The scope of a quantum channel

Keye Martin

Center for High Assurance Computer Systems
Naval Research Laboratory
Washington, DC 20375

kmartin@itd.nrl.navy.mil

Abstract

The capacity of a classical channel is a single number that to a large extent captures its ability to transmit information. Though analogous notions exist for quantum channels, the use of a single number is not particularly informative. For instance, any basis of the state space can be used to represent classical bits, and each representation leads to a classical channel with a capacity all its own. So the ability of a quantum channel to transmit information should minimally depend on all of the different ways classical bits can be represented: instead of a single number, we should measure it with a *set* of numbers. We call this set, which consists of the range of achievable classical capacities, the *scope*.

For unital channels on qubits, we establish that scope is in fact a *compact interval*, provide an exact characterization of it and then show how to systematically calculate it. These results, which rely extensively on the algebraic structure of quantum channels, are then used to design an *adaptive* scheme for communication in which the participants can maximize the information transmitted after first determining the state of the environment (which we show how to do) and then performing a scope calculation. When this technique is applied to quantum cryptography, it becomes possible to minimize the error rate over any time interval where the environment remains stable. For familiar forms of noise, like bit flipping, the error rate is cut in half.

1 Introduction

We understand communication as being the following process: a sender takes some information, represents it in a certain way, sends this representation to a receiver, who then performs some operation on the representation to extract the information. To measure the amount of information being transmitted, one needs some way of measuring the correlation between the sent and the received. Many factors can affect the amount of information that gets transmitted, the most well-known being noise in the environment. This paper is about how our ability to transmit classical information through a noisy quantum channel depends on how that information is represented.

Given a quantum channel f, each basis of the state space defines a particular way to represent classical information. For each such representation, there is an associated classical channel having a capacity all its own. The range of capacities achieved as we vary over all possible representations is called the *scope* of the channel f. It is denoted by $s(f)$. Intuitively, the largest value of scope tells us how we should represent information so that we *maximize* the amount transmitted, while the

smallest value could be used to measure the optimal performance of a method designed to interrupt communication or remove a steganographic message. If all representations are not available to us, then we may want to choose the *best available*, which might mean considering a value in between the maximum and minimum.

In order to calculate scope, we need a simpler representation of quantum channels, and for this we turn in the next section to the Bloch representation. We give a self-contained presentation of many well-known results whose proofs are difficult to find in the literature, as well as some more obscure results that are at times taken for granted. After studying the Bloch representation, we obtain an elementary characterization of the channels that are our principal concern in this paper: the *unital* qubit channels, which occur naturally in communication as conservative models of noise. It turns out that such a channel can always be canonically represented by a convex sum of rotations in $\mathbb{R}^3$. We consider some of their fascinating algebraic properties and use them to establish that qubit unitality is the quantum analogue of a binary symmetric channel.

After our study of unitality, we have the simple language needed to not only characterize scope precisely as the solution to an optimization problem, but also to give a systematic method for solving it: each unital channel can be replaced by a symmetric unital channel with the same scope – but the scope of a symmetric channel can be calculated from its eigenvalues! This is very surprising given that we are then able to show that each symmetric channel is a convex sum of *four* involutive rotations which collectively comprise a copy of the Klein-four group. We give scope calculations for projective measurements, bit flipping, bit-phase flipping, phase flipping, depolarization, the two-Pauli channel, the intercept-resend attack in quantum cryptography and general unitary evolution.

Finally, we focus on the significance of scope to quantum information itself. For symmetric channels, we show that the largest value in the scope coincides with the Holevo capacity and that this is also true for unitary channels. We show how the ability to perform scope calculations can be used to minimizing the error rate in protocols like QKD, leading to a new idea we call *adaptive quantum communication*. Along the way, it becomes clear that scope can be used to classify physical effects according to the degree with which they disturb a system.

2 The Bloch representation

Let $\mathcal{H}^2$ denote a two dimensional complex Hilbert space with specified inner product $\langle\cdot|\cdot\rangle$.

Definition 2.1 A *quantum state* is a density operator $\rho : \mathcal{H}^2 \to \mathcal{H}^2$, i.e., a self-adjoint, positive, linear operator with $\mathrm{tr}(\rho) = 1$. The quantum states on $\mathcal{H}^2$ are denoted Ω^2.

We also sometimes call density operators *mixed states*.

Definition 2.2 A quantum state ρ on $\mathcal{H}^2$ is *pure* if $\mathrm{spec}(\rho) \subseteq \{0, 1\}$.

Pure states always have the form $f : \mathcal{H}^2 \to \mathcal{H}^2$ given by

$$f(u) = \langle\psi|u\rangle\psi$$

for some unit vector $\psi \in \mathcal{H}^2$. However, because unit vectors in quantum mechanics are normally written $|\psi\rangle$, the pure state associated to a unit vector is normally denoted $|\psi\rangle\langle\psi|$. Such operators define *projections* on $\mathcal{H}^2$. If $|\psi\rangle$ and $|\phi\rangle$ are orthogonal, then $|\psi\rangle\langle\psi| + |\phi\rangle\langle\phi| = I$.

Definition 2.3 The *spin operators* are denoted $\{I, \sigma_x, \sigma_y, \sigma_z\}$ and given by

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

Each spin operator is self adjoint and unitary. They have many wonderful and valuable properties, including:

- $\sigma_i^2 = I$ (involutivity),
- $\sigma_i\sigma_j + \sigma_j\sigma_i = 0$ (anticommutativity),

for distinct $i, j \in \{x, y, z\}$. Notice that

$$(\forall i \in \{x, y, z\})\ \mathrm{tr}(\sigma_i) = 0 \tag{1}$$

And since $\mathrm{tr}(ab) = \mathrm{tr}(ba)$ for any two matrices, linearity of the trace gives

$$\mathrm{tr}(\sigma_i\sigma_j) = \mathrm{tr}(\sigma_j\sigma_i) = 0 \tag{2}$$

for distinct $i, j \in \{x, y, z\}$. Since the spin operators anticommute and are all self adjoint,

$$\langle \sigma_i(v) | \sigma_j(v) \rangle = -\langle \sigma_j(v) | \sigma_i(v) \rangle \tag{3}$$

for distinct $i, j \in \{x, y, z\}$. These facts will be important to us later.

There is a 1-1 correspondence between density operators on a two dimensional state space and points on the unit ball $\mathbb{B}^3 = \{x \in \mathbb{R}^3 : |x| \leq 1\}$: each density operator $\rho : \mathcal{H}^2 \to \mathcal{H}^2$ can be written uniquely as

$$\rho = \frac{1}{2}(I + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z) := \frac{1}{2}(I + r \cdot \vec{\sigma})$$

where $r = (r_x, r_y, r_z) \in \mathbb{R}^3$ satisfies $|r| = \sqrt{r_x^2 + r_y^2 + r_z^2} \leq 1$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of spin operators. The vector r is called the *Bloch vector* associated to ρ. We sometimes write $r = [\![\rho]\!]$ and denote the *bijection* between Ω^2 and $\mathbb{B}^3$ as $[\![\cdot]\!] : \Omega^2 \to \mathbb{B}^3$. For the sake of completeness:

Lemma 2.4 $[\![\cdot]\!] : \Omega^2 \to \mathbb{B}^3$ *is a bijection.*

Proof. The spin operators form a basis for the vector space $M_{2\times2}(\mathbb{C})$ of 2×2 matrices with complex entries. Thus, any $A \in M_{2\times2}(\mathbb{C})$ can be written as

$$A = t \cdot I + x \cdot \sigma_x + y \cdot \sigma_y + z \cdot \sigma_z$$

where t, x, y and z are all complex. The following are all standard well-known facts:

- A is self-adjoint iff $(t, x, y, z) \in \mathbb{R}^4$
- A is positive iff $|(x, y, z)| \leq t$
- $\mathrm{tr}(A) = 1$ iff $t = 1/2$

Thus, for a density operator ρ, we have $t = 1/2$ and $|(x, y, z)| \leq 1/2$. If we set $r = 2(x, y, z)$, then

$$\rho = \frac{1}{2}(I + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z) := \frac{1}{2}(I + r \cdot \vec{\sigma})$$

with $r \in \mathbb{B}^3$. That such an r is unique follows from the fact that spin matrices form a basis for $M_{2\times 2}(\mathbb{C})$. This proves that $[\![\cdot]\!]$ is an injective function. To see that it is surjective, we simply take $r \in \mathbb{B}^3$ and form the matrix $\rho = \frac{1}{2}(I + r \cdot \vec{\sigma})$, for which we have $[\![\rho]\!] = r$. $\square$

Proposition 2.5 *Let ρ and σ be density operators with respective Bloch vectors r and s.*

(i) *The eigenvalues of ρ are $(1 \pm |r|)/2$,*

(ii) *The base two von Neumann entropy of ρ is $S\rho = H((1 + |r|)/2) = H((1 - |r|)/2)$, where $H : [0, 1] \to [0, 1]$ is the base two Shannon entropy,*

(iii) *If ρ and σ are pure states and $r + s = 0$, then ρ and σ are orthogonal, and thus form a basis for the state space. Conversely, the Bloch vectors associated to a pair of orthogonal pure states form antipodal points on the sphere.*

(iv) *The Bloch vector for a convex sum of mixed states is the convex sum of the Bloch vectors.*

(v) *The Bloch vector for the completely mixed state $I/2$ is $0 = (0, 0, 0)$.*

Proof. For (i), we simply use the fact that

$$\rho = \frac{1}{2}(I + r \cdot \vec{\sigma}) = \frac{1}{2}\begin{pmatrix} 1 + r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix}$$

where r is the Bloch vector for ρ. Assertion (ii) is then immediate since $S\rho$ is equal to the Shannon entropy of the eigenvalues of ρ.

For (iii), since ρ and σ are pure states, there are unit vectors $|\psi\rangle$ and $|\phi\rangle$ such that $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\phi\rangle\langle\phi|$. Because $r + s = 0$, $\rho + \sigma = I$, so applying this operator to $|\phi\rangle$ gives

$$|\psi\rangle \cdot \langle\psi|\phi\rangle + |\phi\rangle \cdot 1 = |\phi\rangle \quad \Longrightarrow \quad |\psi\rangle \cdot \langle\psi|\phi\rangle = 0$$

But $|\psi\rangle$ is a unit vector, so this means $\langle\psi|\phi\rangle = 0$, proving that ρ and σ are orthogonal pure states. Conversely, suppose that $|\psi\rangle$ and $|\phi\rangle$ are orthogonal pure states with Bloch vectors r and s. Then their associated density operators are

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2}(I + r \cdot \vec{\sigma}) \quad \& \quad \sigma = |\phi\rangle\langle\phi| = \frac{1}{2}(I + s \cdot \vec{\sigma})$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of spin operators. We then see that $\rho + \sigma = I + ((r + s) \cdot \vec{\sigma})/2$. Applying this operator to $|\psi\rangle$ gives $v := ((r+s)\cdot\vec{\sigma})|\psi\rangle = 0$. Taking the inner product of v with itself, applying linearity and conjugate linearity of the inner product, and repeatedly applying equation (3) gives $(r_1 + s_1)^2 + (r_2 + s_2)^2 + (r_3 + s_3)^2 = 0$, which implies that $r + s = 0$, proving that r and s are antipodal points on the sphere.

For (iv), let ρ and σ be density operators and $x \in [0,1]$. Then

$$\begin{aligned} x \cdot \rho + (1-x) \cdot \sigma &= \frac{x}{2} \cdot (I + r \cdot \vec{\sigma}) + \frac{1-x}{2} \cdot (I + s \cdot \vec{\sigma}) \\ &= \frac{1}{2} (I + (xr + (1-x)s) \cdot \vec{\sigma}) \end{aligned}$$

which proves that the Bloch vector of a convex sum is the convex sum of the Bloch vectors.

(v) follows from (iii) and (iv): the completely mixed state can be written

$$I/2 = \frac{1}{2} |0\rangle\langle 0| + \frac{1}{2} |1\rangle\langle 1|$$

where $|0\rangle$ and $|1\rangle$ are orthogonal pure states. Since $I/2$ is a convex sum of density operators, its Bloch vector r is the convex sum of the associated Bloch vectors by (iv). However, by (iii), the sum of these Bloch vectors has to be zero since the vectors used to define them are orthogonal. □

The fact that antipodal points on the Bloch sphere correspond to orthogonal pure states is something we will make repeated use of in this paper.

Definition 2.6 A *qubit channel* is a trace preserving function $\varepsilon : \Omega^2 \to \Omega^2$ that is convex linear and completely positive.

Each qubit channel $\varepsilon : \Omega^2 \to \Omega^2$ can be represented by a function $f_\varepsilon : \mathbb{B}^3 \to \mathbb{B}^3$ on the *Bloch sphere* $\mathbb{B}^3$. The map f_ε is called the *Bloch representation* of ε and is the unique map which makes the following diagram commute:

$$\begin{array}{ccc} \Omega^2 & \xrightarrow{\varepsilon} & \Omega^2 \\ {\scriptstyle [\![\cdot]\!]}\downarrow & & \downarrow{\scriptstyle [\![\cdot]\!]} \\ \mathbb{B}^3 & \xrightarrow[f_\varepsilon]{} & \mathbb{B}^3 \end{array}$$

Formally, f_ε is defined by $f_\varepsilon(r) = [\![\varepsilon([\![r]\!]^{-1})]\!]$ and satisfies $[\![\varepsilon(\rho)]\!] = f_\varepsilon([\![\rho]\!])$ for all $\rho \in \Omega^2$.

Proposition 2.7 *Let ε be a qubit channel and let f_ε be its Bloch representation.*

(i) *The function f_ε is convex linear.*

(ii) *Composition of qubit channels corresponds to composition of Bloch representations: for channels $\varepsilon_1, \varepsilon_2$, we have $f_{\varepsilon_1 \circ \varepsilon_2} = f_{\varepsilon_1} \circ f_{\varepsilon_2}$.*

(iii) *Convex sum of qubit channels corresponds to convex sum of Bloch representations: for channels $\varepsilon_1, \varepsilon_2$ and $x \in [0,1]$, we have $f_{x\varepsilon_1 + \bar{x}\varepsilon_2} = x f_{\varepsilon_1} + \bar{x} f_{\varepsilon_2}$.*

Proof. Throughout this proof, $x \in [0,1]$ and ρ and σ are density operators with Bloch vectors r and s respectively. Let us also recall in general that a qubit channel ε is related to its Bloch representation f_ε by $[\![\varepsilon\rho]\!] = f_\varepsilon([\![\rho]\!])$.

(i) The function f_ε is convex linear:

$$\begin{aligned} f_\varepsilon(xr + (1-x)s) &= f_\varepsilon(x[\![\rho]\!] + (1-x)[\![\sigma]\!]) \\ &= f_\varepsilon([\![x\rho + (1-x)\sigma]\!]) \qquad \text{(Prop. 2.5(iv))} \\ &= [\![\varepsilon(x\rho + (1-x)\sigma)]\!] \\ &= [\![x\varepsilon(\rho) + (1-x)\varepsilon(\sigma)]\!] \\ &= x[\![\varepsilon(\rho)]\!] + (1-x)[\![\varepsilon(\sigma)]\!] \\ &= xf_\varepsilon([\![\rho]\!]) + (1-x)f_\varepsilon([\![\sigma]\!]) \\ &= xf_\varepsilon(r) + (1-x)f_\varepsilon(s) \end{aligned}$$

(ii) The representation of a composition is the composition of the representations:

$$f_{\varepsilon_1\varepsilon_2}(r) = f_{\varepsilon_1\varepsilon_2}([\![\rho]\!]) = [\![\varepsilon_1(\varepsilon_2(\rho))]\!] = f_{\varepsilon_1}([\![\varepsilon_2(\rho)]\!]) = f_{\varepsilon_1}(f_{\varepsilon_2}([\![\rho]\!])) = f_{\varepsilon_1}(f_{\varepsilon_2}(r)).$$

(iii) The representation of a convex sum is the convex sum of the representations:

$$\begin{aligned} f_{x\varepsilon_1+(1-x)\varepsilon_2}(r) &= f_{x\varepsilon_1+(1-x)\varepsilon_2}([\![\rho]\!]) \\ &= [\![x\varepsilon_1(\rho) + (1-x)\varepsilon_2(\rho)]\!] \\ &= x[\![\varepsilon_1(\rho)]\!] + (1-x)[\![\varepsilon_2(\rho)]\!] \qquad \text{(Prop. 2.5(iv))} \\ &= xf_{\varepsilon_1}([\![\rho]\!]) + (1-x)f_{\varepsilon_2}([\![\rho]\!]) \\ &= xf_{\varepsilon_1}(r) + (1-x)f_{\varepsilon_2}(r) \end{aligned}$$

□

If we now denote f_ε by $[\![\varepsilon]\!]$, we see that the following desirable "semantic properties" all hold:

- $[\![\bot]\!] = 0$, where $\bot$ is the constant qubit channel $\rho \mapsto I/2$,
- $[\![I]\!] = I$
- $[\![f \circ g]\!] = [\![f]\!] \circ [\![g]\!]$
- $[\![xf + (1-x)g]\!] = x[\![f]\!] + (1-x)[\![g]\!]$

These properties allow one to calculate Bloch representations of channels that can be written as compositions and convex sums of simpler channels. The predominant example of a 'simple' channel is a *unitary channel.*

Definition 2.8 A qubit channel ε is *unitary* if there is a unitary operator $U : \mathbb{C}^2 \to \mathbb{C}^2$ such that $\varepsilon(\rho) = U\rho U^\dagger$ for all $\rho \in \Omega^2$.

We will now show how to calculate the Bloch representation of a unitary channel.

Definition 2.9 Let $r_x(\theta), r_y(\theta), r_z(\theta) : \mathbb{B}^3 \to \mathbb{B}^3$ denote the rotations about the x, y and z axes:

$$r_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \qquad r_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \qquad r_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

An orthogonal matrix M is an invertible matrix such that $M^{-1} = M^t$. An orthogonal matrix M with $\det(M) = +1$ is called a *rotation.*

Every rotation M on $\mathbb{R}^3$ can be written in the form $M = r_x(\alpha) \cdot r_y(\beta) \cdot r_z(\theta)$ for angles α, β, θ. Though we are not aware of any proofs in the literature, the relationship between unitary qubit channels and rotations is well-known and is as follows:

Proposition 2.10

(i) *If ε is a unitary qubit channel, then f_ε is a rotation.*

(ii) *For every rotation M on $\mathbb{R}^3$, there is a unitary qubit channel ε such that $f_\varepsilon = M$.*

Proof. (i) By exponentiating the spin operators σ_y and σ_z, we obtain unitary operators

$$\hat{r}_y(\theta) = e^{-i\theta\sigma_y/2} = \begin{pmatrix} \cos\theta/2 & -\sin\theta/2 \\ \sin\theta/2 & \cos\theta/2 \end{pmatrix} \quad \text{and} \quad \hat{r}_z(\theta) = e^{-i\theta\sigma_z/2} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}$$

A standard fact [6] is that any unitary operator U can be written as

$$U = e^{i\lambda}\, \hat{r}_z(\alpha)\, \hat{r}_y(\beta)\, \hat{r}_z(\theta)$$

where $\lambda, \alpha, \beta, \theta$ are real numbers. The unitary channel $\varepsilon(\rho) = U\rho U^\dagger$ defined by U is the same as the one defined by the unitary operator $e^{-i\lambda}U$ since $(e^{i\lambda}U)^\dagger = e^{-i\lambda}U$. Since $\varepsilon = \varepsilon_\alpha \circ \varepsilon_\beta \circ \varepsilon_\theta$, where ε_α, ε_β and ε_θ are the unitary channels associated to $\hat{r}_z(\alpha)$, $\hat{r}_y(\beta)$ and $\hat{r}_z(\theta)$, respectively, Prop. 2.7(iii) tells us that its Bloch representation is

$$f_\varepsilon = f_{\varepsilon_\alpha} \circ f_{\varepsilon_\beta} \circ f_{\varepsilon_\theta}$$

However, calculations you wouldn't wish on your worst enemy reveal that for any density operator ρ written in terms of its Bloch vector $r = (r_x, r_y, r_z)$,

$$\varepsilon_\beta(\rho) = \frac{1}{2}\begin{pmatrix} 1 + (r_z\cos\beta - r_x\sin\beta) & (r_z\sin\beta + r_x\cos\beta) - ir_y \\ (r_z\sin\beta + r_x\cos\beta) + ir_y & 1 - (r_z\cos\beta - r_x\sin\beta) \end{pmatrix}$$

and

$$\varepsilon_\theta(\rho) = \frac{1}{2}\begin{pmatrix} 1 + r_z & (r_x\cos\theta - r_y\sin\theta) - i(r_x\sin\theta + r_y\cos\theta) \\ (r_x\cos\theta - r_y\sin\theta) + i(r_x\sin\theta + r_y\cos\theta) & 1 - r_z \end{pmatrix}$$

from which it follows that

$$f_{\varepsilon_\alpha} = r_z(\alpha) \quad \& \quad f_{\varepsilon_\beta} = r_y(\beta) \quad \& \quad f_{\varepsilon_\theta} = r_z(\theta)$$

and hence that f_ε is a composition of rotations.

(ii) Let M be a rotation on $\mathbb{R}^3$. Then there are real numbers α, β, θ such that $M = r_x(\alpha)r_y(\beta)r_z(\theta)$. By exponentiating the spin matrix σ_x, we obtain a unitary operator

$$\hat{r}_x(\theta) = e^{-i\theta\sigma_x/2} = \begin{pmatrix} \cos\theta/2 & -i\sin\theta/2 \\ -i\sin\theta/2 & \cos\theta/2 \end{pmatrix}$$

As in the previous proof, $r_x(\theta)$ is the Bloch representation of the unitary channel associated to $\hat{r}_x(\theta)$. Using the angles in the decomposition of M into basic rotations, we define a unitary operator $U = \hat{r}_x(\alpha)\hat{r}_y(\beta)\hat{r}_z(\theta)$ and let ε be its associated unitary channel. Since $\varepsilon = \varepsilon_\alpha \circ \varepsilon_\beta \circ \varepsilon_\theta$, where ε_α,

ε_β and ε_θ are the unitary channels associated to $\hat{r}_x(\alpha)$, $\hat{r}_y(\beta)$ and $\hat{r}_z(\theta)$, respectively, Prop. 2.7(iii) tells us that its Bloch representation is

$$f_\varepsilon = f_{\varepsilon_\alpha} \circ f_{\varepsilon_\beta} \circ f_{\varepsilon_\theta} = r_x(\alpha) r_y(\beta) r_z(\theta) = M$$

which finishes the proof. □

The proof of the last result gives a way to calculate the Bloch representation of a unitary channel: if we can write U in the form

$$U = e^{i\lambda}\, \hat{r}_x(\alpha)\, \hat{r}_y(\beta)\, \hat{r}_z(\theta)$$

then its Bloch representation is the rotation $r_x(\alpha) r_y(\beta) r_z(\theta)$. Let us now apply this technique to a very important class of unitary channels, the *spin channels*.

Lemma 2.11 *Let ε_x, ε_y, ε_z denote the unitary channels associated to the spin matrices σ_x, σ_y, σ_z. Then*

$$[\![\varepsilon_x]\!] = r_x(\pi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad [\![\varepsilon_y]\!] = r_y(\pi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad [\![\varepsilon_z]\!] = r_z(\pi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Proof. We have

$$\hat{r}_x(\pi) = e^{-i\pi\sigma_x/2} = -i\sigma_x, \quad \hat{r}_y(\pi) = e^{-i\pi\sigma_y/2} = -i\sigma_y, \quad \hat{r}_z(\pi) = e^{-i\pi\sigma_z/2} = -i\sigma_z.$$

The unitary channels associated to $\hat{r}_x(\pi)$, $\hat{r}_y(\pi)$ and $\hat{r}_z(\pi)$ are the spin channels ε_x, ε_y and ε_z. Thus, the Bloch representations of the spin channels are $r_x(\pi)$, $r_y(\pi)$ and $r_z(\pi)$. □

The matrices in the last lemma are all involutions: an *involution* on a set X is a function $f : X \to X$ for which $f \circ f = 1_X$ i.e. a function that is its own inverse. They are very important:

Definition 2.12 The Bloch representations of the spin channels are denoted s_x, s_y, s_z:

$$s_x := r_x(\pi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad s_y := r_y(\pi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad s_z := r_z(\pi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let us now use s_x, s_y and s_z along with Prop. 2.7 to calculate the representations of some very important channels:

Example 2.13 Let $I : \Omega^2 \to \Omega^2$ denote the identity channel, $\perp$ denote the constant channel $\rho \mapsto I/2$ and let $p \in [0, 1]$ be the probability that some form of noise affects a state.

(i) The Bloch representation of the *bit flipping channel* $\varepsilon = (1-p)I + p\,\varepsilon_x$ is

$$[\![\varepsilon]\!] = (1-p)[\![I]\!] + p[\![\varepsilon_x]\!] = (1-p)I + ps_x$$

(ii) The Bloch representation of the *phase flipping* channel $\varepsilon = (1-p)I + p\,\varepsilon_z$ is

$$[\![\varepsilon]\!] = (1-p)[\![I]\!] + p[\![\varepsilon_z]\!] = (1-p)I + ps_z$$

(iii) The Bloch representation of the *bit-phase flip* channel $\varepsilon = (1-p)I + p\,\varepsilon_y$ is

$$[\![\varepsilon]\!] = (1-p)[\![I]\!] + p[\![\varepsilon_y]\!] = (1-p)I + ps_y$$

(iv) The Bloch representation of the *depolarization channel* $d = p \cdot \perp + (1-p)I$ is

$$[\![d]\!] = p[\![\perp]\!] + (1-p)[\![I]\!] = p \cdot 0 + (1-p)I = (1-p)I$$

Because of the close correspondence between qubit channels and their Bloch representations, from here on we shall refer to both $\varepsilon : \Omega^2 \to \Omega^2$ and its Bloch representation $[\![\varepsilon]\!] = f_\varepsilon : \mathbb{B}^3 \to \mathbb{B}^3$ as being *qubit channels*.

Definition 2.14 The Bloch representation $f_\varepsilon : \mathbb{B}^3 \to \mathbb{B}^3$ of a qubit channel $\varepsilon : \Omega^2 \to \Omega^2$ is also referred to as a *qubit channel*.

For instance, ε_x, ε_y, ε_z are spin channels, as are s_x, s_y and s_z. We now turn to the class of qubit channels that will be our principal concern in this paper.

3 Unitality

The channels in Example 2.13 are among the most important examples of noise. Each arises as a convex sum of unitary channels – a nondeterministic choice between rotations. Such channels are examples of what are called *unital channels*. In this section, we use a simple characterization of unital channels to first establish some of their crucial properties and then to reconcile them with what we feel are their classical counterparts, the *binary symmetric channels*.

Definition 3.1 A qubit channel ε is *unital* if $\varepsilon(I/2) = I/2$. Equivalently, when $f_\varepsilon(0) = 0$.

In addition to the unital channels given in Example 2.13, any *projective measurement*

$$\varepsilon(\rho) = P_0\rho P_0 + P_1\rho P_1$$

for projections $P_0 + P_1 = I$ defines a unital channel. Another example was mentioned above: any convex sum of unitary channels is unital. In fact, it turns out that this is the only way to construct a unital qubit channel: the set of unital qubit channels is both compact and convex and its set of extreme points is exactly the set of unitary channels [3]. We thus obtain an inductive characterization of unital qubit channels in familiar and elementary terms:

Theorem 3.2 *The class of unital channels $\mathcal{U}$ is the smallest class of functions of type $\mathbb{B}^3 \to \mathbb{B}^3$ such that*

- *For each angle θ, $r_x(\theta), r_y(\theta), r_z(\theta) \in \mathcal{U}$,*
- *If $f, g \in \mathcal{U}$, then $f \circ g \in \mathcal{U}$, and*
- *If $f, g \in \mathcal{U}$ and $p \in [0,1]$, then $pf + (1-p)g \in \mathcal{U}$.*

Proof. Let $\mathcal{U}$ denote the smallest set of functions satisfying the three closure properties given in the statement of the theorem. This collection is a subset of the class of all unital channels since all maps in $\mathcal{U}$ take 0 to 0.

For the converse, all rotations on $\mathbb{R}^3$ belong to $\mathcal{U}$ since each can be written as the composition $r_x(\alpha) \cdot r_y(\beta) \cdot r_z(\theta)$. Thus, by Prop. 2.10, the Bloch representations of unitary channels belong to $\mathcal{U}$. Finally, the unitaries are the set of extreme points in the class of unital quantum operations [3], so every unital channel is a convex sum of rotations and thus a member of $\mathcal{U}$. □

Corollary 3.3 *If $f_i \in \mathcal{U}$ and $x_i \in [0,1]$ for $1 \leq i \leq n$ with $\sum_{i=1}^n x_i = 1$, then $\sum_{i=1}^n x_i f_i \in \mathcal{U}$.*

Proof. This is just the usual inductive observation about finite probability distributions:

$$\sum_{i=1}^n x_i f_i = x_1 f_1 + (1 - x_1) \left(\sum_{k=2}^n \frac{x_k f_k}{\sum_{i=2}^n x_i} \right)$$

when $x_1 < 1$. □

Corollary 3.4 *If $f \in \mathcal{U}$, then $|f(x)| \leq |x|$, for all $x \in \mathbb{B}^3$.*

Proof. The proof is a straightforward induction on the set $\mathcal{U}$ of unital channels. □

We now consider some valuable closure properties possessed by the class of unital channels.

Proposition 3.5

- *The maps $\mathrm{id}(x) = x$ and $z(x) = 0$ belong to $\mathcal{U}$,*
- *If $f \in \mathcal{U}$ and $p \in [0,1]$, then $pf \in \mathcal{U}$,*
- *If $f \in \mathcal{U}$, then $f^t \in \mathcal{U}$.*

Proof. The identity map $\mathrm{id}(x) = x$ can be written as $\mathrm{id} = r_x(0)$, so it is unital. The zero map z belongs to $\mathcal{U}$ since it can be written as a convex sum of the spin channels

$$z = \frac{1}{4}\mathrm{id} + \frac{1}{4}s_x + \frac{1}{4}s_y + \frac{1}{4}s_z.$$

If $f \in \mathcal{U}$ and $p \in [0,1]$, then $pf = pf + (1-p)z \in \mathcal{U}$.

To prove closure under transposition, we use induction as follows. First, in the base case, f is a rotation, and then $f^t = f^{-1} \in \mathcal{U}$, since f^{-1} is also a rotation, and $\mathcal{U}$ contains all rotations. If $f^t, g^t \in \mathcal{U}$, then $(f \circ g)^t = g^t \circ f^t \in \mathcal{U}$ and $(pf + (1-p)g)^t = pf^t + (1-p)g^t \in \mathcal{U}$. □

Corollary 3.6 *If $x_i \in [0,1]$ and $f_i \in \mathcal{U}$ for $1 \leq i \leq n$ with $\sum_{i=1}^n x_i \leq 1$, then $\sum_{i=1}^n x_i f_i \in \mathcal{U}$.*

Proof. The zero map z belongs to $\mathcal{U}$ so

$$\sum_{i=1}^n x_i f_i = \sum_{i=1}^n x_i f_i + \left(1 - \sum_{i=1}^n x_i\right) \cdot z \in \mathcal{U}$$

by Corollary 3.3. □

Let us take a first look at the structure of $\mathcal{U}$. In particular, we will see examples of linear operators that do *not* belong to the class of unital channels. To begin, we prove the valuable *trace lemma*:

Lemma 3.7 (The trace lemma) *For any unital channel $f \in \mathcal{U}$, we have* $\mathrm{tr}(f) \in [-1, 3]$.

Proof. Each 3×3 rotation r is a normal matrix, so by Theorem 3.3 of [8], we can find an orthogonal matrix s such that srs^t is block diagonal, each block being either a 1×1 matrix consisting of a real eigenvalue of r or a 2×2 matrix of the form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

Since we are in dimension three, only two cases are possible: either we have all eigenvalues, or we have a matrix of the form

$$srs^t = \begin{pmatrix} c & 0 & 0 \\ 0 & a & b \\ 0 & -b & a \end{pmatrix}$$

where $c = \pm 1$ is a real eigenvalue of r. Using $\det(s) = 1/\det(s^t)$ since $s^t = s^{-1}$,

$$\det(srs^t) = \det(s)\det(r)\det(s^t) = \det(r) = 1 = c(a^2 + b^2)$$

so we see that $c = 1$ and that (a, b) is a point on the unit circle. Then

$$\mathrm{tr}(r) = \mathrm{tr}(I \cdot r) = \mathrm{tr}(ss^t r) = \mathrm{tr}(srs^t) = 1 + 2a \in [-1, 3].$$

In the case of all eigenvalues, we get either a trace of -1 or 3. For an arbitrary unital f written as a convex sum of rotations (r_i),

$$\mathrm{tr}(f) = \mathrm{tr}\left(\sum x_i r_i\right) = \sum x_i \cdot \mathrm{tr}(r_i) \in [-1, 3]$$

by the linearity of the trace. □

Corollary 3.8 *The antipodal map $a(x) = -x$ is not unital.*

Proof. By Lemma 3.7, the antipodal map is not unital because $\mathrm{tr}(a) = -3 \notin [-1, 3]$. □

The antipodal map a takes any qubit x and "flips" it to $a(x) = -x$. That a is not a qubit channel says that there is no *single physical operation* capable of flipping an arbitrary qubit: 'universal' bit flipping is impossible. There is also something here of mathematical interest: a map that negates only one coordinate cannot be unital (if it were, then composing with a spin channel would imply that a is unital). Thus, we can negate any two of (r_x, r_y, r_z), but not only one and not all three.

Proposition 3.9

- *If $f \in \mathcal{U}$ and $f^{-1} \in \mathcal{U}$ exists, then $-f \notin \mathcal{U}$.*
- *If an orthogonal matrix belongs to $\mathcal{U}$, it must be a rotation.*
- *For $f \in \mathcal{U}$, $f^{-1} \in \mathcal{U}$ iff f is a rotation.*

Thus, no orthogonal matrix on $\mathbb{R}^3$ is unital unless it is a rotation. By contrast, on $\mathbb{R}^2$, the antipodal map is a rotation.

Proof. Let $f \in \mathcal{U}$ and suppose that it has an inverse $f^{-1} \in \mathcal{U}$. If $-f \in \mathcal{U}$, then $a = (-f) \circ f^{-1} \in \mathcal{U}$, which is impossible by Corollary 3.8. Thus, $-f \notin \mathcal{U}$.

Suppose now that some $f \in \mathcal{U}$ is defined by an orthogonal matrix. Then it is either a rotation or has a determinant of -1. In the latter case, $a \circ f$ is then a rotation, since it is orthogonal and has a determinant of $\det(a \circ f) = \det(a) \cdot \det(f) = -1 \cdot -1 = 1$. Thus, $a \circ f$ belongs to $\mathcal{U}$, and so does its inverse $f^{-1} \circ a$, since it is also a rotation (having $\det(f^{-1} \circ a) = +1$). By the previous result then, $-(a \circ f) \notin \mathcal{U}$ i.e. $f \notin \mathcal{U}$.

Finally, let $f \in \mathcal{U}$ with $f^{-1} \in \mathcal{U}$. By Corollary 3.4,

$$|x| = |f^{-1}(f(x))| \leq |f(x)| \leq |x|$$

which means that $|f(x)| = |x|$. Then f is a linear isometry and hence an orthogonal matrix. By the previous result, it must be a rotation. □

We know that unital channels can be thought of as nondeterministically choosing between rotations – but can we understand them in relation to *classical* channels? We now establish a few senses in which unitality is the quantum analogue of a "binary symmetric channel." A classical channel is *binary symmetric* when its noise matrix has the form

$$\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$$

where p is the probability that a bit is flipped. Let $\Delta^2 = \{(x, y) : x, y \in [0, 1] \ \& \ x + y = 1\}$.

Proposition 3.10 *The class of binary symmetric channels $\mathcal{B}$ is the smallest class of functions of type $\Delta^2 \to \Delta^2$ such that*

- *The channel that flips all bits belongs to $\mathcal{B}$ i.e.*

$$\text{flip} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathcal{B}$$

- *If $f, g \in \mathcal{B}$, then $f \circ g \in \mathcal{B}$, and*
- *If $f, g \in \mathcal{B}$ and $p \in [0, 1]$, then $pf + (1-p)g \in \mathcal{B}$.*

Proof. Every binary symmetric channel belongs to $\mathcal{B}$ as follows. By closure under composition,

$$1 = \text{flip}^2 = \text{flip} \circ \text{flip} \in \mathcal{B}$$

and this means every binary symmetric channel belongs to $\mathcal{B}$ since

$$\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} = (1-p) \cdot 1 + p \cdot \text{flip} \in \mathcal{B}.$$

For the converse, the flip channel is binary symmetric, and the binary symmetric channels are closed under composition and convex sum. □

One connection between $\mathcal{U}$ and $\mathcal{B}$ is that each can be characterized as the set of functions which arise by taking convex sums of channels that preserve entropy:

Proposition 3.11

- *The entropy preserving classical channels on bits are the identity channel and the flip channel. Each channel in $\mathcal{B}$ is the convex sum of such channels.*
- *The entropy preserving quantum channels on qubits are the rotations. Each channel in $\mathcal{U}$ is the convex sum of such channels.*

Proof. First consider the classical case. If f preserves entropy, then in particular, $Hf(\bot) = \bot$, which means $f(\bot) = \bot$. Using parameters (a, b) for the noise matrix of f, we have

$$\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = (1/2, 1/2) \cdot \begin{pmatrix} a & 1-a \\ b & 1-b \end{pmatrix}$$

and so $a + b = 1$ i.e. f is a binary symmetric channel. Then the noise matrix is determined by a single parameter $p = b$. Because f preserves entropy, $H(x) = H((1-2p)x + p)$ for all $x \in [0, 1]$. This is clearly true when $p = 0$ and $p = 1$. If p is another value that validates this equation for all x, then for $x = 1$, we get

$$H(1-p) = H(1) = 0$$

which implies that either $p = 0$ or $p = 1$. Thus, the only entropy preserving classical binary channels are the identity and the flip channel. Every binary symmetric channel is a convex sum of these two channels, as seen in the proof of Prop. 3.10.

In the quantum case, all we need to show is that the entropy preserving quantum channels are the rotations, since we know that each unital channel is a convex sum of rotations. Because every rotation preserves the Euclidean norm and the eigenvalues of a density operator with Bloch vector r are $(1 \pm |r|)/2$, we see that rotations preserve entropy. Now suppose f is any operator that preserves entropy. Since f preserves entropy, we have

$$S(x) = H((1 + |x|)/2) = H((1 + |f(x)|)/2) = S(f(x))$$

On the interval $[1/2, 1]$, entropy is injective, so $|x| = |f(x)|$. This implies that $f(0) = 0$ and so f is unital and hence linear. Then f is a linear isometry in the Euclidean norm, which means f is an orthogonal matrix, and hence a rotation by Prop. 3.9. □

Another connection between $\mathcal{B}$ and $\mathcal{U}$ is that each is precisely the collection of *entropy increasing* channels: the set of channels whose output state never has entropy strictly less than that of its input state. This follows easily from the result above using the convexity of entropy and the fact that there is a unique state of maximal entropy. Yet another connection is that $\mathcal{B}$ and $\mathcal{U}$ arise as the Scott continuous channels with a Scott closed set of fixed points in the Bayesian and spectral orders on Δ^2 and Ω^2 respectively [4].

With so many similarities between unital channels and binary symmetric channels, a fair question then becomes: what is the significance of a binary symmetric channel? Suppose we are communicating in a noisy environment, sending bits with equal frequency and suffering an error rate of p. Then some channel with probability of error p models the environment – but which one? We should assume the worst, and take the channel that has *minimal capacity*. As explained in [5], such a channel is always binary symmetric. Thus, binary symmetric channels – and entropy increasing channels in general – can be thought of as providing *conservative* models of noise.

4 The definition of scope

Each basis of the state space provides a different way to represent information. For each representation, there is an associated classical channel whose capacity measures the amount of information that can be transmitted through a quantum channel when that particular representation is used. The *scope* of a quantum channel is the range of classical capacities achieved as the sender and receiver vary over all possible representations. Let us illustrate the idea for qubits:

Example 4.1 *Scope.* Alice and Bob fix a basis $\{|\psi\rangle, |\phi\rangle\}$ of the state space:

- The state $|\psi\rangle = a|0\rangle + b|1\rangle$ represents '0'
- The state $|\phi\rangle = c|0\rangle + d|1\rangle$ represents '1'

This choice of basis defines a classical channel:

- Alice sends a qubit $|*\rangle$ representing '0' or '1' to Bob.
- As the qubit $|*\rangle$ travels, it interacts with the environment, changing to $\varepsilon(|*\rangle\langle *|)$
- Bob receives and measures the qubit in the $\{|\psi\rangle, |\phi\rangle\}$ basis, obtaining a '0' or '1'.

This classical channel in turn has a capacity given by

$$\mathrm{C}(x, y) = \log_2 \left(2^{\frac{\bar{x}H(y) - \bar{y}H(x)}{x-y}} + 2^{\frac{yH(x) - xH(y)}{x-y}} \right)$$

where $x = P(0|0)$, $y = P(0|1)$, $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the base two entropy and $\bar{x} = 1 - x$. For instance, if the environment is modelled by the bit flipping channel

$$\varepsilon = (1-p)I + p \cdot \varepsilon_x$$

of Example 2.13, then these probabilities are given by

$$x = P(0|0) = (1-p)|a|^4 + (2p+2)|a|^2|b|^2 + (1-p)|b|^4$$

$$y = P(0|1) = 1 - ((1-p)|c|^4 + (2p+2)|c|^2|d|^2 + (1-p)|d|^4)$$

The *scope* of ε is the range of capacities achieved as $\{|\psi\rangle, |\phi\rangle\}$ varies over *all* bases of the state space.

As is clear, calculating the scope of a quantum channel is not a trivial matter. However, by switching to the Bloch representation, not only can we see how to calculate scope for examples like bit flipping, we can develop a *systematic method* for calculating the scope of *any* unital qubit channel. To take our first step toward this, we need to understand how to calculate the classical channels associated to a qubit channel in its Bloch representation:

Lemma 4.2 *If $\rho, \sigma \in \Omega^2$ are mixed states with respective Bloch vectors r and s, then*

$$\mathrm{tr}(\rho \cdot \sigma) = \frac{1 + (r, s)}{2}$$

where (r, s) is the Euclidean inner product on $\mathbb{R}^3$.

Proof. With $r = (r_x, r_y, r_z)$ and $s = (s_x, s_y, s_z)$, we can write

$$\rho = \frac{I}{2} + \frac{r_x\sigma_x + r_y\sigma_y + r_z\sigma_z}{2} \quad \& \quad \sigma = \frac{I}{2} + \frac{s_x\sigma_x + s_y\sigma_y + s_z\sigma_z}{2}$$

where σ_x, σ_y and σ_z are the Pauli spin operators. Multiplying ρ and σ and taking the trace gives

$$\mathrm{tr}(\rho \cdot \sigma) = \frac{1}{2} + \frac{1}{4} \cdot \mathrm{tr}[(r_x\sigma_x + r_y\sigma_y + r_z\sigma_z)(s_x\sigma_x + s_y\sigma_y + s_z\sigma_z)]$$

using linearity of the trace, $\mathrm{tr}(I/4) = 1/2$ and equation 1. Setting $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, we abbreviate the expression on the right as $\mathrm{tr}(\,\langle r, \vec{\sigma}\rangle \cdot \langle s, \vec{\sigma}\rangle\,)$ and find that

$$\begin{aligned}
\mathrm{tr}(\,\langle r, \vec{\sigma}\rangle \cdot \langle s, \vec{\sigma}\rangle\,) &= \mathrm{tr}(r_x s_x \cdot I + r_x s_y \sigma_x\sigma_y + r_x s_z \sigma_x\sigma_z) \\
&+ \mathrm{tr}(r_y s_x \sigma_y\sigma_x + r_y s_y \cdot I + r_y s_z \sigma_y\sigma_z) \\
&+ \mathrm{tr}(r_z s_x \sigma_z\sigma_x + r_z s_y \sigma_z\sigma_y + r_z s_z \cdot I) \\
&= r_x s_x \cdot \mathrm{tr}(I) + 0 + 0 \\
&+ 0 + r_y s_y \cdot \mathrm{tr}(I) + 0 \\
&+ 0 + 0 + r_z s_z \cdot \mathrm{tr}(I) \qquad \text{(Using equation 2)} \\
&= 2 \cdot (r, s)
\end{aligned}$$

Substituting this into the original equation gives the desired result. □

Suppose Alice attempts to send Bob a qubit represented by ρ. As the qubit travels, it suffers the effect of noise described by the quantum channel ε. Bob then receives $\varepsilon(\rho)$ and performs a measurement in *some* basis $\{|0\rangle, |1\rangle\}$. The measurement operators in this case are the projections $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$ and form a complete set since $P_0 + P_1 = I$, so by standard quantum mechanics, the probability that Bob obtains the result 0 is

$$p_0 = \mathrm{tr}(P_0^\dagger P_0 \cdot \varepsilon(\rho))) = \mathrm{tr}(P_0 P_0 \cdot \varepsilon(\rho)) = \mathrm{tr}(P_0 \cdot \varepsilon(\rho))$$

while the probability that Bob obtains the result 1 is

$$p_1 = \mathrm{tr}(P_1^\dagger P_1 \cdot \varepsilon(\rho))) = \mathrm{tr}(P_1 P_1 \cdot \varepsilon(\rho)) = \mathrm{tr}(P_1 \cdot \varepsilon(\rho))$$

where we recall that projections satisfy $P_i^2 = P_i$. Now both projections P_0 and P_1, being density operators, also have a Bloch vector associated with them, given by s and t, respectively. If r is the Bloch vector for ρ and f is the Bloch representation of ε, then the probabilities p_0 and p_1 can be succinctly written as

$$p_0 = \frac{1 + (s, f(r))}{2} \quad \& \quad p_1 = \frac{1 + (t, f(r))}{2}$$

Further, since $|0\rangle$ and $|1\rangle$ form a basis for the state space, $s + t = 0$, which helps us see that $p_0 + p_1 = 1$. Finally, if Alice and Bob use the same basis $\{r, -r\}$ to represent information and attempt to communicate in the presence of unital noise f, then

$$P(0|0) = \frac{1 + (r, f(r))}{2} = \frac{1 + (-r, f(-r))}{2} = P(1|1).$$

These probabilities define a binary channel (a, b) with $a = P(0|0)$ and $b = P(0|1)$ whose capacity is given by

$$\mathrm{C}(a,b) = \log_2\left(2^{\frac{\bar{a}H(b)-\bar{b}H(a)}{a-b}} + 2^{\frac{bH(a)-aH(b)}{a-b}}\right)$$

where $\mathrm{C}(a,a) := 0$ and $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the base two entropy.

Definition 4.3 Let f be a unital channel. For each $r \in \partial\mathbb{B}^3$, the associated classical channel is

$$x_f(r) = \left(\frac{1+(r,f(r))}{2}, \frac{1-(r,f(r))}{2}\right)$$

where $\partial\mathbb{B}^3 = \{x \in \mathbb{B}^3 : |x| = 1\}$ is the set of pure states.

The question we want to answer is: what range of capacities is achieved as $\{r, -r\}$ varies over *all* bases of the state space?

Theorem 4.4 *If f is a unital channel, then the set of achievable capacities is given by*

$$\{\mathrm{C}(x_f(r)) : r \in \mathbb{B}^3, |r| = 1\} = \left[1 - H\left(\frac{1+m^-}{2}\right), 1 - H\left(\frac{1+m^+}{2}\right)\right]$$

where

$$m^+ = \sup_{|x|=1} |(x, f(x))| \quad \& \quad m^- = \inf_{|x|=1} |(x, f(x))|.$$

Proof. Let $\{r, -r\}$ be the Bloch vectors of a basis used to represent the classical bits '0' and '1' respectively. Then the associated classical channel is described by

$$a = P(0|0) = \frac{1+(r,f(r))}{2} \quad \& \quad b = P(0|1) = \frac{1+(r,f(-r))}{2} = \frac{1-(r,f(r))}{2}$$

By the *nonexpansivity* of f in Corollary 3.4, we have $a, b \in [0,1]$. The channel (a, b) is binary symmetric so its capacity is $1 - H(a)$. Let us begin by rewriting the expression for capacity. The function $\varphi : [-1,1] \to [0,1]$ defined by

$$\varphi(x) = 1 - H\left(\frac{1+x}{2}\right)$$

restricts to a bijection on either of the halves $[-1,0]$ or $[0,1]$: it is strictly decreasing on the former and strictly increasing on the latter. Now define $p_f : \partial\mathbb{B}^3 \to [-1,1]$ by $p_f(r) = (r, f(r))$. The capacity achieved when information is represented in the $\{r, -r\}$ basis is then given by $\varphi(p_f(r))$.

To find the largest and smallest values of the function $\varphi \circ p_f$, we proceed as follows. First, $|p_f| : \partial\mathbb{B}^3 \to [0,1]$ is continuous on the compact set $\partial\mathbb{B}^3$, so there are points u and v such that $|p_f(u)| = m^+$ and $|p_f(v)| = m^-$. Since φ is strictly increasing on $[0,1]$ and $|p_f|$ maps into $[0,1]$, the smallest value of $\varphi \circ |\cdot| \circ p_f$ is

$$\varphi(|p_f(v)|) = \inf_{|x|=1} \varphi(|p_f(x)|) = 1 - H\left(\frac{1+m^-}{2}\right)$$

while its largest value is

$$\varphi(|p_f(u)|) = \sup_{|x|=1} \varphi(|p_f(x)|) = 1 - H\left(\frac{1+m^+}{2}\right)$$

However, φ is symmetric about zero: $\varphi(x) = \varphi(-x)$ for all $x \in [-1,1]$. Thus, $\varphi(p_f(x)) = \varphi(|p_f(x)|)$ for all x. This proves that all achievable capacities lie in the indicated range, that $\{v, -v\}$ is a basis for achieving the smallest value of capacity and that $\{u, -u\}$ is a basis for achieving the largest capacity. Finally, because $\partial\mathbb{B}^3$ is a *connected* set, all capacities in between the maximum and minimum are also achievable. □

Notice that we were able to completely characterize the range of capacities achievable by a unital channel using *binary symmetric* channels. This is another reason they seem to provide a classical counterpart to unitality.

Definition 4.5 For a unital channel $f \in \mathcal{U}$, we define

$$f^+ = 1 - H\left(\frac{1+m^+}{2}\right) \quad \& \quad f^- = 1 - H\left(\frac{1+m^-}{2}\right)$$

where $m^+ = \sup_{|x|=1} |(x, f(x))|$ and $m^- = \inf_{|x|=1} |(x, f(x))|$. We define

$$s(f) = [f^-, f^+]$$

and call this the *scope* of f.

The scope $s(f)$ of a channel f is the range of its achievable capacities. It measures how much the capacity of f is capable of varying as different bases are used to represent classical bits, and thus how representative a particular value of capacity, such as f^+, is of a channel's behavior. Let us pause to consider an interesting class of channels whose scope *always* has maximum length.

Example 4.6 *Projective measurements.* Any *projective measurement*

$$\varepsilon(\rho) = P_0 \rho P_0 + P_1 \rho P_1$$

for projections $P_0 + P_1 = I$ defines a unital channel ε that is *idempotent*: $\varepsilon^2 = \varepsilon$. Thus, its Bloch representation f_ε satisfies

$$f_\varepsilon^2 = f_\varepsilon \circ f_\varepsilon = [\![\varepsilon]\!] \circ [\![\varepsilon]\!] = [\![\varepsilon \circ \varepsilon]\!] = [\![\varepsilon]\!] = f_\varepsilon$$

and so is also idempotent. Let us now calculate the scope of an idempotent unital channel.

Given a unital channel $f : \mathbb{B}^3 \to \mathbb{B}^3$ with $f^2 = f$, there are three possibilities: either

(a) $f = I$,

(b) $f = 0$ or

(c) $(\exists x, y \in \mathbb{B}^3)\ f(x) \neq 0\ \&\ f(y) \neq y$.

In the first two cases, we have $s(f) = [1,1]$ and $s(f) = [0,0]$ respectively. In the last, we set $s = y - f(y)/|y - f(y)|$ and $t = f(x)/|f(x)|$. Then $s, t \in \partial\mathbb{B}^3$ are both pure states. By the idempotence of f,

$$m^- \leq |(s, f(s))| = \frac{|(s, f(y) - f^2(y))|}{|y - f(y)|} = |(s, 0)| = 0$$

and

$$|(t, f(t))| = \frac{|(f(x), f^2(x))|}{|f(x)|^2} = \frac{|(f(x), f(x))|}{|f(x)|^2} = \frac{|(f(x), f(x))|}{|f(x)|^2} = 1 \leq m^+$$

so the scope of f is

$$s(f) = [0,1]$$

by Theorem 4.4.

Of course, we were only able to calculate the scope in the preceding example because the values m^+ and m^- could be deduced from the fact that the channel is question was idempotent. Is there a way to systematically calculate the scope of *any* unital channel?

5 The calculation of scope

A real $n \times n$ matrix A is *symmetric* when $A = A^t$. The eigenvalues of a symmetric matrix are real and there are exactly n of them: $(\lambda_1, \ldots, \lambda_n)$, though they are not all necessarily distinct. A standard fact about symmetric matrices is that

$$\sup_{|x|=1} (x, Ax) = \max_{1 \leq i \leq n} \lambda_i \quad \& \quad \inf_{|x|=1} (x, Ax) = \min_{1 \leq i \leq n} \lambda_i. \tag{4}$$

This is quite fortunate for us.

First, each $f \in \mathcal{U}$ can be written as $f(x) = Mx$ where M is a real 3×3 matrix. If the matrix M happens to be symmetric, then by Theorem 4.4 we can calculate the scope of f by simply finding the largest and smallest eigenvalues of M. This in turn requires solving the characteristic equation of M, which means finding the zeroes of a third degree polynomial with real coefficients, and even a formula exists for these[1].

Definition 5.1 A unital channel is called *symmetric* when $f = f^t$. The class of symmetric unital channels is denoted $\mathcal{S}$.

Proposition 5.2 *Let f be a symmetric unital channel with eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$. Then*

$$s(f) = \left[\frac{(1 + \operatorname{sgn}(\lambda_1 \lambda_3))}{2} \left(1 - H\left(\frac{1 + \min |\lambda_i|}{2} \right) \right), 1 - H\left(\frac{1 + \max |\lambda_i|}{2} \right) \right]$$

where $\operatorname{sgn}(x) = x/|x|$ *for* $x \neq 0$ *and* $\operatorname{sgn}(0) = 0$.

Proof. By Theorem 4.4, we know that

$$s(f) = \left[1 - H\left(\frac{1 + m^-}{2} \right), 1 - H\left(\frac{1 + m^+}{2} \right) \right]$$

[1]There is a formula for fourth degree polynomials too, but not five.

It is a standard result of linear algebra that $m^+ = \max|\lambda_i|$ when f is symmetric. Thus, to finish the proof we only need to compare left endpoints of the intervals. Let x and y be eigenvectors for λ_1, λ_3 respectively.

If $\operatorname{sgn}(\lambda_1\lambda_3) = 0$, then either $\lambda_1 = 0$ or $\lambda_3 = 0$, and we get

$$0 \leq m^- \leq \min\{|(x, f(x))|, |(y, f(y))|\} = \min\{|\lambda_1|, |\lambda_3|\} = 0$$

which means $m^- = 0 = \min|\lambda_i|$, so the formula in the statement of the theorem holds in this case. Then we can assume $\lambda_i \neq 0$ for all i.

If $\operatorname{sgn}(\lambda_1\lambda_3) = -1$, then λ_1 and λ_3 have opposite signs and the formula has a left endpoint of zero. By the continuity of $x \mapsto (x, f(x))$ and the connectedness of the unit sphere,

$$(x, f(x)) = \lambda_1 \;\;\&\;\; (y, f(y)) = \lambda_3 \implies (\exists z)\,(z, f(z)) = 0 \implies m^- = 0$$

so we see that the formula gives the correct value of $s(f)$.

If $\operatorname{sgn}(\lambda_1\lambda_3) = +1$, then λ_1 and λ_3 have the same signs, and λ_2 has the same sign since it is the middle eigenvalue. If the overall sign is positive, then the formula gives $s(f)$ by equation (4). If it is negative, then $g = -f$ is a symmetric matrix with positive eigenvalues $|\lambda_1|$, $|\lambda_2|$, $|\lambda_3|$ so

$$\inf_{|x|=1}(x, g(x)) = \min|\lambda_i|$$

by equation (4). We claim that this quantity is equal to m^-. To prove this, first note that

$$\min|\lambda_i| = \inf_{|x|=1}(x, g(x)) = \inf_{|x|=1} -(x, f(x)) = -\sup_{|x|=1}(x, f(x))$$

which implies that $(x, f(x)) < 0$ for all x. Then $(x, g(x)) = -(x, f(x)) = |(x, f(x))|$, which gives $m^- = \min|\lambda_i|$ as desired. $\square$

Example 5.3 *The bit flipping channels.* Consider the bit flipping channel

$$f_x = (1-p)I + ps_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-2p & 0 \\ 0 & 0 & 1-2p \end{pmatrix}$$

from Example 2.13. It is diagonal hence symmetric and its eigenvalues are simply the elements along the diagonal. Thus, by Prop. 5.2, its scope is

$$s(f_x) = \begin{cases} [1-H(p), 1] & \text{if } p \leq 1/2 \\ [0, 1] & \text{if } p \geq 1/2 \end{cases}$$

where we use the equality $H(1-p) = H(p)$. The answer is the same for phase flipping and bit-phase flipping. For the depolarization channel

$$d = (1-p)I = \begin{pmatrix} 1-p & 0 & 0 \\ 0 & 1-p & 0 \\ 0 & 0 & 1-p \end{pmatrix}$$

of Example 2.13, the scope is

$$s(d) = [1-H(p), 1-H(p)]$$

a single value of capacity, regardless of which basis is used to represent information.

Classically, if the environment flips bits with equal probability, then it is not possible to transmit any information: the resulting channel

$$\frac{1}{2}\cdot I + \frac{1}{2}\cdot \text{flip} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

has capacity zero. What happens in the quantum case? If one of the spin channels is applied with probability 1/2, is it still possible to transmit information?

Example 5.4 *Random qubit flipping*? Returning to the bit flipping channel f_x of the previous example with $p = 1/2$, we have

$$f_{1/2} = \frac{I + s_x}{2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

This channel is idempotent and is neither I nor 0, so from Example 4.6, we know it has scope $s(f_{1/2}) = [0, 1]$. Thus, there is a way to represent information so that *any* classical capacity can be achieved through the channel $f_{1/2}$.

In particular, if bits are coded in the $\{|0\rangle, |1\rangle\}$ basis, which has Bloch vectors $\pm e_3 = (0, 0, \pm 1)$, then $f_{1/2}$ has the same effect that it has classically: no information can be transmitted. However, if we use the basis $\{|+\rangle, |-\rangle\}$, which has Bloch vectors $\pm e_1 = (\pm 1, 0, 0)$, then we can transmit information *perfectly*.

The problem of "randomly flipping a qubit," so that no information can be transmitted in any basis, can be solved by applying the channel $f_{1/2}$ *followed by* the channel $(I + s_y)/2$. Intuitively, it is not enough to 'flip bits' – one must also 'flip phases'. This is explained in more detail in [5].

Example 5.5 *The intercept-resend attack in quantum cryptography.* In this attack, an eavesdropper randomly chooses between the $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ bases and then uses this choice to perform a projective measurement on a qubit being sent from Alice to Bob.

As shown in [5], $f = (I + s_x)/2$ is the Bloch representation of a measurement in the $\{|+\rangle, |-\rangle\}$ basis, while $g = (I + s_y)/2$ is the Bloch representation of a measurement in the $\{|0\rangle, |1\rangle\}$ basis. Thus, the action of the eavesdropper causes noise of the form

$$\frac{1}{2}f + \frac{1}{2}g = \frac{1}{2}I + \frac{1}{4}s_x + \frac{1}{4}s_y = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and since the resulting channel is symmetric, it has scope

$$[0, 1 - H(3/4)] = [0, 1 - H(1/4)]$$

In particular, while the act of eavesdropping causes noise which reduces the amount of information that can be transmitted between sender and receiver, it *does not* reduce it to zero.

The next channel has the interesting property that entangled states can be used to reduce the probability of error when transmitting classical information through it:

Example 5.6 *The two-Pauli channel.* The two-Pauli channel [1] is

$$f_x = x \cdot I + \left(\frac{1-x}{2}\right) s_x + \left(\frac{1-x}{2}\right) s_y$$

where $x \in [0,1]$, so we have

$$f_x = \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & 2x-1 \end{pmatrix}$$

By Prop 5.2,

$$s(f_x) = \begin{cases} [0, 1 - H(x/2)] & \text{if } 0 \le x \le 1/3 \\ [0, 1 - H((1+x)/2)] & \text{if } 1/3 \le x \le 1/2 \\ [1 - H(x), 1 - H((1+x)/2)] & \text{if } 1/2 < x \le 1 \end{cases}$$

doing a case-by-case analysis.

For symmetric channels, Prop. 5.2 gives us a systematic way to calculate scope. But not all unital channels are symmetric. How do we calculate the scope of an arbitrary unital channel? A natural idea is to try and show that every unital channel f can be represented by a symmetric channel $\varphi(f)$ that has the same scope as f. We would like this representation to be natural as follows:

- If f is symmetric, then there is no need to represent it differently, so we would like $\varphi(f) = f$.
- When a channel has been decomposed into a convex sum of simpler parts, we would like it to be easy to calculate φ, so we would like

  $$\varphi(pf + (1-p)g) = p\,\varphi(f) + (1-p)\varphi(g).$$

  That is, φ should preserve convex sums.
- Because $(x, f(x)) = (x, f^t(x))$, f and f^t define the same classical channel in any given basis, so we should have

  $$\varphi(f^t) = \varphi(f)$$

  and by the symmetry of $\varphi(f)$, we then have $\varphi(f^t) = \varphi(f) = \varphi(f)^t$. That is, φ should preserve the transpose operation.

Not only is the problem of obtaining such a φ solvable – it has a *unique* solution.

Theorem 5.7 *Let X be a nonempty subset of a real vector space that is closed under convex sums, and let $* : X \to X$ be a convex linear involution with $\mathrm{fix}(*) := \{x \in X : x^* = x\}$. Then there is a convex linear retraction $\varphi : X \to \mathrm{fix}(*)$ which preserves $*$. Furthermore, there is only one function which has these properties, and it is given by*

$$\varphi(x) = \frac{1}{2}\,x + \frac{1}{2}\,x^*$$

for $x \in X$.

Proof. It is clear that φ as defined in the statement of the theorem has all the properties stated; what we need to do is establish its uniqueness. To this end, suppose φ is a convex linear retraction of X onto $\mathrm{fix}(*)$ that preserves $*$. Let $x \in X$. First notice that by the convex linearity of $*$,

$$\left(\frac{1}{2}\,x + \frac{1}{2}\,x^*\right)^* = \frac{1}{2}\,x^* + \frac{1}{2}\,(x^*)^* = \frac{1}{2}\,x^* + \frac{1}{2}\,x = \frac{1}{2}\,x + \frac{1}{2}\,x^*$$

is a fixed point of $*$, and therefore a fixed point of φ, since φ is a retraction onto $\mathrm{fix}(*)$. Then

$$\begin{aligned}
\varphi(x) &= \varphi\left(\frac{1}{2}\,x + \frac{1}{2}\,x\right) \\
&= \frac{1}{2}\,\varphi(x) + \frac{1}{2}\,\varphi(x) && \text{(Convex Linearity)} \\
&= \frac{1}{2}\,\varphi(x) + \frac{1}{2}\,\varphi(x)^* && (\varphi(x) \in \mathrm{fix}(*)) \\
&= \frac{1}{2}\,\varphi(x) + \frac{1}{2}\,\varphi(x^*) && (\varphi \text{ preserves } *) \\
&= \varphi\left(\frac{1}{2}\,x + \frac{1}{2}\,x^*\right) && \text{(Convex Linearity)} \\
&= \frac{1}{2}\,x + \frac{1}{2}\,x^* && \text{(Fixed point of } \varphi)
\end{aligned}$$

which finishes the proof. □

The above result applies for instance to a subset of real matrices closed under convex sum and the transpose operation. In particular, we have

Corollary 5.8 *There is a unique function from $\mathcal{U}$ to $\mathcal{S}$ that preserves convex sum, transpose and retracts $\mathcal{U}$ onto $\mathcal{S}$. It is given by*

$$\varphi(f) = \frac{1}{2}\,f + \frac{1}{2}\,f^t$$

for $f \in \mathcal{U}$.

Proof. For $f \in \mathcal{U}$, $f^t \in \mathcal{U}$ by Prop. 3.5. We now apply Theorem 5.7 with X the subset of unital channels and $*$ the transpose operation. □

From an information theoretic viewpoint, f and $\varphi(f)$ are identical:

Theorem 5.9 *Let $f \in \mathcal{U}$ be an arbitrary unital channel. Then*

- *For each $x \in \mathbb{B}^3$, $(x, f(x)) = (x, \varphi(f)x)$. Thus, in any basis, f and $\varphi(f)$ define the same classical channel.*
- *The channels f and $\varphi(f)$ have the same scope.*

The scope of a unital channel f is found by calculating the largest and smallest eigenvalues of $\varphi(f)$.

Proof. The entire result is obvious once we prove the first equality:

$$\begin{aligned}(x,\varphi(f)x) &= \frac{1}{2}(x,f(x))+\frac{1}{2}(x,f^t(x)) \\ &= \frac{1}{2}(x,f(x))+\frac{1}{2}(f(x),x) && \text{(Property of the transpose)} \\ &= \frac{1}{2}(x,f(x))+\frac{1}{2}(x,f(x)) && \text{(Real inner product)} \\ &= (x,f(x))\end{aligned}$$

□

The function $\varphi(f)$ extracts all the classical information about f as a channel into a form where we can then systematically obtain it as a routine eigenvalue calculation. An example will make the value of this technique clear.

Example 5.10 Consider the basic rotation $r_x(\theta)$, which is only symmetric if it is an involution.

$$\varphi(r_x(\theta)) = \frac{1}{2}\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & 0\\ 0 & 0 & \cos\theta\end{pmatrix}$$

so the largest eigenvalue of $\varphi(f)$ is 1 and the smallest is $\cos\theta$, meaning that the scope of $r_x(\theta)$ can be any interval of the form $[a,1]$ where $a \in [0,1]$ depends on θ.

It is remarkable that the scope of any unital channel f can be found by calculating the scope of a symmetric channel $\varphi(f)$. To illustrate why, let us point out that while any convex sum of rotations gives rise to a unital channel, only a convex sum of *involutive* rotations can give rise to a symmetric unital channel. The proof of this requires delving deeper into the structure of unital channels themselves: we start by using the trace lemma (Lemma 3.7) to give a new and elementary proof of the well-known "complete positivity" conditions [2]:

Lemma 5.11 *A diagonal matrix*

$$\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{pmatrix}$$

is unital if and only if $|\lambda_i| \leq 1$ *for each* $i \in \{1,2,3\}$ *and if the following four inequalities are satisfied:*

(i) $1+\lambda_1+\lambda_2+\lambda_3 \geq 0$

(ii) $1+\lambda_1-\lambda_2-\lambda_3 \geq 0$

(iii) $1-\lambda_1+\lambda_2-\lambda_3 \geq 0$

(iv) $1-\lambda_1-\lambda_2+\lambda_3 \geq 0$

Proof. ($\Rightarrow$): By induction, each unital channel f is nonexpansive ($|f(x)| \leq |x|$), so $|\lambda_i| \leq 1$. By the trace lemma, the channel

$$f = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

must have

$$\mathrm{tr}(f) = \lambda_1 + \lambda_2 + \lambda_3 \geq -1$$

which gives inequality (i). The last three inequalities follow from the fact that $s_x \cdot f$, $s_y \cdot f$, $s_z \cdot f$ are unital and also have a trace that exceeds -1.

($\Leftarrow$) Denoting the nonnegative expressions in inequalities (i)-(iv) by x_i, we then have

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \sum_{i=1}^{4} \frac{x_i}{4} \cdot r_i$$

where (r_1, r_2, r_3, r_4) are the spin channels (I, s_x, s_y, s_z). $\square$

Proposition 5.12

(a) *A channel $f \in \mathcal{U}$ is an involution iff there is a rotation r and an $s \in \{I, s_x, s_y, s_z\}$ such that $f = r \cdot s \cdot r^{-1} = r \cdot s \cdot r^t$. In particular, all unital involutions are symmetric.*

(b) *A unital channel $f \in \mathcal{U}$ is symmetric iff it is a convex sum of involutive rotations.*

Proof. (a) Since $f \in \mathcal{U}$ is an involution, $f^{-1} = f \in \mathcal{U}$, so by Prop. 3.9, f is a rotation. But as a rotation, $f^{-1} = f^t$, while as an involution, $f = f^{-1}$, thus $f = f^t$ is symmetric. Since f is symmetric, there is an orthogonal matrix r

$$rfr^t = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

Because r is an odd dimensional matrix, we can assume r is a rotation, by replacing r with $-r$ if necessary. Each λ_i must be either 1 or -1 since they are all real and each is an eigenvalue of a rotation. However, their product must be one, since it is the determinant of f. This proves the desired result.

(b) ($\Rightarrow$) If f is symmetric, then as in the proof of (a), we can find a rotation r such that

$$rfr^t = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

Because rfr^t is unital, the diagonal map on its right is unital and so must be a convex sum of spin channels as seen in the proof of Lemma 5.11. Conjugating both sides of this equation by r now shows that f is a convex sum of four involutions, the four involutions forming a copy of the Klein four group.

($\Leftarrow$) If f is a convex sum of involutions

$$f = \sum_{i=1} x_i \cdot r_i$$

then by part (a), each r_i is symmetric and we see that

$$f^t = \sum_{i=1}^{n} x_i \cdot r_i^t = \sum_{i=1}^{n} x_i \cdot r_i = f$$

f is also symmetric. $\square$

Since every symmetric channel is a convex sum of four involutions which comprise a copy of the Klein four group, the same is true of $\varphi(f) = (f + f^t)/2$ for *any* unital channel f.

6 Scope and the Holevo capacity

A standard way of measuring the capacity of a quantum channel in quantum information theory is the Holevo capacity; it is sometimes called the product state capacity since input states are not allowed to be entangled across two or more uses of the channel.

Definition 6.1 For a trace preserving quantum operation f, the *Holevo capacity* is given by

$$\mathrm{C}(f) = \sup_{\{x_i,\rho_i\}} \left[S\left(f\left(\sum_i x_i \rho_i \right) \right) - \sum_i x_i \cdot S(f(\rho_i)) \right]$$

where the supremum is taken over all ensembles $\{x_i, \rho_i\}$ of possible input states ρ_i to the channel.

The possible input states ρ_i to the channel are in general mixed and the x_i are probabilities with $\sum_i x_i = 1$.

Theorem 6.2 *If f is a unital channel with scope $s(f) = [f^-, f^+]$, then $f^+ \leq \mathrm{C}(f)$. If f is symmetric, then $f^+ = \mathrm{C}(f)$.*

Proof. Remembering that f is map on the Bloch sphere $\mathbb{B}^3$, and using the correspondence between density operators and the Bloch sphere, the Holevo capacity of f is given by

$$\mathrm{C}(f) = \sup_{\{x_i,r_i\}} \left[H\left(\frac{1 + |f\left(\sum_i x_i r_i\right)|}{2} \right) - \sum_i x_i \cdot H\left(\frac{1 + |f(r_i)|}{2} \right) \right]$$

where r_i are Bloch vectors for density operators in an ensemble, and we recall that eigenvalues of a density operator with Bloch vector r are $(1 \pm |r|)/2$. By the continuity of $|f|$, there is a pure state $r \in \mathbb{B}^3$ for which

$$|f(r)| = \sup_{|x|=1} |f(x)|$$

In order to keep this proof self-contained, we first repeat the proof from [4] that

$$\mathrm{C}(f) = 1 - H((1 + |f(r)|)/2)$$

Setting $r_1 = r$, $r_2 = -r$ and $x_1 = x_2 = 1/2$ defines an ensemble for which the expression maximized in the definition of $\mathrm{C}(f)$ reduces to $1 - H((1 + |f(r)|)/2)$. Notice that in this step we explicitly make use of the fact that f is unital: $f(0) = 0$. This proves $1 - H((1 + |f(r)|)/2) \leq \mathrm{C}(f)$.

For the reverse inequality, any term in the supremum is clearly bounded from above by

$$1 - \sum_i x_i \cdot H\left(\frac{1 + |f(r_i)|}{2}\right)$$

since $H(x) \leq 1$. Because

$$|f(r_i)| \leq \sup_{x \in \mathbb{B}^3} |f(x)| = \sup_{|x|=1} |f(x)| = |f(r)|$$

we have

$$H\left(\frac{1 + |f(r_i)|}{2}\right) \geq H\left(\frac{1 + |f(r)|}{2}\right)$$

which then gives $\mathrm{C}(f) \leq 1 - H((1 + |f(r)|)/2)$ and thus that these two expressions are equal.

To prove $f^+ \leq \mathrm{C}(f)$, note that for $|x| = 1$, we have

$$|(x, f(x))| = |x| \cdot |f(x)| \cdot |\cos\theta| \leq |f(x)|$$

so that

$$m^+ = \sup_{|x|=1} |(x, f(x))| \leq \sup_{|x|=1} |f(x)| = |f(r)|$$

and thus $f^+ = 1 - H((1 + m^+)/2) \leq \mathrm{C}(f)$. If f is symmetric, then $m^+ = |f(r)|$, so $f^+ = \mathrm{C}(f)$. □

As the last proof makes clear, $f^+ = \mathrm{C}(f)$ iff $m^+ = \sup |f(x)|$, so for instance, we also have equality for rotations r, since r on $\mathbb{R}^3$ has at least one real eigenvalue λ with $|\lambda| = 1$. It is worth pointing out that there are unital channels f for which $f^+ < \mathrm{C}(f)$. Any nonzero skew-symmetric unital channel has positive capacity but scope $[0, 0]$, as we now show:

Proposition 6.3

(i) *A unital channel f has scope $s(f) = [0, 0]$ if and only if $f^t = -f$.*

(ii) *A unital channel f has scope $s(f) = [1, 1]$ if and only if $f = I$.*

Proof. (i) First suppose that f is symmetric. If it achieves capacity zero in all bases, then its largest and smallest eigenvalues must be zero. Thus, all its eigenvalues are zero, which means that it is the zero matrix.

Now consider an arbitrary f that achieves zero capacity in every basis. Since $\varphi(f) = (f + f^t)/2$ and f have the same scope, $\varphi(f)$ is symmetric and achieves capacity zero in all bases, so $\varphi(f) = 0$, which finishes the proof.

(ii) Suppose f is symmetric. Then all of its eigenvalues have magnitude one. If two of its eigenvalues had opposite signs, f would achieve capacity zero in some basis by continuity. Thus, its eigenvalues are either all 1 or all -1. In the first case, $f = I$, while in the second $f = -I$, the antipodal map, which is impossible by Corollary 3.8.

For an arbitrary f with scope $[1,1]$, we know $\varphi(f)$ also has scope $[1,1]$, so by the previous remark $\varphi(f) = I$. For $|x| = 1$,

$$(x, f(x)) = (x, \varphi(f)(x)) = (x, x) = 1$$

so $1 = |(x, f(x))| \leq |f(x)| \leq 1$ gives $|f(x)| = 1$. Since f preserves the norm for unit vectors and is a linear mapping, it does so for all vectors in $\mathbb{R}^3$ and must be a linear unital isometry and hence a rotation. As in the proof of Lemma 3.7, take an orthogonal matrix r such that

$$rfr^t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & -b & a \end{pmatrix}$$

with $a^2 + b^2 = 1$. The trace of f is then

$$\operatorname{tr}(f) = \operatorname{tr}(rfr^t) = 1 + 2a$$

On the other hand, since $\varphi(f) = I$,

$$\operatorname{tr}(f) = \frac{\operatorname{tr}(f) + \operatorname{tr}(f^t)}{2} = \operatorname{tr}(\varphi(f)) = \operatorname{tr}(I) = 3.$$

Equating both terms, $1 + 2a = 3$ and so $a = 1$, $b = 0$ and $f = I$. $\square$

Of course, we have not yet seen that $\mathcal{U}$ contains any *nonzero* skew symmetric matrices, but they do actually exist:

Proposition 6.4 *Let E_{ij} denote the 3×3 matrix with a one in the (i,j) position and a zero in every other location.*

(i) *The matrices E_{ij} and $-E_{ij}$ are unital for each $i, j \in \{1,2,3\}$.*

(ii) *The skew-symmetric matrix*

$$\begin{pmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{pmatrix}$$

is unital provided $a, b, c \geq 0$ and $a + b + c \leq 1/2$.

Proof. (i) The three matrices

$$f = r_x(\pi/2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad g = r_y(\pi/2) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad h = r_z(\pi/2) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

as well as their transposes are all unital, since each is a rotation. Each E_{ii} is unital since

$$E_{11} = (f + f^t)/2, \quad E_{22} = (g + g^t)/2, \quad E_{33} = (h + h^t)/2$$

and each $-E_{ii}$ is unital since

$$-E_{11} = E_{11} \cdot s_y, \quad -E_{22} = E_{22} \cdot s_x, \quad -E_{33} = E_{33} \cdot s_y$$

where s_x and s_y are the spin channels. Noting that $M \cdot E_{ii}$ gives us the matrix with the same i^{th} column as M and zeroes elsewhere, we see that the matrices $\pm E_{1j}$ are unital since

$$E_{12} = h^t \cdot E_{22}, \quad -E_{12} = h \cdot E_{22}, \quad E_{13} = g \cdot E_{33}, \quad -E_{13} = g^t \cdot E_{33},$$

the $\pm E_{2j}$ are unital since

$$E_{21} = h \cdot E_{11}, \quad -E_{21} = h^t \cdot E_{11}, \quad E_{23} = f^t \cdot E_{33}, \quad -E_{23} = f \cdot E_{33},$$

and the $\pm E_{3j}$ are unital since

$$E_{31} = g^t \cdot E_{11}, \quad -E_{31} = g \cdot E_{11}, \quad E_{32} = f \cdot E_{22}, \quad -E_{32} = f^t \cdot E_{22}.$$

(ii) We can write this skew-symmetric matrix as

$$\begin{pmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{pmatrix} = a \cdot E_{21} + a(-E_{12}) + b \cdot E_{31} + b \cdot (-E_{13}) + c \cdot E_{32} + c \cdot (-E_{23})$$

which must be unital by Corollary 3.6. □

Thus, any nonzero skew-symmetric unital channel f has $f^+ = 0 < \mathrm{C}(f)$ since $\mathrm{C}(f) = 0$ only when $f \equiv 0$. We also have the following interesting corollary:

Corollary 6.5 *If $f \in \mathcal{U}$ and $p \in [-1/3, 1]$, then $pf \in \mathcal{U}$. Further, this is the largest range over which scalar multiplication is possible: for any $p \in \mathbb{R}$, if $pf \in \mathcal{U}$ for all $f \in \mathcal{U}$, then $p \in [-1/3, 1]$.*

Proof. Let $f \in \mathcal{U}$. First note that

$$-\frac{1}{3}f = \frac{1}{3}f(-E_{11}) + \frac{1}{3}f(-E_{22}) + \frac{1}{3}f(-E_{33})$$

is unital, as a convex sum of unital channels. Given $p \in [-1/3, 0]$, there is $q \in [0, 1]$ with $p = -q/3$ so $pf = q(-f/3) \in \mathcal{U}$ by Prop. 3.5. To see that this is the largest range, we must have

$$3p = \mathrm{tr}(p \cdot I) \in [-1, 3]$$

by Lemma 3.7, so $p \in [-1/3, 1]$, finishing the proof. □

7 Adaptive quantum communication

In quantum cryptography, the number of bits we can transmit requires a key of the same size, so the speed at which we transmit information depends on how fast we can generate keys. If the error rate within a session of QKD is too high, we have to start over: this slows the key generation rate. Key generation rates are important not only because it is desirable to communicate as fast as possible, but also because there are times when it is *the only* way for communication to be possible: for instance, in freespace, we have at best 5-6 minutes to transmit information to a satellite before it is out of reach. By *minimizing the error rate*, we can avoid restarting and speed up the rate at which information is transmitted. We now indicate how the theory of scope can be used to develop a method for minimizing the error rate in quantum cryptography – we call it *adaptive quantum cryptography.*

(i) For each $i \in \{1,2,3\}$, Alice sends many 0's prepared in the e_i basis to Bob

(ii) For each $i \in \{1,2,3\}$, Bob measures one third of them in the e_1 basis, one third in the e_2 basis and one third in the e_3 basis

(iii) Bob calculates the channel f which governs the noise in the environment. First, he uses the measurement results from (ii) to estimate the probability p_{ij} that e_i is received when e_j is sent. But since the element f_{ij} of f located at position (i,j) is related to p_{ij} via

$$p_{ij} = \frac{1 + (e_i, f(e_j))}{2} = \frac{1 + f_{ij}}{2}$$

these probabilities allow Bob to construct f!

(iv) Bob calculates the scope $s(f)$ and an eigenvector r associated to f^+,

(v) Alice and Bob agree to engage in QKD using r and a forty five degree rotation of r – this requires use of a private key so that Bob can transmit the information about r to Alice, but once done, all future communication will be at a faster rate assuming a stable environment.

Example 7.1 Consider a typical case like bit flipping with σ_y

$$f = \begin{pmatrix} 1-2p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1-2p \end{pmatrix}$$

In QKD, random bits are randomly coded in the bases e_1 and e_3, so the error rate is

$$\frac{1}{2}\left(\frac{1 + (\mp e_1, f(\pm e_1))}{2}\right) + \frac{1}{2}\left(\frac{1 + (\mp e_3, f(\pm e_3))}{2}\right) = \frac{1}{2}\cdot p + \frac{1}{2}\cdot p = p,$$

while with adaptive communication, the bases used are e_2 and without loss of generality e_1, so the error rate is

$$\frac{1}{2}\left(\frac{1 + (\mp e_1, f(\pm e_1))}{2}\right) + \frac{1}{2}\left(\frac{1 + (\mp e_2, f(\pm e_2))}{2}\right) = \frac{1}{2}\cdot p + \frac{1}{2}\cdot 0 = \frac{p}{2}.$$

Of course, depolarization $f(r) = p \cdot r$ for $p \in [0,1]$ is an example of an effect where the error rate cannot be improved upon, but in many cases, it will be. For such a scheme to be physically realized, numerous questions must be answered:

(a) We can test for unitality using the complete positivity conditions (Lemma 5.11), but what do we do if the noise in the environment is not unital? How necessary is a theory of scope for non-unital channels?

(b) To calculate scope, we need to be able to solve a cubic equation – there are well known formulas for doing so, is this the best way to solve it, or are there new techniques that can be applied in this particular case?

(c) How many 'test bits' does Alice have to send Bob in step (i) in order to make sure that Bob obtains a channel that accurately models the environment?

(d) Can we *quantify* the degree of improvement in the error rate for a given quantum channel when using adaptive quantum communication?

Other issues perhaps also of interest include: when the eavesdropper knows we are adaptively communicating, they will adapt too, now what? Or: if we are trying to perform QKD from the ground to a low earth orbit satellite via entanglement, it is possible that relativistic effects could adversely affect our ability to communicate – or that relativistic effects could be used to provide a new way to help prevent eavesdropping

8 Closing

In general, a qubit channel $f : \mathbb{B}^3 \to \mathbb{B}^3$ has the form $f(x) = Mx + b$ where M is a 3×3 real matrix and $b \in \mathbb{B}^3$. Such a channel is unital iff $b = 0$, so there are many qubit channels that are not unital, amplitude damping being a notable example. Developing a method for calculating the scope of a nonunital qubit channel is a difficult but high priority. As a first step toward this, the algebraic structure of nonunital qubit channels has to be uncovered.

Part of the value of the definition of scope is that when a certain capacity is achieved, we can point to a definite procedure the sender and receiver should follow in order to achieve that capacity: prepare in basis r, send and then measure in basis r. It is possible though to challenge the definition of scope. For instance, what if the sender prepares in basis r, sends and then the receiver measures in basis s? With a definition of scope that allows for *two* different bases, the least value of capacity ceases to have meaning: for a given unital f, the receiver can choose a pure state s that is Euclidean orthogonal to $f(r)/|f(r)|$ ensuring that $(1 + (s, f(r)))/2 = 1/2$, so the smallest capacity achievable is *always* zero. For symmetric channels, it can be shown that the upper bound in such a definition can be achieved with a *single* basis and is equal to f^+. For nonsymmetric channels, like non-zero skew-symmetric channels, one can achieve a higher capacity than f^+.

Aside from adaptive quantum communication, whose precise operational details will be the subject of future research, there are other uses of scope that may be possible. One is the role of scope in classifying physical effects according to the degree that they disturb the state of a system. For instance, a 'weak effect' would be a channel with scope close to $[1, 1]$. One reason for this is given in Prop. 6.3: the only channel with scope $[1, 1]$ is the identity. Other examples are *projective measurements*: they always have scope $[0, 1]$, which is a maximum distance from $[1, 1]$, indicating that the disturbance caused by such an effect is extreme. By contrast, the Holevo capacity does not distinguish between channels in as precise a manner as scope: for instance, it assigns the value 1 to any rotation, whereas the scope of a rotation always depends on the angles involved.

9 Acknowledgements

We thank Tanner Crowder for helping us find the complete positivity conditions in the literature, and for many valuable discussions while this paper was being written; Johnny Feng, who first asked if the antipodal map was a qubit channel; and Marco Lanzagorta, who upon hearing that it was not a qubit channel would not stop talking about it until we both realized something really cool: that 'universal' bit flipping is physically impossible.

References

[1] C. H. Bennett, C. A. Fuchs, and J. A. Smolin, *Entanglement-enhanced classical communication on a noisy quantum channel*, in Quantum Communication, Computing and Measurement, O. Hirota, A. S. Holevo, and C. M. Caves, Eds. New York: Plenum, 1997, pp. 7988.

[2] P.S. Bourdon and H.T. Williams, *Unital quantum operations on the Bloch ball and Bloch region*, Physical Review A, Vol. 69, Article 022314, 2004.

[3] L. J. Landau and R. F. Streater, *On Birkhoff's theorem for doubly stochastic completely positive maps of matrix algebras*, Linear Algebra and its Applications **193**, p. 107–127, 1993.

[4] K. Martin. *A domain theoretic model of qubit channels.* ICALP 2008, Lecture Notes in Computer Science, Vol. 5126, p. 283–297, 2008.

[5] K. Martin. *How to randomly flip a quantum bit.* Electronic Notes in Theoretical Computer Science, Volume 270, Issue 1, p. 81–97, Elsevier Science, 2011.

[6] M. Nielsen and I. Chuang, *Quantum computation and quantum information.* Cambridge University Press, 2000.

[7] M. B. Ruskai, S. Szarek and E. Werner. *An analysis of completely-positive trace-preserving maps on* $\mathcal{M}_2$. Linear Algebra and its Applications, Volume 347, Issues 1-3, p. 159–187, 2002.

[8] D. Serre. *Matrices: Theory and applications.* Springer-Verlag, Graduate Texts in Mathematics, 2000.

[9] C. E. Shannon. *A mathematical theory of communication.* Bell Systems Technical Journal 27, 379–423 and 623–656, 1948.