The Capacity of the Quantum Multiple Access Channel

Andreas Winter*
SFB 343, Fakultät für Mathematik
Universität Bielefeld, Postfach 100131, 33501 Bielefeld
(July 6, 1998)

We define quantum multiway channels for transmission of classical information, after recent work by Allahverdyan and Saakian. Bounds on the capacity region are derived in a uniform way, which are analogous to the classically known ones. For the single receiver case the exact capacity region is determined.
Finally we comment on the case of noisy quantum channels with arbitrary input states.

I. INTRODUCTION

The multiway channel with s senders and r receivers in classical information theory was already studied by Shannon [1], Ahlswede [2,3] first determined its capacity region. For a good overview on multiuser communication theory consult [4]. In the present paper we will define the corresponding quantum channel (after recent work by Allahverdyan and Saakian [5]), extending Holevo's definition of a classical–quantum channel in [6]. We will bound the capacity region, the actual bounds being obtainable from the classical case by formally replacing Shannon entropy by von Neumann entropy in the expressions, thus following a general principle or feeling in physics. We will also prove the direct coding theorem for the multiple access channel ($r = 1$), using the result of Holevo [8], and Schumacher and Westmoreland [9] for the single–sender case, (after Hausladen et.al. [7], who presented a quantum version of the classical random coding method).

The outline of the paper is as follows: section II recalls the necessary definitions from [10], in particular quantum multiway channels. In the following section III we essentially give the upper bounds for the capacity region of [10]. Sections IV and V are taken from [11] and are included for the sake of self–containedness. In section VI comes the main contribution: the direct coding theorem for the quantum multiple access channel. The last section VII comments on the quantum–quantum multiway channel which may be fed with arbitrary input states.

Notational issues: Finite sets will be denoted $\mathcal{A}, \mathcal{B}, \ldots$, C*–algebras $\mathfrak{X}, \mathfrak{Y}, \ldots$. Hilbert spaces will be denoted $\mathcal{H}, \mathcal{H}_1, \mathcal{H}_2, \ldots$, density operators (states) by $\rho, \sigma, \ldots$. Probability distributions will be written $P, Q, \ldots$, and stochastic matrices (classical channels) $V, W, \ldots$ (and also classical quantum channels), general quantum operations will be written $\varphi_*, \psi_*, \ldots$.
H is the usual Shannon entropy, as well as the von Neumann entropy for density operators ρ: $H(\rho) = -\mathrm{Tr}\,(\rho \log \rho)$. The functions log and exp are to basis 2.

II. QUANTUM MULTIWAY CHANNELS

The following we recall from [10], using also the notation introduced there. This is the simplest situation of multi–user communication in general: consider s independent senders, sender i using an alphabet $\mathcal{X}_i$, say with an a priori probability distribution P_i. We describe this by the quantum state $\sigma_i = \sum_{x_i \in \mathcal{X}_i} P_i(x_i) x_i$ on the commuative C*–algebra $\mathfrak{X}_i = \mathbb{C}\mathcal{X}_i$ generated by the x_i which are mutually orthogonal idempotents (to distinguish these as generators of this algebra we will sometimes write $[x_i]$). The channel is then a map

$$W : \mathcal{X}_1 \times \cdots \times \mathcal{X}_s \to \mathfrak{S}(\mathfrak{Y})$$

with any (finite dimensional) C*–algebra $\mathfrak{Y}$, which connects the input $(x_1, \ldots, x_s)$ with the output $W_{x_1 \ldots x_s}$. By linear extension we may view W as a completely positive, trace preserving map from $\mathfrak{T}(\mathfrak{X}_1 \otimes \cdots \otimes \mathfrak{X}_s)$ to $\mathfrak{T}(\mathfrak{Y})$. The receivers

*Electronic address: winter@mathematik.uni-bielefeld.de

are modelled by compatibel $*$–subalgebras $\mathfrak{Y}_j$. If all the $W_{x_1\dots x_s}$ commute with each other (hence have a common diagonalization) the channel is called *classical.*
For fixed a priori distributions we have the *channel state*

$$\gamma = \sum_{\forall i\ x_i \in \mathcal{X}_i} P_1(x_1)\cdots P_s(x_s)[x_1]\otimes\cdots\otimes[x_s]\otimes W_{x_1\dots x_s}$$

on $\mathfrak{X}_1\otimes\cdots\otimes\mathfrak{X}_s\otimes\mathfrak{Y}$ (this serves as the quantum analogue of the joint distribution of the random variables representing input and output letters in the classical case. Note that we are dealing only with memoryless channels).
For a $J\subset[s]$ denote $P_J = \bigotimes_{i\in J} P_i$, i.e. $P_J(x_i|i\in J) = \prod_{i\in J} P_i(x_i)$, and $\mathcal{X}(J) = \prod_{i\in J}\mathcal{X}_i$ (similarly $\mathfrak{X}(J) = \bigotimes_{i\in J}\mathfrak{X}_i$).
Further define a *reduced channel* $P_{J^c}W : \mathcal{X}(J)\to\mathfrak{S}(\mathfrak{Y})$ by

$$P_{J^c}W : (x_i|i\in J)\longmapsto \sum_{\forall i\in J^c:\ x_i\in\mathcal{X}_i} P_{J^c}(x_i|i\in J^c)W_{x_1\dots x_s}$$

Note that

$$\mathrm{Tr}_{\,\mathfrak{X}(J^c)}\gamma = \sum_{\forall i\in J:\ x_i\in\mathcal{X}_i} [x_i|i\in J]\otimes(P_{J^c}W)_{(x_i|i\in J)}$$

An *n–block–code* is a collection $(f_1,\dots,f_s)$ of maps $f_i:\mathcal{M}_i\to\mathcal{X}_i^n$, together with a collection of decoding observables $(D_1,\dots,D_r)$ with $D_j\subset\mathfrak{Y}_j^{\otimes n}$, indexed by $\mathcal{M}'_1\times\cdots\times\mathcal{M}'_s\supset\mathcal{M}_1\times\cdots\times\mathcal{M}_s$. There are r (average) *error probabilities* of the code, the probability that the receiver j guesses wrongly any one of the sent words, taken over the uniform distribution on the codebooks:

$$\bar{e}_j(f_1,\dots,f_s,D_j) = 1-\frac{1}{|\mathcal{M}_1|\cdots|\mathcal{M}_s|}\sum_{\forall i: m_i\in\mathcal{M}_i}\mathrm{Tr}\left(W^{\otimes n}(f(m_1),\dots,f(m_s))D_{j,m_1\dots m_s}\right)$$

We call $(f_1,\dots,f_s,D_1,\dots,D_r)$ an $(n,\bar\epsilon)$–code if all $\bar{e}_j(f_1,\dots,f_s,D_j)$ are at most $\bar\epsilon$.
The *rates* of the code are the $R_i=\frac{1}{n}\log|\mathcal{M}_i|$. A tuple $(R_1,\dots,R_s)$ is said to be *achievable*, if for any $\bar\epsilon,\delta>0$ there exist for any large enough n a $(n,\bar\epsilon)$–code with i–th rate at least $R_i-\delta$. The set of all achievable tuples (which is clearly closed) is called the *capacity region* of the channel.
From [10] we use the definitions of various information quantities for observables and $*$–subalgebras. With the definitions above we have the (easily checked) identities:

$$\begin{aligned} I(\mathfrak{X}(J)\wedge\mathfrak{Y}|\mathfrak{X}(J^c)) &= I(\mathfrak{X}(J)\wedge\mathfrak{Y}\mathfrak{X}(J^c)) \\ &= H(\mathfrak{Y}|\mathfrak{X}(J^c)) - H(\mathfrak{Y}|\mathfrak{X}_1\cdots\mathfrak{X}_s) \\ &= H(P_JW|P_{J^c}) - H(W|P_{[s]}) \end{aligned}$$

III. UPPER CAPACITY BOUNDS

We recall from [10] the following theorem (which we call the weak converse because of theorem VI.1 and note 2):

Theorem 1 (Weak converse) *The capacity region of the quantum multiway channel is contained in the closure of all nonnegative* $(R_1,\dots,R_s)$ *satisfying*

$$\forall J\subset[s], j\in[r] \qquad R(J)=\sum_{i\in J}R_i\le\sum_u q_u I_{\gamma_u}\left(\mathfrak{X}(J)\wedge\mathfrak{Y}_j|\mathfrak{X}(J^c)\right)$$

for some channel states γ_u *(belonging to appropriate input distributions) and* $q_u\ge 0$, $\sum_u q_u = 1$.

Proof. Consider any $(n,\bar\epsilon)$–code $(f_1,\dots,f_s,D_1,\dots,D_r)$ with rate tuple $(R_1,\dots,R_s)$. Then the uniform distribution on the codewords induces a channel state γ on $(\mathfrak{X}_1\cdots\mathfrak{X}_s\mathfrak{Y})^{\otimes n}$. Its restriction to the u–th copy in this tensor power will be denoted γ_u. Let $j\in[r]$, $J\subset[s]$. By Fano inequality in the form of [10] we have

$$H(\mathfrak{X}^{\otimes n}(J)|\mathfrak{Y}_j^{\otimes n}\mathfrak{X}^{\otimes n}(J^c))\le 1+\bar\epsilon\cdot nR(J)$$

With

$$\begin{aligned} H(\mathfrak{X}^{\otimes n}(J)|\mathfrak{Y}_j^{\otimes n}\mathfrak{X}^{\otimes n}(J^c)) &= H(\mathfrak{X}^{\otimes n}(J)) - I(\mathfrak{X}^{\otimes n}(J)\wedge\mathfrak{Y}_j^{\otimes n}\mathfrak{X}^{\otimes n}(J^c)) \\ &= nR(J) - I(\mathfrak{X}^{\otimes n}(J)\wedge\mathfrak{Y}_j^{\otimes n}\mathfrak{X}^{\otimes n}(J^c)) \end{aligned}$$

we conclude (with subadditivity of mutual information, see [10])

$$\begin{aligned}(1-\bar{\epsilon})R(J) &\leq \frac{1}{n}+\frac{1}{n}I_\gamma(\mathfrak{X}^{\otimes n}(J)\wedge\mathfrak{Y}_j^{\otimes n}\mathfrak{X}^{\otimes n}(J^c)) \\ &\leq \frac{1}{n}+\frac{1}{n}\sum_{u=1}^{n}I_{\gamma_u}(\mathfrak{X}(J)\wedge\mathfrak{Y}_j\mathfrak{X}(J^c))\end{aligned}$$

□

Note 2 *In the case of classical channels the region described in the theorem is the exact capacity region (i.e. all the rates there are achievable), as was first proved by Ahlswede [2,3].*

Note 3 *The numeric computation of the above regions is not yet possible from the given description: we need a bound on the number of different single–letter channel states one has to consider in the convex combinations. For the multiple access channel ($r=1$) this is easy: by Caratheodory's theorem s will suffice. For general r there are also classical bounds, which carry over unchanged to the quantum case (since the quantum mutual information has properties similar to those of classical mutual information): $r(2^s-1)$ always suffice, cf. [12].*

IV. DISTORTION AND FIDELITY

This section compares the distortion measure D from $||\cdot||_1$ (the trace norm)

$$D(\rho,\sigma)=\frac{1}{2}||\rho-\sigma||_1$$

on the states of a quantum system with the (pure state) fidelity F. First note that $||\cdot||_1$ is indeed a distortion measure:

$$||\rho_1\otimes\rho_2-\sigma_1\otimes\sigma_2||_1\leq||\rho_1-\sigma_1||_1+||\rho_2-\sigma_2||_1$$

In contrast, the fidelity has a multiplicative law under tensor products, so we might consider the log of it:

$$-\log F(\rho_1\otimes\rho_2,\sigma_1\otimes\sigma_2)=-\log F(\rho_1,\sigma_1)-\log F(\rho_2,\sigma_2)$$

But since $-\log F$ is unbounded as $F\to 0$ we prefer to use $1-F$. Obviously

$$1-F(\rho_1\otimes\rho_2,\sigma_1\otimes\sigma_2)\leq 1-F(\rho_1,\sigma_1)+1-F(\rho_2,\sigma_2)$$

so $1-F$ is also a distortion measure.

Lemma 1 (Pure state) *Let $\rho=|\psi\rangle\langle\psi|$ and $\sigma=|\phi\rangle\langle\phi|$ be pure states. Then*

$$1-F(\rho,\sigma)=D(\rho,\sigma)^2$$

Proof. W.l.o.g. we may assume $|\psi\rangle=\alpha|0\rangle+\beta|1\rangle$ and $|\phi\rangle=\alpha|0\rangle-\beta|1\rangle$ ($|\alpha|^2+|\beta|^2=1$). A straightforward calculation shows $F=(|\alpha|^2-|\beta|^2)^2$, and $D=2|\alpha\beta|$. Now

$$\begin{aligned}1-F&=1-(|\alpha|^2-|\beta|^2)^2=(1+|\alpha|^2-|\beta|^2)(1-|\alpha|^2+|\beta|^2)\\&=4|\alpha\beta|^2=D^2\end{aligned}$$

□

Lemma 2 (Mixed state) *Let σ an arbitrary mixed state (and ρ pure as above). Then*

$$D\geq 1-F\geq D^2$$

Proof. Write $\sigma=\sum_j q_j|j\rangle\langle j|$. Then

$$\begin{aligned}1-F(\rho,\sigma)&=\sum_j q_j\left(1-F(\rho,|j\rangle\langle j|)\right)=\sum_j q_jD(\rho,|j\rangle\langle j|)^2\\&\geq\left(\sum_j q_jD(\rho,|j\rangle\langle j|)\right)^2\geq D(\rho,\sigma)^2\end{aligned}$$

Conversely: extend $|\psi\rangle$ to an orthonormal basis $(|r\rangle)_r$ ($|0\rangle = |\psi\rangle$), and consider the trace preserving quantum operation

$$\kappa_* : \sigma \longmapsto \sum_{r\geq 0} |r\rangle\langle r|\sigma|r\rangle\langle r|$$

Then with the monotonicity of the trace norm:

$$2D = ||\rho - \sigma||_1 \geq ||\kappa_*\rho - \kappa_*\sigma||_1 = ||\rho - \kappa_*\sigma||_1$$

(since $\rho = \kappa_*\rho$). Hence with $F = \langle\psi|\sigma|\psi\rangle$

$$\begin{aligned} 2D &\geq ||(1-F)|\psi\rangle\langle\psi| - \sum_{r\geq 1}\langle r|\sigma|r\rangle \cdot |r\rangle\langle r|||_1 \\ &= (1-F) + \sum_{r\geq 1}\langle r|\sigma|r\rangle \\ &= 2(1-F) \end{aligned}$$

□

Observe that the inequalities of this lemma still hold if only $\sum_j q_j \leq 1$.

Lemma 3 (Ensembles) *Let ρ_i pure states, σ_i arbitrary states, and p_i probabilities. Then the average fidelity $\bar{F} = \sum_i p_i F(\rho_i, \sigma_i)$ and the average distortion $\bar{D} = \sum_i p_i D(\rho_i, \sigma_i)$ satisfy*

$$\bar{D} \geq 1 - \bar{F} \geq \bar{D}^2$$

Proof. From lemma 2 the inequality $\bar{D} \geq 1 - \bar{F}$ is immediate, and also the inequality $1 - \bar{F} \geq \bar{D}^2$ (with the convexity of the square function). □

As a consequence, the average (pure state) fidelity is close to one, if and only if the average trace norm distortion is small. In particular the conditions that $1 - \bar{F}$, $\bar{D}$ tend to zero, resp., are equivalent. Thus if comparing pure states with mixed states we may freely use either concept, but we will prefer D, as it easily generalizes to mixed states (whereas F is more involved, see e.g. [13]), and has some nice properties. E.g. it does not increase under trace preserving quantum operations, as we already used above.

V. DECODING AS A QUANTUM OPERATION

In this section we will show that good measurements disturbe their ensemble little. Recall from [10] the definition of D_{tot} and D_{int}.

Lemma 1 *Let ρ be a state, and X a positive operator with $X \leq \mathbb{1}$ and $1 - \mathrm{Tr}(\rho X) \leq \epsilon < 1$. Then*

$$||\rho - \sqrt{X}\rho\sqrt{X}||_1 \leq \sqrt{8\epsilon}$$

Proof. Diagonalize ρ: $\rho = \sum_k p_k \pi_k$ with mutually orthogonal one–dimensional projectors π_k. Now with $Y = \sqrt{X}$

$$\begin{aligned} ||\rho - Y\rho Y||_1^2 &\leq \left(\sum_k p_k ||\pi_k - Y\pi_k Y||_1\right)^2 \\ &\leq \sum_k p_k ||\pi_k - Y\pi_k Y||_1^2 \\ &\leq 4\sum_k p_k (1 - \mathrm{Tr}(\pi_k Y \pi_k Y)) \\ &\leq 8\sum_k p_k (1 - \mathrm{Tr}(\pi_k Y)) \\ &= 8(1 - \mathrm{Tr}(\rho Y)) \\ &\leq 8(1 - \mathrm{Tr}(\rho X)) = 8\epsilon \end{aligned}$$

by triangle inequality, convexity of $x \mapsto x^2$, lemma IV.2, $1 - x^2 \leq 2(1-x)$, and $X \leq Y$. □

Lemma 2 (Tender measurement) *Let ρ_a ($a \in \mathcal{A}$) be a set of states on $\mathfrak{A} = \mathfrak{L}(\mathcal{H})$, and D an observable indexed by $\mathcal{B}$. Let further $\varphi : \mathcal{A} \longrightarrow \mathcal{B}$ be any map and $\epsilon > 0$ such that*

$$\forall a \in \mathcal{A} \qquad 1 - \mathrm{Tr}\,(\rho_a D_{\varphi(a)}) \leq \epsilon$$

(I.e. the observable recognizes $\varphi(a)$ from ρ_a with maximal error probability ϵ). Then the quantum operation $D_{\mathrm{int}} : \mathfrak{A} \to \mathfrak{A}$ with adjoint

$$D_{\mathrm{int}*} : \rho \longmapsto \sum_{b \in \mathcal{B}} \sqrt{D_b} \rho \sqrt{D_b}$$

disturbes the states ρ_a only a little:

$$\forall a \in \mathcal{A} \qquad \|\rho_a - D_{\mathrm{int}*}\rho_a\|_1 \leq \sqrt{8\epsilon} + \epsilon$$

Further the quantum operation $D_{\mathrm{tot}} : \mathbb{C}\mathcal{B} \otimes \mathfrak{A} \to \mathfrak{A}$ with adjoint

$$D_{\mathrm{tot}*} : \rho \longmapsto \sum_{b \in \mathcal{B}} [b] \otimes \sqrt{D_b} \rho \sqrt{D_b}$$

has the property that

$$\forall a \in \mathcal{A} \qquad \|[\varphi(a)] \otimes \rho_a - D_{\mathrm{tot}*}\rho_a\|_1 \leq \sqrt{8\epsilon} + \epsilon$$

Proof. It suffices to prove the second statement since the first inequality is obtained from it by a partial trace which does not increase $\|\cdot\|_1$. This is an easy calculation:

$$\begin{aligned}
\|[\varphi(a)] \otimes \rho_a - D_{\mathrm{tot}*}\rho_a\|_1 &\leq \|\rho_a - \sqrt{D_{\varphi(a)}}\rho_a\sqrt{D_{\varphi(a)}}\|_1 + \sum_{b \neq \varphi(a)} \|\sqrt{D_b}\rho_a\sqrt{D_b}\|_1 \\
&= \|\rho_a - \sqrt{D_{\varphi(a)}}\rho_a\sqrt{D_{\varphi(a)}}\|_1 + \sum_{b \neq \varphi(a)} \mathrm{Tr}\,(\rho_a D_b) \\
&\leq \sqrt{8(1 - \mathrm{Tr}\,(\rho_a D_{\varphi(a)}))} + (1 - \mathrm{Tr}\,(\rho_a D_{\varphi(a)})) \\
&\leq \sqrt{8\epsilon} + \epsilon
\end{aligned}$$

using triangle inequality and lemma 1. □

Note 3 *If we modify the statement of the lemma to that the average error in recognizing $\varphi(a)$ from ρ_a should be at most $\bar{\epsilon}$ (relative a distribution on $\mathcal{A}$), then also the distortion bound of the lemmas hold — on average (see lemma IV.3).*

VI. QUANTUM MULTIPLE ACCESS CHANNEL

Throughout this section we will assume $r = 1$ and else notation as before.

Theorem 1 *Let $R_1, \ldots, R_s$ be nonnegative, satisfying for some a priori distributions P_i on the $\mathcal{X}_i$ the constraints*

$$\forall J \subset [s] \qquad \sum_{i \in J} R_i \leq I\left(\mathfrak{X}(J) \wedge \mathfrak{Y} | \mathfrak{X}(J^c)\right)$$

Then for every $\bar{\epsilon}, \delta > 0$ and all sufficiently large n there are $(n, \bar{\epsilon})$–codes with rates $\dfrac{1}{n} \log |\mathcal{M}_i| \geq R_i - \delta$.

Proof. It is sufficient by the time sharing principle (see e.g. [12]) to prove the assertion only for the upper extremal points of the region described, and by symmetry we may assume (for $i = 1, \ldots, s$) that

$$\begin{aligned}
R_i &= I(\mathfrak{X}_i \wedge \mathfrak{Y}\mathfrak{X}_1 \cdots \mathfrak{X}_{i-1}) \\
&= H(\mathfrak{Y}|\mathfrak{X}_1 \cdots \mathfrak{X}_{i-1}) - H(\mathfrak{Y}|\mathfrak{X}_1 \cdots \mathfrak{X}_i) \\
&= H(P_{\{>i-1\}}W|P_{\{\leq i-1\}}) - H(P_{\{>i\}}W|P_{\{\leq i\}})
\end{aligned}$$

(that these are indeed the upper extremal points will be proved in appendix A).
Let $\epsilon, \delta > 0$, and consider s codebooks $\mathcal{C}_i$ of size $L_i = |\mathcal{C}_i| = 2^{\lceil n(R_i-\delta)\rceil}$ drawn independently from $\mathcal{X}_i^n$ according to the a priori distribution $P_i^{\otimes n}$.
Now fix i and define the following channel: for $\mathbf{x}_i \in \mathcal{X}_i^n$

$$\rho:\ \mathbf{x}_i \longmapsto \rho^{(i)}_{\mathbf{x}_i} = \frac{1}{L_1\cdots L_{i-1}} \sum_{\forall j<i:\ \mathbf{c}_j\in\mathcal{C}_j} \left(\mathbf{c}_1 \otimes \cdots \otimes \mathbf{c}_{i-1} \otimes \frac{1}{L_{i+1}\cdots L_s} \sum_{\forall j>i:\ \mathbf{c}_j\in\mathcal{C}_j} W^n_{\mathbf{c}_1\ldots\mathbf{x}_i\ldots\mathbf{c}_s} \right)$$

(we denote these word states with ρ, in contrast to the letter states W).
The code average of $\rho^{(i)}_{\mathbf{x}_i}$ over choice of $\mathcal{C}_1, \ldots, \mathcal{C}_{i-1}, \mathcal{C}_{i+1}, \ldots, \mathcal{C}_s$ is indeed a product state

$$\begin{aligned} V^n_{\mathbf{x}_i} &= \sum_{\forall j\neq i:\ \mathbf{x}_j\in\mathcal{X}_j^n} P_{i^c}(\mathbf{x}_j|j\neq i)[\mathbf{x}_1\ldots\mathbf{x}_{i-1}] \otimes W^n_{\mathbf{x}_1\ldots\mathbf{x}_s} \\ &= \sum_{\forall j<i:\ \mathbf{x}_j\in\mathcal{X}_j^n} P_{\{<i\}}(\mathbf{x}_j|j<i)[\mathbf{x}_1\ldots\mathbf{x}_{i-1}] \otimes (P_{\{>i\}}W)^n_{\mathbf{x}_1\ldots\mathbf{x}_i} \end{aligned}$$

with

$$V_{x_i} = \sum_{\forall j<i:\ x_j\in\mathcal{X}_j} [x_1\ldots x_{i-1}] \otimes (P_{\{>i\}}W)_{x_1\ldots x_i}$$

Construct a decoding observable D_i for the channel V and codebook $\mathcal{C}_i$ (note that this way D_i will be independent from the other codes — and from their decoding observables!) as [8] or [9] indexed by $\mathcal{C}_i$ (note that the decoder thus has always to make a guess). Then from [8] or [9] it follows that the average error probability $\bar{e}_{V^n}(\mathcal{C}_i, D_i)$ of decoding incorrectly $\mathbf{c}_i \in \mathcal{C}_i$ from the signal V^n, averaged over the choice of $\mathcal{C}_i$, is at most ϵ/s for all large enough n (because $I(P_i, V) = I(\mathfrak{X}_i \wedge \mathfrak{Y}\mathfrak{X}_1\cdots\mathfrak{X}_{i-1}) = R_i$). Note that the technique of proof in [8] and [9] carries over without change to the present case of an arbitrary C*–algebra instead of the full operator algebra of a Hilbert space: the only thing to check is that in the more general context there are still the "typical projectors" from [14], built up from a Schatten decomposition of the state in question inside $\mathfrak{Y}$.
Now we confront the decoder with the signal ρ, obtaining an average error probability

$$\bar{e}_i = \bar{e}_\rho(\mathcal{C}_i, D_i) = 1 - \frac{1}{L_i}\sum_{\mathbf{c}_i\in\mathcal{C}_i} \mathrm{Tr}\,(\rho_{\mathbf{c}_i} D_{i,\mathbf{c}_i})$$

Obviously the code average over choice of $\mathcal{C}_j,\ j\neq i$ is exactly $\bar{e}_{V^n}(\mathcal{C}_i, D_i)$, so from the previous

$$< \bar{e}_1 + \ldots + \bar{e}_s >_{\mathcal{C}_1\ldots\mathcal{C}_s} \leq \epsilon$$

This means that there exist codebooks $\mathcal{C}_1, \ldots, \mathcal{C}_s$ with all the $\bar{e}_i \leq \epsilon$. Now consider the corresponding total operations $D_{i,\mathrm{tot}*}$ of the observables D_i. By the tender measurement lemma V.2 for each i:

$$\frac{1}{L_1\cdots L_s} \sum_{\forall j:\mathbf{c}_j\in\mathcal{C}_j} \|[\mathbf{c}_1\ldots\mathbf{c}_i]\otimes W^n_{\mathbf{c}_1\ldots\mathbf{c}_s} - D_{i,\mathrm{tot}*}\left([\mathbf{c}_1\ldots\mathbf{c}_{i-1}]\otimes W^n_{\mathbf{c}_1\ldots\mathbf{c}_s}\right)\|_1 \leq \sqrt{8\epsilon}+\epsilon$$

Their concatenation $D_{\mathrm{tot}*} = D_{s,\mathrm{tot}*}\circ\cdots\circ D_{1,\mathrm{tot}*}$, will be the decoder process, with distortion at most $s(\sqrt{8\epsilon}+\epsilon)$ (since each step introduces a distortion of $\sqrt{8\epsilon}+\epsilon$, and does not increase trace norm). This quantity also upper bounds the error probability of $\mathrm{Tr}_{\mathfrak{Y}} \circ D_{\mathrm{tot}*}$ (which is the adjoint map of an observable indexed by $\mathcal{C}_1\times\cdots\times\mathcal{C}_s$), i.e. of reading off $\mathbf{c}_1, \ldots, \mathbf{c}_s$ after application of $D_{\mathrm{tot}*}$. □

Now we may combine the (weak) converse theorem III.1 and the foregoing direct coding theorem 1 (together with the time sharing principle by which the capacity region is convex) to obtain

Theorem 2 *The capacity region of the quantum multiple access channel is the convex closure of all nonnegative* $(R_1, \ldots, R_s)$ *satisfying*

$$\forall J\subset[s] \qquad \sum_{i\in J} R_i \leq I\left(\mathfrak{X}(J)\wedge\mathfrak{Y}|\mathfrak{X}(J^c)\right)$$

for some input distributions. □

VII. QUANTUM–QUANTUM MULTIWAY CHANNELS

Allahverdyan and Saakian in [5] defined from the outset the quantum multiway channel in more general form: namely allowing in the definition of section II general C*–algebras $\mathfrak{X}_i$ (and not only commutative ones as we did). Each sender i may then use any set of states on $\mathfrak{X}_i$ for transmission. Formally, let

$$\varphi_* : \mathfrak{T}(\mathfrak{X}_1) \otimes \cdots \otimes \mathfrak{T}(\mathfrak{X}_s) \longrightarrow \mathfrak{T}(\mathfrak{Y})$$

any completely positive trace preserving linear map, i.e. a quantum operation on states. A code for sender i now consists of a map $V_i : \mathcal{M}_i \to \mathfrak{S}(\mathfrak{X}_i^{\otimes n})$, the decoders are still observables in the algebras $\mathfrak{Y}_j$. This allows for definition of error probabilities and rates, thus we get also a capacity region.
By composition we can view this as a classical–quantum channel on n–blocks:

$$(m_1, \ldots, m_s) \longmapsto \varphi_*^{\otimes n}(V_1(m_1) \otimes \cdots \otimes V_s(m_s))$$

If we now make the restriction that in coding the senders have to use *product states*, i.e. $V_i(m_i) = V_{i,m_i,1} \otimes \cdots \otimes V_{i,m_i,n}$, we see immediately that the argument from section III carries over unchanged to give a general upper bound theorem (with the slight difference that now in the formation of the channel states γ_u instead of the states $W_{f_1(m_1)_u \ldots f_s(m_s)_u}$ there occur the output states $\varphi_*(V_{1,m_1,u} \otimes \cdots \otimes V_{s,m_s,u})$).[1]
Conversely: those input states and a distribution on them being chosen we can apply the coding theorem VI.1 to see that these bounds can be approached — at least for the multiple access channel.

We have thus obtained the capacity region $\mathbf{R}^{(1)}$ in the *1–separable* case, i.e. for input states that are separable (it is easy to see that best performance is for product states). Of course we could allow the sender to use blocks of ν entangled states, getting a possibly larger region $\mathbf{R}^{(\nu)}$,[2] and we leave open the question of the ultimate capacity region $\mathbf{R} = \bigcup_\nu \mathbf{R}^{(\nu)}$ (which even for the single sender case is unsolved): whether it be equal or strictly larger than $\mathbf{R}^{(1)}$.

VIII. CONCLUSION

We introduced quantum multiway channel and bounded its capacity region. In the case of the quantum multiple access channel this led to complete characterization of the capacity region, formally identical to the result in the classical case. This confirms the validity of the principle that in the theory of classical information transmission over a quantum channel capacities are obtained from the classical formulas by replacing Shannon entropy by von Neumann entropy.
The next step is of course to prove the direct coding theorem in the case of several receivers. It seems that for this a refinement of the methods used here or even a different approach is needed. Here a reduction showed that it is enough to be able to decode the different senders sequentially, in each step using the information obtained so far. Even in the classical case of two senders and two receivers this is no longer possible: one has to decode simultaneously.

ACKNOWLEDGMENTS

I would like to thank Peter Löber for discussions on the subject, and Prof. R. Ahlswede for comments on the proof of the classical theorem.

---

[1] We note that the special argument in [9] for this situation is not correct, because it uses a modified — wrong — Fano inequality. What is more: it is unnecessary. It only appeared necessary because in the previous section of that work too weak a use of Fano inequality had been made.
[2] This is because now the subadditivity of the mutual information, used in the upper bound in section III, does not hold any longer.

APPENDIX A: EXTREMAL POINTS

We shall prove: any upper extremal point of the region of $(R_1,\ldots,R_s)$ which satisfy for all $J\subset[s]$

$$R(J)\leq I(\mathfrak{X}(J)\wedge\mathfrak{Y}\mathfrak{X}(J^c)) \qquad (J)$$

for fixed channel state is one of the $(R_1,\ldots,R_s)$ with

$$R_{\pi(i)}=I(\mathfrak{X}_{\pi(i)}\wedge\mathfrak{Y}\mathfrak{X}_{\pi(1)}\cdots\mathfrak{X}_{\pi(i-1)})$$

for some permutation π of the set $[s]$, and any such tuple belongs to this region.

Let $(R_1,\ldots,R_s)$ an upper extremal point, hence s of the inequalities (for nonempty subsets $K_1,\ldots,K_s$ of $[s]$) are met with equality. We are done if we can prove that these can be chosen to form a strictly increasing chain $\emptyset\neq K_1\subset K_2\subset\ldots\subset K_s=[s]$, w.l.o.g. $K_i=\{s,s-1,\ldots,s-i+1\}$, because then

$$\begin{aligned}R_i&=I(\mathfrak{X}_i\cdots\mathfrak{X}_s\wedge\mathfrak{Y}\mathfrak{X}_1\cdots\mathfrak{X}_{i-1})-I(\mathfrak{X}_{i+1}\cdots\mathfrak{X}_s\wedge\mathfrak{Y}\mathfrak{X}_1\cdots\mathfrak{X}_i)\\&=I(\mathfrak{X}_i\wedge\mathfrak{Y}\mathfrak{X}_1\cdots\mathfrak{X}_{i-1})\end{aligned}\qquad\text{(A1)}$$

Choose any one such set, say $K\neq\emptyset$: $R(K)=I(\mathfrak{X}(K)\wedge\mathfrak{Y}\mathfrak{X}(K^c))$. We would like to employ induction to extend this K in both directions to a chain. That we can do so follows from the

Lemma 1 *If $R_1,\ldots,R_s$ are such that with $\emptyset\neq K\subset[s]$: $R(K)=I(\mathfrak{X}(K)\wedge\mathfrak{Y}\mathfrak{X}(K^c))$ and (J) holds for all $J\subset K$ and for all $K\subset J\subset[s]$. Then (J) holds for all $J\subset[s]$*

Proof. First let $K\subset J\subset[s]$. Then (J) together with (K) met with equality implies

$$\begin{aligned}R(J\setminus K)&\leq H(\mathfrak{Y}|\mathfrak{X}(J^c))-H(\mathfrak{Y}|\mathfrak{X}(K^c))\\&\leq H(\mathfrak{Y}|\mathfrak{X}((J\setminus K)^c))-H(\mathfrak{Y}|\mathfrak{X}([s]))\\&=I(\mathfrak{X}(J\setminus K)\wedge\mathfrak{Y}|\mathfrak{X}((J\setminus K)^c))\end{aligned}$$

(the second estimate by strong subadditivity applied to $\mathfrak{A}_1=\mathfrak{X}(J\setminus K)$, $\mathfrak{A}_2=\mathfrak{Y}\mathfrak{X}(J^c)$, $\mathfrak{A}_3=\mathfrak{X}(K)$, see [10]).

Now let $\subset[s]$ arbitrary and write $J=J_1\dot\cup J_2$ with $J_1=J\cap K$, $J_2=J\setminus K$. Then by assumption and the previous

$$\begin{aligned}R(J)=R(J_1)+R(J_2)&\leq H(\mathfrak{Y}|\mathfrak{X}(J_1^c))-H(\mathfrak{Y}|\mathfrak{X}([s]))+H(\mathfrak{Y}|\mathfrak{X}(J_2^c\cap K^c))-H(\mathfrak{Y}|\mathfrak{X}(K^c))\\&=H(\mathfrak{Y}\mathfrak{X}(J_1^c))+H(\mathfrak{Y}\mathfrak{X}(K^c\cap J_2^c))-H(\mathfrak{Y}\mathfrak{X}(K^c))-\\&\quad-H(\mathfrak{X}((K\setminus J_1)\dot\cup(K^c\setminus J_2)))-H(\mathfrak{Y}|\mathfrak{X}([s]))\\&\leq H(\mathfrak{Y}\mathfrak{X}(J_1^c\cap J_2^c))-H(\mathfrak{X}(J_1^c\cap J_2^c))-H(\mathfrak{Y}|\mathfrak{X}([s]))\\&=H(\mathfrak{Y}|\mathfrak{X}(J_1^c\cap J_2^c))-H(\mathfrak{Y}|\mathfrak{X}([s]))\\&=I(\mathfrak{X}(J)\wedge\mathfrak{Y}|\mathfrak{X}(J^c))\end{aligned}$$

(the second estimate again by strong subadditivity, with $\mathfrak{A}_1=\mathfrak{X}(J_2)$, $\mathfrak{A}_2=\mathfrak{Y}\mathfrak{X}(K^c\cap J_2^c)$, $\mathfrak{A}_3=\mathfrak{X}(K\setminus J_1)$). $\square$

It follows that if (K') is met with equality for any K', also $(K'\cap K)$ and $(K'\cup K)$ are met with equality. It is easily seen that both below and above K there must occur equalities (except we are already on lowest or highest level).
It remains to show that $(R_1,\ldots,R_s)$ with A1 belongs to the region. But this is clear by the same argument as above, used top–down: the conjunction of all (J) with $J\subset[s]$ is implied by the conjunction of all with $J\subset[s-1]$, this in turn inductively by all with $J\subset[s-2]$ etc.

---

[1] C. E. Shannon, in: *Proc. Fourth Berkeley Symposium Probability and Statistics* (ed. J. Neyman), Berkeley 1961, 611–644. Reprinted in: *Claude Elwood Shannon Collected Papers* (eds. N. J. A. Sloane, A. D. Wyner), New York 1993, 351–384
[2] R. Ahlswede, in: *Second International Symposium on Information Theory*, Hungarian Academy of Sciences 1971, 23–52
[3] R. Ahlswede, Ann. Prob. **2**,5(1974), 805–814
[4] A. El Gamal, T. Cover, Proc. IEEE **68**(1980), 1466–1483
[5] A. E. Allahverdyan, D. B. Saakian, eprint quant-ph/9712034

[6] A. S. Holevo, Rep. Math. Phys. **12**,2(1977), 273–278
[7] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, W. K. Wootters, Phys. Rev. A **54**,3(1996), 1869–1876
[8] A. S. Holevo, IEEE Trans. Inf. Theory **44**,1(1998), 269–273 (also available as eprint `quant-ph/9611023`)
[9] B. Schumacher, M. Westmoreland, Phys. Rev. A **56**,1(1997), 131–138
[10] A. Winter, "Languages of Quantum Information Theory", eprint `quant-ph/9807008`
[11] A. Winter, "Coding Theorems for Discrete Memoryless Quantum Systems", manuscript 1998
[12] I. Csiszár, J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, New York 1981
[13] R. Jozsa, J. Mod. Optics **41**,12(1994), 2315–2323
[14] B. Schumacher, Phys. Rev. A **51**,4(1995), 2738–2747