Message Transmission over Classical Quantum Channels with a Jammer with Side Information

*Holger Boche  
Lehrstuhl für Theoretische Informationstechnik,  
Technische Universität München,  
Munich, Germany,  
boche@tum.de

*Minglai Cai  
Lehrstuhl für Theoretische Informationstechnik,  
Technische Universität München,  
Munich, Germany,  
minglai.cai@tum.de

Ning Cai  
School of Information Science and Technology,  
ShanghaiTech University,  
Shanghai, China,  
ningcai@shanghaitech.edu.cn

Abstract—In this paper we propose a new model for arbitrarily varying classical-quantum channels. In this model a jammer has side information. We consider two scenarios. In the first scenario the jammer knows the channel input, while in the second scenario the jammer knows both the channel input and the message. The transmitter and receiver share a secret random key with a vanishing key rate. We determine the capacity for both average and maximum error criteria. We prove that additionally revealing the message to the jammer does not change the capacity.

I. INTRODUCTION

Communication models including a jammer who tries to disturb the legal parties’ communication have received a lot of attention in recent years. These publications concentrated on the model of an arbitrarily varying channel where the jammer may change his input in every channel use. This model captures completely all possible jamming attacks and is not restricted to use a repetitive probabilistic strategy. The arbitrarily varying channel was introduced in [8]. In the model of message transmission over arbitrarily varying channels it is understood that the sender and the receiver have to select their coding scheme first. In the conventional model it is assumed that this coding scheme is known by the jammer, and he may choose the most advantageous jamming attacking strategy depending on his knowledge, but the jammer has neither knowledge about the transmitted codeword nor knowledge about the message. Ahlswede showed in [1] the surprising result, that either the deterministic capacity of an arbitrarily varying channel is zero or it is equal to its random correlated capacity (Ahlswede dichotomy). Ahlswede dichotomy demonstrates the importance of resources (shared randomness) in a very clear form. It is required that both sender and receiver have access to a perfect copy of the outcome of a random experiment, and thus we should assume an additional perfect channel. The legal channel users’ knowledges about the shared randomness is very helpful for message transmission through an arbitrarily varying channel (random correlated capacity), where we assume that the resource is only known by the legal channel users, since otherwise it will be completely useless (cf. [11]).

In this work we consider classical-quantum channels, i.e., the sender’s inputs are classical data and the receiver’s outputs are quantum systems. The capacity of classical-quantum channels has been determined in [15], [19], and [20]. The capacity of arbitrarily varying classical-quantum channels has been delivered in [4]. An alternative proof of [4]’s result and a proof of the strong converse have been given in [6]. In [3] Ahlswede dichotomy for the arbitrarily varying classical-quantum channels was established, and a sufficient and necessary condition for the zero deterministic capacity has been given. In [12] a simplification of this condition was delivered. These results are basis tools for secure communication over arbitrarily varying wiretap channels. An arbitrarily varying wiretap channel is a channel with both a jammer and an eavesdropper. Classical arbitrarily varying wiretap channels have been studied extensively in the context of classical information theory. The secrecy capacity of arbitrarily varying wiretap channels has been determined in [11]. The deterministic capacities of classical arbitrarily varying channel under maximal error criterion and under the average error criterion are in general, not equal, where the deterministic capacity formula of classical arbitrarily varying channels under maximal error criterion is still an open problem. Interestingly, [12] shows that the deterministic capacities of arbitrarily varying quantum channel under maximal error criterion and under the average error criterion are equal. See also [16] and [17] for a classical quantum channel model with a benevolent third channel user.

In all the above mentioned works it is assumed that the jammer knows the coding scheme, but has neither side information about the codeword nor side information about the message of the legal transmitters. In many applications, especially for secure communications, it is too optimistic to assume this. Thus in this paper we want to consider two scenarios, where the jammer has side information: In the first one the jammer knows both coding scheme and input codeword. In the second one the jammer knows additionally the message (cf. Figure 1 and 2). The jammer can make use of this knowledge in each scenario to advance his attacking strategy. We require that information transmission can be guaranteed even in the worst case, when the jammer chooses the most advantageous attacking strategy according to his knowledge. For classical
arbitrarily varying channels this was first considered by [18]. In this paper we extend this result to arbitrarily varying classical-quantum channels, where we use techniques different to these used in [18] (cf. Section IV). In this work we consider for both scenarios the random correlated capacities under average and maximal error criteria. Detailed descriptions for both scenarios are given in Section II. In Section III the message transmission capacities for both scenarios and both error criteria are completely characterized. In Section IV we deliver sketches of proofs for the capacities results for both scenarios and both error criteria. A vanishing rate of the key is sufficient for our codes since the resource we use here is only of polynomial size of the code length (cf. Remark III.2, and also [12] and [10] for a discussion about the difference between various forms of shared randomness).

II. PROBLEM FORMULATION

Basic definitions

For a finite-dimensional complex Hilbert space $\mathcal{H}$, we denote the space of density operators on $\mathcal{H}$ by $\mathcal{S}(\mathcal{H})$. In the following, we introduce the formal definitions of our channel models and our channel code concepts.

Definition II.1. A classical-quantum channel is a mapping $W : \mathcal{X} \rightarrow \mathcal{S}(\mathcal{H})$, specified by a set of quantum states $\{\rho(x), x \in \mathcal{X}\} \subset \mathcal{S}(\mathcal{H})$, indexed by "input letters" $x$ in a finite set $\mathcal{X}$. $\mathcal{X}$ and $\mathcal{H}$ are called input alphabet and output space respectively. We define the $n$-th extension of classical-quantum channel $W$ as follows. The channel outputs a quantum state $\rho^{\otimes n}(x) := \rho(x_1) \otimes \rho(x_2) \otimes \ldots \otimes \rho(x_n)$, in the $n$th tensor power $\mathcal{H}^{\otimes n}$ of the output space $\mathcal{H}$, when an input codeword $x = (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n$ of length $n$ is input into the channel.

For a quantum state $\rho \in \mathcal{S}(\mathcal{H})$ we denote the von Neumann entropy of $\rho$ by $S(\rho) = -\text{tr}(\rho \log \rho)$. For a classical-quantum channel $W$ and a probability distribution $P$ on $\mathcal{X}$, the Holevo quantity $\chi$ is defined as $\chi(P; W) = S(\sum_x P(x)\rho(x)) - \sum_x \rho(x)S(\rho(x))$.

Definition II.2. A classical quantum arbitrarily varying channel (CQAVC) $W$ is specified by a set $\{\{\rho(x, s), x \in \mathcal{X}\}, s \in \mathcal{S}\}$ of classical quantum channels with a common input alphabet $\mathcal{X}$ and output space $\mathcal{H}$, which are indexed by elements $s$ in a finite set $\mathcal{S}$. Elements $s \in \mathcal{S}$ usually are called the states of the channel. $W$ outputs a quantum state $\rho^{\otimes n}(x, s) := \rho(x_1, s_1) \otimes \rho(x_2, s_2) \otimes \ldots \otimes \rho(x_n, s_n)$, if an input codeword $x = (x_1, x_2, \ldots, x_n)$ is input into the channel, and the channel is governed by a state sequence $s = (s_1, s_2, \ldots, s_n)$, while the state varies from symbol to symbol in an arbitrary manner.

We assume that the channel state $s$ is under control of the jammer. Without loss of generality we also assume that the jammer always chooses the most advantageous attacking strategy according to his knowledge.

Definition II.3. A code $\gamma := \{U, \{D(i), i \in \mathcal{I}\}\}$ of length $n$ for a classical quantum channel consists of its codebook $U$ and decoding measurement $\{D(i), i \in \mathcal{I}\}$, where the codebook $U := \{u(i), i \in \mathcal{I}\}$ is a subset of input alphabet $\mathcal{X}^n$ indexed by messages $i$ in the message set $\mathcal{I}$, and the decoding measurement $\{D(i), i \in \mathcal{I}\}$ is a quantum measurement in the output space $\mathcal{H}^{\otimes n}$ that is, $D(i) \geq 0$ for all $i \in \mathcal{I}$ and $\sum_{i \in \mathcal{I}} D(i) = I_{\mathcal{H}}$.

Definition II.4. A random correlated code $\Gamma$ for a CQAVC $W$ is uniformly distributed random variable taking values in a set of codes $\{(U(k), \{D(j, k), j \in J\}), k \in K\}$ with a message set $J$, where $U(k) := \{u(j, k), j \in J\}$ and $\{D(j, k), j \in J\}$ are the codebook and decoding measurement of the $k$th code in the set respectively.

Remark II.5. Usually a random correlated code is defined as any random variable taking values in a set of codes. Here we restrict ourselves to uniformly distributed random variables, since it is sufficiently for our purpose (cf. [21]).

We now distinguish two scenarios depending on the jammer’s knowledge (cf. Figure 1 and 2). We consider for each scenario both average and maximum error criteria.

Scenario 1

In this scenario the jammer knows the input codeword, but not the message.

Definition II.6. By assuming that the random message $J$ is uniformly distributed, the state sequence chosen by jammer is a function $\theta : \mathcal{X}^n \rightarrow \mathcal{S}^n$. We define the average probability of error in scenario 1 by

$$p_a(\Gamma) = \max_{\theta} \mathbb{E}_{\{u(j, K), (\mathbb{I}_{\mathcal{H}} - D(J, K))\}} \left[ \frac{1}{|J|} \sum_{j \in J} \sum_{k \in K} \left| \text{tr}[\rho^{\otimes n}(u(i, k), (\mathbb{I}_{\mathcal{H}} - D(j, k))] \right| \right].$$

The maximum probability of error is defined as

$$p_m(\Gamma) = \max_{\theta} \max_{j \in J} \mathbb{E}_{\{u(j, K), (\mathbb{I}_{\mathcal{H}} - D(j, K))\}} \left[ \frac{1}{|K|} \sum_{s \in \mathcal{S}^n} \text{tr}[\rho^{\otimes n}(u(j, K), s)] \right].$$

Definition II.7. A non-negative number $R$ is an achievable rate for the arbitrarily varying classical-quantum channel.
There is an random correlated code under the average error criterion and under the maximal error criterion if for every \( \delta > 0 \) and \( \epsilon > 0 \), if \( n \) is sufficiently large, there is an random correlated code \( \Gamma \) of length \( n \) such that
\[
\frac{\log |\mathcal{J}|}{n} > R - \delta, \quad \text{and} \quad p_n(\Gamma) < \epsilon \quad \text{and} \quad p_m(\Gamma) < \epsilon, \quad \text{respectively.}
\]
The supremum on achievable rate under random correlated coding of \( \mathcal{W} \) under the average error criterion and under the maximal error criterion in scenario 1, denoted \( C^* (\mathcal{W}) \) and \( C^*_m (\mathcal{W}) \), respectively.

**Scenario 2**

Now the jammer has more benefit and he can choose the state sequence according to both input codeword and message which sender wants to transmit, or a function \( \psi : \cup_{k \in \mathcal{K}} U (k) \times \mathcal{J} \to \mathcal{S}^n \).

**Definition II.8.** We define the average probability of error in scenario 2 by
\[
p_n^* (\Gamma) = \max_{\mathcal{J} \in \mathcal{J}} \frac{1}{|\mathcal{J}|^n} \text{Tr} \left[ \rho^{\otimes n} (u(j, K), \psi(u(j, K), j)) (\mathcal{H} - D(j, K)) \right].
\]
The maximum probability of error in scenario 2 is defined as
\[
p_m^* (\Gamma) = \max_{\mathcal{J} \in \mathcal{J}} \text{Tr} \left[ \rho^{\otimes n} (u(j, K), \psi(u(j, K), j)) (\mathcal{H} - D(j, K)) \right].
\]

**Definition II.9.** A non-negative number \( R \) is an achievable rate for the arbitrarily varying classical-quantum channel \( \mathcal{W} \) under random correlated coding in scenario 2 under the average error criterion and under the maximal error criterion if for every \( \delta > 0 \) and \( \epsilon > 0 \), if \( n \) is sufficiently large, there is an random correlated code \( \Gamma \) of length \( n \) such that
\[
\frac{\log |\mathcal{J}|}{n} > R - \delta, \quad \text{and} \quad p_n^* (\Gamma) < \epsilon \quad \text{and} \quad p_m^* (\Gamma) < \epsilon, \quad \text{respectively.}
\]
The supremum on achievable rate under random correlated coding of \( \mathcal{W} \) under the average error criterion and under the maximal error criterion in scenario 2 is called the random correlated capacity of \( \mathcal{W} \) under the average error criterion and under the maximal error criterion in scenario 2, denoted by \( C^* (\mathcal{W}) \) and \( C^*_m (\mathcal{W}) \), respectively.

It is easy to show that \( C^* (\mathcal{W}) = C^*_m (\mathcal{W}) \), which is also equal to \( C^*_m (\mathcal{W}) \), thus
\[
C^* (\mathcal{W}) \geq C^*_m (\mathcal{W}) = C^* (\mathcal{W}) = C^*_m (\mathcal{W}).
\]

**III. MAIN RESULTS**

For a given CQAVC \( \mathcal{W} = \{\{\rho(x, s), x \in \mathcal{X}\}, s \in \mathcal{S}\} \) with set of state \( \mathcal{S} \), let \( \mathcal{W} := \{\{\tilde{\rho}(x) := \sum_s Q(s|x) \rho(x, s), x \in \mathcal{X}\}, \text{ for all } Q : \mathcal{X} \to \mathcal{S}\} \).

**Theorem III.1.** (Direct Coding Theorem for Scenario 1)

Given a CQAVC \( \mathcal{W} = \{\{\rho(x, s), x \in \mathcal{X}\}, s \in \mathcal{S}\} \) and a type \( P_X \), for all \( \epsilon > 0 \), and \( \lambda > 0 \), there is a \( \delta > 0 \), such that for all sufficiently large \( n \), there exists a code \( \Gamma \) of length \( n \) with a rate larger than \( \min_{\tilde{\rho}(\cdot) \in \tilde{\mathcal{W}}} \chi (P_X, \tilde{\rho}(\cdot)) - \epsilon \), average probability of error in scenario 1 smaller than \( \lambda \), and size of the random correlated code smaller then \( \binom{2^{2n}}{2^n} \). Moreover codewords of codebooks in support set of the random correlated code \( \Gamma \) are all in \( T^n_X \).

**Remark III.2.** In particular, there is a constant \( a > 0 \) such that for any sequence of positive real numbers \( \{\lambda_n\} \), with \( \lim_{n \to \infty} \lambda_n = 0 \), there exists a sequence of random correlated codes with a rate larger than \( \min_{\tilde{\rho}(\cdot) \in \tilde{\mathcal{W}}} \chi (P_X, \tilde{\rho}(\cdot)) - \epsilon \), average probability of error smaller than \( \lambda_n \) and size upper bounded by \( \frac{a n^2}{X^n} \). This means a vanishing rate of the key is sufficient, i.e., the shared randomness we use is a uniformly distributed random variable on a set which has a size (our key size) asymptotically negligible compared to the length of the codeword.

**Theorem III.3.** (Converse Coding Theorem for Scenario 1)

\[
C^* (\mathcal{W}) \leq \max_{\tilde{\rho}(\cdot) \in \tilde{\mathcal{W}}} \chi (P, \tilde{\rho}(\cdot)).
\]

Now one may concern the same question in scenario 2. This is answered by the following Theorem, which can be proven by modifying the proof of Theorem III.1.

**Theorem III.4.** The same conclusion for scenario 2, as that for scenario 1 in Theorem III.1, holds.

The above three Theorems yield the coding theorem:

**Corollary III.5.**

\[
C^* (\mathcal{W}) = C^*_m (\mathcal{W}) = C^*_m (\mathcal{W}) = C^*_m (\mathcal{W}) = C^* (\mathcal{W}) = C^*_m (\mathcal{W}.
\]

Moreover the both capacity \( C^* (\mathcal{W}) \) and \( C^*_m (\mathcal{W}) \) can be achieved by codes with vanishing key rates.

**IV. SKETCHES OF PROOFS FOR THE MAIN RESULTS**

Although coding for classical arbitrarily varying channels is already a challenging topic with a lot of open problems, coding for CQA VC is even much harder. Due to the non-commutativity of quantum operators, many techniques, concepts and methods of classical information theory, for instance, non-standard decoder and list decoding, may not be extended to quantum information theory. Sarwate used in [18] list decoding to prove the coding theorem for classical arbitrarily varying channels when the jammer knows input codeword. However since how to apply list decoding for quantum channels is still an open problem, the technique for classical channels in [18] can not be extended to CQA VC. We need a different approach for our scenario 1. For the complete proof we suggest our readers to view our full version [9].

**Sketch of the Proof of Theorem III.1:**

If the jammer would have some information about the outcome \( k \) of the random key through the input codeword, to which he has access in scenario 1, he could apply a strategy against the 4th deterministic coding for CQA VC by choosing the worst state sequence to attack the communication, which
Step 1: To derive a useful auxiliary results

For a classical quantum channel \( W \) (a set of classical quantum channels with a common input alphabet \( \mathcal{X} \) and a common output Hilbert space \( \mathcal{H} \) when the channel states do not change with every use as in case of CQAVC, but are stationary over the time) and any input codeword \( x \in T^S_\mathcal{X} \), where \([5]\) and \([14]\) there exits a projection \( P(x) \) in \( \mathcal{H} \) such that for all \( \tilde{\rho}(x) \in W, \nu > 0 \), sufficiently large \( n \) and \( \eta > 0 \)

\[
tr(\tilde{\rho}^\otimes n(x)P(x)) > 1 - 2^{-\eta n}, \quad \text{and} \quad tr((\tilde{\rho}^\otimes nP(x))^\otimes n) < 2^{-n[\log_{1/\nu}(\chi(\tilde{\rho}(\cdot))) - \nu - \xi]},
\]

where \( \tilde{\rho}_X := \sum_{x \in X} P_X(x) \tilde{\rho}(x) \). Moreover, \( P(x) \) keeps invariant when permutation \( \pi \) acts on coordinates of nth tensor power \( \mathcal{H}^n \) of Hilbert space \( \mathcal{H} \).

By (3), (4), and some calculation we show that for a CQAVC \( W = \{ (\tilde{\rho}(\cdot), s) = (\tilde{\rho}(x), s) \in \mathcal{X}, s \in S \} \), and \( x \in T^S \mathcal{X} \), when \( X \) is randomly and uniformly distributed on \( T^S_\mathcal{X} \), there is a projection \( P(x) \) in \( \mathcal{H} \) such that

\[
\mathbb{E}tr(\tilde{\rho}^\otimes n(x,s)P(x)) \leq 2^{-n[\log_{1/\nu}(\chi(\tilde{\rho}(\cdot))) - \nu - \xi]},
\]

for all \( \xi > 0 \) and sufficiently large \( n \).

Step 2: Generation of ground set for the codebooks

For having a vanishing key rate we generate our codebook instead of from a typical set \( T^S_\mathcal{X} \) (as usually for random-encoding) randomly from a ground set \( B \) with a cardinality \( |B| \) “slightly” (polynomially) larger that \( |J| \), our desired size of codebooks \( U(k) \), which is defined below. We first randomly generate \( B = \{ (x, i) \in \mathcal{J} \} \) on \( T^n_\mathcal{X} \) with a cardinality \( |B| > n \log_2 \frac{|J|}{|\mathcal{J}|} \) where \( A_n \geq 2^{-n[\log_{1/\nu}(\chi(\tilde{\rho}(\cdot))) - \nu - \xi]} \). By (5) and the Chernoff bound we show that with a positive probability there is a \( B \) such that \( \sum_{(x,i) \in B} tr(\tilde{\rho}^\otimes n(x,s)P(x(i))) \leq 3A_n |\mathcal{J}| \) for all \( x \in T^n_\mathcal{X} \), \( s \in S^n \).

Step 3: The construction of the code

We choose finite sets \( \mathcal{J} \) and \( K_n \) such that \( |\mathcal{J}| \leq A_n^{-1} \) and \( |K_n| = Poly(n) \). We generate \( K_n \) codebooks \( \{ U(j,k) \in J, k \in K_n \} \), \( k \in K_n \) randomly uniformly and independently from \( \{ (x(i_1), x(i_2), \ldots, x(i_{|J|})) \} \), \( i_j \in \mathcal{J} \), for \( j = 1, 2, \ldots, |J| \), with \( i_j \neq i_{j'} \) for \( j \neq j' \).

By simple calculation we conclude that for all \( i \in \mathcal{J}, s \in S^n \), \( j, j' \in \mathcal{J}, n \) with \( j \neq j' \) and \( k \in K_n \)

\[
\mathbb{E}tr(\tilde{\rho}^\otimes n(U(j,k), s)P(U(j',k))) = \chi(i)] \leq 3A_n |\mathcal{J}| \frac{|\mathcal{J}|}{|K_n|} = 1.
\]

Next we define \( \mathcal{E}(i, s, k; \mu_n) \) as the random event that there exists a \( j \in \mathcal{J} \) such that \( U(j,k) = x(i) \) and

\[
\sum_{j' \in \mathcal{J} \setminus \{j\}} tr(\tilde{\rho}^\otimes n(x(i), s)P(U(j',k))) > \mu_n.
\]

Then from above inequality it is not hard to obtain

\[
Pr\{ \mathcal{E}(i, s, k; \mu_n) \} < \frac{3A_n |\mathcal{J}|}{|\mathcal{J}|},
\]

for all \( i \in \mathcal{J}, s \in S, k \in K_n \) and \( \mu_n \in (0, 1) \).

For all fixed \( i \) and \( s \) we define the random set \( \mathcal{R}_0(i, s) := \{ k, \text{there exists a } j \text{ with } U(j, k) = x(i) \text{ and } \sum_{j' \in \mathcal{J} \setminus \{j\}} tr(\tilde{\rho}^\otimes n(x(i), s)P(U(j',k))) > \mu_n \} \). By (6) and the Chernoff Bound we show that

\[
Pr\{ \cup_{s \in S} \cup_{j \in \mathcal{J}} |\mathcal{R}_0(i, s)\} \leq \frac{3A_n |K_n|^2}{|\mathcal{J}|^2} \leq 1/2
\]

for \( \lambda_n := \frac{A_n |\mathcal{J}|}{\mu_n} \), where \( \mu_n \in (0, 1) \). Similarly for all \( i \in \mathcal{J} \), we let \( \mathcal{R}(i) := \{ k, \text{there exists a } j \in \mathcal{J} \text{ with } U(j, k) = x(i) \} \), we show that

\[
Pr\{ \cup_{i \in \mathcal{J}} |\mathcal{R}(i)| \leq \frac{|K_n|^2}{2|\mathcal{J}|^2} \}
\]

for any realization \( \{ (u(j,k) \} \) of \( \{ U(j,k) \} \) we define \( K(i) := \{ k, \text{there exists a } j \in \mathcal{J} \text{ with } U(j,k) = x(i) \} \) and \( K_0(i, s) := \{ k, \text{there exists a } j \text{ with } U(j,k) = x(i) \text{ and } \sum_{j' \in \mathcal{J} \setminus \{j\}} tr(\tilde{\rho}^\otimes n(x(i), s)P(U(j',k))) > \mu_n \} \). By (7) and (8), \( \{ U(j,k) \in J \}, k \in K_n \) has a realization \( U(k) := \{ u(j,k) \in \mathcal{J}, k \in K_n \}, \) such that for all \( k \in K_n \) and \( i, j \neq i', j \in \mathcal{J} \), and \( s \in S^n \)

\[
[u(j,k) = x(i), u(j',k) = x(i')] \Rightarrow i \neq i', \quad |K(i)| \geq \frac{|K_n|^2}{2|\mathcal{J}|^2}, \quad |K_0(i, s)| \leq \frac{9|K_n|^2|\lambda_n|}{2|\mathcal{J}|^2}.
\]

Now we define the decoding measurement \( \{ D(j,k) = \sum_{j' \in \mathcal{J}} P(U(j',k)) \} \) by setting \( D(j,k) := \sum_{j' \in \mathcal{J}} P(U(j',k)) |\sum_{j' \in \mathcal{J}} P(U(j',k))|^{-1/2} \sum_{j' \in \mathcal{J}} P(U(j',k)) \) for every \( j \) and \( k \) and let the random code \( \Gamma \) be randomly uniformly generated from the set of codes \( \{ \gamma(k), k \in K_n \} \).

Step 4 Error Analysis

By Hayashi-Nagaoka inequality it is sufficient for us to show that

\[
\lim_{n \to \infty} tr(\tilde{\rho}^\otimes n(x,s)P(x)) = 1
\]

for all \( x \) and \( s \) and

\[
\lim_{n \to \infty} 1 \frac{|K_0(i, s)|}{|K(i)|} \sum_j tr(\tilde{\rho}^\otimes n(x(i), s)P(U(j',k))) = 0
\]

for all \( i, s \). To show (13) we split the sum over \( K(i) \) into two parts: \( K_0(i, s) \) and \( K_1(i, s) := K(i) \setminus K_0(i, s) \). By the definition of \( K_0(i, s) \) it is not hard to show that the second part goes to 0 if we let \( \mu_n \to 0 \). To let first part go to 0, we may choose the parameter properly such that \( \frac{|K_0(i, s)|}{|K(i)|} \leq 9\lambda_n \) converges to 0. To proof (12) we first use (3) to show that

\[
\sum_{s \in S^n} P^3_{|\mathcal{J}|}(s'|u(j,k))tr(\tilde{\rho}^\otimes n(u(j,k), s')P(u(j,k)))
\]
Based on (14) and the fact that the value of the Holevo quantity to obtain
\[
\rho_{j,k}(x) = \sum_{s' \in S} P_{S|X}(s'|x) \rho(x, s').
\]
we apply Holevo bound and the subadditivity of von Neumann inequality and the data processing inequality, of the state sequence randomly according to \( Q \) for an arbitrary but fixed conditional distribution \( Q \). By the Fano inequality and the data processing inequality, \( R \) cannot exceed \( \frac{1}{n} \left[ I(X; Y) + \delta(\lambda) \right] \) for a \( \delta(\lambda) > 0 \) with \( \delta(\lambda) \to 0 \) as \( \lambda \to 0 \). We apply Holevo bound and the subadditivity of von Neumann entropy to show that
\[
r R \leq \chi(P_{X}, \rho^\otimes_n) + n \delta(\lambda)
\leq \sum_{t=1}^{n} \chi(P_{X}, \rho_t) + n \delta(\lambda).
\]
Finally we complete the proof by choosing a \( Q \) to minimize the Holevo quantity to obtain
\[
R \leq \frac{1}{n} \sum_{t=1}^{n} \min_{P} \chi(P_{X}, \rho_t) + \delta(\lambda) \leq \max_{P} \min_{\rho} \chi(P, \rho_t) + \delta(\lambda).
\]

**Sketch of the Proof of Theorem III.3:**

Suppose that we are given a random correlated code \( \Gamma \) with rate \( R \). Let \( J \) be the (classical) uniformly distributed random message, \( X \) be the random classical input codeword and \( Y \) be the classical outcome of the decoding measurement.

It is easy to show that a randomizing strategy may not enlarge the probability of error, without loss of generality we may assume that the jammer chooses the \( t \)th component \( s_t \) of the state sequence randomly according to \( Q(s|x) \) for an arbitrary but fixed conditional distribution \( Q \). By the Fano inequality and the data processing inequality, \( R \) cannot exceed \( \frac{1}{n} \left[ I(X; Y) + \delta(\lambda) \right] \) for a \( \delta(\lambda) > 0 \) with \( \delta(\lambda) \to 0 \) as \( \lambda \to 0 \).

\[ nR \leq \chi(P_{X}, \rho^\otimes_n) + n \delta(\lambda) \leq \sum_{t=1}^{n} \chi(P_{X}, \rho_t) + n \delta(\lambda). \]

Finally we complete the proof by choosing a \( Q \) to minimize the Holevo quantity to obtain
\[
R \leq \frac{1}{n} \sum_{t=1}^{n} \min_{P} \chi(P_{X}, \rho_t) + \delta(\lambda) \leq \max_{P} \min_{\rho} \chi(P, \rho_t) + \delta(\lambda).
\]

**Sketch of the Proof of Theorem III.4:**

The proof will be done by modification of the proof of Theorem III.1. We partition \( I_n \) into \( |J| \) subsets \( I_n(j), j \in J \) with equal size. Let \( B(j) = \{ x(i), i \in I_n(j) \} \) for \( j \in J \).

For all \( k \in K_n \) let \( U'(j, k) \) be independently and uniformly generated from \( B(j) \) for every \( j \in J \). Similar to the proof of Theorem III.1, \( U'(j, k), j \in J, k \in K_n \) has a realization \( u'(j, k), j \in J, k \in K_n \) with \( u'(j, k) \in B(j) \) such that for all \( j \in J \) and \( k \in K_n \)
\[
|K'(j)| \geq \frac{|K_n||J|}{2|I_n|}
\]
and
\[
|K'_0(i(x(s)), s)| \leq \frac{9|J_n||K_n|\chi^{'n}_n}{2|I_n|},
\]
where \( K'(i(x(s))) := \{ k, u(j, k) = x(i(x(s))) \} \) and \( K'_0(i(x(s)), s) := \{ k, u(j, k) = x(i(x(s))) \} \) and \( \chi^{'n}_n = \{ i, u(j, k) = x(i(x(s))) \} \) and \( \sum_{j=1}^{n} \rho^\otimes_j(x(i(x(s))), s)|\rho^\otimes_j(x(i(x(s)), s))| > \mu_n \). Now the scenario 1 here, for which we have now constructed a code, is actually scenario 2, too, because at this setting, that the jammer knows the codeword \( u'(j, k) \) implies that he knows the message \( j \) as well.

**References**

[1] R. Ahlswede, Elimination of correlation in random codes for arbitrarily varying channels, Z. Wahrscheinlichkeitstheorie verw. Gebiete, Vol. 44, pp. 159-175, 1978.

[2] R. Ahlswede, The maximal error capacity of arbitrarily varying channels for constant list sizes, IEEE Trans. Inform. Theory, Vol. IT-39, pp. 1416-1417, 1993.

[3] R. Ahlswede, I. Bjelaković, H. Boche, and J. Nötzel, Quantum capacity under adversarial quantum noise: arbitrarily varying quantum channels, Comm. Math. Phys. A, Vol. 317, No. 1, pp. 103-156, 2013.

[4] R. Ahlswede and V. Blinovsky, Classical capacity of classical-quantum arbitrarily varying channels, IEEE Trans. Inform. Theory, Vol. 53, No. 2, pp. 526-533, 2007.

[5] I. Bjelaković and H. Boche, Classical capacities of averaged and compound quantum channels. IEEE Trans. Inform. Theory, Vol. 57, No. 7, pp. 3360-3374, 2009.

[6] I. Bjelaković, H. Boche, G. Janßen, and J. Nötzel, Arbitrarily varying and compound classical-quantum channels and a note on quantum zero-error capacities. Information Theory, Combinatorics, and Search Theory, in Memory of Rudolf Ahlswede, H. Aydinian, F. Cicalese, and C. Deppe eds., LNCS Vol.7777, pp. 247-283, arXiv:1209.6325, 2012.

[7] I. Bjelaković, H. Boche, and J. Sommerfeld, Capacity results for arbitrarily varying wiretap channels, Information Theory, Combinatorics, and Search Theory, in Memory of Rudolf Ahlswede, H. Aydinian, F. Cicalese, and C. Deppe eds., LNCS Vol.7777, pp. 114-129, arXiv:1209.5213, 2012.

[8] D. Blackwell, L. Breiman, and A. J. Thomasian, The capacities of a certain channel classes under random coding, Ann. Math. Statist. Vol. 31, No. 3, pp. 558-567, 1960.

[9] H. Boche, M. Cai, N. Cai, Message transmission over classical quantum channels with a jammer with side information: message transmission capacity and resources, on arXiv, 2018.

[10] H. Boche, M. Cai, C. Deppe, and J. Nötzel, Classical-quantum arbitrarily varying wiretap channel: Classical randomness assisted code and continuity. Quantum Information Processing, Vol. 16, No. 1, pp. 1-48, 2016.

[11] H. Boche, M. Cai, C. Deppe, and J. Nötzel, Classical-quantum arbitrarily varying wiretap channel: secret message transmission under jamming attacks. Journal of Mathematical Physics, Vol. 58, pp. 102203, 2017.

[12] H. Boche and J. Nötzel, Arbitrarily small amounts of correlation for arbitrarily varying quantum channel, J. Math. Phys., Vol. 54, No. 11, pp. 112202, arXiv 1301.6063, 2013.

[13] N. Cai, Localized error correction in projective space, IEEE Trans. Inform Theory, Vol.59, pp. 3282-3294, 2013.

[14] M. Hayashi, Universal coding for classical-quantum channel, Comm. Math. Phys., Vol. 289, No. 3, pp. 1087-1098, 2009.

[15] A. S. Holevo, The capacity of quantum channel with general signal states, IEEE Trans. Inform. Theory, Vol. 44, pp. 269-273, 1998.

[16] S. Karumanchi, S. Mancini, A. Winter, and D. Yang, Quantum channel capacities with passive environment assistance, IEEE Trans. Inf. Theory, Vol. 62, No. 4, pp. 1733-1747, arXiv: 1407.8160v2, 2016.

[17] S. Karumanchi, S. Mancini, A. Winter, and D. Yang, Classical capacities of quantum channels with environment assistance, Problems Inf. Transm., Vol. 52, No. 3, pp. 214-238, arXiv:1602.02036v2, 2016.

[18] A. Sarwate, Robust and adaptive communication under uncertain interference. Technical Report No. UCB/EECS-2008-86, University of California at Berkeley, 2008.

[19] B. Schumacher and M. A. Nielsen, Quantum data processing and error correction, Phys. Rev. A, Vol. 54, pp. 2629, 1996.

[20] B. Schumacher and M. D. Westmoreland, Sending classical information via noisy quantum channels, Phys. Rev., Vol. 56, pp. 131-138, 1997.

[21] M. Wiese, J. Nötz, and H. Boche, The arbitrarily varying wiretap channel—deterministic and correlated random coding capacities under the strong secrecy criterion, IEEE Trans. Inform. Theory, Vol. 62, No. 7, pp. 3844-3862, arXiv:1410.8078, 2016.