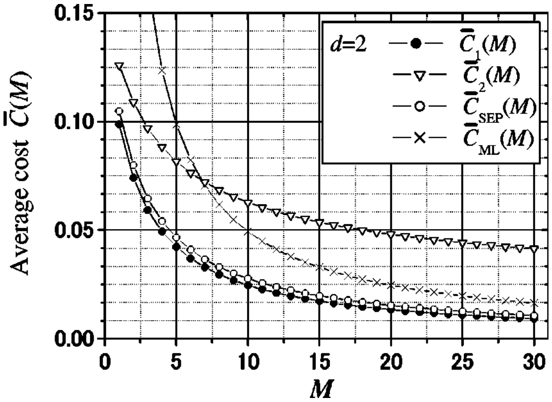

FIG. 3. The average costs as a function of the number of pairs.

In case (c) we use

$$\hat{a}_0 \equiv |0\rangle\langle 0|, \quad \hat{a}_1 \equiv \hat{I} - |0\rangle\langle 0|, \tag{51}$$

and

$$f_0(\theta) = \theta + \frac{1-\theta}{d}, \quad f_1(\theta) = \frac{1-\theta}{d}. \tag{52}$$

The minimum average cost is

$$\overline{C}_{\mathrm{SEP}}(M) = \sum_{m=0}^{2M} \left[\omega_m^{(2)} - \frac{(\omega_m^{(1)})^2}{\omega_m^{(0)}} \right] \binom{2M}{m} (d-1)^m. \tag{53}$$

The three costs $\overline{C}_1(M)$, $\overline{C}_2(M)$, and $\overline{C}_{\mathrm{SEP}}(M)$ are plotted in Fig. 3 $(d=2)$, Fig. 4 $(d=3)$, and Fig. 5 $(d=10)$. In the figures another average cost $\overline{C}_{\mathrm{ML}}(M)$ is also plotted. This cost is by the strategy belonging to case (c), but unlike the one attaining $\overline{C}_{\mathrm{SEP}}(M)$, the estimator is made by the maximum likelihood principle for which

$$\theta_m = \frac{md}{2M(d-1)}, \tag{54}$$

instead of Eq. (47), and leads to the analytic expression

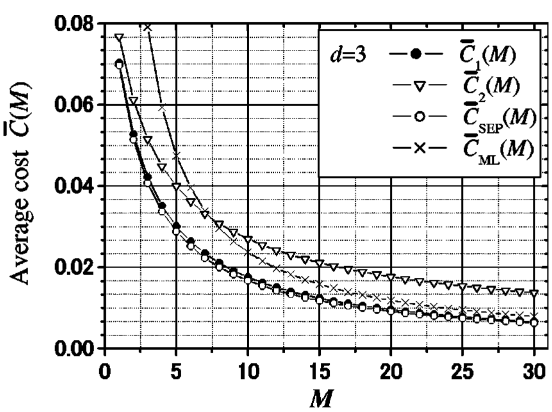

FIG. 4. The average costs as a function of the number of pairs.

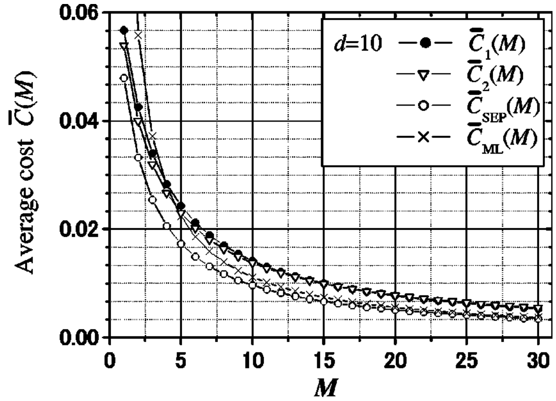

FIG. 5. The average costs as a function of the number of pairs.

$$\overline{C}_{\mathrm{ML}}(M) = \frac{1}{2M} \frac{d^5(d+3)}{6(d^2-1)^3}. \tag{55}$$

It is this strategy that was used in Ref. [4] for the case of $d=2$.

For $d \geqslant 3$, the minimum average cost is always attained by a separable probe state. Only in the two-dimensional case is it the bipartite entangled probe that attains the minimum average cost. It is worth mentioning the depolarizing channel with the narrower parameter region $0 \leqslant \theta \leqslant 1$, which is a more commonly used model with a well-defined interpretation of randomized *probability* of θ. We found that the best probe in this model is always a separable state. In this sense a separable state is generally an adequate probe state for the depolarizing channel estimation as far as the comparison with a bipartite entangled probe state is concerned.

IV. DISCUSSION

We have considered an estimation problem of the depolarizing channel on a d-dimensional system in a scenario where we have a finite number of probe particles, $2M$, at our disposal and we can prepare them in M pairs of the maximally entangled states if necessary. The depolarizing papparameter θ is assumed to be in a full range, $-1/(d^2-1) \leqslant \theta \leqslant 1$, allowed by the complete positivity of the channel action.

For $d \geqslant 3$, the best way is to input the probe particles in the separable state, then to apply the binary measurement $\{\hat{a}_0, \hat{a}_1\}$ [Eq. (51)] on each output particle separately, and finally to establish the estimate as one of the $2M+1$ candidates $\{\theta_m\}$ corresponding to the symmetric subspace spanned by $\{\hat{a}_0, \hat{a}_1\}$. Thus any bipartite entanglement is not necessary.

In contrast, for $d=2$, the best way is to prepare the M maximally entangled pairs, $|\Psi\rangle\langle\Psi|^{\otimes M}$, to input one qubit of each pair into the channel keeping the other untouched, and to apply the binary measurement $\{|\Psi\rangle\langle\Psi|, \hat{I} - |\Psi\rangle\langle\Psi|\}$ on each pair separately. The estimate consists of the $M+1$ candidates $\{\theta_m\}$ corresponding to the symmetric subspace spanned by $\{|\Psi\rangle\langle\Psi|, \hat{I} - |\Psi\rangle\langle\Psi|\}$. In conjunction with the analysis in Sec. II, we can show that this is the optimal estimation strategy for double optimization with respect to an input state and a measurement. If the channel parameter is restricted to the narrower region $0 \leqslant \theta \leqslant 1$ for which the interpretation of randomized *probability* of θ applies, the best probe state is again the separable state.

Thus an entanglement is useful only in the case of qubits. Even in this case, the amount of the cost reduction obtained by replacing the separable input probe with the entangled probe is rather small. As seen in Figs. 3–5, for higher dimension d, the strategies with the entangled probes become less effective (black circles and triangles) compared with the other separable probe strategies. In this sense an entanglement is not necessary for the estimation of the depolarizing channel. This is also the case for other decoherence channels, such as an amplitude damping channel and a dephasing channel. For an amplitude damping channel, for example, a probe particle in its ground state is insensitive to the damping, and the best probe state is the most highly excited state. An entangled probe is wasteful for the same reason, that is, this includes the state components which are less sensitive to the damping.

For the depolarizing channel, things may be more complicated. The superiority of the separable probe strategies to the entangled probe ones may be due to the fact that the former can take more estimate points in the parameter space than the latter. Given $2M$ probe particles at our disposal, the separable probe strategies have $2M+1$ outcomes while the entangled probe strategies have $M+1$ outcomes. More estimate points simply result in higher accuracy. The minimum required number of estimates is basically determined by the dimensionality of the particle. It is well known that in estimating unknown quantum states of a finite dimensional system, increasing the number of POM elements beyond a certain critical number never helps to increase the estimation accuracy, and this criterion increases as the dimensionality of the system. This means that for a higher dimensional system, a POM with more elements is preferable. Therefore the separable probe strategies are generally better in higher dimensional cases. For a lower dimentional system, the quantum nature of polarization dominates, and the minimum number of estimates becomes smaller while the ambiguity of estimation increases. For the qubit case, the difference in the number of outcomes between the separable and entangled probes is not a primary factor, but the other aspect influences the estimation accuracy. It should be noted that the parameter can take the negative value in $-\frac{1}{3} \leqslant \theta \leqslant 0$ in the case where the entangled probe is best. Negative values of θ have a different physical meaning to positive values. For negative values the Bloch vector of an input state shrinks reversing its direction, that is, the output state is more likely the opposite spin, while for positive values the Bloch vetcor shrinks in the same direction. If we are to discriminate the possible θ including all these cases, then in the optimal strategy, one qubit of the entangled probe pair that is not input into the channel can be a good reference for the spin flipping action, while in the other strategies [case (b) in Sec. II] where both qubits udergo the depolarizing channel, the ability to detect the spin flipping is weaker. This is why the optimal strategy for the channel with the parameter region $-\frac{1}{3} \leqslant \theta \leqslant 1$ is to use the maximally entangled pairs with one qubit of each pair untouched to the channel and why the optimal strategy for the channel with the narrower region $0 \leqslant \theta \leqslant 1$ is the separable probe strategy.

It is worth noting that the collective measurement on M identical output pairs or $2M$ identical output particles is not necessary. The action of the depolarizing channel on a maximally entangled state always results in a statistical mixture between the input state and its orthogonal complement [Eq. (40)]. Estimating the channel parameter is nothing but determining this mixing ratio, which is a *classical* distribution. Therefore the optimal measurement is realized by a separable type constructed by the binary orthogonal projectors $\{\hat{a}_0, \hat{a}_1\}$ according to Eq. (44). In the case where the output state includes the channel parameter as a quantum distribution, that is, the parameter appears in the off diagonal components in the density matrix, the optimal measurement would be a collective measurement. Such channels include a unitary channel. It might then be desirable to apply a single multiqubit state as the probe and the best *collective* measurement on the whole system. Channel estimation for such cases is a future problem.

Finally it might also be interesting to study the multiparameter case, such as the Pauli channel estimation. We may then ask how to optimize (in Bayesian sense) the simultaneous measurement on the noncommuting observables as well as searching for appropriate probe states.

ACKNOWLEDGMENTS

We are grateful to K. Usami, Dr. Y. Tsuda, and Dr. K. Matsumoto for helpful discussions. This work was supported, in part, by the British Council, the Royal Society of Edinburgh, and by the Scottish Executive Education and Lifelong Learning Department.

APPENDIX A: DERIVATION OF EQ. (20)

For obtaining the minimizing operator Θ in Eq. (20), we first diagonalize $\hat{W}^{(0)}$ by $\hat{U}_0 = \hat{u}_0 \oplus \hat{I}_\phi$ where

$$\hat{u}_0 = \begin{bmatrix} \cos\gamma_0 & -\sin\gamma_0 \\ \sin\gamma_0 & \cos\gamma_0 \end{bmatrix}, \tag{A1}$$

with $r_0 = \sqrt{1-3x(1-x)}$ and

$$\cos\gamma_0 = \sqrt{\frac{r_0 - 1 + 2x}{2r_0}}, \quad \sin\gamma_0 = \sqrt{\frac{r_0 + 1 - 2x}{2r_0}}. \tag{A2}$$

The spectral decomposition

$$\hat{W}^{(0)} = \sum_{i=1}^{4} \omega_i |\omega_i\rangle\langle\omega_i| \tag{A3}$$

is given by

$$
\begin{aligned}
|\omega_1\rangle &= \hat{u}_0|\mu_1\rangle, \quad \omega_1=(1+r_0)/3,\\
|\omega_2\rangle &= \hat{u}_0|\mu_2\rangle, \quad \omega_2=(1-r_0)/3,\\
|\omega_3\rangle &= |\nu_1\rangle, \quad \omega_3=(1-x)/3,\\
|\omega_4\rangle &= |\nu_2\rangle, \quad \omega_4=x/3.
\end{aligned} \tag{A4}
$$

We then calculate

$$\tilde{\Theta}=\sum_{i,j=1}^{4}\frac{2}{\omega_i+\omega_j}|i\rangle\langle\omega_i|\hat{W}^{(1)}|\omega_j\rangle\langle j|, \tag{A5}$$

where

$$
\begin{aligned}
|1\rangle &= |\mu_1\rangle,\\
|2\rangle &= |\mu_2\rangle,\\
|3\rangle &= |\nu_1\rangle,\\
|4\rangle &= |\nu_2\rangle.
\end{aligned} \tag{A6}
$$

This gives

$$\tilde{\Theta}=\tilde{\Theta}_{\psi}\oplus\frac{1}{9}\hat{I}_{\phi}, \tag{A7}$$

where

$$\tilde{\Theta}_{\psi}=\frac{1}{9}\begin{bmatrix}\dfrac{4r_0(1+r_0)+3x(1-x)}{r_0(1+r_0)} & -\dfrac{3(1-2x)\sqrt{x(1-x)}}{r_0}\\ -\dfrac{3(1-2x)\sqrt{x(1-x)}}{r_0} & \dfrac{4r_0(1-r_0)-3x(1-x)}{r_0(1-r_0)}\end{bmatrix}. \tag{A8}$$

The minimizing operator is given by $\hat{\Theta}=\hat{U}_0\tilde{\Theta}\hat{U}_0^{\dagger}$ which results in Eq. (20).

APPENDIX B: DERIVATION OF EQ. (31)

The unitary operator for diagonalizing $\hat{W}^{(0)}$ in Eq. (28) is $\hat{U}_0=\hat{u}_0\oplus\hat{I}_{\phi}$ where

$$\hat{u}_0=\begin{bmatrix}\cos\gamma_0 & -\sin\gamma_0\\ \sin\gamma_0 & \cos\gamma_0\end{bmatrix}, \tag{B1}$$

with $r_0=\sqrt{81-128x(1-x)}$ and

$$\cos\gamma_0=\sqrt{\frac{r_0-9(1-2x)}{2r_0}}, \quad \sin\gamma_0=\sqrt{\frac{r_0+9(1-2x)}{2r_0}}. \tag{B2}$$

The spectral decomposition

$$\hat{W}^{(0)}=\sum_{i=1}^{4}\omega_i|\omega_i\rangle\langle\omega_i| \tag{B3}$$

is given by

$$
\begin{aligned}
|\omega_1\rangle &= \hat{u}_0|\mu_1\rangle, \quad \omega_1=(17+r_0)/54,\\
|\omega_2\rangle &= \hat{u}_0|\mu_2\rangle, \quad \omega_2=(17-r_0)/54,\\
|\omega_3\rangle &= |\nu_1\rangle, \quad \omega_3=5/27,\\
|\omega_4\rangle &= |\nu_2\rangle, \quad \omega_4=5/27.
\end{aligned} \tag{B4}
$$

We then have

$$\tilde{\Theta}=\tilde{\Theta}_{\psi}\oplus\frac{1}{5}\hat{I}_{\phi}, \tag{B5}$$

where

$$\tilde{\Theta}_{\psi}=\begin{bmatrix}\dfrac{7[r_0+9-16x(1-x)]}{r_0(17+r_0)} & \dfrac{8(1-2x)\sqrt{x(1-x)}}{17r_0}\\ \dfrac{8(1-2x)\sqrt{x(1-x)}}{17r_0} & \dfrac{7[r_0-9+16x(1-x)]}{r_0(17-r_0)}\end{bmatrix}. \tag{B6}$$

Substituting this to $\hat{\Theta}=\hat{U}_0\tilde{\Theta}\hat{U}_0^{\dagger}$, we have Eq. (31).

[1] A. Acín, E. Jané, and G. Vidal, Phys. Rev. A **64**, 050302(R) (2001).
[2] A. Fujiwara, Phys. Rev. A **63**, 042304 (2001).
[3] D.G. Fischer, H. Mack, M.A. Cirone, and M. Freyberger, Phys. Rev. A **64**, 022309 (2001).
[4] M.A. Cirone, A. Delgado, D.G. Fischer, M. Freyberger, H. Mack, and M. Mussinger, e-print quant-ph/0108037.
[5] O.E. Barndorff-Nielsen and R.D. Gill, e-print quant-ph/9808009.
[6] R.D. Gill and S. Massar, Phys. Rev. A **61**, 042312 (2000).

[7] M. Hayashi, e-print quant-ph/0202003.
[8] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[9] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
[10] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic, Dordrecht, 1993), p. 279.
[11] A.S. Holevo, J. Multivariate Anal. **3**, 337 (1973).
[12] H.P. Yuen, R.S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **IT-21**, 125 (1975).
[13] S.D. Personick, IEEE Trans. Inf. Theory **IT-17**, 240 (1971).