Arbitrarily Small Amounts of Correlation for Arbitrarily Varying Quantum Channels

Holger Boche* and Janis Nötzel*
*Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Germany
Email: {boche, janis.noetzel}@tum.de

***Abstract*—As our main result we show that, in order to achieve the randomness assisted message - and entanglement transmission capacities of a finite arbitrarily varying quantum channel it is not necessary that sender and receiver share (asymptotically perfect) common randomness. Rather, it is sufficient that they each have access to an unlimited amount of uses of one part of a correlated bipartite source. This access might be restricted to an arbitrary small (nonzero) fraction per channel use, without changing the main result.**
We investigate the notion of common randomness. It turns out that this is a very costly resource - generically, it cannot be obtained just by local processing of a bipartite source. This result underlines the importance of our main result.
Also, the asymptotic equivalence of the maximal- and average error criterion for classical message transmission over finite arbitrarily varying quantum channels is proven.
At last, we prove a simplified symmetrizability condition for finite arbitrarily varying quantum channels.

I. Introduction

An arbitrarily varying quantum channel (we will use the shorthand 'AVQC' henceforth) is defined by a set $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ of quantum channels which all have the same input- and output system. These systems are controlled by a sender $\mathfrak{S}$ and a receiver $\mathfrak{R}$, who wish to transmit either entanglement or classical messages.
They do so by l-fold usage of the AVQC, which is partially under the control of a third party, called the aversary $\mathfrak{A}$ who is able to select either one of the channels $\mathcal{N}_{s^l} := \mathcal{N}_{s_1}\otimes\ldots\otimes\mathcal{N}_{s_l}$ for which $s^l \in \mathbf{S}^l$. It is understood that $\mathfrak{S}$ and $\mathfrak{R}$ have to select their protocol first, after that $\mathfrak{A}$ makes his choice of channel sequence s^l. We consider the case $l\to\infty$.
This scenario can also be understood as an attack on the communication between $\mathfrak{S}$ and $\mathfrak{R}$, where the only aim of $\mathfrak{A}$ is to jam the communication.
Recent work [4] provided a formula for the 'random entanglement transmission capacity' $\mathcal{A}_{\mathrm{r}}(\mathfrak{I})$ of such a channel when $\mathfrak{S}$ and $\mathfrak{R}$ are allowed to use an unlimited amount of shared randomness in order to perform a possibly correlated randomization over their encoding and decoding strategies. They also showed that already a polynomial (in the number of channel uses) amount of *common randomness* suffices to achieve rates arbitrarily close to $\mathcal{A}_{\mathrm{r}}(\mathfrak{I})$.
Using this result, they then showed that it was sufficient for $\mathfrak{S}$ and $\mathfrak{R}$ to be able to establish common randomness first by sending classical messages, then use a few *deterministic* entanglement transmission codes afterwards. This led to their 'Quantum Ahlswede Dichotomy', stating that the deterministic entanglement transmission capacity $\mathcal{A}_{\mathrm{d}}(\mathfrak{I})$ of an AVQC $\mathfrak{I}$ equals its random entanglement transmission capacity, if its deterministic message transmission capacity $\overline{C}_{\mathrm{d}}(\mathfrak{I})$ is greater than zero. Since $\mathcal{A}_{\mathrm{d}}(\mathfrak{I}) \leq \overline{C}_{\mathrm{d}}(\mathfrak{I})$, the very same statement holds true with $\overline{C}_{\mathrm{d}}(\mathfrak{I})$ replaced by $\mathcal{A}_{\mathrm{d}}(\mathfrak{I})$.
They conjectured that $\mathcal{A}_{\mathrm{d}}(\mathfrak{I}) = \mathcal{A}_{\mathrm{r}}(\mathfrak{I})$ holds f.a. AVQCs $\mathfrak{I}$.
We start an investigation on that conjecture by looking for the least amount of randomness that enables entanglement transmission at $\mathcal{A}_{\mathrm{r}}(\mathfrak{I})$. We distinguish between four different types of entanglement- and message transmission codes for AVQCs. Each class requires a stronger resource than the one before.
1) Deterministic codes. $\mathfrak{S}$ and $\mathfrak{R}$ agree on one encoding- and decoding scheme, $\mathfrak{A}$ selects a channel sequence, and the transmission (of either entanglement or messages at a certain rate) starts. It is successfull, if it is asymptotically perfect for *every* choice of infinite channel sequence $(\hat{s}^l)_{l\in\mathbb{N}}$.
2) $((X,Y),r)-$**correlated codes.** In addition to 1), an i.i.d. random variable (X,Y) with values in some finite set $\mathbf{X}\times\mathbf{Y}$ is given. $\mathfrak{S}$ observes X and $\mathfrak{R}$ observes Y. Every r-th channel use, they obtain one pair of realizations of (X,Y). Their encoding and decoding may depend on the outcomes of (X,Y), which are hidden from $\mathfrak{A}$. In order to avoid trivialities, we assume that $I(X,Y)>0$ holds. Apart from that, (X,Y) will be arbitrary but fixed from now on. $((X,Y),r)$ will also be named 'correlation'.
3) Common randomness assisted codes. $\mathfrak{S}$ and $\mathfrak{R}$ each have access to one part of a random variable that, for l-fold usage of the channel, outputs pairs of elements taken from a set $\Gamma_l\times\Gamma_l$. It is guaranteed that the probability distribution according to which these elements are chosen converges to the equidistribution on the subset of pairs of identical elements and that $\lim_{l\to\infty}|\Gamma_l| = \infty$. Such sequence is equivalently spoken of as 'common randomness'.
4) Random codes. This is the most general class. It consists of the whole set of probability measures on the set of encoding and decoding schemes. It contains all the other classes as special cases.
Our results are the following.
Generically, common randomness is a strictly stronger resource than correlation, *but* $((X,Y),r)-$correlated codes are already enough to achieve either $\mathcal{A}_{\mathrm{r}}(\mathfrak{I})$ or the random message transmission capacity $\overline{C}_{\mathrm{r}}(\mathfrak{I})$, regardless of the value of r.
Two additional results are that the message transmission capacities of an AVQC are independent from the choice of either

978-1-4799-0446-4/13/$31.00 ©2013 IEEE

maximal- or average error criterion and a simpler version of the symmmetrizability conditions stated in [4].

II. Notation and conventions

All Hilbert spaces are assumed to have finite dimension and are over the field $\mathbb{C}$. The set of linear operators on Hilbert space $\mathcal{H}$ is denoted $\mathcal{B}(\mathcal{H})$. $\mathcal{S}(\mathcal{H})$ is the set of states, i.e. positive semi-definite operators with trace (denoted tr) 1 in $\mathcal{B}(\mathcal{H})$.

For $N \in \mathbb{N}$, $[N]$ is the shortcut for the set $\{1, ..., N\}$. For a finite set $\mathbf{X}$ the notation $\mathfrak{P}(\mathbf{X})$ denotes the set of probability distributions on $\mathbf{X}$, and $|\mathbf{X}|$ its cardinality. If $l \in \mathbb{N}$, we define $\mathbf{X}^l := \{x^l = (x_1, \ldots, x_l) : x_i \in \mathbf{X}\ \forall i \in [l]\}$.

The set of completely positive trace preserving (CPTP) maps (quantum channels) from $\mathcal{B}(\mathcal{H})$ to $\mathcal{B}(\mathcal{K})$ is denoted by $\mathcal{C}(\mathcal{H}, \mathcal{K})$. Closely related is the set of classical-quantum channels (abbreviated as 'cq-channels') with finite input alphabet $\mathbf{Z}$ and outputs in $\mathcal{B}(\mathcal{K})$, that arises from $\mathcal{C}(\mathcal{H}, \mathcal{K})$ by setting $d = |\mathbf{Z}|$ and restricting the inputs to matrices that are diagonal in a specific basis. It is denoted $CQ(\mathbf{Z}, \mathcal{K})$.

Writing $\mathcal{B}_+(\mathcal{H})$ for the nonnegative elements of $\mathcal{B}(\mathcal{H})$, the set of measurements (POVMs) on $\mathcal{H}$ with $N \in \mathbb{N}$ outcomes is written $\mathcal{M}_N(\mathcal{H}) := \{\mathbf{D} = (D_i)_{i=1}^N : \sum_i D_i = \mathbb{1}_{\mathcal{H}}\} \cap \mathcal{B}_+(\mathcal{H})$. The symbol $\log(\cdot)$ denotes the base two logarithm which is used throughout the paper.

Given a bipartite random variable (X, Y), its mutual information $I(X, Y)$ is given by $I(X, Y) := H(X) + H(Y) - H(X, Y)$, where $H(\cdot)$ is the usual Shannon entropy.

For $\rho \in \mathcal{S}(\mathcal{H})$ and $\mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{H})$ the entanglement fidelity is given by $F_e(\rho, \mathcal{N}) := \mathrm{tr}\{\psi(id_{\mathcal{B}(\mathcal{H})} \otimes \mathcal{N})(\psi)\}$, with $\psi \in \mathcal{H} \otimes \mathcal{H}$ being an arbitrary purification of the state ρ.

For a finite set $\mathcal{W} = \{W_s\}_{s \in \mathbf{S}} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$ or $\mathcal{W} = \{W_s\}_{s \in \mathbf{S}} \subset CQ(\mathbf{Z}, \mathcal{K})$ we denote its convex hull by $\mathrm{cnv}(\mathcal{W})$. In the cases considered here, $\mathrm{cnv}(\mathcal{W})$ is given by $\mathrm{cnv}(\mathcal{W}) = \left\{W_q : W_q = \sum_{s \in \mathbf{S}} q(s) W_s,\ q \in \mathfrak{P}(\mathbf{S})\right\}$.

If $F \subset V$ is a convex subset of some finite dimensional normed space over the field of real or complex numbers, its relative interior is denoted $\mathrm{ri}\, F$.

III. Definitions

For the rest of this paper, let $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$ denote a finite AVQC. We also follow the convention of [4], using the term 'the AVQC $\mathfrak{I}$' as a linguistic shortcut for the mathematical object $(\{\mathcal{N}_{s^l}\}_{s^l \in \mathbf{S}^l})_{l \in \mathbb{N}}$.

We will now define the some capacities of an AVQC.

Definition 1. *An $(l, k_l)-$random entanglement transmission code for $\mathfrak{I}$ is a probability measure μ_l on $(\mathcal{C}(\mathcal{F}_l, \mathcal{H}^{\otimes l}) \times \mathcal{C}(\mathcal{K}^{\otimes l}, \mathcal{F}_l'), \sigma_l)$, where $\mathcal{F}_l, \mathcal{F}_l'$ are Hilbert spaces, $\dim \mathcal{F}_l = k_l$, $\mathcal{F}_l \subset \mathcal{F}_l'$ and the sigma-algebra σ_l is chosen such that the function $(\mathcal{P}_l, \mathcal{R}_l) \mapsto F_e(\pi_{\mathcal{F}_l}, \mathcal{R}_l \circ \mathcal{N}_{s^l} \circ \mathcal{P}_l)$ 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 number $R \geq 0$ is said to be an achievable entanglement transmission rate for $\mathfrak{I}$ 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) $\lim_{l \to \infty} \min_{s^l \in \mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) = 1$. The random entanglement transmission capacity of $\mathfrak{I}$ is defined by*

$$\mathcal{A}_{\mathrm{r}}(\mathfrak{I}) := \sup\{R : \begin{array}{l} R \text{ is achievable entanglement trans-} \\ \text{mission rate for } \mathfrak{I} \text{ with random codes} \end{array}\}.$$

Having defined random codes and capacity for entanglement transmission we are in the position to introduce the corresponding deterministic quantities: 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 $(\mathcal{P}^l, \mathcal{R}^l)$ 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 number $R \geq 0$ is a deterministically achievable entanglement transmission rate for $\mathfrak{I}$ if it is achievable in the sense of Definition 2 for random codes with* point measures *μ_l. The deterministic entanglement transmission capacity of $\mathfrak{I}$ is given by*

$$\mathcal{A}_{\mathrm{d}}(\mathfrak{I}) := \sup\{R : \begin{array}{l} R \text{ is achievable entanglement trans-} \\ \text{mission rate for } \mathfrak{I} \text{ with det. codes} \end{array}\}.$$

Definition 4. *Let $l \in \mathbb{N}$. A random code for message transmission over $\mathfrak{I}$ is given by a probability measure γ_l on the set $(CQ(M_l, \mathcal{H}^{\otimes l}) \times \mathcal{M}_{M_l}, \Sigma_l)$, where Σ_l again denotes any $\sigma-$algebra containing all singleton sets. A deterministic code is given by a random code γ_l, where γ_l is a point measure.*

Definition 5. *$R \geq 0$ is called achievable with random codes under average error criterion if there exists a sequence $(\gamma_l)_{l \in \mathbb{N}}$ of random codes and a sequence $\varepsilon_l \searrow 0$ satisfying f.a. $l \in \mathbb{N}$ 1) $\min_{s^l \in \mathbf{S}^l} \int \sum_i \frac{1}{M_l} \mathrm{tr}\{D_i \mathcal{N}_{s^l}(\mathcal{P}(i))\} d\gamma_l(\mathcal{P}, \mathbf{D}) \geq 1 - \varepsilon_l$ and 2) $\liminf_{l \to \infty} \frac{1}{l} \log M_l \geq R$. It is called achievable with random codes under the maximal error criterion if instead of* 1) *even holds, for all $l \in \mathbb{N}$,*

3) $\min_{s^l \in \mathbf{S}^l} \min_{i \in [M_l]} \int \mathrm{tr}\{D_i \mathcal{N}_{s^l}(\mathcal{P}(i))\} d\gamma_l(\mathcal{P}, \mathbf{D}) \geq 1 - \varepsilon_l$. If the sequence $(\gamma_l)_{l \in \mathbb{N}}$ can be chosen to consist of point measures only, then R is called achievable with deterministic codes under the average error criterion if 3) *and* 2) *hold and it is called achievable with deterministic codes under the maximal error criterion if* $1'$) *and* 2) *hold.*

Definition 6. *The corresponding capacities are defined as*

$$\overline{C}_{\mathrm{d}}(\mathfrak{I}) := \sup\left\{R : \begin{array}{l} R \text{ is achievable with deterministic} \\ \text{codes under average error criterion} \end{array}\right\},$$

$$\overline{C}_{\mathrm{r}}(\mathfrak{I}) := \sup\left\{R : \begin{array}{l} R \text{ is achievable with random codes} \\ \text{under average error criterion} \end{array}\right\},$$

and $C_{\mathrm{d}}(\mathfrak{I})$, $C_{\mathrm{r}}(\mathfrak{I})$ denote the corresponding capacities w.r.t. maximal error criterion.

From [1], [4] and [5] it is clear that common randomness is a useful resource. Readers with a deeper interest in the topic

will find it fruitful to take a look at Theorem 1 and Theorem 2 a) in [1]) or Theorem 32 and Lemma 37 in [4] or Lemma 9 and Lemma 10 in [5] for applications to AVQCs and AVcqCs. The proofs given there rely on the possibility to establish common randomness between $\mathfrak{S}$ and $\mathfrak{R}$. We will define what we mean by common randomness now.

Definition 7. *A source of common randomness $CR \geq 0$ is given by a sequence $(\gamma_l)_{l\in\mathbb{N}}$ of probability distributions, where $\gamma_l \in \mathfrak{P}(\Gamma_l \times \Gamma_l)$ for every $l \in \mathbb{N}$ and, asymptotically, we have $\liminf_{l\to\infty} \frac{1}{l}\log|\Gamma_l| = CR$ and $\limsup_{l\to\infty} \|\gamma_l - \bar{\delta}_l\|_1 = 0$, where $\bar{\delta}_l \in \mathfrak{P}(\Gamma_l \times \Gamma_l)$ denotes the normalized delta function, $\bar{\delta}_l(i,j) = 1/|\Gamma_l|$ if $i = j$ and $\bar{\delta}_l(i,j) = 0$ else.*

Definition 8. *(X,Y) is said to have common randomness $CR \geq 0$ if there exists a sequence $(f_l, g_l)_{l\in\mathbb{N}}$ of functions $f_l : \mathbf{X}^l \mapsto \Gamma_l$, $g_l : \mathbf{Y}^l \mapsto \Gamma_l$ with Γ_l being a finite set and $\liminf_{l\to\infty} \frac{1}{l}\log|\Gamma_l| = CR$ and*
$\limsup_{l\to\infty} \|(f_l \times g_l) \circ p^{\otimes l} - \bar{\delta}_l\|_1 = 0$.
$CR(p) := \sup\{CR : p \text{ has common randomness } CR\}$ is called the common randomness of the $p \in \mathfrak{P}(\mathbf{X}\times\mathbf{Y})$ that (X,Y) is distributed according to.

Definition 9. *For $r, l \in \mathbb{N}$, an $((X,Y),r)-$correlated code $\mathfrak{C}_l$ for message transmission is given by a set*
$\mathfrak{C}_l = \{(M_l, \mathcal{P}_{x^{n_l}}, \mathcal{D}_{y^{n_l}})\}_{x^{n_l}\in\mathbf{X}^{n_l}, y^{n_l}\in\mathbf{Y}^{n_l}}$, *where: $M_l \in \mathbb{N}$, $\mathcal{P}_{x^{n_l}} \in CQ([M_l], \mathcal{H}^{\otimes l})$ for all $x^{n_l} \in \mathbf{X}^{n_l}$,*
$\mathcal{D}_{y^{n_l}} = \{D_{y^{n_l},1}, \ldots, D_{y^{n_l},M_l}\} \in \mathcal{M}_{[M_l]}(\mathcal{K}^{\otimes l})\ \forall\ y^{n_l} \in \mathbf{Y}^{n_l}$, *and $n_l := \lfloor l/r \rfloor$.*

Definition 10. *A number $R \geq 0$ is said to be an achievable $((X,Y),r)$ rate if there exists a sequence $(\mathfrak{C}_l)_{l\in\mathbb{N}}$ of $((X,Y)r)-$correlated codes such that $\liminf_{l\to\infty}\frac{1}{l}\log M_l \geq R$ holds and, with*
$P_e(s^l, x^{n_l}, y^{n_l}) := \frac{1}{M_l}\sum_{i=1}^{M_l} \mathrm{tr}\{D_{y^{n_l},i}\mathcal{N}_{s^l}(\mathcal{P}_{x^{n_l}}(i))\}$,
$$\lim_{l\to\infty} \min_{s^l\in\mathbf{S}^l} \sum_{x^{n_l}\in\mathbf{X}^{n_l}} \sum_{y^{n_l}\in\mathbf{Y}^{n_l}} p^{\otimes n_l}(x^{n_l}, y^{n_l}) P_e(s^l, x^{n_l}, y^{n_l}) = 1.$$

Definition 11. *The $((X,Y),r)$ message transm. capacity is*

$$\overline{C}(\mathfrak{I}, r, (X,Y)) := \sup\{R : R \text{ is achiev. } ((X,Y),r) \text{ rate}\}.$$

The $((X,Y)),r)$ entanglement transmission capacity $\mathcal{A}(\mathfrak{I}, r, (X,Y))$ is defined analogously in the obvious way.

In the proof of Theorem 6, the following will be important:

Definition 12. *For $n \in \mathbb{N}$ let $\mathcal{H}_{Y^n}$ be a Hilbert space of dimension $|\mathbf{Y}|^n$ and $\{\hat{\rho}_{y^n}\}_{y^n\in\mathbf{Y}^n} \subset \mathcal{S}(\mathcal{H}_{Y^n})$ be a set of pairwise orthogonal and pure states. For a set $\mathfrak{S}_K = \{\rho_1, \ldots, \rho_K\} \subset \mathcal{S}(\mathcal{H}^{\otimes n})$, the associated AVcqC $\mathbb{W}_{\mathfrak{S}_K} = \{W_{s^n}\}_{s^n\in\mathbf{S}^n} \subset CQ(\mathbb{F}(\mathbf{X}^n, [K]), \mathcal{H}_{\mathbf{Y}^n} \otimes \mathcal{K}^{\otimes n})$ is defined by setting, for $s^n \in \mathbf{S}^n$,*

$$W_{s^n}(f) := \sum_{x^n, y^n} p^{\otimes n}(x^n, y^n)\hat{\rho}_{y^n} \otimes \mathcal{N}_{s^n}(\rho_{f(x^n)}). \quad (1)$$

$\mathbb{F}(\mathbf{X}^n, [K])$ denotes the functions on $\mathbf{X}^n$ with values in $[K]$.

Remark 1. *The exact definition of an AVcqC can be read off from [2] or [5]. It can be thought of as an AVQC with fixed encoding operations.*

IV. Main Results

Theorem 1. *Let (X,Y) be distributed according to $p \in \mathrm{ri}\mathfrak{P}(\mathbf{X}\times\mathbf{Y})$. There is no sequence $(f_l, g_l)_{l\in\mathbb{N}}$ of functions $f_l : \mathbf{X}^l \to \Gamma_l$, $g_l : \mathbf{Y}^l \to \Gamma_l$ ($l \in \mathbb{N}$) satisfying*
(1) $\lim_{l\to\infty}|\Gamma_l| = \infty$,
(2) $\lim_{l\to\infty} p^{\otimes l}(\{(x^l, y^l) : f_l(x^l) = g_l(y^l)\}) = 1$,
(3) $\lim_{l\to\infty} p^{\otimes l}_{\mathbf{X}}(\{x^l : f_l(x^l) = k_l\}) = 0$,
$\lim_{l\to\infty} p^{\otimes l}_{\mathbf{Y}}(\{y^l : g_l(x^l) = k_l\}) = 0\ \forall\ (k_l)_{l\in\mathbb{N}} \subset \times_{l=1}^{\infty}\Gamma_l$.
Further, the set of probability distributions $p \in \mathfrak{P}(\mathbf{X}\times\mathbf{Y})$ satisfying $CR(p) > 0$ is closed.

Remark 2. *Every point at which the function CR with domain $\mathfrak{P}(\mathbf{X}\times\mathbf{Y})$ is positive is also a discontinuity point, and the set of $p \in \mathfrak{P}(\mathbf{X}\times\mathbf{Y})$ with $CR(p) > 0$ is highly exceptional: Its complement is open and dense in $\mathfrak{P}(\mathbf{X}\times\mathbf{Y})$, its Lebesgue measure zero. Operationally it is highly unstable w.r.t. small perturbations. And if $CR(p) = 0$, not even a* polynomially *small amount of common randomness can be extracted!*

The importance of this statement stems from the strategy of proof that is used in [1], [4], [5] to establish the Ahlswede Dichotomy (find the original in [1], Theorem 1) in its various forms. An example of the advantage that random codes offer over deterministic ones can be found in [2] for AVcqCs. It is conjectured [4] that $\mathcal{A}_{\mathrm{d}}(\mathfrak{I}) = \mathcal{A}_{\mathrm{r}}(\mathfrak{I})$ holds for every AVQC $\mathfrak{I}$ and, at last, our belief that the case $\overline{C}_{\mathrm{r}}(\mathfrak{I}) > \overline{C}_{\mathrm{r}}(\mathfrak{I})$ occurs. This underlines the importance of our main results:

Theorem 2. *F.a. $r \in \mathbb{N}$ it holds $\overline{C}(\mathfrak{I}, r, (X,Y)) = \overline{C}_{\mathrm{r}}(\mathfrak{I})$.*

Theorem 3. *F.a. $r \in \mathbb{N}$ we have $\mathcal{A}(\mathfrak{I}, r, (X,Y)) = \mathcal{A}_{\mathrm{r}}(\mathfrak{I})$.*

Another important result is the equivalence of maximal - and average error criterion for AVQCs. It should be compared to the results of [1]: The capacity of an arbitrarily varying classical channel generally depends on which of the two criteria one uses. For randomized encoding, the two capacities coincide. The codes that are used in our definition of the two capacities allow for a randomized encoding, since the signal states that get fed into the channel at senders side are allowed to be mixed. Taking this into account, the following result is not too surprising:

Theorem 4. *It holds $\overline{C}_{\mathrm{d}}(\mathfrak{I}) = C_{\mathrm{d}}(\mathfrak{I})$.*

In order to find out whether the case $\overline{C}_{\mathrm{r}}(\mathfrak{I}) > 0$ but $\overline{C}_{\mathrm{d}}(\mathfrak{I}) = 0$ occurs it might be necessary to prove that it is $l-$symmetrizable f.a. $l \in \mathbb{N}$ (see [4], Definition 39, Theorem 40). We provide a simplified formula for l-symmetrizability here, that only requires to find a finite number of probability distributions instead of infinitely many:

Theorem 5. *Let $l \in \mathbb{N}$ and $A_1, \ldots, A_K$ be such that $\mathcal{S}(\mathcal{H}^{\otimes l}) \subset \mathrm{cnv}(\{A_i\}_{i=1}^K)$. $\mathfrak{I} = \{\mathcal{N}_s\}_{s\in\mathbf{S}}$ is $l-$symmetrizable if and only if there is a set $\{p_i\}_{i=1}^K \subset \mathfrak{P}(\mathbf{S}^l)$ such that f.a. $i, j \in [K]$: $\sum_{s^l\in\mathbf{S}^l} q_j(s^l)\mathcal{N}_{s^l}(A_i) = \sum_{s^l\in\mathbf{S}^l} q_i(s^l)\mathcal{N}_{s^l}(A_j)$.*

Theorem 5 provides the more handy criterion, it also allows a more geometric view on the symmetrizability condition.

V. Scetch of proofs

In a first step, we give minimal requirements for the kind of common randomness we can put to use in order to achieve $\overline{C}_{\mathrm{r}}(\mathfrak{I})$ or $\mathcal{A}_{\mathrm{r}}(\mathfrak{I})$ when $\overline{C}_{\mathrm{d}}(\mathfrak{I}) = 0$, an assumption that can safely be made since otherwise we already achieve $\mathcal{A}_{\mathrm{r}}(\mathfrak{I})$ by [4], Theorem 5, even without using common randomness.

Lemma 1. *Let $\mathfrak{I}$ be symmetrizable. $(\gamma_l)_{l\in\mathbb{N}}$, each $\gamma_l \in \mathfrak{P}(\Gamma_l \times \Gamma_l)$ and $|\Gamma_l| < \infty$, can lead to a positive message transmission rate over $\mathfrak{I}$ only if it satisfies 1) $\liminf_{l\to\infty} \gamma_l(k_l, k_l') = 0$, 2) $\liminf_{l\to\infty} \gamma_{S,l}(k_l) = 0$ and 3) $\liminf_{l\to\infty} \gamma_{R,l}(k_l') = 0$ f.a. sequences $(k_l, k_l')_{l\in\mathbb{N}}$ satisfying $k_l, k_l' \in \Gamma_l$ f.a. $l \in \mathbb{N}$ and with $\gamma_{S,l}$ and $\gamma_{R,l}$ being the marginal distributions of the γ_l.*

Remark 3. *It follows from Lemma 1 that no finite amount of common randomness is sufficient to achieve a positive message transmission capacity over a symmetrizable AVQC. This also holds for entanglement transmission, since $\mathcal{A}_{\mathrm{r}}(\mathfrak{I}) \le \overline{C}_{\mathrm{r}}(\mathfrak{I})$.*

Proof. Suppose there is a sequence $(k_l, k_l')_{l\in\mathbb{N}}$ satisfying $k_l, k_l' \in \Gamma_l$ f.a. $l \in \mathbb{N}$ s.t. $\liminf_{l\to\infty} \gamma_l(k_l, k_l') = c > 0$ and $M_l \ge 2$ for large enough $l \in \mathbb{N}$. W.l.o.g., $k_l = k_l' = 1$ f.a. $l \in \mathbb{N}$. Given that $\mathfrak{I}$ is symmetrizable, the proof of Theorem 40 in [4] shows that there exists a sequence $(\mathbf{s}^l)_{l\in\mathbb{N}}$ s.t. $\frac{1}{M_l}\sum_{i=1}^{M_l} \mathrm{tr}\{D_{1,i}\mathcal{N}_{\mathbf{s}^l}(\mathcal{P}_1(i))\} \le 3/4$ holds, hence

$$\liminf_{l\to\infty} \sum_{k_l,k_l'\in\Gamma_l} \frac{\gamma_l(k_l,k_l')}{M_l} \sum_{i=1}^{M_l} \mathrm{tr}\{D_{k_l',i}\mathcal{N}_{\mathbf{s}^l}\circ\mathcal{P}_{k_l}(i)\} \le 1-\frac{c}{4} < 1.$$

Statements 2) and 3) are proven in a similar fashion. □

Proof of Theorem 1. Assume there is such sequence of functions. For $k \in \Gamma_l$ ($l \in \mathbb{N}$ arbitrary), use the abbreviations
$a_l(k) := p_{\mathbf{X}}^{\otimes l}(\{x^l : f_l(x^l) = k\})$,
$b_l(k) := p_{\mathbf{Y}}^{\otimes l}(\{y^l : g_l(y^l) = k\})$,
$c_l(k) := p^{\otimes l}(\{(x^l, y^l) : f_l(x^l) = g_l(y^l) = k\})$.
Let $\varepsilon > 0$ and $l \in \mathbb{N}$ satisfy
$1 - \varepsilon \le \sum_{k=1}^{|\Gamma_l|} c_l(k) \le 1,\ a_l(k) \le \varepsilon,\ b_l(k) \le \varepsilon\ \forall k \in \Gamma_l$.
By choosing l large enough, we can make ε arbitrarily small. Take the monotone increasing sequences $(A_m)_{m=1}^{|\Gamma_l|}, (B_m)_{m=1}^{|\Gamma_l|}$ defined by $A_m := \sum_{k=1}^{m} a_l(k),\ B_m := \sum_{k=1}^{m} b_l(k)$. Let $0 < \sigma < 1/2$ and $\hat{m}$. It can be shown that

$$\sigma \le \min\{A_{\hat{m}}, B_{\hat{m}}\} \le \max\{A_{\hat{m}}, B_{\hat{m}}\} \le \sigma + 2\varepsilon. \quad (2)$$

Define $\Theta_l : \Gamma_l \to \{0,1\}$ by $\Theta_l(k) := 1$, if $1 \le k \le \hat{m}$ and 0, else. Then $p^{\otimes l}(\Theta_l \circ f_l = \Theta_l \circ g_l) = \sum_{k=1}^{|\Gamma_l|} c_l(k) \ge 1 - \varepsilon$ and

$$p^{\otimes l}(\Theta_l \circ f_l \ne \Theta_l \circ g_l) \le \varepsilon. \quad (3)$$

Also, for ε small enough ($\varepsilon \le (1-2\sigma)/2$ is sufficient) we get

$$p_{\mathbf{X}}^{\otimes l}(\Theta_1 \circ f_l = 1),\quad p_{\mathbf{Y}}^{\otimes l}(\Theta_1 \circ g_l = 1) \subset [\sigma, 1-\sigma]. \quad (4)$$

But this is impossible by ([6], 1. Problem statement), since $p \in \mathrm{ri}\mathfrak{P}(\mathbf{X}\times\mathbf{Y})$ implies that there is $\varepsilon > 0$ s.t. $p(A) > \varepsilon$ f.a. $A \subset \mathbf{X}\times\mathbf{Y}$. It follows that a sequence $(f_l, g_l)_{l\in\mathbb{N}}$ satisfying the three conditions in Lemma 1 cannot exist.
Showing that the set of $p \in \mathfrak{P}(\mathbf{X}\times\mathbf{Y})$ satisfying $CR(p) > 0$ is closed is comparably easy, a proof can be found in [8]. □

In the proofs of Theorems 6, 3 and 2 we assume w.l.o.g. $\mathbf{X} = \mathbf{Y} = \{0,1\}$, since for every (X', Y') with $I(X', Y') > 0$, there are functions $f : \mathbf{X} \to [2]$, $g : \mathbf{Y} \to [2]$ s.t. $I(f(X), g(Y)) > 0$. These are incorporated into the code.
The following is crucial to the proof of Theorem 2:

Theorem 6. *If $\overline{C}(\mathfrak{I}, r, (X, Y)) = 0$ for any $r \in \mathbb{N}$, then for every $l \in \mathbb{N}$ and set $\mathfrak{S}_K \subset \mathcal{S}(\mathcal{H}^{\otimes rl})$ the associated AVcqC $\mathbb{W}^l_{\mathfrak{S}_K}$ satisfies $\overline{C}(\mathbb{W}^l_{\mathfrak{S}_K}) = 0$, and this implies $\overline{C}_{\mathrm{r}}(\mathfrak{I}) = 0$.*

Proof. First, for clarity of the argument, assume $r = 1$. For every sequence $(\mathfrak{C}_l)_{l\in\mathbb{N}}$ of correlated codes with $\liminf_{l\to\infty} \frac{1}{l}\log M_l > 0$, it holds that

$$\lim_{l\to\infty} \min_{s^l\in\mathbf{S}^l} \sum_{x^l\in\mathbf{X}^l} \sum_{y^l\in\mathbf{Y}^l} p^{\otimes l}(x^l, y^l) P_e(s^l, x^l, y^l) < 1. \quad (5)$$

Let $\{\rho_1, \dots, \rho_K\} \subset \mathcal{S}(\mathcal{H}^{\otimes l})$ and $f_j^{(i)} : \mathbf{X} \to [K]$ $(i = 1, \dots, M_l,\ j = 1, \dots, l)$. Define the encoding maps $\mathcal{P}_{x^l}(i) := \otimes_{j=1}^{l} \rho_{f_j^{(i)}(x_j)}$. To every $\mathbf{D} \in \mathcal{M}_{M_l}(\mathcal{H}_{Y^l} \otimes \mathcal{K}^{\otimes l})$, assign POVMs $\{\mathbf{D}_{y^l}\}_{y^l\in\mathbf{Y}^l} \subset \mathcal{M}_{M_l}(\mathcal{K}^{\otimes l})$ through $D_{y^l,i} := \mathrm{tr}_{\mathcal{H}_{Y^l}}\{\hat{\rho}_{y^l} \otimes \mathbb{1}_{\mathcal{K}^{\otimes l}} D_i\}$. Then for $\mathbb{W}_{\mathfrak{S}_K}$:

$$\sum_{i=1}^{M_l} \mathrm{tr}\{D_i W_{s^l}(\bigtimes_{j=1}^{l} f_j^{(i)})\} \quad (6)$$

$$= \sum_{i=1}^{M_l} \sum_{y^l\in\mathbf{Y}^l} \sum_{x^l\in\mathbf{X}^l} p^{\otimes l}(x^l, y^l) \mathrm{tr}\{D_{y^l,i}\mathcal{N}_{s^l}(\otimes_{j=1}^{l}\rho_{f_j^{(i)}(x_j)})\}. \quad (7)$$

Thus, no code for the AVcqC $\mathbb{W}_{\mathfrak{S}_K}$ can have asymptotically vanishing average error *and* positive rate at the same time, hence $\overline{C}(\mathbb{W}_{\mathfrak{S}_K}) = 0$ for every such AVcqC.
This proves the first part of the statement.
For the second, consider Theorem 1 from [2]: Every of the AVcqC's $\mathbb{W}_{\mathfrak{S}_K}$ has zero capacity for transmission of messages using average error criterion, hence is symmetrizable in the sense of [2], and it follows that for every such $\mathbb{W}_{\mathfrak{S}_K}$ there exists $\{\tau_f\}_{f\in\mathbb{F}(\mathbf{X},[K])} \subset \mathfrak{P}(\mathbf{S})$ s.t. $\forall f, f' \in \mathbb{F}(\mathbf{X}, [K])$:

$$\sum_{s\in\mathbf{S}} \tau_f(s) W_s(f') = \sum_{s\in\mathbf{S}} \tau_{f'}(s) W_s(f). \quad (8)$$

Define $f_1, \dots, f_K, f^* : \mathbf{X} \to [K]$ by $f_i(0) = i,\ f_i(1) = i \oplus 1$ ($\oplus$ denotes addition mod K) and $f^*(0) = f^*(1) = 1$.
Inserting these, we get for $i = 1, \dots, K$ the following:

$$\sum_{s\in\mathbf{S}} \tau_{f_i}(s) W_s(f^*) = \sum_{s\in\mathbf{S}} \tau_{f^*}(s) W_s(f_i), \quad (9)$$

and with $p(\cdot|0),\ p(\cdot|1) \in \mathfrak{P}(\mathbf{X})$ defined by $p(\cdot|y) := p(\cdot, y)/(p(0,y)+p(1,y))$, $y = 0, 1$ (and $I(X, Y) > 0$ implying $p(\cdot|0) \ne p(\cdot|1)$) it follows for $y = 0, 1,\ i \in [K]$,

$$\sum_{s\in\mathbf{S}} \tau_{f_i}(s) \mathcal{N}_s(\rho_1) = \sum_{x\in\mathbf{X}} p(x|y) \sum_{s\in\mathbf{S}} \tau_{f^*}(s) \mathcal{N}_s(\rho_{x\oplus i}), \quad (10)$$

hence setting, for all $i \in [K]$,

$$\tilde{\sigma}_i := \sum_{s\in\mathbf{S}} \tau_{f_i}(s) \mathcal{N}_s(\rho_1),\ \sigma_i := \sum_{s\in\mathbf{S}} \tau_{f^*}(s) \mathcal{N}_s(\rho_i), \quad (11)$$

we know that for every $i \in \{1, \dots, K\}$

$$p(0|1)\sigma_i + p(1|1)\sigma_{i\oplus 1} = \tilde{\sigma}_i = p(0|0)\sigma_i + p(1|0)\sigma_{i\oplus 1}. \quad (12)$$

This can only be if σ_i is independent of $i \in [K]$, hence the cq-channel $W^*(i) := \sum_{s\in\mathbf{S}} \tau_{f^*}(s)\mathcal{N}_s(\rho_i) \in CQ([K], \mathcal{K})$ is constant. Define the AVcqC $\{W_s\}_{s\in\mathbf{S}} \subset CQ([K], \mathcal{K})$ by $W_s(i) := \mathcal{N}_s(\rho_i)$. Since $\mathrm{cnv}(\{W_s\}_{s\in\mathbf{S}})$ contains the constant channel W^* this implies that $\{W_s\}_{s\in\mathbf{S}}$ satisfies $\overline{C}_{\mathrm{r}}(\{W_s\}_{s\in\mathbf{S}}) = 0$ by [2], Theorem 1. This holds regardless of the set $\{\rho_1, \dots, \rho_K\}$ that was chosen, only the $\tau_\cdot(\cdot)$ change for every such set.
The whole argument can be gone through for every $\mathfrak{I}^{\otimes l} := \{\mathcal{N}_{s^l}\}_{s^l\in\mathbf{S}^l}$, $l \in \mathbb{N}$. It follows that every $\{W_{s^l}\}_{s^l\in\mathbf{S}^l}$ with $W_{s^l} \in CQ([K], \mathcal{H}^{\otimes l})$ defined by $W_{s^l}(i) := \mathcal{N}_{s^l}(\rho_i)$ via a set $\{\rho_1, \dots, \rho_K\} \subset \mathcal{S}(\mathcal{H}^{\otimes l})$ satisfies $\overline{C}_{\mathrm{r}}(\{W_{s^l}\}_{s^l\in\mathbf{S}^l}) = 0$.
But this implies $\overline{C}_{\mathrm{r}}(\mathfrak{I}) = 0$. For a detailed proof, see [8].
Now let r be arbitrary. Then the above argument, applied to $\mathfrak{I}^r := \{\mathcal{N}_{s^r}\}_{s^r\in\mathbf{S}^r}$ shows that $\overline{C}_{\mathrm{r}}(\mathfrak{I}^r) = 0$ has to hold. But $C_{\mathrm{r}}(\mathfrak{I}^r) = r \cdot C_{\mathrm{r}}(\mathfrak{I})$, hence $\overline{C}_{\mathrm{r}}(\mathfrak{I}) = 0$. □

Proof of Theorem 2. By Theorem 6, $\overline{C}(\mathfrak{I}, r, (X, Y)) = 0 \Rightarrow \overline{C}_{\mathrm{r}}(\mathfrak{I}) = 0$. Let $\overline{C}(\mathfrak{I}, r, (X, Y) > 0$. There is a sequence $(\mathfrak{C}_l)_{l\in\mathbb{N}}$ of $((X, Y), r)$-correlated codes s.t. $N_l := \lfloor 2^{l\frac{2}{l}\log(l)} \rfloor$ messages are transmitted ($l \in \mathbb{N}$), and for some sequence $\delta_l \searrow 0$., f.a. $s^{n_l} \in \mathbf{S}^{n_l}$

$$\sum_{x^{n_l}, y^{n_l}} p^{\otimes n_l}(x^{n_l}, y^{n_l}) P_e(s^{n_l}, x^{n_l}, y^{n_l}) \geq 1 - \delta_l. \quad (13)$$

Take a sequence of codes $(\{(M_l, \mathcal{D}_i, \mathcal{P}_i)\}_{i=1}^{l^2})_{l\in\mathbb{N}}$ for message transmission over $\mathfrak{I}$ such that, for some sequence $\varepsilon_l \searrow 0$,

$$\min_{s^l\in\mathbf{S}^l} \sum_{i=1}^{l^2} \frac{1}{l^2} \frac{1}{M_l} \sum_{j=1}^{M_l} \mathrm{tr}\{D_{i,j}\mathcal{N}_{s^l} \circ \mathcal{P}_i^l(j)\} \geq 1 - \varepsilon_l, \quad (14)$$

$$\liminf \frac{1}{l} \log M_l \geq \overline{C}_{\mathrm{r}}(\mathfrak{I}) - \eta, \ \eta > 0 \text{ arbitrary}. \quad (15)$$

Such codes exist by Lemma 9 and Lemma 10 in [5]. Concatenation of the two codes yields a correlated code at a rate $\overline{C}_{\mathrm{r}}(\mathfrak{I}) - \eta$, proving our claim. □

Proof of Theorem 3. This proof is done in a way so similar to the one of Theorem 2, that we totally omit it. □

The following theorem is essential for proving Theorem 5.

Theorem 7. *Let $\{N_s\}_{s\in\mathbf{S}}$ be a finite set of linear maps from $\mathbb{C}^n$ to $\mathbb{C}^m$, $\mathfrak{S} = \{t_i\}_{i=1}^K \subset \mathbb{C}^n$ and $p_1 \dots, p_K \subset \mathfrak{P}(\mathbf{S})$ s.t.*

$$\sum_{s\in\mathbf{S}} p_i(s)N_s(t_j) = \sum_{s\in\mathbf{S}} p_j(s)N_s(t_i) \qquad \forall i, j \in [K]. \quad (16)$$

To $\{t_i\}_{i=K+1}^N \subset \mathrm{cnv}(\mathfrak{S})$, $N \geq K$, $\exists\ \{p_i\}_{i=K+1}^N \subset \mathfrak{P}(\mathbf{S})$ s.t.

$$\sum_{s\in\mathbf{S}} p_i(s)N_s(t_j) = \sum_{s\in\mathbf{S}} p_j(s)N_s(t_i) \qquad \forall i, j \in [N]. \quad (17)$$

Proof. Take a cc-channel $r : [N] \to \mathfrak{P}[N]$ such that $t_i = \sum_{j=1}^N r(j|i)t_j\ \forall i \in [N]$. We may choose r such that $r(i|j) = \delta(i, j)$ for all $i \in [N]$ and $j \in [K]$. Define $p_{K+1}, \dots, p_N$ by $p_i := \sum_{j=1}^K r(j|i)p_j$, then it follows that (17) holds. □

Proof of Theorem 5. If $\mathfrak{I}$ is l-symmetrizable according to Theorem 5, then by Theorem 7 it is l-symmetrizable in the sense of [4]. If it is not l-symmetrizable in the sense of [4], then it can, by Theorem 7, not be l-symmetrizable according to Theorem 5. □

Proof of Theorem 4. In the first part of the proof we show that $\overline{C}_{\mathrm{d}}(\mathfrak{I}) > 0$ if and only if $C_{\mathrm{d}}(\mathfrak{I}) > 0$. Then, we show that $C_{\mathrm{d}}(\mathfrak{I}) > 0$ implies $C_{\mathrm{d}}(\mathfrak{I}) = C_{\mathrm{r}}(\mathfrak{I})$ and, at last, $\overline{C}_{\mathrm{r}}(\mathfrak{I}) = C_{\mathrm{r}}(\mathfrak{I})$.
It is clear that $C_{\mathrm{d}}(\mathfrak{I}) > 0$ implies $\overline{C}_{\mathrm{d}}(\mathfrak{I}) > 0$. On the other hand, if $\overline{C}_{\mathrm{d}}(\mathfrak{I}) > 0$ then for every $\varepsilon > 0$ there is $l \in \mathbb{N}$ and $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{H}^{\otimes l})$ as well as a POVM consisting of D_1, D_2 such that

$$\frac{1}{2}(\mathrm{tr}\{D_1\mathcal{N}_{s^l}(\rho_1) + D_2\mathcal{N}_{s^l}(\rho_2)\}) \geq 1 - \varepsilon\ \forall\ s^l \in \mathbf{S}^l. \quad (18)$$

It follows $\mathrm{tr}\{D_i\mathcal{N}_{s^l}(\rho_i)\} \geq 1 - 2\varepsilon\ \forall\ s^l \in \mathbf{S}^l,\ i = 1, 2$. Assuming $C_{\mathrm{d}}(\mathfrak{I}) = 0$ and using Theorem 42 in [4] can be shown to lead to a contradiction, hence $C_{\mathrm{d}}(\mathfrak{I}) > 0$.
Proving $C_{\mathrm{d}}(\mathfrak{I}) > 0 \Rightarrow C_{\mathrm{d}}(\mathfrak{I}) = C_{\mathrm{r}}(\mathfrak{I})$ is straight along the lines of the proof in the classical case [1]. It rests on an adaption of Lemma 37 in [4] to the maximal error criterion.
$C_{\mathrm{r}}(\mathfrak{I}) \leq \overline{C}_{\mathrm{r}}(\mathfrak{I})$ is trivially true, it remains to show the reverse. For every sequence $(\mu_l)_{l\in\mathbb{N}}$ of random codes, just define a new sequence of random codes $(\hat{\mu}_l)_{l\in\mathbb{N}}$ by setting, f.a. $A \subset \Sigma_l$,

$$\hat{\mu}_l(A) := \sum_{\tau\in\mathrm{Prm}([M_l])} \mu_l(\tau(A)) \frac{1}{M_l}. \quad (19)$$

$\mathrm{Prm}([M_l])$ is the set of permutations on M_l. For $(\mathcal{P}, \mathbf{D}) \in CQ([M_l], \mathcal{H}^{\otimes l}) \times \mathcal{M}_{M_l}(\mathcal{K}^{\otimes l})$ we define $\tau(\mathcal{P}, \mathbf{D}) := (\mathcal{P} \circ \tau, (D_{\tau(1)}, \dots, D_{\tau(M_l)}))$. Every $\hat{\mu}_l$ has the same performance w.r.t. maximal error criterion that μ_l had for average error. □

Acknowledgements. We thank A. Winter for stimulating discussions and remarks on the topic. This work was supported by the DFG via grant BO 1734/20-1 (H.B.) and by the BMBF via grant 01BQ1050 (H.B., J.N.).

References

[1] R. Ahlswede, "Elimination of Correlation in Random Codes for Arbitrarily Varying Channels", *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 44, 159-175 (1978)
[2] R. Ahlswede, V. Blinovsky, "Classical capacity of classical-quantum arbitrarily varying channels", IEEE Trans. Inf. Theory, Vol. 53, No. 2, 526-533.
[3] R. Ahlswede, N. Cai, "Correlated sources help the transmission over AVC", *IEEE Trans. Inf. Th.*, Vol. 43, No. 4, 1254-1255 (1997)
[4] R. Ahlswede, I. Bjelakovic, H. Boche, J. Nötzel "Quantum capacity under adversarial noise: arbitrarily varying quantum channels", *Comm. Math. Phys.*, Vol. 317, Iss. 1, 103-156, 10.1007/s00220-012-1613-x, (2013)
[5] I. Bjelakovic, H. Boche, G. Janßen, J. Nötzel, "Arbitrarily varying and compound classical-quantum channels and a note on quantum zero-error capacities", Ahlswede Festschrift, LNCS 7777, 247-283, (2013)
[6] H. S. Witsenhausen, "On sequences of pairs of dependent random variables", SIAM J. Appl. Math. Vol. 28, No. 1, (1975)
[7] 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)
[8] H. Boche, J. Nötzel, "Arbitrarily Small Amounts of Correlation for Arbitrarily Varying Quantum Channels", arXiv:1301.6063