Capacity of Quantum Arbitrarily Varying Channels

Rudolf Ahlswede*and Vladimir Blinovsky†

We prove that the average error capacity C_q of a quantum arbitrarily varying channel (QAVC) equals 0 or else the random code capacity $\bar{C}$ (Ahlswede's dichotomy). We also establish a necessary and sufficient condition for $C_q > 0$.

I. Introduction and Results

We define the QAVC by the double indexed finite set of density operators $\rho_{x,s},\ x \in \mathcal{X},\ s \in \mathcal{S}$ on $\mathbb{C}^d$. $\mathcal{X}$ is the set of code symbols and $\mathcal{S}$ is the set of states of the QAVC. As usual we consider the scheme of n uses of the channel. A code $\mathcal{C}$ of cardinality N and length n is a set of pairs $\{(c_i^n, I_i),\ i = 1, \dots N\}$, where $c_i^n = (c_{i,1}, \dots, c_{i,n}),\ c_{i,j} \in \mathcal{X}$ and $\{I_i\}$ is the resolution of the identity in $\left(\mathbb{C}^d\right)^{\bigotimes n}$. More precisely, the set $\{I_i\}$ has the property that $\{I_0 = I - \sum_{i=1}^N I_i, I_i\}$ is the resolution of identity. We omit this explanation later, and just say that $\{I_i\}$ is the resolution of identity. The probability $P_e(\mathcal{C}, s^n)$ of the avarage decoding error of the code $\mathcal{C}$, when the state (sequence of states) of the QAVC is $s^n = (s_1, \dots, s_n)$, is defined as follows:

$$P_e(\mathcal{C}, s^n) = 1 - \frac{1}{N} \sum_{i=1}^{N} Tr\left(\rho_{c_i^n, s^n} I_i\right). \tag{1}$$

Here $\rho_{c_i^n, s^n} = \rho_{c_{i,1}, s_1} \bigotimes \dots \bigotimes \rho_{c_{i,n}, s_n}$.

*The author was supported by the 'Finite Structures' Marie Curie Host Fellowship for the Transfer of knowledge project carried out by Alfred Renyi Institute of Mathematics, in the framework of the European Community's Structuring the European Research Area programme.

†This work partially support by RFFI grants No 03-01-00592 and 03-01-00098

The QAVC includes the prototype in classical Information Theory. Classical the arbitrarily varying channel (AVC) is defined by the set of double indexed conditional probabilities on $\{W(y|x,s):\ x\in\mathcal{X},\ s\in\mathcal{S}\}$ on finite set $\mathcal{Y}$. Set $\mathcal{S}$ is called the set of states of the AVC. Actually one should consider the n uses of the AVC and the probability that the output of the AVC is $y^n=(y_1,\ldots,y_n)\in\mathcal{Y}^n$ when $x^n=(x_1,\ldots,x_n)\in\mathcal{X}^n$ was sent and the state of the AVC is $s^n=(s_1,\ldots,s_n)\in\mathcal{S}^n$:

$$W(y^n|x^n,s^n)=\prod_{i=1}^{n}W(y_i|x_i,s_i).$$

If one uses the (classical) code $\mathcal{K}=\{(x_i^n,\mathcal{D}_i):\ i=1,\ldots,N\}$, $\bigcup_{i=1}^{N}\mathcal{D}_i=\mathcal{Y}^n$, $\mathcal{D}_i\bigcap\mathcal{D}_j=\emptyset,\ i\neq j$, then the probability $P_{ec}(\mathcal{K},s^n)$ of the average decoding error, when the state of the AVC is s^n is defined as follows

$$P_{ec}(\mathcal{K},s^n)=\frac{1}{N}\sum_{i=1}^{N}\sum_{y^n\notin\mathcal{D}_i}W(y_i^n|x_i^n,s^n).$$

Let for some $R>0$, for every $\epsilon>0,\delta>0$ and sufficiently large n exist a code $\mathcal{K}$ of cardinality N such that

$$\frac{\log N}{n}>R-\delta,\ \max_{s^n\in\mathcal{S}^n}P_{ec}(\mathcal{K},s^n)<\epsilon.$$

The supremum over all such R is called the capacity C of the AVC.

Classical AVC were studied for average errors in several papers. It suffices here to refer to [3], [1], [6], and [5], where further references can be found. An AVC is symmetrizable if there exists a parameterized set of distributions on $\mathcal{S}$ $\{U(s|x):\ x\in\mathcal{X}\}$ such that for all $x,x'\in\mathcal{X},\ y\in\mathcal{Y}$ the following equalities are valid

$$\sum_{s\in\mathcal{S}}W(y|x,s)U(s|x')=\sum_{s\in\mathcal{S}}W(y|x',s)U(s|x).$$

In [6] was proved that the capacity C of the AVC is 0 if it is symmetrizable and if $C>0$, then in [5] was proved

$$C=\max_{P}\min_{Q}I_{P,Q}(X;Y),$$

where

$$I_{P,Q}(X;Y)=\sum_{x\in\mathcal{X},\ y\in\mathcal{Y}}P(x)W_Q(y|x)\frac{W_Q(y|x)}{\sum_{x\in\mathcal{X}}P(x)W_Q(y|x)}$$

and

$$W_Q(y|x) = \sum_{s\in\mathcal{S}} Q(s)W(y|x,s),$$

P, Q are distributions on X and S respectively.

Now we return to the investigation of the QAVC. We define the capacity C_q of the QAVC in the same way as in the classical case. Let there for some $R > 0$, for every $\epsilon > 0, \delta > 0$ and sufficiently large n exist a quantum code $\mathcal{C}$ of cardinality N such that

$$\frac{\log N}{n} > R - \delta, \ \max_{s^n\in S^n} P_e(\mathcal{C}, s^n) < \epsilon.$$

The supremum over all such R is called the capacity C_q of the QAVC. We say that the QAVC is symmetrizable if there exists a parameterized set of distributions $\{U(s|x) : x \in \mathcal{X}\}$ on $\mathcal{S}$ such that for all $x, x' \in \mathcal{X}$ the following equalities are valid

$$\sum_{s\in\mathcal{S}} U(s|x')\rho_{x,s} = \sum_{s\in\mathcal{S}} U(s|x)\rho_{x',s}.$$

The following simple fact is valid.

Statement 1 *If QAVC is symmetrizable, then* $C_q = 0$.

Proof. The proof is analogous to the proof in [6]. Let QAVC be symmetrizable and

$$e(i, s^n) = 1 - Tr(\rho_{c_i^n,s^n}, I_i) = \sum_{j\neq i} Tr(\rho_{c_i^n,s^n}, I_j),$$

then

$$\begin{aligned}
\mathbb{E}(e(i, S_k^n)) &\stackrel{\Delta}{=} \sum_{j\neq i}\sum_{s^n\in\mathcal{S}^n} U(s^n|c_k^n)Tr(\rho_{c_i^n,s^n}, I_j) \\
&= \sum_{j\neq i} Tr\left(\bigotimes_{m=1,\ldots,n}\sum_{s_m\in\mathcal{S}} U(s_m, c_{k,m})\rho_{c_{i,m},s_m}, I_j\right) \\
&= \sum_{j\neq i} Tr\left(\bigotimes_{m=1,\ldots,n}\sum_{s_m\in\mathcal{S}} U(s_m, c_{i,m})\rho_{c_{k,m},s_m}, I_j\right) \\
&= \sum_{j\neq i}\sum_{s^n\in\mathcal{S}^n} U(s^n|c_i^n)Tr(\rho_{c_k^n,s^n}, I_j).
\end{aligned}$$

Note that $\mathbb{E}(e(c^n, S_j^n))$ is the mathematical expectation based on the distribution of random variable S_j^n :

$$P(S_j^n = s^n) = \prod_{m=1}^{n} U(s_m|c_{j,m}).$$

Since

$$\mathbb{E}(e(k, S_i^n)) = \sum_{j\neq k} \sum_{s^n\in\mathcal{S}^n} U(s^n|c_i^n) Tr(\rho_{c_k^n,s^n}, I_j)$$

we have for $i \neq k$

$$\mathbb{E}(e(k, S_i^n)) + \mathbb{E}(e(i, S_k^n)) \geq 1.$$

Thus

$$\begin{aligned} & \frac{1}{N}\sum_{j=1}^{N} \mathbb{E}(P_e(\mathcal{C}, S_j^n)) \\ = \quad & \frac{1}{N^2}\sum_{k,j=1}^{N} \mathbb{E}(e(k, S_j^n)) \geq \frac{1}{N^2}\binom{N}{2} \geq \frac{1}{4}. \end{aligned}$$

It follows that there exists a $j \in \{1, \ldots, N\}$ such that

$$\mathbb{E}P_e(\mathcal{C}, S_j^n) \geq \frac{1}{4}. \tag{2}$$

From (2) follows that there exists an $s^n \in \mathcal{S}^n$ such that

$$P_e(\mathcal{C}, s^n) \geq \frac{1}{4}$$

and this proves Statement 1.

Hence necessary for $C_q > 0$ is that the QAVC not symmetrizable. The main contribution of the present work is the proof that if QAVC is not symmetrizable then actually $C_q = \bar{C} > 0$, where

$$\bar{C} = \max_P \min_Q H(P, Q),$$

$$\begin{aligned} H(P, Q) &= H\left(\sum_{x\in\mathcal{X}} P(x)\rho_x^Q\right) - \sum_{x\in\mathcal{X}} P(x) H(\rho_x^Q), \\ \rho_x^Q &= \sum_{s\in\mathcal{S}} Q(s)\rho_{x,s} \end{aligned}$$

and $H(\rho)$ is the entropy of the density ρ : $H(\rho) = -Tr(\rho \log \rho)$.

Theorem 1 *If QAVC is not symmetrizable, then* $C_q = \bar{C} > 0$.

The plan of the proof of Theorem 1 is as follows: first we prove the fact, that $C_q = \bar{C}$ or else zero (Ahlswede's dichotomy). Then we will prove, that if QAVC is not symmetrizable, then $C_q > 0$.

To prove Theorem 1 we need some results about random quantum codes. Let $\Gamma = (H^n)^N \bigotimes \left((\mathbb{C}^{d^2})^{\otimes n}\right)^{\otimes N}$. A random quantum code $\left(\{\mathcal{C}^\gamma : \gamma \in \Gamma\}, G\right)$ consists of the (possibly infinite) family of sets of N pairs $\mathcal{C}^\gamma = \{(c_i^\gamma, I_i^\gamma)\colon\ i = 1, \ldots, N\},\ \gamma \in \Gamma$, together with the distribution G on Γ. Because for given n the number of $c_i^{n\,\gamma}$ is finite and I_i^γ are matrices with the linearity restrictions on their elements the natural property of measurability of sets, which we consider, follows. Here $I_i^\gamma,\ i = 1, \ldots, N$ is as before the resolution of identity, $c_i^{n\,\gamma} = (c_{i,1}^{n\,\gamma}, \ldots, c_{i,n}^{n\,\gamma}) \in \mathcal{X}^n$ and operators are considered in $\left(\mathbb{C}^d\right)^{\otimes n}$. The capacity of a QAVC under random quantum coding is defined in the same manner as for usual quantum codes but with

$$P_{er} = \inf_G \max_{s^n \in \mathcal{S}^n} \sum_{\gamma \in \Gamma} G(\gamma) P_e(\mathcal{C}^\gamma, s^n)$$

instead of $\inf_{\mathcal{C}} \max_{s^n \in \mathcal{S}^n} P_e(\mathcal{C}, s^n)$. We denote this capacity by $\tilde{C}$.

To prove Theorem 1 we need the following two lemmas.

Lemma 1 *The following equality is valid*

$$\tilde{C} = \bar{C}.$$

Lemma 2 *If* $C_q > 0$, *then*

$$C_q = \bar{C}.$$

II. Proofs of Lemmas 1, 2 and Theorem 1

We start with the proof of the Lemma 1. Denote

$$\rho^{P,Q} = \sum_{x \in \mathcal{X},\ s \in \mathcal{S}} P(x) Q(s) \rho_{x,s}$$

and in the basis, where $\rho^{P,Q}$ is diagonal, we have

$$\rho^{P,Q} = \sum_j \lambda_j^{P,Q} |e_j^{P,Q} >< e_j^{P,Q}|,$$

$\lambda_j^{P,Q} > 0$, $\sum_j \lambda_j^{P,q} = 1$. Then for $P, Q^n = (Q_1, \ldots, Q_n)$, where P, Q_j is the distribution on $\mathcal{X}$ or $\mathcal{S}$ correspondingly, we have

$$\rho^{P,Q^n} = \rho^{P,Q_1} \bigotimes \ldots \bigotimes \rho^{P,Q_n} = \sum_{j^n} \lambda_{j^n}^{P,Q^n} |e_{j^n}^{P,Q^n} >< e_{j^n}^{P,Q^n}|,$$

where

$$\begin{aligned} \lambda_{j^n}^{P,Q^n} &= \lambda_{j_1}^{P,Q_1} \ldots \lambda_{j_n}^{P,Q_n}, \\ |e_{j^n}^{P,Q^n} > &= \bigotimes |e_{j_1}^{P,Q_1} > \bigotimes \ldots \bigotimes |e_{j_n}^{P,Q_n} > . \end{aligned}$$

Denote for some $\delta > 0$

$$\begin{aligned} \mathcal{F} &\stackrel{\Delta}{=} \left\{ j^n;\ 2^{-\sum_{i=1}^n H(\rho^{P,Q_i}) - \delta n} < \lambda_{j^n}^{P,Q^n} < 2^{-\sum_{i=1}^n H(\rho^{P,Q_i}) + \delta n} \right\}, \\ \Pi^{P,Q^n} &\stackrel{\Delta}{=} \sum_{j^n \in \mathcal{F}} |e_{j^n}^{P,Q^n} >< e_{j^n}^{P,Q^n}|. \end{aligned}$$

Then by Chebyshev's inequality for given P, Q^n, the distribution $\lambda_{j^n}^{P,Q^n}$ on j^n has the property

$$\begin{aligned} K_n &\stackrel{\Delta}{=} Pr(\mathcal{F}^c) \qquad (3) \\ &= Pr\left(\left| \sum_{i=1}^n \log \lambda_{j_i}^{P,Q_i} + \sum_{i=1}^n H\left(\rho^{P,Q_i}\right) \right| > n\delta \right) \\ &< \frac{\max_Q \sum_j \lambda_j^{P,Q} \log^2 \lambda_j^{P,Q}}{n\delta^2} \stackrel{n\to\infty}{\to} 0. \end{aligned}$$

Thus

$$P(j^n \in \mathcal{F}) = 1 - K_n \stackrel{n\to\infty}{\to} 1. \qquad (4)$$

In the same manner for $x \in \mathcal{X}$, $x^n = (x_1, \ldots, x_n) \in \mathcal{X}^n$ we define

$$\begin{aligned} \rho_x^Q &\stackrel{\Delta}{=} \sum_{s\in\mathcal{S}} Q(s)\rho_{x,s} = \sum_j \lambda_{x,j}^Q |e_{x,j}^Q >< e_{x,j}^Q|, \\ \rho_{x^n}^{Q^n} &\stackrel{\Delta}{=} \rho_{x_1}^{Q_1} \bigotimes \ldots \bigotimes \rho_{x_n}^{Q_n} = \sum_{j^n} \lambda_{x^n,j^n}^{Q^n} |e_{x^n,j^n}^{Q^n} >< e_{x^n,j^n}^{Q^n}|, \\ \mathcal{F}_{x^n} &\stackrel{\Delta}{=} \left\{ j^n : 2^{-\sum_{i=1}^n \bar{H}(\rho_\cdot^{Q_i}) - \delta n} < \lambda_{x^n,j^n}^{Q^n} < 2^{-\sum_{i=1}^n \bar{H}(\rho_\cdot^{Q_i}) + \delta n} \right\} \\ \Pi_{x^n}^{Q^n} &\stackrel{\Delta}{=} \sum_{j^n \in \mathcal{F}_{x^n}} |e_{x^n,j^n}^{Q^n} >< e_{x^n,j^n}^{Q^n}|. \end{aligned}$$

Here

$$\begin{aligned} j^n &= (j_1, \ldots, j_n) \\ \lambda^{Q^n}_{x^n, j^n} &= \lambda^{Q_1}_{i_1, j_1} \ldots \lambda^{Q_n}_{i_n, j_n} \\ \bar{H}(\rho^Q_\cdot) &= \sum_{x \in \mathcal{X}} P(x) H(\rho^Q_x). \end{aligned}$$

The distribution $\lambda^{Q^n}_{x^n, j^n}$ on $\{j^n\}$ has analogously to (3) the property that

$$Pr(\mathcal{F}_{x^n}) \stackrel{n \to \infty}{\to} 1$$

which also has an analogous proof. Define the operator

$$I^{Q^n}_i \triangleq \left(\sum_{j=1}^N \Pi^{P,Q^n} \Pi^{Q^n}_{c^n_j} \Pi^{P,Q^n} \right)^{-1/2} \Pi^{P,Q^n} \Pi^{Q^n}_{c^n_i} \Pi^{P,Q^n} \left(\sum_{j=1}^N \Pi^{P,Q^n} \Pi^{Q^n}_{c^n_j} \Pi^{P,Q^n} \right)^{-1/2}.$$

Here $H^{-1/2} \equiv 0$ on the space $\mathbb{F} = \left\{ r^n \in (\mathbb{C}^d)^{\bigotimes n} : \ Hr^n = 0 \right\}$ and $\left(H^{-1/2}\right)^2 H = I$ on $(\mathbb{C}^d)^{\bigotimes n} - \mathbb{F}$. Operators $I^{Q^n}_{c^n_i}$ are positive,

$$\sum_{i=1}^N I^{Q^n}_{c^n_i} \leq I$$

and $\mathcal{C}(Q^n) = \{(c_i, I^{Q^n}_{c^n_i}), \ i = 1, \ldots, N\}$ is the quantum code.

Now we consider the distributions Q_i on S which take values from the set $m/M, \ m = 0, \ldots, M$ for some sufficiently large M. The number of such distributions is finite for given M. This gives us a finite number of resolutions of identity $I^{Q^n}_i$, the property which we will use later. Now we take the average of the probability of decoding errors (1) over some distributions T_i on $\mathcal{S}$ and obtain the average probability of decoding error

$$P_e(\mathcal{C}, Q^n, T^n) = 1 - \frac{1}{N} \sum_{i=1}^N Tr(\rho^{T^n}_{c^n_i} I^{Q^n}_i),$$

where $T^n = (T_1, \ldots, T_n)$, each T_i is the distribution on S and

$$\rho^{T^n}_{c^n_i} = \sum_{s_j \in \mathcal{S}} T_1(s_1) \ldots T_n(s_n) \rho_{c^n_i, s^n}.$$

It is clear that for an arbitrary distribution T on S and $\epsilon > 0$ there exists a distribution from $Q(M)$ such that

$$\max_{s\in\mathcal{S}} |T(s) - Q(s)| < \epsilon.$$

Now we approximate each T_i by some $Q_i \in Q(M)$ in the above sense.

We are going to show that for every $\epsilon_1 > 0$ and $T^n = (T_1, \ldots, T_n)$, for sufficiently large M there exists a code $\mathcal{C}(Q^n) = \{(c_i^n, I_i^{Q^n}) : \ i = 1, \ldots, N\}$ such that

$$P_e(\mathcal{C}, Q^n, T^n) < \epsilon_1. \tag{5}$$

First of all we will show that we can approximate the probability in the LHS of (5) by the value $P_e(\mathcal{C}(Q^n), Q^n)$. We have

$$Tr(\rho_{c_i^n}^{T^n} I_i^{Q^n}) = Tr(\rho_{c_i^n}^{Q^n} I_i^{Q^n}) - Tr((\rho_{c_i^n}^{Q^n} - \rho_{c_i^n}^{T^n}) I_i^{Q^n})$$

and

$$\left| Tr((\rho_{c_i^n}^{Q^n} - \rho_{c_i^n}^{T^n}) I_i^{Q^n}) \right| \leq Tr(I_i^{Q^n}) ||\rho_{c_i^n}^{Q^n} - \rho_{c_i^n}^{T^n}|| \leq qn^2 |\mathcal{S}|^n \epsilon.$$

Here we use the inequalities

$$\begin{aligned} Tr(I_i^{Q^n}) &\leq qn, \\ ||\rho c_i^{n Q^n} - \rho_{c_i^n}^{T^n}|| &\leq n|\mathcal{S}|^n \max_{s\in\mathcal{S}}, i = 1, \ldots, n |T_i(s) - Q_i(s)|. \end{aligned}$$

Thus we have

$$P_e(\mathcal{C}, Q^n, T^n) \leq \pi_N^{Q^n} + \gamma_{n,N} \tag{6}$$

where

$$\pi_N^{Q^n} = 1 - \frac{1}{N} \sum_{i=1}^{N} Tr(I_i^{Q^n} \rho_{c_i^n}^{Q^n}) \tag{7}$$

$$\gamma_{n,M} \overset{M\to\infty}{\to} 0.$$

Next we use the considerations of the work [8], which we do not repeat here, and which allow us to come from the formula (7) to the estimation

$$\begin{aligned} \pi_N^{Q^n} \leq \frac{1}{N} \sum_{i=1}^{N} \Bigg(& 3Tr(\rho_{c_i^n}^{Q^n} (I - \Pi^{P,Q^n})) \\ & + \sum_{j\neq i} Tr \Pi^{P,Q^n} \rho_{c_i^n}^{Q^n} \Pi^{P,Q^n} \Pi_{c_j^n}^{Q^n} + Tr(\rho_{c_i^n}^{Q^n} (I - \Pi_{c_i^n}^{Q^n})) \Bigg). \end{aligned} \tag{8}$$

Next we use the random coding argument which consists in choosing the code strings $c_i^n = (c_{i,1}, \ldots, c_{i,n})$ independently with probability $P(c_i^n) = P(c_{i,1}) \ldots P(c_{i,n})$. Taking the average over both sides of (8) we obtain

$$\begin{aligned}\mathbb{E}\pi_N^{Q^n} &\leq 3Tr\rho^{P,Q^n}(I - \Pi^{P,Q^n}) + NTr\Pi^{P,Q^n}\rho^{P,Q^n}\Pi^{P,Q^n}\mathbb{E}\Pi_{c_j^n}^{Q^n} + \mathbb{E}Tr\rho_{c_i^n}^{Q^n}(I - \Pi_{c_i^n}^{Q^n}) \\ &\leq 4\epsilon_2 + N||\Pi^{P,Q^n}\rho^{P,Q^n}\Pi^{P,Q^n}||Tr\mathbb{E}\Pi_{c_i^n}^{Q^n}.\end{aligned}$$

Next we have

$$||\Pi^{P,Q^n}\rho^{P,Q^n}\Pi^{P,Q^n}|| \leq 2^{-\sum_{i=1}^n H(\rho^{P,Q_i})+\delta n}$$

and

$$Tr\mathbb{E}\Pi_{c_i^n}^{Q^n} = \mathbb{E}Tr\Pi_{c_i^n}^{Q^n}) \leq 2^{\sum_{i=1}^n \bar{H}(\rho_\cdot^{Q_i})+\delta n}.$$

From these inequalities it follows that

$$\begin{aligned}\mathbb{E}\pi_N^{Q^n} &\leq 4\epsilon_2 + N2^{2n\delta - \sum_{i=1}^N (H(\rho^{P,Q_i}) - \bar{H}(\rho_\cdot^{Q_i}))} \\ &\leq 4\epsilon_2 + N2^{n\delta - n\min_Q (H(\rho^{P,Q}) - \bar{H}(\rho_\cdot^{Q}))}.\end{aligned}$$

Thus if

$$\frac{\log N}{n} < \min_Q (H(\rho^{P,Q}) - \bar{H}(\rho_\cdot^Q)) + \delta$$

for some $\delta \overset{n\to\infty}{\to} 0$, then for each n and T^n we can choose M and code $\mathcal{C} = \{(c_i^n, I_i^{Q^n}) : \ i = 1, \ldots, N\}$ such that

$$P_e(\mathcal{C}, T^n) \leq 5\epsilon_2 + \gamma_{n,M} \overset{n\to\infty}{\to} 0.$$

We have freedom in the choice of the distribution P. We will choose it to maximize the value

$$H(\rho^{P,Q}) - \bar{H}(\rho_\cdot^Q).$$

What we show is that for each n we have the finite family of sets of distributions $Q^n = (Q_1, \ldots, Q_n)$ on S such that for each set of distributions $T^n = (T_1, \ldots, T_n)$ on $\mathcal{S}$ we have

$$\min_{\mathcal{C}} \sum_{s^n \in \mathcal{S}^n} T^n(s^n) P_e(\mathcal{C}, s^n) < \epsilon, \tag{9}$$

where min is taken over the finite set of codes $\mathcal{C} = \{(c_i^n, I_i^{Q^n}) : \ i = 1, \ldots, N\}$ and $T^n(s^n) = T_1(s_1), \ldots, T_n(s_n)$.

Now we use the key idea of [3], namely, the Mini-Max Theorem for mixed strategies, which are the distributions T^n on the states s^n of the QAVC and the set of distributions Ω on the given finite set of codes $\mathcal{C}$. It says here that

$$\max_{T^n}\min_{G\in\Omega}\sum_{s^n\in\mathcal{S}^n,\ \mathcal{C}} T^n(s^n)G(\mathcal{C})P_e(\mathcal{C},s^n)=\min_{G\in\Omega}\max_{T^n}\sum_{s^n\in\mathcal{S}^n,\ \mathcal{C}} T^n(s^n)G(\mathcal{C})P_e(\mathcal{C},s^n). \tag{10}$$

From (9) and (10) it follows that there exists a distribution $G\in\Omega$ such that

$$\max_{s^n\in\mathcal{S}^n}\sum_{\mathcal{C}} P_e(\mathcal{C},s^n)G(\mathcal{C})<\epsilon.$$

In other words we show that for the distribution $G(\mathcal{C})$ on the finite set of codes (random code) the average (over this distribution) probability of the decoding error can be made arbitrary small uniformly in the choice of the state of QAVC. This proves the direct part of Lemma 1 (that $\tilde{C}\geq\bar{C}$).

However this scheme has the disadvantage that the sender and the receiver should know the code which is used in the particular transmission. Later we will show how to eliminate this difficulty.

Next we show, that

$$\tilde{C}\leq\bar{C}. \tag{11}$$

Consider another quantum channel which has the following density operators: for given n we fix n distributions $Q^n=(Q_1,\ldots,Q_n)$ on $\mathcal{S}$ and to the code string c^n corresponds the density $\rho_{c^n}^{Q^n}=\sum_{s^n\in\mathcal{S}^n}\rho_{c^n,s^n}Q_1(s_1)\ldots Q_n(s_n)$. The set of complex elements of the matrices of the resolutions of the identity which defined the quantum channel is a compact set in the natural space of vectors with complex components. Obviously the value

$$P_e(Q^n)\stackrel{\Delta}{=}\inf_G\int_\Gamma P_e(\mathcal{C}^\gamma,\bar{Q})dG(\gamma)$$

is achieved on the sequence of distributions $\{G_m;\ m=1,\ldots\}$, each of which has support on a finite number of codes. Thus for given n for each $\epsilon>0$ there exists a distribution on the codes G' with mass, concentrated on a finite number of codes $\mathcal{C}^\gamma,\ \gamma\in\tilde{\Gamma}$, such that

$$P_e(Q^n)\geq\sum_{\gamma\in\tilde{\Gamma}}P_e(\mathcal{C}^\gamma,Q^n)G'(\gamma)-\epsilon_1.$$

From this it follows that if we have the bound

$$\sum_{\gamma\in\tilde{\Gamma}} P_e(\mathcal{C}^\gamma, Q^n)G'(\gamma) > C_1, \tag{12}$$

then we have the lower bound

$$P_e(Q^n) > C_1 - \epsilon_1.$$

This for given small ϵ_1 and large $C_1 = const$ implies the impossibility of transmission with arbitrary small probability of decoding error.

Now we prove (12). It is easy to see that the average probability $\tilde{P}_e(Q^n) \stackrel{\Delta}{=} \sum_{\gamma\in\tilde{\Gamma}} P_e(\mathcal{C}^\gamma, Q^n)G'(\gamma)$ of the decoding error in such a channel satisfies the inequality

$$\begin{aligned} \tilde{P}_e(Q^n) &\leq \inf_G \int_\Gamma P_e(\mathcal{C}^\gamma, Q^n)dG(\gamma) + \epsilon_1 \\ &\leq \inf_{\mathcal{C}} \max_{s^n\in\mathcal{S}^n} P_e(\mathcal{C}, s^n) + \epsilon_1. \end{aligned} \tag{13}$$

We denote the capacity under the criterium of the probability of error $P_e(Q^n)$ by $\tilde{C}_a$.

From (13) follows that the upper bound for the capacity $\tilde{C}_a$ is at the same time the upper bound for C_q. Next we find the proper upper bound for $\tilde{C}_a$. We fix n and G', defined before.

Let $W = \{1, \ldots, 2^{nR}\}$ be the set of messages with uniform distribution on it. If W' is the estimation of W, then by the Fano inequality [4]

$$H(W|W') \leq \mathcal{H}(\tilde{P}_e(Q^n)) + \tilde{P}_e(Q^n)\log(N-1),$$

where $\mathcal{H}(x) = -x\log x - (1-x)\log(1-x)$. At the same time

$$H(W|W') = H(W) - I(W;W') = nR - I(W;W').$$

Thus $\tilde{P}_e(Q^n)$ can be made an arbitrary small as $n \to \infty$ if

$$nR < I(W;W'). \tag{14}$$

Since $I(W;W') < \max_{\mathcal{C}\in supp(G')} I_{\mathcal{C}}(X;Y)$, where $I_{\mathcal{C}}(X;Y)$ is the mutual information between the input and output (for given code $\mathcal{C}$ from the support G') we make the condition (14) weaker, if we allow the following inequality

$$nR < \max_{\mathcal{C}\in supp(G')} I_{\mathcal{C}}(X;Y). \tag{15}$$

For given set $Q^n = (Q_1, \dots, Q_n)$ of n distributions on S we choose the set $Q^n(P^n) = (Q_1(P_1), \dots, Q_n(P_n))$ of n distributions on $\mathcal{S}$ in such a way that

$$Q_j(P_j) = \arg\min_Q H(P_j, Q).$$

If $\mathcal{C} = \{(c_i^n, I_i) : \ i = 1, \dots, N\}$, then, as it was shown in [7],

$$\sup_{\{\{I_i\}\}} I_{\mathcal{C}}(X;Y) \le H(P^n, Q^n(P^n)).$$

Here sup in the LHS can be taken over all resolutions of identity for given set of codewords from $\mathcal{C}$, P^n in the RHS is the distribution on $\mathcal{X}^n$, generated by the uniform distribution on this set of codewords and $Q^n(P^n)(s^n) = \prod_{i=1}^n Q_j(P_j)(s_j)$. Next, as it was shown in [7],

$$H(P^n(Q^n), Q^n) \le \sum_{j=1}^n H(P_j, Q_j(P_j)).$$

Furthermore we have

$$\begin{aligned} I_{\mathcal{C}}(X;Y) &\le \max_{P^n} \sup_{\{\{I_i\}\}} I_{\mathcal{C}}(X;Y) \le \max_{P^n} H(P^n, Q^n(P^n)) \\ &\le n \max_P H(P, Q(P)) = n \max_P \min_Q H(P, Q). \end{aligned}$$

This together with (15) completes the proof of Lemma 1 .

Now we will prove Lemma 2. Assume that $C_q \neq 0$, then $\bar{C} \neq 0$. We have already shown that $C_q \le \bar{C}$. For given $\epsilon > 0$, there exists a random code $\big(\{\mathcal{C}^\gamma : \gamma \in \Gamma\}, G\big)$

$$\mathcal{C} = \{(c_i^\gamma, I_i^\gamma) : \ \gamma \in \bar{\Gamma}\}$$

with finite $\bar{\Gamma}$ such that

$$\max_{s^n \in \mathcal{S}^n} \sum_{\gamma \in \bar{\Gamma}} P_e(\mathcal{C}^\gamma, s^n) G(\gamma) < \epsilon$$

and $\log N/n > \bar{C} - \delta$. Next we consider n^2 i.i.d. random variables $Z_1, \dots, Z_{n^2}$ with values in $\bar{\Gamma}$ such that $P(Z_i = \gamma) = G(\gamma)$. Then for given $s^n \in \mathcal{S}^n$, $\pi_j = P_e(Z_j, s^n)$ are also i.i.d. random variables and

$$\mathbb{E} P_e(Z_j, s^n) < \epsilon.$$

Here $\mathbb{E}(\cdot)$ is the average in the current sense and does not coincide with the concept in the same notation before. By Chebyshev's inequality

$$\pi^n \stackrel{\Delta}{=} G\left(\sum_{j=1}^{n^2}\pi_j > \lambda n\right) < e^{-\alpha\lambda n^2}\prod_{j=1}^{n^n}\mathbb{E}e^{\alpha\pi_j},\ \alpha > 0.$$

But because $0 \le \pi_j^k \le \pi_j \le 1,\ k \ge 0$, we have

$$\mathbb{E}e^{\alpha\pi_j} \le 1 + \sum_{k=1}^{\infty}\frac{\alpha^k}{k!}\mathbb{E}\pi_j \le 1 + \epsilon\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}.$$

Thus if $\alpha = 2$, and $\epsilon > 0$ is sufficiently small, then

$$\begin{aligned}\pi^n &< e^{-\alpha\lambda n^2}\left(1+\epsilon e^{\alpha}\right)^{n^2} \\ &< e^{-2\lambda n^2}e^{\lambda n^2} = e^{-\lambda n^2}.\end{aligned} \tag{16}$$

Since π^n decreases faster than exponentially with n we conclude that $\pi^n \stackrel{n\to\infty}{\to} 0$ uniformly with the choice of $s^n \in \mathcal{S}^n$.

Since $C_q > 0$, there exists the code $\mathcal{C} = \{(c_i^n, I_i),\ i = 1,\ldots,n^2\}$ of length $\nu(n) = o(n)$ with probability of decoding error $\max_{s^n\in\mathcal{S}^n} P_e(\mathcal{C}, s^n) < \lambda$. Now we construct a new code of length $\nu(n) + n$ with Nn^2 codewords. This code is $\mathcal{D} = \{(c_\gamma^n c_i^\gamma, I_\gamma \bigotimes I_i^\gamma),\ i = 1,\ldots,N,\ \gamma = 1,\ldots,n^2\}$.

Now we use one simple fact, that if $\alpha = (\alpha_1,\ldots,\alpha_R)$ and $\beta = (\beta_1,\ldots,\beta_R)$ are two sequences, $\alpha_i,\ \beta_i \in [0,1]$ such that

$$\frac{1}{R}\sum_{i=1}^{R}\alpha_i \ge 1-\epsilon,\ \frac{1}{R}\sum_{i=1}^{R}\beta_i \ge 1-\epsilon, \epsilon\in(0,1),$$

then

$$\frac{1}{R}\sum_{i=1}^{R}\alpha_i\beta_i \ge 1-2\epsilon. \tag{17}$$

Now we have ($s^n = s_1^n s_2^n$)

$$\begin{aligned}&\frac{1}{n^2}\frac{1}{N}\sum_{\gamma=1}^{n^2}\sum_{i=1}^{N}Tr(\rho_{c_\gamma^n,s_1^n}\bigotimes\rho_{c_i^\gamma,s_2^n}I_\gamma\bigotimes I_i^\gamma) \\ = \ &\frac{1}{n^2}\frac{1}{N}\sum_{\gamma=1}^{n^2}\sum_{i=1}^{N}Tr(\rho_{c_\gamma^n,s_1^n}I_\gamma)Tr(\rho_{c_i^\gamma,s_2^n}I_i^\gamma).\end{aligned}$$

Denoting

$$\alpha_i = Tr(\rho_{c_\gamma^n, s_1^n} I_\gamma),\ \beta_i = \frac{1}{N} Tr(\rho_{c_i^\gamma, s_2^n} I_i^\gamma)$$

we see from (17) that

$$\frac{1}{n^2} \sum_{i=1}^{n^2} \alpha_i \beta_i \geq 1 - 2\lambda.$$

Thus the concatenated code $\mathcal{D}$ has the probability of decoding error less than 2λ and rate $R = \frac{\log N^2 N}{f(n)+n} \geq \frac{\log N}{n}(1-\delta)$. This completes the proof of Lemma 2.

The next Statement 2 we will use to prove that, if QAVC is not symmetrizable, i.e. if for each distribution $\mathcal{U} = \{U(s|x),\ x \in \mathcal{X}\}$ on $\mathcal{S}$ for some $x, x' \in \mathcal{X}$

$$\sum_{s \in \mathcal{S}} U(s|x')\rho_{x,s} \neq \sum_{s \in \mathcal{S}} U(s|x)\rho_{x',s}, \tag{18}$$

then $C_q > 0$.

Statement 2 *If for each set of distributions $\mathcal{U}$ for some $x, x' \in \mathcal{X}$ the relation (18) is valid, then there exists the resolution of identity $\{L_i,\ i = 1, \dots d^2+1\}$ in $\mathbb{C}^d$ such that for each set $\mathcal{U}$ for some $x, x' \in X$ and $i \in \{1, \dots, d^2\}$ the following relation is valid*

$$\sum_{s \in \mathcal{S}} U(s|x')Tr(\rho_{x,s}, L_i) \neq \sum_{s \in \mathcal{S}} U(s|x)Tr(\rho_{x',s}, L_i). \tag{19}$$

The proof of this statement is quite simple. We choose the set of d^2 linear independent nonegative operators $L_i,\ i = 1, \dots, d^2$ from the cone of nonegative operators in $\mathbb{C}^d$ and norm them in such a way that

$$\sum_{i=1}^{d^2} L_i \leq I.$$

Then $\{L_i,\ I - \sum_{i=1}^{d^2} L_i;\ i = 1, \dots d^2\}$ is the resolution of the identity. Because $\{L_i,\ i = 1, \dots, d^2\}$ is the basis and (18) is valid we have for some i relation (19). This proves Statement 2.

Now we consider the classical AVC with $Y = \{1, \dots, d^2\}$ and transition probabilities

$$w(i|x, s) = Tr(\rho_{x,s}, L_i).$$

From the Statement 2 it follows that if the initial QAVC is not symmetrizable, then such a related AVC is not symmetrizable in the classical sense and from [5] it follows that the capacity C_a of this AVC is positive. On the other hand obviously $C_q \geq C_a$ and hence $C_q > 0$. Note, that from Lemma 2 in this case $C_q = \bar{C}$. This complete the proof of Theorem 1.

Final Remarks

It is interesting to note, that in our proof of Theorem 1 we essentially use the elimination technique (an early candidate of what is now called derandomization in Computer Sciences) from [1], which gives Lemma 2. This is the analogue of the main result of [1]. There a necessary and sufficient condition for positivity of the capacity was given, if the set of transmission matrices is row-convex closed– that is under a practically satisfactory assumption of robustness. The mathematical problem of characterizing positivity without this assumption in terms of symmetrizability was started in [6] and completely solved in [5] with a non-standard decoding rule and without use of the elimination technique of [1]. (Using this technique and proving directly that non-symmetrizability implies positive capacity is a basic problem, which is open for more than 20 years!)

On the other hand in the present quantum case we have not found a suitable decoding rule and follow the elemination technique (Lemma 2). Analogously the positivity problem for the QAVC can be settled by reducing it to a related classical AVC to which then the result of [5] can be applied.

We emphasize that the very hard maximal error capacity problem for AVC (including Shannon's zero error capacity problem as special case) is based on a more realistic communication model. It was solved for a nice class of channels in [2], where for the first time in the area of AVC a non standard decoding rule was used. Extension to QAVC constitutes a challenging problem!

References

[1] R. Ahlswede, Elimination of correlation in random codes for arbitrarily varying channels, Z. Wahrscheinlichkeitstheorie u. verw. Gebiete, 44, 159-175, 1978.

[2] R. Ahlswede, A method of coding and its application to arbitrarily varying channels, J. Combinatorics, Information and System Sciences, Vol. 5, No. 1, 10-35, 1980.

[3] D. Blackwell, L. Breiman and A.J. Thomaisan, The capacities of certain channel classes under random coding, The Ann. of Math., Statistics, Vol.31, No.3, 558-567, 1960.

[4] T. Cover and J. Thomas, Elements of Information Theory, John Wiley & Sons Inc., N.Y., 1991.

[5] I. Csiszár and P. Narayan, The capacity of the arbitrarily varying channel revisited: positivity constraints, IEEE Transactions of IT, Vol. 34, No, 2, 181-193, 1988.

[6] T. Ericson, Exponential error bounds for random codes in the arbitrarily varying channel, IEEE Transactions on IT, Vol. 31, No. 1, 42-48, 1985.

[7] A.S. Holevo, Bounds for the quantity of information transmitted by a quantum communication channel, English: Problems of Information Transmission, Vol. 9, No. 3, 177-184, 1973.

[8] A.S. Holevo, The capacity of the quantum channel with general signal states, IEEE Transactions of IT, Vol. 44, No. 1, 269-273, 1998.