Rate and Error Estimates for Quantum Channels

Andreas Winter

The author is with SFB 343, Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany. Electronic address: `winter@mathematik.uni-bielefeld.de`

Abstract

For the (finite) discrete memoryless quantum channel we obtain strong bounds on the maximal size of codes for transmission of classical information, and bound the reliability function from above and below.

Keywords

Quantum channel, classical capacity, reliability function.

I. Introduction

We study discrete memoryless quantum channels in continuation of the work begun in [8]. Here we restrict ourselves to the finite and stationary case, formally:

Let $W : \mathcal{X} \to \mathfrak{S}(\mathfrak{L}(\mathcal{H}))$ a finite q–DMC, mapping $x \in \mathcal{X}$ to the state W_x, with a set $\mathcal{X}$, say of cardinality $|\mathcal{X}| = a < \infty$, for a fixed complex Hilbert space $\mathcal{H}$ of finite dimension d (this follows [7], in slight modification of [8]). We will have occasion to consider other channels, say V, implicitely all with the same $\mathcal{X}$. Else we follow the notation and definitions of [8], in particular:

An n*–block code* for a quantum channel W is a pair (f, D), where f is a mapping from a finite set $\mathcal{M}$ into $\mathcal{X}^n$ and D is an observable on $\mathfrak{L}(\mathcal{H})^{\otimes n}$ indexed by $\mathcal{M}' \supset \mathcal{M}$, i.e. a partition of $\mathbb{1}$ into positive operators D_m, $m \in \mathcal{M}'$.

The (maximum) error probability of the code is defined as

$$e(f, D) = \max\{1 - \mathrm{Tr}\,(W_{f(m)} D_m) | m \in \mathcal{M}\}$$

(For $f(m) = x^n \in \mathcal{X}^n$ we adopt the convention $W_{f(m)} = W_{x^n} = W_{x_1} \otimes \cdots \otimes W_{x_n}$). We call (f, D) an (n, λ)*–code*, if $e(f, D) \leq \lambda$.

The *rate* of an n–block code is defined as $\frac{1}{n} \log |\mathcal{M}|$. Define $N(n, \lambda)$ as the maximal size (i.e. $|\mathcal{M}|$) of an (n, λ)–code, and

$$e_{\min}(n, R) = \min\{e(f, D) | (f, D) \text{ is } n\text{–block code}, \quad |\mathcal{M}| \geq \exp(nR)\}$$

Finally let $I(P, W) = H(PW) - H(W|P)$ and

$$C(W) = \max_{P \text{ p.d. on } \mathcal{X}} I(P, W)$$

where H denotes the von Neumann entropy, and

$$D(\nu \| \rho) = \mathrm{Tr}\left(\nu(\log \nu - \log \rho)\right)$$
$$D(V \| W | P) = \sum_{x \in \mathcal{X}} P(x) D(V_x \| W_x)$$

the (conditional) quantum I–divergence, where log (as well as exp) is to basis 2 (see [6]).

The rewards of our restriction are strong estimates on $N(n, \lambda)$ (section III), and — more interestingly — upper and lower bounds on $e_{\min}(n, R)$, which lead to nontrivial lower and upper bounds on the reliability function of the channel (section IV). This extends results of Burnashev, Holevo [1] from pure state to general channels, and thus gives (partial) answers to two problems posed by Holevo in [4].

II. Types and typical subspaces

To begin with define the *type* of a sequence $x^n \in \mathcal{X}^n$ to be the empirical distribution of its letters. Further define for a p.d. P on $\mathcal{X}$ the set of typical (or *generated*) sequences in the sense of Wolfowitz [10]

$$\mathcal{T}^n_{V,P,\delta} = \{x^n : \forall x\ |N(x|x^n) - nP(x)| \leq \sigma_x \delta \sqrt{n}\}$$

with $\sigma_x = \sqrt{P(x)(1-P(x))}$ and $N(x|x^n) = |\{i : x = x_i\}|$.

Now we are ready to construct *variance–typical projectors* $\Pi^n_{V,\rho,\delta}$: for a diagonalization $\rho = \sum_j q_j \pi_j$ let $\sigma_j = \sqrt{q_j(1-q_j)}$ and

$$\mathcal{T}^n_{V,\rho,\delta} = \{(j_1, \dots, j_n) : \forall j\ |N(j|j^n) - nq_j| \leq \sigma_j \delta \sqrt{n}\}$$

with $N(j|j^n) = |\{i : j = j_i\}|$, and define

$$\Pi^n_{V,\rho,\delta} = \sum_{(j_1,\dots,j_n) \in \mathcal{T}^n_{V,\rho,\delta}} \pi_{j_1} \otimes \cdots \otimes \pi_{j_n}$$

Define *conditional variance–typical projectors* $\Pi^n_{V,W,\delta}(x^n)$ with $x^n \in \mathcal{X}^n$ by

$$\Pi^n_{V,W,\delta}(x^n) = \bigotimes_{x \in \mathcal{X}} \Pi^{I_x}_{V,W_x,\delta}$$

where $I_x = \{i \in [n] : x_i = x\}$.

Lemma 1: For every state ρ and $n > 0$

$$\mathrm{Tr}\,(\rho^{\otimes n} \Pi^n_{V,\rho,\delta}) \geq 1 - \frac{d}{\delta^2}$$

For every $x^n \in \mathcal{X}^n$

$$\mathrm{Tr}\,(W_{x^n} \Pi^n_{V,W,\delta}(x^n)) \geq 1 - \frac{ad}{\delta^2}$$

Proof: The second inequality obviously follows from the first, which is d applications of Chebyshev's inequality. ■

Lemma 2 (Weak law) Let $\delta > 0$ and $x^n \in \mathcal{X}^n$ of type P. Then

$$\mathrm{Tr}\,(W_{x^n} \Pi^n_{V,PW,\delta\sqrt{a}}) \geq 1 - \frac{ad}{\delta^2}$$

Proof: Diagonalize $PW = \sum_j q_j \pi_j$, and let $\kappa_* : \mathfrak{L}(\mathcal{H})_* \to \mathfrak{L}(\mathcal{H})_*$ the *corset* map defined by $\kappa_*(\sigma) = \sum_j \pi_j \sigma \pi_j$. We claim that

$$\Pi^n_{V,PW,\delta\sqrt{a}} \geq \Pi^n_{V,\kappa_* W,\delta}(x^n)$$

Indeed let $\pi_{j_1} \otimes \cdots \otimes \pi_{j_n}$ one of the product states constituting $\bigotimes_{x\in\mathcal{X}} \Pi^{I_x}_{V,\kappa_*(W_x),\delta}$, i.e. with $\kappa_*(W_x) = \sum_j q_{j|x}\pi_j$

$$\forall x \in \mathcal{X}\, \forall j \;\; \left|N(j|j^{I_x}) - q_{j|x}|I_x|\right| \leq \delta\sqrt{|I_x|}\sqrt{q_{j|x}(1-q_{j|x})}$$

Hence (with $|I_x| = P(x)n$)

$$\begin{aligned} |N(j|j^n) - q_j n| &\leq \sum_{x\in\mathcal{X}} \left|N(j|j^{I_x}) - q_{j|x}|I_x|\right| \\ &\leq \sum_{x\in\mathcal{X}} \delta\sqrt{n}\sqrt{P(x)}\sqrt{q_{j|x}(1-q_{j|x})} \\ &\leq \delta\sqrt{a}\sqrt{n}\sqrt{\sum_{x\in\mathcal{X}} P(x) q_{j|x}(1-q_{j|x})} \\ &\leq \delta\sqrt{a}\sqrt{n}\sqrt{q_j(1-q_j)} \end{aligned}$$

the last inequality by concavity of the map $x \mapsto x(1-x)$, and $q_j = \sum_{x\in\mathcal{X}} P(x) q_{j|x}$.

Thus we can estimate

$$\begin{aligned} \mathrm{Tr}\,(W_{x^n}\Pi^n_{V,PW,\delta\sqrt{a}}) &= \mathrm{Tr}\,\left((\kappa_*^{\otimes n} W_{x^n})\Pi^n_{V,PW,\delta\sqrt{a}}\right) \\ &\geq \mathrm{Tr}\,\left((\kappa_*^{\otimes n} W_{x^n})\Pi^n_{V,\kappa_* W,\delta}(x^n)\right) \\ &\geq 1 - \frac{ad}{\delta^2} \end{aligned}$$

The last line by lemma 1. ■

Every positive operator $B \leq \mathbb{1}$ with $\mathrm{Tr}\,(\rho^n B) \geq \eta$ is called an η–*shadow* of ρ^n. The estimation of η–shadows is the core of the present section. In the following lemmata $K = 2\frac{\log e}{e}$.

Lemma 3 (Cf. [10]) With $\Pi^n = \Pi^n_{V,\rho,\delta}$

$$\begin{aligned} \Pi^n \exp\left(-nH(\rho) - Kd\delta\sqrt{n}\right) \leq \Pi^n \rho^{\otimes n} \Pi^n \leq \\ \leq \Pi^n \exp\left(-nH(\rho) + Kd\delta\sqrt{n}\right) \end{aligned}$$

$$\begin{aligned} \mathrm{Tr}\,\Pi^n_{V,\rho,\delta} &\leq \exp\left(nH(\rho) + Kd\delta\sqrt{n}\right) \\ \mathrm{Tr}\,\Pi^n_{V,\rho,\delta} &\geq \left(1 - \frac{d}{\delta^2}\right)\exp\left(nH(\rho) - Kd\delta\sqrt{n}\right) \end{aligned}$$

Every η–shadow B of $\rho^{\otimes n}$ satifies

$$\mathrm{Tr}\, B \geq \left(\eta - \frac{d}{\delta^2}\right) \exp\left(nH(\rho) - Kd\delta\sqrt{n}\right)$$

Proof: The first estimate is the key: to prove it let $\pi^n = \pi_{j_1} \otimes \cdots \otimes \pi_{j_n}$ one of the eigenprojections of $\rho^{\otimes n}$ which contributes to $\Pi^n_{V,\rho,\delta}$. Then

$$\mathrm{Tr}\,(\rho^{\otimes n}\pi^n) = q_{j_1} \cdots q_{j_n} = \prod_j q_j^{N(j|j^n)}$$

Taking logs and using the defining relation for the $N(j|j^n)$ we find

$$\begin{aligned}\left|\sum_j -N(j|j^n)\log q_j - nH(\rho)\right| &\leq \sum_j -\log q_j |N(j|j^n) - nq_j| \\ &\leq \sum_j -\delta\sqrt{n}\sqrt{q_j}\log q_j \\ &= 2\delta\sqrt{n}\sum_j -\sqrt{q_j}\log\sqrt{q_j} \\ &\leq 2d\frac{\log e}{e}\delta\sqrt{n}\end{aligned}$$

Now the bounds on $\mathrm{Tr}\,\Pi^n_{V,\rho,\delta}$ follow from

$$1 - \frac{d}{\delta^2} \leq \mathrm{Tr}\,(\rho^{\otimes n}\Pi^n_{V,\rho,\delta}) \leq 1$$

Let B an η–shadow of $\rho^{\otimes n}$. Then

$$\begin{aligned}\mathrm{Tr}\,\left(\Pi^n\rho^{\otimes n}\Pi^n B\right) &= \mathrm{Tr}\,(\rho^{\otimes n}B) - \\ &\quad - \mathrm{Tr}\,\left((\rho^{\otimes n} - \Pi^n\rho^{\otimes n}\Pi^n)B\right) \\ &\geq \eta - \frac{d}{\delta^2}\end{aligned}$$

From the upper bound on $\Pi^n\rho^{\otimes n}\Pi^n$ the last inequality of the lemma follows. ■

A simple consequence is

Lemma 4: For every $x^n \in \mathcal{X}^n$ of type P, and with $\Pi^n = \Pi^n_{V,W,\delta}(x^n)$

$$\begin{aligned}\Pi^n \exp\left(-nH(W|P) - Kd\sqrt{a}\delta\sqrt{n}\right) &\leq \Pi^n W_{x^n}\Pi^n \leq \\ &\leq \Pi^n \exp\left(-nH(W|P) + Kd\sqrt{a}\delta\sqrt{n}\right)\end{aligned}$$

$$\mathrm{Tr}\,\Pi^n_{V,W,\delta}(x^n) \leq \exp\left(nH(W|P) + Kd\sqrt{a}\delta\sqrt{n}\right)$$
$$\mathrm{Tr}\,\Pi^n_{V,W,\delta}(x^n) \geq \left(1 - \frac{ad}{\delta^2}\right) \exp\left(nH(W|P) - Kd\sqrt{a}\delta\sqrt{n}\right)$$

Every η–shadow B of W_{x^n} satifies

$$\operatorname{Tr} B \geq \left(\eta - \frac{ad}{\delta^2}\right) \exp\left(nH(W|P) - Kd\sqrt{a}\delta\sqrt{n}\right)$$

Proof: The first inequality is just a times the estimate from the previous lemma. The estimate for $\Pi^n_{V,W,\delta}(x^n) W_{x^n} \Pi^n_{V,W,\delta}(x^n)$ follows from piecing together the estimates for the $\Pi^{I_x}_{V,W_x,\delta}$ in the previous lemma (using $\sum_{x\in\mathcal{X}} \sqrt{P(x)} \leq \sqrt{a}$). The rest is analogous to the previous proof. ■

Of particular interest are the variance–typical projectors with $\delta = 0$, i.e. the $\Pi^n_\rho = \Pi^n_{V,\rho,0}$ and $\Pi^n_W(x^n) = \Pi^n_{V,W,0}(x^n)$, which we call *exact types.*

For the following fix diagonalizations $\rho = \sum_j q_j \pi_j$ and $W_x = \sum_j q_{j|x} \pi_{xj}$. The commutative algebras $\mathbb{C}[\pi_j | j]$ and $\mathbb{C}[\pi_{xj} | j]$ (which are maximal commutative subalgebras of the *commutants* $\mathbb{C}[\rho]'$ and $\mathbb{C}[W_x]'$) will be important below.

Lemma 5: For $\nu \in \mathbb{C}[\rho]'$

$$\Pi^n_\nu \rho^{\otimes n} \Pi^n_\nu = \Pi^n_\nu \exp\left(-nD(\nu\|\rho) - nH(\nu)\right)$$

$$(n+1)^{-d} \exp(nH(\nu)) \leq \operatorname{Tr} \Pi^n_\nu \leq \exp(nH(\nu))$$

For $V_x \in \mathbb{C}[W_x]'$ and $x^n \in \mathcal{X}^n$ of type P

$$\begin{aligned} \Pi^n_V(x^n) W_{x^n} \Pi^n_V(x^n) &= \\ &= \Pi^n_V(x^n) \exp\left(-nD(V\|W|P) - nH(V|P)\right) \end{aligned}$$

$$(n+1)^{-ad} \exp(nH(V|P)) \leq \operatorname{Tr} \Pi^n_V(x^n) \leq \exp(nH(V|P))$$

Proof: The first equation is straightforward. To estimate $\operatorname{Tr} \Pi^n_\nu$ let $\rho = \nu$ and note that

$$(n+1)^{-d} \leq \operatorname{Tr}\left(\nu^{\otimes n} \Pi^n_\nu\right) \leq 1$$

There the upper bound is trivial, whereas the lower bound is by *type counting*, i.e. the observation that in the decomposition

$$\mathbb{1} = \sum_{\hat\nu \in \mathbb{C}[\pi_j | j]} \Pi^n_{\hat\nu}$$

there appear at most $(n+1)^d$ nonzero terms, and the fact that for such $\hat\nu$ the quantity $\operatorname{Tr}\left(\nu^{\otimes n} \Pi^n_{\hat\nu}\right)$ is maximized with $\hat\nu = \nu$. This last statement is directly from [2, lemma 1.2.3]. The second part of the lemma follows from the first by collecting positions of equal letters in x^n. ■

Corollary 6: If $\nu \in \mathbb{C}[\rho]'$ and $\Pi^n_\nu \neq 0$ then

$$(n+1)^{-d} \exp(-nD(\nu\|\rho)) \leq \operatorname{Tr}\left(\rho^{\otimes n} \Pi^n_\nu\right) \leq \exp(-nD(\nu\|\rho))$$

■

Define for a state ρ, channel W, $x^n \in \mathcal{X}^n$ of type P, and a real number L

$$\Pi^n_{\rho,H(\cdot)\leq L} = \sum_{\nu\in\mathbb{C}[\pi_j\,|j],H(\nu)\leq L} \Pi^n_\nu$$

$$\Pi^n_{\rho,H(\cdot)\geq L} = \sum_{\nu\in\mathbb{C}[\pi_j\,|j],H(\nu)\geq L} \Pi^n_\nu$$

$$\Pi^n_{W,H(\cdot|P)\leq L}(x^n) = \sum_{V_x\in\mathbb{C}[\pi_{xj}\,|j],H(V|P)\leq L} \Pi^n_V(x^n)$$

$$\Pi^n_{W,H(\cdot|P)\geq L}(x^n) = \sum_{V_x\in\mathbb{C}[\pi_{xj}\,|j],H(V|P)\geq L} \Pi^n_V(x^n)$$

Lemma 7: For ρ, W, $x^n \in \mathcal{X}^n$ of type P, and L as above

$$\mathrm{Tr}\,\left(\Pi^n_{W,H(\cdot|P)\leq L}(x^n)\right) \leq (n+1)^{ad}\exp(nL)$$

$$\mathrm{Tr}\,\left(W_{x^n}\Pi^n_{W,H(\cdot|P)\leq L}(x^n)\right) \geq$$
$$\geq 1-(n+1)^{ad}\exp(-n\cdot \inf_{H(V|P)>L} D(V\|W|P))$$

$$\mathrm{Tr}\,\left(\rho^{\otimes n}\Pi^n_{\rho,H(\cdot)\geq L}\right) \geq 1-(n+1)^{d}\exp(-n\cdot \min_{H(\nu)\leq L} D(\nu\|\rho))$$

Proof: The inequalities all follow from lemma 5 and corollary 6 together with type counting. ■

Lemma 8: For ρ and L as above

$$\Pi^n_{\rho,H(\cdot)\geq L}\rho^{\otimes n}\Pi^n_{\rho,H(\cdot)\geq L} \leq$$
$$\leq \Pi^n_{\rho,H(\cdot)\geq L}\exp\left(-n\cdot \min_{H(\nu)\geq L}(H(\nu)+D(\nu\|\rho))\right)$$
$$= \Pi^n_{\rho,H(\cdot)\geq L}\exp\left(-nL-n\cdot \min_{H(\nu)=L} D(\nu\|\rho)\right)$$
$$\leq \Pi^n_{\rho,H(\cdot)\geq L}\exp\left(-nL-n\cdot \min_{H(\nu)\leq L} D(\nu\|\rho)\right)$$

For an η–shadow B of $\rho^{\otimes n}$

$$\mathrm{Tr}\,B \geq \left(\eta-(n+1)^{d}\exp(-n\cdot \min_{H(\nu)\leq L} D(\nu\|\rho))\right)\cdot$$
$$\cdot\exp\left(nL+n\cdot \min_{H(\nu)\leq L} D(\nu\|\rho)\right)$$

Proof: The first estimate is directly from lemma 5. To see that the required min is assumed at the boundary of the (convex) region where $H(\nu)\geq L$ observe that the minimized quantity is linear in ν.

For the η–shadow B: note that by lemma 7 with $\Pi^n = \Pi^n_{\rho,H(\cdot)\geq L}$

$$\mathrm{Tr}\left(\rho^{\otimes n}\Pi^n B\Pi^n\right) \geq \eta - (n+1)^d \exp\left(-n\cdot\min_{H(\nu)\leq L} D(\nu\|\rho)\right)$$

and the rest follows by the estimate on $\Pi^n\rho^{\otimes n}\Pi^n$. ■

III. Code bounds

Theorem 9: For every $\lambda\in(0,1)$ there is a constant $K(a,d,\lambda)$ such that for every q–DMC W

$$N(n,\lambda)\geq\exp\left(nC(W)-K(a,d,\lambda)\sqrt{n}\right)$$

Proof: Let P a p.d. on $\mathcal{X}$ with $C(W)=H(PW)-H(W|P)$. Let (f,D) a maximal (n,λ)–code with the property

$$\forall m\in\mathcal{M}\qquad f(m)\in\mathcal{T}^n_{V,P,\sqrt{2a}},\ \mathrm{Tr}\,D_m\leq\mathrm{Tr}\,\Pi^n_{V,W,\delta}(f(m))$$

with $\delta=\sqrt{\frac{2ad}{\lambda}}$. In particular (by lemma II.4)

$$\mathrm{Tr}\,D_m\leq\exp\left(nH(W|P)+(Kd\sqrt{a}\delta+Ka\sqrt{2a}\log d)\sqrt{n}\right)$$

Let $B=\sum_{m\in\mathcal{M}}D_m$, we claim that for all $x^n\in\mathcal{T}^n_{V,P,\sqrt{2a}}$

$$\mathrm{Tr}\,(W_{x^n}B)\geq\eta=\min\{1-\lambda,\lambda^2/32\}$$

This is clear if x^n is a code word, and true else, for otherwise we could extend our code with the word x^n and decoding operator

$$D'=\sqrt{\mathbb{1}-B}\,\Pi^n_{V,W,\delta}(x^n)\sqrt{\mathbb{1}-B}$$

This is exactly as in [8]. Thus we arrive at

$$\mathrm{Tr}\,\left((PW)^{\otimes n}B\right)\geq\eta/2$$

which implies by lemma II.3 the estimate

$$\mathrm{Tr}\,B\geq\left(\frac{\eta}{2}-\frac{d}{\delta_0^2}\right)\exp\left(nH(PW)-Kd\delta_0\sqrt{n}\right)$$

Choosing $\delta_0=\sqrt{\frac{4d}{\eta}}$ the proof is complete. ■

Theorem 10: For every $\lambda\in(0,1)$ there is a constant $K(a,d,\lambda)$ such that for every q–DMC W and every (n,λ)–code (f,D)

$$|\mathcal{M}|\leq(n+1)^a\exp\left(nC(W)+K(a,d,\lambda)\sqrt{n}\right)$$

Proof: We will prove even more: under the additional assumption that all code words are of the same type P (such codes are called *constant composition*) one has

$$|\mathcal{M}| \leq \exp\left(nI(P,W) + K(a,d,\lambda)\sqrt{n}\right)$$

(From which the theorem clearly follows).

To see this modify the decoder as follows: let

$$D'_m = \Pi^n_{V,PW,\delta} D_m \Pi^n_{V,PW,\delta}$$

with $\delta = \frac{\sqrt{32ad}}{1-\lambda}$. Then (f, D') is an $(n, \frac{1+\lambda}{2})$–code:

$$\begin{aligned}
\mathrm{Tr}\,(W_{f(m)}D'_m) &= \mathrm{Tr}\,(W_{f(m)}\Pi^n_{V,PW,\delta} D_m \Pi^n_{V,PW,\delta}) \\
&= \mathrm{Tr}\,(W_{f(m)}D_m) - \\
&\qquad - \mathrm{Tr}\,\left((W_{f(m)} - \Pi^n W_{f(m)} \Pi^n)D_m\right) \\
&\geq 1 - \lambda - \frac{1-\lambda}{2}
\end{aligned}$$

(the last line by lemma II.2 and the *tender operator* lemma V.8 of [8]). Now from lemma II.4

$$\begin{aligned}
\mathrm{Tr}\,D'_m &\geq \left(\frac{1-\lambda}{2} - \frac{ad}{\delta^2}\right)\exp(nH(W|P) - Kd\sqrt{a}\delta\sqrt{n}) \\
&\geq \frac{1-\lambda}{4}\exp(nH(W|P) - Kd\sqrt{a}\delta\sqrt{n})
\end{aligned}$$

On the other hand $\sum_{m\in\mathcal{M}} D'_m \leq \Pi^n_{V,PW,\delta}$, hence by lemma II.3

$$\sum_{m\in\mathcal{M}} \mathrm{Tr}\,D'_m \leq \exp\left(nH(PW) + Kd\delta\sqrt{n}\right)$$

and we are done. ■

IV. Reliability function

For the finite q–DMC W with capacity $C(W)$ the *reliability function* $E(R)$ is defined

$$E(R) = \liminf_{n\to\infty} -\frac{1}{n}\log e_{\min}(n, R)$$

From the previous section we see that $E(R) = 0$ for $R > C(W)$. On the other hand define the *greedy* bound

$$E_{\mathrm{g}}(R,P) = \max\{\min\{\mu_{\mathrm{i}}(L,P), \frac{1}{2}\mu_{\mathrm{c}}(L',P)\} : R \leq L' - L\}$$

with the *individual exponent* (which may be $+\infty$)

$$\mu_{\mathrm{i}}(L,P)=\inf\{D(V\|W|P):H(V|P)>L\}$$

and the *collective exponent* (which is finite)

$$\mu_{\mathrm{c}}(L',P)=\min\{D(\rho\|PW):H(\rho)\leq L'\}$$

Then we have

Theorem 11: For $n>0$, a type P and $R<I(P,W)$ there exist constant composition n-block codes (f,D) of type P with

$$|\mathcal{M}|\geq(n+1)^{d-ad}\exp(nR)$$

and error probability

$$e(f,D)\leq 8(n+1)^{ad}\exp(-nE_{\mathrm{g}}(R,P))$$

if $n\geq n_0(a,d,P)$.

Proof: Let L,L' a pair of numbers with $R\leq L'-L$ and

$$E_{\mathrm{g}}(R,P)=\min\{\mu_{\mathrm{i}}(L,P),\frac{1}{2}\mu_{\mathrm{c}}(L',P)\}$$

It is easily seen that we may assume $\mu_{\mathrm{i}}(L,P)\geq\frac{1}{2}\mu_{\mathrm{c}}(L',P)$. Also that in this case $L'<H(PW)$ and $L\geq H(W|P)$, in particular $E_{\mathrm{g}}(R,P)>0$.

Define $\lambda=8(n+1)^{ad}\exp(-nE_{\mathrm{g}}(R,P))$ and assume n large enough such that $\eta=\frac{\lambda^2}{32}\leq 1-\lambda$. Let (f,D) a maximal (n,λ)–code with the additional requirement

$$\forall m\in\mathcal{M}\qquad\mathrm{Tr}\,D_m\leq(n+1)^{ad}\exp(nL)$$

We claim that with $B=\sum_{m\in\mathcal{M}}D_m$

$$\forall x^n\text{ of type }P\qquad\mathrm{Tr}\,(W_{x^n}B)\geq\eta$$

For else we could extend our code by an exceptional x^n and corresponding decoding operator

$$D'=\sqrt{\mathbb{1}-B}\,\Pi^n_{W,H(\cdot|P)\leq L}(x^n)\sqrt{\mathbb{1}-B}$$

The argument is as in [8]: observe that $\Pi^n_{W,H(\cdot|P)\leq L}(x^n)$ and hence D' satisfies the trace requirement, and

$$\mathrm{Tr}\left(W_{x^n}\Pi^n_{W,H(\cdot|P)\leq L}(x^n)\right)\geq 1-(n+1)^{ad}\exp(-n\mu_{\mathrm{i}}(L,P))$$

Consequently

$$\mathrm{Tr}\left((PW)^{\otimes n}B\right)\geq\eta(n+1)^{-a}$$

and by lemma II.8

$$\begin{aligned}\operatorname{Tr} B &\geq \big(\eta(n+1)^{-a} - (n+1)^d \exp(-n\mu_{\mathrm{c}}(L', P))\big) \cdot \\ &\qquad\qquad\qquad \cdot \exp(nL' + n\mu_{\mathrm{c}}(L', P)) \\ &\geq (n+1)^d \exp(nL')\end{aligned}$$

from which the estimate on $|\mathcal{M}|$ follows immediately. ■

Corollary 12: For $0 \leq R \leq C(W)$

$$E(R) \geq E_{\mathrm{g}}(R) = \max_{P \text{ p.d.: } R \leq I(P,W)} E_{\mathrm{g}}(R, P)$$

Proof: It suffices to check lower semicontinuity of E_{g} in P, i.e.

$$\liminf_{Q \to P} E_{\mathrm{g}}(R, Q) \geq E_{\mathrm{g}}(R, P)$$

(the ingredients are lower semicontinuity of μ_{i} and μ_{c}) and continuity in R from the right, i.e.

$$\lim_{\delta \searrow 0} E_{\mathrm{g}}(R + \delta, P) = E_{\mathrm{g}}(R, P)$$

(which is quite obvious). ■

Conversely, defining the *sphere packing* bound

$$E_{\mathrm{sp}}(R, P) = \min_{V \text{ channel: } I(P,V) \leq R} D(V \| W | P)$$

we have

Theorem 13: For $R \geq 0$ and $n > 0$ let (f, D) a constant composition n–block code (of type P) with

$$|\mathcal{M}| \geq \exp(n(R + \delta))$$

Then the error probability

$$e(f, D) \geq \frac{1}{2} \exp(-nE_{\mathrm{sp}}(R, P)(1 + \delta))$$

if $n \geq n_0(a, d, \delta)$.

Proof: We can directly apply the original idea of Haroutunian [3]: consider a channel $V : \mathcal{X} \to \mathfrak{S}(\mathfrak{L}(\mathcal{H}))$ with $I(P, V) \leq R$. From the proof of the strong converse theorem III.10 we see that $e(f, D) \geq 1 - \frac{\delta}{2}$ if n is large enough (we assume $\delta < 1$). I.e. for some message $m \in \mathcal{M}$ and $S_m = \mathbb{1} - D_m$

$$\operatorname{Tr}(V_{f(m)} S_m) \geq 1 - \frac{\delta}{2}$$

Now generally for two states ρ, σ and complementary positive operators S, D (i.e. $S + D = \mathbb{1}$) one has

$$\operatorname{Tr}(\rho S) \log \frac{\operatorname{Tr}(\rho S)}{\operatorname{Tr}(\sigma S)} + \operatorname{Tr}(\rho D) \log \frac{\operatorname{Tr}(\rho D)}{\operatorname{Tr}(\sigma D)} \leq D(\rho \| \sigma)$$

This follows immediately from Uhlmann's monotonicity of quantum I–divergence [9], applied to the completely positive, trace preserving map

$$\begin{aligned}\mathfrak{L}(\mathcal{H})_* &\longrightarrow \mathbb{C}^2 \\ \alpha &\longmapsto \mathrm{Tr}\,(\alpha S)e_1 + \mathrm{Tr}\,(\alpha D)e_2\end{aligned}$$

From this we get by elementary operations

$$\mathrm{Tr}\,(\sigma S) \geq \exp\left(-\frac{D(\rho\|\sigma) + h(\mathrm{Tr}\,(\rho S))}{\mathrm{Tr}\,(\rho S)}\right)$$

Applying this to $\rho = V_{f(m)}$, $\sigma = W_{f(m)}$ and $S = S_m$, $D = D_m$ we find

$$\begin{aligned}\mathrm{Tr}\,(W_{f(m)}S_m) &\geq \exp\left(-\frac{nD(V\|W|P) + h\left(1-\frac{\delta}{2}\right)}{1-\frac{\delta}{2}}\right) \\ &\geq \frac{1}{2}\exp(-nD(V\|W|P)(1+\delta))\end{aligned}$$

if only δ is small enough (which is no real restriction). Now we choose V such that $D(V\|W|P)$ is minimal. ■

Corollary 14: For $0 \leq R \leq C(W)$ (with possible exception of the leftmost finite value of E_{sp})

$$E(R) \leq E_{\mathrm{sp}}(R) = \max_{P \text{ p.d.}} E_{\mathrm{sp}}(R, P)$$

Proof: To apply the theorem we have just to note the continuity of E_{sp} in R, which follows from its convexity. ■

Remark 15: The proof obviously also works for infinite input alphabet, if only we have a strong converse which indeed we have: see [8].

Remark 16: The reader may wish to apply the techniques of the previous proofs to show that $e(f, D)$ tends to 1 exponentially for rates above the capacity. The results however yield nothing of interest beyond the analysis of Ogawa, Nagaoka [5].

V. Conclusion

We showed that for transmission of classical information through quantum channels one has code bounds and exponential error convergence similar to classical channels: in particular we have the sphere packing bound and a lower bound on the reliability function which at least shows its positivity for rates below the capacity. For the pure state channel this is matched by random coding and expurgated lower bounds in [1]. Unfortunately in this case our sphere packing bound is trivial.

We leave as open problems: the proof of a random coding lower bound in the general case (which would enable us to determine the reliability function above a critical rate), and in the pure state case to find a suitable modification of the sphere packing bound (as the present formulation does not take into account possible noncommutativity).

References

[1] M. V. Burnashev, A. S. Holevo, "On Reliability Function of Quantum Communication Channel", LANL eprint `quant-ph/9703013`

[2] I. Csiszár, J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, New York 1981

[3] E. A. Haroutunian, "Estimates of the error exponent for the semicontinuous memoryless channel" (in Russian), Problemy Peredachi Informatsii **4**,4 (1968), 37–48

[4] A. S. Holevo, "Coding Theorems for Quantum Channels", Tamagawa University Research Review no.4, 1998 (an extended version as LANL eprint `quant-ph/9809023`)

[5] T. Ogawa, H. Nagaoka, "Strong Converse to the Quantum Channel Coding Theorem", LANL eprint `quant-ph/9808063`

[6] M. Ohya, D. Petz, *Quantum Entropy and its Use*, Berlin, Heidelberg, New York 1993

[7] A. Winter, "Languages of Quantum Information Theory", Preprint Ergänzungsreihe 98–009 of SFB 343 "Diskrete Strukturen in der Mathematik", Universität Bielefeld, 1998 (also LANL eprint `quant-ph/9807008`)

[8] A. Winter, "Coding Theorem and Strong Converse for Nonstationary Quantum Channels", to appear. See also: "Coding Theorem and Strong Converse for Quantum Channels", Preprint 98–074 of SFB 343 "Diskrete Strukturen in der Mathematik", Universität Bielefeld, 1998

[9] A. Uhlmann,"Relative Entropy and the Wigner–Yanase–Dyson–Lieb Concavity in an Interpolation Theory", Comm. Math. Phys. **54** (1977), 21–32

[10] J. Wolfowitz, *Coding Theorems of Information Theory*, Berlin 1964