A general formula for the classical capacity of a general quantum channel

Masahito Hayashi
Lab. for Mathematical Neuroscience
Brain Science Institute, RIKEN
Wako, Saitama, 351-0198, Japan
e-mail: masahito@brain.riken.go.jp

Hiroshi Nagaoka
Graduate School of Information Systems
University of Electro-Communications
Chouhu-shi, Tokyo, 182-8585, Japan
e-mail: nagaoka@is.uec.ac.jp

Abstract — **We derive a general formula of the channel capacity for any (classical-) quantum channel. It can be regarded as a quantum version of Verdú and Han[1]'s result. Our results contains Holevo [2],Schumacher and Westmoreland [3]'s results as the stationary and memoryless case.**

A quantum communication channel is described by a map $W:\mathcal{X}\to\mathcal{S}(\mathcal{H})$. We call a triple (N,φ,Y) a *code* for W when it consists of a natural number (size) N, a mapping (encoding) $\varphi:\{1,\dots,N\}\to\mathcal{X}$ and a POVM (decoding) $Y=\{Y_i\}_{i=1}^N$ on $\mathcal{H}$ such that $\sum_i Y_i\le I$, where $I-\sum_i Y_i$ corresponds to the failure of decoding, and denote the totality of such codes by $\mathcal{C}(\Delta)$. For a code $\Phi=(N,\varphi,Y)\in\mathcal{C}(\Delta)$, the code size and the average error probability are represented as $|\Phi|\stackrel{\rm def}{=}N$ and $\mathrm{P_e}[\Phi]\stackrel{\rm def}{=}\frac{1}{N}\sum_{i=1}^N(1-\mathrm{Tr}\,[W_{\varphi(i)}Y_i])$.

Now let us proceed to the asymptotic setting. Suppose that we are given a sequence $\vec{\mathcal{H}}=\{\mathcal{H}^{(n)}\}_{n=1}^\infty$ of Hilbert spaces and a sequence $\vec{W}=\{W^{(n)}\}_{n=1}^\infty$ of channels $W^{(n)}:\mathcal{X}^{(n)}\to\mathcal{S}(\mathcal{H}^{(n)})$, and consider the following quantities:

$$\sup\{R\,|\,\exists\{\Phi^{(n)}\}\in\vec{\mathcal{C}}_{\vec{W}},\underline{\lim}\frac{\log|\Phi^{(n)}|}{n}\ge R,\lim\mathrm{P_e}[\Phi^{(n)}]=0\},$$
$$\inf\{R\,|\,\forall\{\Phi^{(n)}\}\in\vec{\mathcal{C}}_{\vec{W}},\underline{\lim}\frac{\log|\Phi^{(n)}|}{n}\ge R\Rightarrow\lim\mathrm{P_e}[\Phi^{(n)}]=1\},$$

which are denoted by $C^\dagger(\vec{W})$ and $C(\vec{W})$, respectively, where $\vec{\mathcal{C}}_{\vec{W}}$ denotes the totality of sequences of codes $\vec{\Phi}=\{\Phi^{(n)}\}_{n=1}^\infty$ such that $\Phi^{(n)}\in\mathcal{C}(W^{(n)})$ for all n.

For a self-adjoint trace-class operator A with the spectral decomposition $A=\sum_i\lambda_iE_i$, where $\{\lambda_i\}$ are the eigenvalues and $\{E_i\}$ are the orthogonal projections onto the corresponding eigenspaces, we define the projection $\{A\ge 0\}\stackrel{\rm def}{=}\sum_{i:\lambda_i\ge 0}E_i$. The projections $\{A>0\}$, $\{A\le 0\}$ and $\{A<0\}$ are defined similarly.

For any set $\mathcal{X}$, let $\mathcal{P}(\mathcal{X})$ be the totality of probability distributions on $\mathcal{X}$ with finite supports. That is, an element P of $\mathcal{P}(\mathcal{X})$ is a function $\mathcal{X}\to[0,1]$ such that its support $\mathrm{supp}(x)\stackrel{\rm def}{=}\{x\,|\,P(x)>0\}$ is a finite set and that $\sum_{x\in\mathcal{X}}P(x)=1$. Let the totality of sequences $\vec{P}=\{P^{(n)}\}_{n=1}^\infty$ of $P^{(n)}\in\mathcal{P}(\mathcal{X}^{(n)})$ be denoted by $\vec{\mathcal{P}}(\vec{\mathcal{X}})$, and the totality of $\vec{\sigma}=\{\sigma^{(n)}\}_{n=1}^\infty$ of $\sigma^{(n)}\in\mathcal{S}(\mathcal{H})$ by $\vec{\mathcal{S}}(\vec{\mathcal{H}})$. Given $\vec{P}\in\vec{\mathcal{P}}(\vec{\mathcal{X}})$ and $\vec{\sigma}\in\vec{\mathcal{S}}(\vec{\mathcal{H}})$, we denote

$$\inf\left\{a\left|\lim\sum_{x^n}P^{(n)}(x^n)\mathrm{Tr}\left[W_{x^n}^{(n)}\{W_{x^n}^{(n)}-e^{na}\sigma^{(n)}>0\}\right]=0\right.\right\},$$
$$\sup\left\{a\left|\lim\sum_{x^n}P^{(n)}(x^n)\mathrm{Tr}\left[W_{x^n}^{(n)}\{W_{x^n}^{(n)}-e^{na}\sigma^{(n)}>0\}\right]=1\right.\right\},$$

by $\overline{J}(\vec{P},\vec{\sigma},\vec{W})$ and $\underline{J}(\vec{P},\vec{\sigma},\vec{W})$, respectively. Letting

$$\overline{I}(\vec{P},\vec{W})\stackrel{\rm def}{=}\overline{J}(\vec{P},\vec{W}_{\vec{P}},\vec{W}),\quad\underline{I}(\vec{P},\vec{W})\stackrel{\rm def}{=}\underline{J}(\vec{P},\vec{W}_{\vec{P}},\vec{W}),$$

we have the following theorem, where $\vec{W}_{\vec{P}}$ denotes the sequence $\{W_{P^{(n)}}^{(n)}\}_{n=1}^\infty$ of $W_{P^{(n)}}^{(n)}\stackrel{\rm def}{=}\sum_{x^n\in\mathcal{X}^{(n)}}P^{(n)}(x^n)\,W_{x^n}^{(n)}\in\mathcal{S}(\mathcal{H}^{(n)})$.

Theorem 1

$$C(\vec{W})=\max_{\vec{P}\in\vec{\mathcal{P}}(\vec{\mathcal{X}})}\underline{I}(\vec{P},\vec{W})=\max_{\vec{P}\in\vec{\mathcal{P}}(\vec{\mathcal{X}})}\min_{\vec{\sigma}\in\vec{\mathcal{S}}(\vec{\mathcal{H}})}\underline{J}(\vec{P},\vec{\sigma},\vec{W}),$$
$$C^\dagger(\vec{W})=\max_{\vec{P}\in\vec{\mathcal{P}}(\vec{\mathcal{X}})}\overline{I}(\vec{P},\vec{W})=\max_{\vec{P}\in\vec{\mathcal{P}}(\vec{\mathcal{X}})}\min_{\vec{\sigma}\in\vec{\mathcal{S}}(\vec{\mathcal{H}})}\overline{J}(\vec{P},\vec{\sigma},\vec{W}).$$

Combining this formula and quantum Stein's lemma in quantum hypothesis testing, we can derive the capacity formula in the stationary and memoryless case[4].

As is proven by Hayashi and Nagaoka [4], this theorem is yielded from three inequalities. The first one is the key operator inequality to prove the second, while the second and third lemmas are directly used to prove the theorem. Throughout this paper the generalized inverse of a nonnegative operator A is simply denoted by A^{-1}; i.e., A^{-1} is the nonnegative operator such that $AA^{-1}=A^{-1}A=P_A=P_{A^{-1}}$ where P_A and $P_{A^{-1}}$ denote the orthogonal projections onto the ranges of A and A^{-1}. For any operators $0\le S\le I$ and $T\ge 0$, we have

$$I-\sqrt{S+T}^{-1}S\sqrt{S+T}^{-1}\le 2(I-S)+4T.$$

For any $n\in\mathbb{N}$, $a\in\mathbb{R}$, $N\in\mathbb{N}$ and $P^{(n)}\in\mathcal{P}(\mathcal{X}^{(n)})$, there exists a code $\Phi^{(n)}\in\mathcal{C}(W^{(n)})$ such that $|\Phi^{(n)}|=N$ and

$$\mathrm{P_e}[\Phi^{(n)}]\le 2\sum_{x^n}P^{(n)}(x^n)\mathrm{Tr}\left[W_{x^n}^{(n)}\{W_{x^n}^{(n)}-e^{na}W_{P^{(n)}}^{(n)}\le 0\}\right]+4e^{-na}N.$$

For any $n\in\mathbb{N}$ and any code $\Phi^{(n)}\in\mathcal{C}(W^{(n)})$ with $|\Phi^{(n)}|=N$, there exists a probability distribution $P^{(n)}\in\mathcal{P}(\mathcal{X}^{(n)})$ such that for any $a\in\mathbb{R}$ and $\sigma^{(n)}\in\mathcal{S}(\mathcal{H}^{(n)})$

$$\mathrm{P_e}[\Phi^{(n)}]\ge\sum_{x^n}P^{(n)}(x^n)\mathrm{Tr}\left[W_{x^n}^{(n)}\{W_{x^n}^{(n)}-e^{na}\sigma^{(n)}\le 0\}\right]-\frac{e^{na}}{N}.$$

References

[1] S. Verdú and T.S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, vol.40, 1147–1157, (1994).

[2] A.S. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Trans. Inform. Theory*, vol.44, 269–273, (1998). e-print quant-ph/9611023, (1996).

[3] B. Schumacher and M.D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol.56, 131–138, (1997).

[4] M. Hayashi and H. Nagaoka, "A general quantum formula for channel capacity," in preparation, (2002).