Quantum capacity under adversarial quantum noise: arbitrarily varying quantum channels

Rudolf Ahlswede [1*], Igor Bjelaković [2], Holger Boche [3], Janis Nötzel [2]

Electronic addresses: {igor.bjelakovic, boche, janis.noetzel}@tum.de

[1] Fakultät für Mathematik, Universität Bielefeld,
Universitätsstr. 25, 33615 Bielefeld, Germany

[2] Theoretische Informationstechnik, Technische Universität München,
80291 München, Germany

[3] Lehrstuhl für Theoretische Informationstechnik, Technische Universität München,
80291 München, Germany

February 4, 2011

Abstract

We investigate entanglement transmission over an unknown channel in the presence of a third party (called the adversary), which is enabled to choose the channel from a given set of memoryless but non-stationary channels without informing the legitimate sender and receiver about the particular choice that he made. This channel model is called arbitrarily varying quantum channel (AVQC).
We derive a quantum version of Ahlswede's dichotomy for classical arbitrarily varying channels. This includes a regularized formula for the common randomness-assisted capacity for entanglement transmission of an AVQC. Quite surprisingly and in contrast to the classical analog of the problem involving the maximal and average error probability, we find that the capacity for entanglement transmission of an AVQC always equals its strong subspace transmission capacity.
These results are accompanied by different notions of symmetrizability (zero-capacity conditions) as well as by conditions for an AVQC to have a capacity described by a single-letter formula. In he final part of the paper the capacity of the erasure-AVQC is computed and some light shed on the connection between AVQCs and zero-error capacities. Additionally, we show by entirely elementary and operational arguments motivated by the theory of AVQCs that the quantum, classical, and entanglement-assisted zero-error capacities of quantum channels are generically zero and are discontinuous at every positivity point.

Contents

[*] Tragically, Rudolf Ahlswede passed away during the preparation of the final version of the present paper in December 2010. We, the remaining authors, are thankful to have had the opportunity to experience and enjoy his boundless enthusiasm for science and his lively spirit.

1 Introduction

System uncertainty is a basic feature of many information processing systems, regardless whether classical or quantum mechanical, which has a significant impact on structure and performance of protocols used to cope with limited system knowledge. While in case of quantum communication through unknown quantum channels several important techniques, including channel detection and quantum channel tomography, have been developed to gain at least partial system knowledge, the assumptions needed for these techniques to work satisfactory seem to be rather limiting. Either involved channels have to be stationary (and memoryless) or there has to be additional assistance by a noiseless two-way classical side channel of potentially unlimited capacity or both.
An alternative approach consists of following the successful paradigm of classical information theory according to which one develops the techniques, tailored to clearly specified channel models, for identification of optimal communication parameters, e.g. achievable rates, without making any attempt to reduce system uncertainty. However, channel detection and/or tomography can be seen as auxiliary steps which help to specify the channel model prior to actual information processing. In this paper we follow this information-theoretic strategy, and consider the problem of entanglement transmission over adversarially selected quantum channels.

The basic setup consists of a set of quantum channels $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ which is known to both the sender and receiver. The goal of the sender is to transmit one half of a maximally entangled pure state ψ, suitably encoded, by l-fold usage of the (unknown) channel. An entity, which we call the adversary for simplicity, can choose a sequence $s^l = (s_1, \ldots, s_l) \in \mathbf{S}^l$ at her/his will which results in the selection of the channel $\mathcal{N}_{s^l} = \otimes_{i=1}^{l}\mathcal{N}_{s_i}$. The encoded version of ψ is then fed into $\mathcal{N}_{s^l}$ and the receiver's goal is to recover the input state, of course without knowing the sequence s^l being selected by the adversary. Implicit in this informal description of the communication scenario is that we suppose that the adversary knows the code which is used for entanglement transmission. Therefore, the communicators are forced to use entanglement transmission protocols that are reliable for the whole family $\mathfrak{I}^{(l)} = \{\mathcal{N}_{s^l}\}_{s^l\in\mathbf{S}^l}$ of memoryless and partly non-stationary channels. In other words, the desired entanglement transmission protocol should be resistant to the effect of arbitrarily varying noise represented by the family $\mathfrak{I}^{(l)} = \{\mathcal{N}_{s^l}\}_{s^l\in\mathbf{S}^l}$. Even in the simplest non-trivial case of a finite set $\mathfrak{I}$ with $|\mathfrak{I}| > 1$ we have to deal for each block length l with exponentially many quantum channels simultaneously.

The main contribution of this paper is a generalization of Ahlswede's dichotomy [2] which can be stated as follows:

First, the common-randomness-assisted quantum capacity of the AVQC $(\mathfrak{I}^{(l)})_{l\in\mathbb{N}}$ is equal to the entanglement transmission capacity of the compound channel built up from conv($\mathfrak{I}$), i.e. the uncountable family of stationary, memoryless channels that lie in the convex hull of $\mathfrak{I}$ (cf. [9], [10] for more information on compound quantum channels).

Second, if the deterministic capacity for transmission of messages with asymptotically vanishing average error over an AVQC is greater than zero, its capacity for transmission of entanglement with deterministic codes is equal to its common-randomness-assisted capacity for transmission of entanglement.

The proof of the direct part as well as the proof of the converse rely substantially on the corresponding results for compound quantum channels developed in [9], [10]. The link between the compound and arbitrarily varying channel models needed in the achievability proofs is given by the powerful robustification technique of [3] and [4].

The idea behind the second part of the theorem is the following. If the deterministic capacity for message transmission, with average error probability as the success criterion, of $\mathfrak{I}$ is greater than zero, then sender and receiver can use a few (sub-exponentially many) bits to derandomize a given common-randomness-assisted code for transmission of entanglement.

As a supplement to the coding theorem, we derive a multi-letter necessary and sufficient condition for the deterministic capacity, with average error, for message transmission of a (finite) AVQC to be zero in Section 8. For sake of completeness, we also include a necessary and sufficient condition for the deterministic capacity for message transmission with *maximal* error probability to be equal to zero. Moreover, we present a first attempt to derive a non-trivial sufficient condition for the common-randomness-assisted capacity for transmission of entanglement to be zero, which we call qc-symmetrizability. Our feeling in this matter is that the definition of that kind of symmetrizability is too narrow to have any chance to be necessary and sufficient. This is basically because according to that definition the adversary does not use all the freedom he is given by the channel model to prevent the common-randomness-assisted entanglement transmission.

We find a striking difference to the classical theory: entanglement transmission with entanglement fidelity as the criterion of success is widely acknowledged as a fully quantum counterpart to message transmission with average error as a criterion for failure of transmission, while the counterpart of strong subspace transmission should be maximal error probability.

The two classical criteria have been proven to be asymptotically equivalent e.g. for single memoryless channels. For transmission over an AVC they lead to different capacities, as can be seen from Example 2 in [2]. The AVC given there has zero capacity for message transmission with asymptotically vanishing maximal error probability, but from Theorem 3, part a) it can be seen that it has positive capacity for message transmission with asymptotically vanishing average error.

In the quantum case, asymptotic equivalence of entanglement and strong subspace transmission for single quantum channels has already been proven in [7]. Our results show, that they are - in contrast to the classical theory - also (asymptotically) equivalent criteria w.r.t. AVQCs.

It is no surprise then, that the connection between arbitrarily varying channels and zero-error capacities that is valid in the classical case [1] only partly survives in the quantum regime. This connection is explored in the last part of the paper. Additionally, we show that quantum, classical, and entanglement-assisted zero-error capacities of quantum channels are generically zero and are discontinuous at every positivity point. This is obvious for the classical zero-error capacity in Shannon's original setting [32]. In the quantum case we employ some simple facts from convex geometry combined with methods motivated by the theory of arbitrarily varying channels to obtain this conclusion in an extremely simple way directly from the corresponding definitions of zero-error quantum capacities. It should be mentioned at this point that these results can as well be obtained rather easily using the concept of non-commutative graphs (again accompanied by some convex geometry) that has been systematically explored in the recent work [16]. The fact that the quantum zero-error capacity is generically zero shows that the channels for which it is possible to satisfy the Knill-Laflamme condition [24] on a subspace of dimension greater or equal than 2 are exceptional.

We also list two properties that lead to a single-letter capacity formula of an AVQC and compute the (deterministic) entanglement transmission capacity of an erasure AVQC.

1.1 Related Work

The model of an arbitrarily varying channel has been introduced by Blackwell, Breiman and Thomasian [11] in 1960. They derived a formula for the capacity of an AVC with random codes and asymptotically vanishing average error probability. They also wrote down an explicit example of an AVC whose deterministic capacity is zero, while having nonzero capacity when using random codes.

Later landmarks in the development of coding theorems for AVCs have been the papers by Kiefer and Wolfowitz [22], who found a necessary and sufficient condition for an AVC to have nonzero capacity with deterministic codes and asymptotically vanishing maximal error probability.

The maximal error probability criterion was further investigated in [6] by Ahlswede and Wolfowitz, who completely determined the capacity of AVCs with binary output alphabet under that criterion. A solution for arbitrarily large alphabets does not seem to exist until now. It should be mentioned that such a solution would include the solution to Shannon's zero error capacity problem [32], as pointed out in [1].

In our approach we use the powerful elimination technique developed by the first author in 1978 [2] that, together with the random coding results of [11] enabled him to prove the following dichotomy result for AVCs: It stated that the capacity of an AVC (under the average error probability criterion) is either zero or equals its random coding capacity. Together with the robustificaion technique [3, 4] of the first author, the elimination technique led to a rather streamlined approach that, in this combination, has first been successfully used by Ahlswede in [4].

After the discoveries of [2], an important open question was, when exactly the deterministic capacity with vanishing average error is equal to zero. In 1985, a first step towards a solution was made by Ericson [17], who came up with a sufficient condition that was proven to be necessary by Csiszar and Narayan [14] in 1989.

The model of an arbitrarily varying channel with classical input and quantum output has first been considered in 2007 by the first author together with Blinovsky [5]. They considered the transmission of messages under the average error criterion and gave a complete solution of the problem, i.e. a single-letter capacity formula, including a necessary and sufficient condition for the case of zero capacity.

1.2 Outline

The notation we freely use throughout the paper is summarized in Section 2. The definitions of codes and capacities that are needed in the sequel are given in Section 3. This section also contains our main result, a quantum version of Ahlswede's dichotomy.

A perhaps surprising result is proven in Section 4: As can be seen from an application of a concentration inequality, the capacities for entanglement and strong subspace transmission are identical. This is in sharp contrast to the classical case, where their analogs - average and maximal error criterion - lead to different capacities [2].

The main part of the paper, up to Section 8, is mostly devoted to the proof of the main result and is organized as follows.

In Section 5 we are concerned with the upper bound to the common-randomness-assisted capacity for entanglement transmission, i.e. with the converse part. Here, the basic problem is that we cannot employ the Minimax-Theorem (which has originally been proven by von Neumann [33] and later put into a more general context by Kakutani [21]) like in the classical case to reduce the converse part to that of a single channel. We circumvent this obstacle by noting that the desired result follows from the optimal upper bound on the common-randomness-assisted capacity for entanglement transmission for *compound* quantum channels. The latter is easily shown using the methods from [10].

Section 6 contains the achievability proofs for $\mathcal{A}_{\text{random}}(\mathfrak{I})$. As we already mentioned above, we are in the pleasant situation of having at our disposal the coding results for compound quantum channels from [10] and the robustification technique from [3, 4]. Since the latter is a central tool for our results and because there is a short and simple proof of it in [4], we have decided to include the full account of the robustification technique. The technique on its own operates as follows: We start with a "good" code for the compound quantum channel built up from $\text{conv}(\mathfrak{I})$, then applying permutations to the encoding and decoding operations of that code we obtain a "good" random code for the AVQC. The source of common randomness now helps coordinating the selection of permutations at sender's and receiver's side.

The last part of our main theorem is proven in Section 7. The first step in the proof is to show that not that much common randomness is needed to achieve $\mathcal{A}_{\text{random}}(\mathfrak{I})$. Basically for each block length l we need roughly $O(\log l)$ random bits. This is shown by a slight modification of the elimination technique of [2]. If we assume now that $C_{\text{det}}(\mathfrak{I}) > 0$ holds, the sender and receiver can generate the required $O(\log l)$ random bits by sending (classical) messages over the AVQC by spending negligible block length compared to l and are therefore able to simulate reliable random codes by deterministic ones.

Section 8 summarizes attempts to address the question when exactly a given (finite) AVQC has a capacity (for various types of transmission and criteria of success) equal to zero. Most important for our present work is a necessary and sufficient condition for the message transmission capacity with deterministic codes and average error criterion to be greater than zero. Together with our results on random entanglement transmission codes it enables us to prove our main theorem. We also give a necessary and sufficient condition for the classical deterministic capacity with maximal error criterion to be greater than zero and end the section with a first attempt to find non-trivial conditions for the common-randomness-assisted capacity for transmission of entanglement of an AVQC to be equal to zero.

Section 9 is devoted to single letter characterizations of $\mathcal{A}_{\text{det}}(\mathfrak{I})$. We give two conditions that lead to single letter formulas, both of which demand certain properties to be valid all over the convex hull $\text{conv}(\mathfrak{I})$ of $\mathfrak{I}$.

Finally in Section 10 we explicitly calculate the capacities for entanglement transmission of the erasure-AVQC and exploit the connection between AVQC's and zero-error capacities.

2 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. If $\mathcal{F} \subset \mathcal{H}$ is a subspace of $\mathcal{H}$ then we write $\pi_{\mathcal{F}}$ for the maximally mixed state on $\mathcal{F}$, i.e. $\pi_{\mathcal{F}} = \frac{p_{\mathcal{F}}}{\text{tr}(p_{\mathcal{F}})}$ where $p_{\mathcal{F}}$ stands for the projection onto $\mathcal{F}$. $\mathcal{B}(\mathcal{H})$ denotes the set of linear operators acting on $\mathcal{H}$. For a finite set A the notation $\mathfrak{P}(A)$ is reserved for the set of probability distributions on A.

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 $\mathcal{C}(\mathcal{H}, \mathcal{K})$. 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

$$I_c(\rho, \mathcal{N}) := S(\mathcal{N}(\rho)) - S((id_{\mathcal{B}(\mathcal{H})} \otimes \mathcal{N})(|\psi\rangle\langle\psi|)),$$

where $\psi \in \mathcal{H} \otimes \mathcal{H}$ is an arbitrary purification of the state ρ. Following the usual conventions we let $S_e(\rho, \mathcal{N}) := S((id_{\mathcal{B}(\mathcal{H})} \otimes \mathcal{N})(|\psi\rangle\langle\psi|))$ denote the entropy exchange.

As a measure of closeness between two states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ we use the fidelity $F(\rho, \sigma) := ||\sqrt{\rho}\sqrt{\sigma}||_1^2$. The fidelity is symmetric in the input and for a pure state $\rho = |\phi\rangle\langle\phi|$ we have $F(|\phi\rangle\langle\phi|, \sigma) = \langle \phi, \sigma\phi\rangle$.

A closely related quantity is the entanglement fidelity. For $\rho \in \mathcal{S}(\mathcal{H})$ and $\mathcal{N} \in \mathcal{C}^{\downarrow}(\mathcal{H}, \mathcal{K})$ it is given by

$$F_e(\rho, \mathcal{N}) := \langle \psi, (id_{\mathcal{B}(\mathcal{H})} \otimes \mathcal{N})(|\psi\rangle\langle\psi|)\psi\rangle,$$

with $\psi \in \mathcal{H} \otimes \mathcal{H}$ being an arbitrary purification of the state ρ.

We use the diamond norm $|| \cdot ||_\lozenge$ as a measure of closeness in the set of quantum channels, which is given by

$$||\mathcal{N}||_\lozenge := \sup_{n \in \mathbb{N}} \max_{a \in \mathcal{B}(\mathbb{C}^n \otimes \mathcal{H}), ||a||_1 = 1} ||(id_n \otimes \mathcal{N})(a)||_1, \tag{1}$$

where $id_n : \mathcal{B}(\mathbb{C}^n) \to \mathcal{B}(\mathbb{C}^n)$ is the identity channel, and $\mathcal{N} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ is any linear map, not necessarily completely positive. The merits of $|| \cdot ||_\lozenge$ are due to the following facts (cf. [23]). First, $||\mathcal{N}||_\lozenge = 1$ for all $\mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$. Thus, $\mathcal{C}(\mathcal{H}, \mathcal{K}) \subset S_\lozenge$, where $S_\lozenge$ denotes the unit sphere of the normed space $(\mathcal{B}(\mathcal{B}(\mathcal{H}), \mathcal{B}(\mathcal{K})), || \cdot ||_\lozenge)$. Moreover, $||\mathcal{N}_1 \otimes \mathcal{N}_2||_\lozenge = ||\mathcal{N}_1||_\lozenge ||\mathcal{N}_2||_\lozenge$ for arbitrary linear maps $\mathcal{N}_1, \mathcal{N}_2 : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$. Finally, the supremum in (1) needs only be taken over n that range over $\{1, 2, \ldots, \dim \mathcal{H}\}$.

We further use the diamond norm to define the function $D_\lozenge(\cdot, \cdot)$ on $\{(\mathfrak{I}, \mathfrak{I}') : \mathfrak{I}, \mathfrak{I}' \subset \mathcal{C}(\mathcal{H}, \mathcal{K})\}$, which is for $\mathfrak{I}, \mathfrak{I}' \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ given by

$$D_\lozenge(\mathfrak{I}, \mathfrak{I}') := \max\{\sup_{\mathcal{N} \in \mathfrak{I}} \inf_{\mathcal{N}' \in \mathfrak{I}'} ||\mathcal{N} - \mathcal{N}'||_\lozenge, \sup_{\mathcal{N}' \in \mathfrak{I}'} \inf_{\mathcal{N} \in \mathfrak{I}} ||\mathcal{N} - \mathcal{N}'||_\lozenge\}.$$

For $\mathfrak{I} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ let $\bar{\mathfrak{I}}$ denote the closure of $\mathfrak{I}$ in $|| \cdot ||_\lozenge$. Then $D_\lozenge$ defines a metric on $\{(\mathfrak{I}, \mathfrak{I}') : \mathfrak{I}, \mathfrak{I}' \subset \mathcal{C}(\mathcal{H}, \mathcal{K}),\ \mathfrak{I} = \bar{\mathfrak{I}},\ \mathfrak{I}' = \bar{\mathfrak{I}}'\}$ which is basically the Hausdorff distance induced by the diamond norm.

Obviously, for arbitrary $\mathfrak{I}, \mathfrak{I}' \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$, $D_\lozenge(\mathfrak{I}, \mathfrak{I}') \le \epsilon$ implies that for every $\mathcal{N} \in \mathfrak{I}$ ($\mathcal{N}' \in \mathfrak{I}'$) there exists $\mathcal{N}' \in \mathfrak{I}'$ ($\mathcal{N} \in \mathfrak{I}$) such that $||\mathcal{N} - \mathcal{N}'||_\lozenge \le 2\epsilon$. If $\mathfrak{I} = \bar{\mathfrak{I}},\ \mathfrak{I}' = \bar{\mathfrak{I}}'$ holds we even have $||\mathcal{N} - \mathcal{N}'||_\lozenge \le \epsilon$. In this way $D_\lozenge$ gives a measure of distance between sets of channels.

For any set $\mathfrak{I} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ and $l \in \mathbb{N}$ we set

$$\mathfrak{I}^{\otimes l} := \{\mathcal{N}^{\otimes l} : \mathcal{N} \in \mathfrak{I}\}.$$

For an arbitrary set $\mathbf{S}$, $\mathbf{S}^l := \{(s_1, \ldots, s_l) : s_i \in \mathbf{S}\ \forall i \in \{1, \ldots, l\}\}$. We also write s^l for the elements of $\mathbf{S}^l$.

For an arbitrary set $\mathfrak{I}$ of CPTP maps we denote by $\mathrm{conv}(\mathfrak{I})$ its convex hull (see [34] for the definition) and note that in case that $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$ is a finite set we have

$$\mathrm{conv}(\mathfrak{I}) = \left\{ \mathcal{N}_q \in \mathcal{C}(\mathcal{H}, \mathcal{K}) : \mathcal{N}_q = \sum_{s \in \mathbf{S}} q(s)\mathcal{N}_s,\ q \in \mathfrak{P}(\mathbf{S}) \right\}, \tag{2}$$

an equality that we will make use of in the approximation of infinite AVQC's by finite ones.

Finally, we need some simple topological notions for convex sets in finite dimensional normed space $(V, ||\cdot||)$ over the field of real or complex numbers which we borrow from [34]. Let $F \subset V$ be convex. $x \in F$ is said to be a relative interior point of F if there is $r > 0$ such that $B(x, r) \cap \mathrm{aff}\, F \subset F$. Here $B(x, r)$ denotes the open ball of radius r with the center x and $\mathrm{aff}\, F$ stands for the affine hull of F. The set of relative interior points of F is called the relative interior of F and is denoted by $\mathrm{ri}\, F$.

The relative boundary of F, $\mathrm{rebd}\, F$, is the set difference between the closure of F and $\mathrm{ri}\, F$.

For a set $A \subset V$ and $\delta \ge 0$ we define the parallel set or the blow-up $(A)_\delta$ of A by

$$(A)_\delta := \{x \in V : ||x - y|| \le \delta \text{ for some } y \in A\}.$$

3 Basic definitions and main results

In this section we define the quantities that we will be dealing with in the rest of the paper: Arbitrarily varying quantum channels and codes for transmission of entanglement and subspaces. Since they will be of importance for our derandomization arguments, we will also include definitions of the capacities for message transmission with average and maximal error probability criterion.
Our most basic object is the arbitrarily varying quantum channel (AVQC). It is generated by a set $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ of CPTP maps with input Hilbert space $\mathcal{H}$ and output Hilbert space $\mathcal{K}$ and given by the family of CPTP maps $\{\mathcal{N}_{s^l} : \mathcal{B}(\mathcal{H})^{\otimes l} \to \mathcal{B}(\mathcal{K})^{\otimes l}\}_{l\in\mathbb{N}, s^l\in\mathbf{S}^l}$, where

$$\mathcal{N}_{s^l} := \mathcal{N}_{s_1} \otimes \ldots \otimes \mathcal{N}_{s_l} \qquad (s^l \in \mathbf{S}^l).$$

Thus, even in the case of a finite set $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$, showing the existence of reliable codes for the AVQC determined by $\mathfrak{I}$ is a non-trivial task: For each block length $l \in \mathbb{N}$ we have to deal with $|\mathfrak{I}|^l$, i.e. exponentially many, memoryless partly non-stationary quantum channels simultaneously. In order to relieve ourselves of the burden of complicated notation we will simply write $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ for the AVQC.

3.1 Entanglement transmission

For the rest of this subsection, let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be an AVQC.

Definition 1. *An $(l, k_l)-$*random entanglement transmission code *for $\mathfrak{I}$ is a probability measure μ_l on $(\mathcal{C}(\mathcal{F}_l, \mathcal{H}) \times \mathcal{C}(\mathcal{K}, \mathcal{F}_l'), \sigma_l)$, where $\dim \mathcal{F}_l = k_l$, $\mathcal{F}_l \subset \mathcal{F}_l'$ and the sigma-algebra σ_l is chosen such that $F_e(\pi_{\mathcal{F}_l}, (\cdot) \circ \mathcal{N}_{s^l} \circ (\cdot))$ is measurable w.r.t. σ_l for every $s^l \in \mathbf{S}^l$. Moreover, we assume that σ_l contains all singleton sets. An example of such a sigma-algebra σ_l is given by the product of sigma-algebras of Borel sets induced on $\mathcal{C}(\mathcal{F}_l, \mathcal{H})$ and $\mathcal{C}(\mathcal{K}, \mathcal{F}_l')$ by the standard topologies of the ambient spaces.*

Definition 2. *A non-negative number R is said to be an achievable entanglement transmission rate for the AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ with random codes if there is a sequence of $(l, k_l)-$random entanglement transmission codes such that*

1. $\liminf_{l\to\infty} \frac{1}{l} \log k_l \geq R$ *and*
2. $\lim_{l\to\infty} \inf_{s^l\in\mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) = 1.$

The random entanglement transmission capacity $\mathcal{A}_{\text{random}}(\mathfrak{I})$ of $\mathfrak{I}$ is defined by

$$\mathcal{A}_{\text{random}}(\mathfrak{I}) := \sup\{R : R \textit{ is an achievable entanglement transmission rate for } \mathfrak{I} \textit{ with random codes}\}.$$

Having defined random codes and random code capacity for entanglement transmission we are in the position to introduce their deterministic counterparts: An $(l, k_l)-$code for entanglement transmission over $\mathfrak{I}$ is an $(l, k_l)-$random code for $\mathfrak{I}$ with $\mu_l(\{(\mathcal{P}^l, \mathcal{R}^l)\}) = 1$ for some encoder-decoder pair $(\mathcal{P}^l, \mathcal{R}^l)$[1] and $\mu_l(A) = 0$ for any $A \in \sigma_l$ with $(\mathcal{P}^l, \mathcal{R}^l) \notin A$. We will refer to such measures as point measures in what follows.

Definition 3. *A non-negative number R is a deterministically achievable entanglement transmission rate for the AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ if it is achievable in the sense of Definition 2 for random codes with* point measures *μ_l.*
The deterministic entanglement transmission capacity $\mathcal{A}_{\text{det}}(\mathfrak{I})$ of $\mathfrak{I}$ is given by

$$\mathcal{A}_{\text{det}}(\mathfrak{I}) := \sup\{R : R \textit{ is a deterministically achievable entanglement transmission rate for } \mathfrak{I}\}.$$

[1] This explains our requirement on σ_l to contain all singleton sets.

Finally, we shall need the notion of the classical deterministic capacity $C_{\mathrm{det}}(\mathfrak{I})$ of the AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ with *average* error criterion.

Definition 4. *An (l, M_l)-(deterministic) code for message transmission is a family of pairs $\mathfrak{C}_l = (\rho_i, D_i)_{i=1}^{M_l}$ where $\rho_1, \dots, \rho_{M_l} \in \mathcal{S}(\mathcal{H}^{\otimes l})$, and positive semi-definite operators $D_1, \dots, D_{M_l} \in \mathcal{B}(\mathcal{K}^{\otimes l})$ satisfying $\sum_{i=1}^{M_l} D_i = \mathbf{1}_{\mathcal{K}^{\otimes l}}$.*
The worst-case average probability of error of a code $\mathfrak{C}_l$ is given by

$$\bar{P}_{e,l}(\mathfrak{I}) := \sup_{s^l\in\mathbf{S}^l} \bar{P}_e(\mathfrak{C}_l, s^l), \tag{3}$$

where for $s^l \in \mathbf{S}^l$ we set

$$\bar{P}_e(\mathfrak{C}_l, s^l) := \frac{1}{M_l}\sum_{i=1}^{M_l}(1 - \mathrm{tr}(\mathcal{N}_{s^l}(\rho_i)D_i)).$$

The achievable rates and the classical deterministic capacity $C_{det}(\mathfrak{I})$ of $\mathfrak{I}$, with respect to the error criterion given in (3), are then defined in the usual way.

For any AVQC finite or infinite, the compound quantum channel generated by the *infinite* set conv$(\mathfrak{I})$ (cf. [10] for the relevant definitions) shall play the crucial role in our derivation of the coding results below. Our main result, a quantum version of Ahlswede's dichotomy for finite AVQCs, goes as follows:

Theorem 1. *Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be an AVQC.*

1. *With conv$(\mathfrak{I})$ denoting the convex hull of $\mathfrak{I}$ we have*

$$\mathcal{A}_{\mathrm{random}}(\mathfrak{I}) = \lim_{l\to\infty}\frac{1}{l}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})}\inf_{\mathcal{N}\in\mathrm{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes l}). \tag{4}$$

2. *Either $C_{\mathrm{det}}(\mathfrak{I}) = 0$ or else $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) = \mathcal{A}_{\mathrm{random}}(\mathfrak{I})$.*

Proof. The claim made in (4) follows from Theorem 6 and Corollary 10.
The proof that $C_{\mathrm{det}}(\mathfrak{I}) > 0$ implies $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) = \mathcal{A}_{\mathrm{random}}(\mathfrak{I})$ requires a derandomization argument which is presented in Section 7. □

We conclude this section with some explaining remarks:
1. It is clear that $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) \le C_{\mathrm{det}}(\mathfrak{I})$, so that $C_{\mathrm{det}}(\mathfrak{I}) = 0$ implies $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) = 0$. Therefore, Theorem 1 gives a regularized formula for $\mathcal{A}_{\mathrm{det}}(\mathfrak{I})$ in form of (4), and the question remains when $C_{\mathrm{det}}(\mathfrak{I}) = 0$ happens. We derive a non-single-letter necessary and sufficient condition for the latter in Section 8.
2. Continuous dependence of the coherent information on the channel reveals that for each $l \in \mathbb{N}$ and $\rho \in \mathcal{S}(\mathcal{H}^{\otimes l})$

$$\inf_{\mathcal{N}\in\mathrm{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes l}) = \min_{\mathcal{N}\in\tilde{\mathfrak{I}}} I_c(\rho, \mathcal{N}^{\otimes l}),$$

where

$$\tilde{\mathfrak{I}} := \overline{\mathrm{conv}(\mathfrak{I})}^{||\cdot||_\diamond}$$

is the closure of conv$(\mathfrak{I})$ with respect to $||\cdot||_\diamond$, a fact that we will use repeatedly throughout the paper.

3.2 Strong subspace transmission

Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be an AVQC. An $(l, k_l)-$*random strong subspace transmission code* for $\mathfrak{I}$ is a probability measure μ_l on $(\mathcal{C}(\mathcal{F}_l, \mathcal{H}) \times \mathcal{C}(\mathcal{K}, \mathcal{F}_l'), \sigma_l)$, where $\dim \mathcal{F}_l = k_l$, $\mathcal{F}_l \subset \mathcal{F}_l'$ and the sigma-algebra σ_l is chosen such that $F_e(\pi_{\mathcal{F}_l}, (\cdot)\circ\mathcal{N}_{s^l}\circ(\cdot))$ is measurable w.r.t. σ_l for every $s^l \in \mathbf{S}^l$. Again, we assume that σ_l contains all singleton sets.

Definition 5. *A non-negative number R is said to be an achievable strong subspace transmission rate for the AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ with random codes if there is a sequence of $(l, k_l)-$random strong subspace transmission codes such that*

1. $\liminf_{l\to\infty} \frac{1}{l}\log k_l \geq R$ *and*
2. $\lim_{l\to\infty} \inf_{s^l\in\mathbf{S}^l} \min_{\psi\in\mathcal{F}_l} \int F(|\psi\rangle\langle\psi|, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l(|\psi\rangle\langle\psi|))d\mu_l(\mathcal{P}^l, \mathcal{R}^l) = 1.$

The random strong subspace transmission capacity $\mathcal{A}_{\mathrm{s,random}}(\mathfrak{I})$ of $\mathfrak{I}$ is defined by

$$\mathcal{A}_{\mathrm{s,random}}(\mathfrak{I}) := \sup\{R : R \textit{ is an achievable strong subspace transmission rate for } \mathfrak{I} \textit{ with random codes}\}.$$

As before we also define deterministic codes: A *deterministic* $(l, k_l)-$strong subspace transmission code for $\mathfrak{I}$ is an $(l, k_l)-$random strong subspace transmission code for $\mathfrak{I}$ with $\mu_l(\{(\mathcal{P}^l, \mathcal{R}^l)\}) = 1$ for some encoder-decoder pair $(\mathcal{P}^l, \mathcal{R}^l)$ and $\mu_l(A) = 0$ for any $A \in \sigma_l$ with $(\mathcal{P}^l, \mathcal{R}^l) \notin A$. We will refer to such measures as point measures in what follows.

Definition 6. *A non-negative number R is a deterministically achievable strong subspace transmission rate for the AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ if it is achievable in the sense of Definition 5 for random codes with* point measures μ_l.
The deterministic capacity $\mathcal{A}_{\mathrm{s,det}}(\mathfrak{I})$ for strong subspace transmission over an AVQC $\mathfrak{I}$ is given by

$$\mathcal{A}_{\mathrm{s,det}}(\mathfrak{I}) := \sup\{R : R \textit{ is a deterministically achievable rate for } \mathfrak{I}\}.$$

If we want to transmit classical messages, then the error criterion that is most closely related to strong subspace transmission is that of maximal error probability. It leads to the notion of classical deterministic capacity with maximal error:

Definition 7. *Let $\mathfrak{C}_l$ be an (l, M_l)-(deterministic) code for message transmission as given in Definition 4. The worst-case maximal probability of error of the code $\mathfrak{C}_l$ is given by*

$$P_{e,l}(\mathfrak{I}) := \sup_{s^l\in\mathbf{S}^l} P_e(\mathfrak{C}_l, s^l), \tag{5}$$

where for $s^l \in \mathbf{S}^l$ we set

$$P_e(\mathfrak{C}_l, s^l) := \max_{i\in M_l}\left(1 - \mathrm{tr}(\mathcal{N}_{s^l}(\rho_i)D_i)\right).$$

The achievable rates and the classical deterministic capacity $C_{\mathrm{det,max}}(\mathfrak{I})$ of $\mathfrak{I}$, with respect to the error criterion given in (5), are then defined in the usual way.

The perhaps surprising result is that the strong subspace transmission capacity of a (finite) AVQC always equals its entanglement transmission capacity:

Theorem 2. *For every AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ we have the equalities*

$$\mathcal{A}_{\mathrm{s,random}}(\mathfrak{I}) = \mathcal{A}_{\mathrm{random}}(\mathfrak{I}), \tag{6}$$

$$\mathcal{A}_{\mathrm{s,det}}(\mathfrak{I}) = \mathcal{A}_{\mathrm{det}}(\mathfrak{I}). \tag{7}$$

3.3 Zero-error capacities

In this subsection we only give definitions of zero-error capacities. Through the ideas of [1] these capacities are connected to arbitrarily varying channels, though this connection is not as strong as in the classical setting.
Results concerning these capacities are stated in subsections 10.2 and 10.3.

Definition 8. *An (l,k) zero-error quantum code (QC for short) $(\mathcal{F},\mathcal{P},\mathcal{R})$ for $\mathcal{N}\in\mathcal{C}(\mathcal{H},\mathcal{K})$ consists of a Hilbert space $\mathcal{F}$, $\mathcal{P}\in\mathcal{C}(\mathcal{F},\mathcal{H}^{\otimes l})$, $\mathcal{R}\in\mathcal{C}(\mathcal{K}^{\otimes l},\mathcal{F})$ with $\dim\mathcal{F}=k$ such that*

$$\min_{x\in\mathcal{F},||x||=1}\langle x,\mathcal{R}\circ\mathcal{N}^{\otimes l}\circ\mathcal{P}(|x\rangle\langle x|)x\rangle=1. \tag{8}$$

For fixed block length $l\in\mathbb{N}$ define

$$k(l,\mathcal{N}):=\max\{\dim\mathcal{F}:\exists(l,k)\ \textit{zero-error QC for}\ \mathcal{N}\}. \tag{9}$$

The zero-error quantum capacity $Q_0(\mathcal{N})$ of $\mathcal{N}\in\mathcal{C}(\mathcal{H},\mathcal{K})$ is then defined by

$$Q_0(\mathcal{N}):=\lim_{l\to\infty}\frac{1}{l}\log k(l,\mathcal{N}). \tag{10}$$

The existence of the limit follows from standard arguments based on Fekete's Lemma.

Next we pass to the zero-error classical capacities of quantum channels.

Definition 9. *Let $\sigma_{\mathcal{F}\mathcal{F}'}$ be a bipartite state on $\mathcal{F}\otimes\mathcal{F}'$ where $\mathcal{F}'$ denotes a unitary copy of the Hilbert space $\mathcal{F}$. An (l,M) entanglement assisted code (ea-code for short) $(\sigma_{\mathcal{F}\mathcal{F}'},\{\mathcal{P}_m,D_m\}_{m=1}^M)$ consists of a bipartite state $\sigma_{\mathcal{F}\mathcal{F}'}$, $\mathcal{P}_m\in\mathcal{C}(\mathcal{F},\mathcal{H}^{\otimes l})$, $m=1,\ldots,M$, and a POVM $\{D_m\}_{m=1}^M$ on $\mathcal{F}'\otimes\mathcal{K}^{\otimes l}$. A given (l,M) entanglement assisted code $(\sigma_{\mathcal{F}\mathcal{F}'},\{\mathcal{P}_m,D_m\}_{m=1}^M)$ is a zero-error code for $\mathcal{N}\in\mathcal{C}(\mathcal{H},\mathcal{K})$ if*

$$\operatorname{tr}((\mathcal{N}^{\otimes l}\circ\mathcal{P}_m\otimes id_{\mathcal{F}'})(\sigma_{\mathcal{F}\mathcal{F}'})D_m)=1 \tag{11}$$

holds for all $m\in[M]:=\{1,\ldots,M\}$. For $l\in\mathbb{N}$ we set

$$M_{EA}(l,\mathcal{N}):=\max\{M:\exists\ \ \textit{zero-error}\ (l,M)\ \textit{ea-code for}\ \mathcal{N}\}. \tag{12}$$

Definition 10. *The entanglement assisted classical zero-error capacity $C_{0EA}(\mathcal{N})$ of $\mathcal{N}\in\mathcal{C}(\mathcal{H},\mathcal{K})$ is given by*

$$C_{0EA}(\mathcal{N}):=\lim_{l\to\infty}\frac{1}{l}\log M(l,\mathcal{N}). \tag{13}$$

If we restrict the definition of zero-error ea-code to states $\sigma_{\mathcal{F}\mathcal{F}'}$ with $\dim\mathcal{F}=\dim\mathcal{F}'=1$ we obtain the perfomance parameter $M(l,\mathcal{N})$ as a special case of $M_{\mathrm{EA}}(l,\mathcal{N})$ in (12) and the classical zero-error capacity $C_0(\mathcal{N})$ of a quantum channel $\mathcal{N}$.

Definition 11. *Given a bipartite state $\rho\in\mathcal{S}(\mathcal{H}_A\otimes\mathcal{H}_B)$. An (l,k_l) zero-error entanglement distillation protocol (EDP for short) for ρ consists of an LOCC operation $\mathcal{D}\in\mathcal{C}(\mathcal{H}_A^{\otimes l}\otimes\mathcal{H}_B^{\otimes l},\mathbb{C}^{k_l}\otimes\mathbb{C}^{k_l})$ and a maximally entangled state vector $\varphi_{k_l}=\frac{1}{\sqrt{k_l}}\sum_{i=1}^{k_l}e_i\otimes e_i\in\mathbb{C}^{k_l}\otimes\mathbb{C}^{k_l}$ with an orthonormal basis $\{e_1,\ldots,e_{k_l}\}$ of $\mathbb{C}^{k_l}$ such that*

$$\langle\varphi_{k_l},\mathcal{D}(\rho^{\otimes l})\varphi_{k_l}\rangle=1. \tag{14}$$

Let for $l\in\mathbb{N}$

$$d(l,\rho):=\max\{k_l:\exists(l,k_l)\ \textit{zero-error EDP for}\ \rho\}, \tag{15}$$

and we define the zero-error distillable entanglement of $\rho\in\mathcal{S}(\mathcal{H}_A\otimes\mathcal{H}_B)$ as

$$D_0(\rho):=\lim_{l\to\infty}\frac{1}{l}\log d(l,\rho). \tag{16}$$

4 Equivalence of strong subspace and entanglement transmission

We will now use results from convex high-dimensional geometry to show that every sequence of asymptotically perfect (random) codes for entanglement transmission for $\mathfrak{I}$ yields another sequence of (random) codes that guarantees asymptotically perfect strong subspace transmission.
First, we state the following theorem which is the complex version of a theorem that can essentially be picked up from [29], Theorem 2.4 and Remark 2.7:

Theorem 3. *For $\delta, \Theta > 0$ and an integer n let $k(\delta, \Theta, n) = \lfloor \delta^2(n-1)/(2\log(4/\Theta)) \rfloor$. Let $f : S(\mathbb{C}^n) \to \mathbb{R}$ be a continuous function and ν_k the uniform measure induced on the Grassmannian $G_{n,k}$ by the normalized Haar measure on the unitary group on $\mathbb{C}^n$ then, for all $\delta, \Theta > 0$, the measure of the set $E_k \subset G_{n,k}$ of all subspaces $E \subset \mathbb{C}^n$ satisfying the three conditions*

1. $\dim E = k(\delta, \Theta, n)$
2. *There is a $\Theta-$net N in $S(E) = S(\mathbb{C}^n) \bigcap E$ such that $|f(x) - M_f| \leq \omega_f(\delta)$ for all $x \in N$*
3. $|f(x) - M_f| \leq \omega_f(\delta) + \omega_f(\Theta)$ *for all $x \in S(E)$*

satisfies $\nu_k(E_k) \geq 1 - \sqrt{2/\pi} e^{-\delta^2(n-1)/2}$.

Here, $S(\mathbb{C}^n)$ is the unit sphere in $\mathbb{C}^n$, $\omega_f(\delta) := \sup\{|f(x) - f(y)| : D(x,y) \leq \delta\}$ is the modulus of continuity, D the geodesic metric on $S(\mathbb{C}^n)$ and M_f the median of f, which is the number such that with ν the Haar measure on $S(\mathbb{C}^n)$ both $\nu(\{x : f(x) \leq M_f\}) \geq 1/2$ and $\nu(\{x : f(x) \geq M_f\}) \geq 1/2$ hold.

Remark 1. *The proof of Theorem 3 uses the identification $\mathbb{C}^n \simeq \mathbb{R}^{2n}$ under the map $\sum_{i=1}^n z_i e_i \mapsto \sum_{i=1}^n (\mathfrak{Re}\{z_i\} e_i + \mathfrak{Im}\{z_i\} e_{i+n})$, where $\{e_1, \ldots, e_n\}$ and $\{e_1, \ldots, e_{2n}\}$ denote the standard bases in $\mathbb{C}^n$ and $\mathbb{R}^{2n}$.*

Second, we use the following lemma which first appeared in [20]:

Lemma 1. *Let $\mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$ and $\mathcal{G}$ a d-dimensional subspace of $\mathcal{H}$. Then*

$$\int_{\mathfrak{U}(\mathcal{G})} \langle U\phi, \mathcal{N}(U|\phi\rangle\langle\phi|U^*) U\phi \rangle dU = \frac{1}{d+1}(d \cdot F_e(\pi_{\mathcal{G}}, \mathcal{N}) + 1). \tag{17}$$

Third, we need a well behaving relation between the median and the expectation of a function $f : S(\mathbb{C}^n) \to \mathbb{R}$. This is given by a Lemma taken from [28]:

Lemma 2. *Let $f : S(\mathbb{C}^n) \to \mathbb{R}$ be Lipschitz with constant one (w.r.t. the geodesic metric). Then*

$$|M_f - \mathbb{E}(f)| \leq \frac{12}{\sqrt{2(n-1)}}.$$

Remark 2. *Obviously, this implies $|M_f - \mathbb{E}(f)| \leq \frac{12 \cdot L}{\sqrt{2(n-1)}}$ for Lipschitz functions with constant $L \in \mathbb{R}_+$.*

The function that we will apply Lemma 2 to is given by the following:

Lemma 3. *Let $\Lambda \in \mathcal{C}(\mathcal{H}, \mathcal{H})$. Define $f_\Lambda : S(\mathcal{H}) \to \mathbb{R}$ by*

$$f_\Lambda(x) := \langle x, \Lambda(|x\rangle\langle x|) x \rangle, \ \ x \in S(\mathcal{H}).$$

Then f_Λ is Lipschitz with constant $L = 4$ (w.r.t. the geodesic metric).

Proof. Let $x, y \in S(\mathcal{H})$. Then by Hölder's inequality,

$$\begin{aligned} |f_\Lambda(x) - f_\Lambda(y)| &= |\mathrm{tr}(|x\rangle\langle x|\Lambda(|x\rangle\langle x|)) - \mathrm{tr}(|y\rangle\langle y|\Lambda(|y\rangle\langle y|))| && (18)\\ &= |\mathrm{tr}(|x\rangle\langle x|\Lambda(|x\rangle\langle x| - |y\rangle\langle y|))| + |\mathrm{tr}((|x\rangle\langle x| - |y\rangle\langle y|)\Lambda(|y\rangle\langle y|))| && (19)\\ &\leq \| \, |x\rangle\langle x| \, \|_\infty \cdot \|\Lambda(|x\rangle\langle x| - |y\rangle\langle y|)\|_1 + \| \, |x\rangle\langle x| - |y\rangle\langle y| \, \|_1 \cdot \|\Lambda(|y\rangle\langle y|)\|_\infty && (20)\\ &\leq \|\Lambda(|x\rangle\langle x| - |y\rangle\langle y|)\|_1 + \||x\rangle\langle x| - |y\rangle\langle y|\|_1 && (21)\\ &\leq 2\| \, |x\rangle\langle x| - |y\rangle\langle y| \, \|_1. && (22) \end{aligned}$$

It further holds

$$\| \, |x\rangle\langle x| - |y\rangle\langle y| \, \|_1 \leq 2\|x - y\| \leq 2D(x, y). \tag{23}$$

□

We now state the main ingredient of this section.

Lemma 4. *Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be a finite set of channels and $(\mu_l)_{l\in\mathbb{N}}$ any sequence of (random or deterministic) entanglement transmission codes that satisfies*

A1 $\min_{s^l\in\mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{R}^l, \mathcal{P}^l) = 1 - f_l,$

where $(f_l)_{l\in\mathbb{N}}$ is any sequence of real numbers in the interval $[0, 1]$.
Let $(\varepsilon_l)_{l\in\mathbb{N}}$ be a sequence with $\varepsilon_l \in (0, 1]$ $\forall l \in \mathbb{N}$ satisfying

A2 There is $\hat{l} \in \mathbb{N}$ such that $|\mathbf{S}|^l \sqrt{2/\pi} e^{-\varepsilon_l^2(k_l-1)/128} < 1$ and $k_l \geq 2$ hold for all $l \geq \hat{l}$.

Then for any $l \geq \hat{l}$ there is a subspace $\hat{\mathcal{F}}_l \subset \mathcal{F}_l$ with the properties

P1 $\dim \hat{\mathcal{F}}_l = \lfloor \frac{\varepsilon_l^2}{256 \log(32/\varepsilon_l)} \cdot k_l \rfloor,$

P2 $\min_{s^l\in\mathbf{S}^l} \min_{\phi\in\mathcal{S}(\hat{\mathcal{F}}_l):\phi^2=\phi} \int \langle \phi, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l(|\phi\rangle\langle\phi|)\phi\rangle d\mu_l(\mathcal{R}^l, \mathcal{P}^l) \geq 1 - f_l - \frac{4\cdot 12}{\sqrt{2(k_l-1)}} - \varepsilon_l.$

Proof. Let $l \in \mathbb{N}$. For an arbitrary $s^l \in \mathbf{S}^l$ define $f_{s^l} : S(\mathcal{F}_l) \to \mathbb{R}$ by

$$f_{s^l}(\phi) := \int \langle \phi, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l(|\phi\rangle\langle\phi|)\phi\rangle d\mu_l(\mathcal{R}^l, \mathcal{P}^l) \quad (\phi \in S(\mathcal{F}_l)). \tag{24}$$

Since f_{s^l} is an affine combination of functions with Lipschitz-constant $L = 4$, it is itself Lipschitz with $L = 4$.
Also, by the Theorem of Fubini, Lemma 1 and our assumption *A1* we have

$$\begin{aligned} \mathbb{E}(f_s{}^l) &= \int_{\mathfrak{U}(\mathcal{F}_l)} f_{s^l}(U\phi) dU && (25)\\ &= \int_{\mathfrak{U}(\mathcal{F}_l)} [\int \langle U\phi, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l(|U\phi\rangle\langle U\phi|)U\phi\rangle d\mu_l(\mathcal{R}^l, \mathcal{P}^l)] dU && (26)\\ &= \int \int_{\mathfrak{U}(\mathcal{F}_l)} [\langle U\phi, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l(|U\phi\rangle\langle U\phi|)U\phi\rangle dU] d\mu_l(\mathcal{R}^l, \mathcal{P}^l) && (27)\\ &= \int \frac{k_l F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) + 1}{k_l + 1} d\mu_l(\mathcal{R}^l, \mathcal{P}^l) && (28)\\ &\geq \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{R}^l, \mathcal{P}^l) && (29)\\ &= 1 - f_l. && (30) \end{aligned}$$

By Lemma 2 and Lemma 3 we now get a good lower bound on the median of f_{s^l}:

$$M_{f_{s^l}} \geq \mathbb{E}(f_{s^l}) - \frac{4 \cdot 12}{\sqrt{2(k_l - 1)}} \tag{31}$$

$$\geq 1 - f_l - \frac{4 \cdot 12}{\sqrt{2(k_l - 1)}}. \tag{32}$$

We now apply Theorem 3 with $n = k_l$ and $\delta = \Theta = \varepsilon_l/8$ to f_{s^l}. Then $k \geq \lfloor \frac{\varepsilon_l^2}{256 \log(32/\varepsilon_l)} k_l \rfloor$ holds due to the second estimate in *A2*, and we may set $k_l' := \lfloor \frac{\varepsilon_l^2}{256 \log(32/\varepsilon_l)} k_l \rfloor$, establishing *P1*. Since the fact that f_{s^l} is 4-Lipschitz implies $\omega_{f_{s^l}}(\delta) \leq 4\delta$ we get the following:

$$\nu_k(\{E \in G_{k_l, k_l'} : |f_{s^l}(\phi) - M_{f_{s^l}}| \leq \varepsilon_l \ \forall \phi \in E\}) \geq 1 - \sqrt{2/\pi} e^{-\varepsilon_l^2 (k_l - 1)/128}. \tag{33}$$

The last inequality is valid for each choice of s^l, so we can conclude that

$$\nu_k(\{E \in G_{k_l, k_l'} : |f_{s^l}(\phi) - M_{f_{s^l}}| \leq \varepsilon_l \ \forall \phi \in E, \ \forall s^l \in \mathbf{S}^l\}) \geq 1 - |\mathbf{S}|^l \sqrt{2/\pi} e^{-\varepsilon_l^2 (k_l - 1)/128}. \tag{34}$$

Thus for all $l \geq \hat{l}$ we have

$$\nu_k(\{E \in G_{k_l, k_l'} : |f_{s^l}(\phi) - M_{f_{s^l}}| \leq \varepsilon_l \ \forall \phi \in E, \ \forall s^l \in \mathbf{S}^l\}) > 0 \tag{35}$$

by assumption *A2*, implying the existence of a subspace $E \subset \mathcal{F}_l$ of dimension $\dim E = k_l'$ such that $|f_{s^l}(\phi) - M_{f_{s^l}}| \leq \varepsilon_l \ \forall \ \phi \in E, \ s^l \in \mathbf{S}^l$.
By equation (32) this establishes *P2*:

$$f_{s^l}(\phi) \geq 1 - f_l - \frac{4 \cdot 12}{\sqrt{2(k_l - 1)}} - \varepsilon_l \ \ \forall \ \phi \in E, \ \forall s^l \in \mathbf{S}^l. \tag{36}$$

□

Proof of Theorem 2. First, assuming that $R > 0$ is an achievable rate for entanglement transmission over a *finite* AVQC $\mathfrak{I}$ (with random codes), we show that it is also an achievable strong subspace transmission rate (with random codes) for $\mathfrak{I}$. The proof does not depend on the form of the sequence of probability distributions assigned to the codes, so it applies to the case of deterministically achievable rates as well. So, let there be a sequence of (l, k_l) random entanglement transmission codes with

$$\liminf_{l \to \infty} \frac{1}{l} \log k_l \geq R, \tag{37}$$

$$\min_{s^l \in \mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) = 1 - f_l, \text{ where } f_l \searrow 0. \tag{38}$$

Thus, there is $l' \in \mathbb{N}$ such that $k_l \geq 2^{lR}$ for all $l \geq l'$. Choose $\varepsilon_l = \frac{1}{l}$ (this is just one of many possible choices). Obviously, since R is *strictly* greater than zero there is $\hat{l} \in \mathbb{N}$ such that $|\mathbf{S}|^l \sqrt{2/\pi} e^{-\frac{1}{l^2}(k_l - 1)/8} < 1 \ \forall l \geq \hat{l}$ holds. Application of Lemma 4 then yields a sequence of subspaces $\hat{\mathcal{F}}_l$ with dimensions $\hat{k}_l$ such that

$$\liminf_{l \to \infty} \frac{1}{l} \log \hat{k}_l = \liminf_{l \to \infty} \frac{1}{l} \log k_l \geq R, \tag{39}$$

$$\min_{s^l\in\mathbf{S}^l}\min_{\phi\in\mathcal{S}(\hat{\mathcal{F}}_l)}\int\langle\phi,\mathcal{R}^l\circ\mathcal{N}_{s^l}\circ\mathcal{P}^l(|\phi\rangle\langle\phi|)\phi\rangle d\mu_l(\mathcal{R}^l,\mathcal{P}^l)\geq 1-f_l-\frac{4\cdot 12}{\sqrt{2(k_l-1)}}-\frac{1}{l}\ \ \forall\ l\geq\max\{l',\hat{l}\}. \tag{40}$$

Since the right hand side of (40) goes to zero for l going to infinity, we have shown that R is an achievable rate for strong subspace transmission (with random codes).
In case that $|\mathfrak{I}|=\infty$ holds we have to take care of some extra issues that arise from approximating $\mathfrak{I}$ by a finite AVQC. Such an approximation is carried out in detail in the proof of Lemma 9.
Now let $R=0$ be an achievable rate for entanglement transmission with (random) codes. We show that it is achievable for strong subspace transmission by demonstrating that we can *always* achieve a strong subspace transmission rate of zero:
Choose any sequence $(|x_l\rangle\langle x_l|)_{l\in\mathbb{N}}$ of pure states such that $|x_l\rangle\langle x_l|\in\mathcal{S}(\mathcal{H}^{\otimes l})\ \forall l\in\mathbb{N}$. Set $\mathcal{F}_l:=\mathbb{C}\cdot x_l$ $(l\in\mathbb{N})$. Define a sequence of recovery operations by $\mathcal{R}^l(a):=\mathrm{tr}(a)\cdot|x_l\rangle\langle x_l|$ $(a\in\mathcal{B}(\mathcal{K}^{\otimes l}),\ l\in\mathbb{N})$. Then $F_e(\pi_{\mathcal{F}_l},\mathcal{R}^l\circ\mathcal{N}_{s^l})=1$ for all $l\in\mathbb{N},\ s^l\in\mathbf{S}^l$ and $\liminf_{l\to\infty}\frac{1}{l}\log(\dim\mathcal{F}_l)=0$.

Now let $R\geq 0$ be an achievable rate for strong subspace transmission over some AVQC $\mathfrak{I}$ (with random codes). Thus, there exists a sequence of (random) strong subspace transmission codes with

$$\liminf_{l\to\infty}\frac{1}{l}\log k_l\geq R, \tag{41}$$

$$\inf_{s^l\in\mathbf{S}^l}\min_{\phi\in\mathcal{S}(\mathcal{F}_l)}\int\langle\phi,\mathcal{R}^l\circ\mathcal{N}_{s^l}\circ\mathcal{P}^l(|\phi\rangle\langle\phi|)\phi\rangle d\mu_l(\mathcal{R}^l,\mathcal{P}^l)=1-f_l\ \forall l\in\mathbb{N},\ \text{where}\ f_l\searrow 0. \tag{42}$$

Now consider, for every $l\in\mathbb{N}$ and $s^l\in\mathbf{S}^l$, the channels $\int\mathcal{R}^l\circ\mathcal{N}_{s^l}\circ\mathcal{P}^l d\mu_l(\mathcal{R}^l,\mathcal{P}^l)$. Then (42) implies that for these channels we have the estimate

$$\inf_{s^l\in\mathbf{S}^l}\min_{\phi\in\mathcal{S}(\mathcal{F}_l)}\langle\phi,\int\mathcal{R}^l\circ\mathcal{N}_{s^l}\circ\mathcal{P}^l d\mu_l(\mathcal{R}^l,\mathcal{P}^l)(|\phi\rangle\langle\phi|)\phi\rangle=1-f_l, \tag{43}$$

and by a well-known result ([7], Theorem 2) we get

$$\inf_{s^l\in\mathbf{S}^l}F_e(\pi_{\mathcal{F}_l},\int\mathcal{R}^l\circ\mathcal{N}_{s^l}\circ\mathcal{P}^l d\mu_l(\mathcal{R}^l,\mathcal{P}^l))\geq 1-\frac{3}{2}f_l, \tag{44}$$

which by convex-linearity of the entanglement fidelity in the channel implies

$$\inf_{s^l\in\mathbf{S}^l}\int F_e(\pi_{\mathcal{F}_l},\mathcal{R}^l\circ\mathcal{N}_{s^l}\circ\mathcal{P}^l)d\mu_l(\mathcal{R}^l,\mathcal{P}^l)\geq 1-\frac{3}{2}f_l. \tag{45}$$

But $\lim_{l\to\infty}\frac{3}{2}f_l=0$ by assumption, implying that R is an achievable rate for entanglement transmission (with random codes) as well. □

5 Proof of the converse part

The basic technical obstacle we are faced with is that the converse part of the coding theorem for an AVQC cannot be reduced immediately to that of the single stationary memoryless quantum channel via Minimax Theorem (cf. [11] and [13]). In order to circumvent this problem we derive a relation between $\mathcal{A}_{\mathrm{random}}(\mathfrak{I})$ and the corresponding random capacity of a suitable compound channel.
To be explicit, let us consider a finite AVQC $\mathfrak{I}=\{\mathcal{N}_s\}_{s\in\mathbf{S}}$ and let $(\mu_l)_{l\in\mathbb{N}}$ be a sequence of random $(l,k_l)-$codes for the AVQC $\mathfrak{I}$ with

$$\lim_{l\to\infty}\inf_{s^l\in\mathbf{S}^l}\int F_e(\pi_{\mathcal{F}_l},\mathcal{R}^l\circ\mathcal{N}_{s^l}\circ\mathcal{P}^l)d\mu_l(\mathcal{P}^l,\mathcal{R}^l)=1. \tag{46}$$

On the other hand, for the infinite channel set $\mathrm{conv}(\mathfrak{I})$, defined in (2), and each $\mathcal{N}_q \in \mathrm{conv}(\mathfrak{I})$ we obtain

$$\begin{aligned}\int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_q^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) &= \sum_{s^l \in \mathbf{S}^l} q(s_1) \cdot \ldots \cdot q(s_l) \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) \\ &\geq \inf_{s^l \in \mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l). \qquad (47)\end{aligned}$$

Consequently, (46) and (47) imply

$$\lim_{l\to\infty} \inf_{q \in \mathfrak{P}(\mathbf{S})} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_q^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) = 1. \qquad (48)$$

Defining the random entanglement transmission capacity $Q_{\text{comp, random}}(\mathrm{conv}(\mathfrak{I}))$ for the *compound quantum channel* (cf. [10]) built up from $\mathrm{conv}(\mathfrak{I})$ in a similar fashion to $\mathcal{A}_{\text{random}}(\mathfrak{I})$ we can infer from the considerations presented above that

$$\mathcal{A}_{\text{random}}(\mathfrak{I}) \leq Q_{\text{comp, random}}(\mathrm{conv}(\mathfrak{I})). \qquad (49)$$

Since the inequality $\mathcal{A}_{\det}(\mathfrak{I}) \leq \mathcal{A}_{\text{random}}(\mathfrak{I})$ is obvious, we obtain the following basic lemma.

Lemma 5. *Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be any finite set of channels and let $\mathrm{conv}(\mathfrak{I})$ be the associated infinite set given in (2). Then*

$$\mathcal{A}_{\det}(\mathfrak{I}) \leq \mathcal{A}_{\text{random}}(\mathfrak{I}) \leq Q_{\text{comp, random}}(\mathrm{conv}(\mathfrak{I})). \qquad (50)$$

Thus, our remaining task is to show that right-most capacity in (50) is upper bounded by the last term in (4). This is done in the following two subsections for finite and infinite AVQCs respectively.

5.1 Converse for the finite AVQC

First, we prove the converse to the coding theorem for finite compound quantum channels with random codes.

Theorem 4 (Converse Part: Compound Channel, $|\mathfrak{I}| < \infty$)**.** *Let $\mathfrak{I} = \{\mathcal{N}_1, \ldots, \mathcal{N}_N\} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ be a finite compound channel. The capacity $Q_{\text{comp, random}}(\mathfrak{I})$ of $\mathfrak{I}$ is bounded from above by*

$$Q_{\text{comp, random}}(\mathfrak{I}) \leq \lim_{l\to\infty} \max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes l})} \min_{\mathcal{N}_i \in \mathfrak{I}} \frac{1}{l} I_c(\rho, \mathcal{N}_i^{\otimes l}).$$

Proof. Let for arbitrary $l \in \mathbb{N}$ a random (l, k_l) code for a compound channel $\mathfrak{I} = \{\mathcal{N}_1, \ldots, \mathcal{N}_N\}$ with the property

$$\min_{1\leq i\leq N} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_i^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) \geq 1 - \varepsilon_l$$

be given, where $\varepsilon_l \in [0, 1]$ and $\lim_{l\to\infty} \varepsilon_l = 0$. Obviously, the above code then satisfies

$$\begin{aligned}\int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \frac{1}{N}\sum_{i=1}^N \mathcal{N}_i^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) &= \frac{1}{N}\sum_{i=1}^N \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_i^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) \\ &\geq 1 - \varepsilon_l. \qquad (51)\end{aligned}$$

This implies the existence of at least one pair $(\mathcal{R}^l, \mathcal{P}^l)$ such that

$$F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \frac{1}{N}\sum_{i=1}^N \mathcal{N}_i^{\otimes l} \circ \mathcal{P}^l) \geq 1 - \varepsilon_l.$$

The rest of the proof is identical to that of Theorem 10 in [10]. □

Using the approximation techniques developed in [10], we will now prove the converse for random codes and general compound channels.

Theorem 5 (Converse Part: Compound Channel)**.** *Let $\mathfrak{I} \subset \mathcal{C}(\mathcal{H},\mathcal{K})$ be an arbitrary compound quantum channel. The capacity $Q_{\text{comp, random}}(\mathfrak{I})$ of $\mathfrak{I}$ is bounded from above by*

$$Q_{\text{comp, random}}(\mathfrak{I}) \leq \lim_{l\to\infty} \max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} \inf_{\mathcal{N}\in\mathfrak{I}} \frac{1}{l} I_c(\rho, \mathcal{N}^{\otimes l}).$$

For the proof of this theorem, we will need the following lemma:

Lemma 6 (Cf. [10])**.** *Let $\mathcal{H},\mathcal{K}$ be finite dimensional Hilbert spaces. There is a function $\nu : [0,1] \to \mathbb{R}_+$ with $\lim_{x\to 0}\nu(x) = 0$ such that for every $\mathfrak{I},\mathfrak{I}' \subseteq \mathcal{C}(\mathcal{H},\mathcal{K})$ with $D_\Diamond(\mathfrak{I},\mathfrak{I}') \leq \tau \leq 1/2$ and every $l \in \mathbb{N}$ we have the estimate*

$$\left|\frac{1}{l}\inf_{\mathcal{N}\in\mathfrak{I}} I_c(\rho,\mathcal{N}^{\otimes l}) - \frac{1}{l}\inf_{\mathcal{N}'\in\mathfrak{I}'} I_c(\rho,\mathcal{N}'^{\otimes l})\right| \leq \nu(2\tau) \quad \forall \rho \in \mathcal{S}(\mathcal{H}^{\otimes l})$$

The function ν is given by $\nu(x) = x + 8x\log(\dim\mathcal{K}) + 4h(x)$. Here, $h(\cdot)$ denotes the binary entropy.

Proof of Theorem 5. Let a sequence $(l,k_l)_{l\in\mathbb{N}}$ of random codes for $\mathfrak{I}$ be given such that

- $\liminf_{l\to\infty} \log\dim\mathcal{F}_l = R$
- $\inf_{\mathcal{N}\in\mathfrak{I}} \int F_e(\mathcal{F}_l, \mathcal{R}^l \circ \mathcal{N}^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l,\mathcal{R}^l) = 1 - \varepsilon_l$,

where the sequence $(\varepsilon_l)_{l\in\mathbb{N}}$ satisfies $\lim_{l\to\infty}\varepsilon_l = 0$. It is well-known (see for example [9] and references therein) that we can always choose a $\frac{\tau}{2}$-net $\{\mathcal{N}_i\}_{i=1}^{N(\frac{\tau}{2})}$ in $\mathcal{C}(\mathcal{H},\mathcal{K})$ such that $N(\frac{\tau}{2}) \leq (\frac{6}{\tau})^{2(\dim\mathcal{H}\dim\mathcal{K})^2}$.
We will now construct a τ-net for $\mathfrak{I}$ that lies entirely within $\mathfrak{I}$. Let the numbering of the elements of the net be chosen such that there exists an $M_\tau \in \{0,\ldots,N(\frac{\tau}{2})\}$ such that $B_\Diamond(\mathcal{N}_i,\frac{\tau}{2}) \cap \mathfrak{I} = \emptyset \Leftrightarrow i > M_\tau$ holds for all $i \in \{1,\ldots,N(\frac{\tau}{2})\}$.
For $1 \leq i \leq M_\tau$, choose $\mathcal{N}_i' \in B_\Diamond(\mathcal{N}_i,\frac{\tau}{2}) \cap \mathfrak{I}$. Then $\mathcal{N}_i' \in \mathfrak{I}$ and $||\mathcal{N}_i' - \mathcal{N}_i||_\Diamond \leq \frac{\tau}{2}$.
Let $\mathcal{N} \in B_\Diamond(\mathcal{N}_i,\frac{\tau}{2}) \cap \mathfrak{I}$ be arbitrary. By the preceding observations, we get the estimate

$$||\mathcal{N} - \mathcal{N}_i'||_\Diamond \leq ||\mathcal{N} - \mathcal{N}_i||_\Diamond + ||\mathcal{N}_i - \mathcal{N}_i'||_\Diamond \leq \frac{\tau}{2} + \frac{\tau}{2},$$

thus $\mathcal{N} \in B_\Diamond(\mathcal{N}_i',\tau)$.
Therefore, $\mathcal{M}_\tau := \{\mathcal{N}_i'\}_{i=1}^{M_\tau}$ defines a τ-net of cardinality $|\mathcal{M}_\tau| = M_\tau \leq N(\frac{\tau}{2}) \leq (\frac{6}{\tau})^{2(\dim\mathcal{H}\dim\mathcal{K})^2}$.
By $\mathcal{M}_\tau \subset \mathfrak{I}$ we get, for every $\tau > 0$, the following result:

- $\liminf_{l\to\infty} \log\dim\mathcal{F}_l = R$
- $\min_{\mathcal{N}_i\in\mathcal{M}_\tau} \int F_e(\mathcal{F}_l, \mathcal{R}^l \circ \mathcal{N}_i^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l,\mathcal{R}^l) \geq 1 - \varepsilon_l$.

By Theorem 4, this immediately implies

$$R \leq \lim_{l\to\infty} \frac{1}{l} \max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} \min_{\mathcal{N}_i\in\mathcal{M}_\tau} I_c(\rho,\mathcal{N}_i^{\otimes l}).$$

From Lemma 6 we get, by noting that $D_\Diamond(\mathfrak{I},\mathcal{M}_\tau) \leq \tau$ the estimate

$$R \leq \lim_{l\to\infty} \frac{1}{l} \max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} \inf_{\mathcal{N}\in\mathfrak{I}} I_c(\rho,\mathcal{N}^{\otimes l}) + \nu(2\tau).$$

Taking the limit $\tau \to 0$ proves the theorem. □

Theorem 6 (Converse: finite AVQC)**.** *Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be a finite AVQC. Then*

$$\mathcal{A}_{\text{random}}(\mathfrak{I}) \leq Q_{\text{comp, random}}(\text{conv}(\mathfrak{I})) = \lim_{l\to\infty} \frac{1}{l} \max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} \inf_{\mathcal{N}\in\text{conv}(\mathfrak{I})} I_c(\rho,\mathcal{N}^{\otimes l}).$$

Proof. Just combine Lemma 5 and Theorem 5 applied to conv$(\mathfrak{I})$.

□

5.2 Case $|\mathfrak{I}| = \infty$

The proof of the converse part of Theorem 1 requires just a bit of additional work. Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be an arbitrary AVQC and let $\mathfrak{P}_{\text{fin}}(S)$ denote the set of probability distributions on S with *finite* support. Then

$$\text{conv}(\mathfrak{I}) = \left\{ \mathcal{N}_q \in \mathcal{C}(\mathcal{H},\mathcal{K}) : \mathcal{N}_q := \sum_{s\in\mathbf{S}} q(s)\mathcal{N}_s, \text{ and } q \in \mathfrak{P}_{\text{fin}}(S) \right\},$$

and let

$$\tilde{\mathfrak{I}} := \overline{\text{conv}(\mathfrak{I})}^{||\cdot||_\diamondsuit} \tag{52}$$

where $\overline{\text{conv}(\mathfrak{I})}^{||\cdot||_\diamondsuit}$ denotes the closure of conv($\mathfrak{I}$) with respect to $||\cdot||_\diamondsuit$. The argument that led us to the inequality (47) accompanied by the continuity of the entanglement fidelity with respect to $||\cdot||_\diamondsuit$ and an application of the dominated convergence theorem show that for each $\mathcal{N} \in \tilde{\mathfrak{I}}$

$$\int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) \geq \inf_{s^l\in\mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l)$$

holds. Then Lemma 5 holds *mutatis mutandis* with conv($\mathfrak{I}$) replaced by $\tilde{\mathfrak{I}}$. Additionally, if we apply Theorem 5 to $\tilde{\mathfrak{I}}$ we are led to the following theorem.

Theorem 7 (Converse: general AVC)**.** *Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be an arbitrary AVQC. Then*

$$\mathcal{A}_{\text{random}}(\mathfrak{I}) \leq \lim_{l\to\infty} \frac{1}{l} \max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} \min_{\mathcal{N}\in\tilde{\mathfrak{I}}} I_c(\rho, \mathcal{N}^{\otimes l}).$$

6 Achievability of entanglement transmission rate I: Random codes

We show in this section how the achievability results for compound quantum channels from our previous paper [10] imply existence of reliable random codes for AVQC via Ahlswede's robustification technique [3].

Let $l \in \mathbb{N}$ and let Perm$_l$ denote the set of permutations acting on $\{1,\ldots,l\}$. Let us further suppose that we are given a finite set $\mathbf{S}$. Then each permutation $\sigma \in \text{Perm}_l$ induces a natural action on $\mathbf{S}^l$ by $\sigma : \mathbf{S}^l \to \mathbf{S}^l$, $\sigma(s^l)_i := s_{\sigma(i)}$. Moreover, let $T(l,\mathbf{S})$ denote the set of types on $\mathbf{S}$ induced by the elements of $\mathbf{S}^l$, i.e. the set of empirical distributions on $\mathbf{S}$ generated by sequences in $\mathbf{S}^l$. Then Ahlswede's robustification can be stated as follows.

Theorem 8 (Robustification technique, cf. Theorem 6 in [3])**.** *If a function $f : \mathbf{S}^l \to [0,1]$ satisfies*

$$\sum_{s^l\in\mathbf{S}^l} f(s^l) q(s_1) \cdot \ldots \cdot q(s_l) \geq 1 - \gamma \tag{53}$$

for all $q \in T(l,\mathbf{S})$ and some $\gamma \in [0,1]$, then

$$\frac{1}{l!} \sum_{\sigma\in\text{Perm}_l} f(\sigma(s^l)) \geq 1 - (l+1)^{|\mathbf{S}|} \cdot \gamma \qquad \forall s^l \in \mathbf{S}^l. \tag{54}$$

Remark 3. *Ahlswede's original approach in [3] gives*

$$\frac{1}{l!} \sum_{\sigma\in\text{Perm}_l} f(\sigma(s^l)) \geq 1 - 3 \cdot (l+1)^{|\mathbf{S}|} \cdot \sqrt{\gamma} \qquad \forall s^l \in \mathbf{S}^l.$$

The better bound (54) is from [4].

Proof. Because the result of Theorem 8 is a central tool in our paper and the proof given in [4] is particularly simple we reproduce it here in full for reader's convenience.
Notice first that (53) is equivalent to

$$\sum_{s^l \in \mathbf{S}^l} (1 - f(s^l)) q(s_1) \cdot \ldots \cdot q(s_l) \leq \gamma \qquad \forall q \in T(l, \mathbf{S}),$$

which in turn is equivalent to

$$\sum_{s^l \in \mathbf{S}^l} (1 - f(\sigma(s^l))) q(s_{\sigma(1)}) \cdot \ldots \cdot q(s_{\sigma(l)}) \leq \gamma \qquad \forall q \in T(l, \mathbf{S}),$$

and $\sigma \in \mathrm{Perm}_l$, since σ is bijective. Clearly, we have

$$q(s_{\sigma(1)}) \cdot \ldots \cdot q(s_{\sigma(l)}) = q(s_1) \cdot \ldots \cdot q(s_l) \qquad \forall \sigma \in \mathrm{Perm}_l, \forall s^l \in \mathbf{S}^l,$$

and therefore, we obtain

$$\sum_{s^l \in \mathbf{S}^l} \left(1 - \frac{1}{l!} \sum_{\sigma \in \mathrm{Perm}_l} f(\sigma(s^l))\right) q(s_1) \cdot \ldots \cdot q(s_l) \leq \gamma \qquad \forall q \in T(l, \mathbf{S}). \tag{55}$$

Now, for $q \in T(l, \mathbf{S})$ let $T_q^l \subset \mathbf{S}^l$ denote the set of sequences whose empirical distribution is q. Since f takes values in $[0,1]$ we have $1 - \frac{1}{l!} \sum_{\sigma \in \mathrm{Perm}_l} f(\sigma(s^l)) \geq 0$ and thus from (55)

$$\sum_{s^l \in T_q^l} \left(1 - \frac{1}{l!} \sum_{\sigma \in \mathrm{Perm}_l} f(\sigma(s^l))\right) q(s_1) \cdot \ldots \cdot q(s_l) \leq \gamma \qquad \forall q \in T(l, \mathbf{S}). \tag{56}$$

It is clear from definition that for each $s^l \in T_q^l$ we have $\bigcup_{\sigma \in \mathrm{Perm}_l} \{\sigma(s^l)\} = T_q^l$ and, consequently, $\sum_{\sigma \in \mathrm{Perm}_l} f(\sigma(s^l))$ does not depend on $s^l \in T_q^l$. Therefore, from (56) we obtain

$$\left(1 - \frac{1}{l!} \sum_{\sigma \in \mathrm{Perm}_l} f(\sigma(s^l))\right) q^{\otimes l}(T_q^l) \leq \gamma \qquad \forall q \in T(l, \mathbf{S}),\ \forall s^l \in T_q^l. \tag{57}$$

On the other hand

$$q^{\otimes l}(T_q^l) \geq \frac{1}{(l+1)^{|\mathbf{S}|}} \qquad \forall q \in T(l, \mathbf{S}) \tag{58}$$

holds (cf. [13] page 30), which, by (57), implies

$$\left(1 - \frac{1}{l!} \sum_{\sigma \in \mathrm{Perm}_l} f(\sigma(s^l))\right) \leq (l+1)^{|\mathbf{S}|} \cdot \gamma \qquad \forall q \in T(l, \mathbf{S}),\ \forall s^l \in T_q^l.$$

This is the inequality we aimed to prove since $\mathbf{S}^l = \bigcup_{q \in T(l, \mathbf{S})} T_q^l$. □

The function f appearing in Theorem 8 will be built up from the entanglement fidelities of the channels constituting a finite AVQC that approximates our AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$.
As another ingredient for the arguments to follow we need an achievability result for compound channels.

Lemma 7. *Let $k \in \mathbb{N}$ and $\mathfrak{T} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$. Suppose that*

$$\max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes k})} \inf_{\mathcal{N} \in \mathfrak{T}} I_c(\rho, \mathcal{N}^{\otimes k}) > 0$$

holds. Then for each sufficiently small $\eta > 0$ there is a sequence of (l, k_l)-codes $(\mathcal{P}^l, \mathcal{R}^l)_{l\in\mathbb{N}}$ such that for all $l \geq l_0(\eta)$

$$F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}^{\otimes l} \circ \mathcal{P}^l) \geq 1 - 2^{-lc} \qquad \forall \mathcal{N} \in \mathfrak{T}, \tag{59}$$

and

$$\frac{1}{l} \log \dim \mathcal{F}_l \geq \frac{1}{k} \max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes k})} \inf_{\mathcal{N}\in\mathfrak{T}} I_c(\rho, \mathcal{N}^{\otimes k}) - \eta, \tag{60}$$

hold with a real constant $c = c(k, \dim \mathcal{H}, \dim \mathcal{K}, \mathfrak{T}, \eta) > 0$.

Proof. We will give the details for the case $k = 1$ only. The proof for arbitrary k follows by an almost identical argument.

According to the compound BSST Lemma (cf. [10]) to any $\eta > 0$ we can find $m = m(\mathfrak{T}, \eta) \in \mathbb{N}$ and a subspace $\mathcal{G} \subset \mathcal{H}^{\otimes m}$ such that

$$\frac{1}{m} \inf_{\mathcal{N}\in\mathfrak{T}} I_c(\pi_{\mathcal{G}}, \mathcal{N}^{\otimes m}) \geq \max_{\rho\in\mathcal{S}(\mathcal{H})} \inf_{\mathcal{N}\in\mathfrak{T}} I_c(\rho, \mathcal{N}) - \frac{\eta}{3}. \tag{61}$$

Let us consider the compound quantum channel built up from $\{\mathcal{N}^{\otimes m} : \mathcal{N} \in \mathfrak{T}\}$. By a suitable choice of the relevant parameters in the proof of Lemma 9 in [10] we obtain for each sufficiently small $\eta > 0$ a sequence of (t, k_t)-codes $(\bar{\mathcal{P}}^t, \bar{\mathcal{R}}^t)_{t\in\mathbb{N}}$, $\bar{\mathcal{P}}^t \in \mathcal{C}(\mathcal{F}_t, \mathcal{H}^{\otimes mt})$, $\bar{\mathcal{R}}^t \in \mathcal{C}(\mathcal{K}^{\otimes mt}, \mathcal{F}'_t)$ with

$$\inf_{\mathcal{N}\in\mathfrak{T}} F_e(\pi_{\mathcal{F}_t}, \bar{\mathcal{R}}^t \circ \mathcal{N}^{\otimes mt} \circ \bar{\mathcal{P}}^t) \geq 1 - 2^{-tc'}, \tag{62}$$

and

$$\frac{1}{t} \log k_t = \frac{1}{t} \log \dim \mathcal{F}_t \geq \inf_{\mathcal{N}\in\mathfrak{T}} I_c(\pi_{\mathcal{G}}, \mathcal{N}^{\otimes m}) - \frac{\eta}{3} \tag{63}$$

for all $t \geq t_1(\eta)$ with a positive constant $c' = c'(\dim \mathcal{H}, \dim \mathcal{K}, \mathfrak{T}, \eta)$.

For $t, l \in \mathbb{N}$ let $r \in \{0, 1, \ldots, m-1\}$ be the unique non-negative integer such that $l = mt + r$. Furthermore, let us choose for each $r \in \{0, 1, \ldots, m-1\}$ a state vector $x_r \in \mathcal{H}^{\otimes r}$ and set

$$\mathcal{F}_l := \mathcal{F}_t \otimes \mathbb{C} \cdot \{x_r\}. \tag{64}$$

Then

$$\pi_{\mathcal{F}_l} = \pi_{\mathcal{F}_t} \otimes |x_r\rangle\langle x_r|. \tag{65}$$

Moreover we set

$$\mathcal{P}^l := \bar{\mathcal{P}}^t \otimes id_{\mathcal{B}(\mathcal{H}^{\otimes r})} \quad \text{and} \quad \mathcal{R}^l := \bar{\mathcal{R}}^t \otimes T^r, \tag{66}$$

where $T^r \in \mathcal{C}(\mathcal{K}^{\otimes r}, \mathcal{H}^{\otimes r})$ is given by $T^r(a) := \mathrm{tr}(a)|x_r\rangle\langle x_r|$. Then it is clear that

$$\begin{aligned} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}^{\otimes l} \circ \mathcal{P}^l) &= F_e(\pi_{\mathcal{F}_t}, \bar{\mathcal{R}}^t \circ \mathcal{N}^{\otimes mt} \circ \bar{\mathcal{P}}^t) \\ &\geq 1 - 2^{-tc'} \\ &= 1 - 2^{-\frac{l-r}{m}c'} \\ &\geq 1 - 2^{-lc} \qquad \forall \mathcal{N} \in \mathfrak{T} \end{aligned} \tag{67}$$

for all $l \geq l_1(\eta)$ with $c := \frac{c'}{2m}$, and where in the second line we have used (62).

On the other hand, from equations (61), (63) and (64) we obtain for $t \geq t(\eta)$

$$\begin{aligned} \frac{1}{l} \log \dim \mathcal{F}_l &= \frac{1}{tm + r} \log \dim \mathcal{F}_t \\ &\geq \frac{1}{1 + \frac{r}{tm}} \Big(\max_{\rho\in\mathcal{S}(\mathcal{H})} \inf_{\mathcal{N}\in\mathfrak{T}} I_c(\rho, \mathcal{N}) - \frac{\eta}{3} - \frac{\eta}{3m} \Big) \\ &\geq \max_{\rho\in\mathcal{S}(\mathcal{H})} \inf_{\mathcal{N}\in\mathfrak{T}} I_c(\rho, \mathcal{N}) - \eta \end{aligned} \tag{68}$$

if t and consequently l is sufficiently large. Therefore there is an $l_0(\eta) \in \mathbb{N}$ such that (67) and (68) hold simultaneously for all $l \geq l_0(\eta)$ which concludes the proof in the case $k = 1$. □

In the next step we will combine the robustification technique and Lemma 7 to prove the existence of good random codes for the AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$.
Recall that there is a canonical action of Perm_l on $\mathcal{B}(\mathcal{H})^{\otimes l}$ given by $A_{\sigma,\mathcal{H}}(a_1 \otimes \ldots \otimes a_l) := a_{\sigma^{-1}(1)} \otimes \ldots \otimes a_{\sigma^{-1}(l)}$. It is easy to see that $A_{\sigma,\mathcal{H}}(a) = U_\sigma a U_\sigma^*$, $(a \in \mathcal{B}(\mathcal{H})^{\otimes l})$ with the unitary operator $U_\sigma : \mathcal{H}^{\otimes l} \to \mathcal{H}^{\otimes l}$ defined by $U_\sigma(x_1 \otimes \ldots \otimes x_l) = x_{\sigma^{-1}(1)} \otimes \ldots \otimes x_{\sigma^{-1}(l)}$.

Theorem 9 (Conversion of compound codes)**.** *Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be an AVQC. For each $k \in \mathbb{N}$ and any sufficiently small $\eta > 0$ there is a sequence of codes $(\mathcal{P}^l, \mathcal{R}^l)_{l\in\mathbb{N}}$, $\mathcal{P}^l \in \mathcal{C}(\mathcal{F}_l, \mathcal{H}^{\otimes l})$, $\mathcal{R}^l \in \mathcal{C}(\mathcal{K}^{\otimes l}, \mathcal{F}_l')$, for the compound channel built up from* $\mathrm{conv}(\mathfrak{I})$ *(cf. (2)) satisfying*

$$\frac{1}{l}\log\dim\mathcal{F}_l \geq \frac{1}{k}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes k})}\inf_{\mathcal{N}\in\mathrm{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes k}) - 2\cdot h(8\eta), \tag{69}$$

$$\frac{1}{l!}\sum_{\sigma\in\mathrm{Perm}_l} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ A_{\sigma^{-1},\mathcal{K}} \circ \mathcal{N}_{s^l} \circ A_{\sigma,\mathcal{H}} \circ \mathcal{P}^l) \geq 1 - (l+1)^{N_\eta}\cdot 2^{-lc} \qquad \forall s^l \in \mathbf{S}^l \tag{70}$$

for all sufficiently large l with a positive number $c = c(k, \dim\mathcal{H}, \dim\mathcal{K}, \mathrm{conv}(\mathfrak{I}), \eta)$, $\nu : [0,1] \to \mathbb{R}$ defined by $\nu(x) := x + 8x\log(d_\mathcal{K}) + 4\cdot h(x)$ ($h(\cdot)$ being the binary entropy) and an integer N_η which depends on the set $\mathfrak{I}$ as well.

The idea of the proof is the following. We want to approximate the set $\mathrm{conv}(\mathfrak{I})$ from the outside by using a polytope P_η with N_η extreme points. Then, our results for compound codes and an application of the robustification technique yield a sequence of codes which have asymptotically optimal performance for the AVQC P_η. Since $\mathrm{conv}(\mathfrak{I}) \subset P_\eta$, they will also have asymptotically optimal performance for $\mathfrak{I}$.
A problem occurs if $\mathrm{conv}(\mathfrak{I})$ touches the boundary of the set of quantum channels because parts of that boundary are curved and the approximating polytope P_η may contain maps that are not channels. Therefore, an intermediate step consists of slightly moving $\mathrm{conv}(\mathfrak{I})$ away from the boundary. This may be seen as application of a completely positive map and can therefore be absorbed into the recovery operation.
During the proof we are going to make use of the following Lemma, that will be proven first:

Lemma 8. *Let A, B be compact convex sets in $\mathbb{C}^n$ with $A \subset B$ and*

$$d(\mathrm{rebd}\,B, A) := \inf\{||b - a|| : b \in \mathrm{rebd}\,B, a \in A\} = t > 0, \tag{71}$$

where $||\cdot||$ denotes any norm.
Let $P \supset A$ be a polytope with $D(A, P) \leq \delta$, where $\delta \in (0, t]$ and D is the Hausdorff distance induced by $||\cdot||$. Then $P' := P \cap \mathrm{aff}\,B$ is a polytope and $P' \subset B$.

Proof of Lemma 8. The assertion that P' is a polytope is clear. Suppose $\exists p \in P'\backslash B$. Then since $D(A, P) \leq \delta$ we have $P \subset (A)_\delta$ (cf. [34], Theorem 2.7.3). But this means, since $P' \subset P$, that to our $p \in P'\backslash B$ we can find $a_\delta \in A$ with

$$||p - a_\delta|| \leq \delta. \tag{72}$$

For $\lambda \in [0,1]$ define

$$x_\lambda := (1-\lambda)a_\delta + \lambda p. \tag{73}$$

Then there is $\lambda^* \in (0,1)$ such that

$$x := x^*_\lambda \quad \in \quad \mathrm{rebd}\,B. \tag{74}$$

This is seen as follows: Since $d(\mathrm{rebd}\,B, A) = t > 0$ we have $A \subset \mathrm{ri}\,B$. Set

$$L := \{\lambda \in (0,1] : (1-\lambda)a_\delta + \lambda p \in B\}. \tag{75}$$

From $a_\delta \in \operatorname{ri} B$ it follows that $L \neq \emptyset$ and from the fact that B is compact and convex we then get that $L = (0, \lambda^*]$. Now,

$$\begin{aligned} ||x - a_\delta|| &= ||(1-\lambda^*)a_\delta + \lambda^* p - (1-\lambda^*)a_\delta - \lambda^* a_\delta|| && (76)\\ &= \lambda^* ||p - a_\delta|| && (77)\\ &\leq \lambda^* \cdot \delta && (78)\\ &< t, && (79) \end{aligned}$$

where the last line follows from $\lambda \in (0,1)$. This is a contradiction to $d(\operatorname{rebd} B, A) = t$. □

Proof of Theorem 9. We can suppose that

$$\max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes k})} \inf_{\mathcal{N} \in \operatorname{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes k}) > 0, \tag{80}$$

because otherwise our claim is obviously true. We will further assume that $\mathfrak{I}$, and therefore $\operatorname{conv}(\mathfrak{I})$ as well, is compact. Since the Hausdorff-distance of $\operatorname{conv}(\mathfrak{I})$ to its closure (in $||\cdot||_\lozenge$) is zero, this does not change the left hand side of equation (80), due to the estimates in Lemma 6. Since $\mathfrak{I}$ is a subset of its norm-closure, good codes for the norm-closure will also work for $\mathfrak{I}$. Thus, our assumption is a pure technicality and, indeed, without loss of generality.

Now let us, for $\varepsilon \leq 1$, by $\mathfrak{D}_\varepsilon$ denote the operation $\mathfrak{D}_\varepsilon(\cdot) := (1-\varepsilon)\mathrm{id}_{\mathcal{B}(\mathcal{K})}(\cdot) + \frac{\varepsilon}{\dim \mathcal{K}} \mathbf{1}_{\mathcal{K}} \mathrm{tr}(\cdot)$. If $\varepsilon \geq 0$, this is nothing but a depolarizing channel.

By Lemma 2.3.3 in [34] and since $\mathfrak{D}_1 \circ \mathcal{N} \notin \operatorname{rebd} \mathcal{C}(\mathcal{H},\mathcal{K})$ for arbitrary $\mathcal{N} \in \mathcal{C}(\mathcal{H},\mathcal{K})$, we have

$$\mathfrak{D}_\eta(\operatorname{conv}(\mathfrak{I})) \subset \operatorname{ri} \mathcal{C}(\mathcal{H},\mathcal{K}). \tag{81}$$

Since $\mathfrak{D}_\eta(\operatorname{conv}(\mathfrak{I}))$ is compact, we know that

$$c' := \min\{||\mathcal{N} - \mathcal{N}'||_\lozenge : \mathcal{N} \in \mathfrak{D}_\eta(\operatorname{conv}(\mathfrak{I})), \mathcal{N}' \in \operatorname{rebd} \mathcal{C}(\mathcal{H},\mathcal{K})\} \tag{82}$$

satisfies $c' > 0$. Thus, by Lemma 8 and Theorem 3.1.6 in [34] there exists a polytope $P_\eta \subset \mathcal{C}(\mathcal{H},\mathcal{K})$ such that $\mathfrak{D}_\eta(\operatorname{conv}(\mathfrak{I})) \subset P_\eta$ and

$$D_\lozenge(\mathfrak{D}_\eta(\operatorname{conv}(\mathfrak{I})), P_\eta) \leq 2\eta. \tag{83}$$

The set of extremal points of P_η we denote by $\operatorname{ext}(P_\eta) = \{\mathcal{N}_e\}_{e \in E_\eta}$, where E_η is a finite set indexing the extremal points, the number of which we label N_η. Consider the compound quantum channel (the papers [9, 10] give proper definitions of this object) P_η. It follows from Lemma 7 that there exists a sequence of (l, k_l)-codes $(\mathcal{P}^l, \mathcal{R}^l)_{l \in \mathbb{N}}$ such that for all $l \geq l_0(\eta)$

$$F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}^{\otimes l} \circ \mathcal{P}^l) \geq 1 - 2^{-lc} \qquad \forall \mathcal{N} \in P_\eta, \tag{84}$$

and

$$\frac{1}{l} \log \dim \mathcal{F}_l \geq \frac{1}{k} \max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes k})} \inf_{\mathcal{N} \in P_\eta} I_c(\rho, \mathcal{N}^{\otimes k}) - \eta, \tag{85}$$

with a positive number $c = c(k, \dim \mathcal{H}, \dim \mathcal{K}, \operatorname{conv}(\mathfrak{I}), \eta)$.

Let us define $f : E_\eta^l \to [0,1]$ by

$$f(e^l) := F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{e^l} \circ \mathcal{P}^l).$$

Then (84) implies that

$$\sum_{e^l \in E_\eta^l} f(e^l) q(s_1) \cdot \ldots \cdot q(s_l) \geq 1 - 2^{-lc} \qquad \forall q \in T(l, E_\eta). \tag{86}$$

But (86) and Theorem 8 yield

$$\frac{1}{l!}\sum_{\sigma\in\mathrm{Perm}_l} F_e(\pi_{\mathcal{F}_l},\mathcal{R}^l\circ\mathcal{N}_{\sigma(e^l)}\circ\mathcal{P}^l)\ge 1-(l+1)^{N_\eta}\cdot 2^{-lc}\qquad \forall e^l\in E_\eta^l. \tag{87}$$

By (85) and (87) we are guaranteed the existence of a good random code for P_η if we can somehow consider permutations as part of the encoding and recovery procedure. More precisely, we will now show that

$$\mathcal{N}_{\sigma(e^l)}=A_{\sigma^{-1},\mathcal{K}}\circ\mathcal{N}_{e^l}\circ A_{\sigma,\mathcal{H}}\qquad \forall e^l\in E_\eta^l. \tag{88}$$

To this end, let $\psi=\psi_1\otimes\ldots\otimes\psi_l,\ \varphi=\varphi_1\otimes\ldots\otimes\varphi_l\in\mathcal{H}^{\otimes l}$. Then

$$\begin{aligned}
A_{\sigma^{-1},\mathcal{K}}\circ\mathcal{N}_{e^l}\circ A_{\sigma,\mathcal{H}}(|\psi\rangle\langle\varphi|) &= (A_{\sigma^{-1},\mathcal{K}}\circ\mathcal{N}_{e^l})(|\psi_{\sigma^{-1}(1)}\rangle\langle\varphi_{\sigma^{-1}(1)}|\otimes\ldots\otimes|\psi_{\sigma^{-1}(l)}\rangle\langle\varphi_{\sigma^{-1}(l)}|)\\
&= A_{\sigma^{-1},\mathcal{K}}(\otimes_{i=1}^l\mathcal{N}_{s_i}(|\psi_{\sigma^{-1}(i)}\rangle\langle\varphi_{\sigma^{-1}(i)}|))\\
&= \otimes_{i=1}^l\mathcal{N}_{s_{\sigma(i)}}(|\psi_i\rangle\langle\varphi_i|)\\
&= \mathcal{N}_{\sigma(e^l)}(\otimes_{i=1}^l|\psi_i\rangle\langle\varphi_i|)\\
&= \mathcal{N}_{\sigma(e^l)}(|\psi\rangle\langle\varphi|).
\end{aligned}$$

Therefore,

$$A_{\sigma^{-1},\mathcal{K}}\circ\mathcal{N}_{e^l}\circ A_{\sigma,\mathcal{H}}=\mathcal{N}_{\sigma(e^l)}.$$

By construction of P_η we know that for every $\mathcal{N}_s\in\mathfrak{I}$ there exists a probability distribution $q(\cdot|s)\in\mathfrak{P}(E_\eta)$ such that

$$\mathfrak{D}_\eta\circ\mathcal{N}_s=\sum_{e\in E_\eta}q(e|s)\mathcal{N}_e \tag{89}$$

holds. We define

$$\tilde{\mathcal{R}}_\sigma^l:=\mathcal{R}^l\circ A_{\sigma^{-1},\mathcal{K}}\circ\mathfrak{D}_\eta^{\otimes l},\qquad \tilde{\mathcal{P}}_\sigma^l:=A_{\sigma,\mathcal{H}}\circ\mathcal{P}^l. \tag{90}$$

Combining the equations (87),(88,),(89),(90) we get for every $s^l\in\mathbf{S}^l$:

$$\begin{aligned}
\sum_{\sigma\in\mathrm{Perm}_l}F_e(\pi_{\mathcal{F}_l},\tilde{\mathcal{R}}_\sigma^l\circ\mathcal{N}_{s^l}\circ\tilde{\mathcal{P}}_\sigma^l) &= \sum_{\sigma\in\mathrm{Perm}_l}F_e(\pi_{\mathcal{F}_l},\mathcal{R}^l\circ A_{\sigma^{-1},\mathcal{K}}\circ\mathfrak{D}_\eta^{\otimes l}\circ\mathcal{N}_{s^l}\circ A_{\sigma,\mathcal{H}}\circ\mathcal{P}^l)\\
&= \sum_{\sigma\in\mathrm{Perm}_l}F_e(\pi_{\mathcal{F}_l},\mathcal{R}^l\circ A_{\sigma^{-1},\mathcal{K}}\circ\sum_{e^l\in E_\eta^l}\prod_{i=1}^l q(e_i|s_i)\otimes_{j=1}^l\mathcal{N}_{e_j}\circ A_{\sigma,\mathcal{H}}\circ\mathcal{P}^l)\\
&= \sum_{e^l\in E_\eta^l}\prod_{i=1}^l q(e_i|s_i)\sum_{\sigma\in\mathrm{Perm}_l}F_e(\pi_{\mathcal{F}_l},\mathcal{R}^l\circ A_{\sigma^{-1},\mathcal{K}}\circ\mathcal{N}_{e^l}\circ A_{\sigma,\mathcal{H}}\circ\mathcal{P}^l)\\
&= \sum_{e^l\in E_\eta^l}\prod_{i=1}^l q(e_i|s_i)\sum_{\sigma\in\mathrm{Perm}_l}F_e(\pi_{\mathcal{F}_l},\mathcal{R}^l\circ\mathcal{N}_{\sigma(e^l)}\circ\mathcal{P}^l)\\
&\ge (l!)(1-(l+1)^{N_\eta}\cdot 2^{-lc})
\end{aligned} \tag{91}$$

Now, defining a discretely supported probability measure μ_l, $l\in\mathbb{N}$ by

$$\mu_l:=\frac{1}{l!}\sum_{\sigma\in\mathrm{Perm}_l}\delta_{(\tilde{\mathcal{R}}_\sigma^l,\tilde{\mathcal{P}}_\sigma^l)},$$

where $\delta_{(\tilde{\mathcal{R}}_\sigma^l,\tilde{\mathcal{P}}_\sigma^l)}$ denotes the probability measure that puts measure 1 on the point $(\tilde{\mathcal{R}}_\sigma^l,\mathcal{P}_\sigma^l)$, we obtain for each $k\in\mathbb{N}$ a sequence of (l,k_l)-random codes for $\mathfrak{I}$ achieving

$$\frac{1}{k}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes k})}\inf_{\mathcal{N}\in P_\eta}I_c(\rho,\mathcal{N}^{\otimes k})-\eta.$$

It remains to show that this last number is close to (69). This in turn is true mostly because, by construction, $D_\Diamond(P_\eta, \mathfrak{D}_\eta(\mathrm{conv}(\mathfrak{I}))) \leq 2\eta$ holds and, as will be shown, $D_\Diamond(\mathrm{conv}(\mathfrak{I}), \mathfrak{D}_\eta(\mathrm{conv}(\mathfrak{I}))) \leq 2\eta$ holds.
We start with the upper bound on $D_\Diamond(\mathrm{conv}(\mathfrak{I}), \mathfrak{D}_\eta(\mathrm{conv}(\mathfrak{I})))$, which will be derived in a slightly more general way. For arbitrary $s \leq 1$ and a compact $A \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$

$$D_\Diamond(\mathfrak{D}_s(A), A) \leq |s| \cdot \max_{x \in A} ||x - \mathfrak{D}_1 \circ x|| \leq 2|s| \tag{92}$$

holds, where the second inequality follows from $A \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ in an obvious way and we only prove the first one:

$$\begin{aligned}
\max_{x \in \mathfrak{D}_s(A)} \min_{y \in A} ||x - y||_\Diamond &= \max_{x \in A} \min_{y \in A} ||\mathfrak{D}_s(x) - y||_\Diamond &(93)\\
&= \max_{x \in A} \min_{y \in A} ||(1-s)x + s\mathfrak{D}_1 \circ x - (1-s)y - sy||_\Diamond &(94)\\
&\leq \max_{x \in A} \min_{y \in A} (||(1-s)x - (1-s)y||_\Diamond + ||sy - s\mathfrak{D}_1 \circ x||_\Diamond) &(95)\\
&\leq \max_{x \in A} |s| \cdot ||x - \mathfrak{D}_1 \circ x||_\Diamond. &(96)
\end{aligned}$$

A similar calculation leads to

$$\max_{x \in A} \min_{y \in \mathfrak{D}_s(A)} ||x - y||_\Diamond \leq |s| \cdot \max_{x \in A} \cdot ||x - \mathfrak{D}_1 \circ x||_\Diamond. \tag{97}$$

Application of the triangle inequality for $D_\Diamond$ gives us the estimate

$$D_\Diamond(P_\eta, \mathrm{conv}(\mathfrak{I})) \leq 4\eta. \tag{98}$$

Lemma 16 in [10] (originating back to [26]), finally makes the connection between our set-theoretic approximations and the capacity formula:

$$|\frac{1}{k} \max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes k})} \inf_{\mathcal{N} \in P_\eta} I_c(\rho, \mathcal{N}^{\otimes k}) - \frac{1}{k} \max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes k})} \inf_{\mathcal{N} \in \mathrm{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes k})| \leq \nu(8\eta) \tag{99}$$

with $\nu(x) = x + 8x \log(d_\mathcal{K}) + 4h(x)$. It is obvious that $-\eta \geq -\nu(8\eta)$ holds, therefore for l large enough

$$\frac{1}{l} \log \dim \mathcal{F}_l \geq \frac{1}{k} \max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes k})} \inf_{\mathcal{N} \in \mathrm{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes k}) - 2\nu(8\eta). \tag{100}$$

□

This leads to the following corollary to Theorem 9.

Corollary 10. *For any AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$ we have*

$$\mathcal{A}_{\mathrm{random}}(\mathfrak{I}) \geq \lim_{l \to \infty} \frac{1}{l} \max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes l})} \inf_{\mathcal{N} \in \mathrm{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes l}).$$

Together with Theorem 5 this proves the first part of Theorem 1.

7 Achievability of entanglement transmission rate II: Derandomization

In this section we will prove the second claim made in Theorem 1 by following Ahlswede's elimination technique. The main result of this section is the following Theorem.

Theorem 11. *Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be an AVQC. Then $C_{\mathrm{det}}(\mathfrak{I}) > 0$ implies $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) = \mathcal{A}_{\mathrm{random}}(\mathfrak{I})$.*

The proof of Theorem 11 is based mainly on the following lemma, which shows that not much of common randomness is needed to achieve $\mathcal{A}_{\mathrm{random}}(\mathfrak{I})$.

Lemma 9 (Random Code Reduction)**.** *Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be an AVQC, $l \in \mathbb{N}$, and μ_l an (l, k_l)-random code for the AVQC $\mathfrak{I}$ with*

$$e(\mu_l, \mathfrak{I}) := \inf_{s^l\in\mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) \geq 1 - 2^{-la} \tag{101}$$

for some positive constant $a \in \mathbb{R}$.
Let $\varepsilon \in (0, 1)$. Then for all sufficiently large $l \in \mathbb{N}$ there exist l^2 codes $\{(\mathcal{P}^l_i, \mathcal{R}^l_i) : i = 1, \dots, l^2\} \subset \mathcal{C}(\mathcal{F}_l, \mathcal{H}^{\otimes l}) \times \mathcal{C}(\mathcal{K}^{\otimes l}, \mathcal{F}'_l)$ such that

$$\frac{1}{l^2} \sum_{i=1}^{l^2} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l_i \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l_i) > 1 - \varepsilon \qquad \forall s^l \in \mathbf{S}^l. \tag{102}$$

Proof. Before we get into the details, we should note that the whole proof can be read much more easily if one restricts to the case $|\mathfrak{I}| < \infty$ and sets each of the approximating sets occurring in the sequel equal to $\mathfrak{I}$.
Let (Λ_i, Ω_i), $i = 1, \dots, K$, be independent random variables with values in $\mathcal{C}(\mathcal{F}_l, \mathcal{H}^{\otimes l}) \times \mathcal{C}(\mathcal{K}^{\otimes l}, \mathcal{F}'_l)$ which are distributed according to $\mu_l^{\otimes K}$. Let $(P_l)_{l\in\mathbb{N}}$ be a sequence of polytopes with, for all $l \in \mathbb{N}$, the properties

1. $P_l \subset \mathrm{conv}(\mathfrak{I})$
2. $D_\Diamond(P_, \mathrm{conv}(\mathfrak{I})) \leq 1/l^2$.

Denote by $ext(P_l)$ the extremal points of P_l. Consider an indexing such that we can write $ext(P_l) = \{\mathcal{N}_e\}_{e\in E_l}$ and note that the polytope P_l can be chosen in such a way that $N_l := |E_l|$ satisfies $N_l \leq (6l)^{4\dim(\mathcal{H})^2 \dim(\mathcal{K})^2}$ (see, for example, Lemma 5.2 in [9]).
For every $e^l \in E^l_l$ and corresponding channel $\mathcal{N}_{e^l}$, an application of Markov's inequality yields for any $\varepsilon \in (0, 1)$ and any $\gamma > 0$ the following:

$$\begin{aligned} \mathbb{P}\left(1 - \frac{1}{K}\sum_{i=1}^{K} F_e(\pi_{\mathcal{F}_l}, \Lambda_i \circ \mathcal{N}_{e^l} \circ \Omega_i) \geq \varepsilon/2\right) &= \mathbb{P}\left(2^{K\gamma - \gamma\sum_{i=1}^K F_e(\pi_{\mathcal{F}_l}, \Lambda_i\circ\mathcal{N}_{e^l}\circ\Omega_i)} \geq 2^{K\gamma(\varepsilon/2)}\right) \\ &\leq 2^{-K\gamma(\varepsilon/2)} \cdot \mathbb{E}\left(2^{\gamma(K - \sum_{i=1}^K F_e(\pi_{\mathcal{F}_l}, \Lambda_i\circ\mathcal{N}_{e^l}\circ\Omega_i))}\right). \end{aligned} \tag{103}$$

We will derive an upper bound on the expectation in the preceding line:

$$\begin{aligned} \mathbb{E}\left(2^{\gamma(K - \sum_{i=1}^K F_e(\pi_{\mathcal{F}_l}, \Lambda_i\circ\mathcal{N}_{e^l}\circ\Omega_i))}\right) &= \mathbb{E}\left(2^{\gamma(\sum_{i=1}^K (1 - F_e(\pi_{\mathcal{F}_l}, \Lambda_i\circ\mathcal{N}_{e^l}\circ\Omega_i)))}\right) \\ &\overset{(a)}{=} \left[\mathbb{E}\left(2^{\gamma(1 - F_e(\pi_{\mathcal{F}_l}, \Lambda_1\circ\mathcal{N}_{e^l}\circ\Omega_1))}\right)\right]^K \\ &\overset{(b)}{\leq} [\mathbb{E}(1 + 2^\gamma(1 - F_e(\pi_{\mathcal{F}_l}, \Lambda_1 \circ \mathcal{N}_{e^l} \circ \Omega_1)))]^K \\ &\overset{(c)}{\leq} [1 + 2^\gamma 2^{-la}]^K. \end{aligned} \tag{104}$$

We used (a) independence of the (Λ_i, Ω_i), (b) the inequality $2^{\gamma t} \leq (1 - t)2^{\gamma\cdot 0} + t2^\gamma \leq 1 + t2^\gamma,\ t \in [0, 1]$, where the first inequality is simply the convexity of $[0, 1] \ni t \mapsto 2^{\gamma t}$, (c) holds by (101) and by $P_l \subset \mathrm{conv}(\mathfrak{I})$.
Now, for $K = l^2$, $\gamma = 2$ there is an $l_0(\varepsilon) \in \mathbb{N}$ such that for all $l \geq l_0(\varepsilon)$ we have

$$(1 + 2^2 2^{-la})^{l^2} \leq 2^{l^2(\varepsilon/2)}. \tag{105}$$

Therefore, we obtain from (103), (104), and (105) that for all sufficiently large $l \in \mathbb{N}$

$$\mathbb{P}\left(1 - \frac{1}{l^2}\sum_{i=1}^{l^2} F_e(\pi_{\mathcal{F}_l}, \Lambda_i \circ \mathcal{N}_{e^l} \circ \Omega_i) \geq (\varepsilon/2)\right) \leq 2^{-l^2(\varepsilon/2)} \tag{106}$$

uniformly in $e^l \in E_l^l$. It follows from (106) that

$$\mathbb{P}\left(\frac{1}{l^2}\sum_{i=1}^{l^2} F_e(\pi_{\mathcal{F}_l}, \Lambda_i \circ \mathcal{N}_{e^l} \circ \Omega_i) > 1 - \varepsilon/2 \ \forall e^l \in E_l^l\right) \quad \geq \quad 1 - N_l^l \cdot 2^{-l^2(\varepsilon/2)}$$

implying the existence of a realization $(\mathcal{P}_i^l, \mathcal{R}_i^l)_{i=1}^{l^2}$ with

$$\frac{1}{l^2}\sum_{i=1}^{l^2} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}_i^l \circ \mathcal{N}_{e^l} \circ \mathcal{P}_i^l) > 1 - \varepsilon/2 \qquad \forall e^l \in \mathbf{E}_l^l$$

whenever $N_l^l \cdot 2^{-l^2\varepsilon} < 1$, which is clearly fulfilled for all sufficiently large $l \in \mathbb{N}$.
Finally, we note that for every $l \in \mathbb{N}$ and $\mathcal{N}_s \in \mathfrak{I}$ there is $\mathcal{N}_e \in E_l$ such that $||\mathcal{N}_s - \mathcal{N}_e||_\diamond \leq \frac{1}{l^2}$ and, therefore, to every $\mathcal{N}_{s^l}$ there exists $\mathcal{N}_{e^l}$ (with each $\mathcal{N}_{e_i} \in E_l$) such that (see the proof of Lemma 5.2 in [9] for details)

$$||\mathcal{N}_{s^l} - \mathcal{N}_{e^l}||_\diamond \leq \sum_{i=1}^{l} ||\mathcal{N}_{s_i} - \mathcal{N}_{e_i}||_\diamond \leq \frac{1}{l}, \tag{107}$$

and therefore for every $s^l \in \mathbf{S}^l$ we have, for a maybe even larger l as before (satisfying $1/l < \varepsilon$, additionally),

$$\frac{1}{l^2}\sum_{i=1}^{l^2} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}_i^l \circ \mathcal{N}_{e^l} \circ \mathcal{P}_i^l) > 1 - \varepsilon \qquad \forall s^l \in \mathbf{S}^l.$$

□

We proceed with the proof of Theorem 11. Since $C_{\det}(\mathfrak{I}) > 0$ according to the assumption of the theorem, there is an (m_l, l^2)-deterministic code $\mathfrak{C}_{m_l} = (\rho_i, D_i)_{i=1}^{l^2}$ with $\rho_1, \ldots, \rho_{l^2} \in \mathcal{S}(\mathcal{H}^{\otimes m_l})$, $D_1, \ldots, D_{l^2} \in \mathcal{B}(\mathcal{K}^{\otimes m_l})$ with $m_l = o(l)$ and

$$\bar{P}_{e,m_l} = \sup_{s^{m_l} \in \mathbf{S}^{m_l}} P_e(\mathfrak{C}_{m_l}, s^{m_l}) \leq \varepsilon. \tag{108}$$

On the other hand, let us consider an (l, k_l)-random code as in Lemma 9, i.e. with

$$\frac{1}{l^2}\sum_{i=1}^{l^2} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}_i^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}_i^l) > 1 - \varepsilon \qquad \forall s^l \in \mathbf{S}^l. \tag{109}$$

Define CPTP maps $\mathcal{P}^{l+m_l} \in \mathcal{C}(\mathcal{F}_l, \mathcal{H}^{\otimes l+m_l})$, $\mathcal{R}^{l+m_l} \in \mathcal{C}(\mathcal{K}^{\otimes l+m_l}, \mathcal{F}_l')$ by

$$\mathcal{P}^{l+m_l}(a) := \frac{1}{l^2}\sum_{i=1}^{l^2} \mathcal{P}_i^l(a) \otimes \rho_i \quad \text{and} \quad \mathcal{R}^{l+m_l}(b \otimes d) := \sum_{i=1}^{l^2} \operatorname{tr}(D_i d)\mathcal{R}_i^l(b). \tag{110}$$

Then for each $s^{l+m_l} = (v^l, u^{m_l},) \in \mathbf{S}^{l+m_l}$

$$\begin{aligned} F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^{l+m_l} \circ (\mathcal{N}_{v^l} \otimes \mathcal{N}_{u^{m_l}}) \circ \mathcal{P}^{l+m_l}) &= \frac{1}{l^2}\sum_{i,j=1}^{l^2} \operatorname{tr}(D_j \mathcal{N}_{u^{m_l}}(\rho_i)) F_e(\pi_{\mathcal{F}_l}, \mathcal{R}_j^l \circ \mathcal{N}_{v^l} \circ \mathcal{P}_i^l) \\ &\geq \frac{1}{l^2}\sum_{i=1}^{l^2} \operatorname{tr}(D_i \mathcal{N}_{u^{m_l}}(\rho_i)) F_e(\pi_{\mathcal{F}_l}, \mathcal{R}_i^l \circ \mathcal{N}_{v^l} \circ \mathcal{P}_i^l), \end{aligned} \tag{111}$$

where in the last line we have used that all involved terms are non-negative. In order to show that the fidelity on the left-hand side of (111) is at least $1-2\varepsilon$ we need the following lemma from [2].

Lemma 10. *Let $K \in \mathbb{N}$ and real numbers $a_1, \ldots, a_K, b_1, \ldots, b_K \in [0,1]$ be given. Assume that*

$$\frac{1}{K}\sum_{i=1}^{K} a_i \geq 1-\varepsilon \qquad \text{and} \qquad \frac{1}{K}\sum_{i=1}^{K} b_i \geq 1-\varepsilon,$$

hold. Then

$$\frac{1}{K}\sum_{i=1}^{K} a_i b_i \geq 1-2\varepsilon.$$

Applying this lemma with $K = l^2$,

$$a_i = \operatorname{tr}(D_i \mathcal{N}_{u^{m_l}}(\rho_i)), \quad \text{and} \quad b_i = F_e(\pi_{\mathcal{F}_l}, \mathcal{R}_i^l \circ \mathcal{N}_{v^l} \circ \mathcal{P}_i^l)$$

along with (108), (109), and (111) shows that

$$F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^{l+m_l} \circ (\mathcal{N}_{v^l} \otimes \mathcal{N}_{u^{m_l}}) \circ \mathcal{P}^{l+m_l}) \geq 1-2\varepsilon. \tag{112}$$

On the other hand we know from Theorem 9 that for each sufficiently small $\eta > 0$ there is a random code μ_l for the AVQC $\mathfrak{I}$ with

$$\frac{1}{l}\log\dim\mathcal{F}_l \geq \frac{1}{k}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes k})}\inf_{\mathcal{N}\in\operatorname{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes k}) - \eta, \tag{113}$$

and

$$e(\mu_l, \mathfrak{I}) = \inf_{s^l\in\mathbf{S}^l}\int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l)d\mu_l(\mathcal{P}^l, \mathcal{R}^l) \geq 1-(l+1)^{N_\eta}2^{-lc}$$

for all sufficiently large l with $c = c(k, \dim\mathcal{H}, \dim\mathcal{K}, \operatorname{conv}(\mathfrak{I}), \eta)$ and $N_\eta \in \mathbb{N}$. Thus the arguments that led us to (112) show that for all sufficiently large l there is a deterministic $(l+m_l, k_l)$-code for the AVQC $\mathfrak{I}$ with

$$F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^{l+m_l} \circ (\mathcal{N}_{v^l} \otimes \mathcal{N}_{u^{m_l}}) \circ \mathcal{P}^{l+m_l}) \geq 1-2\varepsilon,$$

and

$$\frac{1}{l+m_l}\log\dim\mathcal{F}_l \geq \frac{1}{k}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes k})}\inf_{\mathcal{N}\in\operatorname{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes k}) - 2\eta$$

by (113) and since $m_l = o(l)$. This shows that $\mathcal{A}_{\text{det}}(\mathfrak{I}) \geq \mathcal{A}_{\text{random}}(\mathfrak{I})$. Since the reverse inequality is trivially true we are done.

8 Zero-capacity-conditions: Symmetrizability

The most basic quality feature of a information processing system is whether it can be used for communication at a positive rate or not. This applies especially to such rather complex systems as AVCs or AVQCs. The notion of symmetrizability stems from the theory of classical AVCs and it addresses exactly that question. A classical AVC has deterministic capacity for message transmission equal to zero if and only if it is symmetrizable (with the definition of symmetrizability adjusted to the two different scenarios 'average error criterion' and 'maximal error criterion') [17] and [14], [22].
Of course, a similar statement for $\mathcal{A}_{\text{det}}$ would be of great interest.

In this section we give three different conditions for three different capacities of an AVQC to be equal to zero. We restrict ourselves to the case $|\mathfrak{I}| < \infty$. Starting with the statement that has the weakest

information theoretic consequences, we proceed to stronger statements. The case $|\mathfrak{I}| = \infty$ requires some involved continuity issues which shall be carried out elsewhere.
All three conditions have in common that they enable the adversary to simulate, on average over some probability distribution, a different output at the receiver side than the one that was originally put into the channel by the sender. The first two conditions, dealing with message transmission, exhibit a possibly nonlinear dependence between message set and probability distribution. They are direct (but not single-letter) analogs of their classical counterparts.
The third one is a sufficient condition for $\mathcal{A}_{\mathrm{random}}$ to be equal to zero. It employs a linear dependence between input state and probability distribution. If this condition is valid, the adversary is not only able to simulate a wrong output, he can also simulate an entanglement breaking channel between sender and receiver. In contrast to the first two criteria, this third one is a single-letter criterion.
There is a fourth and, at first sight, trivial condition, given by the following: An AVQC $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ has (deterministic *and* random) capacity for transmission of entanglement equal to zero if there is an $s \in \mathbf{S}$ such that $\mathcal{N}_s$ has zero capacity for transmission of entanglement.
We note that this fourth condition is nontrivial only because of the following reason: there is, until now, no way of telling exactly when a given (memoryless) quantum channel has a capacity greater than zero (except for calculating (4) for a single channel, an awkward task in general). This is in sharp contrast to the classical case, where the question can be trivially answered: A classical memoryless channel has a nonzero capacity if and only if there are at least two input states that lead to different output states.
Since our results do not answer the question whether it can happen that $C_{\mathrm{det}}(\mathfrak{I}) = 0$, $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) = 0$ and $\mathcal{A}_{\mathrm{random}}(\mathfrak{I}) > 0$ hold simultaneously for a given AVQC $\mathfrak{I}$, we are left with two interesting and intimately related questions:
First, there is the zero-capacity question for single memoryless channels. Second, we need to find a criterion telling us exactly when $\mathcal{A}_{\mathrm{det}}$ is equal to zero.

8.1 Classical capacity with deterministic codes and average error

We now introduce a notion of symmetrizability which is a sufficient and necessary condition for $C_{\mathrm{det}}(\mathfrak{I}) = 0$. Our approach is motivated by the corresponding concept for arbitrarily varying channels with classical input and quantum output (cq-AVC) given in [5].

Definition 12. *Let* $\mathbf{S}$ *be a finite set and* $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ *an AVQC.*

1. $\mathfrak{I}$ *is called* l*-symmetrizable,* $l \in \mathbb{N}$*, if for each finite set* $\{\rho_1, \ldots, \rho_K\} \subset \mathcal{S}(\mathcal{H}^{\otimes l})$*,* $K \in \mathbb{N}$*, there is a map* $p : \{\rho_1, \ldots, \rho_K\} \to \mathfrak{P}(\mathbf{S}^l)$ *such that for all* $i, j \in \{1, \ldots, K\}$
$$\sum_{s^l\in\mathbf{S}^l} p(\rho_i)(s^l)\mathcal{N}_{s^l}(\rho_j) = \sum_{s^l\in\mathbf{S}^l} p(\rho_j)(s^l)\mathcal{N}_{s^l}(\rho_i) \tag{114}$$
holds.

2. *We call* $\mathfrak{I}$ *symmetrizable if it is* l*-symmetrizable for all* $l \in \mathbb{N}$*.*

We now state the main statement of this section.

Theorem 12. *Let* $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$*,* $|\mathbf{S}| < \infty$*, be an AVQC. Then* $\mathfrak{I}$ *is symmetrizable if and only if* $C_{\mathrm{det}}(\mathfrak{I}) = 0$*.*

Proof. 1. "Symmetrizability implies $C_{\mathrm{det}}(\mathfrak{I}) = 0$".
The proof follows closely the corresponding arguments given in [17], [14], and [5]. We give the full proof for reader's convenience. Suppose that $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ is symmetrizable and let $(\rho_i, D_i)_{i=1}^M$, $M \geq 2$, be a code for transmission of messages over $\mathfrak{I}$ with $\{\rho_1, \ldots, \rho_M\} \subset \mathcal{S}(\mathcal{H}^{\otimes l})$ and POVM $\{D_i\}_{i=1}^M$ on $\mathcal{H}^{\otimes l}$. Since $\mathfrak{I}$ is symmetrizable there is a map $p : \{\rho_1, \ldots, \rho_M\} \to \mathfrak{P}(\mathbf{S}^l)$ such that for all $i, j \in \{1, \ldots, M\}$
$$\sum_{s^l\in\mathbf{S}^l} p(\rho_i)(s^l)\mathcal{N}_{s^l}(\rho_j) = \sum_{s^l\in\mathbf{S}^l} p(\rho_j)(s^l)\mathcal{N}_{s^l}(\rho_i) \tag{115}$$

For $s^l \in \mathbf{S}^l$ and $i \in \{1, \ldots, M\}$ we set

$$e(i, s^l) := 1 - \mathrm{tr}(\mathcal{N}_{s^l}(\rho_i) D_i) = \sum_{\substack{j=1 \\ j \neq i}}^{M} \mathrm{tr}(\mathcal{N}_{s^l}(\rho_i) D_j). \tag{116}$$

For $k \in \{1, \ldots, M\}$ let S_k^l be a random variable taking values in $\mathbf{S}^l$ and which is distributed according to $(p(\rho_k)(s^l))_{s^l \in \mathbf{S}^l}$. Then using relation (116) we can write

$$\begin{aligned} \mathbb{E}(e(i, S_k^l)) &= \sum_{s^l \in \mathbf{S}^l} \sum_{\substack{j=1 \\ j \neq i}}^{M} p(\rho_k)(s^l) \mathrm{tr}(\mathcal{N}_{s^l}(\rho_i) D_j) \\ &= \sum_{\substack{j=1 \\ j \neq i}}^{M} \mathrm{tr}\{ \sum_{s^l \in \mathbf{S}^l} p(\rho_k)(s^l) \mathcal{N}_{s^l}(\rho_i) D_j \} \\ &= \sum_{\substack{j=1 \\ j \neq i}}^{M} \mathrm{tr}(\sum_{s^l \in \mathbf{S}^l} p(\rho_i)(s^l) \mathcal{N}_{s^l}(\rho_k) D_j) \\ &= \sum_{\substack{j=1 \\ j \neq i}}^{M} \sum_{s^l \in \mathbf{S}^l} p(\rho_i)(s^l) \mathrm{tr}(\mathcal{N}_{s^l}(\rho_k) D_j), \end{aligned} \tag{117}$$

where the third line is by (115). On the other hand we have

$$\mathbb{E}(e(k, S_i^l)) = \sum_{s^l \in \mathbf{S}^l} \sum_{\substack{j=1 \\ j \neq k}}^{M} p(\rho_i)(s^l) \mathrm{tr}(\mathcal{N}_{s^l}(\rho_k) D_j). \tag{118}$$

Since $\{D_i\}_{i=1}^M$ is a POVM (117) and (118) imply that for $i \neq k$

$$\mathbb{E}(e(i, S_k^l)) + \mathbb{E}(e(k, S_i^l)) \geq 1 \tag{119}$$

holds. Let us abbreviate $\mathfrak{C} := (\rho_i, D_i)_{i=1}^M$, then with

$$\bar{P}_e(\mathfrak{C}, s^l) = \frac{1}{M} \sum_{k=1}^{M} (1 - \mathrm{tr}(\mathcal{N}_{s^l}(\rho_k) D_k))$$

for $s^l \in \mathbf{S}^l$ we obtain

$$\begin{aligned} \mathbb{E}(\bar{P}_e(\mathfrak{C}, S_j^l)) &= \sum_{s^l \in \mathbf{S}^l} p(\rho_j)(s^l) \frac{1}{M} \sum_{k=1}^{M} (1 - \mathrm{tr}(\mathcal{N}_{s^l}(\rho_k) D_k)) \\ &= \frac{1}{M} \sum_{k=1}^{M} \mathbb{E}(e(k, S_j^l)). \end{aligned} \tag{120}$$

(119) and (120) yield

$$\begin{aligned} \frac{1}{M} \sum_{j=1}^{M} \mathbb{E}(\bar{P}_e(\mathfrak{C}, S_j^l)) &= \frac{1}{M^2} \sum_{i,j=1}^{M} \mathbb{E}(e(k, S_j^l)) \\ &\geq \frac{1}{M^2} \binom{M}{2} \\ &= \frac{M-1}{2M} \geq \frac{1}{4} \end{aligned}$$

for $M \geq 2$. Thus it follows that there is at least one $j \in \{1, \ldots, M\}$ with

$$\mathbb{E}(\bar{P}_e(\mathfrak{C}, S_j^l)) \geq \frac{1}{4}$$

and consequently there is at least one $s^l \in \mathbf{S}^l$ with

$$\bar{P}_e(\mathfrak{C}, s^l) \geq \frac{1}{4}$$

implying that $C_{\mathrm{det}}(\mathfrak{I}) = 0$.
2. "$C_{\mathrm{det}}(\mathfrak{I}) = 0$ implies symmetrizability".
Suppose that $\mathfrak{I}$ is non-symmetrizable. Then there is an $\hat{l} \in \mathbb{N}$ and a finite set $\{\rho_x\}_{x \in \mathcal{X}} \subset \mathcal{S}(\mathcal{H}^{\otimes \hat{l}})$ such that for no map $p : \{\rho_x\}_{x \in \mathcal{X}} \to \mathfrak{P}(\mathbf{S}^{\hat{l}})$ the relation (114) holds. Let us define for each $s^{\hat{l}} \in \mathbf{S}^{\hat{l}}$ a cq-channel $\mathcal{X} \ni x \mapsto W_{s^{\hat{l}}}(x) := \mathcal{N}_{s^{\hat{l}}}(\rho_x) \in \mathcal{S}(\mathcal{K}^{\otimes \hat{l}})$, and consider the cq-AVC generated by the set $\mathfrak{I}_{cq} := \{W_{s^{\hat{l}}}\}_{s^{\hat{l}} \in \mathbf{S}^{\hat{l}}}$. Then, due to the assumed non-symmetrizability of $\mathfrak{I}$, our new cq-AVC $\mathfrak{I}_{cq}$ is non-symmetrizable in the sense of [5].
Since $\mathfrak{I}_{cq}$ is non-symmetrizable the reduction argument from [5] to the results of [14] show that the cq-AVC $\mathfrak{I}_{cq}$ has positive capacity. This implies the existence of a sequence $(K_m, f_m, D_m, \varepsilon_m)_{m \in \mathbb{N}}$, where $K_m \in \mathbb{N}$, $f_m : \{1, \ldots, K_m\} \to \mathcal{X}^m$, $D_m \in \mathcal{B}_+(\mathcal{H}^{\otimes l \cdot m})$, $\lim_{m \to \infty} \varepsilon_m \searrow 0$, $\liminf_{m \to \infty} \frac{1}{m} \log K_m = c > 0$ and $\frac{1}{K_m} \sum_{i=1}^{K_m} (1 - \mathrm{tr}(D_i W_{s^{\hat{l}}}^m (f(i)))) = \varepsilon_m$.
We may use this sequence to construct another sequence $(\rho_i, D_i)_{i=1}^{M_l}$ of deterministic codes for message transmission over $\mathfrak{I}$, thereby achieving a capacity of $\frac{1}{\hat{l}} c > 0$. A similar construction is carried out explicitly at the end of the proof of the following Theorem 14. □

Corollary 13. *If the AVQC $\mathfrak{I} = \{\mathcal{N}\}_{s \in \mathbf{S}}$ is symmetrizable then $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) = 0$.*

Proof. Note that $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) \leq C_{\mathrm{det}}(\mathfrak{I})$ and apply Theorem 12. □

8.2 Classical capacity with deterministic codes and maximal error

We will now investigate, when exactly it is possible to send classical messages at positive rate over a finite AVQC, with the error criterion being that of maximal rather than average error.

Theorem 14. *Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ be a finite AVQC. The classical deterministic maximal error capacity $C_{\mathrm{det,max}}(\mathfrak{I})$ of $\mathfrak{I}$ is equal to zero if and only if for every $l \in \mathbb{N}$ and every set $\{\rho_1, \rho_2\} \subset \mathcal{S}(\mathcal{H}^{\otimes l})$ we have*

$$\mathrm{conv}(\{\mathcal{N}_{s^l}(\rho_1)\}_{s^l \in \mathbf{S}^l}) \cap \mathrm{conv}(\{\mathcal{N}_{s^l}(\rho_2)\}_{s^l \in \mathbf{S}^l}) \neq \emptyset. \tag{121}$$

Proof. We closely follow the line of proof given in [22]. Let us begin with the 'if' part. Let $K, l \in \mathbb{N}$, $\{\rho_1, \ldots, \rho_K\} \subset \mathcal{S}(\mathcal{H}^{\otimes l})$ and $D_1, \ldots, D_K \in \mathcal{B}_+(\mathcal{K}^{\otimes l})$ with $\sum_{i=1}^K D_i = \mathbf{1}_{\mathcal{K}^{\otimes l}}$ be a code for transmission of classical messages over $\mathfrak{I}$.
We show that the maximal error probability of this code is bounded away from zero for large enough l. Let, without loss of generality, l be such that

$$\mathrm{tr}(D_1 \mathcal{N}_{s^l}(\rho_1)) > 1/2 \qquad \forall\ s^l \in \mathbf{S}^l \tag{122}$$
$$\mathrm{tr}(D_2 \mathcal{N}_{s^l}(\rho_2)) > 1/2 \qquad \forall\ s^l \in \mathbf{S}^l. \tag{123}$$

We show that there is a contradiction between (122) and (123). By assumption, there exist probability distributions $p_1, p_2 \in \mathfrak{P}(\mathbf{S}^l)$ such that

$$\sum_{s^l \in \mathbf{S}^l} p_1(s^l) \mathcal{N}_{s^l}(\rho_1) = \sum_{s^l \in \mathbf{S}^l} p_2(s^l) \mathcal{N}_{s^l}(\rho_2). \tag{124}$$

Of course, (122) implies

$$\sum_{s^l\in\mathbf{S}^l} p_1(s^l)\mathrm{tr}(D_1\mathcal{N}_{s^l}(\rho_1)) > 1/2. \tag{125}$$

Together with (124) this leads to

$$\begin{aligned}
1/2 &< \sum_{s^l\in\mathbf{S}^l} p_1(s^l)\mathrm{tr}(D_1\mathcal{N}_{s^l}(\rho_1))\\
&= \sum_{s^l\in\mathbf{S}^l} p_2(s^l)\mathrm{tr}(D_1\mathcal{N}_{s^l}(\rho_2))\\
&\leq \sum_{s^l\in\mathbf{S}^l} p_2(s^l)\mathrm{tr}((D_1+\sum_{i=3}^{K} D_i)\mathcal{N}_{s^l}(\rho_2))\\
&= \sum_{s^l\in\mathbf{S}^l} p_2(s^l)\mathrm{tr}((\mathbf{1}-D_2)\mathcal{N}_{s^l}(\rho_2))\\
&= 1-\sum_{s^l\in\mathbf{S}^l} p_2(s^l)\mathrm{tr}(D_2\mathcal{N}_{s^l}(\rho_2))\\
&< 1-1/2,
\end{aligned} \tag{126}$$

a clear contradiction. Thus, for every code the maximal error probability is bounded from below by 1/2.
Let us turn to the 'only if' part.
Assume there is an $\hat{l}\in\mathbb{N}$ and a set $\{\rho_1,\rho_2\}\subset\mathcal{S}(\mathcal{H}^{\otimes\hat{l}})$ such that

$$\mathrm{conv}(\{\mathcal{N}_{s^{\hat{l}}}(\rho_1)\}_{s^{\hat{l}}\in\mathbf{S}^{\hat{l}}})\cap\mathrm{conv}(\{\mathcal{N}_{s^{\hat{l}}}(\rho_2)\}_{s^{\hat{l}}\in\mathbf{S}^{\hat{l}}})=\emptyset. \tag{127}$$

Thus, there exists a self adjoint operator $A\in\mathcal{B}(\mathcal{K}^{\otimes\hat{l}})$ such that

$$\mathrm{tr}(A\rho)<0\ \ \forall\,\rho\in\mathrm{conv}(\{\mathcal{N}_{s^{\hat{l}}}(\rho_1)\}_{s^{\hat{l}}\in\mathbf{S}^{\hat{l}}}),\qquad \mathrm{tr}(A\rho)>0\ \ \forall\,\rho\in\mathrm{conv}(\{\mathcal{N}_{s^{\hat{l}}}(\rho_2)\}_{s^{\hat{l}}\in\mathbf{S}^{\hat{l}}}). \tag{128}$$

Let A have a decomposition $A=\sum_{x=1}^d a_xA_x$, where a_x are real numbers (including the possibility of $a_x=0$ for some x) and A_x are one dimensional projections fulfilling $\sum_{x=1}^d A_x=\mathbf{1}_{\mathcal{K}^{\otimes\hat{l}}}$. For every $m\in\mathbb{N}$, define

$$P_1^m := \sum_{x^m:\frac{1}{m}\sum_{i=1}^m a_{x_i}<0} A_{x_1}\otimes\ldots\otimes A_{x_m},\qquad P_2^m := \sum_{x^m:\frac{1}{m}\sum_{i=1}^m a_{x_i}\geq 0} A_{x_1}\otimes\ldots\otimes A_{x_m}. \tag{129}$$

Then $P_1^m+P_2^m=\mathbf{1}_{\mathcal{K}^{\otimes\hat{l}\cdot m}}$. Let us denote elements of $\mathbf{S}^{\hat{l}m}$ by $s^{\hat{l}m}=(s_1^{\hat{l}},\ldots,s_m^{\hat{l}})$, where each $s_i^{\hat{l}}\in\mathbf{S}^{\hat{l}}$.
To every $s^{\hat{l}}\in\mathbf{S}^{\hat{l}}$, define probability distributions $p_{s^{\hat{l}}},q_{s^{\hat{l}}}\in\mathcal{S}(\{1,\ldots,d\})$ according to

$$p_{s^{\hat{l}}}(x):=\mathrm{tr}(A_x\mathcal{N}_{s^{\hat{l}}}(\rho_1)),\qquad q_{s^{\hat{l}}}(x):=\mathrm{tr}(A_x\mathcal{N}_{s^{\hat{l}}}(\rho_2)),\ \ \forall\,x\in\{1,\ldots,d\} \tag{130}$$

and to every $s^{\hat{l}m}\in\mathbf{S}^{\hat{l}m}$ we associate two real numbers $\bar{A}_{s^{\hat{l}m}}(\rho_1),\bar{A}_{s^{\hat{l}m}}(\rho_2)$ by

$$\bar{A}_{s^{\hat{l}m}}(\rho_1):=\sum_{s^{\hat{l}}\in\mathbf{S}^{\hat{l}}}\frac{1}{m}N(s^{\hat{l}}|s^{\hat{l}m})\mathrm{tr}(A\mathcal{N}_{s^{\hat{l}}}(\rho_1)),\qquad \bar{A}_{s^{\hat{l}m}}(\rho_2):=\sum_{s^{\hat{l}}\in\mathbf{S}^{\hat{l}}}\frac{1}{m}N(s^{\hat{l}}|s^{\hat{l}m})\mathrm{tr}(A\mathcal{N}_{s^{\hat{l}}}(\rho_2)), \tag{131}$$

with natural numbers $N(s^{\hat{l}}|s^{\hat{l}m}):=|\{i:s_i^{\hat{l}}=s^{\hat{l}},\ i\in\{1,\ldots,m\}\}|$ for every $s^{\hat{l}m}\in\mathbf{S}^{\hat{l}m}$ and $s^{\hat{l}}\in\mathbf{S}^{\hat{l}}$.
Obviously, $\bar{A}_{s^{\hat{l}m}}(\rho_1)<0$ and $\bar{A}_{s^{\hat{l}m}}(\rho_2)>0$. Setting

$$C:=\max_{(s^{\hat{l}},X)\in\mathbf{S}^{\hat{l}}\times\{\rho_1,\rho_2\}}(\mathrm{tr}(A\mathcal{N}_{s^{\hat{l}}}(X))/2)^{-2}(\mathrm{tr}(A^2\mathcal{N}_{s^{\hat{l}}}(X))-\mathrm{tr}(A\mathcal{N}_{s^{\hat{l}}}(X))^2) \tag{132}$$

we arrive, by application of Chebyshev's inequality and for every $s^{\hat{l}m} = (s^{\hat{l}}_1, \ldots, s^{\hat{l}}_m) \in \mathbf{S}^{\hat{l}m}$ at

$$\begin{aligned}
\mathrm{tr}(P_1^m \mathcal{N}_{s^{\hat{l}m}}(\rho_1^{\otimes m})) &= \sum_{x^m : \frac{1}{m}\sum_{i=1}^m a_{x_i} < 0} \mathrm{tr}(A_{x_1} \otimes \ldots \otimes A_{x_m} \mathcal{N}_{s^{\hat{l}}_1}(\rho_1) \otimes \ldots \otimes \mathcal{N}_{s^{\hat{l}}_m}(\rho_1)) \\
&= \sum_{x^m : \frac{1}{m}\sum_{i=1}^m a_{x_i} < 0} p_{s^{\hat{l}}_1}(x_1) \cdot \ldots \cdot p_{s^{\hat{l}}_m}(x_m) \\
&\geq \sum_{x^m : |\frac{1}{m}\sum_{i=1}^m a_{x_i} - \bar{A}_{s^{\hat{l}m}}(\rho_1)| \leq |\bar{A}_{s^{\hat{l}m}}(\rho_1)/2|} p_{s^{\hat{l}}_1}(x_1) \cdot \ldots \cdot p_{s^{\hat{l}}_m}(x_m) \\
&\geq 1 - \frac{1}{m}(\bar{A}_{s^{\hat{l}m}}(\rho_1)/2)^{-2} \sum_{s^{\hat{l}} \in \mathbf{S}^{\hat{l}}} \frac{1}{m} N(s^{\hat{l}}|s^{\hat{l}m})(\mathrm{tr}(A^2 \mathcal{N}_{s^{\hat{l}}}(\rho_1)) - \mathrm{tr}(A\mathcal{N}_{s^{\hat{l}}}(\rho_1))^2) \\
&\geq 1 - \frac{1}{m} \max_{s^{\hat{l}}}(\mathrm{tr}(A\mathcal{N}_{s^{\hat{l}}}(\rho_1))/2)^{-2}(\mathrm{tr}(A^2\mathcal{N}_{s^{\hat{l}}}(\rho_1)) - \mathrm{tr}(A\mathcal{N}_{s^{\hat{l}}}(\rho_1))^2) \\
&\geq 1 - \frac{1}{m} \cdot C. \qquad (133)
\end{aligned}$$

In the very same way, we can prove that

$$\begin{aligned}
\mathrm{tr}(P_2^m \mathcal{N}_{s^{\hat{l}m}}(\rho_2^{\otimes m})) &= \sum_{x^m : \frac{1}{m}\sum_{i=1}^m a_{x_i} \geq 0} \mathrm{tr}(A_{x_1} \otimes \ldots \otimes A_{x_m} \mathcal{N}_{s^{\hat{l}}_1}(\rho_2) \otimes \ldots \otimes \mathcal{N}_{s^{\hat{l}}_m}(\rho_2)) \\
&= \sum_{x^m : \frac{1}{m}\sum_{i=1}^m a_{x_i} \geq 0} q_{s^{\hat{l}}_1}(x_1) \cdot \ldots \cdot q_{s^{\hat{l}}_m}(x_m) \\
&\geq \sum_{x^m : |\frac{1}{m}\sum_{i=1}^m a_{x_i} - \bar{A}_{s^{\hat{l}m}}(\rho_2)| \leq |\bar{A}_{s^{\hat{l}m}}(\rho_2)|/2} q_{s^{l}_1}(x_1) \cdot \ldots \cdot q_{s^{\hat{l}}_m}(x_m) \\
&\geq 1 - \frac{1}{m} \max_{s^{\hat{l}}}(\mathrm{tr}(A\mathcal{N}_{s^{\hat{l}}}(\rho_2))/2)^{-2}(\mathrm{tr}(A^2\mathcal{N}_{s^{\hat{l}}}(\rho_2)) - \mathrm{tr}(A\mathcal{N}_{s^{\hat{l}}}(\rho_2))^2) \\
&\geq 1 - \frac{1}{m} \cdot C. \qquad (134)
\end{aligned}$$

Take any $0 < \varepsilon < 1/4$. Let $m' = \min\{m \in \mathbb{N} : \frac{1}{m} \cdot C < \varepsilon\}$. Then

$$\mathrm{tr}(P_1^{m'} \mathcal{N}_{s^{\hat{l}m'}}(\rho_1^{\otimes m'})) \geq 1 - \varepsilon \qquad\qquad \mathrm{tr}(P_2^{m'} \mathcal{N}_{s^{\hat{l}m'}}(\rho_2^{\otimes m'})) \geq 1 - \varepsilon \qquad (135)$$

hold. Consider the classical AVC given by the family $J := \{c_{\nu,\delta}\}_{\delta,\nu \in [3/4,1]}$ of classical channels $c_{\nu,\delta} : \{0,1\} \to \{0,1\}$ with stochastic matrices defined via $c_{\nu,\delta}(1|1) := 1 - \nu,\ c_{\nu,\delta}(2|2) := 1 - \delta$. Clearly, J is a convex set and, for every $c_{\nu,\delta} \in J$ we have that

$$\begin{aligned}
\max_{p \in \mathfrak{P}(\{0,1\})} I(p, c_{\nu,\delta}) &\geq 1 - \frac{1}{2}(h(\nu) + h(\delta)) \\
&\geq 1 - h(3/4) \\
&> 0, \qquad (136)
\end{aligned}$$

where $I(p, c_{\nu,\delta})$ is the mutual information of the probability distribution q on $\{1,2\} \times \{1,2\}$ which is generated by p and $c_{\nu,\delta}$ through $q(i,j) := p(i)c_{\nu,\delta}(j|i)$ $((i,j) \in \{1,2\} \times \{1,2\})$. The lower bound given here is calculated using an equidistributed input. Note further that for this special AVC, with notation taken from [6], $\bar{\bar{J}} = \mathrm{conv}(J) = J$.

At this point in their proof of the classical zero-capacity-condition for AVCs [22], Kiefer and Wolfowitz made reference to a result by Gilbert [18], who proved existence of codes that achieve a positive rate.

Kiefer and Wolfowitz used these codes for message transmission over an AVC with binary input and output alphabet. Our strategy of proof is to use the existence of codes for AVCs with binary input and output that is guaranteed by Theorem 1 of [6] instead. Together with (136) this theorem gives us the existence of a number $C' > 0$, a function $\kappa : \mathbb{N} \to \mathbb{R}$ with $\lim_{r\to\infty} \kappa(r) = 0$ and a sequence $(M^r, f^r, \varepsilon_r, (D_1^r, \ldots, D_{|M^r|}^r))_{r\in\mathbb{N}}$ where for each $r \in \mathbb{N}$:

1. $M^r = \{1, \ldots, N\}$ is a finite set of cardinality $N = |M^r| = 2^{r(C'-\kappa(r))}$,
2. $f^r : M^r \to \{1, 2\}^r$,
3. $\varepsilon_r \geq 0$ and $\lim_{r\to\infty} \varepsilon_r = 0$,
4. $D_1^r, \ldots, D_{|M^r|}^r \subset \{1, 2\}^n$ are pairwise disjoint and
5. for every sequence $x^r \in ([3/4, 1] \times [3/4, 1])^r$ and every $i \in M^r$ we have that
$$\sum_{y^n \in D_i^r} \prod_{j=1}^{r} c_{x_j}(y_j | f^r(i)_j) \geq 1 - \varepsilon_r. \tag{137}$$

For $n \in \mathbb{N}$, take the unique numbers $r \in \mathbb{N}$, $t \in \{0, \ldots, m'-1\}$ such that $n = m'r + t$ holds. The code for $\mathfrak{I}$ is then defined as follows:

$$\begin{aligned}
M_n &:= M^r,\\
f_n(i) &:= (\rho_{f^r(i)_1})^{\otimes m'} \otimes \ldots \otimes (\rho_{f^r(i)_r})^{\otimes m'} \otimes \sigma^{\otimes t},\\
P_i^n &:= \sum_{y^r \in D_i^r} P_{y_1}^{m'} \otimes \ldots \otimes P_{y_r}^{m'} \otimes \mathbf{1}_{\mathcal{K}}^{\otimes t}.
\end{aligned}$$

Let, for every $s^{m'} \in \mathbf{S}^{m'}$, $x = (\nu, \delta) \in [3/4, 1]^2$ be such that

$$c_{\nu,\delta}(0|0) := \mathrm{tr}(P_1^{m'} \mathcal{N}_{s^{nm'}}(\rho_1^{\otimes m'})) = 1 - \nu \qquad c_{\nu,\delta}(1|1) := \mathrm{tr}(P_1^{m'} \mathcal{N}_{s^{nm'}}(\rho_2^{\otimes m'})) = 1 - \delta. \tag{138}$$

Then for every $s^n \in \mathbf{S}^n$ we use the decomposition $s^n = (s_1^{m'}, \ldots, s_r^{m'}, s^t)$ and get, using equation (137) and the definition (138), for every $i \in M_n$,

$$\begin{aligned}
\mathrm{tr}\{P_i^n f_n(i)\} &= \mathrm{tr}\{[\sum_{y^r \in D_i^r} P_{y_1}^{m'} \otimes \ldots \otimes P_{y_r}^{m'} \otimes \mathbf{1}_{\mathcal{K}}^{\otimes t}](\rho_{f^r(i)_1})^{\otimes m'} \otimes \ldots \otimes (\rho_{f^r(i)_r})^{\otimes m'} \otimes \sigma^{\otimes t}\}\\
&= \sum_{y^r \in D_i^r} \prod_{j=1}^{r} \mathrm{tr}\{P_{y_j}^{m'} (\rho_{f^r(i)_j})^{\otimes m'}\}\\
&= \sum_{y^r \in D_i^r} \prod_{j=1}^{r} c_{\nu,\delta}(y_j | f^r(i)_j)\\
&\geq 1 - \varepsilon_r.
\end{aligned} \tag{139}$$

Obviously, this implies

$$\lim_{n\to\infty} \min_{s^n \in \mathbf{S}^n} \max_{i \in M_n} \mathrm{tr}\{P_i^n f_n(i)\} = 0. \tag{140}$$

Together with

$$\lim_{n\to\infty} \frac{1}{n} \log |M_n| = \frac{1}{m'} C' > 0 \tag{141}$$

we have shown that $C_{det,max}(\mathfrak{I}) > 0$ holds. □

Notice that the statements made in (124) and (121) equivalent and a glance at Definition 12 reveals that the assertion of (124) is nothing else than the symmetrizability restricted to sets of states consisting of two elements.

8.3 Entanglement transmission capacity with random codes

The final issue in this section is a sufficient condition for $\mathcal{A}_{\mathrm{random}}(\mathfrak{I}) = 0$ which is based on the notion of qc-symmetrizability.
Let $\mathfrak{F}_{\mathbb{C}}(\mathbf{S})$ stand for the set of $\mathbb{C}$-valued functions defined on $\mathbf{S}$ in what follows and we consider the set of channels with quantum input and classical output (qc-channels)[2]

$$\mathrm{QC}(\mathcal{H}, \mathbf{S}) := \{T : \mathcal{B}(\mathcal{H}) \to \mathfrak{F}_{\mathbb{C}}(\mathbf{S}) : T \text{ is linear, positive, and trace preserving}\}.$$

The condition that $T \in \mathrm{QC}(\mathcal{H}, \mathbf{S})$ is trace preserving means that

$$\sum_{s \in \mathbf{S}} [T(b)](s) = \mathrm{tr}(b)$$

holds for all $b \in \mathcal{B}(\mathcal{H})$. By Riesz' representation theorem there is a one-to-one correspondence between elements $T \in \mathrm{QC}(\mathcal{H}, \mathbf{S})$ and (discrete) positive operator-valued measures (POVM) $\{E_s\}_{s \in \mathbf{S}}$.
For a given finite set of quantum channels $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$ and $T \in \mathrm{QC}(\mathcal{H}, \mathbf{S})$ we define a CPTP map $\mathcal{M}_{T,\mathbf{S}} : \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ by

$$\begin{aligned} \mathcal{M}_{T,\mathbf{S}}(a \otimes b) &:= \sum_{s \in \mathbf{S}} [T(a)](s) \mathcal{N}_s(b) \\ &= \sum_{s \in \mathbf{S}} \mathrm{tr}(E_s a) \mathcal{N}_s(b), \end{aligned} \tag{142}$$

where $\{E_s\}_{s \in \mathbf{S}}$ is the unique POVM associated with T.

Definition 13. *An arbitrarily varying quantum channel, generated by a finite set $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$, is called qc-symmetrizable if there is $T \in QC(\mathcal{H}, \mathbf{S})$ such that for all $a, b \in \mathcal{B}(\mathcal{H})$*

$$\mathcal{M}_{T,\mathbf{S}}(a \otimes b) = \mathcal{M}_{T,\mathbf{S}}(b \otimes a) \tag{143}$$

holds, where $\mathcal{M}_{T,\mathbf{S}} : \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ is the CPTP map defined in (142).

The best illustration of the definition of qc-symmetrizability is given in the proof of our next theorem.

Theorem 15. *If an arbitrarily varying quantum channel generated by a finite set $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$ is qc-symmetrizable, then for any sequence of (l, k_l)-random codes $(\mu_l)_{l \in \mathbb{N}}$ with $k_l = \dim \mathcal{F}_l \geq 2$ for all $l \in \mathbb{N}$ we have*

$$\inf_{s^l \in \mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{R}^l, \mathcal{P}^l) \leq \frac{1}{2},$$

for all $l \in \mathbb{N}$. Thus

$$\mathcal{A}_{\mathrm{random}}(\mathfrak{I}) = 0,$$

and consequently

$$\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) = 0.$$

Proof. We have to show that for the codes $(\mathcal{P}^l, \mathcal{R}^l)$ with the properties as stated in the lemma the inequality

$$\inf_{s^l \in \mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{R}^l, \mathcal{P}^l) \leq \frac{1}{2} \tag{144}$$

holds for all $l \in \mathbb{N}$.
Let $\psi_l \in \mathcal{S}(\mathcal{F}_l \otimes \mathcal{F}_l)$ be a purification of $\pi_{\mathcal{F}_l}$ which is, clearly, maximally entangled. Inequality (144) can then be equivalently reformulated as

$$\inf_{s^l \in \mathbf{S}^l} \int \langle \psi_l, (id_{\mathcal{F}_l} \otimes (\mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l))(|\psi_l\rangle\langle\psi_l|)\psi_l \rangle d\mu_l(\mathcal{R}^l, \mathcal{P}^l) \leq \frac{1}{2}. \tag{145}$$

[2] Mere positivity is sufficient here because $\mathfrak{F}_{\mathbb{C}}(\mathbf{S})$ is commutative, cf. [30].

We fix $\sigma \in \mathcal{S}(\mathcal{H})$ and define CPTP maps $E_1, E_2 : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ by

$$E_1(a) := \mathcal{M}_{T,\mathbf{S}}(\sigma \otimes a) = \sum_{s\in\mathbf{S}} \mathrm{tr}(E_s\sigma)\mathcal{N}_s(a) \tag{146}$$

and

$$E_2(a) := \mathcal{M}_{T,\mathbf{S}}(a \otimes \sigma) = \sum_{s\in\mathbf{S}} \mathrm{tr}(E_s a)\mathcal{N}_s(\sigma). \tag{147}$$

Then

$$\begin{aligned}\int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ E_1^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{R}^l, \mathcal{P}^l) &= \sum_{s^l\in\mathbf{S}^l} \mathrm{tr}(E_{s^l}\sigma^{\otimes l}) \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{R}^l, \mathcal{P}^l) \\ &\geq \inf_{s^l\in\mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{R}^l, \mathcal{P}^l), \end{aligned} \tag{148}$$

where $E_{s^l} := E_{s_1} \otimes \ldots \otimes E_{s_l}$. Therefore, we are done if we can show that

$$\int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ E_1^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{R}^l, \mathcal{P}^l) \leq \frac{1}{2} \tag{149}$$

for all $l \in \mathbb{N}$.
On the other hand, choosing bases $\{e_{i,j}\}_{i,j=1}^{k_l}$ and $\{f_{k,m}\}_{k,m=1}^{d^l}$ of $\mathcal{B}(\mathcal{F}_l)$ and $\mathcal{B}(\mathcal{H})^{\otimes l}$ respectively, we can write

$$id_{\mathcal{F}_l} \otimes \mathcal{P}^l(|\psi_l\rangle\langle\psi_l|) =: \rho_l = \sum_{i,j,k,m} \rho_{i,j,k,m} e_{i,j} \otimes f_{k,m},$$

and obtain

$$\begin{aligned} id_{\mathcal{F}_l} \otimes (\mathcal{R}^l \circ E_1^{\otimes l})(\rho_l) &= \sum_{i,j,k,m} \rho_{i,j,k,m} e_{i,j} \otimes \mathcal{R}^l(\mathcal{M}_{T,\mathbf{S}}^{\otimes l}(\sigma^{\otimes l} \otimes f_{k,m})) \\ &= \sum_{i,j,k,m} \rho_{i,j,k,m} e_{i,j} \otimes \mathcal{R}^l(\mathcal{M}_{T,\mathbf{S}}^{\otimes l}(f_{k,m} \otimes \sigma^{\otimes l})) \\ &= \sum_{i,j,k,m} \rho_{i,j,k,m} e_{i,j} \otimes \mathcal{R}^l(E_2^{\otimes l}(f_{k,m})) \\ &= id_{\mathcal{F}_l} \otimes (\mathcal{R}^l \circ E_2^{\otimes l})(\rho_l), \end{aligned} \tag{150}$$

where the second equality follows from the assumed qc-symmetrizability. Thus, we end up with

$$F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ E_1^{\otimes l} \circ \mathcal{P}^l) = F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ E_2^{\otimes l} \circ \mathcal{P}^l), \tag{151}$$

for any encoding operation $\mathcal{P}^l$ and any recovery operation $\mathcal{R}^l$. Consequently, by (151) and (148) we have to show that for all $l \in \mathbb{N}$

$$\int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ E_2^{\otimes l} \circ \mathcal{P}^l) d\mu_l(\mathcal{R}^l, \mathcal{P}^l) \leq \frac{1}{2} \tag{152}$$

holds. But the channel

$$E_2(a) = \sum_{s\in S} \mathrm{tr}(E_s a)\mathcal{N}_s(\sigma) \qquad (a \in \mathcal{B}(\mathcal{H}))$$

is entanglement breaking implying that the state

$$(id_{\mathcal{F}_l} \otimes \mathcal{R}^l \circ E_2^{\otimes l} \circ \mathcal{P}^l)(|\psi_l\rangle\langle\psi_l|)$$

is separable. A standard result from entanglement theory implies that

$$\langle\psi_l, (id_{\mathcal{F}_l}\otimes\mathcal{R}^l\circ E_2^{\otimes l}\circ\mathcal{P}^l)(|\psi_l\rangle\langle\psi_l|)\psi_l\rangle \le \frac{1}{k_l} \tag{153}$$

holds for any $\mathcal{R}^l$ and $\mathcal{P}^l$ since ψ_l is maximally entangled with Schmidt rank k_l. Now, our assumption that for each $l\in\mathbb{N}$ the relation $k_l\ge 2$ holds implies along with (153) that for all $l\in\mathbb{N}$

$$\int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l\circ E_2^{\otimes l}\circ\mathcal{P}^l)d\mu_l(\mathcal{R}^l,\mathcal{P}^l) = \int\langle\psi_l, (id_{\mathcal{F}_l}\otimes\mathcal{R}^l\circ E_2^{\otimes l}\circ\mathcal{P}^l)(|\psi_l\rangle\langle\psi_l|)\psi_l\rangle d\mu_l(\mathcal{R}^l,\mathcal{P}^l) \le \frac{1}{2},$$

and by (151) and (148) we obtain

$$\inf_{s^l\in\mathbf{S}^l}\int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l\circ\mathcal{N}_{s^l}\circ\mathcal{P}^l)d\mu_l(\mathcal{R}^l,\mathcal{P}^l) \le \frac{1}{2}$$

which concludes the proof.

□

Our Definition 13 addresses the notion of qc-symmetrizability for block length $l=1$. Thus the question arises whether a less restrictive requirement, as stated in the following definition, gives a better sufficient condition for an arbitrarily varying quantum channel to have capacity 0.

Definition 14. *An arbitrarily varying quantum channel, generated by a finite set $\mathfrak{I}=\{\mathcal{N}_s\}_{s\in\mathbf{S}}$, is called l-qc-symmetrizable, $l\in\mathbb{N}$, if there is $T\in QC(\mathcal{H}^{\otimes l},\mathbf{S}^l)$ such that for all $a,b\in\mathcal{B}(\mathcal{H})^{\otimes l}$*

$$\mathcal{M}^l_{T,\mathbf{S}}(a\otimes b) = \mathcal{M}^l_{T,\mathbf{S}}(b\otimes a) \tag{154}$$

holds, where $\mathcal{M}^l_{T,\mathbf{S}}:\mathcal{B}(\mathcal{H})^{\otimes l}\otimes\mathcal{B}(\mathcal{H})^{\otimes l}\to\mathcal{B}(\mathcal{K})^{\otimes l}$ is the CPTP map defined by

$$\mathcal{M}^l_{T,\mathbf{S}}(a\otimes b) := \sum_{s^l\in\mathbf{S}^l} tr(E_{s^l}a)\mathcal{N}_{s^l}(b), \tag{155}$$

and $\{E_{s^l}\}_{s^l\in\mathbf{S}^l}$ is the unique POVM corresponding to $T\in QC(\mathcal{H}^{\otimes l},\mathbf{S}^l)$.

Obviously, qc-symmetrizability implies l-qc-symmetrizability for all $l\in\mathbb{N}$. The next lemma states that the reverse implication is true too.

Lemma 11. *For any finitely generated AVQC given by $\mathfrak{I}=\{\mathcal{N}_s\}_{s\in\mathbf{S}}$ l-qc-symmetrizability implies qc-symmetrizability for any $l\in\mathbb{N}$.*

Proof. For a given finite set of quantum channels $\mathfrak{I}=\{\mathcal{N}_s\}_{s\in\mathbf{S}}$ and $l\in\mathbb{N}$ let $T\in\text{QC}(\mathcal{H}^{\otimes l},\mathbf{S}^l)$ be such that for all $a,b\in\mathcal{B}(\mathcal{H})^{\otimes l}$

$$\mathcal{M}^l_{T,\mathbf{S}}(a\otimes b) = \mathcal{M}^l_{T,\mathbf{S}}(b\otimes a), \tag{156}$$

where $\mathcal{M}^l_{T,\mathbf{S}}$ is defined in (155).
Let $b\in\mathcal{B}(\mathcal{H})$ and for each $s\in\mathbf{S}$ define a linear functional

$$\phi_s(b) := \text{tr}\left(\left(b\otimes\left(\frac{1}{\dim\mathcal{H}}\mathbf{1}_{\mathcal{H}}\right)^{\otimes l-1}\right)\sum_{s_2^l\in\mathbf{S}^{l-1}} E_{ss_2^l}\right),$$

where $ss_2^l := (s,s_2,\dots,s_l)\in\mathbf{S}^l$. Clearly, ϕ_s is positive. Consequently, Riesz' representation theorem shows that there is a unique positive $\tilde{E}_s\in\mathcal{B}(\mathcal{H})$ with

$$\phi_s(b) = \text{tr}(\tilde{E}_s b) \qquad (b\in\mathcal{B}(\mathcal{H})). \tag{157}$$

Obviously, $\{\tilde{E}_s\}_{s\in\mathbf{S}}$ is a POVM and let $\tilde{T}\in \mathrm{QC}(\mathcal{H},\mathbf{S})$ denote the associated qc-channel. Some simple algebra shows that for each $a,b\in\mathcal{B}(\mathcal{H})$

$$\mathcal{M}_{\tilde{T},\mathbf{S}}(a\otimes b)=\mathrm{tr}_{\mathcal{H}^{\otimes l-1}}\mathcal{M}^l_{T,\mathbf{S}}\left(\left(a\otimes\left(\frac{1}{\dim\mathcal{H}}\mathbf{1}_{\mathcal{H}}\right)^{\otimes l-1}\right)\otimes\left(b\otimes\left(\frac{1}{\dim\mathcal{H}}\mathbf{1}_{\mathcal{H}}\right)^{\otimes l-1}\right)\right) \tag{158}$$

where $\mathrm{tr}_{\mathcal{H}^{\otimes l-1}}$ denotes the partial trace over the last $l-1$ tensor factors. The relation (158) immediately implies that for all $a,b\in\mathcal{B}(\mathcal{H})$

$$\mathcal{M}_{\tilde{T},\mathbf{S}}(a\otimes b)=\mathcal{M}_{\tilde{T},\mathbf{S}}(b\otimes a), \tag{159}$$

and

$$\begin{aligned}\mathcal{M}_{\tilde{T},\mathbf{S}}(a\otimes b) &= \sum_{s^l\in\mathbf{S}^l}\mathrm{tr}\left(\left(b\otimes\left(\frac{1}{\dim\mathcal{H}}\mathbf{1}_{\mathcal{H}}\right)^{\otimes l-1}\right)E_{s^l}\right)\mathcal{N}_{s_1}(a)\\ &= \sum_{s_1\in\mathbf{S}}\mathrm{tr}\left(\left(b\otimes\left(\frac{1}{\dim\mathcal{H}}\mathbf{1}_{\mathcal{H}}\right)^{\otimes l-1}\right)\sum_{s_2^l\in\mathbf{S}^{l-1}}E_{s^l}\right)\mathcal{N}_{s_1}(a)\\ &= \sum_{s_1\in\mathbf{S}}\phi_{s_1}(b)\mathcal{N}_{s_1}(a)\\ &= \sum_{s_1\in\mathbf{S}}\mathrm{tr}(b\tilde{E}_{s_1})\mathcal{N}_{s_1}(a). \end{aligned}\tag{160}$$

Equations (159) and (160) show that $\mathfrak{I}$ is qc-symmetrizable.

□

9 Conditions for single-letter-capacities

In this section we give two conditions on the structure of a finite AVQC which guarantee that their quantum capacity is given by a single-letter formula. The first one is empty in the case of a single channel, while the second one generalizes the degradability condition from [15].

Lemma 12. *Let $\mathfrak{I}=\{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be a non-symmetrizable AVQC. We have a single-letter formula for $\mathcal{A}_{\mathrm{det}}(\mathfrak{I})$ in any of the following two cases:*

1. *There is $\mathcal{N}_*\in\mathrm{conv}(\mathfrak{I})$ such that for any $\mathcal{N}\in\mathrm{conv}(\mathfrak{I})$ there is $\mathcal{D}_{\mathcal{N}}\in\mathcal{C}(\mathcal{K},\mathcal{K})$ with the property $\mathcal{N}_*=\mathcal{D}_{\mathcal{N}}\circ\mathcal{N}$ and, additionally, $Q(\mathcal{N}_*)=\max_{\rho\in\mathcal{S}(\mathcal{H})}I_c(\rho,\mathcal{N}_*)$ holds for the entanglement transmission capacity $Q(\mathcal{N}_*)$ of the memoryless channel $\mathcal{N}_*$.*
2. *Each $\mathcal{N}\in\mathrm{conv}(\mathfrak{I})$ is degradable.*

Proof. 1. It is clear that

$$\mathcal{A}_{\mathrm{det}}(\mathfrak{I})\le Q(\mathcal{N}_*)=\max_{\rho\in\mathcal{S}(\mathcal{H})}I_c(\rho,\mathcal{N}_*). \tag{161}$$

Since $\mathfrak{I}$ is non-symmetrizable we have

$$\mathcal{A}_{\mathrm{det}}(\mathfrak{I})=\mathcal{A}_{\mathrm{random}}(\mathfrak{I})=\lim_{l\to\infty}\frac{1}{l}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})}\min_{\mathcal{N}\in\mathrm{conv}(\mathfrak{I})}I_c(\rho,\mathcal{N}^{\otimes l}). \tag{162}$$

On the other hand by application of the data-processing inequality [31] we have, for all $\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})$, $\mathcal{N}\in\mathrm{conv}(\mathfrak{I})$ and $l\in\mathbb{N}$,

$$\begin{aligned} I_c(\rho,\mathcal{N}^{\otimes l}) &\ge I_c(\rho,\mathcal{D}_{\mathcal{N}}^{\otimes l}\circ\mathcal{N}^{\otimes l}) &(163)\\ &= I_c(\rho,\mathcal{N}_*^{\otimes l}). &(164)\end{aligned}$$

It follows that

$$\frac{1}{l}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} I_c(\rho,\mathcal{N}^{\otimes l}) \geq \frac{1}{l}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} I_c(\rho,\mathcal{N}_*^{\otimes l}) \tag{165}$$

and by (162):

$$\begin{aligned}\mathcal{A}_{\mathrm{det}}(\mathfrak{J}) &\geq \lim_{l\to\infty}\frac{1}{l}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} I_c(\rho,\mathcal{N}_*^{\otimes l}) &(166)\\ &= Q(\mathcal{N}_*) &(167)\\ &= \max_{\rho\in\mathcal{S}(\mathcal{H})} I_c(\rho,\mathcal{N}_*). &(168)\end{aligned}$$

Equations (161) and (168) give us the desired result:

$$\mathcal{A}_{\mathrm{det}}(\mathfrak{J}) = \max_{\rho\in\mathcal{S}(\mathcal{H})} I_c(\rho,\mathcal{N}_*). \tag{169}$$

2. It is well known that the following three properties are valid:

P1 If a $\mathcal{N}\in\mathcal{C}(\mathcal{H},\mathcal{K})$ is degradable, then the map $\rho\mapsto I_c(\rho,\mathcal{N})$ is concave [35].

P2 For every fixed $\rho\in\mathcal{S}(\mathcal{H})$, $\mathcal{N}\mapsto I_c(\rho,\mathcal{N})$ is convex (see [27]).

P3 Let $\mathcal{N}\in\mathcal{C}(\mathcal{H},\mathcal{K})$ be degradable. For an arbitrary $l\in\mathbb{N}$, write $\mathcal{H}^{\otimes l}=\mathcal{H}_1\otimes\ldots\otimes\mathcal{H}_l$ with $\mathcal{H}_i:=\mathcal{H}$ for every $i\in\mathbb{N}$. Let $\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})$ with marginal states $\rho_i:=\mathrm{tr}_{\mathcal{H}_1\otimes\ldots\otimes\mathcal{H}_{i-1}\otimes\mathcal{H}_{i+1}\otimes\ldots\otimes\mathcal{H}_l}(\rho)$. Then the inequality $I_c(\rho,\mathcal{N}^{\otimes l})\leq\sum_{i=1}^n I_c(\rho_i,\mathcal{N})$ holds [15].

P4 The coherent information is continuous in both of its entries.

By the minimax-theorem [33, 21], properties P1, P2 and P4 imply that

$$\max_{\rho\in\mathcal{S}(\mathcal{H})}\min_{\mathcal{N}\in\mathrm{conv}(\mathfrak{J})} I_c(\rho,\mathcal{N}) = \min_{\mathcal{N}\in\mathrm{conv}(\mathfrak{J})}\max_{\rho\in\mathcal{S}(\mathcal{H})} I_c(\rho,\mathcal{N}). \tag{170}$$

Suppose now, that each $\mathcal{N}\in\mathrm{conv}(\mathfrak{J})$ is degradable. It then holds, for every $l\in\mathbb{N}$,

$$\begin{aligned}\frac{1}{l}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})}\min_{\mathcal{N}\in\mathrm{conv}(\mathfrak{J})} I_c(\rho,\mathcal{N}^{\otimes l}) &\leq \frac{1}{l}\min_{\mathcal{N}\in\mathrm{conv}(\mathfrak{J})}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})} I_c(\rho,\mathcal{N}^{\otimes l}) &(171)\\ &\leq \min_{\mathcal{N}\in\mathrm{conv}(\mathfrak{J})}\max_{\rho\in\mathcal{S}(\mathcal{H})} I_c(\rho,\mathcal{N}) &(172)\\ &= \max_{\rho\in\mathcal{S}(\mathcal{H})}\min_{\mathcal{N}\in\mathrm{conv}(\mathfrak{J})} I_c(\rho,\mathcal{N}), &(173)\end{aligned}$$

where the second inequality follows from P3 and the equality from P1, P2 via the minimax-theorem. It follows that

$$\mathcal{A}_{\mathrm{det}}(\mathfrak{J})\leq\mathcal{A}_{\mathrm{random}}(\mathfrak{J})\leq\max_{\rho\in\mathcal{S}(\mathcal{H})}\min_{\mathcal{N}\in\mathrm{conv}(\mathfrak{J})} I_c(\rho,\mathcal{N}). \tag{174}$$

By non-symmetrizability of $\mathfrak{J}$, we also have $\mathcal{A}_{\mathrm{det}}(\mathfrak{J})=\mathcal{A}_{\mathrm{random}}(\mathfrak{J})$. The obvious relation $\mathcal{A}_{\mathrm{random}}(\mathfrak{J})\geq\max_{\rho\in\mathcal{S}(\mathcal{H})}\min_{\mathcal{N}\in\mathrm{conv}(\mathfrak{J})} I_c(\rho,\mathcal{N})$ then implies the reverse inequality. □

10 An example and an application to zero-error capacities

10.1 Erasure-AVQC

As an application and illustration of most of the results obtained so far we calculate the quantum capacity of finite AVQC $\mathfrak{I}$ consisting of erasure quantum channels. As expected, we obtain that $\mathcal{A}_{\mathrm{det}}(\mathfrak{I})$ equals the capacity of the worst erasure channel in the set $\mathfrak{I}$.

Lemma 13. *Let $d \in \mathbb{N}$, $d \geq 2$ and denote by $\{e_1, \dots e_d\}$, $\{e_1, \dots, e_{d+1}\}$, the standard basis of $\mathbb{C}^d, \mathbb{C}^{d+1}$. Set $\mathcal{H} = \mathbb{C}^d$, $\mathcal{K} = \mathbb{C}^{d+1}$. Define, for $p \in [0,1]$, the erasure channel $\mathcal{E}_p \in \mathcal{C}(\mathcal{H}, \mathcal{K})$ by*

$$\mathcal{E}_p(x) := (1-p)x + p \cdot \mathrm{tr}(x)|e_{d+1}\rangle\langle e_{d+1}| \qquad \forall x \in \mathcal{B}(\mathcal{H}). \tag{175}$$

Let, for a finite collection $\{p_s\}_{s \in \mathbf{S}} \subset [0,1]$, an AVQC be given by $\mathfrak{I} = \{\mathcal{E}_{p_s}\}_{s \in \mathbf{S}}$.
The following are true.

1. *If $p_s \geq 1/2$ for some $s \in \mathbf{S}$, then $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) = \mathcal{A}_{\mathrm{random}}(\mathfrak{I}) = 0$.*
2. *If $p_s < 1/2$ for every $s \in \mathbf{S}$, then $\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) = \mathcal{A}_{\mathrm{random}}(\mathfrak{I}) = \min_{s \in \mathbf{S}}(1-2p_s)\log(d)$.*

Proof. We start with *2.* by showing the validity of the following properties.

A. For $q \in \mathfrak{P}(\mathbf{S})$, we have $\sum_{s \in \mathbf{S}} q(s)\mathcal{E}_{p_s} = \mathcal{E}_{q(p)}$, where $q(p) := \sum_{s \in \mathbf{S}} q(s)p_s$.

B. There is a set $\{\widehat{\mathcal{E}_{p_s}}\}_{s \in \mathbf{S}}$ of complementary maps given by $\widehat{\mathcal{E}_{p_s}} = \mathcal{E}_{1-p_s}$, $s \in \mathbf{S}$.

C. $\mathcal{E}_p$ is degradable for $p \in [0, 1/2)$.

D. $\{\mathcal{E}_{p_s}\}_{s \in \mathbf{S}}$ is non-symmetrizable if $p_s \in [0,1)$ for all $s \in \mathbf{S}$.

A.: For every $x \in \mathcal{B}(\mathbb{C}^d)$,

$$\sum_{s \in \mathbf{S}} q(s)\mathcal{E}_{p_s}(x) = \sum_{s \in \mathbf{S}} q(s)[(1-p_s)x + p_s \cdot \mathrm{tr}(x)|e_{d+1}\rangle\langle e_{d+1}|] \tag{176}$$

$$= (1 - \sum_{s \in \mathbf{S}} q(s)p_s)x + \sum_{s \in \mathbf{S}} q(s)p_s \cdot \mathrm{tr}(x)|e_{d+1}\rangle\langle e_{d+1}|. \tag{177}$$

B.: Consider an environment defined by $\mathcal{K}_{env} := \mathcal{K}$. For every $p \in [0,1]$ we can give a Stinespring isometry $V_p : \mathcal{H} \to \mathcal{K} \otimes \mathcal{K}_{env}$ of $\mathcal{E}_p$ by

$$V_p u := \sqrt{1-p} \cdot u \otimes e_{d+1} + \sqrt{p} \cdot e_{d+1} \otimes u, \qquad u \in \mathcal{B}(\mathcal{H}). \tag{178}$$

The claim becomes clear by tracing out the first or second subsystem, depending on whether one wants to calculate $\hat{\mathcal{E}}_p$ or $\mathcal{E}_p$.
C.: Set $\mu := \frac{1-2p}{1-p}$ and define $E_\mu \in \mathcal{C}(\mathcal{K}, \mathcal{K})$ by

$$E_\mu(x) := (1-\mu) \cdot x + \mu \cdot |e_{d+1}\rangle\langle e_{d+1}|, \qquad x \in \mathcal{B}(\mathcal{K}). \tag{179}$$

Then by $p \in [0, 1/2)$ we have $\mu \in (0,1]$. We show that $\mathcal{E}_{1-p} = E_\mu \circ \mathcal{E}_p$ holds. Let $x \in \mathcal{B}(\mathcal{H})$, then

$$\begin{aligned} E_\mu \circ \mathcal{E}_p(x) &= (1-p)E_\mu(x) + pE_\mu(|e_{d+1}\rangle\langle e_{d+1}|) && (180)\\ &= (1-p)(1-\mu) \cdot x + \mu(1-p) \cdot |e_{d+1}\rangle\langle e_{d+1}| + p|e_{d+1}\rangle\langle e_{d+1}| && (181)\\ &= (1-p-1+2p) \cdot x + (1-2p)|e_{d+1}\rangle\langle e_{d+1}| + p|e_{d+1}\rangle\langle e_{d+1}| && (182)\\ &= p \cdot x + (1-p) \cdot |e_{d+1}\rangle\langle e_{d+1}| && (183)\\ &= \mathcal{E}_{1-p}(x). && (184) \end{aligned}$$

D.: Let $\rho_1 := |e_1\rangle\langle e_1|,\ \rho_2 := |e_2\rangle\langle e_2| \in \mathcal{S}(\mathcal{H})$. We show by contradiction that there are no two probability distributions $r_1, r_2 \in \mathfrak{P}(\mathbf{S})$ such that

$$\sum_{s\in\mathbf{S}} r_1(s)\mathcal{E}_{p_s}(\rho_1) = \sum_{s\in\mathbf{S}} r_2(s)\mathcal{E}_{p_s}(\rho_2). \tag{185}$$

Assume there are $r_1, r_2 \in \mathfrak{P}(\mathbf{S})$ such that (185) is true. This is equivalent to

$$\sum_{s\in\mathbf{S}}(1-p_s)[r_1(s)|e_1\rangle\langle e_1| - r_2(s)|e_2\rangle\langle e_2|] = 0, \qquad \sum_{s\in\mathbf{S}} p_s[r_1(s) - r_2(s)]\cdot|e_{d+1}\rangle\langle e_{d+1}| = 0. \tag{186}$$

By linear independence of $|e_1\rangle\langle e_1|, |e_2\rangle\langle e_2|$ and since $p_s \in [0,1)$ for every $s \in \mathbf{S}$ the first equality implies $r_1(s) = r_2(s) = 0\ \forall s \in \mathbf{S}$, in clear contradiction to the assumption $r_1, r_2 \in \mathfrak{P}(\mathbf{S})$.
Thus, $\{\mathcal{E}_{p_s}\}_{s\in\mathbf{S}}$ with all $p_s \in [0,1)$ is non-symmetrizable.

Using **A** and the fact that $\{p_s\}_{s\in\mathbf{S}} \subset [0,1/2)$ we see that for an arbitrary $q \in \mathfrak{P}(\mathbf{S})$ we have $\sum_{s\in\mathbf{S}} q(s)\mathcal{E}_{p_s} = \mathcal{E}_{q(p)}$ with $q(p) \in [0,1/2)$.
Now **B** implies that for every $q \in \mathfrak{P}(\mathbf{S})$ the channel $\sum_{s\in\mathbf{S}} q(s)\mathcal{E}_{p_s}$ is degradable.
Thus by Lemma 12, *2.*, the regularization in the identity

$$\mathcal{A}_{\text{random}}(\mathfrak{I}) = \lim_{l\to\infty}\frac{1}{l}\max_{\rho\in\mathcal{S}(\mathcal{H}^{\otimes l})}\inf_{\mathcal{N}\in\text{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}^{\otimes l})$$

is not necessary, so

$$\mathcal{A}_{\text{random}}(\mathfrak{I}) = \max_{\rho\in\mathcal{S}(\mathcal{H})}\inf_{\mathcal{N}\in\text{conv}(\mathfrak{I})} I_c(\rho, \mathcal{N}). \tag{187}$$

Further, for a fixed degradable channel, the coherent information is concave in the input state [35] and thus by the minimax theorem for concave-convex functions [21, 33] we can interchange min and max in (187). Now, to any given $\rho \in \mathcal{S}(\mathcal{H})$, we may write $\rho = \sum_{i=1}^d \lambda_i |v_i\rangle\langle v_i|$ for some set $\{v_1,\ldots,v_d\}$ of orthonormal vectors that satisfy, by standard identification of $\mathbb{C}^d$ and $\mathbb{C}^{d+1}$, $v_i \perp e_{d+1}$ $(1 \le i \le d)$ and write a purification of ρ as $|\psi_\rho\rangle\langle\psi_\rho| = \sum_{i,j=1}^d \lambda_i\lambda_j |v_i\rangle\langle v_j| \otimes |v_i\rangle\langle v_j|$. Then for every $\mathcal{E}_p \in \text{conv}(\mathfrak{I})$ we have

$$\begin{aligned}
\max_{\rho\in\mathcal{S}(\mathcal{H})} I_c(\rho, \mathcal{E}_p) &= \max_{\rho\in\mathcal{S}(\mathcal{H})}(S(\mathcal{E}_p(\rho)) - S(Id_{\mathcal{H}} \otimes \mathcal{E}_p(|\psi_\rho\rangle\langle\psi_\rho|)) &(188)\\
&= \max_{\rho\in\mathcal{S}(\mathcal{H})}(S((1-p)\rho + p|e_{d+1}\rangle\langle e_{d+1}|) - S((1-p)|\psi_\rho\rangle\langle\psi_\rho| + p\rho\otimes|e_{d+1}\rangle\langle e_{d+1}|)) &(189)\\
&= \max_{\rho\in\mathcal{S}(\mathcal{H})}((1-p)S(\rho) + pS(|e_{d+1}\rangle\langle e_{d+1}|) + H(p) &(190)\\
&\quad - (1-p)S(|\psi_\rho\rangle\langle\psi_\rho|) - pS(\rho\otimes|e_{d+1}\rangle\langle e_{d+1}|) - H(p)) &(191)\\
&= \max_{\rho\in\mathcal{S}(\mathcal{H})}((1-p)S(\rho) - pS(\rho\otimes|e_{d+1}\rangle\langle e_{d+1}|)) &(192)\\
&= \max_{\rho\in\mathcal{S}(\mathcal{H})}(1-2p)S(\rho) &(193)\\
&= (1-2p)\log(d). &(194)
\end{aligned}$$

This leads to

$$\mathcal{A}_{\text{random}}(\mathfrak{I}) = \min_{s\in\mathbf{S}}(1-2p_s)\log(d), \tag{195}$$

a formula that was first discovered for the case of a single memoryless channel and $d = 2$ by [8].
From **C** it follows that $\mathcal{A}_{\text{det}}(\mathfrak{I}) = \mathcal{A}_{\text{random}}(\mathfrak{I})$.
We can now prove *1.* by noting that for $p_{\max} := \max_{s\in\mathbf{S}} p_s$ (with $p_{\max} \ge 1/2$), the channel $\mathcal{E}_{\max} := \mathcal{E}_{p_{\max}}$ satisfies $\mathcal{A}_{\text{det}}(\{\mathcal{E}_{\max}\}) = \mathcal{A}_{\text{random}}(\{\mathcal{E}_{\max}\}) = 0$ by (194). Thus, for every $l \in \mathbb{N}$, the adversary can always choose $\mathcal{E}_{\max}^{\otimes l}$ to ensure that transmission of entanglement will fail. □

10.2 Qualitative behavior of zero-error capacities

Let us, first, embark on the connection between AVQCs and zero-error capacities. Classical information theory exhibits an interesting connection between the zero-error capacity of certain channels and the deterministic capacity with asymptotically vanishing maximal error probability criterion. This connection is described in [1].

We give (following closely the lines of [1]) the remaining part of this connection in the quantum case: Let $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ be a finite AVQC. Consider

$$\mathcal{N}_{\mathfrak{I}} := \frac{1}{|\mathbf{S}|}\sum_{s\in\mathbf{S}}\mathcal{N}_s. \tag{196}$$

Assume that $Q_0(\mathcal{N}_{\mathfrak{I}}) > 0$ holds. Thus, there exists an $l \in \mathbb{N}$, a maximally mixed state $\pi_{\mathcal{F}_l}$ with $\frac{1}{l}\log\dim\mathcal{F}_l > 0$ (implying $\dim\mathcal{F}_l > 1$) and a pair $(\mathcal{R}^l, \mathcal{P}^l)$ of recovery and encoding map such that

$$\min_{x\in\mathcal{F}_l, ||x||=1}\langle x, \mathcal{R}^l \circ \mathcal{N}_{\mathfrak{I}}^{\otimes l} \circ \mathcal{P}(|x\rangle\langle x|)x\rangle = 1 \tag{197}$$

holds. But this directly implies

$$\min_{x\in\mathcal{F}_l, ||x||=1}\langle x, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}(|x\rangle\langle x|)x\rangle = 1 \qquad \forall s^l \in \mathbf{S}^l, \tag{198}$$

so $(\pi_{\mathcal{F}_l}, \mathcal{R}, \mathcal{P}^l)$ is a zero-error code for the AVQC $\mathfrak{I}$ as well and therefore

$$\mathcal{A}_{\mathrm{det}}(\mathfrak{I}) \geq Q_0(\mathcal{N}_{\mathfrak{I}}). \tag{199}$$

One may now ask when exactly this is a meaningful (nonzero) lower bound. The answer is given by the proof of Lemma 15: On any face of $\mathcal{C}(\mathcal{H}, \mathcal{K})$ the zero-error capacities are constant and the encoding and recovery maps are *universal*. Thus, if $\mathfrak{I}$ is a subset of a face and $\mathfrak{I} \subset \mathrm{ri}(\mathcal{C}(\mathcal{H}, \mathcal{K})^{\complement}$ then there is good hope to get a nonzero lower bound by means of inequality (199). So far for the connection between AVQCs and zero-error capacities.

Motivated by the above observation, a closer study of zero-error capacities reveals some additional facts that are interesting in their own right.

To be more precise, we investigate continuity of zero-error capacities. This property is a highly desirable property both from the practical and the theoretical point of view. It is of particular importance in situations where full knowledge of the communication system cannot be achieved but only a narrow confidence set containing the unknown channel is given. In [26] it has been shown that the ordinary capacities of stationary memoryless quantum channels are continuous in the finite-dimensional setting and it was demonstrated by examples that these functions become discontinuous in infinite dimensional situations.

In this subsection we show that quantum, entanglement-assisted, and classical zero-error capacities of quantum channels are discontinuous at every positivity point. Our approach is based on two simple observations. The first one is that the zero-error capacities mentioned above of each quantum channel belonging to the relative interior of the set of quantum channels are equal to 0. The second one is the well known fact that the relative interior of any convex set is open and dense in that set, i.e. generic. Hence any channel can be approximated by a sequence belonging to the relative interior implying the discontinuity result.

Similar arguments can be applied to the recently introduced Lovász $\tilde{\theta}$ function and zero-error distillable entanglement as well, leading to analogous conclusions as shall be shown in the last part of this subsection.

We now show that all the zero-error capacities defined in subsection 3.3 are generically equal to 0 and are discontinuous at any positivity point. Then we demonstrate that the zero-error capacities of quantum channels can be thought of as step functions subordinate to the partition built from the relative interiors of the faces of $\mathcal{C}(\mathcal{H}, \mathcal{K})$.

Discontinuity of zero-error capacities

Theorem 16. *Let $\mathcal{N} \in \mathrm{ri}\, \mathcal{C}(\mathcal{H}, \mathcal{K})$. Then $k(l, \mathcal{N}) = M(l, \mathcal{N}) = M_{\mathrm{EA}}(l, \mathcal{N}) = 1$. Consequently, $Q_0(\mathcal{N}) = C_0(\mathcal{N}) = C_{0\mathrm{EA}}(\mathcal{N}) = 0$.*

In the proof of Theorem 16 we shall make use of the following elementary fact:

Lemma 14. *Let F be a non-empty convex set and $\mathcal{N}_0, \mathcal{N}\, \mathrm{ri}\, F$ with $\mathcal{N}_0 \neq \mathcal{N}$. Then there exists $\mathcal{N}_1 \in F$ and $\lambda_0, \lambda_1 \in (0,1)$, $\lambda_0 + \lambda_1 = 1$ with $\mathcal{N} = \lambda_0 \mathcal{N}_0 + \lambda_1 \mathcal{N}_1$.*

Proof of Lemma 14. Let

$$L := \{\mu \in [0, +\infty) : (1-\mu)\mathcal{N}_0 + \mu \mathcal{N} \in F\}. \tag{200}$$

Since F is convex we can conclude that L is convex too. Clearly, $[0,1] \subset L$ since $\mathcal{N} \in F$. Moreover, since $\mathcal{N} \in \mathrm{ri}\, F$ there is $\mu' > 1$ such that

$$\mathcal{N}_1 := (1-\mu')\mathcal{N}_0 + \mu' \mathcal{N} \in F. \tag{201}$$

We define now

$$\lambda_1 := \frac{1}{\mu'} \in (0,1), \quad \lambda_0 := 1 - \lambda_1, \tag{202}$$

and obtain using $\mathcal{N}_1$ given in (201) the desired convex decomposition

$$\mathcal{N} = \lambda_0 \mathcal{N}_0 + \lambda_1 \mathcal{N}_1.$$

□

Proof of Theorem 16. Let $\mathcal{N} \in \mathrm{ri}\, \mathcal{C}(\mathcal{H}, \mathcal{K})$. Observing that the fully depolarizing channel $\mathcal{N}_0(a) = \frac{\mathrm{tr}(a)}{d_{\mathcal{K}}} \mathbf{1}_{\mathcal{K}}$, $a \in \mathcal{B}(\mathcal{H})$, belongs to $\mathrm{ri}\, \mathcal{C}(\mathcal{H}, \mathcal{K})$ we obtain from Lemma 14 a convex decomposition of $\mathcal{N}$ as

$$\mathcal{N} = \lambda_0 \mathcal{N}_0 + \lambda_1 \mathcal{N}_1, \tag{203}$$

where $\lambda_0, \lambda_1 \in (0,1)$, $\lambda_0 + \lambda_1 = 1$.
Clearly, this decomposition implies that

$$\mathcal{N}^{\otimes l} = \sum_{s^l \in \{0,1\}^l} \lambda_{s^l} \mathcal{N}_{s^l}, \tag{204}$$

with $\lambda_{s^l} := \lambda_{s_1} \cdot \ldots \cdot \lambda_{s_l} > 0$ and $\mathcal{N}_{s^l} := \mathcal{N}_{s_1} \otimes \ldots \otimes \mathcal{N}_{s_l}$ for all $s^l \in \{0,1\}^l$. Then for any zero-error (l, M) ea-code $(\sigma_{\mathcal{F}\mathcal{F}'}, \{\mathcal{P}_m, D_m\}_{m=1}^M)$ for $\mathcal{N}$ we get for each $m \in [M]$

$$\begin{aligned} 1 &= \mathrm{tr}((\mathcal{N}^{\otimes l} \circ \mathcal{P}_m \otimes \mathrm{id}_{\mathcal{F}'})(\sigma_{\mathcal{F}\mathcal{F}'}) D_m) \\ &= \sum_{s^l \in \{0,1\}^l} \lambda_{s^l} \mathrm{tr}((\mathcal{N}_{s^l} \circ \mathcal{P}_m \otimes \mathrm{id}_{\mathcal{F}'})(\sigma_{\mathcal{F}\mathcal{F}'}) D_m) \end{aligned} \tag{205}$$

and, consequently, since $\lambda_{s^l} > 0$ for all $s^l \in \{0,1\}^l$

$$\mathrm{tr}((\mathcal{N}_{s^l} \circ \mathcal{P}_m \otimes \mathrm{id}_{\mathcal{F}'})(\sigma_{\mathcal{F}\mathcal{F}'}) D_m) = 1 \qquad \forall s^l \in \{0,1\}^l, \tag{206}$$

for all $m \in [M]$. Choosing $\bar{s}^l = (0, \ldots, 0)$ we obtain from Eqn. (206) that $(\sigma_{\mathcal{F}\mathcal{F}'}, \{\mathcal{P}_m, D_m\}_{m=1}^M)$ is a zero-error ea-code for $\mathcal{N}_0$. Since $M_{\mathrm{EA}}(l, \mathcal{N}_0) = 1$ for all $l \in \mathbb{N}$ we can conclude that $M_{\mathrm{EA}}(l, \mathcal{N}) \leq 1$ and thus $M_{\mathrm{EA}}(l, \mathcal{N}) = 1$ holds. Consequently $C_{0\mathrm{EA}}(\mathcal{N}) = 0$. The other assertions follow from the observation that $1 \leq k(l, \mathcal{N}) \leq M(l, \mathcal{N}) \leq M_{\mathrm{EA}}(l, \mathcal{N})$. □

Corollary 17. *The function $Q_0 : \mathcal{C}(\mathcal{H},\mathcal{K}) \to \mathbb{R}_+$ that assigns the zero-error quantum capacity to each quantum channel is discontinuous at any $\mathcal{N} \in \mathcal{C}(\mathcal{H},\mathcal{K})$ with $Q_0(\mathcal{N}) > 0$. The same conclusion holds true for C_0 and $C_{0\mathrm{EA}}$.*

Proof. If $Q_0(\mathcal{N}) > 0$ holds then necessarily $\mathcal{N} \in \mathrm{rebd}\,\mathcal{C}(\mathcal{H},\mathcal{K})$ by Theorem 16. On the other hand $\mathrm{ri}\,\mathcal{C}(\mathcal{H},\mathcal{K})$ is dense in $\mathcal{C}(\mathcal{H},\mathcal{K})$ (cf. Theorem 2.3.8 in [34]). So there is a sequence of channels $(\mathcal{N}_i)_{i\in\mathbb{N}} \subset \mathrm{ri}\,\mathcal{C}(\mathcal{H},\mathcal{K})$ with $\lim_{i\to\infty} ||\mathcal{N}_i - \mathcal{N}||_\diamondsuit = 0$ and by Theorem 16 we have $Q_0(\mathcal{N}_i) = 0$ for all $i \in \mathbb{N}$. The arguments for C_0 and $C_{0\mathrm{EA}}$ follow the same line of reasoning. □

Relation to the facial structure of the set of quantum channels Here we shall show that the considered zero-error capacities are basically step functions, the underlying partition consisting of the relative interiors of the faces of $\mathcal{C}(\mathcal{H},\mathcal{K})$.

Lemma 15. *Let $F \subset \mathcal{C}(\mathcal{H},\mathcal{K})$ be convex and let $\tilde{\mathcal{N}} \in \mathrm{ri}\,F$. Then for any $\mathcal{N} \in \mathrm{ri}\,F$ $Q_0(\mathcal{N}) = Q_0(\tilde{\mathcal{N}})$, $C_{0\mathrm{EA}}(\mathcal{N}) = C_{0\mathrm{EA}}(\tilde{\mathcal{N}})$, and $C_0(\mathcal{N}) = C_0(\tilde{\mathcal{N}})$ hold.*

Proof. We assume w.l.o.g. that $\mathcal{N} \neq \tilde{\mathcal{N}}$ to avoid trivialities. Then setting $\mathcal{N}_0 := \mathcal{N}$ we can find $\mathcal{N}_1 \in F$ and $\lambda_0, \lambda_1 \in (0,1), \lambda_0 + \lambda_1 = 1$, with

$$\tilde{\mathcal{N}} = \lambda_0\mathcal{N}_0 + \lambda_1\mathcal{N}_1 \tag{207}$$

just by applying Lemma 14 to $\mathcal{N}_0, \tilde{\mathcal{N}} \in \mathrm{ri}\,F$.
Let $(\mathcal{F}_l, \mathcal{P}, \mathcal{R})$ be an (l, k_l) zero-error quantum code for $\tilde{\mathcal{N}}$. Then using the representation (207) we obtain for any $x \in \mathcal{F}_l, ||x|| = 1$

$$\begin{aligned} 1 &= \langle x, \mathcal{R} \circ \tilde{\mathcal{N}}^{\otimes l} \circ \mathcal{P}(|x\rangle\langle x|)x\rangle \\ &= \sum_{s^l \in \{0,1\}^l} \lambda_{s^l} \langle x, \mathcal{R} \circ \mathcal{N}_{s^l} \circ \mathcal{P}(|x\rangle\langle x|)x\rangle \end{aligned} \tag{208}$$

and consequently, since $\lambda_{s^l} > 0$ for all $s^l \in \{0,1\}^l$, we are led to

$$\langle x, \mathcal{R} \circ \mathcal{N}_{s^l} \circ \mathcal{P}(|x\rangle\langle x|)x\rangle = 1 \tag{209}$$

for all $s^l \in \{0,1\}^l$ and all $x \in \mathcal{F}_l, ||x|| = 1$. Choosing the sequence $s^l = (0,\ldots,0)$ and recalling that $\mathcal{N}_0 = \mathcal{N}$ we arrive at

$$Q_0(\mathcal{N}) \geq Q_0(\tilde{\mathcal{N}}). \tag{210}$$

The reverse inequality is derived by interchanging the roles of $\mathcal{N}$ and $\tilde{\mathcal{N}}$. The remaining assertions are shown in the same vein. □

We shall now pass to the set of faces $\mathfrak{F} := \{F : \text{face of } \mathcal{C}(\mathcal{H},\mathcal{K})\}$ of $\mathcal{C}(\mathcal{H},\mathcal{K})$.

Theorem 18. *To each $\mathcal{N} \in \mathcal{C}(\mathcal{H},\mathcal{K})$ there is a unique $F \in \mathfrak{F}$ with $\mathcal{N} \in \mathrm{ri}\,F$. Moreover, each of the capacity functions Q_0, $C_{0\mathrm{EA}}$, and C_0 is constant on $\mathrm{ri}\,F$.*

Proof. According to Theorem 2.6.10 in [34] the family of sets $\{\mathrm{ri}\,F : F \in \mathfrak{F}\}$ forms a partition of $\mathcal{C}(\mathcal{H},\mathcal{K})$. This shows the first assertion of the theorem. The second follows from Lemma 15. □

Notice that the results obtained so far show that the optimal (i.e. capacity achieving) code for any channel $\mathcal{N}$ in the relative interior of any face F of $\mathcal{C}(\mathcal{H},\mathcal{K})$ is also optimal for any other channel in ri F.

10.3 Discontinuity of quantum Lovász $\tilde{\theta}$ function & zero-error distillable entanglement

In this final section we show that our methods are not only bound to the zero-error capacities of quantum channels. They apply to Lovász $\tilde{\theta}$ function from [16] and also to zero-error distillable entanglement.

Discontinuity of quantum Lovász $\tilde{\theta}$ function For a given channel $\mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$ with a corresponding set of Kraus operators $\{E_j\}_{j\in[K]}$ we define the non-commutative confusability graph following [16] by

$$\begin{aligned} S(\mathcal{N}) &:= \operatorname{span}\{E_j^* E_i : i, j \in [K]\} \\ &= \hat{\mathcal{N}}_*(\mathcal{B}(\mathcal{E})), \end{aligned} \tag{211}$$

where $\hat{\mathcal{N}}_*$ is the adjoint of the complementary channel $\hat{\mathcal{N}} \in \mathcal{C}(\mathcal{H}, \mathcal{E})$ defined via the Stinespring isometry $V : \mathcal{H} \to \mathcal{K} \otimes \mathcal{E}$

$$Vx := \sum_{j=1}^{K} E_j x \otimes f_j$$

with an ONB $\{f_1, \ldots, f_K\}$ in $\mathcal{E}$. The representation of $S(\mathcal{N})$ in (211) is from [16].
In the following we shall need the next simple lemma.

Lemma 16. *Let $\mathcal{N} \in \operatorname{ri}\mathcal{C}(\mathcal{H}, \mathcal{K})$. Then $S(\mathcal{N}) = \mathcal{B}(\mathcal{H})$.*

Proof. Again we can represent $\mathcal{N}$ as

$$\mathcal{N} = \lambda_0 \mathcal{N}_0 + \lambda_1 \mathcal{N}_1$$

with $\lambda_0, \lambda_1 \in (0, 1)$, $\lambda_0 + \lambda_1 = 1$, $\mathcal{N}_0$ being the fully depolarizing channel, and $\mathcal{N}_1 \in \mathcal{C}(\mathcal{H}, \mathcal{K})$. The proof is concluded by the following simple observation: Given any two channels $\mathcal{N}_0, \mathcal{N}_1$ and $\lambda_0, \lambda_1 \in (0, 1)$ with $\lambda_0 + \lambda_1 = 1$. Then for the channel $\mathcal{N} := \lambda_0 \mathcal{N}_0 + \lambda_1 \mathcal{N}_1$ it holds that

$$S(\mathcal{N}) \supseteq S(\mathcal{N}_0), S(\mathcal{N}_1).$$

Since in our case $S(\mathcal{N}_0) = \mathcal{B}(\mathcal{H})$ we are done. □

Duan, Severini and Winter [16] have given the following characterization of the quantum Lovász $\tilde{\theta}$ function as a solution to the dual of a semidefinite programme.

Theorem 19 (Theorem 9 in [16])**.** *For any $\mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$ we have*

$$\tilde{\theta}(S(\mathcal{N})) = \min\left\{ ||\operatorname{tr}_{\mathcal{H}} Y|| : Y \in S(\mathcal{N}) \otimes \mathcal{B}(\mathcal{H}'), Y \geq |\Phi\rangle\langle\Phi| \right\}, \tag{212}$$

where $\mathcal{H}'$ is just a copy of $\mathcal{H}$ and $\Phi = \sum_{i=1}^{\dim \mathcal{H}} e_i \otimes e_i'$ with ONBs $\{e_1, \ldots, e_{\dim \mathcal{H}}\}$ and $\{e_1', \ldots, e_{\dim \mathcal{H}}'\}$ of $\mathcal{H}$ and $\mathcal{H}'$.

With this theorem at our disposal we can deduce the following discontinuity result for $\tilde{\theta}$:

Theorem 20. *The function $\tilde{\theta} : \mathcal{C}(\mathcal{H}, \mathcal{H}) \to \mathbb{R}_+$ assigning the number $\tilde{\theta}(S(\mathcal{N}))$ to each quantum channel $\mathcal{N}$ is discontinuous at any $\mathcal{N}$ with $C_{0\mathrm{EA}}(\mathcal{N}) > 0$.*

Proof. Note that for $\mathcal{N} \in \operatorname{ri}\mathcal{C}(\mathcal{H}, \mathcal{K})$ $S(\mathcal{N}) = \mathcal{B}(\mathcal{H})$ by Lemma 16. Hence $|\Phi\rangle\langle\Phi| \in S(\mathcal{N}) \otimes \mathcal{B}(\mathcal{H}') = \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H})$ and $||\operatorname{tr}_{\mathcal{H}}|\Phi\rangle\langle\Phi||| = 1 = \tilde{\theta}(S(\mathcal{N}))$. On the other hand, Lemma 7 and Corollary 10 in [16] show that for any $\mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$

$$C_{0\mathrm{EA}}(\mathcal{N}) \leq \log \tilde{\theta}(S(\mathcal{N})),$$

implying $\tilde{\theta}(S(\mathcal{N})) > 1$ if $C_{0\mathrm{EA}}(\mathcal{N}) > 0$. Since $\operatorname{ri}\mathcal{C}(\mathcal{H}, \mathcal{K})$ is dense in $\mathcal{C}(\mathcal{H}, \mathcal{K})$ and since $\tilde{\theta}(S(\mathcal{N})) = 1$ for each $\mathcal{N} \in \operatorname{ri}\mathcal{C}(\mathcal{H}, \mathcal{K})$ we are done. □

Notice that the arguments given for the Lovász $\tilde{\theta}$ function apply to any other upper bound to the entanglement-assisted zero-error capacity vanishing in the relative interior of $\mathcal{C}(\mathcal{H}, \mathcal{K})$.

Zero-error distillation of entanglement The simple methods employed so far can also be applied to the problem of zero-error distillation of entanglement as we shall briefly indicate below. Assuming that $\rho \in \operatorname{ri} \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ we can find $\lambda_0, \lambda_1 \in (0,1)$, $\lambda_0 + \lambda_1 = 1$ such that

$$\rho = \lambda_0 \rho_0 + \lambda_1 \rho_1, \tag{213}$$

with $\rho_0 = \frac{1}{d_A}\mathbf{1}_{\mathcal{H}_A} \otimes \frac{1}{d_B}\mathbf{1}_{\mathcal{H}_B} \in \operatorname{ri} \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, $d_A = \dim \mathcal{H}_A, d_B = \dim \mathcal{H}_B$, and $\rho_1 \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then

$$\rho^{\otimes l} = \sum_{s^l \in \{0,1\}^l} \lambda_{s^l} \rho_{s^l}, \tag{214}$$

and for any (l, k_l) zero-error EDP $(\mathcal{D}, \varphi_{k_l})$ for ρ we obtain

$$1 = \sum_{s^l \in \{0,1\}^l} \lambda_{s^l} \langle \varphi_{k_l}, \mathcal{D}(\rho_{s^l}) \varphi_{k_l} \rangle, \tag{215}$$

leading to

$$1 = \langle \varphi_{k_l}, \mathcal{D}(\rho_{s^l}) \varphi_{k_l} \rangle \tag{216}$$

for all $s^l \in \{0,1\}^l$. Choosing $s^l = (0, \ldots, 0)$ and noting that due to the fact that $\mathcal{D}$ is a LOCC operation the state $\mathcal{D}(\rho_0^{\otimes l})$ is separable, we obtain from [19]

$$1 = \langle \varphi_{k_l}, \mathcal{D}(\rho_0^{\otimes l}) \varphi_{k_l} \rangle \leq \frac{1}{k_l}. \tag{217}$$

Thus $k_l = 1$ and $d(l, \rho) = 1$ for all $l \in \mathbb{N}$. We collect these observations in the following corollary.

Corollary 21. *Let $\rho \in \operatorname{ri} \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then $d(l, \rho) = 1$ for all $l \in \mathbb{N}$ and $D_0(\rho) = 0$. Moreover, the function D_0 is discontinuous at any $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with $D_0(\rho) > 0$.*

11 Conclusion

We have been able to derive a multi-letter analog of Ahlswede's dichotomy for quantum capacities of arbitrarily varying quantum channels: Either the classical, deterministic capacity of such a channel with average error criterion is zero , or else its deterministic and common-randomness-assisted entanglement transmission capacities are equal. Moreover, we have shown that the entanglement and strong subspace transmission capacities for this channel model are equal. It should be noted, however, that our proof of this does not rely on a strategy of "hiding" randomness in the encoding operation. In fact, by using a probabilistic variant of Dvoretzky's theorem we achieve this equality of capacities just by restricting to an appropriate code subspace of comparable dimension on the exponential scale. Here we have left open the question whether the quantum capacity of arbitrarily varying quantum channels can be achieved with isometric encoding operations.

Simple conditions that guarantee single-letter capacity formulas have been provided. They are generalizations of those for memoryless and stationary quantum channels.

The major unresolved problem of this paper is the question whether there are AVQCs $\mathfrak{I}$ for which $C_{\text{det}}(\mathfrak{I}) = 0$ and $\mathcal{A}_{\text{random}}(\mathfrak{I}) > 0$ can occur. Or to put the question into different words: Does common randomness really help to transmit entanglement through arbitrarily varying quantum channels?

Acknowledgment. We thank Andreas Winter for encouragement and his repeated insistence on launching this research. Several very stimulating discussions with Toby Cubitt and Debbie Leung during the fall program 2010 at the Institut Mittag-Leffler have triggered our interest in zero-error capacities which resulted in the last part of the present paper. We thank both of them for these conversations.

Support by the Institut Mittag-Leffler (Djursholm, Sweden) is gratefully acknowledged

This work is supported by the DFG via grants Bj 57/1-2 (IB), BO 1734/20-1 (IB and HB) and by the BMBF via grants 01BQ1050 (IB, HB, and JN) and 01BQ1052 (RA).

References

[1] R. Ahlswede, “A Note on the Existence of the Weak Capacity for Channels with Arbitrarily Varying Channel Probability Functions and Its Relation to Shannon’s Zero Error Capacity” *The Annals of Mathematical Statistics*, Vol. 41, No. 3. (1970)

[2] R. Ahlswede, “Elimination of Correlation in Random Codes for Arbitrarily Varying Channels”, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 44, 159-175 (1978)

[3] R. Ahlswede, “Coloring Hypergraphs: A New Approach to Multi-user Source Coding-II”, *Journal of Combinatorics, Information & System Sciences* Vol. 5, No. 3, 220-268 (1980)

[4] R. Ahlswede, “Arbitrarily Varying Channels with States Sequence Known to the Sender”, *IEEE Trans. Inf. Th.* Vol. 32, 621-629, (1986)

[5] R. Ahlswede, V. Blinovsky, “Classical Capacity of Classical-Quantum Arbitrarily Varying Channels”, *IEEE Trans. Inf. Th.* Vol. 53, No. 2, 526-533 (2007)

[6] R. Ahlswede, J. Wolfowitz, “The Capacity of a Channel with Arbitrarily Varying Channel Probability Functions and Binary Output Alphabet” Z. Wahrscheinlichkeitstheorie verw. Geb. 15, 186-194 (1970)

[7] H. Barnum, E. Knill and M.A. Nielsen, “On Quantum Fidelities and Channel Capacities”, *IEEE Trans. Inf. Theory*, VOL. 46, NO. 4, (2000)

[8] C.H. Bennett, D.P. DiVincenzo, and J.A. Smolin, “Capacities of Quantum Erasure Channels”, Phys. Rev. Lett. 78, 32173220 (1997)

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

[10] I. Bjelaković, H. Boche, J. Nötzel, “Entanglement transmission and generation under channel uncertainty: Universal quantum channel coding”, *Commun. Math. Phys.* 292, 55-97 (2009) - Available at: http://arxiv.org/abs/0811.4588

[11] D. Blackwell, L. Breiman, A.J. Thomasian, “The capacities of certain channel classes under random coding”, *Ann. Math. Stat.* 31, 558-567 (1960)

[12] M.-D. Choi, “Completely Positive Linear Maps on Complex Matrices”, *Linear Algebra and Its Applications* 10, 285-290 (1975)

[13] I. Csiszar, J. Körner, *Information Theory; Coding Theorems for Discrete Memoryless Systems*, Akadémiai Kiadó, Budapest/Academic Press Inc., New York 1981

[14] I. Csiszar, P. Narayan, “The Capacity of the Arbitrarily Varying Channel Revisited: Positivity, Constraints”, *IEEE Trans. Inf. Th.* Vol. 34, No. 2, 181-193 (1989)

[15] I. Devetak and P.W. Shor, “The Capacity of a Quantum Channel for Simultaneous Transmission of Classical and Quantum Information”, *Commun. Math. Phys.* Vol. 256, Nr. 2 (2005)

[16] R. Duan, S. Severini, A. Winter, “Zero-error communication via quantum channels, non-commutative graphs and a quantum Lovász θ function”, arXiv:1002.2514v2

[17] T. Ericson, “Exponential Error Bounds for Random Codes in the Arbitrarily Varying Channel”, *IEEE Trans. Inf. Th.* Vol. 31, No. 1, 42-48 (1985)

[18] E.N. Gilbert, “A comparison of signaling alphabets”, *Bell System Tech. J.* 31, 504-522. (1952)

[19] M. Horodecki, P. Horodecki, “Reduction criterion of separability and limits for a class of distillation protocols”, *Phys. Rev. A* Vol. 59, No. 6, 4206 (1999)

[20] M. Horodecki, P. Horodecki, R. Horodecki, “General teleportation channel, singlet fraction, and quasidistillation ”, *Phys. Rev. A* 60, 18881898 (1999)

[21] S. Kakutani, “A Generalization of Brouwer’s Fixed Point Theorem”, *Duke Math. J.*, Volume 8, Number 3, 457-459 (1941)

[22] J. Kiefer, J. Wolfowitz, “Channels with arbitrarily varying channel probability functions”, *Information and Control* 5, 44-54 (1962)

[23] A.Y. Kitaev, A.H. Shen, M.N. Vyalyi, *Classical and Quantum Computation*, Graduate Studies in Mathematics 47, American Mathematical Society, Providence, Rhode Island 2002

[24] E. Knill, R. Laflamme, “Theory of quantum error-correcting codes”, *Phys. Rev. A* Vol. 55, No. 2, 900-911 (1997)

[25] J. Körner, A. Orlitsky, “Zero-error Information Theory”, *IEEE Trans. Inf. Theory* Vol. 44, No. 6, 2207-2229 (1998)

[26] D. Leung, G. Smith, “Continuity of quantum channel capacities”, *Commun. Math. Phys.* 292, 201-215, (2009)

[27] E.H. Lieb and M.B. Ruskai, “Proof of the strong subadditivity of quantum-mechanical entropy”, *J. Math. Phys.* 14, 1938 (1973)

[28] J. Matousek, *Lectures on Discrete Geometry*, Graduate Texts in Mathematics, Vol. 212, Springer 2002

[29] V.D. Milman, G. Schechtman *Asymptotic Theory of Finite Dimensional Normed Spaces*, Lecture Notes in Mathematics vol. 1200, Springer-Verlag 1986

[30] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge Studies in Advanced Mathematics vol. 78, Cambridge University Press 2002

[31] B. Schumacher, M.A. Nielsen, Quantum data processing and error correction. *Phys. Rev. A* Vol. 54, No. 4, 2629 (1996)

[32] C. E. Shannon, “The zero error capacity of a noisy channel”. *IRE Trans. Information Theory IT-2*, 8-19 (1956)

[33] J. von Neumann, “Zur Theorie der Gesellschaftsspiele”, *Math. Ann.* Vol. 100, 295-320 (1928)

[34] R. Webster, *Convexity*, Oxford University Press 1994

[35] J. Yard, I. Devetak, P. Hayden, “Capacity theorems for quantum multiple access channels: Classical-quantum and quantum-quantum capacity Regions”, *IEEE Trans. Inf. Theory* 54, 3091 (2008) e-print arXiv:quant-ph/0501045.