Notes on continuity of channel capacities

Egor Larionov (20263767)

University of Waterloo

Friday 14th September, 2012

Contents

Abstract

A collection of research notes accumulated in the summer of 2012. The goal of the research was to prove that the quantum channel assisted by two-way or backward classical communication is continuous.

1 Definitions and Preliminaries

We will use $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ to denote Hilbert spaces of the form $\mathbb{C}^\Sigma$ where Σ is a finite, non-empty index set. We denote the set of linear operators of the form $A : \mathcal{X} \to \mathcal{Y}$ by $\mathrm{L}(\mathcal{X}, \mathcal{Y})$, and $\mathrm{L}(\mathcal{X})$ for short, when $\mathcal{Y}$ is the same as $\mathcal{X}$. We adopt Dirac notation throughout this paper. Now consider the following definitions [1].

Definition 1.1. If $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, then

$$\begin{aligned} \ker(A) &\equiv \{u \in \mathcal{X} : A|u\rangle = 0\} && \text{is the kernel of A,} \\ \mathrm{im}(A) &\equiv \{A|u\rangle : u \in \mathcal{X}\} && \text{is the image of A, and} \\ \mathrm{supp}(A) &\equiv \overline{\{u \in \mathcal{X} : A|u\rangle \neq 0\}} && \text{is the support of A} \end{aligned}$$

Definition 1.2. Let $A \in \mathrm{L}(\mathcal{X})$, and $\{e_a\}_{a\in\Sigma}$ be an orthonormal basis for $\mathcal{X}$, then

$$\begin{aligned} \mathrm{Tr}(A) &\equiv \sum_{a\in\Sigma} \langle e_a|A|e_a\rangle \\ \mathrm{Det}(A) &\equiv \sum_{\sigma\in\mathrm{Sym}(\Sigma)} \mathrm{sgn}(\sigma) \prod_{a\in\Sigma} \langle e_a|A|e_{\sigma(a)}\rangle \end{aligned}$$

Note that $\mathrm{Sym}(\Sigma)$ is the permutation group of Σ.

Remark 1.3. Additionally we can express the above quantities in terms of their spectrum as

$$\mathrm{Tr}(A) = \sum_{\lambda\in\mathrm{spec}(A)} \lambda \quad \text{and} \quad \mathrm{Det}(A) = \prod_{\lambda\in\mathrm{spec}(A)} \lambda$$

Definition 1.4. The inner product on the space $\mathrm{L}(\mathcal{X}, \mathcal{Y})$ is defined as

$$\langle A, B\rangle \equiv \mathrm{Tr}(A^\dagger B)$$

Definition 1.5. For $A, B \in \mathrm{L}(\mathcal{X})$, we define the *Lie bracket* (or *commutator*) $[A, B] \in \mathrm{L}(\mathcal{X})$ by

$$[A, B] \equiv AB - BA$$

1.1 Composite Hilbert Spaces

Given a vector space $\mathcal{X}$, the *Kronecker product* of two vectors $u, v \in \mathcal{X}$ is denoted by $u \otimes v$.

The *tensor product* of two Hilbert spaces say, $\mathcal{X}$ and $\mathcal{Y}$ is defined as

$$\mathcal{X} \otimes \mathcal{Y} \equiv \mathrm{span}\{x \otimes y : x \in \mathcal{X}, y \in \mathcal{Y}\}$$

Definition 1.6. Given two Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$, with bases $\{e_i^\mathcal{X}\}$ and $\{e_j^\mathcal{Y}\}$ repsectively. Then for any $A \in \mathrm{L}(\mathcal{X}\otimes\mathcal{Y})$, we define the *partial trace over* $\mathcal{X}$ to be

$$\mathrm{Tr}_\mathcal{X}(A) \equiv \sum_i \left(\langle e_i^\mathcal{X}| \otimes \mathbb{1}_\mathcal{Y}\right) A \left(|e_i^\mathcal{X}\rangle \otimes \mathbb{1}_\mathcal{Y}\right)$$

And analagously, the partial trace over $\mathcal{Y}$ is defined by

$$\mathrm{Tr}_\mathcal{Y}(A) \equiv \sum_j \left(\mathbb{1}_\mathcal{X} \otimes \langle e_j^\mathcal{Y}|\right) A \left(\mathbb{1}_\mathcal{X} \otimes |e_j^\mathcal{Y}\rangle\right)$$

1.2 Norms on $\mathrm{L}(\mathcal{X}, \mathcal{Y})$

We will first introduce the Shatten p-norm, which gives a more specific norm called the trace norm which is widely used in quantum information.

Definition 1.7. Given $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ and a real number $p \geq 0$ the Shatten p-norm is defined to be

$$\|A\|_p \equiv \left[\mathrm{Tr}\left((A^\dagger A)^{p/2}\right)\right]^{1/p}$$

Definition 1.8. Given $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, we define the *trace norm* of A to be

$$\|A\|_1 \equiv \mathrm{Tr}(\sqrt{A^\dagger A}).$$

Where the subscript 1 indicates that this is the Shatten 1-norm.

1.3 Types of Linear Operators

Definition 1.9. Let $A \in \mathrm{L}(\mathcal{X})$, then it is

$$\textit{normal} \text{ if } \quad AA^\dagger = A^\dagger A,$$
$$\textit{Hermitian} \text{ if } \quad A = A^\dagger.$$

Remark 1.10. $\mathrm{Herm}(\mathcal{X})$ is a real inner product space of Hermitian linear operators.

Definition 1.11. If $A \in \mathrm{L}(\mathcal{X})$, then it is called *positive semidefinite* if

- $A \in \mathrm{Pos}(\mathcal{X})$.
- $A = B^\dagger B$ for some $B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$.
- $\langle u|A|u\rangle \geq 0$ for all $u \in \mathcal{X}$.
- $\langle B, A\rangle \geq 0$ for all $B \in \mathrm{Pos}(\mathcal{X})$.
- A is Hermitian and every eigenvalue of A is non-negative.
- There exists $\{u_a : a \in \Sigma\} \subset \mathcal{Y}$, s.t. $A(a, b) = \langle u_a|u_b\rangle$.

Definition 1.12. If $A \in \mathrm{Pos}(\mathcal{X})$, then it is called *positive definite* if

- A is invertible.
- $\mathrm{Det}(A) \neq 0$.
- $\langle u|A|u\rangle > 0$ for all $u \in \mathcal{X}$.
- A is Hermitian, and every eigenvalue of A is positive.
- A is Hermitian, and there exists $\varepsilon > 0$ s.t. $A \geq \varepsilon \mathbb{1}$

Remark 1.13. For convenience, if $A \in \mathrm{L}(\mathcal{X})$ is positive semidefinite, we write $A \geq 0$, and similarly if A is positive definite, we write $A > 0$

Definition 1.14. If $A \in \mathrm{Pos}(\mathcal{X})$, then it is called a *density operator* if

- $A \in \mathrm{D}(\mathcal{X})$
- $\mathrm{Tr}(A) = 1$

Definition 1.15. If $P \in \mathrm{L}(\mathcal{X})$, then it is called a *projection* if $P^2 = P$. A projection is *orthogonal* if

- it can be written as $P = \sum_{i=1}^k |i\rangle\langle i|$, where $\{|i\rangle\}$ is an orthonormal basis of some subspace of $\mathcal{X}$.
- P is Hermitian.

Definition 1.16. If $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, then it is called a *linear isometry* if

- $\|A|u\rangle\| = \|u\|$ for all $u \in \mathcal{X}$.
- $A^\dagger A = \mathbb{1}_{\mathcal{X}}$.
- $\langle u|A^\dagger A|v\rangle = \langle u|v\rangle$ for all $u, v \in \mathcal{X}$.

Note that a linear isometry in $\mathrm{L}(\mathcal{X})$ is called a *unitary*.

Remark 1.17. If $A \in \mathrm{L}(\mathcal{X})$ is unitary or Hermitian, then it is also normal.

1.4 Useful linear algebra results

Lemma 1.18 (Hölder's inequality)**.** *If $u, v \in \mathbb{C}^n$, and $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$, (where $\frac{1}{\infty} = 0$), then*

$$|\langle u, v\rangle| \leq \|u\|_p \|v\|_q.$$

Theorem 1.19 (First Isomorphism)**.** *If $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, then*

$$\mathrm{im}(A) \cong \mathcal{X} / \ker(A)$$

Remark 1.20. By Theorem 1.19 we can conclude that

$$\dim(\ker(A)) + \mathrm{rank}(A) = \dim(\mathcal{X}).$$

1.5 Analysis on Superoperators

We define a *superoperator* to be a linear operator of the form $\mathcal{N} : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$. We denote the space of such superoperators by

$$\mathrm{T}(\mathcal{X}, \mathcal{Y}) \equiv \mathrm{L}[\mathrm{L}(\mathcal{X}), \mathrm{L}(\mathcal{Y})].$$

We will now introduce a few different types of superoperators.

Definition 1.21. If $\mathcal{N} \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, then it is called *trace-preserving* (t.p.) if for all $A \in \mathrm{L}(\mathcal{X})$, we have

$$\mathrm{Tr}(\mathcal{N}(A)) = \mathrm{Tr}(A)$$

Definition 1.22. If $\mathcal{N} \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, then it is called *hermitian-preserving* (h.p.) if for all $A \in \mathrm{Herm}(\mathcal{X})$, we have that $\mathcal{N}(A) \in \mathrm{Herm}(\mathcal{Y})$.

Definition 1.23. If $\mathcal{N} \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, then it is called *completely positive* (c.p.) if for any Hilbert space $\mathcal{Z}$, and for all $A \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Z})$, we have that

$$\mathcal{N} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}(A) \geq 0$$

It is easy to see that c.p. maps are also h.p. so all of the c.p. maps live in the ambient space of h.p. maps.

We denote the set of trace-preserving and completely positive superoperators (TCP maps) by $\mathrm{TCP}(\mathcal{X}, \mathcal{Y})$, or $\mathrm{TCP}(\mathcal{X})$ when $\mathcal{Y}$ is the same as $\mathcal{X}$.

Since quantum channels are exactly represented by TCP maps, we will choose a specific norm that will not amplify the difference between maps when tensored with the identity.

Definition 1.24. Given $\mathcal{N} \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, we define the *diamond norm*[1] to be

$$\|\mathcal{N}\|_\diamond \equiv \max\{\|\mathcal{N} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}(A)\|_1 : A \in \mathrm{L}(\mathcal{X}^{\otimes 2}), \|A\|_1 = 1\}.$$

Note that the maximum above can be saturated by a pure quantum state. With the given norm, we immediately gain a metric

$$d(\mathcal{N}, \mathcal{M}) \equiv \|\mathcal{N} - \mathcal{M}\|_\diamond$$

and thus, a metric topology on $\mathrm{T}(\mathcal{X}, \mathcal{Y})$.

Definition 1.25. Given $\mathcal{N} \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, we define an *open ball* with radius $\delta > 0$ around $\mathcal{N}$ to be

$$\mathcal{B}_\delta(\mathcal{N}) \equiv \{\mathcal{M} \in \mathrm{T}(\mathcal{X}, \mathcal{Y}) : d(\mathcal{N}, \mathcal{M}) < \delta\}$$

The set $\{\mathcal{B}_\delta(\mathcal{N}) : \mathcal{N} \in \mathrm{T}(\mathcal{X}, \mathcal{Y}), \delta > 0\}$ of open balls, forms a basis for the metric topology on $\mathrm{T}(\mathcal{X}, \mathcal{Y})$, since every open set in $\mathrm{T}(\mathcal{X}, \mathcal{Y})$, can be decomposed into unions and finite intersections of open balls.

We now turn to the notion of continuity of real valued functions on the space of superoperators.

Definition 1.26. Let $E \subset \mathrm{T}(\mathcal{X}, \mathcal{Y})$, and $f : E \to \mathbb{R}$. We say that f is *continuous at the point* $\mathcal{N} \in E$ if given $\epsilon > 0$, there exists a $\delta > 0$ s.t. $\mathcal{M} \in E \cap \mathcal{B}_\delta(\mathcal{N})$ implies $|f(\mathcal{M}) - f(\mathcal{N})| < \epsilon$. Finally, f is *continuous* if it is continuous at each point of E.

1.6 Entropy and Capacity Theorems

Definition 1.27. The classical (or Shannon) entropy of a probability distribution $\{p_x\}$, is defined as

$$H(\{p_x\}) \equiv -\sum_x p_x \log p_x$$

We will refer to "a quantum system", say X, as some isolated physical system, which is completely described by its quantum state, $\rho \in \mathrm{D}(\mathcal{X})$. Naturally two or more systems, say X and Y, can form a composite system and it's described by a quantum state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$. The individual systems X and Y are in the states ρ^X and ρ^Y respectively, where

$$\rho^X = \mathrm{Tr}_{\mathcal{Y}}(\rho) \qquad \text{and} \qquad \rho^Y = \mathrm{Tr}_{\mathcal{X}}(\rho).$$

Definition 1.28. The quantum (or von Neumann) entropy of a quantum system X in a state $\rho \in \mathrm{D}(\mathcal{X})$ is defined as

$$S(X)_\rho \equiv -\mathrm{Tr}(\rho \log \rho)$$

or equivalently

$$S(X)_\rho \equiv -\sum_i \lambda_i \log \lambda_i$$

where λ_i are eigenvalues of ρ.

Theorem 1.29 (Joint entropy theorem)**.** *Suppose the systems X and Y are prepared in the state $\rho = \sum_i p_i |i\rangle\langle i| \otimes \phi_i$ where $\{|i\rangle\}_i$ is an orthonormal basis for $\mathcal{X}$. Then*

$$S\left(\sum_i p_i |i\rangle\langle i| \otimes \phi_i\right) = H(\{p_i\}) + \sum_i p_i S(\phi_i)$$

Proof. See p. 514 in [2]. □

Definition 1.30. The quantum mutual information of two systems A and B is defined to be

$$I(A; B) \equiv S(A) + S(B) - S(A \otimes B)$$

Definition 1.31. The Holevo χ-quantity of an ensemble of states $\mathcal{E} = \{(p_x, \phi_x)\}$ is defined to be

$$\chi(\mathcal{E}) \equiv S\left(\sum_x p_x \phi_x\right) - \sum_x p_x S(\phi_x)$$

Theorem 1.32 (The Holevo bound)**.** *Given an ensemble of states $\mathcal{E} = \{(p_x, \phi_x)\}$ in a system X, then for any POVM measurement $\{E_y\}_y$, on X leaving the measurement outcome in the system Y, we have*

$$I(X; Y) \le \chi(\mathcal{E})$$

or equivalently

$$I_{acc}(\mathcal{E}) \le \chi(\mathcal{E})$$

where $I_{acc}(\mathcal{E}) \equiv \max\{I(X; Y) : \{E_y\}\}$ is the accessible information

Lemma 1.33. *Given two systems X and Y in the state $\rho = \sum_i p_i |i\rangle\langle i| \otimes \phi_i$, we have that*

$$I(X; Y)_\rho = S\left(\sum_x p_x \phi_x\right) - \sum_x p_x S(\phi_x)$$

Proof.

$$I(X; Y)_\rho = S(X)_{\rho^X} + S(Y)_{\rho^Y} - S(XY)_\rho \tag{1.1}$$

Note that:

$$\begin{aligned}
\rho^X &= \mathrm{Tr}_Y\left(\sum_i p_i |i\rangle\langle i| \otimes \phi_i\right) \\
&= \sum_i p_i \mathrm{Tr}_Y(|i\rangle\langle i| \otimes \phi_i) && \text{by linearity of partial trace} \\
&= \sum_i p_i |i\rangle\langle i| && \text{since } \mathrm{Tr}(\phi_i) = 1 \\
\rho^Y &= \sum_i p_i \phi_i && \text{similarily}
\end{aligned}$$

So equation (1.1) becomes

$$\begin{aligned}
I(X; Y)_\rho &= H(\{p_i\}) + S\left(\sum_i p_i \phi_i\right) - S\left(\sum_i p_i |i\rangle\langle i| \otimes \phi_i\right) \\
&= H(\{p_i\}) + S\left(\sum_i p_i \phi_i\right) - H(\{p_i\}) - \sum_i p_i S(\phi_i) \\
&= S\left(\sum_i p_i \phi_i\right) - \sum_i p_i S(\phi_i),
\end{aligned}$$

where the second equality comes from Theorem 1.29. □

[1] The diamond norm is sometimes called the *completely bounded trace norm*

Definition 1.34. The *Holevo capacity*[2] of a channel $\mathcal{N}$ with input system X and output system Y is defined to be

$$C_H(\mathcal{N}) \equiv \max_{p_x,\phi_x} I(X;Y)_\rho$$

where X and Y are prepared in the state $\rho = \sum_x p_x |x\rangle\langle x| \otimes \mathcal{N}(\phi_x)$

It immediately follows from Lemma 1.33 that the Holevo information can also be written as

$$C_H(\mathcal{N}) \equiv \max_{p_x,\phi_x} \left[S\left(\sum_x p_x \mathcal{N}(\phi_x)\right) - \sum_x p_x S\left(\mathcal{N}(\phi_x)\right)\right]$$

From this point on, we will denote $x^{\otimes n}$ by x^n, where x can be a Hilbert space, a TCP map, or a quantum state. Also we let $\mathbb{R}^+ \equiv \{r \in \mathbb{R} : r > 0\}$, and $\mathbb{N}$ be the strictly positive integers. Recall that if $A \subset \mathcal{X}$ where $\mathcal{X}$ is some Hilbert space, then $\mathrm{Conv}(A)$ denotes the convex hull of A.

Definition 1.35 (Classical Capacity)**.** Given $\mathcal{N} \in \mathrm{TCP}(\mathcal{X}, \mathcal{Y})$, a rate $R \in \mathbb{R}^+$ of data transmission through the channel $\mathcal{N}$ is said to be *classically-achievable*, if for all $\epsilon > 0$, there exists $n \in \mathbb{N}$ with a classical code $\{\phi_k \in \mathcal{X}^n\}_{k=1}^{K_n}$ and a t.p. and c.p. decoding operation

$$\mathcal{D}_n \; : \; \mathrm{D}(\mathcal{Y}^n) \; \rightarrow \; \mathrm{Conv}\left(\{|k\rangle\langle k|\}_{k=1}^{K_n}\right)$$

such that $\forall k \in \{1, \dots, K_n\}$, we have

$$\|\mathcal{D}_n(\mathcal{N}^n(\phi_k)) - |k\rangle\langle k|\|_1 < \epsilon, \qquad \text{and} \tag{1.2}$$

$$\log K_n \geq nR. \tag{1.3}$$

The *classical capacity* of $\mathcal{N}$, $C(\mathcal{N})$, is defined to be the supremum over classically-achievable rates.

Theorem 1.36. *The classical capacity satisfies*

$$C(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} C_H(\mathcal{N}^n)$$

Note that the Holevo capacity measures the optimal rate of transmission assuming the input state is a product state. In contrast the classical information doesn't make such an assumption and indeed we have that $C_H(\mathcal{N}) \leq C(\mathcal{N})$ for all channels $\mathcal{N}$.

Definition 1.37 (Coherent Information)**.** Given two systems A and B, in some state ρ, the *coherent information* between A and B is given by

$$I^{\mathrm{coh}}(A\rangle B)_\rho \equiv S(B)_{\rho^B} - S(AB)_\rho$$

Definition 1.38 (Quantum Capacity)**.** Given $\mathcal{N} \in \mathrm{TCP}(\mathcal{X}, \mathcal{Y})$, a rate $R \in \mathbb{R}^+$ of data transmission through the channel $\mathcal{N}$ is said to be *achievable*, if for all $\epsilon > 0$, there exists $n \in \mathbb{N}$ with a subspace $C_n \subset \mathcal{X}^n$ and a decoding operation $\mathcal{D}_n \in \mathrm{TCP}(\mathcal{Y}^n, C_n)$ such that for any $\rho \in \mathrm{D}(C_n)$, we have

$$\|\mathcal{D}_n(\mathcal{N}^n(\rho)) - \rho\|_1 < \epsilon, \qquad \text{and} \tag{1.4}$$

$$\log \dim C_n \geq nR. \tag{1.5}$$

The *quantum capacity* of $\mathcal{N}$, $Q(\mathcal{N})$, is defined to be the supremum over achievable rates.

[2] Also known as the one-shot classical capacity.

Theorem 1.39. *The quantum capacity satisfies*

$$Q(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} I^{\mathrm{coh}}(\mathcal{N}^n)$$

where

$$I^{\mathrm{coh}}(\mathcal{N}) \equiv \max_{\rho \in \mathrm{D}(\mathcal{Z}\otimes\mathcal{X})} I^{\mathrm{coh}}(Z\rangle Y)_{\mathcal{I}\otimes\mathcal{N}(\rho)}$$

We will now describe a general protocol for quantum communication assisted by two-way classical communication and finally define the quantum capacity with two-way classical assistance.

Definition 1.40. Given an $n \in \mathbb{N}$, define an *n-use protocol*, $\mathcal{P}_n$, as a composition of operations performed on an input quantum state by the two communicating parties, and by the channel $\mathcal{N}$ at most n times.

A protocol such as the one described above is shown in Figure 1.1. Here we can see that $\mathcal{A}_i$ is the operation performed by Alice on the code system C_n and her auxiliary system A to prepare for forward classical communication denoted by $\mathcal{M}_{\to 0}$, and quantum communication denoted by $\mathcal{N}$. On the other side, the operation $\mathcal{B}_i$, performed by Bob, receives the sent information from Alice, prepares classical data to send back to Alice via $\mathcal{M}_{\leftarrow i+1}$, and retains the rest of the input to be used later. Both parties are allowed to maintain auxiliary systems A for Alice and B for Bob to aid with communication, for example by accumulating classical data sent throughout the length of the protocol. This process is repeated n times.

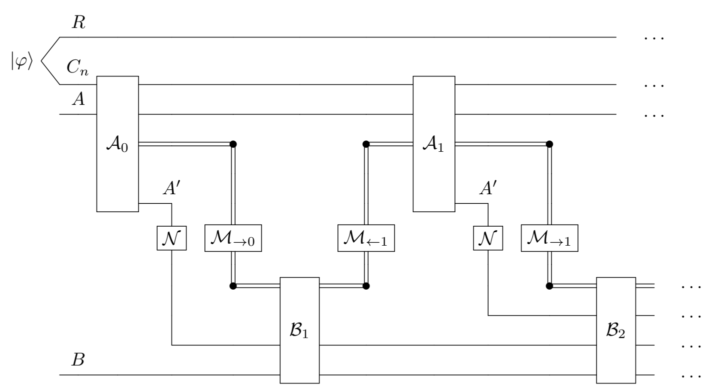

Figure 1.1: Here we show a general protocol implementing quantum communication assisted by two-way classical communication.

Definition 1.41 (Quantum Capacity with two-way classical assistance)**.** Given $\mathcal{N} \in \mathrm{TCP}(\mathcal{X}, \mathcal{Y})$, a rate $R \in \mathbb{R}^+$ of data transmission using the channel $\mathcal{N}$ along with two-way classical communication assistance, is said to be *two-way-achievable*, if for all $\epsilon > 0$, there exists $n \in \mathbb{N}$ with a subspace $C_n \subset \mathcal{X}^n$ and an n-use protocol $\mathcal{P}_n$, such that for any $\rho \in \mathrm{D}(C_n)$, we have that

$$\|\mathcal{P}_n(\rho) - \rho\|_1 < \epsilon, \qquad \text{and} \tag{1.6}$$

$$\log \dim C_n \geq nR. \tag{1.7}$$

The *quantum capacity of $\mathcal{N}$ with two-way classical assistance*, $Q_2(\mathcal{N})$, is defined to be the supremum over two-way-achievable rates.

Remark 1.42. Note that in each of the above definitions (1.35, 1.38 and 1.41), it is possible that no achievable rates exist. In this case the capacity is said to be zero.

2 Peres-Horodecki criterion

Definition 2.1. Two quantum systems are *separable* if they don't share entanglement. More formally, a mixed state ρ of the system $A \otimes B$ is *separable* if there exist probabilities $\{p_k \geq 0\}$, and mixed states $\{\psi_k\}$, and $\{\phi_k\}$, such that

$$\rho = \sum_k p_k \psi_k \otimes \phi_k$$

where $\sum_k p_k = 1$.

Definition 2.2. Given a general state ρ of the composite system $A \otimes B$ as

$$\rho = \sum_{i,j,k,l} p_{i,j,k,l} |i\rangle\langle j| \otimes |k\rangle\langle l|,$$

we define the *partial transpose* of ρ to be

$$\rho^{\Gamma_B} = \sum_{i,j,k,l} p_{i,j,k,l} |i\rangle\langle j| \otimes |l\rangle\langle k|.$$

Theorem 2.3 (Peres-Horodecki criterion)**.** *If the mixed state ρ of the system $A \otimes B$ has* positive partial transpose (PPT), *then ρ is separable. More precisely, if $\rho^{\Gamma_B} \geq 0$, then ρ is separable.*

Proof. This result is thoroughly examined in [7]. □

3 Understanding TCP maps

Note that any map $\mathcal{N} \in \mathrm{TCP}(\mathcal{X}, \mathcal{Y})$, can be extended to a map $\hat{\mathcal{N}} \in \mathrm{TCP}(\mathcal{Z})$, without changing its capacities. So in the following sections we will use TCP to denote $\mathrm{TCP}(\mathcal{X})$ and T to denote $\mathrm{T}(\mathcal{X})$, for some arbitrary Hilbert space $\mathcal{X}$.

Lemma 3.1. *The set of non-t.p. maps is dense in the ambient space of linear superoperators.*

Proof. Take an $\epsilon > 0$. Let $\mathcal{N}$ be t.p. and $\mathcal{M}$, non-t.p. Define $c = \|\mathcal{M} - \mathcal{N}\|_\diamond$, where $0 < c < \infty$ by the definition of the diamond norm. Take $0 < \delta < \epsilon/c$, and construct

$$\Phi_\delta = \delta\mathcal{M} + (1-\delta)\mathcal{N}.$$

Then we have

$$\begin{aligned} \|\Phi_\delta - \mathcal{N}\|_\diamond &= \|\delta\mathcal{M} + (1-\delta)\mathcal{N} - \mathcal{N}\|_\diamond \\ &= \|\delta(\mathcal{M} - \mathcal{N})\|_\diamond \\ &= c\delta \\ &< \epsilon. \end{aligned}$$

So Φ_δ can be made arbitrarily close to $\mathcal{N}$. Now we will show that Φ_δ is not t.p. Take a density matrix $\rho \in \mathrm{D}(\mathcal{X})$, such that $\mathcal{M}(\rho) \neq 1$. We know that such a density matrix exists since $\mathcal{M}$ is not t.p. So we get

$$\begin{aligned} \mathrm{Tr}(\Phi_\delta(\rho)) &= \mathrm{Tr}\left([\delta\mathcal{M} + (1-\delta)\mathcal{N}](\rho)\right) \\ &= \delta \underbrace{\mathrm{Tr}(\mathcal{M}(\rho))}_{\neq 1} + (1-\delta)\underbrace{\mathrm{Tr}(\mathcal{N}(\rho))}_{=1} \neq 1. \end{aligned}$$

Thus we have that the set of non-t.p. maps is dense. □

It follows from the Lemma 3.1, that the interior of TCP maps is empty when taken with respect to the metric topology of T. This motivates us to find a more restricted ambient space of linear superoperators that will give a non-empty interior for the TCP maps.

3.1 Representations of superoperators

We will need a special mapping that gives an operator-vector correspondence, which is described in great detail in [1], though in this paper we allow non-standard bases. For the following definitions, let $\{x_i\}_{i=1}^n$ be a basis for $\mathcal{X}$ and $\{y_k\}_{k=1}^m$ a basis for $\mathcal{Y}$, where $n = \dim \mathcal{X}$ and $m = \dim \mathcal{Y}$.

Definition 3.2. Note that $\{y_k x_i^\dagger\}$ forms a basis for $\mathrm{L}(\mathcal{X}, \mathcal{Y})$. Now define the *operator-vector correspondence* with a linear mapping

$$\mathrm{vec} : \mathrm{L}(\mathcal{X}, \mathcal{Y}) \to \mathcal{Y} \otimes \mathcal{X},$$

given by

$$\mathrm{vec}(y_k x_i^\dagger) = y_k \otimes x_i.$$

Note that this is a bijection since $\{y_k \otimes x_i\}$ is a basis for $\mathcal{Y} \otimes \mathcal{X}$.

Now we will examine two important representations of superoperators (Sections 5.2.1 and 5.2.2 in [1]). We will use slightly more general versions of these representations by allowing an arbitrary basis. That is we will take two arbitrary orthonormal bases for $\mathrm{L}(\mathcal{X})$ and $\mathrm{L}(\mathcal{Y})$, say

$$\mathcal{A} = \{A_a\}_{a=1}^{n^2} \qquad \text{and} \qquad \mathcal{B} = \{B_b\}_{b=1}^{m^2}$$

respectively, and explore these representations with respect to our chosen bases.

Definition 3.3. Take $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, then the *natural representation* of Φ in the given bases is

$$K(\Phi) = \sum_{a,b} \langle B_b, \Phi(A_a)\rangle\, \mathrm{vec}(B_b)\mathrm{vec}(A_a)^\dagger$$

If we take the bases to be normalized with respect to the Frobenius norm, one can easily verify that the natural representation satisfies the equation

$$K(\Phi)\mathrm{vec}(A_a) = \mathrm{vec}(\Phi(A_a)).$$

Hence $K : \mathrm{T}(\mathcal{X}, \mathcal{Y}) \to \mathrm{L}(\mathcal{X}^{\otimes 2}, \mathcal{Y}^{\otimes 2})$ is a linear bijection.

Another useful representation of a linear superoperator is the Choi-Jamiłkowski representation, given by the mapping $J : \mathrm{T}(\mathcal{X}, \mathcal{Y}) \to \mathrm{L}(\mathcal{Y} \otimes \mathcal{X})$ as follows.

Definition 3.4. Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, then the *Choi-Jamiłkowski representation* of Φ is defined by

$$J(\Phi) = \sum_a \Phi(A_a) \otimes A_a$$

It is easy to check that J is a linear bijection.

3.2 Characterizing superoperators

The goal of this section is to analyse the possible supersets of TCP maps. We have already introduced trace-preserving, hermitian-preserving, and completely positive maps. Now consider the generalized Gell-Mann operators (GGMs) along with the identity as a basis on the space $\mathrm{L}(\mathcal{X})$. Let $d = \dim \mathcal{X}$, and we get three types of GGMs:

1. $\frac{d(d-1)}{2}$ symmetric
$$\Lambda_s^{jk} = |j\rangle\langle k| + |k\rangle\langle j|, \text{ for } 1 \leq j < k \leq d;$$

2. $\frac{d(d-1)}{2}$ antisymmetric
$$\Lambda_a^{jk} = -i|j\rangle\langle k| + i|k\rangle\langle j|, \text{ for } 1 \leq j < k \leq d;$$

3. $(d-1)$ diagonal
$$\Lambda^l = \sqrt{\frac{2}{l(l+1)}}\left(\sum_{j=1}^{l} |j\rangle\langle j| - l|l+1\rangle\langle l+1|\right),$$
for $1 \leq l \leq d-1$.

Hence we have a total of d^2-1 Hermitian, traceless, and pairwise orthogonal GGMs. Adding the d-dimentional identity completes the basis. Now if we normalize the above basis and place the identity as the first basis element, we can examine the natural representation with respect to this basis.

Proposition 3.5. *For illustrative purposes, assume that $\mathcal{X} = \mathbb{C}^n$ and $\mathcal{Y} = \mathbb{C}^m$ for some input and output dimensions n and m respectively. Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, then the following statements hold:*

1. *$K(\Phi) \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$.*
2. *Φ is h.p. iff $K(\Phi) \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$.*
3. *Φ is t.p. iff* $K(\Phi) = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline \vdots & \vdots & \vdots & \vdots \end{array}\right] \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$.
4. *Φ is t.p and h.p. iff*
$$K(\Phi) = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline \vdots & \vdots & \vdots & \vdots \end{array}\right] \in \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m).$$

Remark 3.6. A superoperator $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is c.p. if it is h.p. with a few additional constraints. Hence TCP maps live in the space of maps that are both t.p. and h.p.

Sketch proof of Proposition 3.5. We will describe the reasoning behind each point individually:

1. Clear from the bijectivity of K.
2. It is easy to see that any Hermitian operator can be expressed as a real linear combination of the chosen generalized Gell-Mann operators along with the identity. Thus any Hermitian operator is simply a real vector in the given basis. So the hermitian preserving superoperators precisely correspond to the linear operators that take real vectors to real vectors in our chosen basis.
3. Since Φ is t.p. iff $\Phi^\dagger$ is unital, we have that
$$\Phi^\dagger(\mathbb{1}_{\mathrm{L}(\mathcal{X})}) = \mathbb{1}_{\mathrm{L}(\mathcal{Y})}.$$
Now since the identity is our first basis element, we know that
$$K(\Phi^\dagger) = \left[\begin{array}{c|c} 1 & \cdots \\ \hline 0 & \cdots \\ \vdots & \cdots \\ 0 & \cdots \end{array}\right].$$
Notice that $K(\Psi^\dagger) = K(\Psi)^\dagger$ for any $\Psi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, so we are done.
4. This follows from parts 2 and 3.

□

Proposition 3.5 helps us visualize the space of superoperators and its subclasses. It trivializes Lemma 3.1, and clearly shows that the correct ambient space to use when defining the interior of the TCP maps, is the space of t.p. and h.p. maps, which we will denote by THP.

4 Continuity of Q_2

For convenience let us define the set of channels with positive two-way assisted quantum capacity:
$$\mathcal{Q}_2^+ \equiv \{\Phi \in \mathrm{TCP} : Q_2(\Phi) > 0\}.$$

There is a proof [8] of the continuity of the two-way assisted quantum capacity, Q_2, on the interior[3] of $\mathcal{Q}_2^+$, which we denote by $\mathrm{int}(\mathcal{Q}_2^+)$. Our goal is to extend this continuity result to maps that lie outside this interior, but we will exclude the zero-capacity channels from our domain to simplify the problem. Hence we are left to prove the following weak continuity result.

Theorem 4.1 (Continuity of Q_2)**.** *Given $\epsilon > 0$ and $\Phi \in \mathcal{Q}_2^+$, there exists $\delta > 0$ such that*
$$\Psi \in \mathcal{Q}_2^+ \cap \mathcal{B}_\delta(\Phi) \implies |Q_2(\Psi) - Q_2(\Phi)| < \epsilon.$$

In [8], it is shown that Q_2 is continuous on $\mathrm{int}(\mathcal{Q}_2^+)$ so it is enough to prove the following lemma in order to conclude Theorem 4.1.

Lemma 4.2. *Given $\epsilon > 0$, there exists a continuous mapping $f_\epsilon : \mathcal{Q}_2^+ \to \mathrm{int}(\mathcal{Q}_2^+)$, such that for every $\Phi \in \mathcal{Q}_2^+$, we have*
$$|Q_2(f_\epsilon(\Phi)) - Q_2(\Phi)| < \epsilon.$$

Proof of Theorem 4.1 assuming Lemma 4.2. Take $\Phi \in \mathcal{Q}_2^+$, and $\epsilon > 0$, and by Lemma 4.2 we take a mapping, f_ϵ. If $\Phi \in \mathrm{int}(\mathcal{Q}_2^+)$ then we use the proof in [8] and we are done. Otherwise, $\Phi \in \mathcal{Q}_2^+ \backslash \mathrm{int}(\mathcal{Q}_2^+)$, so for any $\delta > 0$, there exists a $\delta' > 0$ such that
$$\Psi \in \mathcal{Q}_2^+ \cap \mathcal{B}_{\delta'}(\Phi) \implies f_\epsilon(\Psi) \in \mathcal{Q}_2^+ \cap \mathcal{B}_\delta(f_\epsilon(\Phi)),$$

[3] The interior of this set is taken with respect to the metric topology of THP maps. See Proposition 3.5 for more details.

by the continuity of f_ϵ. We use the continuity of Q_2 on $\text{int}(\mathcal{Q}_2^+)$ to get a suitable $\delta = \delta_0$ for the given ϵ. Now notice that

$$\begin{aligned}|Q_2(\Phi) - Q_2(\Psi)| &\le |Q_2(\Phi) - Q_2(f_\epsilon(\Phi))| \\ &\quad + |Q_2(f_\epsilon(\Phi)) - Q_2(f_\epsilon(\Psi))| \\ &\quad + |Q_2(f_\epsilon(\Psi)) - Q_2(\Psi)| \\ &< \epsilon + \epsilon + \epsilon.\end{aligned}$$

Here the first and the last ϵ come from Lemma 4.2, and the middle one comes from the continuity of Q_2 on $\text{int}(\mathcal{Q}_2^+)$. This completes the proof of Theorem 4.1. □

Proof of Lemma 4.2. Take an $\epsilon > 0$. Let $d = \dim \mathcal{X}$, and consider the mapping $f_\epsilon : \mathcal{Q}_2^+ \to \text{int}(\mathcal{Q}_2^+)$ defined as

$$f_\epsilon(\Phi) = p_\epsilon \tilde{\Phi} + (1 - p_\epsilon)\Phi, \tag{4.1}$$

where $\tilde{\Phi}$ is the twirled channel Φ, as given by:

$$\begin{aligned}\tilde{\Phi}(\rho) &= \int_{U(d)} \mathrm{d}\mu(U)\, \mathcal{U} \circ \Phi \circ \mathcal{U}^\dagger(\rho) &(4.2)\\ &= \sum_i \Pr(\mathcal{U}_i)\, \mathcal{U}_i \circ \Phi \circ \mathcal{U}_i^\dagger(\rho), &(4.3)\end{aligned}$$

for some finite collection of unitary operations[4] $\{\mathcal{U}_i\}$. The last equality follows from the simulation of the Haar-measure μ by a unitary 2-design [9]... (unfinished)

□

5 Extending the quantum capacity

We will first examine if it makes sense to extend the von Neuman entropy to Hermitian matrices. The end goal is to be able to extend the definition of the quantum capacity and the two-way assisted quantum capacity to THP maps. This will provide us with a possibility to extend the proof of continuity of Q_2 in [8], to a bigger domain.

Recall that the entropy function is given by

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i,$$

for all $\rho \in \mathrm{D}(\mathcal{X})$. Now if we allow ρ to be an arbitrary Hermitian matrix with unit trace, than it may have negative eigenvalues, and so the entropy may take on complex values, if we consider the complex logarithm. So let $\mathrm{HU}(\mathcal{X})$ denote the set of Hermitian operators with unit trace, and consider the von Neumann entropy function with the complex logarithm in the principal branch. That is $S : \mathrm{HU}(\mathcal{X}) \to \mathbb{C}$. It is easy to check that the function $\eta(x) = x \log x$ is continuous on the real domain, hence so is the entropy function on the domain $\mathrm{HU}(\mathcal{X})$.

Now we will check that a particular non-TCP mapping has a quantum capacity that is close to its TCP neighbour. In particular consider the completely random channel $\mathcal{R}$ and the identity channel $\mathcal{I}$. Take $0 < \varepsilon \ll 1$, and define $p = \frac{\varepsilon}{1+\varepsilon}$. So we have that $0 < p \ll 1$. We can express the identity channel as a mixture of a TCP map and a non-TCP map:

$$\mathcal{I} = ((1+\varepsilon)\mathcal{I} - \varepsilon\mathcal{R})(1-p) + \mathcal{R}p.$$

[4] We use $\mathcal{U}(\rho) = U\rho U^\dagger$ to denote unitary operations.

Naturally we expect that mixing in the completely random channel cannot raise the quantum capacity of the channel, so we expect that the extended quantum capacity of

$$\Psi \equiv (1+\varepsilon)\mathcal{I} - \varepsilon\mathcal{R}$$

should not exceed $Q(\mathcal{I})$, but it should not differ from $Q(\mathcal{I})$ much either if we expect Q to be continuous.

Now we take

$$Q(\Phi) = \lim_{n\to\infty} \frac{1}{n} I^{\text{coh}}(\Phi^n)$$

to be the definition of the quantum capacity, where

$$\begin{aligned}I^{\text{coh}}(\mathcal{N}) &= \max_{\rho \in \mathrm{D}(\mathcal{Z}\otimes\mathcal{X})} |I^{\text{coh}}(Z\rangle Y)_{\mathcal{I}\otimes\mathcal{N}(\rho)}| \\ &= \max_{\rho \in \mathrm{D}(\mathcal{Z}\otimes\mathcal{X})} |S(\mathrm{Tr}_{\mathcal{Z}}(\mathcal{I}\otimes\mathcal{N}(\rho))) - S(\mathcal{I}\otimes\mathcal{N}(\rho))|\end{aligned}$$

as before, can be seen as the one shot quantum capacity. We would like to verify that the channel Ψ has a similar capacity to $\mathcal{I}$. We will assume a 2-dimensional Hilbert space $\mathcal{X} = \mathbb{C}^2$ and compute the one shot quantum capacity of Ψ. We will aslo assume that the state that maximizes $I^{\text{coh}}(\mathcal{I})$ also maximizes $I^{\text{coh}}(\Psi)$. In particular if we define

$$\beta = \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right),$$

then we have that

$$\begin{aligned}Q(\mathcal{I}) &= I^{\text{coh}}(\mathcal{I}) \\ &= \max_{\rho \in \mathrm{D}(\mathcal{Z}\otimes\mathcal{X})} |S(\mathrm{Tr}_{\mathcal{Z}}(\mathcal{I}\otimes\mathcal{I}(\rho))) - S(\mathcal{I}\otimes\mathcal{I}(\rho))| \\ &= |S(\mathrm{Tr}_{\mathcal{Z}}(\mathcal{I}\otimes\mathcal{I}(\beta))) - S(\mathcal{I}\otimes\mathcal{I}(\beta))| \\ &= |S(\mathbb{1}_{\mathrm{L}(\mathcal{X})}) - S(\beta)| &(5.1)\\ &= 1 - 0 = 1. &(5.2)\end{aligned}$$

So the coherent information of Ψ will be

$$I^{\text{coh}}(\Psi) = |S(\mathrm{Tr}_{\mathcal{Z}}(\mathcal{I}\otimes\Psi(\beta))) - S(\mathcal{I}\otimes\Psi(\beta))|. \tag{5.3}$$

Notice that

$$\begin{aligned}\mathcal{I}\otimes\Psi(\beta) &= \mathcal{I}\otimes[(1+\varepsilon)\mathcal{I} - \varepsilon R](\beta) \\ &= (1+\varepsilon)\mathcal{I}\otimes\mathcal{I}(\beta) - \varepsilon\mathcal{I}\otimes R(\beta) \\ &= (1+\varepsilon)\beta - \varepsilon(\mathrm{Tr}_{\mathcal{X}}(\beta)\otimes R(\mathrm{Tr}_{\mathcal{Z}}(\beta))) \\ &= (1+\varepsilon)\beta - \varepsilon(\tfrac{1}{2}\mathbb{1}_{\mathrm{L}(\mathcal{Z})}\otimes\phi) &(5.4)\end{aligned}$$

where ϕ is some random density operator given by the completely random map, R, we can express it in terms of its representation in the standard basis:

$$\phi = \begin{pmatrix}\phi_{00} & \phi_{01}\\ \phi_{10} & \phi_{11}\end{pmatrix}.$$

From (5.4), we can also compute the following state

$$\mathrm{Tr}_{\mathcal{Z}}(\mathcal{I}\otimes\Psi(\beta)) = (1+\varepsilon)\tfrac{1}{2}\mathbb{1}_{\mathrm{L}(\mathcal{X})} - \varepsilon\phi. \tag{5.5}$$

We can now compute the eigenvalues of (5.5) and (5.4), to see if their entropies yield a value in (5.3) that is close to (5.2). A direct computation gives us that the eigenvalues of (5.5) are given by

$$\lambda_{\pm} = \tfrac{1}{2} \pm \tfrac{1}{2}\sqrt{1 - [(1+\varepsilon)^2 + 2\varepsilon(1+\varepsilon) + 4\varepsilon^2 \mathrm{Det}(\phi)]}$$

A simple observation (See Appendix A) shows that $\mathrm{Det}(\phi)$ is bounded above since ϕ is a density matrix. Hence $\lambda_{\pm} \to \frac{1}{2}$ as $\varepsilon \to 0$, and so

$$S(\mathrm{Tr}_{\mathcal{Z}}(\mathcal{I} \otimes \Psi(\beta))) \xrightarrow{\varepsilon \to 0} S(\mathbb{1}_{\mathrm{L}(\mathcal{X})}) \tag{5.6}$$

by continuity of the extended entropy. Furthermore we can determine the eigenvalues of (5.4) with the help of some software package such as Mathematica, giving us:

$$\mu_{\pm} = \varepsilon \frac{1}{4}\left(-\mathrm{Tr}(\phi) \pm \sqrt{\kappa(\phi)}\right)$$
$$\nu_{\pm} = \frac{1}{4}\left(2(1+\varepsilon) - \varepsilon\mathrm{Tr}(\phi) \pm \sqrt{4(1+\varepsilon)^2 + \varepsilon^2\kappa(\phi)}\right),$$

where $\kappa(\phi) \equiv 4\phi_{01}\phi_{10} + (\phi_{00} - \phi_{11})^2$. It is easy to see that $\mu_{\pm} \to 0$ as $\varepsilon \to 0$. On the other hand $\nu_{+} \to 1$, but $\nu_{-} \to 0$ when $\varepsilon \to 0$. Hence the eigenvalues approach those of β in the limit of small ε, thus giving us that

$$S(\mathcal{I} \otimes \Psi(\beta)) \xrightarrow{\varepsilon \to 0} S(\beta). \tag{5.7}$$

Now we can conclude from (5.6) and (5.7), that

$$I^{\mathrm{coh}}(\Psi) \xrightarrow{\varepsilon \to 0} I^{\mathrm{coh}}(\mathcal{I}). \tag{5.8}$$

This is an example of how the quantum capacity may be extended to THP maps, and gives us reason to believe that such an extension may also be available for other capacities, such as the two-way assisted quantum capacity.

References

[1] J. Watrous. *Theory of quantum information* (2011). Lecture notes from the course on quantum information hosted at the Institute for Quantum Computing.

[2] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press (2010). ISBN 9781107002173.

[3] A. Holevo. *The capacity of the quantum channel with general signal states*. Information Theory, IEEE Transactions on, 44(1), pp. 269 –273 (jan 1998). ISSN 0018-9448. doi:10.1109/18.651037.

[4] B. Schumacher and M. D. Westmoreland. *Sending classical information via noisy quantum channels*. Phys. Rev. A, 56, pp. 131–138 (Jul 1997). doi:10.1103/PhysRevA.56.131.

[5] I. Devetak. *The private classical capacity and quantum capacity of a quantum channel*. Information Theory, IEEE Transactions on, 51(1), pp. 44 –55 (jan. 2005). ISSN 0018-9448. doi:10.1109/TIT.2004.839515.

[6] P. Shor. *The quantum channel capacity and coherent information* (2002). Lecture notes.

[7] A. Peres. *Separability criterion for density matrices*. Phys. Rev. Lett., 77, pp. 1413–1415 (Aug 1996). doi:10.1103/PhysRevLett.77.1413.

[8] D. Leung and G. Smith. *Continuity of Quantum Channel Capacities*. Communications in Mathematical Physics, 292, pp. 201–215 (Nov. 2009). doi:10.1007/s00220-009-0833-1.

[9] C. Dankert, R. Cleve, J. Emerson, and E. Livine. *Exact and approximate unitary 2-designs and their application to fidelity estimation*. Phys. Rev. A: atomic, molecular, and optical physics, 80(1), p. 12304 (2009).

[10] C. Shannon. *A mathematical theory of communication*. ACM SIGMOBILE Mobile Computing and Communications Review, 5(1), pp. 3–55 (2001).

A Determinant of a density matrix

Given $\phi \in \mathrm{D}(\mathcal{X})$, we have that $\mathrm{Tr}(\phi) = \sum_i \lambda_i = 1$, where λ_i are the eigenvalues of ϕ. This means that $\lambda_i \in [0,1]$ for all i and at least one eigenvalue is non-zero. So observe that

$$\log(\mathrm{Det}(\phi)) = \log\left(\prod_i \lambda_i\right)$$
$$= \sum_i \log \lambda_i \in (-\infty, 0).$$

Thus we can see that

$$\mathrm{Det}(\phi) = 2^{\log(\mathrm{Det}(\phi))} \in (0,1).$$