Entanglement Transmission Capacity of Compound Channels

Igor Bjelaković
Heinrich-Hertz-Lehrstuhl
für Mobilkommunikation (HFT 6)
and Institut für Mathematik
Technische Universität Berlin, Germany
Email: igor.bjelakovic@mk.tu-berlin.de

Holger Boche
Heinrich-Hertz-Lehrstuhl
für Mobilkommunikation (HFT 6)
and Institut für Mathematik
Technische Universität Berlin, Germany
Email: holger.boche@mk.tu-berlin.de

Janis Nötzel
Heinrich-Hertz-Lehrstuhl
für Mobilkommunikation (HFT 6)
Technische Universität Berlin, Germany
Email: janis.noetzel@mk.tu-berlin.de

***Abstract*— We determine the optimal achievable rate at which entanglement can be reliably transmitted when the memoryless channel used during transmission is unknown both to sender and receiver. To be more precise, we assume that both of them only know that the channel belongs to a given set of channels. Thus, they have to use encoding and decoding schemes that work well for the whole set.**

I. Introduction

One of the main goals of quantum Shannon theory is the determination of optimal transmission rates for various quantum communication tasks. In contrast to classical information theory to every quantum channel we can associate various capacities each of which characterizes the optimal rates in a specific communication scenario. In this paper we focus on the determination of the entanglement transmission capacity of quantum compound channels.

The correct formula describing this capacity for a single channel has been identified in [1], [5], [9]. Of particular interest for our work are the later on developments by Klesse [7] and Hayden, Horodecki, Winter and Yard [6] which are based on a decoupling idea that can be traced back to Schumacher and Westmoreland [8].

We use their approach to determine the optimal achievable entanglement transmission rate under channel uncertainty: while sustaining the assumption of memoryless communication, we assume that sender as well as receiver only know that the channel they use belongs to some given set of channels. This describes a somewhat more realistic situation since exact channel knowledge will hardly ever be given in applications.

Due to space limitation we will only give the proof of the direct part of the coding theorem for finite compound channels. The extension to the general case, the proof of the converse part and the relation to the entanglement-generating capacity of compound channels can be picked up in the accompanying paper [4].

The paper is organized as follows: We first fix the notation in section II. In section III we introduce our model and state the main theorem. Section IV contains two results concerning existence of recovery operations of a certain performance and behavior of entanglement fidelity under disturbance of a channel through a projection. The proof of our main theorem further uses some basic properties of typical projections and operations, which are stated in section V. From there we pass on to the proof of our main theorem in Section VI.

II. Notation and Conventions

All Hilbert spaces are assumed to have finite dimension and are over the field $\mathbb{C}$. $\mathcal{S}(\mathcal{H})$ is the set of states, i.e. positive semi-definite operators with trace 1 acting on the Hilbert space $\mathcal{H}$. Pure states are given by projections onto one-dimensional subspaces. A vector of unit length spanning such a subspace will therefore be referred to as a state vector.

The set of completely positive trace preserving (CPTP) maps between the operator spaces $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{K})$ is denoted by $\mathcal{C}(\mathcal{H},\mathcal{K})$. $\mathcal{C}^{\downarrow}(\mathcal{H},\mathcal{K})$ stands for the set of completely positive trace decreasing maps between $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{K})$. $\mathfrak{U}(\mathcal{H})$ will denote in what follows the group of unitary operators acting on $\mathcal{H}$. For a Hilbert space $\mathcal{G}\subset\mathcal{H}$ we will always identify $\mathfrak{U}(\mathcal{G})$ with a subgroup of $\mathfrak{U}(\mathcal{H})$ in the canonical way. For any projection $q\in\mathcal{B}(\mathcal{H})$ we set $q^{\perp}:=\mathbf{1}_{\mathcal{H}}-q$. Each projection $q\in\mathcal{B}(\mathcal{H})$ defines a completely positive trace decreasing map $\mathcal{Q}$ given by $\mathcal{Q}(a):=qaq$ for all $a\in\mathcal{B}(\mathcal{H})$. In a similar fashion any $u\in\mathfrak{U}(\mathcal{H})$ defines a $\mathcal{U}\in\mathcal{C}(\mathcal{H},\mathcal{H})$ by $\mathcal{U}(a):=uau^*$ for $a\in\mathcal{B}(\mathcal{H})$.

We use the base two logarithm which is denoted by $\log$. The von Neumann entropy of a state $\rho\in\mathcal{S}(\mathcal{H})$ is given by $S(\rho):=-\mathrm{tr}(\rho\log\rho)$. The coherent information for $\mathcal{N}\in\mathcal{C}(\mathcal{H},\mathcal{K})$ and $\rho\in\mathcal{S}(\mathcal{H})$ is defined by $I_c(\rho,\mathcal{N}):=S(\mathcal{N}(\rho))-S((id_{\mathcal{H}}\otimes\mathcal{N})(|\psi\rangle\langle\psi|))$, where $\psi\in\mathcal{H}\otimes\mathcal{H}$ is an arbitrary purification of the state ρ. Following the usual conventions we let $S_e(\rho,\mathcal{N}):=S((id_{\mathcal{H}}\otimes\mathcal{N})(|\psi\rangle\langle\psi|))$ denote the entropy exchange.

For $\rho\in\mathcal{S}(\mathcal{H})$ and $\mathcal{N}\in\mathcal{C}^{\downarrow}(\mathcal{H},\mathcal{K})$ the entanglement Fidelity is given by $F_e(\rho,\mathcal{N}):=\langle\psi,(id_{\mathcal{H}}\otimes\mathcal{N})(|\psi\rangle\langle\psi|)\psi\rangle$, with $\psi\in\mathcal{H}\otimes\mathcal{H}$ being an arbitrary purification of the state ρ.

In the following, a compound channel is identified with the set $\mathfrak{I}\subset\mathcal{C}(\mathcal{H},\mathcal{K})$ of its constituents. It is called finite if $\mathfrak{I}$ consists of finitely many elements.

III. Codes, Capacity and Main Result

An $(l,k_l)-$*entanglement transmission code* for the compound channel $\mathfrak{I}$ is a pair $(\mathcal{P}^l,\mathcal{R}^l)$ of CPTP maps $\mathcal{P}^l\in$

$\mathcal{C}(\mathcal{F}_l, \mathcal{H}^{\otimes l})$ where $\mathcal{F}_l$ is a Hilbert space with $k_l = \dim \mathcal{F}_l$ and $\mathcal{R}^l \in \mathcal{C}(\mathcal{K}^{\otimes l}, \mathcal{F}_l')$ with $\mathcal{F}_l \subset \mathcal{F}_l'$.

A nonnegative number R is called an achievable rate for (entanglement transmission through) $\mathfrak{I}$ if there is a sequence of (l, k_l)-entanglement transmission codes such that

1) $\liminf_{l\to\infty} \frac{1}{l} \log k_l \geq R$, and
2) $\lim_{l\to\infty} \inf_{\mathcal{N}\in\mathfrak{I}} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}^{\otimes l} \circ \mathcal{P}^l) = 1$.

The *entanglement transmission capacity* $Q(\mathfrak{I})$ of the compound channel $\mathfrak{I}$ is given by

$$Q(\mathfrak{I}) := \sup\{R \in \mathbb{R}_+ : R \text{ is achievable for } \mathfrak{I}\}.$$

Our main result can now be formulated as follows:

Theorem 3.1: Let $\mathfrak{I} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ be a compound channel. The entanglement transmission capacity of $\mathfrak{I}$ is given by

$$Q(\mathfrak{I}) = \lim_{l\to\infty} \frac{1}{l} \max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} \inf_{\mathcal{N}\in\mathfrak{I}} I_c(\rho, \mathcal{N}^{\otimes l}).$$

Remark. Corresponding results for the entanglement transmission capacity of a compound channel with informed encoder or informed decoder can be found in [4]. It is a remarkable fact that the proof of the coding theorem for an informed encoder is not, as in the classical case, just a trivial modification of the one for Theorem 3.1.

IV. One-Shot Results

This section contains essentially two statements. The first gives an estimate on the performance of universal recovery operations for a given finite set of channels. The second relates the entanglement fidelity of a coding-decoding procedure to that of a disturbed version of the procedure, where disturbance means application of a projection after using the channel.

Both results give rather loose bounds that become sharp enough only in the asymptotic limit.

A. Performance of Recovery Operations

Before we turn our attention to quantum compound channels we will shortly describe a part of recent developments in coding theory for single (i.e. perfectly known) channels as given in [7] and [6]. Both approaches are based on a decoupling idea which is closely related to approximate error correction. In order to state this decoupling lemma we need some notational preparation.

Let $\rho \in \mathcal{S}(\mathcal{H})$ be given and consider any purification $\psi \in \mathcal{H}_a \otimes \mathcal{H}$, $\mathcal{H}_a = \mathcal{H}$, of ρ. According to Stinespring's representation theorem any $\mathcal{N} \in \mathcal{C}^{\downarrow}(\mathcal{H}, \mathcal{K})$ is given by

$$\mathcal{N}(\,\cdot\,) = \mathrm{tr}_{\mathcal{H}_e}((\mathbf{1}_{\mathcal{H}} \otimes p_e) v(\,\cdot\,)v^*), \tag{1}$$

where $\mathcal{H}_e$ is a suitable finite-dimensional Hilbert space, p_e is a projection onto a subspace of $\mathcal{H}_e$, and $v : \mathcal{H} \to \mathcal{K} \otimes \mathcal{H}_e$ is an isometry.

Let us define a pure state on $\mathcal{H}_a \otimes \mathcal{K} \otimes \mathcal{H}_e$ by the formula

$$\psi' := \frac{1}{\sqrt{\mathrm{tr}(\mathcal{N}(\rho))}} (\mathbf{1}_{\mathcal{H}_a\otimes\mathcal{K}} \otimes p_e)(\mathbf{1}_{\mathcal{H}_a} \otimes v)\psi.$$

We set

$$\rho' := \mathrm{tr}_{\mathcal{H}_a\otimes\mathcal{H}_e}(|\psi'\rangle\langle\psi'|), \quad \rho'_{ae} := \mathrm{tr}_{\mathcal{K}}(|\psi'\rangle\langle\psi'|),$$

$$\rho_a := \mathrm{tr}_{\mathcal{K}\otimes\mathcal{H}_e}(|\psi'\rangle\langle\psi'|), \quad \rho'_e := \mathrm{tr}_{\mathcal{H}_a\otimes\mathcal{K}}(|\psi'\rangle\langle\psi'|).$$

The announced decoupling lemma can now be stated as follows.

Lemma 4.1 (Cf. [7], [6]): For any $\mathcal{N} \in \mathcal{C}^{\downarrow}(\mathcal{H}, \mathcal{K})$ there exists a recovery operation $\mathcal{R} \in \mathcal{C}(\mathcal{K}, \mathcal{H})$ with

$$F_e(\rho, \mathcal{R} \circ \mathcal{N}) \geq w - ||w\rho'_{ae} - w\rho_a \otimes \rho'_e||_1,$$

where $w = \mathrm{tr}(\mathcal{N}(\rho))$.

We will make use of this lemma in the proof of the following theorem, which is the heart of the proof of Theorem 3.1. In order to state the theorem, we need to introduce the code entanglement fidelity which is, for $\rho \in \mathcal{S}(\mathcal{H}), \mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$ (referring to ρ as the code) given by

$$F_{c,e}(\rho, \mathcal{N}) := \max_{\mathcal{R}\in\mathcal{C}(\mathcal{K},\mathcal{H})} F_e(\rho, \mathcal{R} \circ \mathcal{N}).$$

Theorem 4.1 (One-Shot Result for Averaged Channel): Let the Hilbert space $\mathcal{H}$ be given and consider subspaces $\mathcal{E} \subset \mathcal{G} \subset \mathcal{H}$ with $\dim \mathcal{E} = k$. For any choice of $\mathcal{N}_1, \ldots, \mathcal{N}_N \in \mathcal{C}^{\downarrow}(\mathcal{H}, \mathcal{K})$ each allowing a representation with n_j Kraus operators, $j = 1, \ldots, N$, and and for any $u \in \mathfrak{U}(\mathcal{G})$ we set

$$\mathcal{N} := \frac{1}{N}\sum_{j=1}^N \mathcal{N}_j, \qquad \mathcal{N}_u := \frac{1}{N}\sum_{j=1}^N \mathcal{N}_j \circ \mathcal{U}. \quad \text{Then}$$

$$\int_{\mathfrak{U}(\mathcal{G})} F_{c,e}(\pi_{\mathcal{E}}, \mathcal{N}_u) du \geq \mathrm{tr}(\mathcal{N}(\pi_{\mathcal{G}})) - 2\sum_{j=1}^N \sqrt{kn_j}||\mathcal{N}_j(\pi_{\mathcal{G}})||_2,$$

where the integration is with respect to the normalized Haar measure on $\mathfrak{U}(\mathcal{G})$ and $\pi_{\mathcal{E}}, \pi_{\mathcal{G}}$ are the maximally mixed states on $\mathcal{E}$ and $\mathcal{G}$.

Remark. The above Theorem gives a lower bound on the code entanglement fidelity of an averaged channel. Since entanglement fidelity is affine in the operation, $F_e(\pi_{\mathcal{F}_l}, \frac{1}{N}\sum_{i=1}^N \mathcal{N}_i^{\otimes l} \circ \mathcal{P}^l) \geq 1 - \epsilon_l$ implies $F_e(\pi_{\mathcal{F}_l}, \mathcal{N}_i^{\otimes l} \circ \mathcal{P}^l) \geq 1 - N\epsilon_l$ for every $i \in \{1, \ldots, N\}$. If $\mathfrak{I}$ is finite and ϵ_l becomes arbitrarily small for good codes, this Theorem gives a sufficient estimate. The case of general $\mathfrak{I}$ exploits the difference in polynomial growth of the number N_l of approximating channels for $\mathfrak{I}$ versus exponential decay of ϵ_l.

For the proof of this Theorem, we shall need the following two lemmata:

Lemma 4.2 (Cf. [3]): Let L and D be $N \times N$ matrices with non-negative entries which satisfy

$$L_{jl} \leq L_{jj}, \quad L_{jl} \leq L_{ll}, \quad \text{and} \quad D_{jl} \leq \max\{D_{jj}, D_{ll}\} \tag{2}$$

for all $j, l \in \{1, \ldots, N\}$. Then

$$\sum_{j,l=1}^N \frac{1}{N}\sqrt{L_{jl}D_{jl}} \leq 2\sum_{j=1}^N \sqrt{L_{jj}D_{jj}}.$$

Lemma 4.3 (Cf. [4]): Let $\mathcal{E}$ and $\mathcal{G}$ be subspaces of $\mathcal{H}$ with $\mathcal{E} \subset \mathcal{G} \subset \mathcal{H}$ where $k := \dim \mathcal{E}$, $d_{\mathcal{G}} := \dim \mathcal{G}$. p and $p_{\mathcal{G}}$ will denote the orthogonal projections onto $\mathcal{E}$ and $\mathcal{G}$. For a

Haar distributed random variable U with values in $\mathfrak{U}(\mathcal{G})$ and $x, y \in \mathcal{B}(\mathcal{H})$ we define a random sesquilinear form

$$b_{UpU^*}(x,y) := \mathrm{tr}(UpU^*x^*UpU^*y) - \frac{1}{k}\mathrm{tr}(UpU^*x^*)\mathrm{tr}(UpU^*y).$$

Then

$$\begin{aligned}\mathbb{E}\{b_{UpU^*}(x,y)\} &= \frac{k^2-1}{d^2-1}\mathrm{tr}(p_{\mathcal{G}}x^*p_{\mathcal{G}}y) \\ &\quad + \frac{1-k^2}{d(d^2-1)}\mathrm{tr}(p_{\mathcal{G}}x^*)\mathrm{tr}(p_{\mathcal{G}}y).\end{aligned}$$

Proof of Theorem 4.1: We can assume without loss of generality that the numbering of the channels is chosen in such a way that $n_1 \leq n_2 \leq \ldots \leq n_N$ holds for the numbers of Kraus operators of the maps $\mathcal{N}_1, \ldots, \mathcal{N}_N$. From Lemma 4.1 we know that for every $u \in \mathfrak{U}(\mathcal{G})$ there is a recovery operation $\mathcal{R}$ such that

$$F_e(\pi_{\mathcal{E}}, \mathcal{R} \circ \mathcal{N}_u) \geq w - ||w\rho'_{ae} - w\rho_a \otimes \rho'_e||_1, \qquad (3)$$

where we have used the notation introduced in the paragraph preceding Lemma 4.1 and the states on the RHS of equation (3) now depend on u.

For each $j \in \{1, \ldots, N\}$ let $\{b_{j,i}\}_{i=1}^{n_j}$ be the set of Kraus operators of $\mathcal{N}_j$. Then $\mathcal{N}_j \circ \mathcal{U}$ has Kraus operators $\{a_{j,i}\}_{i=1}^{n_j}$ given by $a_{j,i} = b_{j,i}u$. Let $\{f_1, \ldots, f_N\}$ and $\{e_1, \ldots, e_{n_N}\}$ be arbitrary orthonormal bases of $\mathbb{C}^N$ and $\mathbb{C}^{n_N}$ with only imposed restriction that $e_1 \otimes f_1 = \psi_e$. Let the projection p_e and unitary v in (1) be chosen in such a way that for each $\phi \in \mathcal{H}$ the relation

$$(\mathbf{1}_{\mathcal{H}} \otimes p_e)v(\phi \otimes e_1 \otimes f_1) = \sum_{j=1}^{N}\sum_{i=1}^{n_j} \frac{1}{\sqrt{N}}(b_{j,i}\phi) \otimes e_i \otimes f_j, \quad (4)$$

holds. For a purification $\psi \in \mathcal{H}_a \otimes \mathcal{H}$ of the state $\pi_{\mathcal{E}}$ we consider a Schmidt representation

$$\psi = \frac{1}{\sqrt{k}} \sum_{m=1}^{k} h_m \otimes g_m,$$

with suitable orthonormal systems $\{h_1, \ldots, h_k\}$ and $\{g_1, \ldots, g_k\}$.

We use this representation to derive explicit representations of the states $\rho'_{ae}, \rho_a, \rho'_e$ in terms of the Kraus operators of the operations $\mathcal{N}_i$ and insert them into (3). If we perform the unitary conjugation induced by the unitary map $x_{s,i,j} = h_s \otimes e_i \otimes f_j \mapsto x'_{s,i,j} = g_s \otimes e_i \otimes f_j$ followed by the complex conjugation of the matrix elements with respect to the matrix units $\{|x'_{s,i,j}\rangle\langle x'_{t,k,l}|\}_{s,i,j,t,k,l}$ we obtain an anti-linear isometry I with respect to the metrics induced by the trace distances on the operator spaces under consideration. A calculation identical to that performed by Klesse [7] and additionally using the triangle inequality for $||\cdot||_1$ as well as the relation $||a||_1 \leq \sqrt{d}||a||_2$, d being the number of non-zero singular values of the operator a shows that

$$F_{c,e}(\pi_{\mathcal{E}}, \mathcal{N}_u) \geq \mathrm{tr}(\mathcal{N}_u(\pi_{\mathcal{E}})) - \sum_{j,l=1}^{N} \frac{1}{N}\sqrt{\frac{1}{k}L_{jl}D_{jl}(u)}, \qquad (5)$$

where

$$D_{jl}(u) := \sum_{i=1,r=1}^{n_j,n_l} (\mathrm{tr}(p(a^*_{j,i}a_{l,r})^* p a^*_{j,i}a_{l,r}) - \frac{1}{k}|\mathrm{tr}(pa^*_{j,i}a_{l,r})|^2)$$

(dependence on u is through $a_{i,j} = b_{i,j}u$) and $L_{jl} := \min\{n_j, n_l\}$.

Let U be a random variable taking values in $\mathfrak{U}(\mathcal{G})$ according to the Haar measure of $\mathfrak{U}(\mathcal{G})$. Then we can infer from (5) that

$$\begin{aligned}\mathbb{E}F_{c,e}(\pi_{\mathcal{E}}, \mathcal{N} \circ \mathcal{U}) &\geq \mathbb{E}\mathrm{tr}(\mathcal{N} \circ \mathcal{U}(\pi_{\mathcal{E}})) \\ &\quad - \sum_{j,l=1}^{N} \frac{1}{N}\sqrt{\frac{1}{k}L_{jl}\mathbb{E}(D_{jl}(U))},\end{aligned} \qquad (6)$$

where we have used concavity of the function $\sqrt{\cdot}$ and Jensen's inequality. Now, setting $D_{jl} := \langle \mathcal{N}_j(\pi_{\mathcal{G}}), \mathcal{N}_l(\pi_{\mathcal{G}})\rangle_{HS}$, where $\langle \cdot\ , \cdot\ \rangle_{HS}$ denotes the Hilbert-Schmidt inner product, and using Lemma 4.3 we obtain

$$\mathbb{E}D_{jl}(U) \leq \mathrm{tr}(\mathcal{N}_j(\pi_{\mathcal{G}})\mathcal{N}_l(\pi_{\mathcal{G}})) = D_{jl}. \qquad (7)$$

It is obvious that $L_{jl} \leq L_{jj}$ and $L_{jl} \leq L_{ll}$ hold. Moreover, the Cauchy-Schwarz inequality for the Hilbert-Schmidt inner product justifies the inequality $D_{jl} \leq \max\{D_{jj}, D_{ll}\}$.

Therefore, an application of Lemma 4.2 allows us to conclude from (6) that

$$\mathbb{E}(F_{c,e}(\pi_{\mathcal{E}}, \mathcal{N} \circ \mathcal{U})) \geq \mathrm{tr}(\mathcal{N}(\pi_{\mathcal{G}})) - 2\sum_{j=1}^{N} \sqrt{kn_j}||\mathcal{N}_j(\pi_{\mathcal{G}})||_2,$$

which is what we aimed to prove. □

B. Projections and Entanglement Fidelity

Lemma 4.4: Let $\rho \in \mathcal{S}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$. Let, for some other Hilbert space $\mathcal{K}$, $\mathcal{A} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$, $\mathcal{D} \in \mathcal{C}(\mathcal{K}, \mathcal{H})$, $q \in \mathcal{B}(\mathcal{K})$ be an orthogonal projection. If for some $\epsilon > 0$ the relation $F_e(\rho, \mathcal{D} \circ \mathcal{Q} \circ \mathcal{A}) \geq 1 - \epsilon$ holds, then

$$F_e(\rho, \mathcal{D} \circ \mathcal{A}) \geq 1 - 3\epsilon. \qquad (8)$$

The following Lemma 4.5 contains an inequality which will be needed in the proof of Lemma 4.4.

Lemma 4.5 (Cf. [4]): Let $\mathcal{D} \in \mathcal{C}(\mathcal{K}, \mathcal{H})$ and $x_1 \perp x_2$, z be state vectors, $x_1, x_2 \in \mathcal{K},\ z \in \mathcal{H}$. Then

$$|\langle z, \mathcal{D}(|x_1\rangle\langle x_2|)z\rangle| \leq \sqrt{|\langle z, \mathcal{D}(\mathcal{P}_{x_1})z\rangle| \cdot |\langle z, \mathcal{D}(\mathcal{P}_{x_2})z\rangle|},$$

where $\mathcal{P}_y := |y\rangle\langle y|$ for arbitrary state vectors $y \in \mathcal{H}, \mathcal{K}$.

Proof of Lemma 4.4. Let $\dim \mathcal{H} = h,\ \dim \mathcal{K} = \kappa$, $|\psi\rangle\langle\psi| \in \mathcal{H}_a \otimes \mathcal{H}$ be a purification of ρ (w.l.o.g. $\mathcal{H}_a = \mathcal{H}$). Set $\tilde{\mathcal{D}} := id_{\mathcal{H}_a} \otimes \mathcal{D},\ \tilde{\mathcal{A}} := id_{\mathcal{H}_a} \otimes \mathcal{A},\ \tilde{q} := \mathbf{1}_{\mathcal{H}_a} \otimes q$ and, as usual, $\tilde{q}^\perp$ the orthocomplement of $\tilde{q}$ within $\mathcal{H}_a \otimes \mathcal{K}$. Obviously,

$$\begin{aligned}&F_e(\rho, \mathcal{D} \circ \mathcal{A}) = \\ &= \langle\psi, \tilde{\mathcal{D}} \circ \tilde{\mathcal{A}}(|\psi\rangle\langle\psi|)\psi\rangle \\ &= \langle\psi, \tilde{\mathcal{D}}(\tilde{q}\tilde{\mathcal{A}}(|\psi\rangle\langle\psi|)\tilde{q})\psi\rangle + \langle\psi, \tilde{\mathcal{D}}(\tilde{q}^\perp\tilde{\mathcal{A}}(|\psi\rangle\langle\psi|)\tilde{q}^\perp)\psi\rangle \\ &\quad + \langle\psi, \tilde{\mathcal{D}}(\tilde{q}\tilde{\mathcal{A}}(|\psi\rangle\langle\psi|)\tilde{q}^\perp)\psi\rangle + \langle\psi, \tilde{\mathcal{D}}(\tilde{q}^\perp\tilde{\mathcal{A}}(|\psi\rangle\langle\psi|)\tilde{q})\psi\rangle \\ &\geq \langle\psi, \tilde{\mathcal{D}}(\tilde{q}\tilde{\mathcal{A}}(|\psi\rangle\langle\psi|)\tilde{q})\psi\rangle - 2|\langle\psi, \tilde{\mathcal{D}}(\tilde{q}\tilde{\mathcal{A}}(|\psi\rangle\langle\psi|)\tilde{q}^\perp)\psi\rangle| \\ &= F_e(\rho, \mathcal{D} \circ \mathcal{Q} \circ \mathcal{A}) - 2|\langle\psi, \tilde{\mathcal{D}}(\tilde{q}\tilde{\mathcal{A}}(|\psi\rangle\langle\psi|)\tilde{q}^\perp)\psi\rangle|. \qquad (9)\end{aligned}$$

We establish a lower bound on the second term on the RHS of (9). Let

$$\tilde{\mathcal{A}}(|\psi\rangle\langle\psi|) = \sum_{i=1}^{\kappa\cdot h} \lambda_i |a_i\rangle\langle a_i|,$$

where $\{a_1, \ldots, a_{\kappa\cdot h}\}$ are assumed to form an orthonormal basis. Now every a_i can be written as $a_i = \alpha_i x_i + \beta_i y_i$ where $x_i \in \mathrm{supp}(\tilde{q})$ and $y_i \in \mathrm{supp}(\tilde{q}^{\perp})$, $i \in \{1, ..., \kappa \cdot h\}$, are state vectors and $\alpha_i, \beta_i \in \mathbb{C}$. Define $\sigma := \tilde{A}(|\psi\rangle\langle\psi|)$, then

$$\begin{aligned}\sigma &= \sum_{j=1}^{\kappa\cdot h} \lambda_j (|\alpha_j|^2 |x_j\rangle\langle x_j| + \alpha_j\beta_j^* |x_j\rangle\langle y_j| \\ &\quad + \beta_j\alpha_j^* |y_j\rangle\langle x_j| + |\beta_j|^2 |y_j\rangle\langle y_j|). \qquad (10)\end{aligned}$$

Set $X := |\langle\psi, \tilde{\mathcal{D}}(\tilde{q}\tilde{\mathcal{A}}(|\psi\rangle\langle\psi|)\tilde{q}^{\perp})\psi\rangle|$. Then, using the decomposition (10) and the abbreviation $\mathcal{P}_w := |w\rangle\langle w|$ (for $w \in \mathcal{K}$ being a state-vector)

$$\begin{aligned}X &= |\langle\psi, \tilde{\mathcal{D}}(\tilde{q}\sigma q^{\perp})\psi\rangle| \\ &\leq \sum_{i=1}^{\kappa\cdot h} |\lambda_i\alpha_i\beta_i^*| \cdot |\langle\psi, \tilde{\mathcal{D}}(|x_i\rangle\langle y_i|)\psi\rangle| \\ &\overset{\mathbf{a}}{\leq} \sum_{i=1}^{\kappa\cdot h} |\alpha_i\beta_i^*| \lambda_i \sqrt{|\langle\psi, \tilde{\mathcal{D}}(\mathcal{P}_{x_i})\psi\rangle\langle\psi, \tilde{\mathcal{D}}(\mathcal{P}_{y_i})\psi\rangle|} \\ &\overset{\mathbf{b}}{\leq} \sum_{i=1}^{\kappa\cdot h} \lambda_i |\alpha_i|^2 \langle\psi, \tilde{\mathcal{D}}(\mathcal{P}_{x_i})\psi\rangle \sum_{j=1}^{\kappa\cdot h} \lambda_j |\beta_j|^2 \langle\psi, \tilde{\mathcal{D}}(\mathcal{P}_{y_j})\psi\rangle. \\ &= F_e(\rho, \mathcal{D}\circ\mathcal{Q}\circ\mathcal{A}) \cdot F_e(\rho, \mathcal{D}\circ\mathcal{Q}^{\perp}\circ\mathcal{A}) \\ &\overset{\mathbf{c}}{\leq} \epsilon. \qquad (11)\end{aligned}$$

Here, **a** follows from utilizing Lemma 4.5, **b** is an application of the Cauchy-Schwarz inequality and **c** is true by assumption. The inequality (11) establishes (8). □

V. Typical Projections and Operations

At this point, we introduce the minimal amount of statements about typical projections and operations that is needed for the proof of Theorem 3.1. The reader interested in more details is referred to [3], [4] and references therein. The basic idea is that we throw away some non-essential information about an object and get nice estimates in return.

Lemma 5.1: There is a real number $c > 0$ such that for every two Hilbert spaces $\mathcal{H}, \mathcal{K}$ the following hold:
There are functions $h : \mathbb{N} \to \mathbb{R}_+$ and $\varphi : (0, 1/2) \to \mathbb{R}_+$ with $h(l) \searrow 0$ and $\varphi(\delta) \searrow 0$ (Setting $d := \dim\mathcal{H}$, $\kappa := \dim\mathcal{K}$, h and φ are given by $h(l) := \frac{d\cdot\kappa}{l}\log(l+1)$ $\forall l \in \mathbb{N}$ and $\varphi(\delta) := -\delta\log\frac{\delta}{d\cdot\kappa}$ $\forall\delta \in (0, 1/2)$) such that
A) For any $\rho \in \mathcal{S}(\mathcal{H})$, $\delta \in (0, 1/2)$, $l \in \mathbb{N}$ there is an orthogonal projection $q_{\delta,l} \in \mathcal{B}(\mathcal{H})^{\otimes l}$ called frequency-typical projection that satisfies

1) $\mathrm{tr}(\rho^{\otimes l} q_{\delta,l}) \geq 1 - 2^{-l(c\delta^2 - h(l))}$,
2) $q_{\delta,l}\rho^{\otimes l} q_{\delta,l} \leq 2^{-l(S(\rho)-\varphi(\delta))} q_{\delta,l}$.

The inequality 2) implies

$$||q_{\delta,l}\rho^{\otimes l} q_{\delta,l}||_2^2 \leq 2^{-l(S(\rho)-\varphi(\delta))}.$$

B) For each $\mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$, $\delta \in (0, 1/2)$, $l \in \mathbb{N}$ and maximally mixed state $\pi_{\mathcal{G}}$ on some subspace $\mathcal{G} \subset \mathcal{H}$ there is an operation $\mathcal{N}_{\delta,l} \in \mathcal{C}^{\downarrow}(\mathcal{H}^{\otimes l}, \mathcal{K}^{\otimes l})$ called reduced operation with respect to $\mathcal{N}$ and $\pi_{\mathcal{G}}$ that satisfies

3) $\mathrm{tr}(\mathcal{N}_{\delta,l}(\pi_{\mathcal{G}}^{\otimes l})) \geq 1 - 2^{-l(c\delta^2 - h(l))}$,
4) $\mathcal{N}_{\delta,l}$ has a Kraus representation with at most $n_{\delta,l} \leq 2^{l(S_e(\pi_{\mathcal{G}},\mathcal{N})+\varphi(\delta)+h(l))}$ Kraus operators.
5) For every state $\rho \in \mathcal{S}(\mathcal{H}^{\otimes l})$ and every two channels $\mathcal{I} \in \mathcal{C}^{\downarrow}(\mathcal{H}^{\otimes l}, \mathcal{H}^{\otimes l})$ and $\mathcal{L} \in \mathcal{C}^{\downarrow}(\mathcal{K}^{\otimes l}, \mathcal{H}^{\otimes l})$ the inequality $F_e(\rho, \mathcal{L}\circ\mathcal{N}_{\delta,l}\circ\mathcal{I}) \leq F_e(\rho, \mathcal{L}\circ\mathcal{N}^{\otimes l}\circ\mathcal{I})$ is fulfilled.

VI. Proof of Theorem 3.1

We will restrict our proof to the case that $\mathfrak{I}$ consists of finitely many elements. Also, we only prove the direct part $Q(\mathfrak{I}) \geq \lim_{l\to\infty}\frac{1}{l}\inf_{\mathcal{N}\in\mathfrak{I}}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} I_c(\rho, \mathcal{N}^{\otimes l})$. The converse part for finite $\mathfrak{I}$ follows from an application of Lemma 6 in [5]. In order to pass on to the case of general $\mathfrak{I}$ one approximates $\mathfrak{I}$ by a sequence $(\mathfrak{I}_l)_{l\in\mathbb{N}}$ of finite compound channels. It has to be taken care that the numbers $N_l := |\mathfrak{I}_l|$ increase subexponentially fast in l. All calculations are carried out in our papers [4] and [3].
Let us consider a compound channel given by a finite set $\mathfrak{I} := \{\mathcal{N}_1, \ldots, \mathcal{N}_N\} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ and a subspace $\mathcal{G} \subset \mathcal{H}$. For every $l \in \mathbb{N}$, we choose a subspace $\mathcal{F}_l \subset \mathcal{G}^{\otimes l}$. As usual, $\pi_{\mathcal{F}_l}$ and $\pi_{\mathcal{G}}$ denote the maximally mixed states on $\mathcal{F}_l$, respectively $\mathcal{G}$ while $k_l := \dim\mathcal{F}_l$ gives the dimension of $\mathcal{F}_l$.
For $j \in \{1, \ldots, N\}$, $\delta \in (0, 1/2)$, $l \in \mathbb{N}$ and states $\mathcal{N}_j(\pi_{\mathcal{G}})$ let $q_{j,\delta,l} \in \mathcal{B}(\mathcal{K})^{\otimes l}$ be the frequency-typical projection of $\mathcal{N}_j(\pi_{\mathcal{G}})$ and $\mathcal{N}_{j,\delta,l}$ be the reduced operation associated with $\mathcal{N}_j$ and $\pi_{\mathcal{G}}$ as given in Lemma 5.1.
For an arbitrary unitary operation $u^l \in \mathcal{B}(\mathcal{H}^{\otimes l})$ we set

$$\hat{\mathcal{N}}^l_{j,u^l,\delta} := \mathcal{Q}_{j,\delta,l}\circ\mathcal{N}_{j,\delta,l}\circ\mathcal{U}^l, \quad \hat{\mathcal{N}}^l_{u^l,\delta} := \frac{1}{N}\sum_{j=1}^N \hat{\mathcal{N}}^l_{j,u^l,\delta},$$

$$\hat{\mathcal{N}}^l_{j,\delta} := \mathcal{Q}_{j,\delta,l}\circ\mathcal{N}_{j,\delta,l}, \quad \hat{\mathcal{N}}^l_{\delta} := \frac{1}{N}\sum_{j=1}^N \hat{\mathcal{N}}^l_{j,\delta}.$$

Let U^l be a random variable taking values in $\mathfrak{U}(\mathcal{G}^{\otimes l})$ which is distributed according to the Haar measure. Application of Theorem 4.1 yields

$$\begin{aligned}\mathbb{E}F_{c,e}(\pi_{\mathcal{F}_l}, \hat{\mathcal{N}}^l_{U^l,\delta}) &\geq \mathrm{tr}(\hat{\mathcal{N}}^l_{\delta}(\pi_{\mathcal{G}}^{\otimes l})) \qquad (12) \\ &\quad -2\sum_{j=1}^N \sqrt{k_l n_{j,\delta,l}}||\hat{\mathcal{N}}^l_{j,\delta}(\pi_{\mathcal{G}}^{\otimes l})||_2,\end{aligned}$$

where $n_{j,\delta,l}$ is the number of Kraus operators of $\mathcal{N}_{j,\delta,l}$. Notice that $\mathcal{Q}_{j,\delta,l}\circ\mathcal{N}_{j,\delta,l}$ has a Kraus representation containing exactly $n_{j,\delta,l}$ elements. We will use inequality (12) in the proof of the following Lemma.

Lemma 6.1 (Direct Part for maximally mixed states): Let $\mathfrak{I} = \{\mathcal{N}_1, ..., \mathcal{N}_N\} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ be a compound channel and $\pi_{\mathcal{G}}$ the maximally mixed state associated to a subspace $\mathcal{G} \subset \mathcal{H}$. Then

$$Q(\mathfrak{I}) \geq \min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\pi_{\mathcal{G}}, \mathcal{N}_i).$$

Proof. We show that for every $\epsilon > 0$ the number $\min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\pi_{\mathcal{G}},\mathcal{N}_i) - \epsilon$ is an achievable rate for $\mathfrak{I}$.
1) If $\min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\pi_{\mathcal{G}},\mathcal{N}_i) - \epsilon \leq 0$, there is nothing to prove.
2) Let $\min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\pi_{\mathcal{G}},\mathcal{N}_i) - \epsilon > 0$.
Choose $\delta \in (0, 1/2)$ and $l_0 \in \mathbb{N}$ satisfying $2\cdot\varphi(\delta) + h(l_0) < \epsilon/2$ with functions φ, h from Lemma 5.1.
For every $l \in \mathbb{N}$ let the dimension of the subspace $\mathcal{F}_l \subset \mathcal{G}^{\otimes l}$ be given by

$$k_l = \lfloor 2^{l(\min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\pi_{\mathcal{G}},\mathcal{N}_i)-\epsilon)} \rfloor.$$

By $S(\pi_{\mathcal{G}}) \geq I_c(\pi_{\mathcal{G}},\mathcal{N}_j)$ (see [1]), this is always possible.
We will now give lower bounds on the terms in (12), thereby making use of Lemma 5.1:

$$\mathrm{tr}(\hat{\mathcal{N}}^l_\delta(\pi_{\mathcal{G}}^{\otimes l})) \geq 1 - 2\cdot 2^{-l(c\delta^2 - h(l))}. \tag{13}$$

A more detailed calculation can be found in [3] or [7]. Further, using that $||A+B||_2^2 \geq ||A||_2^2 + ||B||_2^2$ holds for nonnegative operators $A, B \in \mathcal{B}(\mathcal{K}^{\otimes l})$ (see [7]), we get the inequality

$$||\hat{\mathcal{N}}^l_{j,\delta}(\pi_{\mathcal{G}}^{\otimes l})||_2^2 \quad\leq\quad 2^{-l(S(\mathcal{N}_j(\pi_{\mathcal{G}}))-\varphi(\delta))}. \tag{14}$$

From (12), (13), (14) and our specific choice of k_l we get for every $l \geq l_0$

$$\mathbb{E}F_{c,e}(\pi_{\mathcal{F}_l}, \hat{\mathcal{N}}^l_{U^l,\delta}) \geq 1 - 2\cdot 2^{-l(c\delta^2-h(l))} - 2N\sqrt{2^{-l\epsilon/2}}.$$

This shows the existence of at least one sequence $(\mathcal{W}^l, \mathcal{R}^l)_{l\in\mathbb{N}}$ of $(l, k_l)-$ entanglement transmission codes for $\mathfrak{I}$ and

$$\liminf_{l\to\infty} \frac{1}{l}\log k_l = \min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\pi_{\mathcal{G}},\mathcal{N}_i) - \epsilon$$

as well as (using that entanglement fidelity is affine in the channel), for every $l \in \mathbb{N}$ with $l \geq l_0$

$$\min_{j\in\{1,\ldots,N\}} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \hat{\mathcal{N}}^l_{j,\delta} \circ \mathcal{W}^l) \geq 1 - N\frac{1}{3}\epsilon_l \tag{15}$$

where $\mathcal{W}^l(\cdot) = w^l(\cdot)w^{l*},\ w^l \in \mathfrak{U}(\mathcal{G}^{\otimes l})\ \forall l \in \mathbb{N}$, and

$$\epsilon_l = 3\cdot(2\cdot 2^{-l(c\delta^2-h(l))} + 2N\sqrt{2^{-l\epsilon/2}}). \tag{16}$$

For every $j \in \{1,\ldots,N\}$ and $l \in \mathbb{N}\backslash\{1,\ldots,l_0-1\}$ we thus have, by property 5) of Lemma 5.1, construction of $\hat{\mathcal{N}}^l_{j,w^j,\delta}$, and equation (15),

$$\begin{aligned} & F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{Q}_{j,\delta,l} \circ \mathcal{N}_j^{\otimes l} \circ \mathcal{W}^l) \geq \\ \geq\ & F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{Q}_{j,\delta,l} \circ \mathcal{N}_{j,\delta,l} \circ \mathcal{W}^l) \\ =\ & F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \hat{\mathcal{N}}^l_{j,w^j,\delta}) \\ \geq\ & 1 - N\frac{1}{3}\epsilon_l. \end{aligned}$$

By Lemma 4.4, this immediately implies

$$\min_{\mathcal{N}_j\in\mathfrak{I}} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_j^{\otimes l} \circ \mathcal{W}^l) \geq 1 - N\epsilon_l \quad \forall l \in \mathbb{N}\backslash\{1,\ldots,l_0-1\}.$$

Since $\epsilon > 0$ was arbitrary, we have shown that $\min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\pi_{\mathcal{G}},\mathcal{N}_i)$ is an achievable rate. □

For the proof of Theorem 3.1 we only need one more ingredient, which is a generalization of the well known BSST Lemma of [2]:

Lemma 6.2 (Compound BSST Lemma, Cf. [3]): Let $\mathfrak{I} \subset \mathcal{C}(\mathcal{H},\mathcal{K})$ be an arbitrary set of channels. For any $\rho \in \mathcal{S}(\mathcal{H})$ let $q_{\delta,l} \in \mathcal{B}(\mathcal{H}^{\otimes l})$ be the frequency-typical projection of ρ and set

$$\pi_{\delta,l} := \frac{q_{\delta,l}}{\mathrm{tr}(q_{\delta,l})} \in \mathcal{S}(\mathcal{H}^{\otimes l}).$$

Then there is a positive sequence $(\delta_l)_{l\in\mathbb{N}}$ satisfying $\lim_{l\to\infty}\delta_l = 0$ with

$$\lim_{l\to\infty}\frac{1}{l}\inf_{\mathcal{N}\in\mathfrak{I}} I_c(\pi_{\delta_l,l}, \mathcal{N}^{\otimes l}) = \inf_{\mathcal{N}\in\mathfrak{I}} I_c(\rho,\mathcal{N}).$$

From Lemma 6.1 and the fact that

$$Q(\mathfrak{I}^{\otimes l}) = lQ(\mathfrak{I}) \tag{17}$$

holds for every $l \in \mathbb{N}$ we get independent from the value of l and for every maximally mixed state $\pi_{\mathcal{F}_l} \in \mathcal{S}(\mathcal{H}^{\otimes l})$ supported on a subspace $\mathcal{F}_l \subset \mathcal{H}^{\otimes l}$ the inequality

$$Q(\mathfrak{I}) \geq \frac{1}{l}\min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\pi_{\mathcal{F}_l}, \mathcal{N}_i^{\otimes l}). \tag{18}$$

Let $\rho \in \mathcal{S}(\mathcal{H})$ be arbitrary and $(\delta_l)_{l\in\mathbb{N}}$, $(\pi_{\delta_l,l})_{l\in\mathbb{N}}$ as in Lemma 6.2. Then by (18) and Lemma 6.2 we have

$$\begin{aligned} Q(\mathfrak{I}) &\geq \lim_{l\to\infty}\frac{1}{l}\min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\pi_{\delta_l,l}, \mathcal{N}_i^{\otimes l}) \\ &= \min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\rho,\mathcal{N}_i) \\ &= \min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\rho,\mathcal{N}_i). \end{aligned} \tag{19}$$

Thus, $Q(\mathfrak{I}) \geq \max_{\rho\in\mathcal{S}(\mathcal{H})}\min_{\mathcal{N}_i\in\mathfrak{I}} I_c(\rho,\mathcal{N}_i)$ has to hold. A second application of equation (17) and taking the limit $l\to\infty$ yields the desired result. □

REFERENCES

[1] H. Barnum, E. Knill, and M.A. Nielsen, "On Quantum Fidelities and Channel Capacities", *IEEE Trans. Inf. Th.* 46, 1317-1329 (2000) H. Barnum, M.A. Nielsen, B. Schumacher, "Information transmission through a noisy quantum channel", *Phys. Rev. A* Vol. 57, No. 6, 4153 (1998)

[2] C.H. Bennett, P.W. Shor, J.A. Smolin, and A.V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem", *IEEE Trans. Inf. Th.* 48, 2637-2655 (2002)

[3] I. Bjelaković, H. Boche, J. Nötzel, "Quantum capacity of a class of compound channels", *Phys. Rev. A* 78, 042331, (2008)

[4] I. Bjelaković, H. Boche, J. Nötzel, "Entanglement transmission and generation under channel uncertainty: Universal quantum channel coding", submitted to: *Comm. Math. Phys.* - Available at: http://arxiv.org/abs/0811.4588

[5] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel", *IEEE Trans. Inf. Th.* 51, No.1, 44-55 (2005)

[6] P. Hayden, M. Horodecki, A. Winter, J. Yard, "A decoupling approach to the quantum capacity", *Open. Syst. Inf. Dyn.* 15, 7-19 (2008)

[7] R. Klesse, "Approximate Quantum Error Correction, Random Codes, and Quantum Channel Capacity", *Phys. Rev. A* 75, 062315 (2007)

[8] B. Schumacher, M.D. Westmoreland, "Sending classical information via noisy quantum channels", *Phys. Rev. A* Vol. 56, No. 1, 131-138, (1997)

[9] P. Shor, unpublished talk manuscript. Available at: http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/