Qubit channels that achieve capacity with two states

Dominic W. Berry
Department of Physics, The University of Queensland, Brisbane, Queensland 4072, Australia
(Received 22 December 2004; published 22 March 2005)

This paper considers a class of qubit channels for which three states are always sufficient to achieve the Holevo capacity. For these channels, it is known that there are cases where two orthogonal states are sufficient, two nonorthogonal states are required, or three states are necessary. Here a systematic theory is given which provides criteria to distinguish cases where two states are sufficient, and determine whether these two states should be orthogonal or nonorthogonal. In addition, we prove a theorem on the form of the optimal ensemble when three states are required, and present efficient methods of calculating the Holevo capacity.

DOI: 10.1103/PhysRevA.71.032334 PACS number(s): 03.67.Hk

I. INTRODUCTION

A quantum channel is a completely positive and trace-preserving (CPTP) map on quantum states. The condition that it is completely positive means that the result of the map is a positive operator, and therefore may represent the state of a system, even if the map acts on one part of an entangled system. The condition that it is trace-preserving ensures that the final state is normalized. In contrast to unitary operations, quantum channels can increase the entropy of a state. A quantum channel arises if an ancilla space is added, a unitary operation is performed between the system and the ancilla, and then the ancilla is traced over to obtain the reduced density operator for the system.

Quantum channels are used to model communication channels, and therefore an important quantity to consider for these channels is the amount of classical communication that may be performed. This is often quantified by the Holevo capacity. The Holevo capacity of a quantum channel Φ is given by

$$C(\Phi) = \sup_{p_i,\rho_i} S[\Phi(\overline{\rho})] - \sum_i p_i S[\Phi(\rho_i)], \tag{1}$$

where $\overline{\rho}=\Sigma_i p_i\rho_i$, and $S(\sigma)=-\mathrm{Tr}\,\sigma\log_2\sigma$ is the von Neumann entropy. The p_i are probabilities, and therefore must be non-negative and sum to 1. The Holevo capacity is the asymptotic classical communication that may be achieved using joint measurements on output states, but unentangled inputs [1,2]. In general, determining the Holevo capacity of a channel is a nontrivial task. For the class of channels considered here, it will be shown that the capacity may be determined in a straightforward way.

An important issue is the number of states ρ_i that must be considered in the maximization. It is well known that, for quantum channels that act upon a Hilbert space of dimension d, the number of states in the ensemble need not exceed d^2 [3]. In particular, for a qubit channel no more than four states are required. For the very simple case of unital qubit channels, where $\Phi(\mathbb{1})=\mathbb{1}$, the capacity is achieved for two orthogonal input states [4]. For more general qubit channels, the capacity may be achieved for two nonorthogonal inputs [5]; three states [6] or four states may be required [7].

With the exception of the channels considered in Ref. [7], these results are all for a class of channels that can require at most three states. Here we give simple criteria for these channels that, when satisfied, mean that two states are sufficient. These criteria are not satisfied by the channels that require three states given in [6], but are satisfied by examples given in Refs. [4–6,8] where two states are sufficient. In addition, we give criteria to determine when the input states should be orthogonal or nonorthogonal.

This paper is organized as follows. We present the proof of the criteria in Sec. II. Then, in Sec. III we give applications of the result to results presented in previous work. We consider the form of the optimal ensembles for those cases where three states are required in Sec. IV. In Sec. V, we show how our results may be applied to the calculation of the Holevo capacity. Conclusions are given in Sec. VI.

II. TWO-STATE ENSEMBLES

To obtain the results, we use the representation of the qubit channel on the Bloch sphere. A general qubit density operator may be expressed as

$$\rho = \frac{1}{2}(\mathbb{1} + \vec{r}\cdot\vec{\sigma}), \tag{2}$$

where $\vec{\sigma}$ is the vector of Pauli operators $(\sigma_x,\sigma_y,\sigma_z)^{\mathrm{T}}$. The length of the vector $\vec{r}$ does not exceed 1, and its components give the position of the state in the Bloch sphere. A qubit channel Φ maps the sphere of possible input states to an ellipsoid, and may be expressed as

$$\Phi(\rho) = \frac{1}{2}[\mathbb{1} + (\Lambda\vec{r}+\vec{t})\cdot\vec{\sigma}]. \tag{3}$$

That is, the channel Φ produces the mapping $\vec{r}\mapsto\Lambda\vec{r}+\vec{t}$. Via local unitary operations before and after the map, the transformation matrices Λ and $\vec{t}$ may be brought to the form [4]

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad \vec{t} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}. \tag{4}$$

That is, an arbitrary qubit channel Φ may be expressed as $\Phi=\Gamma_U\circ\Phi_{t,\Lambda}\circ\Gamma_V$, where Γ_U and Γ_V are unitary channels, and

$\Phi_{t,\Lambda}$ is the channel with Λ and $\vec{t}$ given by Eqs. (4). For this study, we consider the restricted case of channels Φ such that the x and y components of $\vec{t}$ are zero, and use the notation $t=t_3$. Hence $\vec{t}$ is given by

$$\vec{t}=\begin{pmatrix}0\\0\\t\end{pmatrix}. \tag{5}$$

In order to evaluate the Holevo capacity, we use an approach similar to that of Ref. [8]. The Holevo capacity may be given by the following expression [9,10]:

$$C(\Phi) = \min_{\psi_0}\max_{\rho_0} D\big[\Phi(\rho_0)\,\|\,\Phi(\psi_0)\big], \tag{6}$$

where D is the relative entropy

$$D(\rho\,\|\,\psi) = \mathrm{Tr}(\rho\log\rho - \rho\log\psi). \tag{7}$$

Throughout this paper, we use the convention that "log" and "exp" are base 2, and logarithms base e are given as "ln." The relative entropy can be evaluated using the following useful result from [8]:

$$D(\rho\,\|\,\psi) = \frac{1}{2}[f(r) - \log(1-q^2) - r\cos(\theta)f'(q)], \tag{8}$$

where

$$f(x) = (1+x)\log(1+x) + (1-x)\log(1-x), \tag{9}$$

$$f'(x) = \log\left(\frac{1+x}{1-x}\right). \tag{10}$$

The Bloch vectors for ρ and ψ are $\vec{r}$ and $\vec{q}$, respectively, and we also define $r=|\vec{r}|$, $q=|\vec{q}|$, $\cos(\theta)=\vec{r}\cdot\vec{q}/rq$.

To evaluate the Holevo capacity, we consider the action of the simplified channel $\Phi_{t,\Lambda}$. This channel has the same capacity as Φ, because unitary operations do not affect the capacity. The set of possible output states from the channel $\Phi_{t,\Lambda}$ forms an ellipsoid centered on the z axis. The ellipsoid has a radius of $|\lambda_1|$ in the x direction, and a radius of $|\lambda_2|$ in the y direction.

The nature of the optimal ensemble may be determined by considering the states in the minmax formula (6). In the following, we take the states $\rho=\Phi_{t,\Lambda}(\rho_0)$ and $\psi=\Phi_{t,\Lambda}(\psi_0)$ to be output states from the simplified channel. If ψ is the average output density operator for an optimal ensemble, the operators ρ_k that maximize $D(\rho_k\|\psi)$ are possible output states for this ensemble. It is necessary that there is some set of p_k such that $\Sigma_k p_k\rho_k=\psi$. The optimal ensemble is not necessarily unique, because there may be different ways of choosing the probabilities such that $\Sigma_k p_k\rho_k=\psi$. However, from Ref. [8], the optimal average output state is unique.

As we are restricting to operations such that $\vec{t}$ lies on the z axis, there are many simplifications due to the symmetry of the system. Many of these simplifications were used in Ref. [8] in the analysis of the amplitude damping channel. We give a general explanation here. First, the optimal state ψ must lie on the z axis. To show this result, for any pair of states ρ and ψ, consider the second pair ρ' and ψ', where $\vec{r}'=(-r_x,-r_y,r_z)^{\mathrm{T}}$ and $\vec{q}'=(-q_x,-q_y,q_z)^{\mathrm{T}}$. Due to symmetry, if ρ and ψ are possible output states, then so are ρ' and ψ'. From the symmetry of the relative entropy, it is evident that $D(\rho\|\psi)=D(\rho'\|\psi')$. This immediately implies that $\max_\rho D(\rho\|\psi)=\max_\rho D(\rho\|\psi')$. Therefore, if ψ minimizes this quantity, then so does ψ'. However, as the optimal average output state is unique, ψ and ψ' must coincide, which implies that ψ lies on the z axis.

In the case that $|\lambda_1|\neq|\lambda_2|$, the ρ_k that maximize the relative entropy will lie in the x-z plane if $|\lambda_1|>|\lambda_2|$, and the y-z plane if $|\lambda_1|<|\lambda_2|$. That is because ψ lies on the z axis, so the relative entropy is symmetric under rotation about the z axis. If $|\lambda_1|>|\lambda_2|$, then the ellipsoid has a radius in the x direction larger than the radius in the y direction. Consider any state ρ that is not in the x-z plane. We can determine a second state ρ' in the x-z plane with Bloch vector $\vec{r}'=(\sqrt{r_x^2+r_y^2},0,r_z)^{\mathrm{T}}$. This state is in the interior of the ellipsoid, and we may obtain a third state on the surface of the ellipsoid, ρ'', by extending outwards in a straight line from ψ. From Ref. [8] (the first lemma in Sec. 5.3),

$$D(\rho''\,\|\,\psi) > D(\rho'\,\|\,\psi) = D(\rho\,\|\,\psi). \tag{11}$$

This implies that ρ does not maximize the relative entropy. Hence, all ρ_k that maximize the relative entropy must be in the x-z plane. Similarly, if $|\lambda_1|<|\lambda_2|$, the ellipsoid has a radius in the y direction larger than the radius in the x direction, and the optimal ρ_k must be in the y-z plane.

In the case that $|\lambda_1|=|\lambda_2|$, the situation is a little more complicated. For each optimal ρ_k, there is a circle of optimal density operators around the z axis. However, in order to obtain an optimal ensemble, it is only necessary to use nonzero probabilities such that $\Sigma_k p_k\rho_k=\psi$. As ψ lies on the z axis, it is sufficient to take ρ_k from a single plane in the Bloch sphere that contains the z axis.

This reasoning means that, regardless of the relative values of $|\lambda_1|$ and $|\lambda_2|$, we may restrict to considering ρ_k that maximize $D(\rho_k\|\psi)$ in a single plane in the Bloch sphere. Caratheodory's theorem implies that there need be no more than three states in the ensemble. This fact was also noted in Ref. [6]. The examples given by Ref. [7] which needed four states used $\vec{t}$ that were not on the z axis.

In fact, in some cases the number of states required is only two [5], though in some cases three are required. Here we give criteria that can show when only two states are required via the following theorem.

Theorem 1. For a CPTP map $\Phi=\Gamma_U\circ\Phi_{t,\Lambda}\circ\Gamma_V$ with Λ given by Eq. (4) and $\vec{t}$ given by Eq. (5), if $\lambda_m=|\lambda_3|$ or $A\notin(0,1/2)$, where

$$A = \frac{t^2\lambda_3^2}{\lambda_m^2-\lambda_3^2} - 1 + \lambda_m^2 + t^2 \tag{12}$$

and $\lambda_m=\max(|\lambda_1|,|\lambda_2|)$, then there is an ensemble that gives the maximum output Holevo information and has two states.

Before we proceed to the proof, we give some explanation of the quantity A. Let us consider the output ellipse in the x-z plane if $|\lambda_1|\geq|\lambda_2|$, or the y-z plane if $|\lambda_1|<|\lambda_2|$. A point on the surface of this ellipse has a distance from the origin r, which is given by Eq. (16) in the proof below. Taking the derivative of r^2 with respect to ϕ gives

$$\frac{d^2}{d\phi^2}(r^2) = 2\sin\phi\,[(\lambda_m^2-\lambda_3^2)\cos\phi - \lambda_3 t]. \tag{13}$$

This expression is zero if $\sin\phi=0$, $\lambda_m^2-\lambda_3^2=\lambda_3 t=0$, or

$$\cos\phi = \frac{\lambda_3 t}{\lambda_m^2-\lambda_3^2}. \tag{14}$$

The third case is only possible if the absolute value of the right-hand side (RHS) does not exceed 1. If it does not, then substituting this expression for $\cos\phi$ into the expression for r gives the extremum

$$r_{\mathrm{ex}}^2 = \frac{t^2\lambda_3^2}{\lambda_m^2-\lambda_3^2} + \lambda_m^2 + t^2 = A+1. \tag{15}$$

Therefore, in this case, A is the difference between the square of an extremum of r and 1. In the case $\lambda_m^2-\lambda_3^2=\lambda_3 t=0$, the radius is independent of ϕ. This possibility will be excluded in the discussion of A, because $\lambda_m=|\lambda_3|$ is an alternative criterion to $A\notin(0,1/2)$, and leads to infinite A.

If A were positive, then r_{ex}^2 would be larger than 1, which is not possible for CPTP maps. Therefore, for any map such that an extremum of r is obtained for $\sin\phi\neq 0$ (and $\lambda_m\neq|\lambda_3|$), the condition $A\notin(0,1/2)$ is automatically satisfied due to the fact that states cannot be mapped outside the Bloch sphere. However, $A\notin(0,1/2)$ is not satisfied for every possible CPTP map, because for some $|\lambda_3 t/(\lambda_m^2-\lambda_3^2)|>1$.

Another case where $A\notin(0,1/2)$ is automatically satisfied is when $\lambda_m<|\lambda_3|$. That is because the condition that the map is CPTP implies that $\lambda_m^2+t^2\leqslant 1$, and if $\lambda_m<|\lambda_3|$, then $t^2\lambda_3^2/(\lambda_m^2-\lambda_3^2)$ is negative. Therefore, from the definition of A, it is clear that $A\leqslant 0$. We now proceed to the proof of the theorem.

Proof. We begin the analysis by mentioning some trivial cases that would otherwise complicate the analysis. If $t=0$, then the channel is unital, and the result in this case was proven in Ref. [4]. If all three of the λ_k are zero, then the channel capacity is zero, and the result is trivial. If two of the λ_k are zero, then the possible output states form a line in the Bloch sphere, and the result follows from the fact that there are only two extremal output states.

The result is also trivial if $\lambda_3=0$. In that case, since we may restrict ourselves to considering states in the x-z or y-z plane, the set of output states that it is sufficient to consider forms a line. The result again follows from the fact that there are only two extremal states. For the remainder of the analysis, we take $t\neq 0$, $\lambda_3\neq 0$, and assume that no more than one of the λ_k is zero. This third assumption means that $\lambda_m\neq 0$.

For the remainder of this proof, we consider the input and output states for the simplified channel $\Phi_{t,\Lambda}$. The input and output states for the total channel Φ will simply be rotated from these states. We take the input state to have $\vec{r}=(\sin\phi,0,\cos\phi)^{\mathrm{T}}$ for $|\lambda_1|\geqslant|\lambda_2|$, or $\vec{r}=(0,\sin\phi,\cos\phi)^{\mathrm{T}}$ for $|\lambda_1|<|\lambda_2|$. The output state will then have $\vec{r}=(\lambda_1\sin\phi,0,t+\lambda_3\cos\phi)^{\mathrm{T}}$ or $\vec{r}=(0,\lambda_2\sin\phi,t+\lambda_3\cos\phi)^{\mathrm{T}}$. The state ψ has $\vec{q}=(0,0,q_z)^{\mathrm{T}}$. In either case, we have for the output

$$r=\sqrt{\lambda_m^2\sin^2\phi+(t+\lambda_3\cos\phi)^2},$$
$$r\cos\theta=(t+\lambda_3\cos\phi)\,\mathrm{sgn}(q_z). \tag{16}$$

To search for the optimal ρ, it is merely necessary to search for the optimal ϕ. Because $\mathrm{sgn}(q_z)f'(q)=f'(q_z)$, we may write the relative entropy as

$$D(\rho\,\|\,\psi)=\frac{1}{2}[f(r)-\log(1-q_z)-(t+\lambda_3\cos\phi)f'(q_z)]. \tag{17}$$

The derivative of $D(\rho\|\psi)$ with respect to ϕ is

$$\begin{aligned}\frac{d}{d\phi}D(\rho\,\|\,\psi) &= \frac{1}{2}\left\{\frac{dr}{d\phi}f'(r)-f'(q_z)\frac{d}{d\phi}(t+\lambda_3\cos\phi)\right\}\\ &= \frac{1}{2}\{[(\lambda_m^2-\lambda_3^2)\cos\phi-t\lambda_3]f'(r)/r\\ &\quad + f'(q_z)\lambda_3\}\sin\phi.\end{aligned} \tag{18}$$

There will be extrema of $D(\rho\|\psi)$ for $\phi=0$ and $\phi=\pi$, as well as when

$$[(\lambda_m^2-\lambda_3^2)\cos\phi-t\lambda_3]f'(r)/r=-f'(q_z)\lambda_3. \tag{19}$$

We will consider the solutions of this equation for ϕ in the interval $(0,\pi)$. Any solution in $(0,\pi)$ will yield a corresponding solution in $(-\pi,0)$ due to symmetry.

Taking the derivative of the left-hand side (LHS) gives

$$\begin{aligned}&\frac{d}{d\phi}[(\lambda_m^2-\lambda_3^2)\cos\phi-t\lambda_3]f'(r)/r\\ &= \left\{-(\lambda_m^2-\lambda_3^2)\frac{f'(r)}{r}+[(\lambda_m^2-\lambda_3^2)\cos\phi-t\lambda_3]^2\right.\\ &\quad\left.\times\frac{1}{r}\frac{d}{dr}\left(\frac{f'(r)}{r}\right)\right\}\sin\phi.\end{aligned} \tag{20}$$

In the case that $|\lambda_m|\neq|\lambda_3|$,

$$[(\lambda_m^2-\lambda_3^2)\cos\phi-t\lambda_3]^2=(\lambda_m^2-\lambda_3^2)(1-r^2+A). \tag{21}$$

We then obtain

$$\begin{aligned}&\frac{d}{d\phi}[(\lambda_m^2-\lambda_3^2)\cos\phi-t\lambda_3]f'(r)/r\\ &= \frac{(\lambda_m^2-\lambda_3^2)\sin\phi}{r}[h(r)+Ag(r)],\end{aligned} \tag{22}$$

where

$$g(r)=\frac{d}{dr}\left(\frac{f'(r)}{r}\right)=\frac{2}{(1-r^2)r\ln 2}-\frac{1}{r^2}\log\left(\frac{1+r}{1-r}\right), \tag{23}$$

$$h(r)=\frac{2}{r}-\frac{f'(r)}{r^2}=\frac{2}{r\ln 2}-\frac{1}{r^2}\log\left(\frac{1+r}{1-r}\right). \tag{24}$$

The functions $g(r)$ and $h(r)$ satisfy the inequalities

$$g(r) > 0, \quad h(r) < 0, \quad 2h(r) + g(r) > 0, \tag{25}$$

for $r \in (0,1)$. If $A \leq 0$, then $h(r)+Ag(r)$ is negative for $r \in (0,1)$. Similarly, if $A \geq 1/2$, then $h(r)+Ag(r)$ is positive for $r \in (0,1)$. In either case, $h(r)+Ag(r)$ has constant sign. We do not need to consider the possibility that $r=0$, because this value is only possible when $\sin\phi=0$ (for $\lambda_m \neq 0$).

The case where $r=1$ is more complicated. It is possible for r to be equal to 1 for $\phi \in (0,\pi)$. In the case where r has a maximum for $\phi \in (0,\pi)$, the maximum value of r is $A+1$. If r is equal to 1 for $\phi \in (0,\pi)$, this must be a maximum, and therefore $A=0$ (as we are taking $\lambda_m \neq |\lambda_3|$). That implies that the expression in square brackets on the LHS of Eq. (19) is proportional to $\sqrt{1-r^2}$. Hence the LHS of Eq. (19) approaches zero as r approaches 1, and is continuous as a function of ϕ for $\phi \in (0,\pi)$. As $h(r)+Ag(r)$ has constant sign for all values of $\phi \in (0,\pi)$ except where $r=1$, and the LHS of Eq. (19) is continuous where $r=1$, the LHS of Eq. (19) is one-to-one in this interval.

For the case $\lambda_m = |\lambda_3|$,

$$\frac{d}{d\phi}[(\lambda_m^2 - \lambda_3^2)\cos\phi - t\lambda_3]f'(r)/r = t^2\lambda_3^2(\sin\phi)g(r)/r. \tag{26}$$

Therefore, the derivative of the LHS of Eq. (19) is nonzero for $\phi \in (0,\pi)$. Note that we are assuming that $t \neq 0$ and $\lambda_3 \neq 0$, so the RHS of Eq. (26) is nonzero. Thus we have shown that, regardless of the relative values of λ_m and λ_3, the LHS of Eq. (19) is a one-to-one function of ϕ, and there can be at most one solution of Eq. (19) in $(0,\pi)$. If there is a solution, it must correspond to an extremum, because a point of inflection would conflict with the fact that the LHS of Eq. (19) is one-to-one.

As $D(\rho \| \psi)$ is symmetric about $\phi=0$, there must be two solutions of Eq. (19) with $\sin\phi \neq 0$ or none. In the case where there are no solutions, there are only two extrema (for $\phi=0$ and π), and only one of these can be a maximum. This is not consistent with ψ being optimal, because the optimal ensemble cannot have only one state. Therefore, if ψ is optimal, then there must be two solutions of Eq. (19). As the maxima and minima alternate, the maxima are either at $\phi=0$ and π, or the solutions of Eq. (19).

In the case that $|\lambda_1| \neq |\lambda_2|$, this result immediately implies that there are only two states in the optimal ensemble. In the case $|\lambda_1| = |\lambda_2|$, if the maxima correspond to the solutions of Eq. (19), optimal ensembles may contain any states in a ring about the z axis. However, as discussed above, it is only necessary to consider ρ_k in one plane in the Bloch sphere in this case, so there is again an optimal ensemble with two members. □

It is also possible to determine simple criteria for when the optimal states in the ensemble are on the z axis, and when the optimal states in the ensemble correspond to the maxima for $\sin\phi \neq 0$. The result is as follows.

Theorem 2. Let $\Phi_{t,\Lambda}$ be a CPTP map with $\Lambda \neq 0$ given by Eq. (4) and $\vec{t}$ given by Eq. (5). The condition that $\lambda_m = |\lambda_3|$ or $A \notin (0,1/2)$ may be expressed as two alternative mutually exclusive conditions:

Condition 1. $\lambda_m \leq |\lambda_3|$ or $A \geq 1/2$.

Condition 2. $\lambda_m > |\lambda_3|$ and $A \leq 0$.

If Condition 1 is satisfied, the optimal ensemble consists of two states on the z axis. If Condition 2 is satisfied, there is an optimal ensemble consisting of two states equidistant from the z axis and lying on a line perpendicular to and intersecting the z axis.

Here we have given the result in terms of the simplified map $\Phi_{t,\Lambda}$, rather than expressing it in terms of the arbitrary map Φ. That is because the ellipse of output states will be rotated for the arbitrary map, so it is not possible to express the result in this way. The statement of this theorem also differs in that Λ is taken to be nonzero. This is to exclude the trivial case where all ensembles give zero Holevo information.

Proof. As was shown above, $\lambda_m < |\lambda_3|$ also implies that $A \leq 0$. Another consequence of this is that, if $A > 0$, then $\lambda_m > |\lambda_3|$. Therefore Condition 1 contains three alternatives:

(i) $\lambda_m = |\lambda_3|$.

(ii) $\lambda_m < |\lambda_3|$ and $A \leq 0$.

(iii) $A \geq 1/2$ and $\lambda_m > |\lambda_3|$.

It is clear that, for each of these three alternatives, the conditions of Theorem 1 must hold. If none of these alternatives apply, but $A \notin (0,1/2)$, then $\lambda_m > |\lambda_3|$ and $A \leq 0$, which is Condition 2 given in the theorem.

To determine which extrema of $D(\rho \| \psi)$ are maxima and which are minima, it is sufficient to consider the point $\phi=0$. At this point, the second derivative of $D(\rho \| \psi)$ is given by

$$\frac{d^2}{d\phi^2}D(\rho \| \psi) = \frac{1}{2}\{[(\lambda_m^2 - \lambda_3^2) - t\lambda_3]f'(r)/r + f'(q_z)\lambda_3\}. \tag{27}$$

We know that the LHS of Eq. (19) is one-to-one, and there must be at least one solution of Eq. (19) if ψ is optimal (otherwise there would be only one possible state for the ensemble).

If $\lambda_m = |\lambda_3|$, then from Eq. (26), the LHS of Eq. (19) is monotonically increasing for $\phi \in (0,\pi)$. If $A \geq 1/2$ and $\lambda_m > |\lambda_3|$, then $h(r)+Ag(r) > 0$, and from Eq. (22) the LHS of Eq. (19) is monotonically increasing. Similarly, if $\lambda_m < |\lambda_3|$ and $A \leq 0$, then $h(r)+Ag(r) < 0$, and the LHS of Eq. (19) is again monotonically increasing. Therefore, for all three alternatives for Condition 1, the LHS of Eq. (19) is monotonically increasing for $\phi \in (0,\pi)$. For Condition 2, $\lambda_m > |\lambda_3|$ and $A \leq 0$, so $h(r)+Ag(r) < 0$, and the LHS of Eq. (19) is monotonically decreasing for $\phi \in (0,\pi)$.

If the LHS of Eq. (19) is monotonically increasing for $\phi \in (0,\pi)$, the LHS of Eq. (19) must be less than the RHS for $\phi=0$, so

$$[(\lambda_m^2 - \lambda_3^2) - t\lambda_3]f'(r)/r + f'(q_z)\lambda_3 < 0. \tag{28}$$

This means that the second derivative of $D(\rho \| \psi)$ is negative for $\phi=0$, and $D(\rho \| \psi)$ is a maximum at this point. Hence, the two maxima are obtained for $\phi=0$ and π, and these values correspond to the states in the optimal ensemble. Thus we see that, for Condition 1, the LHS of Eq. (19) is monotonically increasing and the optimal ensemble consists of two states on the z axis.

Alternatively, for Condition 2, the LHS of Eq. (19) is monotonically decreasing, so the LHS of Eq. (19) is greater than the RHS for $\phi=0$, and less for $\phi=\pi$. This implies that the second derivative of $D(\rho\|\psi)$ is positive for $\phi=0$ and $\phi=\pi$, and these points are minima. Hence, in this case the states in the optimal ensemble correspond to the extrema of $D(\rho\|\psi)$ for $\sin\phi\neq 0$.

In the case that $|\lambda_1|>|\lambda_2|$ or $|\lambda_1|<|\lambda_2|$, the optimal ensemble must be in the x-z plane or y-z plane, respectively. In either case, two maxima are obtained in the appropriate plane for $\phi=\pm\phi_0$, where ϕ_0 maximizes $D(\rho\|\psi)$. These two solutions are equidistant from the z axis, and on a line perpendicular to and intersecting the z axis. If $|\lambda_1|=|\lambda_2|$, then there will be a circle of states about the z axis that maximize the relative entropy. Optimal ensembles may contain any number of these states. However, as discussed above, we may restrict ourselves to states in one plane. This yields an ensemble with two members that again lie on a line perpendicular to and intersecting the z axis. □

Another issue is the position of the optimal average output state. It is possible to use similar techniques as above to show that this state should be further from the center of the Bloch sphere than the output for the maximally mixed state. Specifically, q_z for the optimal average output state should satisfy $q_z/t>1$ for t and λ_3 both nonzero. The case $t=0$ means that the map is unital, and it is known in that case that $q_z=0$ is optimal. If $\lambda_3=0$, then clearly $q_z=t$.

To show this result, let us assume some value for q_z (the other components of $\vec{q}$ are zero), and take a value of ϕ such that $|t+\lambda_3\cos\phi|>|t-\lambda_3\cos\phi|$. We denote the states with $r_z=t\pm\lambda_3\cos\phi$ by $\rho_\pm$. Determining the difference in relative entropies gives

$$\begin{aligned}D(\rho_+\|\psi)-D(\rho_-\|\psi)&=f(r_+)-f(r_-)-2\lambda_3 f'(q_z)\cos\phi\\&>f'(\overline{r})(r_+-r_-)-2\lambda_3 f'(q_z)\cos\phi,\end{aligned}\tag{29}$$

where $r_\pm$ is the magnitude of the Bloch vector for $\rho_\pm$, and $\overline{r}=(r_++r_-)/2$. In the second line, we have used the strict convexity of $f'(r)$ and the Hermite-Hadamard inequality [11]. Now using the fact that $r_+^2-r_-^2=4t\lambda_3\cos\phi$, we have $r_+-r_-=(2t\lambda_3\cos\phi)/\overline{r}$. Therefore, Eq. (29) simplifies to

$$D(\rho_+\|\psi)-D(\rho_-\|\psi)>2t\lambda_3\cos\phi\,[f'(\overline{r})/\overline{r}-f'(q_z)/t].\tag{30}$$

We have chosen ϕ such that $t\lambda_3$ is positive, and both $f'(x)$ and $f'(x)/x$ are monotonically increasing functions. Also $\overline{r}\geq t$, with equality only if $\lambda_m\sin\phi=0$. Therefore, $q_z/t\leq 1$ implies that

$$D(\rho_+\|\psi)-D(\rho_-\|\psi)>0.\tag{31}$$

This means that, if t is positive and $q_z\leq t$, then all states ρ_- that have z component of their Bloch vector less than t do not maximize the relative entropy. In addition, if $q_z=t$, the relative entropy cannot be maximized for $r_z=t$. In the case $\lambda_m=0$ this is trivial, because the maxima are for $r_z=t+\lambda_3$ and $r_z=t-\lambda_3$. If $\lambda_m\neq 0$, then $f'(r)/r>f'(t)/t$. As we are also taking $\lambda_3\neq 0$, this inequality means that Eq. (19) cannot be satisfied for $\phi=\pi/2$.

Hence, for $q_z\leq t>0$ and $\lambda_3\neq 0$, all ρ_k that maximize the relative entropy must have a z component of their Bloch vector greater than that for ψ, and they cannot give an average equal to ψ. This is not consistent with ψ being the average state for the optimal ensemble, and therefore the average state for the optimal ensemble must satisfy $q_z>t$. Similarly, if t is negative and $\lambda_3\neq 0$, then the average state for the optimal ensemble satisfies $q_z<t$.

With the aid of this result, we can alternatively express Theorem 2 in terms of the orthogonality of the input states. The result is as follows.

Corollary 1. Consider a CPTP map $\Phi=\Gamma_U\circ\Phi_{t,\Lambda}\circ\Gamma_V$ with $\Lambda\neq 0$ given by Eq. (4) and $\vec{t}$ given by Eq. (5). The condition that $\lambda_m=|\lambda_3|$ or $A\notin(0,1/2)$ may be expressed as two alternative mutually exclusive conditions:

Condition 1. $\lambda_m\leq|\lambda_3|$ or $A\geq 1/2$.

Condition 2. $\lambda_m>|\lambda_3|$ and $A\leq 0$.

If $t\neq 0$ and $\lambda_3\neq 0$, the maximum output Holevo information is obtained for two orthogonal input states if Condition 1 is satisfied, and two nonorthogonal input states if Condition 2 is satisfied.

Proof. Note first that unitary operations do not change the orthogonality relations between the states. Therefore, it is sufficient to prove the orthogonality relations for the simplified map $\Phi_{t,\Lambda}$. For Condition 1, the result follows immediately from Theorem 2. The two input states are the extremal states on the z axis, and therefore are $|0\rangle$ and $|1\rangle$, which are orthogonal.

To prove the result for Condition 2, we use the result that, for $t\neq 0$ and $\lambda_3\neq 0$, q_z is not equal to t. If the input states for Condition 2 were orthogonal, then that would lead to $q_z=t$. Therefore, if $t\neq 0$ and $\lambda_3\neq 0$, the input states must be nonorthogonal if Condition 2 holds. □

III. APPLICATIONS

These results allow us to make sense of the results obtained in previous work. In particular, [8] found that only two states in the ensemble were required for the amplitude damping channel, where $\lambda_1=\lambda_2=\sqrt{\mu}$, $\lambda_3=\mu$, and $t=1-\mu$. We find that, in this case, $A=0$, so $A\notin(0,1/2)$ is satisfied and Theorem 1 predicts that the optimal ensemble requires two states. For this channel, $\lambda_m>|\lambda_3|$ and $A\leq 0$, which corresponds to Condition 2 in Theorem 2. Theorem 2 therefore predicts that, for this channel, the optimal ensemble consists of two states at the same distance from the x-y plane, rather than on the z axis. This is what was found in Ref. [8].

Another channel is the shifted depolarizing channel, which was considered in Ref. [6]. For this channel, $\lambda_k=\mu$ and $t=1-\mu$. As $\lambda_m=\lambda_3$, Theorem 1 applies, and the ensemble should require only two states. This result is what was found in [6]. Also, because $\lambda_m=\lambda_3$, Condition 1 in Theorem 2 holds, so Theorem 2 predicts that the states in the optimal ensemble lie on the z axis. This is also consistent with the results of Ref. [6].

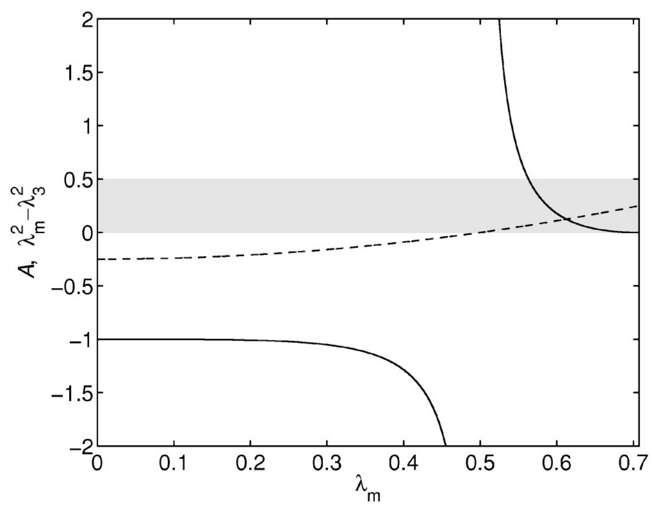

FIG. 1. The values of A (solid line) and $\lambda_m^2-\lambda_3^2$ (dashed line) as a function of λ_m for $\lambda_3=t=1/2$. The shaded region shows the region of values of A such that the optimal ensemble may require three states. Results for $\lambda_m>1/\sqrt{2}$ are not shown, because the maps for $\lambda_m>1/\sqrt{2}$ are not CPTP.

On the other hand, let us consider the examples given in [6] that require three states. For one of these examples, $\lambda_1=\lambda_2=0.6$ and $\lambda_3=t=0.5$, so $A\approx 0.178$. This is in the interval $(0,1/2)$, so it is not surprising that three states are required. Another example is $\lambda_1=t=0.5$ and $\lambda_2=\lambda_3=0.435$; in this case A is about 0.278, which is again in the interval $(0,1/2)$.

In Ref. [6], a strategy used to find channels that require three states was to vary the parameters from a channel such that the optimal states are on the z axis to one where the optimal states are away from the z axis. This strategy can alternatively be explained in terms of Theorem 2. The channel parameters cannot be continuously varied from Condition 1 to Condition 2 without A passing through the interval $(0,1/2)$. That is because it is not possible to continuously vary the channel parameters from $\lambda_m<|\lambda_3|$ to $\lambda_m>|\lambda_3|$ while maintaining the same sign for A.

To take an example from [6], let $\lambda_3=t=1/2$, and vary λ_m. Then the variation of A and $\lambda_m^2-\lambda_3^2$ are as in Fig. 1. It can be seen from this figure that as $\lambda_m^2-\lambda_3^2$ passes through zero, A switches from negative to positive. In fact, the only point where Condition 2 is satisfied is for $\lambda_m=1/\sqrt{2}$. In passing from $\lambda_m=0.5$, where $\lambda_m=\lambda_3$, to $\lambda_m=1/\sqrt{2}$, the value of A passes through $(0,1/2)$.

A case of particular interest is that in which Λ and $\vec{t}$ are given by

$$\Lambda=\begin{pmatrix}\cos\delta & 0 & 0\\ 0 & \cos\gamma & 0\\ 0 & 0 & \cos\gamma\cos\delta\end{pmatrix},\quad \vec{t}=\begin{pmatrix}0\\ 0\\ \sin\gamma\sin\delta\end{pmatrix}. \tag{32}$$

This type of channel arises naturally when considering qubit interactions. If one introduces an ancilla qubit, performs a unitary operation, then traces over this ancilla qubit, the resulting operation is of this form [12]. Maps of this form also arise naturally when considering extremal maps [13]. Also, it is known that all qubit maps with two Kraus operators are of this form [13].

For maps of this form, we find that $A=0$, so the conditions of Theorem 1 are satisfied. Therefore, for maps that arise from a unitary interaction with an ancilla qubit, the optimal ensemble requires only two states. This result was also claimed in Ref. [14], although the complete proof was not given. In addition, $|\lambda_3|<\lambda_m$, so from Theorem 2 the two states for the optimal ensemble are away from the z axis.

IV. THREE-STATE ENSEMBLES

In the case where three states are required for the optimal ensemble, it is possible to show that one of the states needs to be on the z axis. The result is as follows.

Theorem 3. Consider a CPTP map $\Phi_{t,\Lambda}$ with Λ given by Eq. (4) and $\vec{t}$ given by Eq. (5). If the Holevo capacity cannot be achieved with a two-state ensemble, then any optimal ensemble with three states consists of one state on the z axis, and two states equidistant from the z axis and on a line perpendicular to and intersecting the z axis. The optimal input state on the z axis is $|0\rangle$ if $|t+\lambda_3|>|t-\lambda_3|$, and $|1\rangle$ if $|t+\lambda_3|<|t-\lambda_3|$.

Proof. In order to prove the result, we start by considering the expression in square brackets in Eq. (22). Although $h(r)+Ag(r)$ can change sign, it is only zero for one value of r. To show this result, we use the following facts:

$$h(r)<0,\quad g(r)>0,\quad g'(r)>0, \tag{33}$$

$$h'(r)g(r)-g'(r)h(r)>0. \tag{34}$$

These inequalities are all for $r\in(0,1)$, and are easily checked by plotting the functions. If $h(r)+Ag(r)\geq 0$ for $r=r_0$, then $A\geq -h(r_0)/g(r_0)$, so $h'(r_0)+Ag'(r_0)\geq[h'(r_0)g(r_0)-g'(r_0)h(r_0)]/g(r_0)>0$. Therefore, if $h(r)+Ag(r)\geq 0$ for $r=r_0$, then $h(r)+Ag(r)$ is increasing for $r=r_0$. This implies that, if there is a value of r for which $h(r)+Ag(r)=0$, then $h(r)+Ag(r)>0$ for all larger values of r. Hence $h(r)+Ag(r)$ can be zero for only one value of r in (0, 1).

Recall that, if there is an extremum of r for $\sin\phi\neq 0$, then the condition $A\neq(0,1/2)$ is satisfied, and therefore the optimal ensemble requires no more than two states. In the conditions for Theorem 3, the optimal ensemble requires more than two states, so r has no extremum for $\sin\phi\neq 0$. Hence r is a one-to-one function for ϕ in the interval $(0,\pi)$. Combining this result with the above reasoning, the RHS of Eq. (22) can be zero for only one value of ϕ in the interval $(0,\pi)$.

These results imply that the LHS of Eq. (19) can have a turning point for only one value of ϕ in $(0,\pi)$, and therefore there are at most two solutions of Eq. (19) for $\phi\in(0,\pi)$. In turn, this implies that there are no more than two extrema of $D(\rho\|\psi)$ for $\phi\in(0,\pi)$. In fact, there must be exactly two (if ψ is optimal), because if there were only one, then the optimal ensemble would require only two states, which violates the conditions of Theorem 3.

Thus there will be two extrema of $D(\rho\|\psi)$ for $\phi\in(0,\pi)$, two symmetric extrema for $\phi\in(-\pi,0)$, and extrema at $\phi=0$ and π. These extrema must alternate between minima and maxima, and so one of the extrema at $\phi=0$ and π will be a maximum, and the other will be a minimum. To determine which points are minima and which are maxima, consider the second derivative of $D(\rho\|\psi)$ at a solution of Eq. (19),

$$\frac{d^2}{d\phi^2}D(\rho\,\|\,\psi) = \frac{(\lambda_m^2-\lambda_3^2)\sin^2\phi}{2r}[h(r)+Ag(r)]. \tag{35}$$

Recall that, if $h(r)+Ag(r)$ is positive for $r=r_0$, it must also be positive for $r>r_0$. Therefore, for the solution of Eq. (19) with smaller r, $h(r)+Ag(r)$ is negative, and for the solution with larger r, $h(r)+Ag(r)$ is positive.

For maps that require three states to achieve the Holevo capacity, $A>0$. As discussed above, this implies that $\lambda_m>|\lambda_3|$, so $\lambda_m^2-\lambda_3^2$ is positive. Thus multiplication by $\lambda_m^2-\lambda_3^2$ does not change the sign, so the solution of Eq. (19) with smaller r is a maximum, and the solution with larger r is a minimum. As the extrema alternate between maxima and minima, the extremum on the z axis that is closer to the origin must be a minimum. Therefore, if $|t+\lambda_3|$ is greater than $|t-\lambda_3|$, then the optimal output state on the z axis will be at $t+\lambda_3$. This corresponds to an input state of $|0\rangle$. Similarly, if $|t-\lambda_3|$ is greater than $|t+\lambda_3|$, then the optimal output state on the z axis is at $t-\lambda_3$, which corresponds to the input state $|1\rangle$.

The two remaining states in the optimal ensemble will correspond to solutions $\phi=\pm\phi_0$ of Eq. (19). In the case that $|\lambda_1|\neq|\lambda_2|$, these states are in the x-z or y-z plane of the Bloch sphere, depending on whether $|\lambda_1|>|\lambda_2|$ or $|\lambda_1|<|\lambda_2|$. In either case, the states are equidistant from the z axis, on a line that is perpendicular to and intersecting the z axis. If $|\lambda_1|=|\lambda_2|$, then optimal ensembles may contain any states from a circle about the z axis. However, for optimal ensembles with three states, the condition that the mean state is on the z axis restricts the remaining two states to be equidistant from the z axis, and on a line perpendicular to and intersecting the z axis. □

V. CALCULATING CAPACITIES

These results enable us to determine numerically efficient ways of calculating capacities. In the case that the channel satisfies the conditions of Theorem 1, the problem becomes particularly simple. First it is necessary to check whether it is Condition 1 or Condition 2 in Theorem 2 that is satisfied. For Condition 1, the optimal ensemble consists of the two extremal states on the z axis. The probabilities may be determined by the fact that $D(\rho_1\|\psi)=D(\rho_2\|\psi)$. The expression for the relative entropy (8) simplifies to

$$D(\rho\,\|\,\psi) = \frac{1}{2}[f(r_z) - \log(1-q_z^2) - r_z f'(q_z)]. \tag{36}$$

The condition that $D(\rho_1\|\psi)=D(\rho_2\|\psi)$ then becomes

$$f(t+\lambda_3) - (t+\lambda_3)f'(q_z) = f(t-\lambda_3) - (t-\lambda_3)f'(q_z). \tag{37}$$

This may be solved for q_z, yielding

$$q_z = \frac{X-1}{X+1}, \tag{38}$$

where

$$X = \exp\left[\frac{f(t+\lambda_3)-f(t-\lambda_3)}{2\lambda_3}\right]. \tag{39}$$

Recall that we are using notation where "exp" means 2 to the power of the argument. The channel capacity is obtained by substituting Eq. (38) into Eq. (36). Thus the channel capacity may be obtained analytically. The optimal ensemble may also be determined analytically. The optimal states correspond to points on the z axis at $t\pm\lambda_3$, and the probabilities are given by

$$p_\pm = \frac{1}{2} \pm \frac{q_z-t}{2\lambda_3}. \tag{40}$$

For Condition 2 in Theorem 2, the optimal states are away from the z axis. Because ψ must be the average of the two ρ_k, and the z components of the two $\vec{r}_k$ are equal, the z component of $\vec{q}$ must also be equal. If ψ is optimal, for the solution of Eq. (19) the z component of r should be equal to the z component of q. Therefore, the optimal ensemble may be found by finding the solution of Eq. (19) with $q_z=r_z$. Thus finding the capacity in this case reduces to finding the zero of a function of a single real variable, which is easily performed numerically.

As an alternative interpretation of this result, consider the ensemble consisting of two states corresponding to $\phi=\pm\phi_0$. The Holevo information of this ensemble is given by

$$D(\rho_\pm\,\|\,\psi) = \frac{1}{2}[f(r) - f(r_z)], \tag{41}$$

where ψ is the average state. If the optimal ensemble is of this form, then the maximum of this quantity gives the Holevo capacity for the channel. Taking the derivative with respect to ϕ, we find that the maximum will be for a solution of Eq. (19) with $q_z=r_z$.

For the case where Λ and $\vec{t}$ are as given in Eq. (32), the problem of calculating the capacity has been considered in Ref. [15]. For this case, this reference gives an analytic method for calculating the Holevo capacity for a given mean state. Although this method was derived in quite a different way from the method given here, it is equivalent.

In those cases where $A\in(0,1/2)$, it is still possible that two states may be sufficient for the optimal ensemble. In those cases, the ensemble must still consist of either two states on the z axis of the Bloch sphere, or two states corresponding to $\phi=\pm\phi_0$, where ϕ_0 is a root of Eq. (19). This result may be shown by considering $D(\rho\|\psi)$ as a function of ϕ. As was shown in the previous section, there can be at most three maxima of $D(\rho\|\psi)$. If there are only two, then these are at $\phi=0$ and π or $\phi=\pm\phi_0$. In either case, the form of the optimal ensemble is the same as for channels satisfying the conditions of Theorem 1.

If there are three maxima, then one of these is on the z axis, and the other two are for $\phi=\pm\phi_0$. If two states are sufficient for the optimal ensemble, these states must correspond to $\phi=\pm\phi_0$, because otherwise ψ would not be on the z axis. Therefore, regardless of whether there are two maxima or three, if two states are sufficient for the optimal ensemble, then these consist of either two states on the z axis, or two states corresponding to $\phi=\pm\phi_0$.

These results can be used to determine if the optimal ensemble requires three states in cases where $A\in(0,1/2)$. From the "sufficiency of maximal distance property" in [9], we know that the ensemble is optimal if there are no values of ρ that give values of $D(\rho\|\psi)$ greater than the ρ_k in the ensemble. Therefore, in order to determine if the ensemble requires more than two states, determine ψ via the two different methods above. If, for one of them, $D(\rho\|\psi)$ is maximized for the corresponding ρ_k, then the optimal ensemble requires only two states. If neither of these methods gives the optimal ensemble, then we have eliminated all possibilities for optimal two-state ensembles, and the optimal ensemble must require three states.

It is also possible to efficiently determine the Holevo capacity in those cases where the ensemble requires three states. The reason for this is that the only unknowns for the three-state ensemble are the value of ϕ_0 such that $\phi=\pm\phi_0$ for the two off-axis states, and the probabilities for the three states. Given the value of ϕ_0, there is an analytic method to determine the probabilities. Therefore, the problem reduces to a numerical maximization in a single real variable, which is easily performed.

From Theorem 3, the state on the z axis will be at $t+\lambda_3$ if $|t+\lambda_3|>|t-\lambda_3|$, and $t-\lambda_3$ if $|t+\lambda_3|<|t-\lambda_3|$. Taking the other two states to correspond to $\phi=\pm\phi_0$, the condition that the relative entropy $D(\rho_k\|\psi)$ is independent of k becomes

$$f(t\pm\lambda_3)-(t\pm\lambda_3)f'(q_z)=f(r_0)-(t+\lambda_3\cos\phi_0)f'(q_z), \tag{42}$$

where $r_0^2=\lambda_m^2\sin^2\phi_0+(t+\lambda_3\cos\phi_0)^2$. We take the plus sign if $|t+\lambda_3|>|t-\lambda_3|$, and the minus sign if $|t+\lambda_3|<|t-\lambda_3|$. Solving for q_z gives

$$q_z=\frac{X-1}{X+1}, \tag{43}$$

where

$$X=\exp\left[\frac{f(t\pm\lambda_3)-f(r_0)}{\lambda_3(\pm1-\cos\phi_0)}\right]. \tag{44}$$

Note that this solution is reasonable only if the value of q_z obtained is between $t\pm\lambda_3$ and $t+\lambda_3\cos\phi_0$; otherwise, negative probabilities would be required for the ensemble.

Given this solution for q_z, the common value of the relative entropy is given by

$$D(\rho_k\|\psi)=\frac{1}{2}[f(t\pm\lambda_3)-\log(1-q_z^2)-(t\pm\lambda_3)\log X]. \tag{45}$$

By finding the maximum of this (with q_z between $t\pm\lambda_3$ and $t+\lambda_3\cos\phi_0$), the Holevo capacity may be determined.

This method was used to determine the difference between the two-state capacity and the three-state capacity for a range of different maps. This difference is plotted as a function of A in Fig. 2. In addition, the states that maximize this difference were searched for numerically for given values of A; these results are also shown in Fig. 2. It can be seen that the maximum difference in the capacities is still quite small, less than 0.004. Also, the difference can be nonzero in the entire interval $(0,1/2)$. The difference approaches zero quite rapidly as A approaches $1/2$, but is still nonzero. For comparison, two of the examples from Ref. [6] are shown in Fig. 2. It was also found that, regardless of the value of A, there were cases where two states were sufficient for the optimal ensemble.

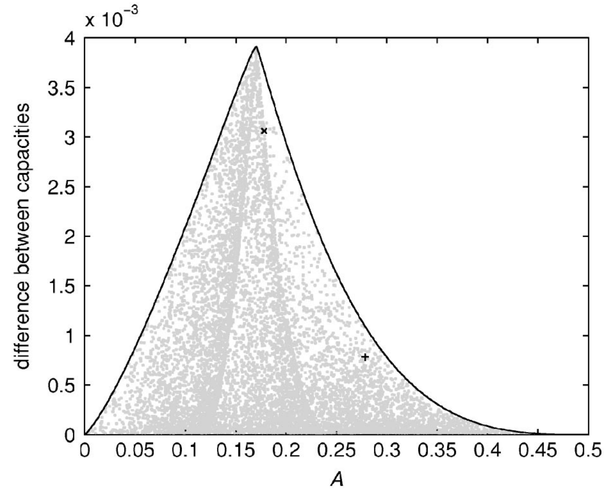

FIG. 2. The difference between the two-state capacity and the three-state capacity vs the value of A. Random samples are shown as gray points, and the numerically obtained upper bound is shown as the solid line. The cross and plus are examples from Ref. [6]. The cross is for $\lambda_1=\lambda_2=0.6$ and $\lambda_3=t=0.5$, and the plus is for $\lambda_1=t=0.5$ and $\lambda_2=\lambda_3=0.435$.

VI. CONCLUSIONS

We have shown a number of results on the form of optimal ensembles for qubit channels. The class of channels considered includes those that can be simplified, via unitary operations before and after the channel, to a form that is symmetric under reflections in the x-z and y-z planes. This class includes extremal channels, and most examples of channels considered in previously published work. For these channels, we have introduced the parameter A, which can be interpreted in some cases in terms of the distance between the output ellipsoid and the unit sphere.

The main result is that if A is not in the interval $(0,1/2)$, then two states are sufficient for the ensemble that maximizes the Holevo capacity. In addition, optimal two-state ensembles must consist of either two states on the z axis of the Bloch sphere, or two states on a line that is perpendicular to and intersecting the z axis. For cases where $A\notin(0,1/2)$, we have presented a simple method to determine which form the optimal ensemble takes. This result also enables us to determine if the input states should be orthogonal or nonorthogonal. Even in cases where $A\in(0,1/2)$, if two states are sufficient for the optimal ensemble, then the ensemble must take one of these two forms.

For cases where three states are necessary for the optimal ensemble, our results show that the optimal three-state ensemble consists of one state on the z axis at the maximum distance from the origin, and two states on a line perpendicular to and intersecting the z axis. This demonstrates that the form of the optimal three-state ensembles found in Ref. [6] is universal.

Last, we have provided a computationally efficient method of determining the Holevo capacity. For cases where the optimal ensemble consists of two states on the z axis, the capacity may be determined analytically. For other cases, the calculation is a numerical maximization of a function of a single real variable, which is easily performed. For the specific case of extremal channels, this method is equivalent to that given in Ref. [15].

ACKNOWLEDGMENTS

This project has been supported by the Australian Research Council and the University of Queensland. The author is grateful for helpful comments from Barry Sanders.

[1] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
[2] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
[3] E. B. Davies, IEEE Trans. Inf. Theory **24**, 596 (1978).
[4] C. King and M. B. Ruskai, IEEE Trans. Inf. Theory **47**, 192 (2001).
[5] C. A. Fuchs, Phys. Rev. Lett. **79**, 1162 (1997).
[6] C. King, M. Nathanson, and M. B. Ruskai, Phys. Rev. Lett. **88**, 057901 (2002).
[7] M. Hayashi, H. Imai, K. Matsumoto, M. B. Ruskai, and T. Shimono, Quantum Inf. Comput. **5**, 13 (2005).
[8] J. Cortese, e-print quant-ph/0207128.
[9] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **63**, 022308 (2001).
[10] M. Ohya, D. Petz, and N. Watanabe, Probab. Math. Stat. **17**, 179 (1997).
[11] J. Hadamard, J. Math. Pures Appl. **58**, 171 (1893); D. S. Mitrinović and I. B. Lacković, Aequ. Math. **28**, 229 (1985).
[12] C.-S. Niu and R. B. Griffiths, Phys. Rev. A **60**, 2764 (1999).
[13] M. B. Ruskai, S. Szarek, and E. Werner, Linear Algebr. Appl. **347**, 159 (2002).
[14] F. Verstraete and H. Verschelde, e-print quant-ph/0202124.
[15] A. Uhlmann, J. Phys. A **34**, 7047 (2001).