Quantum Channels and Simultaneous ID Coding

Dissertation
zur Erlangung des Doktorgrades,
vorgelegt der Fakultät für Mathematik
der Universität Bielefeld

von

Peter Löber

Bielefeld, im April 1999

Die Gutachten wurden von Prof. Dr. R. Ahlswede und Prof. Dr. F. Götze angefertigt.
Die mündliche Prüfung fand am 3. Juli 1999 statt.

(subm. to IEEE Trans. Inf. Th.)

Contents

[...] we always have had (secret, secret, close the doors!) we always have had a great deal of difficulty in understanding the world view that quantum mechanics represents. At least I do [...]

R. P. Feynman ([3])

0 Preface

This thesis on the identification problem for quantum channels reflects my main work during two years of PhD studies at the mathematics department of the university of Bielefeld. Prof. Rudolf Ahlswede proposed me to study this problem, and I'd like to use the opportunity to thank him for this as well as for introducing me to the area of information theory. I enjoy working in his group very much.
Another special thanks is dedicated to my colleague Andreas Winter. We discussed a lot on topics of this thesis, and doing so, he deeply influenced my view on quantum information theory.

Bielefeld, April 1999

1 Introduction

Since 1948 when Shannon ([15]) introduced information theory as a theory of communication there has been a quite big development in this field. People realized that beyond Shannon's original models of communication there are further ones of interest for investigation. One of those is the theory of identification (ID) 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 doubly exponential size in the block length of the 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 ([4]) in 1992. (The discussion of other interesting code models may be found in [2].)

The goal of this thesis is to study the ID scheme for the case of the information being transfered 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 [10] for a list of references), leading to Holevo's famous upper bound ([7]) 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 ([6], [9], [14]). Today it's also known that the strong converse holds ([13], [16]), and that things work for non-stationary quantum channels, too ([17]).

The contents of this thesis essentially coincides with the contents of my preprint [12].

2 Basic Definitions and Main Results

Definition 2.1 Let $A = \{1, ..., a\}$ be a finite set and let $\mathcal{H}$ be a finite dimensional (complex) Hilbert space with $S(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})^*$ its corresponding *state space* [1] (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

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

We call it *memoryless* if $W^n_{x^n} = W^1_{x_1} \otimes \ldots \otimes W^1_{x_n}$ for all $x^n = (x_1, \ldots, x_n) \in A^n$.

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

Definition 2.2 Let $\mathcal{H}$ be a finite dimensional (complex) Hilbert space. A *POM (positive operator measurement) on* $\mathcal{H}$ is a tuple $D = (D_i)_{i=1,\ldots,N}$ 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 $\mathcal{H}$ be a finite dimensional (complex) Hilbert space with orthonormal basis $(\psi_i)_{i=1,\ldots,N}$. Let D_i be the projector on ψ_i. Then the tupel $D = (D_i)_{i=1,\ldots,N}$ 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 an (occasionally) larger system. This is known as Naimark's theorem. (For a rigorous formulation and proof see p. 65 of [8].)

We start with the definitions for the transmission problem because we will present our results on ID capacities relatively to those for transmission:

[1] $\mathcal{L}(\mathcal{H})$ denotes the space of linear operators on $\mathcal{H}$.

Definition 2.5 An (n, M, ε) *Q code* is defined to be 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 an (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 define the following four *capacities:*

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

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 to know exactly which message was sent, in the ID model the receiver only wants to check if it is some (fixed) message i. The sender (of course) does not know 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_is are probability distributions on A^n and the D_is, $\mathbf{0} \leq D_i \leq \mathbf{1}$, denote operators on $\mathcal{H}^{\otimes n}$ such that for all $i, j = 1, \dots, N$ with $i \neq j$:

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

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

We draw attention to the fact that we use random encoding, which 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_is 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 Q-ID codes have a remarkable property 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 allows to identify every message at the same time. 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* if there is a POM $(E_m)_{m\in\mathcal{M}}$ and there are subsets $\mathcal{A}_i \subseteq \mathcal{M}$ $(1 \leq i \leq N)$ such that $D_i = \sum_{m\in\mathcal{A}_i} E_m$.

$(E_m)_m$ should be viewed as a common refinement of the resolutions of unity $(D_i, \mathbf{1}-D_i)_{i=1,\dots,N}$.

We remark at this point that most examples (!) of ID coding schemes require simultaneous ID codes because their real implementation would consist of many receivers (at one time). This holds for the examples given in [1]. On the other hand this is not always the case, e.g. if both, sender and receiver, have a (possibly different) text and they want to check if it is the same one, using an ID code. Here really is only one receiver asking only one question.[2]

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

$$\begin{aligned} C(\lambda_1, \lambda_2) &\triangleq \liminf_{n\to\infty} \frac{\log\log N(n, \lambda_1, \lambda_2)}{n}, \quad \text{and} \\ \bar{C}(\lambda_1, \lambda_2) &\triangleq \limsup_{n\to\infty} \frac{\log\log N(n, \lambda_1, \lambda_2)}{n}. \end{aligned}$$

[2] This example is taken from [11].

Accordingly, 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, following the same lines.

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

We are now able to state the main results of this thesis:

Theorem 2.11 Let $\lambda_1, \lambda_2 > 0$. Then

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

Corollary 2.12 Let $\lambda_1, \lambda_2 > 0$. Then

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

Theorem 2.13 Let $\lambda_1 + \lambda_2 < 1$. Then

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

Corollary 2.14 Since it is known for memoryless (!) quantum channels that all the four transmission capacities of Definition 2.6 are equal to one constant C, we have 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 as usual, C fulfills the 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 [6], or for general input states [9] or [14], for a proof of $C_0 \geq C$, and [13] or [16] 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. [17]).

We shall prove Coding Theorem 2.11 in the next section. At the end of Section 4 there is a proof of the Converse Theorem 2.13. (This proof will be completed by a theorem we prove in Section 5.)

Remark 2.15 It is an open question whether (the analogue of) Converse Theorem 2.13 holds in the general (non-simultaneous) case, too (cf. also Remark 4.14).

3 Direct Part of Simultaneous Q-ID Coding

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

Lemma 3.1 Let $\mathcal{M}$ be a finite set of cardinality M and let $\lambda \in (0,1)$. Let $\varepsilon > 0$ be so small such that $\lambda \log_2(\frac{1}{\varepsilon} - 1) > 2$. Then 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 the cardinalities of the pairwise intersections fulfill

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

Proof: Let N be the maximum number such that there is a family of (distinct) subsets $\mathcal{A}_1, \ldots, \mathcal{A}_N \subset \mathcal{M}$ with the desired properties. Let $a \triangleq \lfloor \varepsilon M \rfloor$. 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

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

Defining $S \triangleq a\binom{M}{a-\lceil \lambda a \rceil} 2^a$ we could add another set to our family of subsets if $\binom{M}{a} > N \cdot S$. Therefore:

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

■

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

Proposition 3.2 Let $\lambda_1, \lambda_2, \delta > 0$, let $\lambda \triangleq \min(\lambda_1, {}^{\lambda_2}\!/_2)$, and let $\varepsilon > 0$ be so small that $\lambda \log(\frac{1}{\varepsilon} - 1) > 2$. Then 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 an (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 than $\lambda \lfloor \varepsilon M \rfloor$. Here we have for the number N of those sets:

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

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

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

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

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

and for $i \neq j$:

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

■

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 $\lambda \triangleq \min(\lambda_1, {}^{\lambda_2}/_2)$.

4 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. [5]). Speaking very loosely it concerns the following question: Say that two input distributions are similar if the variational distance $d_1(P,Q) \triangleq \sum_{x\in A} |P(x) - Q(x)|$ of the corresponding output (!) distributions is small. How small may be a set of input distributions 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. Recall that there are certainly very useless measurements (e.g. trivial ones), which 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 be a finite set and let $\mathcal{H}$ be 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 probability distribution 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 be a probability distribution on A. P is M*-type* ($M \in \mathbb{N}$) if $P(x) \in \{0, \frac{1}{M}, \frac{2}{M}, \dots, \frac{M-1}{M}, 1\}$ for all $x \in A$.

Remark 4.3 Obviously, the number of different M-type distributions is upper bounded by $|A|^M$.

Definition 4.4 Let P be a probability distribution 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 $\sigma \in S(\mathcal{H})$ be a state and let $E = (E_m)_{m=1,\dots,M}$ be a POM on a finite dimensional Hilbert space $\mathcal{H}$. This induces a probability distribution $\sigma(E)$ on $\{1,\dots,M\}$ 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 $\mathbf{W}$ be a quantum channel, let $\mathbf{E}$ be a measurement process on its output space, and let $\varepsilon > 0$. We call $R \geq 0$ 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. the input process $\mathbf{P}$, too, 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 $S(\mathbf{P}, \mathbf{E})$.

The following properties are immediate consequences of the definitions:

Remark 4.7 a) If $\varepsilon \leq \varepsilon'$ then $S_\varepsilon(\mathbf{E}) \geq S_{\varepsilon'}(\mathbf{E})$ and $S_\varepsilon(\mathbf{P}, \mathbf{E}) \geq S_{\varepsilon'}(\mathbf{P}, \mathbf{E})$.
b) $S(\mathbf{E}) = \sup_{\varepsilon > 0} S_\varepsilon(\mathbf{E})$ and $S(\mathbf{P}, \mathbf{E}) = \sup_{\varepsilon > 0} S_\varepsilon(\mathbf{P}, \mathbf{E})$.
c) $S(\mathbf{E}) = \sup_{\mathbf{P}} S(\mathbf{P}, \mathbf{E})$ and $S_\varepsilon(\mathbf{E}) = \sup_{\mathbf{P}} S_\varepsilon(\mathbf{P}, \mathbf{E})$ for all $\varepsilon > 0$.

Next we define the notion of uniform resolution rates which will be a useful tool in the proof of Lemma 4.10.

Definition 4.8 Let $\mathbf{W}$ be a quantum channel, let $\mathbf{E}$ be a measurement process on its output space, and let $\varepsilon > 0$. We call $R \geq 0$ 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 $\tilde{\mathbf{P}}$:

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

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

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

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 $n_k \triangleq n_0(\mathbf{P}_k)$ is strictly monotonically increasing (hence divergent). Define a new process $\mathbf{P}$ by

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

Consider the minimum k for which $n_0(\mathbf{P}) < n_k$. Since for $n_0(\mathbf{P}) \leq n < n_k$ we have $P^n = P_k^n$ there exist probability distributions $\tilde{P}^n$ for which $\frac{\log_2 R(\tilde{P}^n)}{n} < R + \gamma$ and $d_{E^n}(P^n W^n, \tilde{P}^n W^n) \ < \ \varepsilon$. By definition of $n_0(\mathbf{P}_k)$ there are such $\tilde{P}^n$ for $n \geq n_k = n_0(\mathbf{P}_k)$, too. This contradicts $n_0(\mathbf{P}_k)$ being chosen as minimum number. ■

Lemma 4.10 Let $\mathbf{W}$ be a quantum channel and let $\mathbf{E}$ be a measurement process on its output space. Moreover, let $(\mathbf{P}_k)_{k\in\mathbb{N}}$ be a sequence of processes, let $(N_n)_{n\in\mathbb{N}}$ be a sequence of positive integers, and let $\varepsilon > 0$ be with $(\forall\, n \in \mathbb{N})$ $\min_{1\leq k<l\leq N_n} d_{E^n}(P_k^n W^n, P_l^n W^n) \geq 2\varepsilon$. Then $\limsup_{n\to\infty} \frac{\log_2 \log_2 N_n}{n} \leq S_\varepsilon(\mathbf{E})$.

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 for all $n \geq n_0$ and $k \in \mathbb{N}$:

$$\frac{\log_2 R(\tilde{P}_k^n)}{n} \ < \ S_\varepsilon(\mathbf{E}) + \gamma \quad \wedge \quad d_{E^n}(P_k^n W^n, \tilde{P}_k^n W^n) \ < \ \varepsilon \ .$$

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

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

Since 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 $|A|^{nM}$ probability distributions on A^n of type M, and as for all $n \geq n_0$ and $1 \leq k \leq N_n$ the resolution of $\tilde{P}_k^n$ is smaller than $2^{n(S_\varepsilon(\mathbf{E})+\gamma)}$:

$$\begin{aligned} N_n \ &\leq \ \sum_{M=1}^{\lfloor 2^{n(S_\varepsilon(\mathbf{E})+\gamma)} \rfloor} |A|^{nM} \ \leq \ 2^{n(S_\varepsilon(\mathbf{E})+\gamma)} |A|^{n2^{n(S_\varepsilon(\mathbf{E})+\gamma)}} \\ &= \ 2^{n(S_\varepsilon(\mathbf{E})+\gamma)+\log_2|A| n 2^{n(S_\varepsilon(\mathbf{E})+\gamma)}} \ \overset{n\gg 1}{<} \ 2^{2^{n(S_\varepsilon(\mathbf{E})+2\gamma)}} \ . \end{aligned}$$

It follows $\frac{\log_2\log_2 N_n}{n} < S_\varepsilon(\mathbf{E}) + 2\gamma$ (for all $\gamma > 0$), and $\limsup_{n\to\infty} \frac{\log_2\log_2 N_n}{n} \leq S_\varepsilon(\mathbf{E})$.

■

Theorem 4.11 Let $\mathbf{W}$ be a quantum channel and let $\lambda_1, \lambda_2 > 0$ be with $\lambda_1 + \lambda_2 < 1$. 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 $i > N_n$. We have for all $1 \leq k < l \leq N_n$:[3]

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

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

So, Lemma 4.10 implies $\bar{C}(\lambda_1, \lambda_2) = \limsup_{n\to\infty} \frac{\log_2 \log_2 N_n}{n} \leq S_\varepsilon(\mathbf{E})$.

∎

Remark 4.12 If in our definitions we replaced 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 some fact we deal with in the next section:

Proof: By Theorem 4.11 we have $\bar{C}^{sim}(\lambda_1, \lambda_2) \leq S_\varepsilon(\mathbf{E})$. 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 obtain $C^{sim}(\lambda_1, \lambda_2) \leq C_1$ (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 (cf. 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 in our definitions we replaced $d_{\mathbf{E}}(\rho, \sigma)$ by $d(\rho, \sigma) \triangleq \max_{\mathbf{E}} d_{\mathbf{E}}(\rho, \sigma)$. This leads to the (natural) bound $C(\lambda_1, \lambda_2) \leq \log_2 |A|$.

5 Transmission Capacity as Upper Bound of Resolution

This section is devoted to the proof of the following theorem. Like in the previous section we will do this following the ideas of [5].

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

We introduce some basic concepts to prove this theorem:

Definition 5.2 Let $\mathbf{A} = (A_n)_{n\in\mathbb{N}}$ be 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 $(\mathbf{P}, \mathbf{E})$ be a pair of processes for $\mathbf{W}$. For every $n \in \mathbb{N}$ let

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

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

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

be 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 $(\mathbf{P}, \mathbf{E})$ be 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 that 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 – an $(n, M_n, 1-\frac{\alpha}{3})$ Q code, with M_n some integer fulfilling

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

As the first inequality contradicts to the definition of $\bar{C}_1$ 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 $\sum_{y \in D_i} E^n_y$ where $D_i \triangleq D(c_i) \setminus \bigcup_{j<i} D(c_j)$.
The success probability is

$$\begin{aligned} W^n_{c_i}\big(\sum_{y \in D_i} E^n_y\big) \;&\geq\; W^n_{c_i}\big(\sum_{y \in D(c_i)} E^n_y\big) - \sum_{j<i} W^n_{c_i}\big(\sum_{y \in D(c_j)} E^n_y\big) \\ &\geq\; \frac{\alpha}{2} - \sum_{j<i} W^n_{c_i}\big(\sum_{y \in D(c_j)} E^n_y\big)\ . \end{aligned}$$

For the expected value of the last summands

$$\begin{aligned}
\mathbb{E}_{Q^n}\, W^n_{c_i}(\sum_{y\in D(c_j)} E^n_y) &\le \frac{1}{P^n(G)} \sum_{x^n\in A^n} P^n(x^n) W^n_{x^n}(\sum_{y\in D(c_j)} E^n_y) \\
&= \frac{1}{P^n(G)} P^n W^n(\sum_{y\in D(c_j)} E^n_y),
\end{aligned}$$

holds, where

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

Here the inequality holds because

$$\begin{aligned}
y\in D(c_j) \quad &\Leftrightarrow \quad \frac{1}{n}\log_2 \frac{W^n_{c_j}(E^n_y)}{P^n W^n(E^n_y)} > \bar{C}_1+\gamma \\
&\Leftrightarrow \quad P^n W^n(E^n_y) < W^n_{c_j}(E^n_y)\cdot 2^{-n(\bar{C}_1+\gamma)}.
\end{aligned}$$

So, we get for the expected success probability

$$\mathbb{E}_{Q^n}\, W^n_{c_i}(\sum_{y\in D_i} E^n_y) \ge \frac{\alpha}{2} - \frac{1}{P^n(G)} M_n\cdot 2^{-n(\bar{C}_1+\gamma)} \ge \frac{\alpha}{2} - \frac{1}{P^n(G)} 2^{-n\frac{\gamma}{2}} .$$

For some $\zeta \ge 0$ and random variable $Z < 1$ clearly $\mathbb{E}_P\, Z \le P[Z\ge\zeta]+\zeta$ holds. Therefore:

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

Hence

$$\mathbb{E}_{Q^n}\, W^n_{c_i}(\sum_{y\in D_i} E^n_y) \ge \frac{\alpha}{2} - \frac{2}{\alpha} 2^{-n\frac{\delta}{2}} \ge \frac{\alpha}{3}$$

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

Lemma 5.5 [cf. [5], p. 758] Let Q and R be probability distributions on a finite set. Then for every $\mu > 0$:

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

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

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

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

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 $(c_1,\dots,c_M) \in (A^n)^M$ of codewords gives rise to the M-type probability distribution

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

We will show that

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

interpreting the $c_1,\dots,c_M$ as independent outcomes of a random experiment with underlying probability distribution P^n. This directly implies our claim. Recall that by Definition 4.5 it follows that $d_{E^n}(P^nW^n, \tilde{P}^n_{(c_1,\dots,c_M)}W^n) = d_1(P^nW^n(E^n), \tilde{P}^n_{(c_1,\dots,c_M)}W^n(E^n))$, 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:

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

Since all M summands are equal this is just

$$\begin{aligned}
&\sum_{c_2\in A^n}\cdots\sum_{c_M\in A^n} P^n(c_2)\cdots P^n(c_M)\\
&\qquad\qquad \sum_{c_1\in A^n}\sum_{y=1}^{b_n} P_{(P^n,E^n)}(c_1,y)\cdot \mathbf{1}\{\log_2\frac{\tilde{P}^n_{(c_1,\ldots,c_M)}W^n(E^n_y)}{P^nW^n(E^n_y)}>\mu\}\\
&\leq P_{(P^n,E^n)}\left[\frac{1}{M}2^{i_{(P^n,E^n)}}>\tau\right] \qquad\qquad (2)\\
&\qquad\qquad + \mathbb{E}_{P^nW^n(E^n)}P_{(P^n)^{M-1}}\left[\frac{1}{M}\sum_{j=2}^{M}2^{i_{(P^n,E^n)}(c_j,y)}>1+\tau\right]\ .
\end{aligned}$$

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:

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

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, since for every $y \in \{1,\ldots,b_n\}$

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

we can apply Chebychev's Inequality:

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

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

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

■

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

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] R. P. Feynman, "Simulating Physics with Computers," *Int. J. Theor. Phys.*, vol. 6/7, pp. 467-488, 1982.

[4] T. S. Han, S. Verdú, "New Results in the Theory of Identification via Channels," *IEEE Trans. Inform. Theory*, vol. 38, pp. 14-25, 1992.

[5] T. S. Han, S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, vol. 39, pp. 752-772, 1993.

[6] 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.

[7] 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).

[8] A. S. Holevo, *"Probabilistic and Statistical Aspects of Quantum Theory,"* North-Holland Publishing-Company, Amsterdam (1982).

[9] 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).

[10] A. S. Holevo, "Coding Theorems for Quantum Channels," LANL Report no. quant-ph/9809023.

[11] Ch. Kleinewächter, "On Identification," Dissertation, Fakultät für Mathematik, Universität Bielefeld, 1999.

[12] P. Löber, "Quantum Channels and Simultaneous ID Coding," Preprint 99-031, SFB 343, Universität Bielefeld, 1999.

[13] T. Ogawa, H. Nagaoka, "Strong Converse to the Quantum Channel Coding Theorem," LANL Report no. quant-ph/9808063.

[14] B. Schumacher, M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol. 56, no. 1, pp. 131-138, 1997.

[15] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J. 27,* pp. 339-425 and 623-656, 1948.

[16] A. Winter, "Coding Theorem and Strong Converse for Quantum Channels," Preprint 98-074, SFB 343, Universität Bielefeld, 1998.

[17] A. Winter, "Coding Theorem and Strong Converse for Nonstationary Quantum Channels," Preprint 99-033, SFB 343, Universität Bielefeld, 1999.

Index