Quantum Channels and Simultaneous ID Coding

Peter Löber [1]

Abstract

This preprint is on identification of classical information by the use of quantum channels. We focus on simultaneous ID codes which use measurements being useful to identify an arbitrary message. We give a direct and a converse part of the appropriate coding theorem.

I. Introduction

Since 1948 when Shannon ([11]) introduced information theory as a theory of communication there has been a quite big development in this field. People realized that beyond Shannon's communication models there are others interesting for investigation. One of those is the theory of identification via channels introduced in 1989 by Ahlswede/Dueck ([1]). Here the receiver is not interested in the exact message. He only wants to know if the sent message is a special one he is interested in. The authors gave a nice proof that under these constraints there are codes with double exponential size in the block length of the transmitted codewords. Even though this was a big surprise, it is (at least technically) much harder to give a satisfying converse to this coding theorem. A strong converse was given by Han/Verdú ([3]) in 1993. They used the fact that for (randomized) ID codes it is important that two "codewords" (input distributions) induce distinguishable output distributions. (The discussion of other interesting code models may be found in [2].)
The goal of this preprint is to study the ID scheme for the case of the information being transmitted by a quantum channel. We give a coding theorem and a strong converse theorem for ID coding schemes that use measurements which may be used to identify every message. This means that the measurement the receiver has to perform to access the information is not allowed to depend on the message he's actually interested in.
Investigation in quantum channels started in the 1960's (see [8] for a list of references), leading to Holevo's famous upper bound ([5]) which implies immediately a weak converse for the transmission problem of the memoryless quantum channel. Even though it was undoubted that the appropriate coding theorem holds it was not before 1996 when people were able to prove this direct part of the coding theorem ([4], [7], [10]). Today it's also known that the strong converse holds ([9], [12]), and that things work for non-stationary quantum channels, too ([13]).

II. Basic Definitions and Main Results

Definition 2.1: Let $A = \{1, ..., a\}$ a finite set and $\mathcal{H}$ a finite dimensional (complex) Hilbert space with $S(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})^*$ its corresponding *state space* (positive and unity preserving linear forms on $\mathcal{L}(\mathcal{H})$). A *quantum channel* $\mathbf{W} = (W_n)_{n\in\mathbb{N}}$ is a sequence of maps

$$W^n : A^n \to S(\mathcal{H}^{\otimes n}) \qquad x^n \mapsto W^n_{x^n} \ .$$

It is called *memoryless* if, for $x^n = (x_1, \ldots, x_n) \in A^n$, it holds: $W^n_{x^n} = W^1_{x_1} \otimes \ldots \otimes W^1_{x_n}$.

To access the (classical) information of a quantum state we have to perform a measurement on the output space.

[1] Email: loeber@mathematik.uni-bielefeld.de

Definition 2.2: Let $\mathcal{H}$ a finite dimensional (complex) Hilbert space. A *POM (positive operator measurement) on* $\mathcal{H}$ is a finite collection $D = (D_i)_{i=1,\dots,N}$ of non-negative operators D_i on $\mathcal{H}$ such that $\sum_{i=1}^{N} D_i = \mathbf{1}_{\mathcal{H}}$. Here $\mathbf{1}_{\mathcal{H}}$ denotes the unity operator on $\mathcal{H}$.

Remark 2.3: A POM is a kind of resolution of unity. Its practical interpretation is the following: Given a state $\sigma \in S(\mathcal{H})$ the probability that the result of measurement D will be i is $\sigma(D_i)$.

Example 2.4: Let $\mathcal{H}$ a finite dim. (complex) Hilbert space with ON basis $(\psi_i)_{i=1,\dots,N}$. Let D_i the projection on ψ_i. The collection $D = (D_i)_{i=1,\dots,N}$ is a POM, a so called *von Neumann measurement.*
It can be shown that any POM on $\mathcal{H}$ may be interpreted as a von Neumann measurement on a larger system. This is known as Naimark's theorem. (For a rigorous formulation and proof see p. 65 of [6].)

We start with the definitions for the transmission problem because we will present our results in terms of its capacities.

Definition 2.5: A (n, M, ε) *Q code* is a set of pairs $\{(c_m, E_m) : m = 1, \dots, M\}$ with $c_i \in A^n$ and $E \triangleq (E_m)_{m=1,\dots,M}$ a POM on $\mathcal{H}^{\otimes n}$ such that $(\forall\, m = 1, \dots, M)$:

$$W^n_{c_m}(E_m) \geq 1 - \varepsilon \ .$$

Definition 2.6: Given a quantum channel $\mathbf{W}$ denote the maximum M such that there is a (n, M, ε) Q code by $M(n, \varepsilon)$. For $0 < \varepsilon < 1$ we introduce a *pessimistic* and an *optimistic* ε*-error capacity* by:

$$C(\varepsilon) \triangleq \liminf_{n\to\infty} \frac{\log M(n,\varepsilon)}{n} \qquad \text{and} \qquad \bar{C}(\varepsilon) \triangleq \limsup_{n\to\infty} \frac{\log M(n,\varepsilon)}{n} \ .$$

We may write down the following four *capacities:*

$$C_0 \triangleq \inf_{\varepsilon>0} C(\varepsilon) \ , \qquad C_1 \triangleq \sup_{\varepsilon<1} C(\varepsilon) \ , \qquad \bar{C}_0 \triangleq \inf_{\varepsilon>0} \bar{C}(\varepsilon) \ , \qquad \bar{C}_1 \triangleq \sup_{\varepsilon<1} \bar{C}(\varepsilon) \ .$$

We mention that (obviously) $C_0 \leq C_1, \bar{C}_0 \leq \bar{C}_1$, and turn directly to the ID code model: Whereas for Shannon's transmission problem the receiver wants exactly know which message was sent, in the ID model the receiver wants only to check if it is some (fixed) message i. The sender (of course) does not know in which message the receiver is interested in. The canonical model for a quantum version of the ID code model is the following (cf. [1] for further motivation and examples):

Definition 2.7: A *(randomized)* $(n, N, \lambda_1, \lambda_2)$ *Q-ID code* is a set of pairs $\{(P_i, D_i) : i = 1, \dots, N\}$ where the P_i are PDs on A^n and the $\mathbf{0} \leq D_i \leq \mathbf{1}$ denote operators on $\mathcal{H}^{\otimes n}$ such that $(\forall\ i, j = 1, \dots, N : i \neq j)$:

$$P_i W^n(D_i) \geq 1 - \lambda_1 \quad \text{and} \quad P_i W^n(D_j) \leq \lambda_2 \ .$$

We use here and in the following $P_i W^n = \sum_{x^n \in A^n} P_i(x^n) W^n_{x^n} \in S(\mathcal{H}^{\otimes n})$ as an abbreviation.

We draw attention to the fact that we use random encoding, what means that a message is represented by a probability distribution on the possible codewords and (in general) not by a single codeword (cf.[1]).

It is important to realize that for the Q-ID code model (above) the D_i are not supposed to form a POM. Each D_i for itself (together with $\mathbf{1}-D_i$) could be thought of as a POM, namely as the POM the receiver performs asking for message i. Therefore there is a remarkable fact of Q-ID codes that is different from the classical ID case: The receiver can't in general use the same received state to ask for two different messages i and j because asking for message i includes a measurement on this state. To overcome this problem we formulate a second code model for which there has to be one single (simultaneous) measurement which is good to identify every message. This model is also valid if the one who performs the measurement is not the (final) receiver himself and doesn't also know in which message this receiver is interested in.

Definition 2.8: A Q-ID code $\{(P_i, D_i) : 1 \leq i \leq N\}$ is called *simultaneous* iff there is a POM $(E_m)_{m\in\mathcal{M}}$ and subsets $\mathcal{A}_i \subseteq \mathcal{M}$ $(1 \leq i \leq N)$ such that $D_i = \sum_{m\in\mathcal{A}_i} E_m$.

So, $(E_m)_m$ is a common refinement of the resolutions of unity $(D_i, \mathbf{1}-D_i)_{i=1,\ldots,N}$.

Definition 2.9: Denote the maximum N such that there is a $(n, N, \lambda_1, \lambda_2)$ Q-ID code by $N(n, \lambda_1, \lambda_2)$. The *ID capacities* are defined as follows (cf. [1]):

$$C(\lambda_1, \lambda_2) := \liminf_{n\to\infty} \frac{\log\log N(n, \lambda_1, \lambda_2)}{n}, \quad \text{and} \quad \bar{C}(\lambda_1, \lambda_2) := \limsup_{n\to\infty} \frac{\log\log N(n, \lambda_1, \lambda_2)}{n} \ .$$

Following the same lines we define $N^{sim}(n, \lambda_1, \lambda_2)$, $C^{sim}(\lambda_1, \lambda_2)$ and $\bar{C}^{sim}(\lambda_1, \lambda_2)$ for the smaller class of simultaneous Q-ID codes.

Remark 2.10: $N(n, \lambda_1, \lambda_2) \geq N^{sim}(n, \lambda_1, \lambda_2)$, $C(\lambda_1, \lambda_2) \geq C^{sim}(\lambda_1, \lambda_2)$, and $\bar{C}(\lambda_1, \lambda_2) \geq \bar{C}^{sim}(\lambda_1, \lambda_2)$.

Theorem 2.11: For all $\lambda_1, \lambda_2 > 0$ it holds

$$C^{sim}(\lambda_1, \lambda_2) \geq C_0 \qquad \text{and} \qquad \bar{C}^{sim}(\lambda_1, \lambda_2) \geq \bar{C}_0 \ .$$

Corollary 2.12: For all $\lambda_1, \lambda_2 > 0$ it holds

$$C(\lambda_1, \lambda_2) \geq C_0 \qquad \text{and} \qquad \bar{C}(\lambda_1, \lambda_2) \geq \bar{C}_0 \ .$$

Theorem 2.13: For all $\lambda_1 + \lambda_2 < 1$ it holds

$$C^{sim}(\lambda_1, \lambda_2) \leq C_1 \qquad \text{and} \qquad \bar{C}^{sim}(\lambda_1, \lambda_2) \leq \bar{C}_1 \ .$$

Corollary 2.14: As it is known for memoryless (!) quantum channels that all the four transmission capacities of Definition 2.6 are equal to one constant C, it holds for all $\lambda_1, \lambda_2 > 0$ with $\lambda_1 + \lambda_2 < 1$:

$$C^{sim}(\lambda_1, \lambda_2) = \bar{C}^{sim}(\lambda_1, \lambda_2) = C \ .$$

Here C is given by the usual formula

$$C = \max_{P \text{ PD on } A} \left(H(P\hat{W}) - \sum_{x\in A} P(x) H(\hat{W}_x) \right) \ ,$$

with $H(\hat{W}_x) = -\operatorname{tr}(\hat{W}_x \cdot \log \hat{W}_x)$, where for a state $\sigma \in S(\mathcal{H})$ we wrote $\hat{\sigma} \in \mathcal{L}(\mathcal{H})$ for the uniquely defined operator with $\sigma = \operatorname{tr}(\hat{\sigma}\,\cdot\,)$. (See [4], or for general input states [7] or [10], for a proof of $C_0 \geq C$, and [9] or [12] for $\bar{C}_1 \leq C$.) Of course our theorems apply to other quantum channels, too, e.g. to the non-stationary quantum channels (cf. [13]).

We will prove Coding Theorem 2.11 in the next section. At the end of the fourth section there is a proof of the Converse Theorem 2.13. (This proof will be completed by a theorem we prove in the fifth section.)

III. Direct Part of Simultaneous Q-ID Coding

For this section we were fortunately able to follow directly [1]. We formulate a lemma what is up to slight modifications nothing but the main proposition used in that paper:

Lemma 3.1: Let $\mathcal{M}$ a finite set of cardinality M and $\lambda \in (0,1)$. Let $\varepsilon > 0$ so small such that $\lambda \log_2(\frac{1}{\varepsilon} - 1) > 2$. There are at least $N \geq \frac{1}{M} 2^{\lfloor \varepsilon M \rfloor}$ subsets $\mathcal{A}_1, \ldots, \mathcal{A}_N \subset \mathcal{M}$, each of cardinality $\lfloor \varepsilon M \rfloor$, such that it holds for the cardinalities of the pairwise intersections:

$$|\mathcal{A}_i \cap \mathcal{A}_j| < \lambda \lfloor \varepsilon M \rfloor \qquad \forall\ i,j = 1, \ldots, N\ (i \neq j)\ .$$

Proof: Let N the maximum number such that there are subsets $\mathcal{A}_1, \ldots, \mathcal{A}_N \subset \mathcal{M}$ with the desired properties. Let $a := \lfloor \varepsilon M \rfloor$. For each $i = 1, \ldots, N$ we count the number of subsets $\mathcal{A} \subset \mathcal{M}$ with cardinality a but $|\mathcal{A}_i \cap \mathcal{A}| \geq \lambda a$. This number is

$$\sum_{k=\lceil \lambda a \rceil}^{a} \binom{M-a}{a-k}\binom{a}{k} \;\leq\; \sum_{k=\lceil \lambda a \rceil}^{a} \binom{M}{a-k} 2^a \;\leq\; a \binom{M}{a - \lceil \lambda a \rceil} 2^a \;=:\; S\ .$$

If $\binom{M}{a} > N \cdot S$ then we could add another set to our family of subsets, contradicting N being chosen as maximum number. Therefore:

$$N \;\geq\; \frac{1}{S}\binom{M}{a} \;\geq\; \frac{1}{a}\, 2^{-a} \Big(\underbrace{\frac{M-a}{a}}_{\geq \frac{1}{\varepsilon} - 1}\Big)^{\lceil \lambda a \rceil} \;\geq\; \frac{1}{M}\, 2^{-a}\, 2^{\lceil \lambda a \rceil \log_2(\frac{1}{\varepsilon} - 1)} \;\geq\; \frac{1}{M}\, 2^a\ .$$

■

Coding Theorem 2.11 is an immediate consequence of the following proposition (s. also Remark 3.3).

Proposition 3.2: Let $\lambda_1, \lambda_2, \delta > 0$. Let $\lambda \triangleq \min(\lambda_1, \lambda_2/2)$ and $\varepsilon > 0$ so small that $\lambda \log(\frac{1}{\varepsilon} - 1) > 2$. There is a positive number n_0 such that for all $n \geq n_0$ there exists some simultaneous $(n, N, \lambda_1, \lambda_2)$ Q-ID code $\{(P_i, D_i) : i = 1, \ldots, N\}$ with $N \geq 2^{\lfloor \varepsilon 2^{(C_0 - \delta)n} \rfloor - n}$.

Proof: By definition of C_0 there is a (n, M, λ) Q code $\mathcal{C} = \{(c_m, E_m) : m = 1, \ldots, M\}$ with $M \geq 2^{(C_0 - \delta)n}$ if only n is large enough. Using $\mathcal{M} = \{1, \ldots, M\}$ as ground set, Lemma 3.1 provides us with subsets $\mathcal{A}_1, \ldots, \mathcal{A}_N \subset \mathcal{M}$ of cardinality $\lfloor \varepsilon M \rfloor$ with pairwise intersections smaller then $\lambda \lfloor \varepsilon M \rfloor$. Here for the number N of those sets it holds

$$N \geq \frac{1}{M} 2^{\lfloor \varepsilon M \rfloor} \geq 2^{\lfloor \varepsilon 2^{(C_0 - \delta)n} \rfloor - n}\ .$$

We construct a simultaneous Q-ID code $\{(P_i, D_i) : i = 1, \ldots, N\}$ by taking as P_i the uniform distribution on $\mathcal{A}_i$ and as D_i the sum of the corresponding E_ms:

$$P_i(x^n) = \begin{cases} \frac{1}{\lfloor \varepsilon M \rfloor} & \text{if } x^n \in \mathcal{A}_i, \\ 0 & else, \end{cases} \qquad \text{and} \qquad D_i = \sum_{m \in \mathcal{A}_i} E_m \qquad (i = 1, \ldots, N) \ .$$

It's now straight forwards to calculate the errors being small:

$$\begin{aligned} \mathrm{tr}\,(D_i \cdot P_i W^n) &= \frac{1}{\lfloor \varepsilon M \rfloor} \sum_{m \in \mathcal{A}_i} \sum_{m' \in \mathcal{A}_i} \mathrm{tr}\,(E_m \cdot W^n_{c_{m'}}) \\ &\geq \frac{1}{\lfloor \varepsilon M \rfloor} \sum_{m \in \mathcal{A}_i} \mathrm{tr}\,(E_m \cdot W^n_{c_m}) \ \geq \ 1 - \lambda \ \geq \ 1 - \lambda_1 \ , \end{aligned}$$

and for $i \neq j$:

$$\begin{aligned} \mathrm{tr}\,(D_j \cdot P_i W^n) &= \frac{1}{\lfloor \varepsilon M \rfloor} \sum_{m \in \mathcal{A}_j} \sum_{m' \in \mathcal{A}_i} \mathrm{tr}\,(E_m \cdot W^n_{c_{m'}}) \\ &= \frac{1}{\lfloor \varepsilon M \rfloor} \sum_{m \in \mathcal{A}_j} \left(\sum_{m' \in \mathcal{A}_i \cap \mathcal{A}_j} \mathrm{tr}\,(E_m \cdot W^n_{c_{m'}}) \ + \sum_{m' \in \mathcal{A}_i \setminus \mathcal{A}_j} \mathrm{tr}\,(E_m \cdot W^n_{c_{m'}}) \right) \\ &= \frac{1}{\lfloor \varepsilon M \rfloor} \sum_{m' \in \mathcal{A}_i \cap \mathcal{A}_j} \underbrace{\mathrm{tr}\,(\sum_{m \in \mathcal{A}_j} E_m \cdot W^n_{c_{m'}})}_{\leq 1} \ + \ \frac{1}{\lfloor \varepsilon M \rfloor} \sum_{m' \in \mathcal{A}_i \setminus \mathcal{A}_j} \underbrace{\mathrm{tr}\,(\sum_{m \in \mathcal{A}_j} E_m \cdot W^n_{c_{m'}})}_{\leq \lambda} \\ &\leq \frac{1}{\lfloor \varepsilon M \rfloor} \cdot \lambda \lfloor \varepsilon M \rfloor \ + \ \frac{1}{\lfloor \varepsilon M \rfloor} \cdot \lfloor \varepsilon M \rfloor \cdot \lambda \quad = \quad 2\,\lambda \quad \leq \quad \lambda_2 \ . \end{aligned}$$

■

Remark 3.3: It is obvious that the previous proposition still holds for (at least) an infinite number of $n \in \mathbb{N}$ if we replace C_0 by (the possibly larger) $\bar{C}_0$. This shows that $\bar{C}^{sim}(\lambda_1, \lambda_2) \geq \bar{C}_0$. Moreover it's clear that the slightly stronger statements $C^{sim}(\lambda_1, \lambda_2) \geq C(\lambda)$ and $\bar{C}^{sim}(\lambda_1, \lambda_2) \geq \bar{C}(\lambda)$ hold for $\lambda_1, \lambda_2 > 0$ and $\lambda \triangleq \min(\lambda_1, {}^{\lambda_2}\!/_2)$.

IV. A Resolvability Theory for Quantum Channels

In this section we develop a resolvability theory for quantum channels. This theory arises quite naturally from those for classical channels (cf. [3]). Very loosely spoken it concerns the following question: Say that two input distributions are similar if the norm $d_1(P, Q) \triangleq \sum_{x \in A} |P(x) - Q(x)|$ of the corresponding output distributions (!) is small. How small may a set of input distributions be under the constraint that it represents all input distributions up to similarity?

As we focus on simultaneous ID coding we will have to make our definitions dependent on a fixed measurement. Remind that there are certainly very useless measurements (e.g. trivial ones), what means that applications of the results should only be of interest if one uses special (non-trivial) measurements (e.g. the underlying measurement of a "good" simultaneous Q-ID code). Like above we will prefer the notion of probability distributions instead of random variables.

Definition 4.1: Let A a finite set and $\mathcal{H}$ a finite dimensional Hilbert space. A *process* $\mathbf{P}$ on A is a sequence $\mathbf{P} = (P^n)_{n \in \mathbb{N}}$ with P^n a PD on A^n, a *measurement process* $\mathbf{E}$ on $\mathcal{H}$ is

a sequence $\mathbf{E} = (E^n)_{n\in\mathbb{N}}$ with E^n a POM on $\mathcal{H}^{\otimes n}$. We call $(\mathbf{P}, \mathbf{E})$ a *pair of processes.*

Definition 4.2: Let P a PD on A. P is *M-type* ($M \in \mathbb{N}$) if $P(x) \in \{0, \frac{1}{M}, \frac{2}{M}, \ldots, \frac{M-1}{M}, 1\}$ for all $x \in A$.

Remark 4.3: The number of different M-type distributions is upper bounded by $|A|^M$.

Definition 4.4: Let P a PD on A. We call

$$R(P) \triangleq \min\{M \in \mathbb{N} : P \text{ is } M\text{-type}\}$$

the *resolution* of P. (Let $R(P) \triangleq \infty$ if P isn't M-type for all $M \in \mathbb{N}$.)

Definition 4.5: Let $\sigma \in S(\mathcal{H})$ a state and $E = (E_m)_{m\in\{1,\ldots,M\}}$ a POM on a finite dimensional Hilbert space $\mathcal{H}$. This induces a PD $\sigma(E)$ on $\{1, \ldots, M\}$ with $\sigma(E)(m) \triangleq \sigma(E_m)$. Given a second state $\rho \in S(\mathcal{H})$ let

$$d_E(\rho, \sigma) \triangleq d_1(\rho(E), \sigma(E)) \ .$$

Definition 4.6: 1. Let $\mathbf{W}$ a quantum channel, $\mathbf{E}$ a measurement processes on its output space, and $\varepsilon > 0$. $R \geq 0$ is an *ε-achievable resolution rate for* $\mathbf{E}$ if $\forall$ processes $\mathbf{P}$, $\gamma > 0$ $\exists$ process $\tilde{\mathbf{P}}, n_0 \in \mathbb{N}$:

$$\left(\frac{\log_2 R(\tilde{P}^n)}{n} \ < \ R + \gamma \quad \wedge \quad d_{E^n}(P^n W^n, \tilde{P}^n W^n) \ < \ \varepsilon \right) \quad \forall \ n \geq n_0 \ . \tag{1}$$

2. $R \geq 0$ is an *achievable resolution rate for* $\mathbf{E}$ if R is an ε-achievable resolution rate for $\mathbf{E}$ for all $\varepsilon > 0$.

3. Now the channel's *(ε-)resolution for* $\mathbf{E}$ is given as follows:

$$\begin{aligned} S_\varepsilon(\mathbf{E}) &\triangleq \inf\{R \geq 0 : \text{R is an } \varepsilon\text{-achievable resolution rate for } \mathbf{E}\}, \\ S(\mathbf{E}) &\triangleq \inf\{R \geq 0 : \text{R is an achievable resolution rate for } \mathbf{E}\}. \end{aligned}$$

4. Fixing in 1. also the input process $\mathbf{P}$, we say that $R \geq 0$ is an *ε-achievable resolution rate for* $(\mathbf{P}, \mathbf{E})$ if $\forall\, \gamma > 0$ $\exists$ process $\tilde{\mathbf{P}}, n_0 \in \mathbb{N}$:

$$\left(\frac{\log_2 R(\tilde{P}^n)}{n} \ < \ R + \gamma \quad \wedge \quad d(P^n W^n, \tilde{P}^n W^n) \ < \ \varepsilon \right) \quad \forall \ n \geq n_0 \ .$$

Like above we define numbers $S_\varepsilon(\mathbf{P}, \mathbf{E})$ and $S(\mathbf{P}, \mathbf{E})$.

We mention that $S(\mathbf{E}) \triangleq \sup_{\mathbf{P}} S(\mathbf{P}, \mathbf{E})$ and $S_\varepsilon(\mathbf{E}) \triangleq \sup_{\mathbf{P}} S_\varepsilon(\mathbf{P}, \mathbf{E})$ for all $\varepsilon > 0$.

Remark 4.7: a) If $\varepsilon \leq \varepsilon'$ then $S_\varepsilon \geq S_{\varepsilon'}$, and b) $S = \sup_{\varepsilon>0} S_\varepsilon$.

Proof: As a) is obvious we will only prove b):
Since S is an achievable resolution rate it is also an ε-achievable resolution rate, and so $S_\varepsilon \leq S$ for all $\varepsilon > 0$. Therefore $\sup_{\varepsilon>0} S_\varepsilon \leq S$.
Suppose now $\sup_{\varepsilon>0} S_\varepsilon > S$, what means that there is some $\varepsilon > 0$ such that $S_\varepsilon > S$. So,

S is no ε-achievable resolution rate, contradicting to the definition of S.

■

Next we define the notion of uniform resolution rates which will be a useful tool in the proof of Lemma 4.11.

Definition 4.8: Let $\mathbf{W}$ a quantum channel, $\mathbf{E}$ a measurement process on its output space, and $\varepsilon > 0$. $R \geq 0$ is a *uniform ε-achievable resolution rate for* $\mathbf{E}$ if $\forall\ \gamma > 0\ \exists\ n_0 \in \mathbb{N}$: $\forall$ processes $\mathbf{P}\ \exists$ process $\tilde{\mathbf{P}}$:

$$\left(\frac{\log_2 R(\tilde{P}^n)}{n} \ < \ R+\gamma \quad \wedge \quad d_{E^n}(P^nW^n, \tilde{P}^nW^n) \ < \ \varepsilon \right) \quad \forall\ n \geq n_0\ .$$

Remark 4.9: If $R \geq 0$ is an ε-achievable resolution rate for a measurement process $\mathbf{E}$ then R is also a uniform ε-achievable resolution rate for $\mathbf{E}$.

Proof: Let $R \geq 0$ an ε-achievable resolution rate for $\mathbf{E}$, and $\gamma > 0$. For a fixed process $\mathbf{P}$ there is a minimum $n_0(\mathbf{P})$ such that for some process $\tilde{\mathbf{P}}$:

$$\left(\frac{\log_2 R(\tilde{P}^n)}{n} \ < \ R+\gamma \quad \wedge \quad d_{E^n}(P^nW^n, \tilde{P}^nW^n) \ < \ \varepsilon \right) \quad \forall\ n \geq n_0(\mathbf{P})\ .$$

We have to prove that $\sup_{\mathbf{P}} n_0(\mathbf{P}) < \infty$.
Suppose the opposite and let $(\mathbf{P}_k)_k$ be a sequence of processes such that $n_k := n_0(\mathbf{P}_k)$ is strictly monotonically increasing (hence divergent). Define a new process $\mathbf{P}$ by

$$P^n := P_k^n \qquad \text{for } n_{k-1} \leq n \leq n_k\ .$$

Consider the minimum k for which $n_0(\mathbf{P}) < n_k$. As for $n_0(\mathbf{P}) \leq n < n_k$ it holds $P^n = P_k^n$ there are PDs $\tilde{P}^n$ for which $\left(\frac{\log_2 R(\tilde{P}^n)}{n} \ < \ R+\gamma \quad \wedge \quad d_{(E^n)}(P^nW^n, \tilde{P}^nW^n) \ < \ \varepsilon \right)$. By definition of $n_0(\mathbf{P}_k)$ such $\tilde{P}^n$ exist also for $n \geq n_k = n_0(\mathbf{P}_k)$. This contradicts $n_0(\mathbf{P}_k)$ being chosen as minimum number. ■

Lemma 4.10: Let $\mathbf{W}$ a quantum channel, $\mathbf{E}$ a measurement process on its output space, $(\mathbf{P}_k)_{k\in\mathbb{N}}$ a sequence of processes, $(N_n)_{n\in\mathbb{N}}$ a sequence of positive integers, and $\varepsilon > 0$, such that $(\forall\, n \in \mathbb{N})\ \min_{1\leq k<l\leq N_n} d_{E^n}(P_k^nW^n, P_l^nW^n) \geq 2\varepsilon$. Then $\limsup_{n\to\infty} \frac{\log\log N_n}{n} \leq S_\varepsilon(\mathbf{E})$.

Proof: By the previous remark, $S_\varepsilon(\mathbf{E})$ is a uniform ε-achievable resolution rate for $\mathbf{E}$.
So, for $\gamma > 0$ there is some $n_0 \in \mathbb{N}$ and a sequence of processes $(\tilde{\mathbf{P}}_k)_{k\in\mathbb{N}}$ such that:

$$\left(\frac{\log_2 R(\tilde{P}_k^n)}{n} \ < \ S_\varepsilon(\mathbf{E})+\gamma \quad \wedge \quad d_{E^n}(P_k^nW^n, \tilde{P}_k^nW^n) \ < \ \varepsilon \right) \quad \forall\ n \geq n_0 \text{ and } k \in \mathbb{N}\ .$$

For fixed $n \in \mathbb{N}$ let's assume that $\tilde{P}_k^n = \tilde{P}_l^n$ for some $1 \leq k < l \leq N_n$. This leads to

$$d_{E^n}(P_k^nW^n, P_l^nW^n) \ \leq \ d_{E^n}(P_k^nW^n, \tilde{P}_k^nW^n) + d_{E^n}(\tilde{P}_l^nW^n, P_l^nW^n) \ \leq \ 2\varepsilon\ .$$

As this is a contradiction it follows that $\tilde{P}_k^n \neq \tilde{P}_l^n$ for all $1 \leq k < l \leq N_n$.
Now, as there are not more than $|A|^{nM}$ PDs on A^n of type M, and as for all $n \geq n_0, 1 \leq$

$k \leq N_n$ it holds that $\tilde{P}_k^n$ is of type smaller than $2^{n(S_\varepsilon(\mathbf{E})+\gamma)}$:

$$\begin{aligned} N_n &\leq \sum_{M=1}^{\lfloor 2^{n(S_\varepsilon(\mathbf{E})+\gamma)} \rfloor} |A|^{nM} \leq 2^{n(S_\varepsilon(\mathbf{E})+\gamma)} |A|^{n2^{n(S_\varepsilon(\mathbf{E})+\gamma)}} \\ &= 2^{n(S_\varepsilon(\mathbf{E})+\gamma)+\log|A| n 2^{n(S_\varepsilon(\mathbf{E})+\gamma)}} \overset{n\gg 1}{<} 2^{2^{n(S_\varepsilon(\mathbf{E})+2\gamma)}} . \end{aligned}$$

It follows $\frac{\log\log N_n}{n} < S_\varepsilon(\mathbf{E}) + 2\gamma$ (for all $\gamma > 0$), and so $\limsup_{n\to\infty} \frac{\log\log N_n}{n} \leq S_\varepsilon(\mathbf{E})$. ■

Theorem 4.11: Let $\mathbf{W}$ a quantum channel and $\lambda_1, \lambda_2 > 0$ such that $\lambda_1 + \lambda_2 < 1$. Let $\varepsilon \triangleq 1 - \lambda_1 - \lambda_2$. There is a measurement process $\mathbf{E}$ on $\mathbf{W}$'s output space such that $\bar{C}^{sim}(\lambda_1, \lambda_2) \leq S_\varepsilon(\mathbf{E})$.

Proof: For each block length $n \in \mathbb{N}$ let $\{(P_i^n, D_i^n) : i = 1, \ldots, N\}$ be a simultaneous $(n, N, \lambda_1, \lambda_2)$ Q-ID code of maximum size $N_n \triangleq N(n, \lambda_1, \lambda_2)$ and $\mathbf{E} \triangleq (E_n)_{n\in\mathbb{N}}$ with E^n the common refinement of the D_i^n (cf. Def. 2.8). Consider a sequence $(\mathbf{P}_i)_{i\in\mathbb{N}}$ of processes with $\mathbf{P}_i = (P_i^n)_{n\in\mathbb{N}}$ where P_i^n is arbitrary for $i > N_n$. For all $1 \leq i < j \leq N_n$ it holds:[2]

$$\begin{aligned} d_{E^n}(P_k^n W^n, P_l^n W^n) &= d_1(P_k^n W^n(E^n), P_l^n W^n(E^n)) \geq d_1(P_k^n W^n(D_k^n), P_l^n W^n(D_k^n)) \\ &\geq 2\,(P_k^n W^n(D_k^n) - P_l^n W^n(D_k^n)) \geq 2\,(1 - \lambda_1 - \lambda_2) = 2\,\varepsilon\ . \end{aligned}$$

So, Lemma 4.10 implies $\bar{C}(\lambda_1, \lambda_2) = \limsup_{n\to\infty} \frac{\log\log N_n}{n} \leq S_\varepsilon(\mathbf{E})$. ■

Remark 4.12: If we replaced in our definitions the condition that (1) holds for all $n \in \mathbb{N}$ by "(1) holds for an infinite number of $n \in \mathbb{N}$," there would be no need of the notion of 'uniform' resolution rates (as the term "$\exists\, n_0 \in \mathbb{N}$" would disappear in the definitions). Then Lemma 4.10 would state that $\liminf_{n\to\infty} \frac{\log_2\log_2 N_n}{n} \leq S_\varepsilon(\mathbf{E})$, and the result of Theorem 4.11 would be $C^{sim}(\lambda_1, \lambda_2) \leq S_\varepsilon(\mathbf{E})$.

Now we are able to prove the Converse Theorem 2.13 up to a theorem we leave open for the next section:

Proof: [of Theorem 2.13] By Theorem 4.11 it holds $\bar{C}^{sim}(\lambda_1, \lambda_2) \leq S_\varepsilon(\mathbf{E})$. In the following section we will see (cf. Theorem 5.1) that $S_\varepsilon(\mathbf{E}) = \sup_{\mathbf{P}} S_\varepsilon(\mathbf{P}, \mathbf{E}) \leq \bar{C}_1$. Analogously we get $C^{sim}(\lambda_1, \lambda_2) \leq C_1$ (cf. the previous remark and 5.7). ■

Theorem 4.11 and Q-ID coding Theorem 2.11 imply a *converse* of Resolvability Theorem 5.1 (c.f. next section):

Remark 4.13: Let $0 < \varepsilon < 1$. There is a measurement process $\mathbf{E}$ of $\mathbf{W}$ with $\bar{C}_0 \leq S_\varepsilon(\mathbf{E})$.

The method of upperbounding ID capacities by resolutions immediately leads to trivial bounds:

Remark 4.14: It's easy to see that $S_\varepsilon(\mathbf{E}) \leq \log_2 |A|$ always holds, and $\log_2 |A|$ would remain an upper bound of resolution if we replaced in our definitions $d_{\mathbf{E}}(\rho, \sigma)$ by $d(\rho, \sigma) \triangleq \max_{\mathbf{E}} d_{\mathbf{E}}(\rho, \sigma)$. This leads to the (natural) bound $C(\lambda_1, \lambda_2) \leq \log_2 |A|$.

[2] We use that for PDs Q, Q' on a set B it holds $d_1(Q, Q') = 2 \sup_{C\subset B}(Q(C) - Q'(C))$.

V. Shannon Capacity as Upper Bound of Resolution

This section is devoted to prove the following theorem. We will do this following, like in the previous section, the ideas of [3].

Theorem 5.1: Let $(\mathbf{P}, \mathbf{E})$ a pair of processes for $\mathbf{W}$, and $\varepsilon > 0$. Then $S_\varepsilon(\mathbf{P}, \mathbf{E}) \leq \bar{C}_1$.

To prove this theorem we need some basic concepts.

Definition 5.2: Let $\mathbf{A} = (A_n)_{n\in\mathbb{N}}$ a sequence of random variables. Its *limsup in probability* is the number

$$\bar{\mathbf{A}} \triangleq \min\{\beta \in \bar{\mathbb{R}} \,|\, \forall\, \varepsilon > 0 : \lim_{n\to\infty} P[A_n \geq \beta + \varepsilon] = 0\} \ .$$

Definition 5.3: Let $(\mathbf{P}, \mathbf{E})$ a pair of processes for $\mathbf{W}$. For every $n \in \mathbb{N}$ let

$$P_{(P^n,E^n)}(x^n, y) \triangleq P^n(x^n) W^n_{x^n}(E^n_y)$$

the joint distribution of the classical channel that outputs the result of measurement E^n on W^n's output. Let

$$i_{(P^n,E^n)}(x^n, y) \triangleq \log_2 \frac{W^n_{x^n}(E^n_y)}{P^n W^n(E^n_y)}$$

its *information density.* The *sup-information rate* $\bar{\mathbf{I}}(\mathbf{P}, \mathbf{E})$ is defined to be the limsup in probability of the normalized information density $I_n \triangleq \frac{1}{n}\, i_{(P^n,E^n)}$.

Lemma 5.4: Let $(\mathbf{P}, \mathbf{E})$ a pair of processes for $\mathbf{W}$. Then $\bar{\mathbf{I}}(\mathbf{P}, \mathbf{E}) \leq \bar{C}_1$.

Proof: Let's assume that this is false and there is a pair of processes $(\mathbf{P}, \mathbf{E})$ such that for some $\alpha, \gamma > 0$:

$$P_{(P^n,E^n)} \left[\frac{1}{n}\, i_{P^n,E^n} > \bar{C}_1 + \gamma \right] > \alpha$$

for infinitely many integers n. With this assumption we will be able to construct for those integers – if only they are large enough – a $(n, M_n, 1 - \frac{\alpha}{3})$ Q code, with M_n some integer fulfilling

$$\bar{C}_1 + \frac{\gamma}{3} \ \leq \ \frac{1}{n} \log_2 M_n \ \leq \ \bar{C}_1 + \frac{\gamma}{2} \ .$$

As this is a contradiction to the strong converse (for transmission) the lemma will be proved.
Given $E^n = (E^n_1, \ldots, E^n_{b_n})$ let for every $x^n \in A^n$

$$D(x^n) \triangleq \left\{ y \in \{1, \ldots, b_n\} : \frac{1}{n}\, i_{P^n,E^n}(x^n, y) > \bar{C}_1 + \gamma \right\},$$

and then

$$G \triangleq \left\{ x^n \in A^n : W^n_{x^n}\Big(\sum_{y\in D(x^n)} E^n_y \Big) \geq \frac{\alpha}{2} \right\}.$$

We choose the codewords $c_i \in G$ successively by the random selection method with probability

$$Q^n(x^n) \triangleq \begin{cases} \frac{P^n(x^n)}{P^n(G)}, & \text{if } x^n \in G, \\ 0, & \text{otherwise.} \end{cases}$$

The decoding operator of the codeword c_i is defined to be the operator $\sum_{y\in D_i} E_y^n$ where $D_i \triangleq D(c_i) \setminus \bigcup_{j<i} D(c_j)$.
The success probability is

$$W_{c_i}^n(\sum_{y\in D_i} E_y^n) \;\geq\; W_{c_i}^n(\sum_{y\in D(c_i)} E_y^n) - \sum_{j<i} W_{c_i}^n(\sum_{y\in D(c_j)} E_y^n) \;\geq\; \frac{\alpha}{2} - \sum_{j<i} W_{c_i}^n(\sum_{y\in D(c_j)} E_y^n)\;.$$

For the expected value of the last summands it holds

$$\mathbb{E}_{Q^n}\, W_{c_i}^n(\sum_{y\in D(c_j)} E_y^n) \;\leq\; \frac{1}{P^n(G)} \sum_{x^n\in A^n} P^n(x^n) W_{x^n}^n(\sum_{y\in D(c_j)} E_y^n) \;=\; \frac{1}{P^n(G)} P^n W^n(\sum_{y\in D(c_j)} E_y^n),$$

where

$$P^n W^n(\sum_{y\in D(c_j)} E_y^n) \;=\; \sum_{y=1}^{b_n} P^n W^n(E_y^n)\cdot 1_{\{y\in D(c_j)\}} \;<\; 2^{-n(\bar{C}_1+\gamma)}\;.$$

Here the inequality holds because

$$y\in D(c_j) \quad\Leftrightarrow\quad \frac{1}{n}\log_2 \frac{W_{c_j}^n(E_y^n)}{P^n W^n(E_y^n)} > \bar{C}_1+\gamma \quad\Leftrightarrow\quad P^n W^n(E_y^n) < W_{c_j}^n(E_y^n)\cdot 2^{-n(\bar{C}_1+\gamma)}.$$

So, for the expected success probability it holds

$$\mathbb{E}_{Q^n}\, W_{c_i}^n(\sum_{y\in D_i} E_y^n) \;\geq\; \frac{\alpha}{2} - \frac{1}{P^n(G)}\, M_n\cdot 2^{-n(\bar{C}_1+\gamma)} \;\geq\; \frac{\alpha}{2} - \frac{1}{P^n(G)}\, 2^{-n\frac{\delta}{2}}\;.$$

Here, since for $\zeta\in\mathbb{R}$ and random variable $Z<1$ it holds $\mathbb{E}_P\, Z \leq P[Z\geq\zeta]+\zeta$,

$$\begin{aligned}
P^n(G) \;&=\; P^n\left[W_{x^n}^n(\sum_{y\in D(x^n)} E_y^n) \geq \frac{\alpha}{2}\right] \;\geq\; \mathbb{E}_{P^n}\, W_{x^n}^n(\sum_{y\in D(x^n)} E_y^n) - \frac{\alpha}{2} \\
&=\; \sum_{x^n\in A^n} P^n(x^n) W_{x^n}^n(\sum_{y\in D(x^n)} E_y^n) - \frac{\alpha}{2} \;=\; P_{(P^n,E^n)}\left[\frac{1}{n}\, i_{P^n,E^n} > \bar{C}_1+\gamma\right] - \frac{\alpha}{2} \\
&\geq\; \alpha - \frac{\alpha}{2} \qquad (\ldots \text{ by assumption}).
\end{aligned}$$

Hence

$$\mathbb{E}_{Q^n}\, W_{c_i}^n(\sum_{y\in D_i} E_y^n) \;\geq\; \frac{\alpha}{2} - \frac{2}{\alpha}\, 2^{-n\frac{\delta}{2}} \;\geq\; \frac{\alpha}{3}$$

– if only n is large enough –, and there is certainly one codeword $c_i\in G$ with the desired success probability. ■

Lemma 5.5 ([3]) Let Q and R probability distributions on a finite set. Then for every $\mu>0$:

$$d_1(Q,R) \;\leq\; \frac{2}{\log_2 e}\mu + 2Q\left[\log_2\frac{Q}{R} > \mu\right]\;.$$

Lemma 5.6: Let $(\mathbf{P},\mathbf{E})$ a pair of processes for $\mathbf{W}$, and $\varepsilon>0$. Then $S_\varepsilon(\mathbf{P},\mathbf{E}) \leq \bar{\mathbf{I}}(\mathbf{P},\mathbf{E})$.

Proof: Let $\gamma > 0$. We show that by the random selection method there is a process $\tilde{\mathbf{P}}$ such that

$$\frac{\log_2 R(\tilde{P}^n)}{n} \;\leq\; \bar{\mathbf{I}}(\mathbf{P},\mathbf{E}) + \gamma \qquad \text{and} \qquad \lim_{n\to\infty} d_{E^n}(P^nW^n, \tilde{P}^nW^n) \;=\; 0 \;.$$

This works as follows: For fixed n let $M \triangleq \lfloor 2^{n(\bar{\mathbf{I}}(\mathbf{P},\mathbf{E})+\gamma)} \rfloor$. Each M-tuple $(c_1,\ldots,c_M) \in (A^n)^M$ of codewords gives rise to the M-type probability distribution

$$\tilde{P}^n_{(c_1,\ldots,c_M)}(x^n) \;\triangleq\; \frac{1}{M}\sum_{i=1}^{M} \mathbf{1}_{\{x^n=c_i\}} \;.$$

We will show that

$$\lim_{n\to\infty} \mathbb{E}_{P^n}\; d_{E^n}(P^nW^n, \tilde{P}^n_{(c_1,\ldots,c_M)}W^n) \;=\; 0 \;,$$

interpreting the $c_1,\ldots,c_M$ as independent outcomes of a random experiment with underlying probability distribution P^n. This directly implies our claim.
Recall that by Def. 4.5: $d_{E^n}(P^nW^n, \tilde{P}^n_{(c_1,\ldots,c_M)}W^n) = d_1(P^nW^n(E^n), \tilde{P}^n_{(c_1,\ldots,c_M)}W^n(E^n))$, and by the previous lemma it is enough to show that for every $\mu > 0$ the following expression goes to 0 as n tends to infinity:

$$\sum_{c_1\in A^n}\cdots\sum_{c_M\in A^n} P^n(c_1)\cdots P^n(c_M)\sum_{y=1}^{b_n} \tilde{P}^n_{(c_1,\ldots,c_M)}W^n(E^n_y)\cdot \mathbf{1}\{\log_2 \frac{\tilde{P}^n_{(c_1,\ldots,c_M)}W^n(E^n_y)}{P^nW^n(E^n_y)} > \mu\}$$

$$= \frac{1}{M}\sum_{j=1}^{M}\sum_{c_1\in A^n}\cdots\sum_{c_M\in A^n} P^n(c_1)\cdots P^n(c_M)\sum_{y=1}^{b_n} W^n_{c_j}(E^n_y)\cdot \mathbf{1}\{\log_2 \frac{\tilde{P}^n_{(c_1,\ldots,c_M)}W^n(E^n_y)}{P^nW^n(E^n_y)} > \mu\} \;.$$

As all M summands are equal this is just ...

$$\sum_{c_2\in A^n}\cdots\sum_{c_M\in A^n} P^n(c_2)\cdots P^n(c_M)\sum_{c_1\in A^n}\sum_{y=1}^{b_n} P_{(P^n,E^n)}(c_1,y)\cdot \mathbf{1}\{\log_2 \frac{\tilde{P}^n_{(c_1,\ldots,c_M)}W^n(E^n_y)}{P^nW^n(E^n_y)} > \mu\}$$

$$\leq\; P_{(P^n,E^n)}\left[\frac{1}{M}2^{i(P^n,E^n)} > \tau\right] \;+\; \mathbb{E}_{P^nW^n(E^n)}P_{(P^n)^{M-1}}\left[\frac{1}{M}\sum_{j=2}^{M} 2^{i(P^n,E^n)(c_j,y)} > 1+\tau\right] \;. \quad (2)$$

Here $\tau \triangleq \frac{1}{2}(2^\mu - 1) > 0$, and the last inequality holds because

$$\begin{aligned}
\mathbf{1}\{\log_2 \frac{\tilde{P}^n_{(c_1,\ldots,c_M)}W^n(E^n_y)}{P^nW^n(E^n_y)} > \mu\} \;&=\; \mathbf{1}\{\log_2 \frac{\frac{1}{M}\sum_{j=1}^{M} W^n_{c_j}(E^n_y)}{P^nW^n(E^n_y)} > \mu\} \\
&=\; \mathbf{1}\{\log_2 \frac{1}{M}\sum_{j=1}^{M} 2^{i(P^n,E^n)(c_j,y)} > \mu\} \\
&=\; \mathbf{1}\{\frac{1}{M}2^{i(P^n,E^n)(c_1,y)} + \frac{1}{M}\sum_{j=2}^{M} 2^{i(P^n,E^n)(c_j,y)} > 1+2\tau\} \\
&\leq\; \mathbf{1}\{\frac{1}{M}2^{i(P^n,E^n)(c_1,y)} > \tau\} + \mathbf{1}\{\frac{1}{M}\sum_{j=2}^{M} 2^{i(P^n,E^n)(c_j,y)} > 1+\tau\} \;.
\end{aligned}$$

The first summand of (2) is easy to handle:

$$\begin{aligned} P_{(P^n,E^n)}\left[\frac{1}{M}2^{i_{(P^n,E^n)}} > \tau\right] &= P_{(P^n,E^n)}\left[\frac{1}{n}i_{(P^n,E^n)} > \frac{\log_2\tau}{n} + \frac{\log_2 M}{n}\right] \\ &\leq P_{(P^n,E^n)}\left[I_n > \frac{\log_2\tau}{n} + \bar{\mathbf{I}}(\mathbf{P},\mathbf{E}) + \frac{\gamma}{2}\right] \underset{n\to\infty}{\longrightarrow} 0\ . \end{aligned}$$

For the last inequality (which holds for large n) recall that $M = \lfloor 2^{n(\bar{\mathbf{I}}(\mathbf{P},\mathbf{E})+\gamma)}\rfloor$. Now, for the second summand of (2), as it holds for every $y \in \{1,\dots,b_n\}$

$$\begin{aligned} \mathbb{E}_{(P^n)^{M-1}}\frac{1}{M}\sum_{j=2}^{M} 2^{i_{(P^n,E^n)}(c_j,y)} &= \frac{1}{M}\sum_{j=2}^{M}\mathbb{E}_{P^n}\, 2^{i_{(P^n,E^n)}(c_j,y)} \leq \mathbb{E}_{P^n}\, 2^{i_{(P^n,E^n)}(c_2,y)} \\ &= \sum_{c_2\in A^n} P^n(c_2)\frac{W^n_{c_2}(E^n_y)}{P^nW^n(E^n_y)} = 1\ , \end{aligned}$$

we can use Chebychev's Inequality

$$\begin{aligned} P_{(P^n)^{M-1}}\left[\frac{1}{M}\sum_{j=2}^{M} 2^{i_{(P^n,E^n)}(c_j,y)} > 1+\tau\right] &\leq \frac{1}{\tau^2}\operatorname{Var}_{(P^n)^{M-1}}\left[\frac{1}{M}\sum_{j=2}^{M} 2^{i_{(P^n,E^n)}(c_j,y)}\right] \\ &= \frac{1}{\tau^2}\frac{1}{M^2}\sum_{j=2}^{M}\operatorname{Var}_{P^n}\left[2^{i_{(P^n,E^n)}(c_j,y)}\right] \\ &\leq \frac{1}{\tau^2}\frac{1}{M}\operatorname{Var}_{P^n}\left[2^{i_{(P^n,E^n)}(c_2,y)}\right] \\ &\leq \frac{1}{\tau^2}\frac{1}{M}\mathbb{E}_{P^n}\left[(2^{i_{(P^n,E^n)}(c_2,y)})^2\right]\ , \end{aligned}$$

and we get the following as upper bound for the second summand of (2):

$$\begin{aligned} &\frac{1}{\tau^2}\frac{1}{M}\mathbb{E}_{P^nW^n(E^n)}\mathbb{E}_{P^n}\left[(2^{i_{(P^n,E^n)}(c_2,y)})^2\right] \\ &= \frac{1}{\tau^2}\frac{1}{M}\sum_{c_2\in A^n}\sum_{y=1}^{b_n} P^n(c_2)P^nW^n(E^n_y)\left(\frac{W^n_{c_2}(E^n_y)}{P^nW^n(E^n_y)}\right)^2 \\ &= \frac{1}{\tau^2}\mathbb{E}_{(P^n,E^n)}\left[\frac{1}{M}\,2^{i_{(P^n,E^n)}(c_2,y)}\right] \\ &= \frac{1}{\tau^2}\Big(\mathbb{E}_{(P^n,E^n)}\left[\frac{1}{M}\,2^{i_{(P^n,E^n)}(c_2,y)}\cdot\mathbf{1}\{\frac{1}{M}\,2^{i_{(P^n,E^n)}(c_2,y)} \leq 2^{-n\frac{\gamma}{2}}\}\right] \\ &\qquad + \mathbb{E}_{(P^n,E^n)}\left[\frac{1}{M}\,2^{i_{(P^n,E^n)}(c_2,y)}\cdot\mathbf{1}\{\frac{1}{M}\,2^{i_{(P^n,E^n)}(c_2,y)} > 2^{-n\frac{\gamma}{2}}\}\right]\Big) \\ &\leq \frac{1}{\tau^2}\left(2^{-n\frac{\gamma}{2}} + P_{(P^n,E^n)}\left[\frac{1}{M}\,2^{i_{(P^n,E^n)}} > 2^{-n\frac{\gamma}{2}}\right]\right) \\ &\underset{n>>1}{\leq} \frac{1}{\tau^2}\left(2^{-n\frac{\gamma}{2}} + P_{(P^n,E^n)}\left[I_n > \bar{\mathbf{I}}(\mathbf{P},\mathbf{E}) + \frac{\gamma}{3}\right]\right) \underset{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

■

Remark 5.7: If we replaced the limit in Definition 5.2 by a liminf, Lemma 5.4 could be formulated as $\bar{\mathbf{I}}(\mathbf{P},\mathbf{E}) \leq C_1$, and with the changes proposed by Remark 4.12, Lemma 5.6 would still hold. Thus Theorem 5.1 would state that $S_\varepsilon(\mathbf{P},\mathbf{E}) \leq C_1$ (with changed definitions for $S_\varepsilon(\mathbf{P},\mathbf{E})$).

VI. Acknowledgments

I'd like to thank my colleague Andreas Winter for the discussion on topics of this preprint as well as for some ideas he gave me about how to look at several aspects of quantum information theory. I'd also like to thank Prof. Rudolf Ahlswede who introduced me to the area of information theory and who proposed to me the problem of identification by the use of quantum channels.

References

[1] R. Ahlswede, G. Dueck, "Identification via channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 15-29, 1989.
[2] R. Ahlswede, "General Theory of Information Transfer," Preprint 97-118, SFB 343, Universität Bielefeld, 1997.
[3] T. S. Han, S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, vol. 39, pp. 752-772, 1993.
[4] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, W. K. Wootters, "Classical information capacity of a quantum channel," *Phys. Rev. A*, vol. 54, no. 3, pp. 1869-1876, 1996.
[5] A. S. Holevo, "Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel," *Probl. Peredachi Inform.*, vol. 9, no. 3, pp. 3-11, 1973 . (Engl. transl.: *Probl. of Inf. Transm.*, vol. 9, no. 3, pp. 177-183, 1973).
[6] A. S. Holevo, "Probabilistic and Statistical Aspects of Quantum Theory," North-Holland Publishing-Company, Amsterdam (1982).
[7] A. S. Holevo, "The Capacity of the Quantum Channel with General Signal States," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 269-273, 1998 (LANL Report no. quant-ph/9611023).
[8] A. S. Holevo, "Coding Theorems for Quantum Channels," LANL Report no. quant-ph/9809023.
[9] T. Ogawa, H. Nagaoka, "Strong Converse to the Quantum Channel Coding Theorem," LANL Report no. quant-ph/9808063.
[10] B. Schumacher, M. D. Westmoreland, "Sending classical Information via noisy quantum channels," *Phys. Rev. A*, vol. 56, no. 1, pp. 131-138, 1997.
[11] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J. 27,* pp. 339-425 and 623-656, 1948.
[12] A. Winter, "Coding Theorem and Strong Converse for Quantum Channels," Preprint 98-074, SFB 343, Universität Bielefeld, 1998.
[13] A. Winter, "Coding Theorem and Strong Converse for Nonstationary Quantum Channels," Preprint in preparation.