An Analysis of Completely-Positive Trace-Preserving Maps on $\mathcal{M}_2$

Mary Beth Ruskai *
Department of Mathematics
University of Massachusetts Lowell
Lowell, MA 01854 USA
bruskai@cs.uml.edu

Stanislaw Szarek †
Department of Mathematics
Case Western Reserve University
Cleveland, OH 44106 USA
sjs13@po.cwru.edu
and
Equipe d'Analyse, Boite 186, Université Paris VI
4, Place Jussieu, F-75252 Paris, France
szarek@ccr.jussieu.fr

Elisabeth Werner ‡
Department of Mathematics
Case Western Reserve University
Cleveland, OH 44106 USA
emw2@po.cwru.edu
and
Université de Lille 1, UFR de Mathématique
F-59655 Villeneuve d'Ascq, France

7 August 2001

Abstract

We give a useful new characterization of the set of all completely positive, trace-preserving maps $\Phi : \mathcal{M}_2 \to \mathcal{M}_2$ from which one can easily check any trace-preserving map for complete positivity. We also determine explicitly all extreme points of this set, and give a useful parameterization after reduction to a certain canonical form. This allows a detailed examination of an important class of non-unital extreme points that can be characterized as having exactly two images on the Bloch sphere.

We also discuss a number of related issues about the images and the geometry of the set of stochastic maps, and show that any stochastic map on $\mathcal{M}_2$ can be written as a convex combination of two "generalized" extreme points.

Key Words: Completely positive maps, stochastic maps, quantum communication, noisy channels, Bloch sphere, states, .

MR Classification. 47L07, 81P68, 15A99, 46L30, 46L60, 94A40

Contents

*Partially supported by the National Security Agency and Advanced Research and Development Activity under Army Research Office contract DAAG55-98-1-0374 and by the National Science Foundation under grants DMS-9706981 and DMS-0074566.

†Partially supported by a grant from the National Science Foundation.

‡Partially supported by a grant from the National Science Foundation and by a NATO Collaborative Linkage Grant

1 Introduction

1.1 Background

Completely positive, trace-preserving maps arise naturally in quantum information theory and other situations in which one wishes to restrict attention to a quantum system that should properly be considered a subsystem of a larger system with which it interacts. In such situations, the system of interest is described by a Hilbert space $\mathcal{H}_1$ and the larger system by a tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. States correspond to density matrices, where a density matrix ρ is a positive semi-definite operator on $\mathcal{H}_1$ with $\mathrm{Tr}\rho = 1$. The result of the "noisy" interaction with the larger system is described by a *stochastic* map $\Phi : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_1)$ that takes ρ to another density matrix $\Phi(\rho)$. Since $\Phi(\rho)$ should also be a density matrix, Φ must be both trace-preserving and positivity preserving. However, the latter is not enough, since $\Phi(\rho)$ is the result of a positivity-preserving process on the larger space of operators in $\mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2) = \mathcal{B}(\mathcal{H}_1) \otimes \mathcal{B}(\mathcal{H}_2)$. This is precisely the condition that Φ be completely positive.

The notion of complete positivity was introduced in the more general context of linear maps $\Phi : \mathcal{A}_1 \to \mathcal{A}_2$ on C^*-algebras where the condition can be stated as the requirement that $\Phi \otimes I_{\mathcal{M}_n} : \mathcal{A}_1 \otimes \mathcal{M}_n \to \mathcal{A}_2 \otimes \mathcal{M}_n$ is positivity preserving for all n, where $\mathcal{M}_n$ denotes the algebra of complex $n \times n$ matrices. Stinespring showed that a completely positive map always has a representation of the form $\pi_2[\Phi(A)] = V^*\pi_1(A)V$ where π_1 and π_2 are representations of the algebras $\mathcal{A}_1, \mathcal{A}_2$ respectively. Kraus [15, 16] and Choi [4] showed that this leads to the more tangible condition that there exists a set of operators A_k such that

$$\Phi(\rho) = \sum_k A_k^\dagger \rho A_k. \tag{1}$$

(where we henceforth follow the physics convention of using $\dagger$ to denote the adjoint of an operator.) The condition that Φ is trace-preserving can then be written as $\sum_k A_k A_k^\dagger = I$. When this condition is also satisfied, (1) can also be used [16, 19] to find a representation of Φ in terms of a partial trace on a larger space.

The operators in (1) are often referred to as "Kraus operators" because of his influential work [15, 16] on quantum measurement in which he emphasized the role of completely positive maps. Recognition that such maps play a natural role in the description of quantum systems goes back at least to Haag and Kastler [7]. It is worth noting that this representation is highly non-unique, and that this non-uniqueness can not be eliminated by simple constraints. In particular, one can find extreme maps with at least two two different Kraus representations each of which uses only the minimal number of Kraus operators. An example is given in Section 3.4. The representation (1) was obtained independently by Choi [4] in connection with important tests, upon which this paper is based, for complete-positivity and extremality in the case of maps on $\mathcal{M}_n$.

However, Choi's condition and all of the representations discussed above require, at least implicitly, consideration of the map $\Phi \otimes I_{\mathcal{M}_n}$ on a larger space. (In quantum information theory this problem is sometimes avoided by defining a stochastic map, or channel, in terms of its Kraus operators as in (1); however, this approach has thus tended to focus attention on a rather limited set of channels.) One would like to find a simple way to characterize completely positive maps in terms of their action on the algebra $\mathcal{B}(\mathcal{H}_1)$ of the subsystem. The purpose of this paper is to obtain such a characterization in the special case of trace-preserving maps on $\mathcal{B}(\mathcal{H}_1) = \mathcal{M}_2$. This leads to a complete description of their extreme points, and a useful parameterization of the stochastic maps and their extreme points. Although the two-dimensional case $\mathcal{H}_1 = \mathbf{C}^2$ may seem rather special, it is of considerable importance because of its role in quantum computation and quantum communication.

If, in addition to being trace-preserving, a completely positive map Φ is unital, i.e., $\Phi(I) = I$, we call Φ *bistochastic*. This terminology for maps that are both unital and stochastic was introduced in [2].

1.2 Notation

First, we note that for a linear map Φ, its adjoint, which we denote $\widehat{\Phi}$ (to avoid confusion with the operator adjoint of a specific image) can be defined with respect to the Hilbert-Schmidt inner product $\langle A, B\rangle = \mathrm{Tr} A^\dagger B$ so that $\mathrm{Tr}\,[\Phi(A)]^\dagger B = \mathrm{Tr}\, A^\dagger \widehat{\Phi}(B)$. It is easy to verify that the Kraus operators for $\widehat{\Phi}$ are the adjoints of those for Φ so that (1) implies $\widehat{\Phi}(\rho) = \sum_k A_k \rho A_k^\dagger$, and that Φ is trace-preserving if and only if $\widehat{\Phi}(I) = I$, i.e., if $\widehat{\Phi}$ is unital.

In order to state our results in a useful form, we recall that the identity and Pauli matrices $\{I, \sigma_x, \sigma_y, \sigma_z\}$ form a basis for $\mathcal{M}_2$ where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Every density matrix can be written in this basis as $\rho = \frac{1}{2}[I + \mathbf{w} \cdot \sigma]$ with $\mathbf{w} \in \mathbf{R}^3$ and $|\mathbf{w}| \leq 1$. Thus, the set of density matrices, which we shall denote by $\mathcal{D}$, can be identified with the unit ball in $\mathbf{R}^3$ and the pure states (rank one projections) lie on the surface known as the "Bloch sphere." Since Φ is a linear map on $\mathcal{M}_2$, it can also be represented in this basis by a unique 4×4 matrix $\mathbf{T}$, and Φ is trace-preserving if and only if the first row satisfies $t_{1k} = \delta_{1k}$, i.e., $\mathbf{T} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{t} & \mathrm{T} \end{pmatrix}$ where T is a 3×3 matrix (and $\mathbf{0}$ and $\mathbf{t}$ are row and column vectors respectively) so that

$$\Phi(w_0 I + \mathbf{w} \cdot \sigma) = w_0 I + (\mathbf{t} + \mathrm{T}\mathbf{w}) \cdot \sigma. \tag{2}$$

The $\mathbf{T}$-matrix corresponding to $\widehat{\Phi}$ is $\mathbf{T}^\dagger$.

When Φ is also positivity-preserving, it maps the subspace of self-adjoint matrices in $\mathcal{M}_2$ into itself, which implies that $\mathbf{T}$ is real. King and Ruskai [13] showed that any such map can be reduced, via changes of basis in $\mathbf{C^2}$, to the form

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ t_1 & \lambda_1 & 0 & 0 \\ t_2 & 0 & \lambda_2 & 0 \\ t_3 & 0 & 0 & \lambda_3 \end{pmatrix}. \tag{3}$$

Because a unitary change of basis $\rho \to V\rho V^\dagger$ on $\mathbf{C^2}$ is equivalent to a 3-dimensional rotation $\mathbf{w} \to R\mathbf{w}$ acting on the Pauli matrices, this is equivalent to

$$\Phi(\rho) = U\Big[\Phi_{\mathbf{t},\mathbf{\Lambda}}(V\rho V^\dagger)\Big]U^\dagger, \tag{4}$$

where $U, V \in U(2)$ and $\Phi_{\mathbf{t},\mathbf{\Lambda}}$ denotes the map corresponding to (3). The reduction to (3) is obtained by applying a modification of the singular value decomposition to the 3×3 real matrix T corresponding to the restriction of Φ to the subspace of matrices with zero trace. However, the constraint that the diagonalization is achieved using rotations rather than arbitrary real orthogonal matrices forces us to relax the usual requirement that the λ_k be positive (so that one can only say that $|\lambda_k|$ are singular values of T). See Appendix A of [13] for details and Appendix C for some examples and discussion of subtle issues regarding the signs of λ_k. (This decomposition is not unique; the parameters can be permuted and the signs of any *two* of the λ_k can be changed by conjugating with a Pauli matrix. Only the sign of the product $\lambda_1\lambda_2\lambda_3$ is fixed. The canonical example of a positivity preserving map which is not completely positive, the transpose, corresponds to λ_k taking the values $+1, -1, +1$. Hence it may be surprising that the product $\lambda_1\lambda_2\lambda_3$ can be negative, as in Example 2(b) of Section 3.2, and the sign of this product has an impact on the allowed values of the translation parameters t_k.)

We call a stochastic map Φ *unitary* if $\Phi(\rho) = U\rho U^\dagger$ and sometimes write $\Gamma_U(\rho)$. It is easy to check that the Kraus representation of a unitary stochastic map is (essentially) unique, and (4) can be rewritten as $\Phi = \Gamma_U \circ \Phi_{\mathbf{t},\mathbf{\Lambda}} \circ \Gamma_V$.

We are interested here in the (convex) set $\mathcal{S}$ of stochastic maps, i.e., those Φ that satisfy the stronger condition of complete positivity. The crucial point about the reduction (4) is that Φ is completely positive if and only if $\Phi_{\mathbf{t},\mathbf{\Lambda}}$ is. Thus, the question of characterizing stochastic maps reduces to studying matrices of the form (3) under the assumption that $|\lambda_k| \leq 1$ (which is necessary for Φ to be positivity preserving). Of course, this reduction is not necessarily unique when the λ_k's are not distinct; this will be discussed further in section 3.1.

The image of the Bloch sphere of pure states under a map of the form (3) is the ellipsoid

$$\left(\frac{x_1 - t_1}{\lambda_1}\right)^2 + \left(\frac{x_2 - t_2}{\lambda_2}\right)^2 + \left(\frac{x_3 - t_3}{\lambda_3}\right)^2 = 1 \tag{5}$$

so that eigenvalues λ_k define the length of the axes of the ellipsoid and the vector $\mathbf{t}$ its center. The Bloch sphere picture of images as ellipsoids is useful because it allows one to determine geometrically the states that emerge with minimal entropy and, roughly, the states associated with maximal capacity. Note that a trace preserving map is positivity-preserving if and only if it maps the Bloch sphere into the "Bloch ball", defined as the Bloch sphere plus its interior. However, not all ellipsoids contained in the "Bloch ball" correspond to images of a stochastic map Φ. It was shown in [1, 13] that the λ_k's are limited by the inequalities $(\lambda_1 \pm \lambda_2)^2 \leq (1 \pm \lambda_3)^2$. However, even for most allowable choices of λ_k, complete positivity restricts (often rather severely) the extent to which translation of the ellipsoid is possible. Moreover, characterizing the allowable ellipsoids is not equivalent to characterizing all stochastic maps because (5) depends only on $|\lambda_k|$ while the actual conditions restrict the choice of signs of λ_k as well. It is worth emphasizing that complete positivity is an extremely strong condition. In fact whether the map is stochastic or not depends on the position and orientation of that ellipsoid inside the Bloch sphere and there are many ellipsoids within the Bloch sphere that do *not* correspond to a completely positive map.

For bistochastic maps, the extreme points are known [1, 13] to consist of the maps that conjugate by a unitary matrix and, in the λ_k representation, correspond to four corners of a tetrahedron. The maps on the edges of the tetrahedron correspond to ellipsoids that have exactly two points in common with the boundary of the Bloch sphere. These maps play a special role and it is useful to consider them as *quasi-extreme points*. We sometimes call the closure of the set of extreme points the set of the *generalized extreme points* and we then refer to those points that are generalized, but not true, extreme points as quasi-extreme points. We will see that for non-unital maps the extreme points correspond to those maps for which the translation allows the ellipsoid to touch the boundary of the sphere at two points (provided one interprets a single point as a pair of degenerate ones in certain special cases.) This is discussed in more detail in section 3.2.

1.3 Summary of Results

We now summarize our results for maps of the form (3). For such maps, it is easy to verify that a necessary condition for Φ to be positivity-preserving is that $|t_k| + |\lambda_k| \leq 1$ for $k = 1, 2, 3$.

Theorem 1 *A map Φ given by* (3) *for which $|t_3| + |\lambda_3| \leq 1$ is completely positive if and only if the equation*

$$\begin{pmatrix} t_1 + it_2 & \lambda_1 + \lambda_2 \\ \lambda_1 - \lambda_2 & t_1 + it_2 \end{pmatrix} \tag{6}$$
$$= \begin{pmatrix} (1 + t_3 + \lambda_3)^{\frac{1}{2}} & 0 \\ 0 & (1 + t_3 - \lambda_3)^{\frac{1}{2}} \end{pmatrix} R_\Phi \begin{pmatrix} (1 - t_3 - \lambda_3)^{\frac{1}{2}} & 0 \\ 0 & (1 - t_3 + \lambda_3)^{\frac{1}{2}} \end{pmatrix}$$

has a solution R_Φ that is a contraction.

Remark: When $|t_k| + |\lambda_k| < 1$ (6) has the unique solution

$$R_\Phi = \begin{pmatrix} \dfrac{t_1 + it_2}{(1+t_3+\lambda_3)^{1/2}(1-t_3-\lambda_3)^{1/2}} & \dfrac{\lambda_1+\lambda_2}{(1+t_3+\lambda_3)^{1/2}(1-t_3+\lambda_3)^{1/2}} \\ \dfrac{\lambda_1-\lambda_2}{(1+t_3-\lambda_3)^{1/2}(1-t_3-\lambda_3)^{1/2}} & \dfrac{t_1+it_2}{(1+t_3-\lambda_3)^{1/2}(1-t_3+\lambda_3)^{1/2}} \end{pmatrix}. \tag{7}$$

When $|t_3| + |\lambda_3| = 1$ no solution to (6) exists unless $t_1 = t_2 = 0$ and either $\lambda_1 = \lambda_2$ or $\lambda_1 = -\lambda_2$. In either case, it would suffice to let $R_\Phi = \frac{1}{2\sqrt{|\lambda_3|}} \begin{pmatrix} 0 & \lambda_1+\lambda_2 \\ \lambda_1-\lambda_2 & 0 \end{pmatrix}$. However, this matrix is singular. Since the solution to (6) is not unique in this case, it will be more convenient to define R_Φ by the non-singular matrix

$$R_\Phi = \frac{\lambda_1}{\sqrt{|\lambda_3|}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{8}$$

If $\lambda_3 = 0$ and $|t_3| + |\lambda_3| = 1$, then $|t_3| = 1$ in which case Φ will not be completely positive unless $t_1 = t_2 = \lambda_1 = \lambda_2 = 0$; then it is consistent and sufficient to interpret $\frac{\lambda_1}{\sqrt{|\lambda_3|}} = 1$ so that $R_\Phi = \sigma_x$.

We now return to the case $|t_k| + |\lambda_k| < 1$ and analyze the requirement that R_Φ given by (7) is a contraction. The requirement that the diagonal elements of $R_\Phi^\dagger R_\Phi$ and $R_\Phi R_\Phi^\dagger$ must be ≤ 1 then implies that

$$\begin{aligned} (\lambda_1+\lambda_2)^2 &\leq (1+\lambda_3)^2 - t_3^2 - (t_1^2+t_2^2)\left(\frac{1+\lambda_3 \pm t_3}{1-\lambda_3 \pm t_3}\right) \qquad (9) \\ &\leq (1+\lambda_3)^2 - t_3^2 \\ (\lambda_1-\lambda_2)^2 &\leq (1-\lambda_3)^2 - t_3^2 - (t_1^2+t_2^2)\left(\frac{1-\lambda_3 \pm t_3}{1+\lambda_3 \pm t_3}\right) \qquad (10) \\ &\leq (1-\lambda_3)^2 - t_3^2. \end{aligned}$$

This implies that the Algoet-Fujiwara condition [1]

$$(\lambda_1 \pm \lambda_2)^2 \leq (1 \pm \lambda_3)^2 - t_3^2 \tag{11}$$

always holds. This was originally obtained [1] as a necessary condition for complete positivity under the *assumption* that Φ has the form (3) with the additional constraint $t_1 = t_2 = 0$. Moreover, it follows that a *necessary* condition for complete positivity of a map of the form (3) is

$$(\lambda_1 \pm \lambda_2)^2 \leq (1 \pm \lambda_3)^2. \tag{12}$$

Although not obvious from this analysis, (12) holds if and only if it is valid for any permutation of the parameters λ_k.

In addition to a condition on the diagonal, the requirement $R_\Phi^\dagger R_\Phi \leq I$ also implies that $\det(I - R_\Phi^\dagger R_\Phi) \geq 0$. This leads to the condition

$$\begin{aligned} &\left[1 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2) - (t_1^2 + t_2^2 + t_3^2)\right]^2 \\ &\qquad \geq \quad 4\left[\lambda_1^2(t_1^2 + \lambda_2^2) + \lambda_2^2(t_2^2 + \lambda_3^2) + \lambda_3^2(t_3^2 + \lambda_1^2) - 2\lambda_1\lambda_2\lambda_3\right] \end{aligned} \tag{13}$$

Conditions on the diagonal elements and determinant of $I - R_\Phi^\dagger R_\Phi$ suffice to insure R_Φ is a contraction. Moreover, a direct analysis of the case $|t_k| + |\lambda_k| = 1$ allows us to extend these inequalities to yield the following result.

Corollary 2 *A map Φ given by* (3) *for which $|t_3| + |\lambda_3| \leq 1$ is completely positive if and only if* (9)*,* (10) *and* (13) *hold, where* (9) *and* (10) *are interpreted so that $t_1 = t_2 = 0$ when $|t_3| + |\lambda_3| = 1$.*

We now discuss those maps Φ given by (3) that are extreme points of $\mathcal{S}$.

Theorem 3 *Let Φ be a stochastic map induced by a matrix $\mathbf{T}$ of the form* (3)*. Then Φ belongs to the closure of the set of extreme points of $\mathcal{S}$ if and only if the matrix R_Φ [as given by* (7) *or* (8) *above] is unitary.*

If $\Phi \in \mathcal{S}$, then it is unitarily equivalent, in the sense of (4), to a map of the form (3). Moreover, as will be shown in section 2.3, that "equivalence" preserves the property of being a generalized extreme point of $\mathcal{S}$. Thus, in particular, a necessary condition for Φ to be an extreme point of $\mathcal{S}$ is that it is equivalent [in the sense of (4)] to a map of the form (3) for which the corresponding R_Φ is unitary.

Since a cyclic permutation of indices corresponds to a rotation in $\mathbf{R}^3$ which can be incorporated into the maps U and V in (4), stating conditions (for membership of $\mathcal{S}$, or for any form of extremality) in a form in which, say, λ_3 and t_3 play a special role does not lead to any real loss of generality. (Non-cyclic permutations are more subtle; e.g, a rotation of $\pi/2$ around the y-axis is equivalent to the interchange $t_1 \leftrightarrow -t_3, \lambda_1 \leftrightarrow \lambda_3$. However, the sign change on t_3 does not affect the conditions (9)-(13) above nor (14)-(16) below which involve either t_3^2 or both signs $\pm t_3$.) Thus, Φ is in the closure of the extreme points of $\mathcal{S}$ if and only if, up to a permutation of indices, equality holds in (9), (10) and (13). This can be summarized and reformulated in the following useful form.

Theorem 4 *A map Φ belongs to the closure of the set of extreme points of $\mathcal{S}$ if and only if it can be reduced to the form* (3) *so that at most one t_k is non-zero and (with the convention this is t_3)*

$$(\lambda_1 \pm \lambda_2)^2 = (1 \pm \lambda_3)^2 - t_3^2. \tag{14}$$

Moreover, Φ is extreme unless $t_3 = 0$ and $|\lambda_3| < 1$.

Adding and subtracting the conditions (14) yields the equivalent conditions

$$\lambda_3 = \lambda_1\lambda_2 \tag{15}$$
$$t_3^2 = (1-\lambda_1^2)(1-\lambda_2^2) \tag{16}$$

This leads to a useful trigonometric parameterization of the extreme points and their Kraus operators, as given below.

We conclude this summary by noting that we have obtained three equivalent sets of conditions on the closure of the set of extreme points of $\mathcal{S}$ after appropriate reduction to the "diagonal" form (3).

a) Equality in (9), (10) and (13),

b) $t_1 = t_2 = 0$ and (14) with both signs,

c) $t_1 = t_2 = 0$, (15) and (16),

where it is implicitly understood that the the last two conditions can be generalized to other permutations of the indices. However, it should be noted that the requirement $t_1 = t_2 = 0$ is equivalent to requiring that R_Φ is skew diagonal. This is a strong constraint that necessarily excludes many unitary 2×2 matrices. That such a constrained subclass of unitary matrices can eventually yield all extreme points is a result of the fact that we have already made a reduction to the special form (3).

It may be worth noting that when one of the $\lambda_k = 0$, then (15) implies that at least two λ_k are zero. The resulting degeneracy allows extreme points that do not necessarily have the form described in Theorem 4. Although the degeneracy in λ_k permits two t_k to be non-zero, it also permits a reduction to the form of the theorem. This is discussed in Section 3.1 as Case IC.

In section 3.1 we consider the implications of conditions (15) and (16) in detail. For now, we emphasize that the interesting new class of extreme points are those that also satisfy $1 > |\lambda_1| > |\lambda_2| > |\lambda_3| > 0$.

1.4 Trigonometric parameterization

The reformulation of the conditions of Theorem 4 as (15) and (16) implies that when Φ is in the closure of the extreme points, the matrix in (3) can be written in the useful form

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos u & 0 & 0 \\ 0 & 0 & \cos v & 0 \\ \sin u \sin v & 0 & 0 & \cos u \cos v \end{pmatrix} \tag{17}$$

with $u \in [0, 2\pi), v \in [0, \pi)$. Extending the range of u from $[0, \pi)$ to $[0, 2\pi)$ insures that the above parameterization yields both positive and negative values of t_3 as allowed

by (16). It is straightforward to verify that in this parameterization Φ can be realized with the Kraus operators

$$\begin{aligned} A_+ &= [\cos\tfrac{1}{2}v\,\cos\tfrac{1}{2}u]\,I + [\sin\tfrac{1}{2}v\,\sin\tfrac{1}{2}u]\,\sigma_z \\ A_- &= [\sin\tfrac{1}{2}v\,\cos\tfrac{1}{2}u]\,\sigma_x - i\,[\cos\tfrac{1}{2}v\,\sin\tfrac{1}{2}u]\,\sigma_y\,. \end{aligned} \tag{18}$$

The rationale behind the subscripts $\pm$ should be clear when we compute the products $A_\pm A_\pm^\dagger$ in Section 2.

A similar parameterization was obtained by Niu and Griffiths [22] in their work on two-bit copying devices. One generally regards noise as the result of failure to adequately control interactions with the environment. However, even classically, noise can also arise as the result of deliberate "jamming" or as the inadvertent result of eavesdropping as in, e.g., wire-tapping. One advantage to quantum communication is that protocols involving non-orthogonal signals provide protection against undetected eavesdropping. Any attempt to intercept and duplicate the signal results leads to errors that may then appear as noise to the receiver. The work of Niu and Griffiths [22] suggests that the extreme points of the form (17) can be regarded as arising from an eavesdropper trying to simultaneously optimize information intercepted and minimize detectable effects in an otherwise noiseless channel.

This parameterization was also obtained independently by Rieffel and Zalka [25] who, like Niu and Griffiths, considered maps that arise via interactions between a pair of qubits, i.e., maps defined by linear extension of

$$\Phi(E_1) = T_2 U(E_1 \otimes E_2)U^\dagger \tag{19}$$

where T_2 denotes the partial trace, U is a unitary matrix in $\mathcal{M}_4$ and E_j denotes a projection onto a pure state in $\mathbf{C}_2$. Moreover, Niu and Griffiths's construction showed that taking the partial trace T_1 rather than T_2 is equivalent to switching sin and cos in (17).

Since one can obtain all extreme points of $\mathcal{S}$ by a reduction via partial trace on $\mathbf{C}^2 \otimes \mathbf{C}^2$ one might conjecture, as Lloyd [20] did, that any map in $\mathcal{S}$ can be represented in the form (19) if E_2 is replaced with a mixed state ρ_2. However, Terhal, et al [26] showed that this is false and another counter-example was obtained in [25]. This does not contradict Lindblad's representation in [19] because his construction requires that the second Hilbert space have dimension equal to the number of Kraus operators. Thus, a map that requires four Kraus operators is only guaranteed to have a representation of the form (19) on $\mathbf{C}^2 \otimes \mathbf{C}^4$.

2 Proofs

2.1 Choi's results and related work

Our results will be based on the following fundamental result of Choi [4, 23].

Theorem 5 *A linear map $\Omega : \mathcal{M}_n \to \mathcal{M}_n$ is completely positive if and only if $\Omega \otimes I_{\mathcal{M}_n}$ is positivity-preserving on $\mathcal{M}_n \otimes \mathcal{M}_n$ or, equivalently, if and only if the matrix*

$$\begin{pmatrix} \Omega(E_{11}) & \dots & \Omega(E_{1n}) \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \Omega(E_{n1}) & \dots & \Omega(E_{nn}) \end{pmatrix} \tag{20}$$

is positive semi-definite where $(E_{j,k})_{j,k=1}^n$ are the standard matrix units for $\mathcal{M}_n$ so that (20) is in $\mathcal{M}_n(\mathcal{M}_n) = \mathcal{M}_{n^2}$

We are interested in $n = 2$, in which case this condition is equivalent [12] to

$$(\Omega \otimes I)(B_0) \geq 0 \tag{21}$$

where B_0 is the pure state density matrix that projects onto the maximally entangled Bell state ψ_0 which physicists usually write as $\psi_0 = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle)$, i.e., $B_0 = |\psi_0\rangle\langle\psi_0|$. It will be useful to define $\beta(\Omega) \equiv 2(\Omega \otimes I)(B_0)$ to be the matrix in (20) and write

$$\beta(\Omega) = \begin{pmatrix} A & C \\ C^\dagger & B \end{pmatrix}, \tag{22}$$

so that $A = \Omega(E_{11})$, $B = \Omega(E_{22})$ and $C = \Omega(E_{12})$. We now assume that Φ is trace-preserving, so that its adjoint $\widehat{\Phi}$ is unital, $\widehat{\Phi}(I) = I$, and write

$$\beta(\widehat{\Phi}) = \begin{pmatrix} \widehat{\Phi}(E_{11}) & \widehat{\Phi}(E_{12}) \\ \widehat{\Phi}(E_{21}) & \widehat{\Phi}(E_{22}) \end{pmatrix} = \begin{pmatrix} A & C \\ C^\dagger & I - A \end{pmatrix}, \tag{23}$$

where we have exploited the fact that $E_{11} + E_{22} = I$ so that $B = \widehat{\Phi}(E_{22}) = \widehat{\Phi}(I - E_{11}) = I - A$ when $\Omega = \widehat{\Phi}$. Thus, $0 \leq A \leq I$.

Note that $\beta(\Phi) = 0 \Leftrightarrow \Phi(E_{jk}) = 0 \ \forall \ j, k$. Hence β defines an affine isomorphism between $\mathcal{S}$ and the image $\beta(\mathcal{S}) \subset \mathcal{M}_4$. In particular, there is a one-to-one correspondence between extreme points of $\mathcal{S}$ and those of the image $\beta(\mathcal{S})$ or $\beta(\widehat{\mathcal{S}})$ where $\widehat{\mathcal{S}} = \{\widehat{\Phi} : \Phi \in \mathcal{S}\}$ denotes the set of completely positive maps that are unital.

It is also worth noting that the matrix

$$\beta(\widehat{\Phi}) = 2(\widehat{\Phi} \otimes I)(B_0) = \begin{pmatrix} \widehat{\Phi}(E_{11}) & \widehat{\Phi}(E_{12}) \\ \widehat{\Phi}(E_{21}) & \widehat{\Phi}(E_{22}) \end{pmatrix} \tag{24}$$

is obtained from $\beta(\Phi)$ by conjugating with a permutation matrix that exchanges the second and the third coordinate in $\mathbf{C}^4$ and then taking the complex conjugate, i.e.,

$$\beta(\widehat{\Phi}) = \overline{U_{23}^\dagger \, \beta(\Phi) \, U_{23}} \tag{25}$$

where

$$U_{23} = U_{23}^{\dagger} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In particular, (25) shows that $\beta(\widehat{\Phi})$ is positive semi-definite if and only if $\beta(\Phi)$ is.

For maps of the form (3), one easily finds

$$\beta(\Phi) = \frac{1}{2} \begin{pmatrix} 1+t_3+\lambda_3 & t_1-it_2 & 0 & \lambda_1+\lambda_2 \\ t_1+it_2 & 1-t_3-\lambda_3 & \lambda_1-\lambda_2 & 0 \\ 0 & \lambda_1-\lambda_2 & 1+t_3-\lambda_3 & t_1-it_2 \\ \lambda_1+\lambda_2 & 0 & t_1+it_2 & 1-t_3+\lambda_3 \end{pmatrix} \tag{26}$$

or, equivalently,

$$\beta(\widehat{\Phi}) = \frac{1}{2} \begin{pmatrix} 1+t_3+\lambda_3 & 0 & t_1+it_2 & \lambda_1+\lambda_2 \\ 0 & 1+t_3-\lambda_3 & \lambda_1-\lambda_2 & t_1+it_2 \\ t_1-it_2 & \lambda_1-\lambda_2 & 1-t_3-\lambda_3 & 0 \\ \lambda_1+\lambda_2 & t_1-it_2 & 0 & 1-t_3+\lambda_3 \end{pmatrix}. \tag{27}$$

The reason why we concentrate on matrices $\beta(\widehat{\Phi})$ (rather than $\beta(\Phi)$) is that for maps of the form (3) the matrices A and $B = I - A$ are diagonal, which greatly simplifies the analysis.

Our next goal is to characterize elements of $\beta(\widehat{\mathcal{S}})$ and the set of extreme points of $\beta(\widehat{\mathcal{S}})$. We will do this by applying Choi's condition to (27) for which we will need the following result.

Lemma 6 *A matrix* $M = \begin{pmatrix} A & C \\ C^{\dagger} & B \end{pmatrix}$ *is positive semi-definite if and only if* $A \geq 0, B \geq 0$ *and* $C = A^{\frac{1}{2}} R B^{\frac{1}{2}}$ *for some contraction* R*. Moreover, the set of positive semi-definite matrices with fixed* A *and* B *is a convex set whose extreme points satisfy* $C = A^{\frac{1}{2}} U B^{\frac{1}{2}}$*, where* U *is unitary.*

The first part of the lemma is well-known, see, e.g, [11], Lemma 3.5.12. The second part follows easily by using well-known facts that the extreme points of the set of contractions in $\mathcal{M}_n$ are unitary (see, e.g. [24] Lemma 1.4.7, [11] Section 3.1, problem 27 or Section 3.2, problem 4) and that the extreme points of the image of this (compact convex) set under the affine map $R \to A^{\frac{1}{2}} R B^{\frac{1}{2}}$ are images of extreme points.

When the matrix $M = \beta(\widehat{\Phi})$ has the form (27), then $B = I - A$ and the map R coincides with R_Φ defined by (6) (or by (7) if $0 < A < I$). Accordingly, Lemma 6 will imply that a necessary condition that Φ be extreme is that R_Φ is unitary. In fact, we have the following equivalence from which Theorem 3 follows.

Theorem 7 *A map Φ is a generalized extreme point of $\mathcal{S}$ if and only if the corresponding matrix $M = \beta(\widehat{\Phi})$ defined via* (23) *is of the form*

$$M = \begin{pmatrix} A & \sqrt{A}\,U\sqrt{I-A} \\ \sqrt{I-A}\,U^\dagger\sqrt{A} & I-A \end{pmatrix} \tag{28}$$

with $0 \leq A \leq I$ and U unitary.

Since the matrices $\{E_{jk}\}$ form a basis for $\mathcal{M}_2$, any positive semi-definite matrix of the form (23) defines a completely positive unital map $\widehat{\Phi}$ and, hence, a stochastic map Φ. Thus, it remains only to verify the "if" part, i.e., that U being unitary implies that Φ is a generalized extreme point.

To that end, we shall need some additional results. First, we observe that Lemma 6 implies that M is positive semi-definite if and only if it can be written in the form (all blocks are 2×2)

$$M = \begin{pmatrix} A & C \\ C^\dagger & B \end{pmatrix} = \begin{pmatrix} \sqrt{A} & 0 \\ 0 & \sqrt{B} \end{pmatrix} \begin{pmatrix} I & R \\ R^\dagger & I \end{pmatrix} \begin{pmatrix} \sqrt{A} & 0 \\ 0 & \sqrt{B} \end{pmatrix} \tag{29}$$

with R a contraction. Furthermore, it is well-known and easily verified that a matrix R in $\mathcal{M}_2$ is unitary if and only if $E_R := \begin{pmatrix} I & R \\ R^\dagger & I \end{pmatrix}$ has rank two. When A and B are both non-singular, it follows immediately that a positive semi-definite matrix M can be written in the form (29) with R unitary if and only if M has rank two. In the singular case some care must be used, but we still have

Lemma 8 *Let $M \in \mathcal{M}_4$ with $M \geq 0$. Then M admits a factorization* (29) *with R unitary if and only if* $\operatorname{rank}(M) \leq 2$.

Proof: The argument is quite standard (e.g., it can be extracted from results in [11], in fact for blocks of any size) but we include it for completeness. We need to show that if $\operatorname{rank}(M) \leq 2$, then the equation

$$C = \sqrt{A}\,R\,\sqrt{B} \tag{30}$$

admits a unitary solution R. This holds, by the comments above, if A and B are nonsingular and, trivially, if $A = 0$ or $B = 0$. To settle the remaining cases we observe that by conjugating M with a matrix of the form $\begin{pmatrix} V & 0 \\ 0 & W \end{pmatrix}$ where V, W are 2×2 unitaries, we may reduce the question to the case when A and B are diagonal. If $\operatorname{rank}(A) = \operatorname{rank}(B) = 1$, the equation (30) imposes restriction on only one of the entries of R. By the first part of Lemma 6 that entry must be ≤ 1 in absolute value for M to be positive-definite, and the remaining entries can be chosen so that R is unitary. If, say, $\operatorname{rank}(A) = 2$ and $\operatorname{rank}(B) = 1$ (say, only the first diagonal entry of B is $\neq 0$), then ignoring the last row and last column of M (which must be 0) we can think of

equation (30) as involving nonsingular matrices $A \in \mathcal{M}_2$ and $B \in \mathcal{M}_1$ and a 2×1 matrix R. The condition that $\text{rank}(E_R) \leq 2$ then implies that R is a norm one column vector (if R was an $n \times m$ matrix with $m \leq n$, the condition for $\text{rank}(E_R) \leq n$ would be that R is an isometry). Returning to 2×2 blocks we notice that in the present case the equation (30) does not impose any restrictions on the *second* column of R and so that column can be chosen so that R is unitary. (See, e.g., the discussion of unitary dilations on pp. 57-58 of [11].) **QED**

We will also need more results from Choi [4]. First, there is a special case of Theorem 5 in [4] which we state in a form adapted to our situation and notation.

Lemma 9 *A stochastic map is an extreme point of $\mathcal{S}$ if and only if it can be written in the form* (1) *(necessarily with $\sum_k A_k A_k^\dagger = I$), so that the set of matrices $\{A_j A_k^\dagger\}$ is linearly independent.*

In addition to being an important ingredient in the proof of Theorem 7, this result will allow us to distinguish between true extreme points and those in the closure.

The next result is "essentially" implicit in [4].

Lemma 10 *The minimal number of Kraus operators A_k needed to represent a completely positive map in the form* (1) *is* $\text{rank}[\beta(\Phi)]$.

Indeed, Choi showed that one can obtain a set of Kraus operators for any completely positive map Φ from an orthogonal set of eigenvectors of $\beta(\Phi)$ corresponding to non-zero eigenvalues. Hence one can always write Φ in the form (1) using only $\text{rank}[\beta(\Phi)]$ Kraus operators. The other direction is even simpler. It is readily checked (and it also follows from the proof of Choi's Theorem 5) that if $\Phi(\rho) := V^\dagger \rho V$ with $V \neq 0$, then $\text{rank}\, \beta(\Phi) = 1$ and so, in general, $\text{rank}\, \beta(\Phi)$ does not exceed the number of Kraus operators.

The above results together with yet unproved Theorems 7 and 4 allow us to compile a list of conditions characterizing the closure of extreme points of the set $\mathcal{S}$ of stochastic maps on $\mathcal{M}_2$ (the "generalized extreme points"). These conditions are given in Theorem 12 of section 2.3.

2.2 Proof of Results in Section 1.3

Proof of Theorem 1 and Remark: By Theorem 5, it suffices to show that (27) is positive semi-definite. To do this we apply Lemma 6 to (27). Since the corresponding A and B are then diagonal, they are positive semi-definite if and only if each term on the diagonal is non-negative which is equivalent to $|t_3| + |\lambda_3| \leq 1$. If this inequality is strict, then both A and B are invertible and it follows easily that $R = A^{-1/2} C B^{-1/2}$ has the form given by (7).

If either or both of A, B is singular, we use the fact that when a diagonal element of an $n \times n$ matrix is zero, the matrix can be positive semi-definite only if the corresponding row and column are identically zero. Thus, if any of the diagonal elements of (27) is zero, then $t_1 = t_2 = 0$ and one of $\lambda_1 \pm \lambda_2 = 0$. It is then straightforward to check that R_Φ can be chosen to have the form in (8). The cases $\lambda_1 = 0$ and $\lambda_3 = 0$ can also be easily checked explicitly, which suffices to verify the remark following Theorem 1. **QED**

The next proof will use Lemma 9 which requires the following matrix products, which are easily computed from (18).

$$\begin{aligned} 2A_+A_+^\dagger &= [1 + \cos u \cos v]\, I + [\sin u \sin v]\, \sigma_z \\ 2A_-A_-^\dagger &= [1 - \cos u \cos v]\, I - [\sin u \sin v]\, \sigma_z \\ 2A_\pm A_\mp^\dagger &= [\sin v]\, \sigma_x \pm i\, [\sin u]\, \sigma_y. \end{aligned} \tag{31}$$

Note that it follows immediately that $A_+A_+^\dagger + A_-A_-^\dagger = I$ as required for a trace-preserving map Φ.

We also point out that when a completely positive map Φ can be realized with one Kraus operator A, the condition $\sum_k A_k A_k^\dagger = I$ reduces to $AA^\dagger = I$, from which it is elementary that $A^\dagger A = I$ as well. Thus, when a map Φ can be realized with a single Kraus operator A, then Φ is unital $\Leftrightarrow$ Φ is trace-preserving $\Leftrightarrow$ A is unitary.

Proof of Theorems 3 and 4: We argue as above and observe that fixing A, B in Lemma 6 when applied to $M = \beta(\widehat{\Phi})$ is equivalent to fixing $\widehat{\Phi}(\sigma_z)$ or the last row of $\mathbf{T}$ (note also that A and B are diagonal if and only if the last row of $\mathbf{T}$ is as in (3), i.e., if its two middle entries are 0). It then follows immediately from the second part of Lemma 6 that when Φ is an extreme point R_Φ must be unitary. By continuity and compactness, this holds also for Φ's in the closure of the extreme points.

We now show that this unitarity implies condition (14) in Theorem 4. First, we observe that when R_Φ is unitary, equality holds in both (9) and (10). When $t_3 \neq 0$ and $\lambda_3 \neq 0$, this is possible only if $t_1 = t_2 = 0$, which yields (14).

When $|t_3| + |\lambda_3| = 1$ (including the possibilities $t_3 = 0, |\lambda_3| = 1$ and $t_3 = 1, |\lambda_3| = 0$) the necessity of the conditions in Theorem 4 can be checked explicitly using the observations in the previous proof.

When $t_3 = 0$ and $0 < |\lambda_3| < 1$, the necessity of the conditions in Theorem 4 requires more work. The condition that the columns of R_Φ are orthogonal becomes

$$(t_1 + it_2)\frac{\lambda_1 - \lambda_2}{1 - \lambda_3} + (t_1 - it_2)\frac{\lambda_1 + \lambda_2}{1 + \lambda_3} = 0$$

which implies

$$\begin{aligned} t_1 \left[\frac{\lambda_1 - \lambda_2}{1 - \lambda_3} + \frac{\lambda_1 + \lambda_2}{1 + \lambda_3}\right] &= 0 \\ t_2 \left[\frac{\lambda_1 - \lambda_2}{1 - \lambda_3} - \frac{\lambda_1 + \lambda_2}{1 + \lambda_3}\right] &= 0. \end{aligned}$$

Both quantities in square brackets can not simultaneously be zero unless $\lambda_1 = \lambda_2 = 0$. Although we can then have $t_1, t_2 \neq 0$ this is, except for permutation of indices an allowed special case of (14), as discussed in Section 3.1 as case (IC). If only one of the quantities in square brackets is zero, then we again have only one t_k non-zero, and (14) holds after a suitable permutation of indices. Suppose, for example, that $t_2 \neq 0$. Then

$$\frac{\lambda_1 - \lambda_2}{1 - \lambda_3} = \frac{\lambda_1 + \lambda_2}{1 + \lambda_3} \;\Rightarrow\; \lambda_2 = \lambda_1 \lambda_3 \tag{32}$$

and

$$\begin{aligned} 1 = \frac{t_2^2}{1-\lambda_3^2} + \frac{(\lambda_1+\lambda_2)^2}{(1+\lambda_3)^2} &= \frac{t_2^2}{1-\lambda_3^2} + \frac{\lambda_1+\lambda_2}{1+\lambda_3}\frac{\lambda_1-\lambda_2}{1-\lambda_3} \\ &\Rightarrow \quad 1 - \lambda_3^2 = t_2^2 + \lambda_1^2 - \lambda_2^2 = t_2^2 + \lambda_1^2 - \lambda_1^2\lambda_3^2 \\ &\Rightarrow \quad (1-\lambda_3^2)(1-\lambda_1^2) = t_2^2. \end{aligned} \tag{33}$$

Except for interchange of t_2, λ_2 with t_3, λ_3, equations (32) and (33) are equivalent to (15) and (16).

The case $\lambda_3 = 0$ can be treated using an argument similar to that above for $t_3 = 0$. Thus, we have verified that unitarity of R_Φ implies (14) in all cases.

Conversely, one can easily check that under the hypotheses $t_1 = t_2 = 0$ and (14), R_Φ as given by (7) or (8) is always unitary. The remaining question is whether or not every unitary matrix R_Φ gives rise to a generalized extreme point of $\mathcal{S}$. To answer this question we use the fact that the corresponding maps can be written in the form (17) with Kraus operators given by (18). We then make the following series of observations.

I) Because the set $\{I, \sigma_x, \sigma_y, \sigma_z\}$ forms an orthonormal basis (with respect to the Hilbert-Schmidt inner product) for $\mathcal{M}_2$, the operators in (31) are linearly independent if and only if $\sin v \sin u \neq 0$. Thus, by Lemma 9, the corresponding stochastic map Φ is extreme if *both* of $\{\sin v, \sin u\}$ are non-zero or, equivalently, $t_3 \neq 0$ and all $|\lambda_k| < 1$.

II) If *both* of $\{\sin v, \sin u\}$ are zero, then u, v are both of the form $n\pi$ for some integer n. In this case one and only one of the two Kraus operators given by (18) is non-zero and, hence, unitary. In this case, Φ is unitary and obviously extreme, and $\lambda_k = \pm 1$ for $k = 1, 2, 3$. In fact, the non-zero Kraus operator is simply one of $\{I, \sigma_x, \sigma_y, \sigma_z\}$ and, accordingly, either none or exactly two of the λ_k can be negative.

III) If exactly one of $\{\sin v, \sin u\}$ is zero, then it follows from (17) that $t_3 = \sin v \sin u = 0$, and Φ is unital. The extreme points of the unital maps are known [1, 14], and given by (II) above. Hence maps for which exactly one of $\{\sin v, \sin u\}$ is zero can not be extreme. However, it is readily verified that such a Φ is a convex combination of two extreme maps of type (II).

Since these three situations exhaust the possible situations in (17), they also cover all possible unitary maps which arise from an R_Φ of the form under consideration. Moreover, maps of type (III) form a 1-dimensional subset of the 2-dimensional torus given by the parameterization (17) and so they must be in the closure of the set of extreme points. Maps of type (I) and (II) correspond to the non-unital and unital extreme points of $\mathcal{S}$ respectively. **QED**

The argument presented thus far yields the extreme points of $\mathcal{S}$ that are represented by a matrix $\mathbf{T}$ of the form (3). Since every $\Phi \in \mathcal{S}$ is equivalent in the sense of (4) (i.e., after changes of bases in the 2-dimensional Hilbert spaces corresponding to the domain and the range of Φ) to a map of the form (3), this describes *in principle* all extreme points of $\mathcal{S}$ and all points in the closure of the set of extreme points.

Although it is possible to characterize the extreme points of $\mathcal{S}$ without appealing to the special form (3), we chose the present approach for two reasons. One is that the use of the diagonal form considerably simplifies the computations. The other is that it leads to a parameterization that is useful in applications. Nevertheless, this basis dependent approach may obscure some subtle issues that we now point out.

If R_Φ is unitary, but does not have the special form (7), it will still define a stochastic map, albeit *not* one of the "diagonal" form (3). However, even though Theorem 7 implies that the corresponding stochastic map Φ is a generalized extreme point of $\mathcal{S}$, we are not able to deduce that immediately. The difficulty is that we do not know that the property of the matrix R [implicitly defined by (29)] of being unitary is invariant under changes of bases, and in particular whether it yields an unitary R_Φ after reduction to the form (7). In the next subsection, we clarify that issue and complete the proofs of Theorem 7 and Theorem 12 that follows.

2.3 Invariance of conditions under change of basis

We begin by carefully describing the result of a change of basis on $\beta(\Phi)$ or $\beta(\widehat{\Phi})$. It will be convenient to let Γ_A denote conjugation with the matrix A, i.e., $\Gamma_A(\rho) = A\rho A^\dagger$. Then for any pair of unitary matrices U, V, $\Phi(\rho) = U[\Phi_D(V\rho V^\dagger)]U^\dagger$ is equivalent to $\Phi = \Gamma_U \circ \Phi_D \circ \Gamma_V$ or $\Phi_D = \Gamma_{U^\dagger} \circ \Phi \circ \Gamma_{V^\dagger}$. Notice that the map $\Phi = \Gamma_U \circ \Phi_D \circ \Gamma_V$ is an affine isomorphism of $\mathcal{S}$ and in particular it preserves the sets of extreme and quasi-extreme points. We will be primarily interested in the case when this reduces a map Φ to diagonal form Φ_D as in (4). However, our results apply to any pair of positive maps related by such a change of basis.

Lemma 11 *Let Φ be a positive map on $\mathcal{M}_2$ and U, V unitary. Then $\Phi = \Gamma_U \circ \Phi_D \circ \Gamma_V$ if and only if*

$$(\Phi \otimes I)(B_0) = [U \otimes V^T]\,[(\Phi_D \otimes I)(B_0)]\,[U^\dagger \otimes \overline{V}] \tag{34}$$

where where V^T denotes the transpose and [as in (21)] $2B_0$ is the matrix with blocks E_{jk}.

Proof: First observe that, for any set of matrices $\{G_{jk}\}$ in $\mathcal{M}_2$,

$$[(\Gamma_U \circ \Phi_D) \otimes I] \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} = \begin{pmatrix} U & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} \Phi_D(G_{11}) & \Phi_D(G_{12}) \\ \Phi_D(G_{21}) & \Phi_D(G_{22}) \end{pmatrix} \begin{pmatrix} U^\dagger & 0 \\ 0 & U^\dagger \end{pmatrix}.$$

The result then follows from the fact that

$$(\Gamma_V \otimes I)(B_0) = (I \otimes \Gamma_{V^T})(B_0) \tag{35}$$

which can be verified directly in a number of ways. **QED**

It may be worth remarking that (35) extends to the general case in which B_0 is replaced by the $n^2 \times n^2$ matrix with blocks E_{jk} and may characterize maximally entangled states.

We now present the promised list of conditions characterizing the closure of extreme points of the set $\mathcal{S}$ of stochastic maps on $\mathcal{M}_2$.

Theorem 12 *For a map $\Phi \in \mathcal{S}$ the following conditions are equivalent:*

(i) *Φ belongs to the closure of the set of extreme points of $\mathcal{S}$.*

(ii) *$M = \beta(\widehat{\Phi})$ can be written in the form* (28) [or, equivalently, in the form (29)] *with R unitary and $B = I - A$.*

(iii) *Φ can be reduced via changes of bases* (4) *to a map of the form* (17).

(iv) *Φ can be represented in the form* (1) *using not more than two Kraus operators.*

(v) $\operatorname{rank} \beta(\widehat{\Phi}) = \operatorname{rank} \beta(\Phi) \leq 2$.

Proof. We start by carefully reviewing the arguments presented thus far. In the preceding section we showed that, in the special case when Φ is represented by a matrix $\mathbf{T}$ of the form (3), the first three conditions are equivalent. Moreover, the implication (i) $\Rightarrow$ (ii) holds in full generality. Concerning the other conditions *in the general case*, Lemma 8 says that (ii) $\Leftrightarrow \operatorname{rank} \beta(\widehat{\Phi}) \leq 2$, while it follows from (25) that $\operatorname{rank} \beta(\widehat{\Phi}) = \operatorname{rank} \beta(\Phi)$; thus (ii) $\Leftrightarrow$ (v). Similarly, Lemma 10 shows that (iv) $\Leftrightarrow$ (v). In addition, Lemma 11 above implies that (v) is invariant under a change of basis. It thus remains to prove that (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i) in full generality rather than just for maps of the form (3). But this follows from the facts that every stochastic map can be reduced to the form (3) via changes of bases (4) and that all these conditions are invariant under such changes of bases. The latter was already noticed for (i) and is trivial for (iii). To show the (not obvious at the first sight) invariance for (ii) we notice that we have already proved that (ii) $\Leftrightarrow$ (iv) and that (iv) is clearly invariant under changes of bases: if $\Phi_1(\rho) = \sum_k A_k^\dagger \rho A_k$ and $\Phi(\rho) = U\Big[\Phi_1(V\rho V^\dagger)\Big]U^\dagger$, then $\Phi(\rho) = \sum_k (V^\dagger A_k U^\dagger)^\dagger \rho (V^\dagger A_k U^\dagger)$. **QED**

Remark. It is possible to directly characterize the extreme points (rather than generalized extreme points) if the map is not of the form (3). However, the conditions thus obtained are not very transparent.

Since Theorem 7 is "essentially" (i) $\Leftrightarrow$ (ii) above, we have also proved that result. The following result gives another approach to proving Theorem 7 since it easily implies (i) $\Leftrightarrow$ (iv). Although redundant, we include it because its proof (which should be compared to the argument in the preceding section) is of independent interest.

Theorem 13 *A stochastic map Φ that, when written in the form (1), requires two Kraus operators A_k but not more, is either an extreme point of $\mathcal{S}$ or bistochastic.*

Proof: Since Φ trace-preserving implies $\sum_k A_k A_k^\dagger = I$, we can assume without loss of generality that $A_1 A_1^\dagger = D$ and $A_2 A_2^\dagger = I - D$ where D is diagonal and $0 < D < I$. Then we can write $A_1 = \sqrt{D}\, V_1,\ A_2 = \sqrt{I-D}\, V_2$ with V_1, V_2 unitary. By Lemma 9, Φ is extreme if and only if the set $\{A_1 A_1^\dagger,\ A_2 A_2^\dagger,\ A_1 A_2^\dagger,\ A_2 A_1^\dagger\}$ is linearly independent. This is equivalent to linear independence of the set

$$\Big\{\, D,\ I - D,\ \sqrt{D}\, W \sqrt{I-D},\ \sqrt{I-D}\, W^\dagger \sqrt{D}\,\Big\}.$$

where $W = V_1 V_2^\dagger$ One can readily verify that this set is linearly independent unless D is a multiple of the identity or W is diagonal. If $D = \mu I$, then both $\mu^{-1/2} A_1$ and $(1-\mu)^{-1/2} A_2$ are unitary so that Φ is a convex combination of unitary maps. The second exception is more subtle. We first note that the fact that W is diagonal implies $A_1^\dagger A_1 = V_2^\dagger W^\dagger D W V_2 = V_2^\dagger D V_2$. Thus

$$A_1^\dagger A_1 + A_2^\dagger A_2 = V_2^\dagger D V_2 + V_2^\dagger (I - D) V_2 = I$$

so that Φ is a bistochastic map. **QED**

One might think that a map of the form (3) with all three t_k non-zero would require three extreme points. However, this is not the case. In general, two maps in the set of generalized points suffice. If a bistochastic map is written in diagonal form, then it has a unique decomposition as a convex combination of the four unitary maps corresponding to the corners of a tetrahedron; however, it can also be written non-uniquely as a convex combination of two maps on the "edges" of the tetrahedron. For non-unital maps, not only do two extreme points suffice, but they can be chosen so that an arbitrary non-unital map is the "midpoint" of a line connecting two (true) extreme points.

Theorem 14 *Any stochastic map on $\mathcal{M}_2$ can be written as the convex combination of of two maps in the closure of the set of extreme points.*

This is an immediate consequence of Theorem 7 and the following elementary result. Note that in both cases we have proven the somewhat stronger result that the convex combination can be chosen to be a midpoint.

Lemma 15 *Any contraction in $\mathcal{M}_2$ can be written as the convex combination of two unitary matrices.*

Proof: If R is a contraction, its singular value decomposition can be written in the form

$$\begin{aligned} R &= V \begin{pmatrix} \cos\theta_1 & 0 \\ 0 & \cos\theta_2 \end{pmatrix} W^\dagger \\ &= \tfrac{1}{2} V \begin{pmatrix} e^{i\theta_1} & 0 \\ 0 & e^{i\theta_2} \end{pmatrix} W^\dagger + \tfrac{1}{2} V \begin{pmatrix} e^{-i\theta_1} & 0 \\ 0 & e^{-i\theta_2} \end{pmatrix} W^\dagger \end{aligned} \tag{36}$$

where V and W are unitary. **QED**

Note that if $R = VDW^\dagger$ as above, then

$$\begin{pmatrix} A & \sqrt{A}R\sqrt{B} \\ \sqrt{B}R^\dagger\sqrt{A} & B \end{pmatrix}$$
$$= \begin{pmatrix} V & 0 \\ 0 & W \end{pmatrix} \begin{pmatrix} V^\dagger AV & \sqrt{V^\dagger AV}D\sqrt{W^\dagger BW} \\ \sqrt{W^\dagger BW}D^\dagger\sqrt{V^\dagger AV} & W^\dagger BW \end{pmatrix} \begin{pmatrix} V^\dagger & 0 \\ 0 & W^\dagger \end{pmatrix}$$

However, this transformation does *not* correspond to a change of basis of the type considered at the start of this section.

3 Discussion and Examples

3.1 Types of extreme points

We begin our discussion of extreme points by using the trigonometric parameterization (17) to find image points that lie on the Bloch sphere. We consider a pure state of the form $\rho = \frac{1}{2}[I + \mathbf{w}\cdot\sigma]$ with $\mathbf{w} = (\pm\cos\theta, 0, \sin\theta)$ so that $\Phi(\rho) = \frac{1}{2}[I + \mathbf{x}\cdot\sigma]$ with $\mathbf{x} = (\pm\cos\theta\cos u, 0, \sin u\sin v + \sin\theta\cos u\cos v)$. After some straightforward trigonometry, we find

$$\begin{aligned} |\mathbf{x}|^2 &= \cos^2\theta\cos^2 u + \sin^2\theta\cos^2 u\cos^2 v + \sin^2 u\sin^2 v \\ &\qquad + 2\sin\theta\,[\sin u\sin v\cos u\cos v] \\ &= 1 - [\sin u\cos v - \sin\theta\cos u\sin v]^2 \end{aligned} \tag{37}$$

so that $|\mathbf{x}| = 1$ if (and only if) $\sin\theta = \frac{\tan u}{\tan v}$. This will be possible if $|\tan u| \le |\tan v|$ or, equivalently, $|\lambda_1| \ge |\lambda_2|$. In particular, Φ maps the pair of states corresponding to $(\pm\cos\theta, 0, \sin\theta)$ on the Bloch sphere to the pair of states corresponding to $(\pm\cos\omega, 0, \sin\omega)$ when

$$\begin{aligned} \cos\omega &= \cos\theta\,\cos u = \frac{\sqrt{\cos^2 u\sin^2 v - \sin^2 u\cos^2 v}}{\sin v} \\ \sin\omega &= \sin u\sin v + \sin\theta\cos u\cos v = \frac{\sin u}{\sin v}. \end{aligned} \tag{38}$$

We now describe several subclasses in the closure of the extreme points described in Theorem 4, using the categories defined in its proof. We make the additional assumption that $|\lambda_1| \geq |\lambda_2|$ since, in any case, all permutations of indices must be considered to obtain the extreme points of all maps of the form(3). With that understanding, the list below is exhaustive. A general extreme point is then the composition of a map of type (I) or (II) with unitary maps, as in (4).

I) $1 > |\lambda_1| \geq |\lambda_2| > |\lambda_3| > 0$, $t_1 = t_2 = 0$ and $t_3^2 = (1-\lambda_1^2)(1-\lambda_2^2)$. This class includes all non-unital extreme points, and can be subdivided to distinguish some important special cases.

   A) $|\lambda_1| > |\lambda_2| > 0$ can be regarded as the generic situation. The image $\Phi(\mathcal{D})$ is an ellipsoid translated orthogonal to its major axes until exactly two points touch the Bloch sphere, as described above and shown in Figure 1.

   B) $|\lambda_1| = |\lambda_2| > 0$. As $|\lambda_1| \to |\lambda_2|$ the two image points on the Bloch sphere merge so that when $\lambda_1 = \pm\lambda_2$, we recover the amplitude-damping channel with one fixed point at the North or South pole corresponding to $t_3 = \pm(1 - |\lambda_3|)$.

   C) $\lambda_2 = 0 \Rightarrow \lambda_3 = 0$. If $\lambda_1 \neq 0$, the image $\Phi(\mathcal{D})$ is a line segment whose endpoints lie on the Bloch sphere. Moreover, the degeneracy $\lambda_2 = \lambda_3 = 0$ permits a rotation so that $t_2 \neq 0$ provided that $t_2^2 + t_3^2 = 1 - \lambda_1^2$.

   If $\lambda_1 = 0$ as well, we have a completely noisy channel with all $\lambda_k = 0$. The image $\Phi(\mathcal{D})$ consists of a single point on the unit sphere onto which all density matrices are mapped, and the degeneracy permits a translation $\mathbf{t}$ by vector of length $|\mathbf{t}| = 1$ in an arbitrary direction.

II) All $|\lambda_k| = 1$, $\mathbf{t} = 0$. In this case, either zero or two of the $\lambda_k = -1$ and the others $+1$. As discussed in Appendix B of [13], these four possibilities correspond to 4 points of a tetrahedron. Each of these extreme points is a map with exactly one Kraus operator corresponding to the identity or one of the three Pauli matrices. Φ takes $\mathcal{D}$ *onto* itself, i.e., $\Phi(\mathcal{D}) = \mathcal{D}$.

III) $\lambda_1 = \pm 1$, $\lambda_2 = \pm\lambda_3 = \mu$ with $|\mu| < 1$. In this case we must have $\mathbf{t} = 0$, and Φ is not a true extreme point, but a point on an "edge" formed by taking the convex combination of two corners of the tetrahedron (II) above. The image $\Phi(\mathcal{D})$ is an ellipsoid whose major axis has length one. Thus, $\Phi(\mathcal{D})$ has one pair of orthogonal states on the unit sphere, which are also fixed points.

   These maps have two non-zero Kraus operators and, hence, the form $s\Phi_j + (1-s)\Phi_k$ where $j, k \in \{0, 1, 2, 3\}$ and $\Phi_j(\rho) = \sigma_j \rho \sigma_j$ (and $\sigma_0 = I$). For each pair (j, k) we obtain a line between two of the extreme points that form the tetrahedron of bistochastic maps.

Figure 1: Cross section of the Bloch sphere and its image for an extreme point of the form (17). The cross section is shown in the plane for which two points on the ellipsoid touch the sphere; and the axes lengths and shift of the center are indicated under the assumption that $\cos u > \cos v$.

It is worth pointing out that case (IC) is the *only* situation in which one can have extreme points with more than one t_k nonzero. When Φ has the form (17) this can only happen when λ_3 is degenerate [i.e., equal to λ_1 or λ_2.] However, this is precluded by (15) unless $\lambda_3 = 0, \pm 1$. In the latter case we must have a unital channel with all $t_k = 0$. Thus, two nonzero t_k occur only in channels that are so noisy that at least two $\lambda_k = 0$.

3.2 Images of stochastic maps

The discussion at the start of Section 3.1 suggests that, roughly speaking, extreme points correspond to maps for which two (or more) points in the image $\Phi(\mathcal{D})$ lie on the Bloch sphere as shown in Figure 1. However, this statement is correct only if we interpret the single point at a pole as a pair of degenerate images when $|\lambda_1| = |\lambda_2|$.

Now suppose we want to find the map(s) Φ that take the pair of points $(\pm\cos\omega, 0, \sin\omega)$ on the Bloch sphere to the pair $(\pm\cos\theta, 0, \sin\theta)$. Geometric considerations and the arguments leading to (38) imply that this will be possible only if $|\sin\theta| > |\sin\omega|$ and $|\lambda_1| > |\lambda_2|$. For simplicity, we assume $0 < \omega < \theta < \frac{1}{2}\pi$ [corresponding to an upward translation and $u \in (0, \pi)$] and seek solutions satisfying $0 < u < v < \frac{1}{2}\pi$ (corresponding to $\lambda_1 > \lambda_2 > 0$). The conditions

$$\cos u = \frac{\cos\omega}{\cos\theta}, \quad \sin v = \frac{\sin u}{\sin\omega}$$

follow from (38) and imply that the solution is unique.

Note that $\cos\omega > 0,\ \cos\theta < 0$ will yield solutions with $\frac{1}{2}\pi < v < u < \pi$ which includes a rotation of the ellipsoid by π as well as a translation, or $(\pm\cos\omega, 0, \sin\omega)$ to $(\mp|\cos\theta|, 0, \sin\theta)$.

Maps that take pairs of the form $(0, \pm\cos\theta, \sin\theta)$ to $(0, \pm\cos\omega, \sin\omega)$ require $|\lambda_2| \geq |\lambda_1|$ and can be treated similarly. All other situations can be reduced to these after suitable rotations. Hence, it follows that when two distinct extreme points have a common pair of points on the image of the Bloch sphere, their pre-images must be distinct.

Theorem 16 *If Φ is an extreme point of $\mathcal{S}$, then at least one point in the image $\Phi(\mathcal{D})$ is a pure state. If $\Phi(\mathcal{D})$ contains two or more pure states, it must be in the closure of the extreme points of $\mathcal{S}$. Moreover, if the intersection of $\Phi(\mathcal{D})$ with the Bloch sphere consists of exactly two non-orthogonal pure states, then Φ must be a non-unital extreme point; while $\Phi(\mathcal{D})$ contains the entire Bloch sphere if and only if it is a unital extreme point.*

A non-extreme point can have an image point on the boundary of the Bloch sphere only if all its extreme components have the same pre-image for that point. Indeed, if $\Phi = \alpha\Phi_1 + (1-\alpha)\Phi_2$ with $0 < \alpha < 1$ is not extreme and Φ has an image point $\mathbf{x} = \Phi(\mathbf{u})$ on the boundary of the Bloch sphere, then it follows immediately from the strict convexity of the Euclidean unit ball that $\mathbf{x} = \Phi(\mathbf{u}) = \Phi_1(\mathbf{u}) = \Phi_2(\mathbf{u})$.

One can generate a large class of such maps from any pair of extreme points by applying a suitable "rotation" to one of them. Hence, one can find many non-extreme maps that take one (but not two) pure states to the boundary of the Bloch sphere. A few special cases are worth particular mention.

1. Example of non-extreme maps that reach the boundary:

a) $\lambda_1 = \lambda_2 = \lambda_3 = \mu$ with $0 \leq \mu < 1$. This corresponds to a depolarizing channel in which case Φ maps the unit ball into a sphere of radius μ which can then be translated in *any* direction with $|\mathbf{t}| \leq 1 - |\mu|$. Similar remarks hold when one $\lambda_k = +\mu$ and the other two $-\mu$, since conjugating a stochastic map with one of the Pauli matrices, yields another stochastic map with the signs of any *two* λ_k changed.

b) Let $|\lambda_1| = |\lambda_2| = \mu$, $|\lambda_3| \geq \mu^2$, and $|t_3| = (1 \pm \lambda_3)$ so that R_Φ is given by (8). This is possible if, in addition, λ_3 and $\lambda_1\lambda_2$ have the same sign, and $t_1 = t_2 = 0$. Under these conditions R_Φ, is always a contraction, but is unitary only for $|\lambda_3| = \mu^2$. However, the condition $|t_3| = (1 \pm \lambda_3)$ implies that the North or South pole is a fixed point. Hence, even when $|\lambda_3| > \mu^2$ so that Φ is not extreme, the ellipsoid is shifted to the boundary in a manner analogous to an amplitude damping channel.

In general, however, it is not possible to translate the image ellipsoid to the boundary of the unit ball. On the contrary, there are many situations in which the translation is severely limited despite contraction of the ellipsoid.

In discussing this question the pair of inequalities (11) play an important role and it is natural to ask if (11) can be extended to

$$(\lambda_1 \pm \lambda_2)^2 \leq (1 \pm \lambda_3)^2 - |\mathbf{t}|^2 \tag{39}$$

when more than one t_k is non-zero. By rewriting (9) and (10) in the form

$$\begin{aligned}(\lambda_1 + \lambda_2)^2 &\leq (1+\lambda_3)^2 - |\mathbf{t}|^2 - (t_1^2 + t_2^2)\left(\frac{2\lambda_3}{1-\lambda_3 \pm t_3}\right) \\ (\lambda_1 - \lambda_2)^2 &\leq (1-\lambda_3)^2 - |\mathbf{t}|^2 + (t_1^2 + t_2^2)\left(\frac{2\lambda_3}{1+\lambda_3 \pm t_3}\right)\end{aligned}$$

one can see that, depending on the sign of λ_3 one of the pair inequalities in (39) holds. However, in general we do not expect that (39) will hold with both signs.

2. Example of maps with limited translation:

a) If $(\lambda_1 \pm \lambda_2)^2 = (1 \pm \lambda_3)^2$ holds with exactly one sign, then $\lambda_1, \lambda_2, \lambda_3$ lie on the surface of the tetrahedron defined by (II), but on the interior of one of the four faces. However, even when equality holds with only one sign, (11) implies that $\mathbf{t} = 0$ so that no translation is possible. Thus, one can find maps with all $|\lambda_k| < 1$ for which the image $\Phi(\mathcal{D})$ is an ellipsoid strictly contained within the unit ball but no translation in any direction is possible.

b) A map with $\lambda_1 = \lambda_2 = \lambda_3 = \mu < 0$ is *not* completely positive unless $|\mu| \leq \frac{1}{3}$. However, unlike the case with $\mu > 0$, such maps can not be translated to the boundary; in fact, for $\mu = -\frac{1}{3}$, one must have $\mathbf{t} = 0$ so that no translation is possible.

c) When $\lambda_1 = \lambda_2 = \mu$, $\lambda_3 = 0$ which implies $|\mu| \leq \frac{1}{2}$ and the image of Φ is a circle in a plane parallel to the equator. Moreover, the equations following (39) imply

$$4\mu^2 \leq (1 - |\mathbf{t}|^2) \tag{40}$$

when $\lambda_3 = 0$. Thus, when $\mu = \frac{1}{2}$ no translation the image $\Phi(\mathcal{D})$ is possible despite the fact that the ellipsoid has shrunk to a flat disk of radius $\frac{1}{2}$. For $0 < \mu < \frac{1}{2}$ translation is possible, but clearly never reaches the boundary of pure states.

Example (2a) includes maps with $\lambda_1 = \lambda_2 = \mu$, $\lambda_3 = 2\mu - 1$, which may seem to contradict Example (1b). However, the condition $2\mu - 1 = \lambda_3 > \lambda_1^2 = \mu^2$ is never satisfied since if would imply $-\mu^2 + 2\mu - 1 = -(1-\mu)^2 > 0$. Hence the conditions for these two examples do not overlap.

Example (2b) illustrates that the ellipsoid picture, while useful, is incomplete. The eigenvalues $[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$ and $[\frac{1}{3}, \frac{1}{3}, -\frac{1}{3}]$ both yield the same ellipsoid (actually, a sphere of radius $\frac{1}{3}$). The former can be translated in any direction with $|\mathbf{t}| \leq 2/3$; however, the latter can not be translated at all without losing complete positivity.

It is tempting to think of a stochastic map Φ as the composition of a rotation, contraction (to an ellipsoid), translation, and another rotation. However, this is not accurate in the sense that the individual maps in this process will no longer be stochastic. In general, it is *not* possible to write a non-unital map as the composition of a unital map (which contracts the Bloch sphere to an ellipsoid) and a translation. Such a translation would have to be representable in the form $\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{t} & I \end{pmatrix}$ which does not satisfy the conditions for complete positivity.

3.3 Geometry of stochastic maps

In the previous section we discussed the geometry of the images of various types of stochastic maps. We now make some remarks about the geometry of the set of stochastic maps $\mathcal{S}$ itself.

In a fixed basis, the convex set of stochastic maps with $\mathbf{t} = 0$ forms a tetrahedron that we can describe using the λ_k in (3) which we write as $[\lambda_1, \lambda_2, \lambda_3]$. The extreme points of the tetrahedron are $[1,1,1], [1,-1,-1], [-1,1,-1], [-1,-1,1]$. The edges connecting any two of these extreme points have the form $[\pm 1, s, \pm s]$ up to permutation and correspond to quasi-extreme points of type (III). Although this tetrahedron lies in a 3-dimensional space in a fixed basis, an arbitrary bistochastic map requires 9 parameters that could be chosen either as the 9 elements of the real 3×3 matrix T, or as the three λ_k and two sets of Euler angles corresponding to the two rotations required to reduce Φ to diagonal form. In a fixed basis, an arbitrary bistochastic map can be written (uniquely) as a convex combination of 4 true extreme points, or non-uniquely as a convex combination of two quasi-extreme points on the "edges". As the changes of basis rotate the tetrahedron (considered as a set of diagonal 3×3 matrices in $\mathbf{R}^9$), we obtain a much larger set so that (by Theorem 14) any non-extreme map can be written as the midpoint of two "generalized" extreme points in different bases.

We now describe the convex set analogous to the tetrahedron (in a fixed basis) for non-zero $\mathbf{t}$ with $\mathbf{t}$ fixed. We will consider the case when $t_1 = t_2 = 0$ and $t_3 > 0$ is fixed and the λ_k vary. Thus, we are in a 3-dimensional submanifold of $\mathbf{R}^6$. It follows from (17) that the extreme points are given by the curve $[\cos u, \cos v, \cos u \cos v]$ subject to the constraint $\sin u \sin v = t_3$. This can be written in parametric form as a pair of curves

$$[\cos u, \pm \cos\left(\sin^{-1}\left[\frac{t_3}{\sin u}\right]\right), \pm \cos u \, \cos\left(\sin^{-1}\left[\frac{t_3}{\sin u}\right]\right)] \tag{41}$$

with $\sin^{-1}(t_3) \leq u \leq \pi - \sin^{-1}(t_3)$. This curve forms the boundary of Figure 3 below. Letting $\alpha^2 = 1 - t_3$, we see that the curve (41) passes through the four points $[\alpha, \alpha, \alpha^2]$, $[-\alpha, -\alpha, \alpha^2]$, $[\alpha, -\alpha, -\alpha^2]$, $[-\alpha, \alpha, -\alpha^2]$ when $\sin u = \pm \sin v$. Thus, as t_3 moves away from zero, the corners of the tetrahedron are replaced by these points as shown in Figure

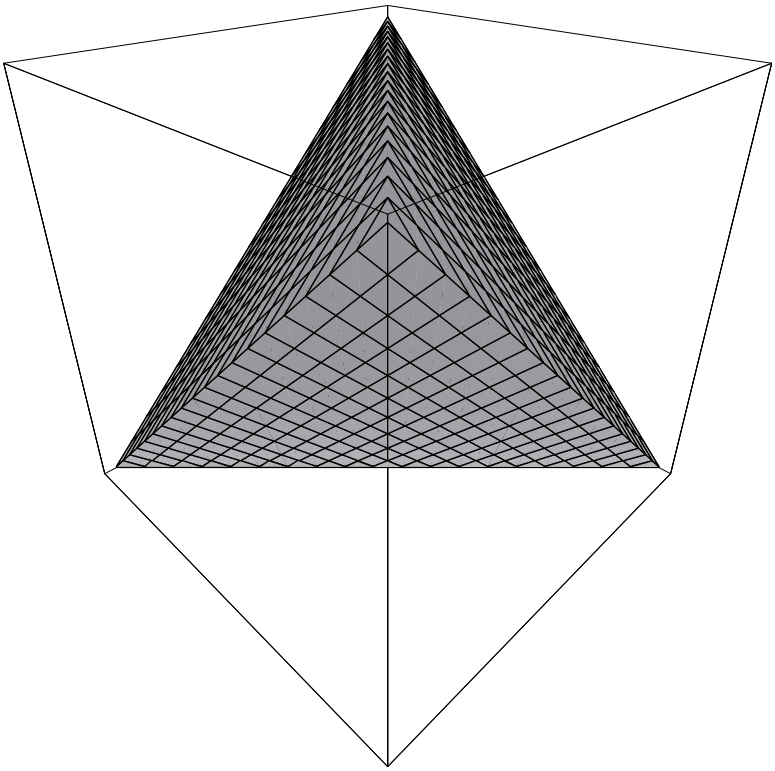


Figure 2: the tetrahedron

3 which might be described as an asymmetric rounded tetrahedron. We have shown the tetrahedron for comparison, placed in such a way that the edge connecting the vertices $[-1,-1,1]$ and $[1,1,1]$ is pointing towards us, and oriented the rounded tetrahedron similarly. The 4 points above split the curve into four pieces, corresponding to four of the six edges of the tetrahedron. In place of the remained two edges there appear two segments (still on the surface of the "rounded" tetrahedron) connecting pairs of the four above points for which λ_3 has the same sign. Unlike the case $t_3 = 0$ these lines are not extreme points, even in a generalized sense.

The inequalities (11) imply that, as for the tetrahedron, the figure in (3b) has rectangular cross sections for fixed λ_3 with $|\lambda_1 \pm \lambda_2| \leq \sqrt{(1+\lambda_3)^2 - t_3^2}$. When $t_3 \neq 0$, the corners depend non-linearly on λ_3, yielding a curve of true extreme points. When $t_3 = 0$, the linearity in λ_3 yields a line segment so that we no longer have "true" extreme points.

The cases $\mathbf{t} = (\sqrt{1-\alpha^2}, 0, 0)$ and $\mathbf{t} = (0, \sqrt{1-\alpha^2}, 0)$ are similar. The only difference is the orientation and displacement, e.g., whether the point [1,1,1] is replaced by $[\alpha^2, \alpha, \alpha]$ or to $[\alpha, \alpha^2, \alpha]$.

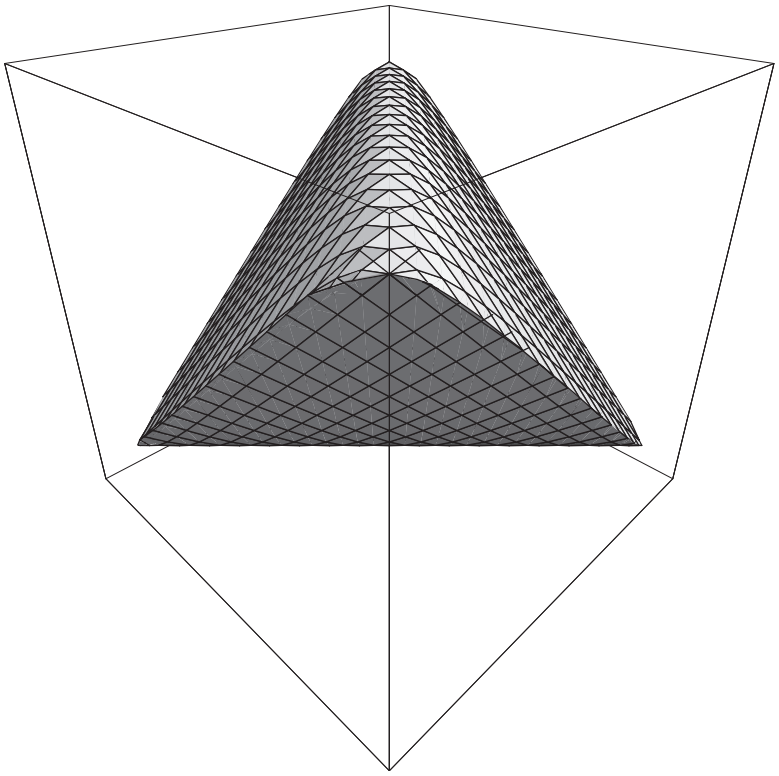


Figure 3: the “rounded tetrahedron”

3.4 Channel capacity

The capacity of a quantum communication channel depends on the particular protocols allowed for transmission and measurement as well as the noise of the channel. See [3, 14] for definitions and discussion of some of these. We consider here only the so-called *Holevo capacity* that corresponds to communication using product signals and entangled measurements and is now believed (primarily on the basis of numerical evidence) to be the maximum capacity associated with communication that does not involve prior entanglements. The Holevo capacity is given by

$$C_{\mathrm{Holv}}(\Phi) = \sup_{\pi_j,\,\rho_j} \left(S[\Phi(\rho)] - \sum_j \pi_j S[\Phi(\rho_j)] \right), \tag{42}$$

where $S(P) = -\mathrm{Tr}\,(P \log P)$ denotes the von Neumann entropy of the density matrix P, ρ_j denotes a set of pure state density matrices, π_j a discrete probability vector, and $\rho = \sum_j \pi_j \rho_j$. It is easy to see that for extreme points of the type (II) and (III), $C_{\mathrm{Holv}}(\Phi) = \log 2$ so that the capacity attains its maximal value, and for the completely noisy channel in (IC), $C_{\mathrm{Holv}}(\Phi) = 0$.

By contrast, the non-unital situation (I) is more interesting because it includes channels for which the capacity (42)is strictly bigger than the classical Shannon capacity, i.e., the capacity for communication restricted to product input and measurements. Such channels demonstrate a definite quantum advantage. Moreover, the capacity is, in general, achieved *neither* with a pair of orthogonal input states, which would yield $h(\sin u \sin v) - h(\sqrt{1 - \sin^2 u \cos^2 v})$, *nor* with a pair of minimal entropy states, which would yield $h(\frac{\sin u}{\sin v})$ where

$$h(t) = -\frac{1+t}{2}\log\frac{1+t}{2} - \frac{1-t}{2}\log\frac{1-t}{2}. \tag{43}$$

This behavior was observed first by Fuchs for a non-extreme channel [6]; subsequently, the amplitude-damping channel was studied by Schumacher and Westmoreland [27]. In both cases the work was numerical, but suggests that this situation is generic for channels that are translated orthogonal to the major axis of the ellipse.

In the case (IC) when the ellipsoid shrinks to a line, the capacity can be computed explicitly as

$$\begin{aligned} C_{\mathrm{Holv}}(\Phi) &= h(\sqrt{t_1^2 + t_2^2} = h(\cos u) \\ &> \log 2 - h(\sin u) \end{aligned} \tag{44}$$

since $h(\cos\theta) + h(\sin\theta) \geq \log 2$. The expression $\log 2 - h(\sin u)$ is the capacity of the unshifted channel. In this case, it is also the classical Shannon capacity. Holevo [8, 9] introduced such channels and showed that they suffice to demonstrate the "quantum advantage" mentioned above, although the both the minimal entropy and capacity are achieved with orthogonal inputs.

Remark: Holevo called channels of the form (IC) "binary" because their Kraus operators can be represented in the form $A_1 = |e_1\rangle\langle\psi_+|,\ A_2 = |e_2\rangle\langle\psi_-|$ where the states $|e_k\rangle$ are orthonormal and $\psi_\pm$ can, in principle, be any pair of states, such as corresponding to $(\pm\cos x, \sin x)$. When e_1, e_2 are the eigenvectors $\frac{1}{\sqrt{2}}(1, \pm 1)$ of σ_x, the resulting Kraus operators are

$$A_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos x & \sin x \\ \cos x & \sin x \end{pmatrix}, \quad A_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -\cos x & \sin x \\ \cos x & -\sin x \end{pmatrix}$$

One can easily check that this yields a map of type (IC) with $u = 2x - \frac{\pi}{2}, v = \frac{1}{2}\pi$ but the Kraus operators do not have the form (18).

Acknowledgment: It is a pleasure to thank Professor Chris King, Professor Denes Petz and Dr. Barbara Terhal for stimulating and helpful discussions, Dr. Christopher Fuchs for bringing reference [22] to our attention, and Professor Roger Horn for clarifying remarks about Lemma 8. After an earlier version of this paper [28] was posted we learned that Dr. Eleanor Rieffel and Dr. C. Zalka had independently obtained some similar results about the extreme points. We are grateful to them for a number of useful comments, including drawing our attention to Choi's criterion for extremality which helped to streamline the proofs, and raising the question of the minimal number of extreme points needed for a given map. MBR would also like to thank Professor Ruedi Seiler for his hospitality at the Technische Universität Berlin, where part of this paper was written. The pictures were done with Mathematica and Adobe Illustrator. We want to thank Joachim Werner for his help in creating them.

References

[1] A. Fujiwara and P. Algoet, "Affine parameterization of completely positive maps on a matrix algebra", preprint. A. Fujiwara and P. Algoet, "One -to-one parameterization of quantum channels", Phys Rev A, vol 59, 3290 –3294, 1999.

[2] G.G. Amosov, A.S. Holevo, and R.F. Werner, "On Some Additivity Problems in Quantum Information Theory" preprint (lanl:quant-ph/0003002).

[3] C. H. Bennett and P.W. Shor, "Quantum Information Theory" *IEEE Trans. Info. Theory* (1998).

[4] M-D Choi, "Completely Positive Linear Maps on Complex Matrices" *Lin. Alg. Appl.* **10**, 285–290 (1975).

[5] M-D Choi, "A Schwarz Inequality for Positive Linear Maps on C^* Algebras" *Ill. J. Math.* **18**, 565–574 (1974).

[6] C. Fuchs "Nonorthogonal quantum states maximize classical information capacity", *Phys. Rev. Lett* **79** 1162–1165 (1997). preprint (lanl: quant-ph/9703043)

[7] R. Haag and D. Kastler, " An Algebraic Approach to Quantum Field Theory", *J. Math. Phys.* **5**, 848–861 (1964).

[8] A. S. Holevo, "On the capacity of quantum communication channel", *Probl. Peredachi Inform.*, **15**, no. 4, 3-11 (1979) (English translation: *Problems of Information Transm.*, **15**, no. 4, 247-253 (1979)).

[9] A. S. Holevo, "Quantum coding theorems", *Russian Math. Surveys*, **53**(6), 1295-1331 (1999). quant-ph/9809023

[10] R.A. Horn and C.R. Johnson, *Matrix Analysis* (Cambridge University press, 1985).

[11] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis* (Cambridge University press, 1991).

[12] M. Horodecki and P. Horodecki, "Reduction criterion of separability and limits for a class of protocols of entanglement distillation" xxx.lanl.gov preprint quant-ph/9708015.

[13] C. King and M.B. Ruskai, "Minimal entropy of states emerging from noisy quantum channels" *IEEE Trans. Info. Theory* **47**, 192–209 (2001).

[14] C. King and M.B. Ruskai, "Capacity of Quantum Channels Using Product Measurements" *J. Math. Phys.* **42**, 87–98 (2001)

[15] K. Kraus, "General state changes in quantum theory" *Ann. Physics* **64**, 311–335 (1971).

[16] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer-Verlag, 1983).

[17] S.-H. Kye, "On the convex set of completely positive linear maps in matrix algebra," Math. Proc. Camb. Phil. Soc. 122, 45–54 (1997).

[18] E. Lieb and M.B. Ruskai "Some Operator Inequalities of the Schwarz Type" *Adv. Math* **12**, 269–273 (1974).

[19] G. Lindblad, "Completely Positive Maps and Entropy Inequalities" *Commun. Math. Phys.* **40**, 147-151 (1975).

[20] S. Lloyd, *Science*, **273**, 1996, p. 1073.

[21] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[22] Niu and Griffiths, "Two Qubit Copying Machine for Economical Quantum Eavesdropping" quant-ph/9810008

[23] V. I. Paulsen, *Completely bounded maps and dilations* Pitman Research Notes in Mathematics Series, 146. (Longman Scientific and Technical, Harlow; John Wiley and Sons, Inc., New York, 1986).

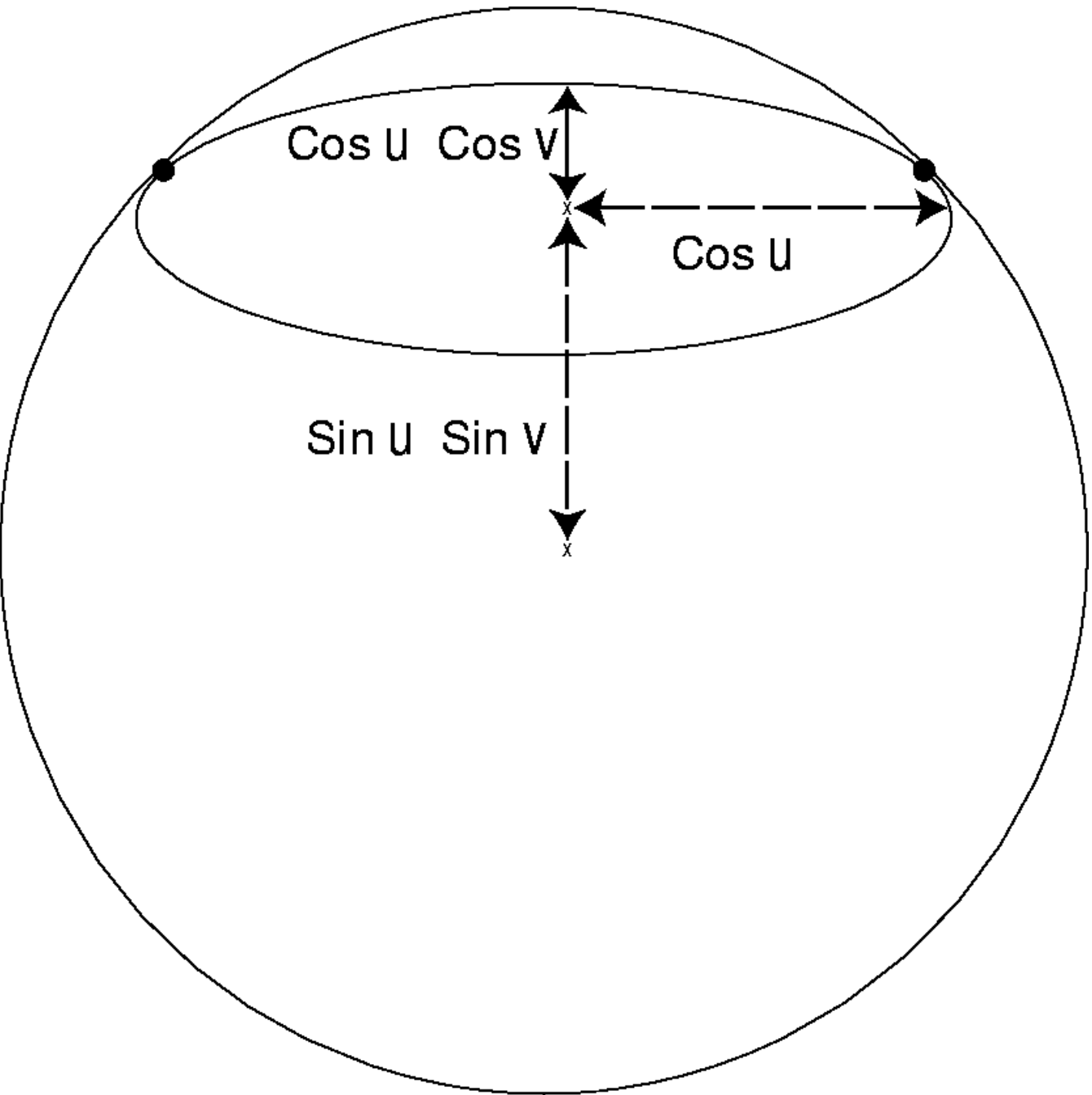
[24] G. Pedersen *C*-algebras and their automorphism groups* (Academic Press, 1979).

[25] E. Rieffel and C. Zalka, private communications and unpublished manuscript.

[26] B. M. Terhal, I. L. Chuang, D. P. DiVincenzo, M. Grassl, and J.A. Smolin, "Simulating quantum operations with mixed environments", quant-ph/9806095.

[27] B. Schumacher and M. D. Westmoreland, "Optimal signal ensembles" xxx.lanl.gov preprint quant-ph/9912122.

[28] M.B. Ruskai S. Szarek and E. Werner xxx.lanl.gov preprint quant-ph/0005004


Cos u Cos V
Cos u
Sin u Sin V