TECHNISCHE UNIVERSITÄT MÜNCHEN
Fakultät für Elektrotechnik und Informationstechnik

# Classical-Quantum Channels: Secret Message Transmission under Attacks

**Minglai Cai**

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

**Doktors der Naturwissenschaften (Dr. rer. nat.)**

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Gerhard Kramer
Prüfer der Dissertation:
1. Prof. Dr. Holger Boche
2. Prof. Dr. Michael M. Wolf

Die Dissertation wurde am 19.04.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 29.10.2018 angenommen.

# Abstract

Quantum information theory is a new field that has a strong impact on both quantum computing and the laws of quantum mechanics. It unifies information theory with quantum mechanics, generalizing classical information theory to the quantum world. This dissertation analyzes quantum channels which are subject to multiple attacks at the same time.

The first part of this work describes quantum compound wiretap channels. A quantum compound wiretap channel is defined as a pair of double indexed finite set of density operators. The first family represents the communication link to the legitimate receiver while the output of the latter is under control of the wiretapper. In the model of classical compound quantum wiretapper channel, the receiver uses classical channels while the wiretapper uses classical-quantum channels. In the model of classical quantum compound wiretapper channel, the receiver and the wiretapper both use classical-quantum channels.

In the first part of this dissertation the secrecy capacity of the compound channel with quantum wiretapper and with channel state information at the transmitter and the secrecy capacity of the compound classical-quantum wiretap channel with channel state information at the encoder are determined. A lower bound on the secrecy capacity of compound channel with quantum wiretapper and without channel state information is delivered as well. The set of the states may be finite or infinite.

These results are used to derive a lower bound on the entanglement generating capacity for the compound quantum channels and the entanglement generating capacity of the compound quantum channels with channel state information at the encoder.

The second part of this work is on quantum arbitrarily varying wiretap channels. The model of quantum arbitrarily varying wiretap channel is subject to two attacks at the same time: one passive (eavesdropping), and one active (jamming).

The Ahlswede dichotomy for arbitrarily varying classical-quantum wiretap channels is established. According to this dichotomy, either the deterministic secrecy capacity of the channel is zero or it equals its randomness-assisted secrecy capacity. An example is given in which the deterministic secrecy capacity of an arbitrarily varying classical-quantum wiretap channel is not equal to its randomness-assisted secrecy capacity. Thus both cases of the Ahlswede dichotomy for arbitrarily varying classical-quantum wiretap channels actually occur.

A phenomenon called "super-activation" is a direct consequence of the Ahlswede dichotomy for arbitrarily varying classical-quantum wiretap channels, i.e., two arbitrarily varying classical-quantum wiretap channels, both with zero deterministic secrecy capacity, if used together allow perfect secure transmission. The sufficient and necessary conditions for the continuity and for the occurrence of super-activation are given.

Furthermore, in the second part of this dissertation the secrecy capacity of these channels when the sender and the receiver use various resources is analyzed. It turns out that randomness, common randomness, and correlation are very helpful resources for achieving a positive secrecy capacity. The secrecy capacity under common randomness assisted coding of arbitrarily varying classical-quantum wiretap channels is given. The determination of the capacity formula follows ideas of [16] and [65] in the classical cases. This entails: At first considering a mixed channel model which is compound from the sender to the legitimate receiver and varies arbitrarily from the sender to the eavesdropper. Together with Ahlswede dichotomy, this secrecy capacity formula yields the formula for deterministic secrecy capacity of the arbitrarily vary-

ing classical-quantum wiretap channel and hence a full description of the arbitrarily varying classical-quantum wiretap channels.

An application of this secrecy capacity formula is the conditions under which the secrecy capacity is a continuous function of system parameters. In other words, when small variations in the underlying model change dramatically the effect of the jammer's actions and when not. Sharing resource is very helpful for the channel stability in the sense that it provides continuity of secrecy capacities.

The strong code concept closes the loop for secrecy capacity of arbitrarily varying classical-quantum wiretap channels. The strategy for robustness and secrecy is to build a two-part code word, which consists of a deterministic secure code word and a common randomness assisted secure code word. The first part is used to create the common randomness for the sender and the legal receiver, and the second part is used to transmit the message to the legal receiver. To determinate the capacity formula under strong code concept the secrecy capacity of a mixed channel model for arbitrarily varying wiretap channels is delivered. Here the wiretap channels are quantum channels while the legal transmission channels are classical channels. The strong code concept grants entire security.

# Acknowledgments

# Contents



# 1 Introduction

In the last few years the developments in modern communication systems produced many results in a short amount of time. Quantum communication systems especially, allow us to exploit new possibilities while at the same time imposing fundamental limitations.

In this realm, a particle is described by its state, encompassing all of the information such as position, polarization, spin, or momentum. Laws of quantum mechanics often make fundamental distinctions between quantum information theory and that of the classical. The unit of quantum information is called the "qubit", the quantum analogue of the classical "bit". Unlike a bit, which is either "0" or "1", a qubit can be in a

"superposition", i.e. both states at the same time. This is a fundamental tool in quantum information and computing.

A quantum channel is a communication channel which can transmit quantum information. In general, there are two ways to represent a quantum channel with linear algebraic tools (cf. e.g. Section 3.4), either as a sum of several transformations, or as a single unitary transformation which explicitly includes the unobserved environment.

Quantum channels can transmit both classical and quantum information. We consider the capacity of quantum channels carrying classical information. This is equivalent to considering the capacity of classical-quantum channels, where the classical-quantum channels are quantum channels whose sender's inputs are classical variables. The classical capacity of quantum channels has been determined in [42], [57], and [58].

Quantum information processing systems provide huge theoretical advantages over their classical counterparts. One of the most prominent ones being secure quantum information transmission using quantum key distributions (see [14] and [13] for two well-known examples). Good one-shot results for quantum channels with a wiretapper who is limited in his actions have been obtained.

However, our goal is to have a more general theory for channel robustness and security in quantum information theory, where message transmission is secure against every possible kind of jamming and eavesdropping. Channel robustness and security are two very desirable features for an information processing system, since many modern communication systems are often not perfect, but are vulnerable to jamming and eavesdropping. Furthermore, we are interested in asymptotic behavior of secret communication where we deliver a large volume of messages by many channel uses.

Therefore, we consider a new paradigm for the design of quantum channel systems, which is referred to as *embedded security*. Instead of the standard approach in secret communication, where a successful transmission of messages is ensured before the implementation of a cryptographic protocol, here, we embed protocols with a guaranteed security right from the start into the physical layer, which is the bottom layer of the model of communication systems. This concept covers both secure message transmission and secure key generation.

Since we allow every possible kind of eavesdropping, we use the Holevo $\chi$ quantity, also simply referred to as the Holevo quantity, as our security criterion (cf. (9)). According to [42] and [58] the wiretapper can never obtain more information asymptotically than the Holevo $\chi$ quantity, no matter which strategy the wiretapper uses. Another widely used security criterion is the variational distance between $p^A p^Z$ and $p^{AZ}$. Here $p^{AZ}$ is the joint probability describing the sender's random variable and the wiretapper's random variable. $p^A$ and $p^Z$ are the marginal probabilities describing the sender's random variable and the wiretapper's random variable, respectively. The Holevo $\chi$ quantity using a strong condition for the security criterion (c.f. Remark 2.23) is stronger than the variational distance between $p^A p^Z$ and $p^{AZ}$ in the sense that if the Holevo $\chi$ quantity between $p^A$ and the wiretap channel's output (using strong condition) goes to zero, then the variational distance between $p^A p^Z$ and $p^{AZ}$ goes to zero as well (cf. [35] and [22]).

This work concentrates on the secure message transmission of two quantum channel models: the model of compound wiretap channel and the model of arbitrarily varying wiretap channel.

Our first goal is to analyze information transmission over a set of indexed channels, which is called a compound channel. The indices are referred to as channel states.

Only one channel in this set is actually used for the information transmission, but the users cannot control which channel in the set will be used.

The compound channel describes channel uncertainty. We envision a communication scenario as follows. The sender and the receiver use a fiber-optic cable for message transmission. Here we assume that the transmitters can, using powerful shielding, gather perfect protection against any environment's interference. However due to measurement inaccuracy they do not have the full information about the actual parameters of the fiber-optic cable. The only knowledge that can be assumed is that these parameters take place in a known set. The task for the transmitters is to build a robust code which works despite the measurement inaccuracies. The capacity of the classical compound channel was determined in [19].

A quantum channel with channel uncertainty is called a compound quantum channel. The classical capacity of the compound quantum channel was determined in [15], [40], and [51].

In the model of a wiretap channel, we consider communication with security. This was first introduced in [68] for classical channels (in this work we consider a security criterion which is stronger than the security criterion used in[68], cf. Remark 2.23). The relation between the different security criteria is discussed, e.g. in [22] with some generality and in [65] with respect to arbitrarily varying channels.

A quantum wiretap channel is described by a map $N$ which maps the set of density operators on a system $G^{\mathfrak{A}}$ to the set of density operators on a composite system $G^{\mathfrak{B}\mathfrak{Z}}$. Here, $G^{\mathfrak{A}}$ is the the system of the sender, $G^{\mathfrak{B}}$ is the system observed by the legal receiver, and $G^{\mathfrak{Z}}$ is the system observed by the wiretapper. A classical-quantum channel with an eavesdropper is called a classical-quantum wiretap channel. The secrecy capacity of classical-quantum wiretap channel has been determined in [34] and [31].

A compound channel with an eavesdropper is called a compound wiretap channel. The transmitters have to solve two main problems. First, the message (a secret key or a secure message) has to be encoded robustly, i.e. such that the legal receiver can be decoded the message correctly despite channel uncertainty. Secondly, the message has to be encoded in such a way that the wiretapper's knowledge of the transmitted classical message can be kept arbitrarily small.

We define a compound wiretap channel as a family of pairs of channels $\{(\mathsf{W}_t, \mathsf{V}_t) : t = 1, \ldots, T\}$ with a common input alphabet and possibly different output alphabets, connecting a sender with two receivers, a legal one and a wiretapper, where $t$ stands for the state of the channel pair $(\mathsf{W}_t, \mathsf{V}_t)$. The legitimate receiver accesses the output of the first channel $\mathsf{W}_t$ in the pair $(\mathsf{W}_t, \mathsf{V}_t)$, while the wiretapper observes the output of the second part $\mathsf{V}_t$ in the pair $(\mathsf{W}_t, \mathsf{V}_t)$, where a state $t$ governs the channel. A code for the channel conveys information to the legal receiver such that the wiretapper's knowledge of the transmitted information can be kept arbitrarily small. This is a generalization of Wyner's classical wiretap channel (cf. [68]) to a case of multiple channel states. In [68], the author required that the wiretapper cannot detect the message using a weak security criterion (cf. Remark 2.23). For the achievable secrecy rate, we use the worst-case interpretation, i.e. we consider that the general secrecy rate is upper bounded by the secrecy rate when the destination has the worst channel state.

We deal with two communication scenarios. In the first one, only the sender is informed about the index $t$, or in other words, has CSI, where CSI is an abbreviation for "channel state information". In the second scenario, neither sender nor receiver has any information about that index at all.

The classical compound wiretap channels were introduced in [47]. A lower bound on the classical secrecy capacity was obtained under the condition that the sender does

not have CSI. In [47], the authors required that the receiver's average error goes to zero and that the wiretapper is not able to detect the message, with respect to the same security criterion as in [68]. The result of [47] was improved in [18] by using a stronger condition for the limit of the legitimate receiver's error, i.e. the maximal error should go to zero, as well as a stronger condition for the security criterion (c.f. Remark 2.23). Furthermore, the secrecy capacity was determined for the case in which the sender had CSI.

In this work, we analyze two variants of compound wiretap quantum channels. The first variant is called the classical compound channel with quantum wiretapper. In this channel model, we assume that the wiretap channels are quantum channels, while the legal transmission channels are classical channels. The second variant is called the compound classical-quantum wiretap channel. In this channel model, we assume that both families of channels are quantum channels, while the sender transmits classical information.

Our results are summarized as follows. Under the condition that the sender has knowledge about the CSI, the secrecy capacity for these two channel models is derived. Additionally, when the sender does not have any knowledge about the CSI, we determine the secrecy capacity of the compound classical-quantum wiretap channel, and give a lower bound for the secrecy capacity of the classical compound channel with quantum wiretapper.

The determination of entanglement generating capacity is an application of the results on quantum channels. This capacity describes the maximal amount of entanglement that we can generate or transmit over a given quantum channel. For the sender and the receiver, the objective is to share a nearly maximally entangled state on a $(2^{nR} \times 2^{nR})$-dimensional Hilbert space by using $n$ instances of the compound quantum channel, where $n$ is a large number. In [11] it is shown how to send a large amount of entangled quantum states through a noisy quantum channel such that the channel does not modify the entanglement. However, the study of entanglement generation allows a noisy quantum channel to modify the entanglement, as long as the transmitters can use a recovery algorithm to restore the entanglement. The entanglement generating capacity of a quantum channel has been determined in [34] and [49]. The entanglement generating capacities of a compound quantum channel with and without CSI have been determined in [17].

In our paper we derive a lower bound on the entanglement generating capacity of the compound quantum channel by using an alternative technique to the method in [17] (cf. Section 3.4). Furthermore, we derive the entanglement generating capacity of the compound quantum channel with CSI at the encoder using an alternative technique.

Our next goal is to investigate communication that takes place over a quantum channel which is, in addition to noise from the environment, subject to the action of a jammer who actively manipulates the states. The messages should also be kept secret from an eavesdropper.

In the model of an arbitrarily varying channel, we consider a channel which is not stationary, but can change with every use of the channel. We interpret it as the attack of a jammer. It is understood that the sender and the receiver have to select their coding scheme first. After that, the jammer makes his choice of the channel state to sabotage the message transmission. However, due to the physical properties, we consider that the jammer's changes only take place in a known set.

The arbitrarily varying channel was first introduced in [20]. [2] showed a surprising result which is now known as the Ahlswede Dichotomy: The capacity of an arbitrarily

varying channel is either zero or equal to its shared randomness assisted capacity. After that discovery, the question of when exactly the deterministic capacity is positive has remained open. In [36] a sufficient condition for that has been given, and in [33] it is proven that this condition is also necessary. In [1], it also has been shown that the capacity of certain arbitrarily varying channels can be equated to the zero-error capacity of related discrete memoryless channels. The idea to show the Ahlswede Dichotomy is to build a two-part code word, the first part is used to create the common randomness for the sender and the legal receiver, the second is used to transmit the message to the legal receiver. The Ahlswede Dichotomy demonstrates the importance of shared randomness for communication in a very clear form.

An arbitrarily varying classical-quantum channel is a classical-quantum channel under control of a malicious third party, where the jammer may change his input in every channel use and is not restricted to use a repetitive probabilistic strategy. In [6] the capacity of arbitrarily varying classical-quantum channels is analyzed. A lower bound of the capacity has been given. An alternative proof of [6]'s result and a proof of the strong converse are given in [16]. In [5] the Ahlswede Dichotomy for the arbitrarily varying classical-quantum channels is established, and a sufficient and necessary condition for the zero deterministic capacity is given. In [28] a simplification of this condition for the arbitrarily varying classical-quantum channels is given. A classical-quantum channel with an eavesdropper is called a classical-quantum wiretap channel. The secrecy capacity of classical-quantum wiretap channel has been determined in [34] and [31].

In this work, we bring two aspects together, namely we investigate the transmission of secret messages from a sending to a receiving party. The messages ought to be kept secret from an eavesdropper. Communication takes place over a quantum channel which is, in addition to noise from the environment, subjected to the action of a jammer which actively manipulates the states.

In the model of an arbitrarily varying wiretap channel, we consider transmission with both a jammer and an eavesdropper. The secrecy capacity of classical arbitrarily varying wiretap channel has been analyzed in [18]. A lower bound of the randomness-assisted secrecy capacity has been given.

In this work we analyze classical-quantum channel with both a jammer and an eavesdropper, which is referred to as an arbitrarily varying classical-quantum wiretap channel. It is defined as a family of pairs of indexed channels $\{(W_t, V_t) : t = 1, \ldots, T\}$ with a common input alphabet and possible different output systems, connecting a sender with two receivers, a legal one and a wiretapper, where $t$ determines the channel state of the channel pair. The legitimate receiver and the wiretapper accesses the output of the first part of the pair, i.e. the first channel $W_t$ in the pair, and the output of the second part, i.e. $V_t$, respectively. A channel state $t$, which varies from symbol to symbol in an arbitrary manner, governs both the legal receiver's channel and the wiretap channel. This is a generalization of compound classical-quantum wiretap channels when the channel states are not stationary, but can change over time.

We are interested in the role that different forms of shared randomness play for the arbitrarily varying classical-quantum wiretap channel. To this end we will distinguish between two kinds of shared randomness: *common randomness* and *correlation*. As mentioned already, the former has been used as a method of proof in [2] and much of the follow-up works for the determination of the random capacity.

Randomness is the strongest resource available in communication tasks: It requires a perfect copy of the outcome of a random experiment, and thus we should assume an additional perfect channel. Moreover, the outcomes of said experiment have to be

distributed uniformly. For this reason the authors of the works [7] and [28] investigate a variant where the randomness is replaced by correlation, which in some sense completely opposites the randomness: It is the weakest resource available in communication tasks.

Assume that a bipartite source, modeled by an i.i.d. random variable $(X, Y)$ with values in a finite product set $\mathbf{X} \times \mathbf{Y}$, is observed by the sender and (legal) receiver. The sender has access to the random variable $X$, and the receiver to $Y$. We call $(X, Y)$ correlated shared randomness whenever the mutual information between $X$ and $Y$ satisfies $I(X;Y) > 0$.

The authors of the work [28] also put emphasis on the quantification of the differences between correlation and common randomness. It can be shown that common randomness is a stronger resource than correlation in the following sense: An example has been given when not even a finite amount of common randomness can be extracted from a given correlation. On the contrary, a sufficiently large amount of common randomness allows the sender and receiver to asymptotically simulate the statistics of any correlation.

It has been shown in [7] that correlated shared randomness is a helpful resource for information transmission through an arbitrarily varying classical channel: The use of mere correlation does allow one to transmit messages at any rate that would be achievable using any form of shared randomness. The capacity of an arbitrarily varying quantum channel assisted by correlated shared randomness as resource has been discussed in [28], where equivalent results were found. In this work, we extend the concept of correlation-assisted coding to the arbitrarily varying classical-quantum wiretap channel.

In this work, we establish the Ahlswede dichotomy for the arbitrarily varying classical-quantum wiretap channels, i.e., either the deterministic secrecy capacity of an arbitrarily varying classical-quantum wiretap channel is zero, or it equals its randomness-assisted secrecy capacity. Furthermorere, we deliver a capacity formula for secure information transmission through an arbitrarily varying classical-quantum wiretap channel using correlation as a resource. Together with the Ahlswede dichotomy for the arbitrarily varying classical-quantum wiretap channels, it yields a formula for deterministic secrecy capacity of the arbitrarily varying classical-quantum wiretap channel.

This formula for deterministic secrecy capacity is shown, as aforementioned, by building two-part deterministic codes. A code word of this code concept is a composition of a public code word to synchronize the second part and a common randomness assisted code word to transmit the message. We only require security for the last part. However this code concept still leaves something to be desired because we have to reduce the generality of the code concept when we explicitly allowe a small part of the code word to be non-secure. As we will show in Corollary 4.15, when the jammer has access to the first part, it will be rendered completely useless. Thus the code concept only works when the jammer is limited in his action, e.g. we have to assume that the eavesdropper cannot send messages towards the jammer. Thus we consider in this work additionally a more general code concept when we construct a code in such a way that every part of it is secure. We show that when the legal channel is not symmetrizabel, the sender can send a small number of secure transmissions which push the secure capacity to the maximally attainable value. Thus, entire security is granted. We call it the strong code concept. This completes our analysis of the arbitrarily varying classical-quantum wiretap channel.

The code concept with weak criterion can be, nevertheless, useful when a small number of public messages are desired, e.g. when the receiver uses it to estimate the

channels.

We also put a focus on the analysis of different forms of shared randomness and their impact on the robustness and security. As a direct consequence of our capacity formula, we show in this paper that a sharing resource is very helpful for the channel stability in the sense that it provides continuity of secrecy capacities.

As an application of our results, we turn to the question: when the secrecy capacity is a continuous function of the system parameters? The analysis of the continuity of capacities of quantum channels was raised from the question of whether small changes in the channel system are able to cause dramatic losses in the performance. The continuity of the message- and entanglement transmission capacity of a stationary memoryless quantum channel was listed as an open problem in [69] and has been solved in [46]. Especially considering channels with active jamming faces new difficulties. The reason is that the capacity in this case is, in general, not specified by entropic quantities. In [29] it has been shown under which conditions the message transmission capacity of an arbitrarily varying quantum channels is continuous. The condition for continuity of message transmission capacity of a classical arbitrarily varying wiretap channel has been given in [65].

We also present a new discovery for the arbitrarily varying classical-quantum wiretap channels which is a consequence of the Ahlswede dichotomy for the arbitrarily varying classical-quantum wiretap channels. This phenomenon is called "superactivation", i.e., two arbitrarily varying classical-quantum wiretap channels, both with zero deterministic secrecy capacity, if used together allow perfect secure transmission.

# 2 Preliminaries

## 2.1 Basic Definitions and Communication Scenarios

For a finite set $\mathbf{A}$ we denote the set of probability distributions on $\mathbf{A}$ by $P(\mathbf{A})$. Let $\rho_1$ and $\rho_2$ be Hermitian operators on a finite-dimensional complex Hilbert space $G$. We say $\rho_1 \geq \rho_2$ and $\rho_2 \leq \rho_1$ if $\rho_1 - \rho_2$ is positive-semidefinite. For a finite-dimensional complex Hilbert space $G$, we denote the (convex) space of density operators on $G$ by

$$\mathcal{S}(G) := \{\rho \in \mathcal{L}(G) : \rho \text{ is Hermitian}, \rho \geq 0_G \, , \mathrm{tr}(\rho) = 1\} \, ,$$

where $\mathcal{L}(G)$ is the set of linear operators on $G$, and $0_G$ is the null matrix on $G$. Note that any operator in $\mathcal{S}(G)$ is bounded.

For finite sets $\mathbf{A}$ and $\mathbf{B}$, we define a (discrete) **classical channel** $\mathsf{V}$*:* $\mathbf{A} \to P(\mathbf{B})$, $P(\mathbf{A}) \ni x \to \mathsf{V}(x) \in P(\mathbf{B})$ *to be a system characterized by a probability transition matrix* $\mathsf{V}(\cdot|\cdot)$*. For* $x \in \mathbf{A}$ *and* $y \in \mathbf{B}$*,* $\mathsf{V}(y|x)$ *expresses the probability of the output symbol* $y$ *when we send the symbol* $x$ *through the channel. The channel is said to be memoryless if the probability distribution of the output depends only on the input at that time and is conditionally independent of previous channel inputs and outputs. Further, we can extend this definition when we define a classical channel to a map* $\mathsf{V}$*:* $P(\mathbf{A}) \to P(\mathbf{B})$ *by denoting* $\mathsf{V}(y|p) := \sum_{x \in \mathbf{A}} p(x)\mathsf{V}(y|x)$*.*

*Let* $n \in \mathbb{N}$*. We define the* $n$*-th memoryless extension of the stochastic matrix* $\mathsf{V}$ *by* $\mathsf{V}^n$*, i.e. for* $x^n = (x_1, \ldots, x_n) \in A^n$ *and* $y^n = (y_1, \ldots, y_n) \in B^n$*,* $\mathsf{V}^n(y^n|x^n) = \prod_{i=1}^n \mathsf{V}(y_i|x_i)$*.*

*For finite-dimensional complex Hilbert spaces* $G$ *and* $G'$ *a* **quantum channel** $N$*:* $\mathcal{S}(G) \to \mathcal{S}(G')$*,* $\mathcal{S}(G) \ni \rho \to N(\rho) \in \mathcal{S}(G')$ *is represented by a completely positive*

*trace-preserving map which accepts input quantum states in $\mathcal{S}(G)$ and produces output quantum states in $\mathcal{S}(G')$.*

*If the sender wants to transmit a classical message of a finite set $A$ to the receiver using a quantum channel $N$, his encoding procedure will include a classical-to-quantum encoder to prepare a quantum message state $\rho \in \mathcal{S}(G)$ suitable as an input for the channel. If the sender's encoding is restricted to transmit an indexed finite set of quantum states $\{\rho_x : x \in \mathbf{A}\} \subset \mathcal{S}(G)$, then we can consider the choice of the signal quantum states $\rho_x$ as a component of the channel. Thus, we obtain a channel $\sigma_x := N(\rho_x)$ with classical inputs $x \in \mathbf{A}$ and quantum outputs, which we call a classical-quantum channel. This is a map $\mathbf{N}$: $\mathbf{A} \to \mathcal{S}(G')$, $\mathbf{A} \ni x \to \mathbf{N}(x) \in \mathcal{S}(G')$ which is represented by the set of $|\mathbf{A}|$ possible output quantum states $\{\sigma_x = \mathbf{N}(x) := N(\rho_x) : x \in \mathbf{A}\} \subset \mathcal{S}(G')$, meaning that each classical input of $x \in \mathbf{A}$ leads to a distinct quantum output $\sigma_x \in \mathcal{S}(G')$. In view of this, we have the following definition.*

*Let $H$ be a finite-dimensional complex Hilbert space. A* **classical-quantum channel** *is a linear map $W : P(\mathbf{A}) \to \mathcal{S}(H)$, $P(\mathbf{A}) \ni P \to W(P) \in \mathcal{S}(H)$. Let $a \in \mathbf{A}$. For a $P_a \in P(\mathbf{A})$, defined by $P_a(a') = \begin{cases} 1 & \text{if } a' = a\ ; \\ 0 & \text{if } a' \neq a \end{cases}$, we write $W(a)$ instead of $W(P_a)$.*

**Remark 2.1.** *In many literature a classical-quantum channel is defined as a map $\mathbf{A} \to \mathcal{S}(H)$, $\mathbf{A} \ni a \to W(a) \in \mathcal{S}(H)$. This is a special case when the input is limited on the set $\{P_a : a \in \mathbf{A}\}$.*

*For any finite set $\mathbf{A}$, any finite-dimensional complex Hilbert space $H$, and $n \in \mathbb{N}$, we define $\mathbf{A}^n := \Big\{(a_1, \cdots, a_n) : a_i \in \mathbf{A}\ \forall i \in \{1, \cdots, n\}\Big\}$, and $H^{\otimes n} := span\Big\{v_1 \otimes \cdots \otimes v_n : v_i \in H\ \forall i \in \{1, \cdots, n\}\Big\}$. We also write $a^n$ for the elements of $\mathbf{A}^n$.*

*Let $n \in \mathbb{N}$. Following [66] we define the $n$-th memoryless extension of the stochastic matrix $\mathsf{V}$ by $\mathsf{V}^n$, i.e. for $x^n = (x_1, \ldots, x_n) \in A^n$ and $y^n = (y_1, \ldots, y_n) \in B^n$, $\mathsf{V}^n(y^n|x^n) = \prod_{i=1}^n \mathsf{V}(y_i|x_i)$.*

*Following [66], we define the $n$-th extension of classical-quantum channel $W$ as follows. Associated with $W$ is the channel map on the n-block $W^{\otimes n}$: $P(\mathbf{A}^n) \to \mathcal{S}(H^{\otimes n})$, such that $W^{\otimes n}(P^n) = W(P_1) \otimes \cdots \otimes W(P_n)$ if $P^n \in P(\mathbf{A}^n)$ can be written as $(P_1, \cdots, P_n)$.*

*Let $\theta := \{1, \cdots, T\}$ be a finite set. Let $\Big\{W_t : t \in \theta\Big\}$ be a set of classical-quantum channels. For $t^n = (t_1, \cdots, t_n)$, $t_i \in \theta$ we define the n-block $W_{t^n}$ such that for $W_{t^n}(P^n) = W_{t_1}(P_1) \otimes \cdots \otimes W_{t_n}(P_n)$ if $P^n \in P(\mathbf{A}^n)$ can be written as $(P_1, \cdots, P_n)$.*

*Let $\mathfrak{P}$ and $\mathfrak{Q}$ be quantum systems, denote the Hilbert space of $\mathfrak{P}$ and $\mathfrak{Q}$ by $H^{\mathfrak{P}}$ and $H^{\mathfrak{Q}}$, respectively. We denote the space of density operators on $H^{\mathfrak{P}}$ and $H^{\mathfrak{Q}}$ by $\mathcal{S}(H^{\mathfrak{P}})$ and $\mathcal{S}(H^{\mathfrak{Q}})$, respectively. A quantum channel $N$: $\mathcal{S}(H^{\mathfrak{P}}) \to \mathcal{S}(H^{\mathfrak{Q}})$, $\mathcal{S}(H^{\mathfrak{P}}) \ni \rho \to N(\rho) \in \mathcal{S}(H^{\mathfrak{Q}})$ is represented by a completely positive trace-preserving map, which accepts input quantum states in $\mathcal{S}(H^{\mathfrak{P}})$ and produces output quantum states in $\mathcal{S}(H^{\mathfrak{Q}})$.*

*Associated with $N$ is the channel maps on the n-block $N^{\otimes n}$: $\mathcal{S}(H^{\mathfrak{P}^{\otimes n}}) \to \mathcal{S}(H^{\mathfrak{Q}^{\otimes n}})$ such that $N^{\otimes n}(\rho^n) = N(\rho_1) \otimes \cdots \otimes N(\rho_n)$ for $\rho^n = \rho_1 \otimes \cdots \otimes \rho_n \in \mathcal{S}(H^{\mathfrak{P}^{\otimes n}})$. For $t^n = (t_1, \cdots, t_n)$, $t_i \in \theta$ we define the $n$-block $N_{t^n}$ such that for $\rho^n = \rho_1 \otimes \cdots \otimes \rho_n \in \mathcal{S}(H^{\mathfrak{P}^{\otimes n}})$ we have $N_{t^n}(\rho^n) = N_{t_1}(\rho_1) \otimes \cdots \otimes N_{t_n}(\rho_n)$.*

*We denote the identity operator on a space $H$ by $\mathrm{id}_H$ and the symmetric group on $\{1, \cdots, n\}$ by $\mathsf{S}_n$.*

*For a probability distribution $P$ on a finite set $\mathbf{A}$ and a positive constant $\delta$, we denote the set of typical sequences by*

$$\mathcal{T}^n_{P,\delta} := \left\{ a^n \in \mathbf{A}^n : \left| \frac{1}{n} N(a' \mid a^n) - P(a') \right| \leq \frac{\delta}{|\mathbf{A}|} \forall a' \in \mathbf{A} \right\} ,$$

*where $N(a' \mid a^n)$ is the number of occurrences of the symbol $a'$ in the sequence $a^n$.*

*For a discrete random variable $X$ on a finite set $\mathbf{A}$ and a discrete random variable $Y$ on a finite set $\mathbf{B}$ we denote the Shannon entropy of $X$ by $H(X) = -\sum_{x \in \mathbf{A}} p(x) \log p(x)$ and the mutual information between $X$ and $Y$ by $I(X;Y) = \sum_{x \in \mathbf{A}} \sum_{y \in \mathbf{B}} p(x,y) \log \left( \frac{p(x,y)}{p(x)p(y)} \right)$. Here $p(x,y)$ is the joint probability distribution function of $X$ and $Y$, and $p(x)$ and $p(y)$ are the marginal probability distribution functions of $X$ and $Y$ respectively, and "$\log$" means logarithm to base 2.*

*For a quantum state $\rho \in \mathcal{S}(H)$ we denote the von Neumann entropy of $\rho$ by*

$$S(\rho) = -\mathrm{tr}(\rho \log \rho) ,$$

*where "$\log$" means logarithm to base 2.*

*Let $\mathfrak{P}$ and $\mathfrak{Q}$ be quantum systems. We denote the Hilbert space of $\mathfrak{P}$ and $\mathfrak{Q}$ by $G^{\mathfrak{P}}$ and $G^{\mathfrak{Q}}$, respectively. Let $\phi^{\mathfrak{P}\mathfrak{Q}}$ be a bipartite quantum state in $\mathcal{S}(G^{\mathfrak{P}\mathfrak{Q}})$. We denote the partial trace over $G^{\mathfrak{P}}$ by*

$$\mathrm{tr}_{\mathfrak{P}}(\phi^{\mathfrak{P}\mathfrak{Q}}) := \sum_l \langle l|_{\mathfrak{P}} \phi^{\mathfrak{P}\mathfrak{Q}} |l\rangle_{\mathfrak{P}} ,$$

*where $\{|l\rangle_{\mathfrak{P}} : l\}$ is an orthonormal basis of $G^{\mathfrak{P}}$.*

*We denote the conditional entropy by*

$$S(\mathfrak{P} \mid \mathfrak{Q})_\phi := S(\phi^{\mathfrak{P}\mathfrak{Q}}) - S(\phi^{\mathfrak{Q}}) .$$

*The quantum mutual information is denoted by*

$$I(\mathfrak{P};\mathfrak{Q})_\phi = S(\phi^{\mathfrak{P}}) + S(\phi^{\mathfrak{Q}}) - S(\phi^{\mathfrak{P}\mathfrak{Q}}) .$$

*Here $\phi^{\mathfrak{Q}} = \mathrm{tr}_{\mathfrak{P}}(\phi^{\mathfrak{P}\mathfrak{Q}})$ and $\phi^{\mathfrak{P}} = \mathrm{tr}_{\mathfrak{Q}}(\phi^{\mathfrak{P}\mathfrak{Q}})$.*

*Let $\mathtt{V}$: $\mathbf{A} \to \mathcal{S}(G)$ be a classical-quantum channel. Following [8], for $P \in P(\mathbf{A})$ the conditional entropy of the channel for $\mathtt{V}$ with input distribution $P$ is denoted by*

$$S(\mathtt{V}|P) := \sum_{x \in \mathbf{A}} P(x) S(\mathtt{V}(x)) .$$

*For quantum states $\rho$ and $\sigma \in \mathcal{S}(G)$, we denote the fidelity of $\rho$ and $\sigma$ by*

$$F(\rho,\sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1^2 ,$$

*where $\|\cdot\|_1$ stands for the trace norm.*

*A quantum state represented by a ensemble $\in \mathcal{S}(G)$ of rank 1 is called a pure state. A quantum state which is not a pure state is called a mixed state.*

*Let $\mathbf{A}$ be a finite set and $G$, $H$ be finite-dimensional complex Hilbert spaces. A purification of a quantum state $\rho \in \mathcal{S}(G)$ in $\mathcal{S}(G \otimes H)$ is a $|\psi\rangle \in G \otimes H$ such that*

$\mathrm{tr}_H(|\psi\rangle\langle\psi|) = \rho$. *We also say* $|\psi\rangle\langle\psi|$ *is a purification of* $\rho$. *Notice that when* $\dim H \geq \dim G$ *there exists for every quantum state in* $\mathcal{S}(G)$ *a purification in* $\mathcal{S}(G \otimes H)$ *(cf. [52]).*

*For a quantum state* $\rho \in \mathcal{S}(G)$ *and a quantum channel* $V$*:* $\mathcal{S}(G) \to \mathcal{S}(H)$ *the coherent information is defined as*

$$I_C(\rho, V) := S(V(\rho)) - S\left((\mathrm{id}_G \otimes V)(|\psi\rangle\langle\psi|)\right) ,$$

*where* $|\psi\rangle\langle\psi|$ *is an arbitrary purification of* $\rho$ *in* $\mathcal{S}(G \otimes G)$.

*Let* $\Phi := \{\rho_x : x \in \mathbf{A}\}$ *be a set of quantum states labeled by elements of* $\mathbf{A}$. *For a probability distribution* $Q$ *on* $\mathbf{A}$, *the Holevo* $\chi$ *quantity is defined as*

$$\chi(Q; \Phi) := S\left(\sum_{x \in \mathbf{A}} Q(x)\rho_x\right) - \sum_{x \in \mathbf{A}} Q(x) S\left(\rho_x\right) .$$

*Note that we can always associate a state* $\rho^{XY} = \sum_x Q(x)|x\rangle\langle x| \otimes \rho_x$ *to* $(Q; \Phi)$ *such that* $\chi(Q; \Phi) = I(X; Y)$ *holds for the quantum mutual information.*

*For a set* $\mathbf{A}$ *and a Hilbert space* $G$ *let* $\mathbf{V}$*:* $\mathbf{A} \to \mathcal{S}(G)$ *be a classical-quantum channel. For a probability distribution* $P$ *on* $\mathbf{A}$ *the Holevo* $\chi$ *quantity of the channel for* $\mathbf{V}$ *with input distribution* $P$ *is defined as*

$$\chi(P; \Phi) := S\left(\mathbf{V}(P)\right) - S\left(\mathbf{V}|P\right) .$$

*Let* $G$ *be a finite-dimensional complex Hilbert space. Let* $n \in \mathbb{N}$ *and* $\alpha > 0$. *We suppose* $\rho \in \mathcal{S}(G)$ *has the spectral decomposition* $\rho = \sum_x P(x)|x\rangle\langle x|$, *its* $\alpha$*-typical subspace is the subspace spanned by* $\left\{|x^n\rangle, x^n \in \mathcal{T}^n_{P,\alpha}\right\}$, *where* $|x^n\rangle := \otimes_{i=1}^n |x_i\rangle$. *The orthogonal subspace projector which projects onto the typical subspace is*

$$\Pi_{\rho,\alpha} = \sum_{x^n \in \mathcal{T}^n_{P,\alpha}} |x^n\rangle\langle x^n| .$$

*Similarly let* $\mathbf{A}$ *be a finite set, and* $G$ *be a finite-dimensional complex Hilbert space. Let* $\mathtt{V}$*:* $\mathbf{A} \to \mathcal{S}(G)$ *be a classical-quantum channel. For* $a \in \mathbf{A}$ *suppose* $\mathtt{V}(a)$ *has the spectral decomposition* $\mathtt{V}(a) = \sum_j V(j|a)|j\rangle\langle j|$ *for a stochastic matrix* $V(\cdot|\cdot)$. *The* $\alpha$*-conditional typical subspace of* $\mathtt{V}$ *for a typical sequence* $a^n$ *is the subspace spanned by* $\left\{\bigotimes_{a \in \mathbf{A}} |j^{\mathtt{I}_a}\rangle, j^{\mathtt{I}_a} \in \mathcal{T}^{\mathtt{I}_a}_{V(\cdot|a),\delta}\right\}$. *Here* $\mathtt{I}_a := \{i \in \{1, \cdots, n\} : a_i = a\}$ *is an indicator set that selects the indices* $i$ *in the sequence* $a^n = (a_1, \cdots, a_n)$ *for which the* $i$*-th symbol* $a_i$ *is equal to* $a \in \mathbf{A}$. *The subspace is often referred to as the* $\alpha$*-conditional typical subspace of the state* $\mathtt{V}^{\otimes n}(a^n)$. *The orthogonal subspace projector which projects onto it is defined as*

$$\Pi_{\mathtt{V},\alpha}(a^n) = \bigotimes_{a \in \mathbf{A}} \sum_{j^{\mathtt{I}_a} \in \mathcal{T}^{\mathtt{I}_a}_{\mathtt{V}(\cdot|a^n),\alpha}} |j^{\mathtt{I}_a}\rangle\langle j^{\mathtt{I}_a}| .$$

*The typical subspace has following properties:*

*For* $\sigma \in \mathcal{S}(G^{\otimes n})$ *and* $\alpha > 0$ *there are positive constants* $\beta(\alpha)$, $\gamma(\alpha)$, *and* $\delta(\alpha)$, *depending on* $\alpha$ *such that*

$$\mathrm{tr}\left(\sigma \Pi_{\sigma,\alpha}\right) > 1 - 2^{-n\beta(\alpha)} , \tag{1}$$

$$2^{n(S(\sigma)-\delta(\alpha))} \leq \mathrm{tr}\left(\Pi_{\sigma,\alpha}\right) \leq 2^{n(S(\sigma)+\delta(\alpha))}\,, \tag{2}$$

$$2^{-n(S(\sigma)+\gamma(\alpha))}\Pi_{\sigma,\alpha} \leq \Pi_{\sigma,\alpha}\sigma\Pi_{\sigma,\alpha} \leq 2^{-n(S(\sigma)-\gamma(\alpha))}\Pi_{\sigma,\alpha}\,. \tag{3}$$

*For $a^n \in \mathcal{T}^n_{P,\alpha}$ there are positive constants $\beta(\alpha)'$, $\gamma(\alpha)'$, and $\delta(\alpha)'$, depending on $\alpha$ such that*

$$\mathrm{tr}\left(\mathtt{V}^{\otimes n}(a^n)\Pi_{\mathtt{V},\alpha}(a^n)\right) > 1 - 2^{-n\beta(\alpha)'}\,, \tag{4}$$

$$\begin{aligned} &2^{-n(S(\mathtt{V}|P)+\gamma(\alpha)')}\Pi_{\mathtt{V},\alpha}(a^n) \leq \Pi_{\mathtt{V},\alpha}(a^n)\mathtt{V}^{\otimes n}(a^n)\Pi_{\mathtt{V},\alpha}(a^n) \\ &\leq 2^{-n(S(\mathtt{V}|P)-\gamma(\alpha)')}\Pi_{\mathtt{V},\alpha}(a^n)\,, \end{aligned} \tag{5}$$

$$2^{n(S(\mathtt{V}|P)-\delta(\alpha)')} \leq \mathrm{tr}\left(\Pi_{\mathtt{V},\alpha}(a^n)\right) \leq 2^{n(S(\mathtt{V}|P)+\delta(\alpha)')}\,. \tag{6}$$

*For the classical-quantum channel $\mathtt{V} : P(\mathbf{A}) \to \mathcal{S}(G)$ and a probability distribution $P$ on $\mathbf{A}$ we define a quantum state $P\mathtt{V} := \mathtt{V}(P)$ on $\mathcal{S}(G)$. For $\alpha > 0$ we define an orthogonal subspace projector $\Pi_{P\mathtt{V},\alpha}$ fulfilling (1), (2), and (3). Let $x^n \in \mathcal{T}^n_{P,\alpha}$. For $\Pi_{P\mathtt{V},\alpha}$ there is a positive constant $\beta(\alpha)''$ such that following inequality holds:*

$$\mathrm{tr}\left(\mathtt{V}^{\otimes n}(x^n)\cdot\Pi_{P\mathtt{V},\alpha}\right) \geq 1 - 2^{-n\beta(\alpha)''}\,. \tag{7}$$

*We give here a sketch of the proof. For a detailed proof please see [66].*

*Proof.* (1) holds because $\mathrm{tr}\left(\sigma\Pi_{\sigma,\alpha}\right) = \mathrm{tr}\left(\Pi_{\sigma,\alpha}\sigma\Pi_{\sigma,\alpha}\right) = P^n(\mathcal{T}^n_{P,\alpha})$. (2) holds because $\mathrm{tr}\left(\Pi_{\sigma,\alpha}\right) = \left|\mathcal{T}^n_{P,\alpha}\right|$. (3) holds because $2^{-n(S(\sigma)+\gamma(\alpha))} \leq P^n(x^n) \leq 2^{-n(S(\sigma)-\gamma(\alpha))}$ for $x \in \mathcal{T}^n_{P,\alpha}$ and a positive $\gamma(\alpha)$. (4), (5), and (6) can be obtained in a similar way. (7) follows from the permutation-invariance of $\Pi_{P\mathtt{V},\alpha}$.

**Definition 2.2.** *Let $\mathfrak{P}$ and $\mathfrak{Q}$ be quantum systems. We denote the Hilbert space of $\mathfrak{P}$ and $\mathfrak{Q}$ by $H^{\mathfrak{P}}$ and $H^{\mathfrak{Q}}$, respectively. Let $\theta := \{1,\ldots,T\}$ be a finite set. For every $t \in \theta$ let $N_t$ be a quantum channel $\mathcal{S}(H^{\mathfrak{P}}) \to \mathcal{S}(H^{\mathfrak{Q}})$.*

*We call the set of the quantum channel $\{(N_t) : t \in \theta\}$ a* **quantum compound channel***. When the channel state is $t$ and the sender inputs a quantum state $\rho^{\mathfrak{P}} \in \mathcal{S}(H^{\mathfrak{P}})$ into the channel, the receiver receives an output quantum state $N_t(\rho^{\mathfrak{P}}) \in \mathcal{S}(H^{\mathfrak{Q}})$.*

**Definition 2.3.** *Let $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ be finite sets. Let $\theta := \{1,\ldots,T\}$ be a finite set. For every $t \in \theta$ let $\mathsf{W}_t$ be a classical channel $P(\mathbf{A}) \to P(\mathbf{B})$ and $\mathsf{V}_t$ be a classical channel $P(\mathbf{A}) \to P(\mathbf{C})$.*

*We call the set of the classical channel pairs $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$ a* **(classical) compound wiretap channel***. When the channel state is $t$, and the sender inputs $p \in P(\mathbf{A})$ into the channel, the receiver receives the output $y \in B$ with probability $\sum_x p(x)\mathsf{W}_t(y|x)$, while the wiretapper receives the output $z \in Z$ with probability $\sum_x p(x)\mathsf{V}_t(z|x)$.*

**Definition 2.4.** *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be finite sets and* $H$ *be a complex Hilbert space. Let* $\theta := \{1, \ldots, T\}$ *be a finite set. For every* $t \in \theta$ *let* $\mathsf{W}_t$ *be a classical channel* $P(\mathbf{A}) \to P(\mathbf{B})$ *and* $\mathsf{V}_t$ *be a classical-quantum channel* $P(\mathbf{A}) \to \mathcal{S}(H)$.

*We call the set of the classical channel and classical-quantum channel pairs* $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$ *a* **compound channel with quantum wiretapper**. *When the channel state is* $t$ *and the sender inputs* $p \in P(\mathbf{A})$ *into the channel, the receiver receives the output* $y \in B$ *with probability* $\sum_x p(x)\mathsf{W}_t(y|x)$, *while the wiretapper receives an output quantum state* $\sum_x p(x)\mathsf{V}_t^{\otimes n}(x) \in \mathcal{S}(H)$.

**Definition 2.5.** *Let* $H$, $H'$, *and* $H''$ *be complex Hilbert spaces. Let* $\theta := \{1, \ldots, T\}$ *be a finite set. For every* $t \in \theta$ *let* $W_t$ *be a quantum channel* $\mathcal{S}(H') \to \mathcal{S}(H'')$ *and* $V_t$ *be a quantum channel* $\mathcal{S}(H') \to \mathcal{S}(H)$.

*We call the set of the quantum channel pairs* $(W_t, V_t)_{t\in\theta}$ *a* **quantum compound wiretap channel**. *When the channel state is* $t$ *and the sender inputs a quantum state* $\rho \in \mathcal{S}(H')$ *into the channel, the receiver receives an output quantum state* $W_t(\rho) \in \mathcal{S}(H'')$, *while the wiretapper receives an output quantum state* $V_t(\rho) \in \mathcal{S}(H)$.

*We distinguish two different scenarios according to the sender's knowledge of the channel state:*

- *the sender has the CSI, i.e. he knows which* $t$ *the channel state actually is,*
- *the sender does not have any CSI.*

*In both cases we assume that the receiver does not have any CSI, but the wiretapper always has the full knowledge of the CSI. Of course we also have the case where both the sender and the receiver have the CSI, but this case is equivalent to the case when we only have one pair of channels* $(W_t, V_t)$, *instead of a family of pairs of channels* $\{(W_t, V_t) : t = 1, \ldots, T\}$.

**Definition 2.6.** *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be finite sets and* $\theta := \{1, \ldots, T\}$ *be a finite set. For every* $t \in \theta$, *let* $\mathsf{W}_t$ *be a classical channel* $P(\mathbf{A}) \to P(\mathbf{B})$. *We call the set of the classical channels* $\{\mathsf{W}_t : t \in \theta\}$ *an* **arbitrarily varying channel** *when the channel state* $t$ *varies from symbol to symbol in an arbitrary manner.*

**Definition 2.7.** *Let* $\mathbf{A}$ *be a finite set. Let* $H$ *be a finite-dimensional complex Hilbert space, and* $\theta := \{1, \cdots, T\}$ *be a finite set. For every* $t \in \theta$, *let* $W_t$ *be a classical-quantum channel* $P(\mathbf{A}) \to \mathcal{S}(H)$. *The set of the classical-quantum channels* $\{W_t : t \in \theta\}$ *defines an* **arbitrarily varying classical-quantum channel** *when the channel state* $t$ *varies from symbol to symbol in an arbitrary manner.*

*Strictly speaking, the set* $\{W_t : t \in \theta\}$ *generates the arbitrarily varying classical-quantum channel* $\{W_{t^n} : t^n \in \theta^n\}$. *When the sender inputs a* $P^n \in P(\mathbf{A}^n)$ *into the channel, the receiver receives the output* $W_{t^n}(P^n) \in \mathcal{S}(H^{\otimes n})$, *where* $t^n = (t_1, t_2, \cdots, t_n) \in \theta^n$ *is the channel state of* $W_{t^n}$.

**Definition 2.8.** *We say that the arbitrarily varying channel* $\{\mathsf{W}_t : t \in \theta\}$ *is* **symmetrizable** *if there exists a parametrized set of distributions* $\{\tau(\cdot \mid a) : a \in \mathbf{A}\}$ *on* $\theta$ *such that for all* $a$, $a' \in \mathbf{A}$, *and* $b \in \mathbf{B}$

$$\sum_{t\in\theta} \tau(t \mid a)\mathsf{W}_t(b \mid a') = \sum_{t\in\theta} \tau(t \mid a')\mathsf{W}_t(b \mid a)\,.$$

*We say that the arbitrarily varying classical-quantum channel* $\{W_t : t \in \theta\}$ *is* **symmetrizable** *if there exists a parametrized set of distributions* $\{\tau(\cdot \mid a) : a \in \mathbf{A}\}$ *on* $\theta$ *such that for all* $a$, $a' \in \mathbf{A}$,

$$\sum_{t\in\theta} \tau(t \mid a) W_t(a') = \sum_{t\in\theta} \tau(t \mid a') W_t(a) \,.$$

**Definition 2.9.** *Let* $\mathfrak{P}$ *and* $\mathfrak{Q}$ *be quantum systems, denote the Hilbert Space of* $\mathfrak{P}$ *and* $\mathfrak{Q}$ *by* $H^{\mathfrak{P}}$ *and* $H^{\mathfrak{Q}}$*, respectively, and let* $\theta := \{1, \cdots, T\}$ *be a finite set. For every* $t \in \theta$*, let* ${W'}_t$ *be a quantum channel* $\mathcal{S}(H^{\mathfrak{P}}) \to \mathcal{S}(H^{\mathfrak{Q}})$*. We call the set of the quantum channels* $\{{W'}_t : t \in \theta\}$ *an* **arbitrarily varying quantum channel** *when the state* $t$ *varies from symbol to symbol in an arbitrary manner. We denote the set of arbitrarily varying quantum channels* $\mathcal{S}(H^{\mathfrak{P}}) \to \mathcal{S}(H^{\mathfrak{Q}})$ *by* $C(H^{\mathfrak{P}}, H^{\mathfrak{Q}})$*.*

**Definition 2.10.** *Let* $\mathbf{A}$ *be a finite set. Let* $H$ *and* $H'$ *be finite-dimensional complex Hilbert spaces. Let* $\theta := \{1, \cdots, T\}$ *be a finite set. For every* $t \in \theta$*, let* $W_t$ *be a classical-quantum channel* $P(\mathbf{A}) \to \mathcal{S}(H)$ *and* $V_t$ *be a classical-quantum channel* $P(\mathbf{A}) \to \mathcal{S}(H')$*. We call the set of the classical-quantum channel pairs* $\{(W_t, V_t) : t \in \theta\}$ *an* **arbitrarily varying classical-quantum wiretap channel***, the legitimate receiver accesses the output of the first channel, i.e.* $W_t$ *in the pair* $(W_t, V_t)$*, and the wiretapper observes the output of the second channel, i.e.* $V_t$ *in the pair* $(W_t, V_t)$*, respectively, when the state* $t$ *varies from symbol to symbol in an arbitrary manner.*

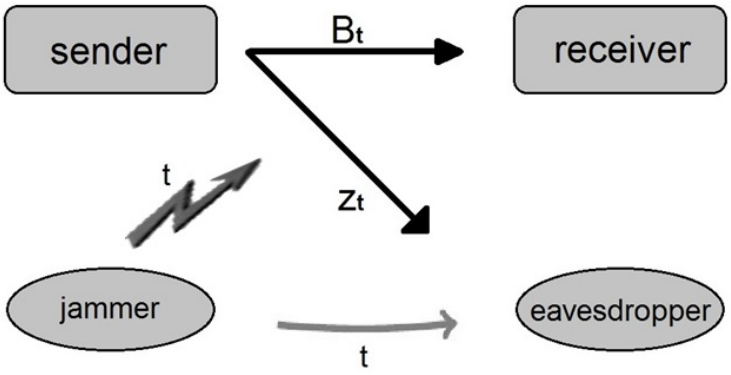


Figure 1: Arbitrarily varying classical-quantum wiretap channel

**Definition 2.11.** *Let* $\mathbf{A}$ *be a finite set. Let* $H$ *and* $H'$ *be finite-dimensional complex Hilbert spaces. Let* $\overline{\theta} := \{1, 2, \cdots\}$ *and* $\theta := \{1, 2, \cdots\}$ *be index sets. For every* $s \in \overline{\theta}$ *let* $\overline{W}_s$ *be a classical-quantum channel* $P(\mathbf{A}) \to \mathcal{S}(H)$*. For every* $t \in \theta$ *let* $V_t$ *be a classical-quantum channel* $P(\mathbf{A}) \to \mathcal{S}(H')$*. We call the set of the classical-quantum channel pairs* $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$ *a* **compound-arbitrarily varying wiretap classical-quantum channel** *when the channel state* $s$ *remains constant over time, but the legitimate users cannot control which* $s$ *in the set* $\overline{\theta}$ *will be used and the state* $t$ *varies from symbol to symbol in an arbitrary manner, while the legal receiver accesses the output of the first channel, i.e.* $\overline{W}_s$ *in the pair* $(\overline{W}_s, V_t)$ *and the wiretapper observes the output of the second channel, i.e.* $V_t$ *in the pair* $(\overline{W}_s, V_t)$*, respectively.*

*When the sender inputs a sequence* $a^n \in \mathbf{A}^n$ *into the channel, the receiver receives the output* $W_{t^n}(a^n) \in \mathcal{S}(H^{\otimes n})$*, where* $t^n = (t_1, t_2, \cdots, t_n) \in \theta^n$ *is the channel state, while the wiretapper receives an output quantum state* $V_{t^n}(a^n) \in \mathcal{S}(H'^{\otimes n})$*.*

**Definition 2.12.** *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be finite sets, and* $H$ *be a finite-dimensional complex Hilbert space. Let* $\theta := \{1, 2, \cdots\}$ *be an index set. For every* $t \in \theta$ *let* $\mathsf{W}_t$ *be a classical channel* $P(\mathbf{A}) \to P(\mathbf{B})$ *and* $V_t$ *be a classical-quantum channel* $P(\mathbf{A}) \to \mathcal{S}(H)$*. We call the set of the classical/classical-quantum channel pairs* $\{(\mathsf{W}_t, V_t) : t \in \theta\}$ *a* **classical arbitrarily varying quantum wiretap channel** *when the state* $t$ *varies from symbol to symbol in an arbitrary manner, while the legitimate receiver accesses the output of the first channel, i.e.* $W_t$ *in the pair* $(\mathsf{W}_t, V_t)$ *and the wiretapper observes the output of the second channel, i.e.* $V_t$ *in the pair* $(\mathsf{W}_t, V_t)$*, respectively.*

## 2.2 Code Concepts and Resources

### 2.2.1 Compound Wiretap Channel

**Definition 2.13.** *An* $(n, J_n)$ **code** *for the compound channel with quantum wiretapper* $\{(\mathsf{W}_t, \mathtt{V}_t) : t \in \theta\}$ *consists of a stochastic encoder* $E : \{1, \ldots, J_n\} \to P(\mathbf{A}^n)$*, specified by a matrix of conditional probabilities* $E(\cdot|\cdot)$*, and a collection of mutually disjoint sets* $\{D_j \subset \mathbf{B}^n : j \in \{1, \ldots, J_n\}\}$ *(decoding sets).*

**Definition 2.14.** *A non-negative number* $R$ *is an achievable* **secrecy rate** *for the compound channel with quantum wiretapper* $\{(\mathsf{W}_t, \mathtt{V}_t) : t \in \theta\}$ *having CSI at the encoder, if for every positive* $\varepsilon$*,* $\delta$*, every* $t \in \theta$*, and a sufficiently large* $n$*, there is an* $(n, J_n)$ *code* $(E_t, \{D_j : j = 1, \ldots, J_n\})$ *such that* $\frac{1}{n} \log J_n \geq R - \delta$*, and*

$$\max_{t \in \theta} \max_{j \in \{1, \ldots, J_n\}} \sum_{x^n \in A^n} E_t(x^n|j) \mathsf{W}_t^n(D_j^c|x^n) \leq \varepsilon \,, \tag{8}$$

$$\max_{t \in \theta} \chi(X_{uni}; Z_t^n) \leq \varepsilon \,, \tag{9}$$

*where we denote the complement of a set* $\Xi$ *by* $\Xi^c$*. Here* $R_{uni}$ *is a random variable uniformly distributed on* $\{1, \ldots, J_n\}$ *and* $Z_t^n$ *are the resulting quantum states at the output of wiretap channels* $\mathtt{V}_t^n$*.*

*A non-negative number* $R$ *is an achievable* **secrecy rate** *for the compound channel with quantum wiretapper* $\{(\mathsf{W}_t, \mathtt{V}_t) : t \in \theta\}$ *having no CSI at the encoder, if for every positive* $\varepsilon$*,* $\delta$ *and a sufficiently large* $n$*, there is an* $(n, J_n)$ *code* $(E, \{D_j : j = 1, \ldots, J_n\})$ *such that* $\frac{1}{n} \log J_n \geq R - \delta$*, and*

$$\max_{t \in \theta} \max_{j \in \{1, \ldots, J_n\}} \sum_{x^n \in A^n} E(x^n|j) \mathsf{W}_t^n(D_j^c|x^n) \leq \varepsilon \,, \tag{10}$$

$$\max_{t \in \theta} \chi(R_{uni}; Z_t^n) \leq \varepsilon \,. \tag{11}$$

**Definition 2.15.** *An* $(n, J_n)$ **code carrying classical information** *for the compound quantum wiretap channel* $\{(W_t, V_t) : t \in \theta\}$ *consists of a family of quantum states* $\{w(j) : j = 1, \ldots, J_n\} \subset \mathcal{S}(H'^{\otimes n})$ *and a collection of positive-semidefinite operators* $\{D_j : j \in \{1, \ldots, J_n\}\}$ *on* $\mathcal{S}(H''^{\otimes n})$ *which is a partition of the identity, i.e.* $\sum_{j=1}^{J_n} D_j = \mathrm{id}_{H''^{\otimes n}}$*.*

**Definition 2.16.** *A non-negative number* $R$ *is an achievable* **secrecy rate with classical input** *for the compound quantum wiretap channel* $\{(W_t, V_t) : t \in \theta\}$ *having CSI at*

*the encoder with average error, if for every positive $\varepsilon$, $\delta$, every $t \in \theta$, and a sufficiently large $n$, there is an $(n, J_n)$ code carrying classical information $(\{w_t(j) : j\}, \{D_j : j\})$ such that $\frac{1}{n} \log J_n \geq R - \delta$, and*

$$\max_{t\in\theta} \frac{1}{J_n} \sum_{j=1}^{J_n} \operatorname{tr}\left((\mathrm{id}_{H''^{\otimes n}} - D_j) W_t^{\otimes n}\left(w_t(j)\right)\right) \leq \varepsilon\,, \tag{12}$$

$$\max_{t\in\theta} \chi(R_{uni}; Z_t^n) \leq \varepsilon\,. \tag{13}$$

*A non-negative number $R$ is an achievable* **secrecy rate with classical input** *for the compound quantum wiretap channel $(W_t, V_t)_{t\in\theta}$ having no CSI at the encoder, if for every positive $\varepsilon$ and $\delta$, and a sufficiently large $n$, there is an $(n, J_n)$ code carrying classical information $(\{w(j) : j\}, \{D_j : j\})$ such that $\frac{1}{n} \log J_n \geq R - \delta$, and*

$$\max_{t\in\theta} \max_{j\in\{1,\ldots,J_n\}} \operatorname{tr}\left((\mathrm{id}_{H''^{\otimes n}} - D_j) W_t^{\otimes n}\left(w(j)\right)\right) \leq \varepsilon\,, \tag{14}$$

$$\max_{t\in\theta} \chi(R_{uni}; Z_t^n) \leq \varepsilon\,. \tag{15}$$

*Instead of "achievable secrecy rate with classical input for the compound quantum wiretap channel ", we say $R$ is an achievable secrecy rate for the compound classical-quantum wiretap channel $(W_t, V_t)_{t\in\theta}$.*

**Definition 2.17.** *An $(n, J_n)$* **code carrying quantum information** *for the compound quantum channel $\{\left(N_t^{\otimes n}\right) : t \in \theta\}$ consists of a Hilbert spaces $H^{\mathfrak{A}}$ such that $\dim H^{\mathfrak{A}} = J_n$, and a general decoding quantum operation $D$, i.e. a completely positive, trace-preserving map $D : \mathcal{S}(H^{\mathfrak{Q}^n}) \to \mathcal{S}(H^{\mathfrak{M}})$, where $H^{\mathfrak{M}}$ is a Hilbert space such that $\dim H^{\mathfrak{M}} = J_n$. The code can be used for entanglement generation in the following way. The sender prepares a pure bipartite quantum state $|\psi\rangle^{\mathfrak{A}\mathfrak{P}^n}$, defined on $H^{\mathfrak{A}} \otimes H^{\mathfrak{P}^n}$, and sends the $\mathfrak{P}^n$ portion of it through the channel $N_t^{\otimes n}$. The receiver performs the general decoding quantum operation on the channel output $D : \mathcal{S}(H^{\mathfrak{Q}^n}) \to \mathcal{S}(H^{\mathfrak{M}})$. The sender and the receiver share the resulting quantum state*

$$\Omega_t^{\mathfrak{A}\mathfrak{M}} := [\mathrm{id}^{\mathfrak{A}} \otimes (D \circ N_t^{\otimes n})] \left(|\psi\rangle\langle\psi|^{\mathfrak{A}\mathfrak{P}^n}\right)\,. \tag{16}$$

**Definition 2.18.** *A non-negative number $R$ is an achievable* **entanglement generating rate** *for the compound quantum channel $\{\left(N_t^{\otimes n}\right) : t \in \theta\}$ if for every positive $\varepsilon$, $\delta$, and a sufficiently large $n$, there is an $(n, J_n)$ code carrying quantum information $\left(H^{\mathfrak{A}}, D\right)$ such that $\frac{1}{n} \log J_n \geq R - \delta$, and*

$$\min_{t\in\theta} F\left(\Omega_t^{\mathfrak{A}\mathfrak{M}}, |\Phi_K\rangle\langle\Phi_K|^{\mathfrak{A}\mathfrak{M}}\right) \geq 1 - \varepsilon\,, \tag{17}$$

*where*

$$|\Phi_K\rangle^{\mathfrak{A}\mathfrak{M}} := \sqrt{\frac{1}{J_n}} \sum_{j=1}^{J_n} |j\rangle^{\mathfrak{A}} |j\rangle^{\mathfrak{M}}\,,$$

*which is the standard maximally entangled state shared by the sender and the receiver. $\{|j\rangle^{\mathfrak{A}}\}$ and $\{|j\rangle^{\mathfrak{M}}\}$ are orthonormal bases for $H^{\mathfrak{A}}$ and $H^{\mathfrak{M}}$, respectively.*

**Definition 2.19.** *Let $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$ be a classical compound wiretap channel.*

*The supremum on achievable secrecy rates of $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$ having CSI at the encoder is called the secrecy capacity of $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$ having CSI at the encoder, denoted by $C_{S,CSI}(\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\})$.*

*The supremum on achievable secrecy rates of $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$ is called the secrecy capacity of $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$, denoted by $C_S(\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\})$.*

*Let $\{(W_t, V_t) : t \in \theta\}$ be a compound quantum wiretap channel.*

*The supremum on achievable secrecy rates of $\{(W_t, V_t) : t \in \theta\}$ having CSI at the encoder is called the secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ having CSI at the encoder, denoted by $C_{S,CSI}(\{(W_t, V_t) : t \in \theta\})$.*

*The supremum on achievable secrecy rates of $\{(W_t, V_t) : t \in \theta\}$ is called the secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$, denoted by $C_S(\{(W_t, V_t) : t \in \theta\})$.*

*Let $\left\{\left(N_t^{\otimes n}\right) : t \in \theta\right\}$ be a compound quantum channel.*

*The supremum on achievable entanglement generating rates for $\left\{\left(N_t^{\otimes n}\right) : t \in \theta\right\}$ is called the entanglement generating capacity of $\left\{\left(N_t^{\otimes n}\right) : t \in \theta\right\}$, denoted by $A\left(\left\{\left(N_t^{\otimes n}\right) : t \in \theta\right\}\right)$.*

### 2.2.2 Arbitrarily Varying Classical-Quantum Wiretap Channel

*Our goal is to see what the effects on the secrecy capacities of an arbitrarily varying classical-quantum wiretap channel are if the sender and the legal receiver have the possibility to use various kinds of resources. We also want to investigate what amount of randomness is necessary for the robust and secure message transmission through an arbitrarily varying classical-quantum wiretap channel. Hence, we consider various kinds of resources, each of them requiring a different amount of randomness, and we consider different codes, each of them requiring a different kind of resource.*

**Definition 2.20.** *An $(n, J_n)$* **(deterministic) code** *$\mathcal{C}$ for the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$ consists of a stochastic encoder $E : \{1, \cdots, J_n\} \to P(\mathbf{A}^n)$, $j \to E(\cdot|j)$, specified by a matrix of conditional probabilities $E(\cdot|\cdot)$, and a collection of positive-semidefinite operators $\{D_j : j \in \{1, \cdots, J_n\}\}$ on $H^{\otimes n}$, which is a partition of the identity, i.e. $\sum_{j=1}^{J_n} D_j = \mathrm{id}_{H^{\otimes n}}$. We call these operators the decoder operators.*

*A code is created by the sender and the legal receiver before the message transmission starts. The sender uses the encoder to encode the message that he wants to send, while the legal receiver uses the decoder operators on the channel output to decode the message.*

**Remark 2.21.** *An $(n, J_n)$ deterministic code $\mathcal{C}$ with deterministic encoder consists of a family of $n$-length strings of symbols $(c_j)_{j \in \{1, \cdots, J_n\}} \in (\mathbf{A}^n)^{J_n}$ and a collection of positive-semidefinite operators $\{D_j : j \in \{1, \cdots, J_n\}\}$ on $H^{\otimes n}$ which is a partition of the identity.*

*The deterministic encoder is a special case of the stochastic encoder when we require that for every $j \in \{1, \cdots, J_n\}$, there is a sequence $a^n \in \mathbf{A}^n$ chosen with probability $1$. The standard technique for message transmission over a channel and robust message transmission over an arbitrarily varying channel is to use the deterministic*

*encoder (cf. [5] and [28]). However, we use the stochastic encoder, since it is a tool for secure message transmission over wiretap channels (cf. [21] and [6]).*

**Definition 2.22.** *A non-negative number $R$ is an achievable* **(deterministic) secrecy rate** *for the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$ if for every $\epsilon > 0$, $\delta > 0$, $\zeta > 0$ and sufficiently large $n$ there exists an $(n, J_n)$ code $\mathcal{C} = \left(E, \{D_j^n : j = 1, \cdots J_n\}\right)$ such that $\frac{\log J_n}{n} > R - \delta$, and*

$$\max_{t^n \in \theta^n} P_e(\mathcal{C}, t^n) < \epsilon \,, \tag{18}$$

$$\max_{t^n \in \theta^n} \chi\left(R_{uni}; Z_{t^n}\right) < \zeta \,, \tag{19}$$

*where $R_{uni}$ is the uniform distribution on $\{1, \cdots J_n\}$. Here $P_e(\mathcal{C}, t^n)$ (the average probability of the decoding error of a deterministic code $\mathcal{C}$, when the channel state of the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$ is $t^n = (t_1, t_2, \cdots, t_n)$), is defined as*

$$P_e(\mathcal{C}, t^n) := 1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \mathrm{tr}(W_{t^n}(E(\,|j))D_j) \,,$$

*and $Z_{t^n} = \left\{V_{t^n}(E(\,|i)) : i \in \{1, \cdots, J_n\}\right\}$ is the set of the resulting quantum state at the output of the wiretap channel when the channel state of $\{(W_t, V_t) : t \in \theta\}$ is $t^n$.*

**Remark 2.23.** *A weaker and widely used security criterion is obtained if we replace (19) with $\max_{t \in \theta} \frac{1}{n} \chi\left(R_{uni}; Z_{t^n}\right) < \zeta$.*

**Remark 2.24.** *When we defined $W_t$ as $\mathbf{A} \to \mathcal{S}(H)$, then $P_e(\mathcal{C}, t^n)$ is defined as $1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{a^n \in \mathbf{A}^n} E(a^n|j) \mathrm{tr}(W_{t^n}(a^n) D_j)$.*

*When deterministic encoder is used, then $P_e(\mathcal{C}, t^n)$ is defined as $1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \mathrm{tr}(W_{t^n}(c_j) D_j)$.*

**Definition 2.25.** *A non-negative number $R$ is an* **enhanced achievable secrecy rate** *for the compound-arbitrarily varying wiretap classical-quantum channel $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$ if for every $\epsilon > 0$, $\delta > 0$, $\zeta > 0$ and sufficiently large $n$ there exists an $(n, J_n)$ code $\mathcal{C} = \left(E^n, \{D_j^n : j = 1, \cdots J_n\}\right)$ such that $\frac{\log J_n}{n} > R - \delta$, and*

$$\max_{s \in \overline{\theta}} P_e(\mathcal{C}, s, n) < \epsilon \,, \tag{20}$$

$$\max_{t^n \in \theta^n} \max_{\pi \in \mathsf{S}_n} \chi\left(R_{uni}; Z_{t^n, \pi}\right) < \zeta \,, \tag{21}$$

*where $R_{uni}$ is the uniform distribution on $\{1, \cdots J_n\}$. Here $P_e(\mathcal{C}, s, n)$ is defined as follows*

$$P_e(\mathcal{C}, s, n) := 1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \mathrm{tr}(\overline{W}_s^{\otimes n}(E^n(\,|j))D_j^n) \,,$$

*and $Z_{t^n, \pi} = \Big\{\sum_{a^n \in \mathbf{A}^n} E^n(\pi(a^n)|1) V^{t^n}(\pi(a^n)), \sum_{a^n \in \mathbf{A}^n} E^n(\pi(a^n)|2) V^{t^n}(\pi(a^n)),$ $\cdots, \sum_{a^n \in \mathbf{A}^n} E^n(\pi(a^n)|J_n) V^{t^n}(\pi(a^n))\Big\}$.*

**Definition 2.26.** *An $(n, J_n)$* **code** *$\mathcal{C}$ for the classical arbitrarily varying quantum wiretap channel $\{(\mathsf{W}_t, V_t) : t \in \theta\}$ consists of a stochastic encoder $E : \{1, \cdots, J_n\} \to P(\mathbf{A}^n)$, $j \to E(\cdot|j)$, specified by a matrix of conditional probabilities $E(\cdot|\cdot)$, and a collection of mutually disjoint sets $\{D_j \subset \mathbf{B}^n : j \in \{1, \ldots, J_n\}\}$ (decoding sets).*

**Definition 2.27.** *A non-negative number $R$ is an achievable* **secrecy rate** *for the classical arbitrarily varying quantum wiretap channel $\{(\mathsf{W}_t, V_t) : t \in \theta\}$ if for every $\epsilon > 0$, $\delta > 0$, $\zeta > 0$ and sufficiently large $n$ there exists an $(n, J_n)$ code $\mathcal{C} = \big(E, \{D_j : j = 1, \cdots J_n\}\big)$ such that $\frac{\log J_n}{n} > R - \delta$, and*

$$\max_{t \in \theta} \max_{j \in \{1,\ldots,J_n\}} \mathsf{W}_t^n(D_j^c|E(\cdot|j)) \leq \varepsilon\,, \tag{22}$$

*and*

$$\max_{t^n \in \theta^n} \chi\left(R_{uni}; Z_{t^n}\right) < \zeta\,. \tag{23}$$

*Now we will define some further coding schemes, where the sender and the receiver use correlation as a resource. We will later show that these coding schemes are very helpful for the robust and secure message transmission over an arbitrarily varying wiretap channel.*

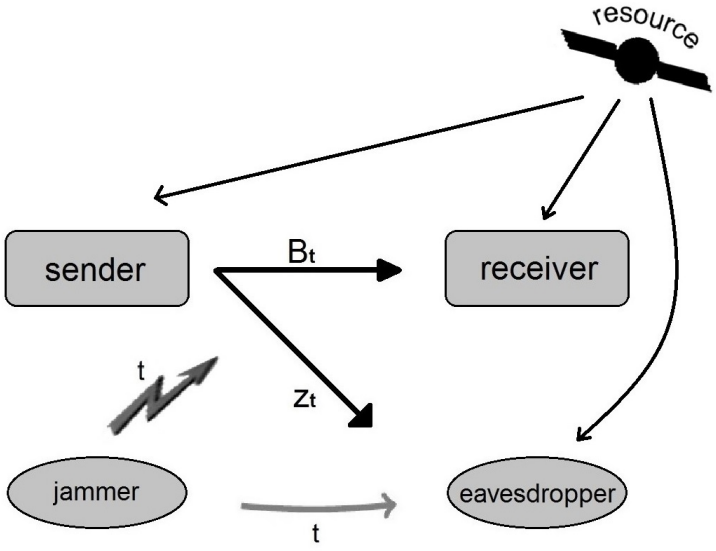


Figure 2: Arbitrarily varying classical-quantum wiretap channel with assistance by shared randomness

**Definition 2.28.** *Let $\mathbf{X}$ and $\mathbf{Y}$ be finite sets. Let $(X, Y)$ be a random variable distributed according to a probability distribution $p \in P(\mathbf{X} \times \mathbf{Y})$.*

*An $(X, Y)$-***correlation assisted** $(n, J_n)$ **code** *$\mathcal{C}(X, Y)$ for the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$ consists of a set of stochastic encoders $\{E_{\mathbf{x}^n} : \{1, \cdots, J_n\} \to P(\mathbf{A}^n) : \mathbf{x}^n \in \mathbf{X}^n\}$, and a set of collections of*

*positive-semidefinite operators* $\left\{\{D_j^{(\mathbf{y}^n)} : j = 1, \cdots, J_n\} : \mathbf{y}^n \in \mathbf{Y}^n\right\}$ *on* $H^{\otimes n}$ *which fulfills* $\sum_{j=1}^{J_n} D_j^{(\mathbf{y}^n)} = \mathrm{id}_{H^{\otimes n}}$ *for every* $\mathbf{y}^n \in \mathbf{Y}^n$.

**Definition 2.29.** *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be finite sets, and let* $(X, Y)$ *be a random variable distributed according to a joint probability distribution* $p \in P(\mathbf{X} \times \mathbf{Y})$.

*A non-negative number* $R$ *is an achievable* $m - a - (X, Y)$ **secrecy rate** *(message transmission under the average error criterion using* $(X, Y)$*-correlation assisted* $(n, J_n)$ *codes) for the arbitrarily varying classical-quantum wiretap channel* $\{(W_t, V_t) : t \in \theta\}$ *if for every* $\epsilon > 0$, $\delta > 0$, $\zeta > 0$ *and sufficiently large* $n$ *there exists an* $(X, Y)$*-correlation assisted* $(n, J_n)$ *code* $\mathcal{C}(X, Y) = \left\{\left(E_{\mathbf{x}^n}, \{D_j^{(\mathbf{y}^n)} : j \in \{1, \cdots, J_n\}\}\right) : \mathbf{x}^n \in \mathbf{X}^n,\ \mathbf{y}^n \in \mathbf{Y}^n\right\}$ *such that* $\frac{\log J_n}{n} > R - \delta$, *and*

$$\max_{t^n \in \theta^n} \sum_{\mathbf{x}^n \in \mathbf{X}^n} \sum_{\mathbf{y}^n \in \mathbf{Y}^n} p(\mathbf{x}^n, \mathbf{y}^n) P_e(\mathcal{C}(\mathbf{x}^n, \mathbf{y}^n), t^n) < \epsilon\,,$$

$$\max_{t^n \in \theta^n} \chi\left(R_{uni}; Z_{t^n, \mathbf{x}^n} \mid X\right) < \zeta\,,$$

*where* $P_e(\mathcal{C}(\mathbf{x}^n, \mathbf{y}^n), t^n)$ *is defined as*

$$P_e(\mathcal{C}(\mathbf{x}^n, \mathbf{y}^n), t^n) := 1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \mathrm{tr}(W_{t^n}(E_{\mathbf{x}^n}(\ |j))D_j^{(\mathbf{y}^n)})\,,$$

$$\chi\left(R_{uni}; Z_{t^n, \mathbf{x}^n} \mid X\right) := \sum_{\mathbf{y}^n \in \mathbf{Y}^n} p(\mathbf{x}^n, \mathbf{y}^n) \chi\left(R_{uni}; Z_{t^n, \mathbf{x}^n}\right)\,,$$

*and* $Z_{t^n, \mathbf{x}^n} = \left\{V_{t^n}(E_{\mathbf{x}^n}(\ |i)) : i \in \{1, \cdots, J_n\}\right\}$, $p(\mathbf{x}^n, \mathbf{y}^n) = \prod_{i=1}^n p(\mathbf{x}_i, \mathbf{y}_i)$. *Here we allowed* $Z_{t^n, \mathbf{x}^n}$, *the resulting quantum state of the wiretapper, to be dependent on* $\mathbf{x}^n$, *this means that we do not require* $(X, Y)$ *to be secure against eavesdropping.*

**Remark 2.30.** *Here we follow [28] and use the definition "*$m - a - (X, Y)$ *secrecy rate" because it is important to point out that here the average error criterion is used. Please see [28] for more discussions on the value of message transmission under the average error criterion and message transmission under the maximum error criterion.*

**Definition 2.31.** *Let* $\left\{C^\gamma = \{(E^\gamma, D_j^\gamma) : j = 1, \ldots, J_n\} : \gamma \in \Lambda\right\}$ *be the the set of* $(n, J_n)$ *deterministic codes, labeled by a set* $\Lambda$.

*An* $(n, J_n)$ **randomness assisted quantum code** *for the arbitrarily varying classical-quantum wiretap channel* $\{(W_t, V_t) : t \in \theta\}$ *is a distribution* $G$ *on* $(\Lambda, \sigma)$, *where* $\sigma$ *is a sigma-algebra, so chosen such that the functions* $\gamma \to P_e(\mathcal{C}^\gamma, t^n)$ *and* $\gamma \to \chi\left(R_{uni}; Z_{\mathcal{C}^\gamma, t^n}\right)$ *are both* $G$*-measurable with respect to* $\sigma$ *for every* $t^n \in \theta^n$, *here for* $t^n \in \theta^n$ *and* $\mathcal{C}^\gamma = \{(w(j)^{n,\gamma}, D_j^\gamma) : j = 1, \ldots, J_n\}$,

$$Z_{\mathcal{C}^\gamma, t^n} := \{V_{t^n}(w(1)^{n,\gamma}), V_{t^n}(w(2)^{n,\gamma}), \ldots, V_{t^n}(w(n)^{n,\gamma})\}\,.$$

**Remark 2.32.** *The randomness assisted code technique is not to be confused with the random encoding technique. For the random encoding technique, only the sender, but not the receiver, randomly chooses a code word in $\mathbf{A}^n$ to encode a message $j$ according to a probability distribution. The receiver should be able to decode $j$ even when he only knows the probability distribution, but not which code word is actually chosen by the sender. For the randomness assisted code technique, the sender randomly chooses a stochastic encoder $E^\gamma$, and the receiver chooses a set of the decoder operators $\{D_j^{\gamma'} : j = 1, \cdots, J_n\}$. The receiver can decode the message if and only if $\gamma = \gamma'$, i.e. when he knows the sender's randomization.*

**Definition 2.33.** *Let $\Lambda$ and $\mathcal{C}^\gamma$, $\gamma \in \Lambda$, be defined as in Definition 2.31. An $(n, J_n)$* **common randomness assisted quantum code** *for the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$ is a finite subset $\left\{C^\gamma = \{(E^\gamma, D_j^\gamma) : j = 1, \cdots, J_n\} : \gamma \in \Gamma\right\}$ of the set of $(n, J_n)$ deterministic codes, labeled by a finite set $\Gamma$.*

**Definition 2.34.** *A non-negative number $R$ is an achievable* **secrecy rate** *for the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$* **under randomness assisted coding** *if for every $\delta > 0$, $\zeta > 0$, and $\epsilon > 0$, if $n$ is sufficiently large, there is an $(n, J_n)$ randomness assisted quantum code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$ such that $\frac{\log J_n}{n} > R - \delta$, and*

$$\max_{t^n \in \theta^n} \int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < \epsilon\,,$$

$$\max_{t^n \in \theta^n} \int_\Lambda \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG(\gamma) < \zeta\,.$$

*Here we allow $Z_{\mathcal{C}^\gamma, t^n}$, the wiretapper's resulting quantum state, to be dependent on $\mathcal{C}^\gamma$. This means that we do not require randomness to be secure against eavesdropping.*

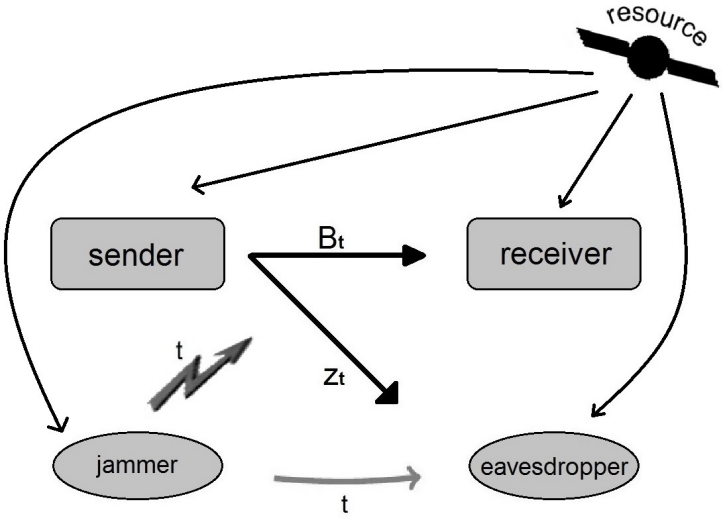


Figure 3: Arbitrarily varying classical-quantum wiretap channel with assistance by shared randomness that is known by the jammer

**Definition 2.35.** *A non-negative number $R$ is an achievable* **secrecy rate** *for the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$* **under non-secure randomness assisted coding** *if for every $\delta > 0$, $\zeta > 0$, and $\epsilon > 0$, if $n$ is sufficiently large, there is an s a distribution $G$ on $(\Lambda, \sigma)$ such that $\frac{\log J_n}{n} > R - \delta$, and*

$$\int_\Lambda \max_{t^n \in \theta^n} P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < \epsilon \,,$$

$$\int_\Lambda \max_{t^n \in \theta^n} \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG(\gamma) < \zeta \,.$$

*Here $\sigma$ is a sigma-algebra, so chosen such that the functions $\gamma \to \max_{t^n \in \theta^n} P_e(\mathcal{C}^\gamma, t^n)$ and $\gamma \to \max_{t^n \in \theta^n} \chi\left(R_{uni}; Z_{\mathcal{C}^\gamma, t^n}\right)$ are both $G$-measurable with respect to $\sigma$.*

**Definition 2.36.** *A non-negative number $R$ is an achievable* **secrecy rate** *for the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$* **under common randomness assisted quantum coding** *if for every $\delta > 0$, $\zeta > 0$, and $\epsilon > 0$, if $n$ is sufficiently large, there is an $(n, J_n)$ common randomness assisted quantum code $(\{\mathcal{C}^\gamma : \gamma \in \Gamma\})$ such that $\frac{\log J_n}{n} > R - \delta$, and*

$$\max_{t^n \in \theta^n} \frac{1}{|\Gamma|} \sum_{\gamma=1}^{|\Gamma|} P_e(\mathcal{C}^\gamma, t^n) < \epsilon \,,$$

$$\max_{t^n \in \theta^n} \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n} \mid \Gamma\right) < \zeta \,,$$

*where*

$$\chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n} \mid \Gamma\right) := \frac{1}{|\Gamma|} \sum_{\gamma=1}^{|\Gamma|} \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) \,.$$

*This means that we do not require the common randomness to be secure against eavesdropping.*

*We may consider the deterministic code, the $(X, Y)$-correlation assisted code, the $((X, Y), r)$-correlation assisted code, the $(X, Y)$-correlation assisted $(n, J_n)$ code, and the common randomness assisted quantum code as special cases of the randomness assisted quantum code. This means that randomness is a stronger resource than both common randomness and the $(X, Y)$-correlation, in the sense that it requires more randomness than common randomness and the $(X, Y)$-correlation. Randomness is therefore a more "costly" resource.*

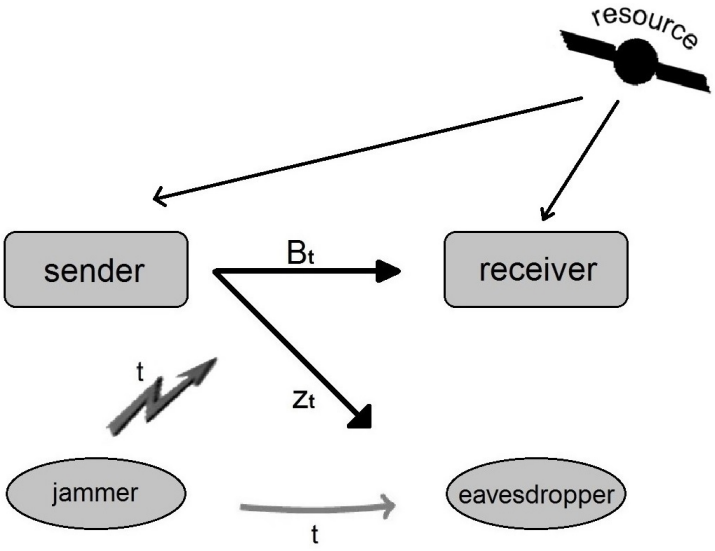


Figure 4: Arbitrarily varying classical-quantum wiretap channel with assistance by shared randomness that is not known by the eavesdropper

**Definition 2.37.** *A non-negative number $R$ is an achievable* **secrecy rate** *for the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$* **under common randomness assisted quantum coding using an amount $g_n$ of secret common randomness**, *where $g_n$ is a non-negative number depending on $n$, if for every $\delta > 0$, $\zeta > 0$, $\epsilon > 0$, and sufficiently large $n$, there is an $(n, J_n)$ common randomness assisted quantum code $(\{\mathcal{C}^\gamma : \gamma \in \Gamma_n\})$ such that $\frac{1}{n} \log |\Gamma_n| = g_n$, $\frac{\log J_n}{n} > R - \delta$, and*

$$\max_{t^n \in \theta^n} \frac{1}{2^{ng_n}} \sum_{\gamma=1}^{2^{ng_n}} P_e(\mathcal{C}^\gamma, t^n) < \epsilon \,,$$

$$\max_{t^n \in \theta^n} \chi\left(R_{uni}, Z_{t^n}\right) < \zeta \,,$$

*where $R_{uni}$ is the uniform distribution on $\{1, \cdots J_n\}$.*

*Unlike in Definition 2.33 we require here that the randomness should be secure against eavesdropping.*

*The code concept for arbitrarily varying classical-quantum wiretap channels is similar to the code concept for arbitrarily varying classical-quantum channels in [6]. We build a two-part code word, the first part is used to create the common randomness for the sender and the legal receiver, the second is used to transmit the message to the legal receiver. We call it the* **weak code concept** when the first part to synchronize the second part is public, and **strong code concept** when the first part is secure.

**Definition 2.38.** *Let $\{(W_t, V_t) : t \in \theta\}$ be an arbitrarily varying classical-quantum wiretap channel.*

*The supremum on achievable (deterministic) secrecy rates of $\{(W_t, V_t) : t \in \theta\}$ under strong code concept is called the (deterministic) secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$,*

*denoted by* $C_s(\{(W_t, V_t) : t \in \theta\})$.

*The supremum on achievable secrecy rates of* $\{(W_t, V_t) : t \in \theta\}$ *under weak code concept is called the (deterministic) secrecy capacity of* $\{(W_t, V_t) : t \in \theta\}$ *under weak code concept, denoted by* $C_{s*}(\{(W_t, V_t) : t \in \theta\})$.

*The supremum on achievable* $m - a - (X, Y)$ *secrecy rates of* $\{(W_t, V_t) : t \in \theta\}$ *is called the* $m - a - (X, Y)$ *secrecy capacity, denoted by* $C_s(\{(W_t, V_t) : t \in \theta\}; corr(X, Y))$.

*The supremum on achievable secrecy rates under random assisted quantum coding of* $\{(W_t, V_t) : t \in \theta\}$ *is called the random assisted secrecy capacity of* $\{(W_t, V_t) : t \in \theta\}$, *denoted by* $C_s(\{(W_t, V_t) : t \in \theta\}; r)$.

*The supremum on achievable secrecy rates for the* $\{(\mathsf{W}_t, V_t) : t \in \theta\}$ *under non-secure randomness assisted coding is called the non-secure randomness assisted secrecy capacity of* $\{(\mathsf{W}_t, V_t) : t \in \theta\}$ *denoted by* $C_s(\{(\mathsf{W}_t, V_t) : t \in \theta\}, r_{ns})$.

*The supremum on achievable secrecy rates under common randomness assisted quantum coding of* $\{(W_t, V_t) : t \in \theta\}$ *is called the common randomness assisted secrecy capacity of* $\{(W_t, V_t) : t \in \theta\}$, *denoted by* $C_s(\{(W_t, V_t) : t \in \theta\}; cr)$.

*The supremum on achievable secrecy rates under random assisted quantum coding using an amount* $g_n$ *of common randomness of* $\{(W_t, V_t) : t \in \theta\}$ *is called the secret random assisted secrecy capacity of* $\{(W_t, V_t) : t \in \theta\}$, *denoted by* $C_{key}(\{(W_t, V_t) : t \in \theta\}; g_n)$.

*Let* $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$ *be a compound-arbitrarily varying wiretap classical-quantum channel.*

*The supremum on the enhanced achievable secrecy rates of* $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$ *is called the enhanced secrecy capacity of* $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$, *denoted by* $\hat{C}_s(\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\})$.

*Let* $\{(\mathsf{W}_t, V_t) : t \in \theta\}$ *be a classical arbitrarily varying quantum wiretap channel.*

*The supremum on achievable secrecy (deterministic) rates of* $\{(W_t, V_t) : t \in \theta\}$ *is called the (deterministic) secrecy capacity of* $\{(\mathsf{W}_t, V_t) : t \in \theta\}$, *denoted by* $C_s(\{(\mathsf{W}_t, V_t) : t \in \theta\})$.

For an arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$ and random variable $(X, Y)$ distributed on finite sets $\mathbf{X}$ and $\mathbf{Y}$, the following facts are obvious and follow from the definitions.

$$\begin{aligned} & C_s(\{(W_t, V_t) : t \in \theta\}) \\ & \leq C_{s*}(\{(W_t, V_t) : t \in \theta\}) \\ & \leq C_s((W_t, V_t)_{t\in\theta}; corr(X, Y)) \\ & \leq C_s(\{(W_t, V_t) : t \in \theta\}; r)\,, \end{aligned} \tag{24}$$

$$\begin{aligned} & C_s(\{(W_t, V_t) : t \in \theta\}) \\ & \leq C_{s*}(\{(W_t, V_t) : t \in \theta\}) \\ & \leq C_s((W_t, V_t)_{t\in\theta}; cr) \end{aligned}$$

$$\leq C_s(\{(W_t, V_t) : t \in \theta\}; r)\,. \tag{25}$$

**Definition 2.39.** *We say super-activation occurs to two arbitrarily varying classical-quantum wiretap channels $\{(W_t, V_t) : t \in \theta\}$ and $\{(W'_t, V'_t) : t \in \theta\}$ when the following hold:*

$$C_{s*}(\{(W_t, V_t) : t \in \theta\}) = 0\,,$$
$$C_{s*}(\{(W'_t, V'_t) : t \in \theta\}) = 0\,,$$

*and*

$$C_{s*}(\{W_t \otimes W'_{t'}, V_t \otimes V'_{t'} : t, t' \in \theta\}) > 0\,.$$

Similar to the entanglement generating capacity of a compound quantum channel we can define the entanglement generating capacity of a given arbitrarily varying quantum channel $\{N_t : t \in \theta\}$, which we denote by $A(\{N_t : t \in \theta\})$.

The entanglement generating capacity of the arbitrarily varying quantum channels has been analyzed in [5]. The authors of [5] made the following Conjecture 2.40, which is still unsolved.

**Conjecture 2.40.** *The entanglement generating capacity of an arbitrarily varying quantum channel is equal to the entanglement generating capacity of an arbitrarily varying quantum channel under shared randomness assisted quantum coding.*

# 3 Secrecy Capacities of Compound Quantum Wiretap Channels

The results in this section was published in [24].

## 3.1 Compound Channels with Quantum Wiretapper

In this section we discuss the classical compound channel with a quantum wiretapper. For the case when the sender has the full knowledge about the CSI, we derive the secrecy capacity. For the case when the sender does not know the CSI, we give a lower bound for the secrecy capacity. In this channel model, the wiretapper uses classical-quantum channels.

Let $A$, $B$, $H$, $\theta$, and $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$ be defined as in Section 2.

**Theorem 3.1.** *The secrecy capacity of the compound channel with quantum wiretapper $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$ in the case with CSI at the transmitter is given by*

$$C_{S,CSI}(\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}) = \min_{t\in\theta} \max_{U \to A \to (BZ)_t} (I(p_U; B_t) - \limsup_{n\to\infty} \frac{1}{n}\chi(p_U; Z_t^n))\,. \tag{26}$$

*The maximum is taken over all random variables that satisfy the Markov chain relationships: $U \to A \to (BZ)_t$. Here $B_t$ are the resulting random variables at the*

*output of legal receiver channels, and $Z_t$ are the resulting random quantum states at the output of wiretap channels. $U$ is a random variable taking values on some finite set* $\mathbf{U}$ *with probability distribution $p_U$.*

*Respectively, in the case without CSI, the secrecy capacity of the compound channel with quantum wiretapper $\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}$ is lower bounded as follows*

$$C_S(\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}) \geq \max_{\mathcal{U} \to A \to (BZ)_t} (\min_{t\in\theta} I(p_U; B_t) - \max_{t\in\theta} \limsup_{n\to\infty} \frac{1}{n} \chi(p_U; Z_t^n)) \,. \tag{27}$$

**Remark 3.2.** *We have only the multi-letter formulas (26) and (27), since we do not have a single-letter formula even for a quantum channel which is neither compound nor has wiretappers.*

*Proof. i) Lower Bound for Case With CSI*

For every $t \in \theta$, fix a probability distribution $p_t$ on $\mathbf{A}^n$. Let $p'_t(x^n) := \begin{cases} \frac{p_t^n(x^n)}{p_t^n(\mathcal{T}^n_{p_t,\delta})} \,, & \text{if } x^n \in \mathcal{T}^n_{p_t,\delta} \,; \\ 0 \,, & \text{else} \,, \end{cases}$

and $X^{(t)} := \{X_{j,l}^{(t)}\}_{j\in\{1,\ldots,J_n\}, l\in\{1,\ldots,L_{n,t}\}}$ be a family of random matrices whose entries are selected i.i.d. according to $p'_t$, where $L_{n,t}$ is a natural number, which will be specified later.

It was shown in [18] that for any positive $\omega$, if we set

$$J_n = \lfloor 2^{n(\min_{t\in\theta}(I(p_t;\mathsf{W}_t) - \frac{1}{n}\log L_{n,t} - \mu)} \rfloor \,,$$

where $\mu$ is a positive constant which does not depend on $j$, $t$, and can be arbitrarily small when $\omega$ goes to 0, the following statement is valid. There are such $\{D_j : j = 1, \ldots, J_n\}$ that for all $t \in \theta$ and for all $L_{n,t} \in \mathbb{N}$

$$\begin{aligned} &Pr\left( \max_{j\in\{1,\ldots,J_n\}} \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} \mathsf{W}_t^n(D_j^c | X_{j,l}^{(t)}) > \sqrt{t} 2^{-n\omega/2} \right) \\ &\leq \sqrt{t} 2^{-n\omega/2} \,. \end{aligned} \tag{28}$$

Since only the error of the legitimate receiver is analyzed, for the result (28) just the channels $\mathsf{W}_t$, but not those of the wiretapper, are regarded. For every $j \in \{1, \ldots, J_n\}$, $l \in \{1, \ldots, L_{n,t}\}$, and $t \in \theta$, $\mathsf{W}_t^n(D_j^c | X_{j,l}^{(t)})$ is a random variable taking values in $]0, 1[$, which depends on $X_{j,l}^{(t)}$, since we defined $X_{j,l}^{(t)}$ as a random variable with value in $\mathbf{A}^n$.

Let

$$Q_t(x^n) := \Pi_{p_t\mathsf{V}_t,\alpha} \Pi_{\mathsf{V}_t^{\otimes n}(x^n),\alpha} \cdot V_t^{\otimes n}(x^n) \cdot \Pi_{\mathsf{V}_t^{\otimes n}(x^n),\alpha} \Pi_{p_t\mathsf{V}_t,\alpha} \,,$$

where $\alpha$ will be defined later.

**Lemma 3.3 (Tender Operator, cf. [67] and [54]).** *Let $\rho$ be a quantum state and $X$ be a positive operator with $X \leq \mathrm{id}$ and $1 - \mathrm{tr}(\rho X) \leq \lambda \leq 1$. Then*

$$\|\rho - \sqrt{X}\rho\sqrt{X}\|_1 \leq \sqrt{2\lambda} \,. \tag{29}$$

Tender Operator was first introduced in [67], where it has been shown that $\|\rho - \sqrt{X}\rho\sqrt{X}\|_1 \leq \sqrt{8\lambda}$. In [54], the result of [67] has been improved, and (29) has been proved.

In view of the fact that $\Pi_{p_t\mathtt{V}_t,\alpha\sqrt{a}}$ and $\Pi_{\mathtt{V}_t,\alpha}(x^n)$ are both projection matrices, by (4), (7), and Lemma 3.3 for any $t$ and $x^n$, it holds that

$$\|Q_t(x^n) - \mathtt{V}_t^{\otimes n}(x^n)\|_1 \leq \sqrt{2^{-n\beta(\alpha)'+1} + 2^{-n\beta(\alpha)''+1}}\,. \tag{30}$$

We set $\Theta_t := \sum_{x^n\in\mathcal{T}^n_{p_t,\delta}} {p'}^n_t(x^n)Q_t(x^n)$. For given $z^n$ and $t$, $\langle z^n|\Theta_t|z^n\rangle$ is the expected value of $\langle z^n|Q_t(x^n)|z^n\rangle$ under the condition $x^n \in \mathcal{T}^n_{p_t,\delta}$.

The following Lemma was first given in [8]. Here we cite the lemma as it was formulated in [66],

**Lemma 3.4 (Covering Lemma).** *Let $\mathcal{V}$ be a finite-dimensional Hilbert space. Let $\mathbf{M}$ be a finite set. Suppose we have an ensemble $\{\rho_m : m \in \mathbf{M}\} \subset \mathcal{S}(\mathcal{V})$ of quantum states. Let $p$ be a probability distribution on $\mathbf{M}$. We define $\rho := \sum_m p(m)\rho_m$.*

*Suppose a total subspace projector $\Pi$ and codeword subspace projectors $\{\Pi_m : m \in \mathbf{M}\}$ exist which project onto subspaces of the Hilbert space in which the states exist, and for all $m \in \mathbf{M}$ there are positive constants $\epsilon \in ]0,1[$, $D$, $d$ such that the following conditions hold:*

$$\begin{aligned}
\mathrm{tr}(\rho_m\Pi) &\geq 1-\epsilon\,,\\
\mathrm{tr}(\rho_m\Pi_m) &\geq 1-\epsilon\,,\\
\mathrm{tr}(\Pi) &\leq D\,,\\
\Pi_m\rho_m\Pi_m &\leq \frac{1}{d}\Pi_m\,.
\end{aligned}$$

*We define a sequence of i.i.d. random variables $X_1,\ldots,X_L$, taking values in $\{\rho_m : m \in \mathbf{M}\}$. If $L \gg \frac{d}{D}$, then*

$$\begin{aligned}
&Pr\left(\|L^{-1}\sum_{i=1}^{L}\Pi\cdot\Pi_m\cdot X_i\cdot\Pi_m\cdot\Pi - \rho\| \leq \epsilon + 4\sqrt{\epsilon} + 24\sqrt[4]{\epsilon}\right)\\
&\geq 1 - 2D\exp\left(-\frac{\epsilon^3 Ld}{2\ln 2D}\right)\,.
\end{aligned} \tag{31}$$

By (2) we have

$$\mathrm{tr}(\Pi_{p_t\mathtt{V}_t,\alpha}) \leq 2^{n(S(V_t(p_t))+\delta(\alpha))}\,. \tag{32}$$

By (6), for all $x^n$ it holds that

$$\Pi_{\mathtt{V}_t^{\otimes n}(x^n),\alpha}\cdot V_t^{\otimes n}(x^n)\cdot\Pi_{\mathtt{V}_t^{\otimes n}(x^n),\alpha} \leq 2^{-n(S(\mathtt{V}_t|p_t)+\delta(\alpha)')}\Pi_{\mathtt{V}_t^{\otimes n}(x^n),\alpha}\,. \tag{33}$$

Let $\lambda = \epsilon$. By applying Lemma 3.4 if $n$ is sufficiently large we have

$$Pr\left(\|\sum_{l=1}^{L_{n,t}}\frac{1}{L_{n,t}}Q_t(X_{j,l}) - \Theta_t\| > \epsilon + 4\sqrt{\epsilon} + 24\sqrt[4]{\epsilon}\right)$$

$$\leq 2^{n(S(V_t(p_t))+\delta(\alpha))} \tag{34}$$
$$\cdot \exp\left(-L_{n,t}\frac{\epsilon^3}{2\ln 2}\lambda \cdot 2^{n(S(\mathtt{V}_t|p_t)-S(V_t(p_t)))+\delta(\alpha)+\delta(\alpha)'}\right)$$
$$= 2^{n(S(V_t(p_t))+\delta(\alpha))}$$
$$\cdot \exp\left(-L_{n,t}\frac{\epsilon^3}{2\ln 2}\lambda \cdot 2^{n(-\chi(p_t;Z_t))+\delta(\alpha)+\delta(\alpha)'}\right)$$
$$\leq \exp\left(-L_{n,t}\cdot 2^{-n(\chi(p_t;Z_t)+\zeta)}\right) , \tag{35}$$

where $\zeta$ is some suitable positive constant which does not depend on $j$, $t$, and can be arbitrarily small when $\epsilon$ is close to 0. The equality holds since

$$\begin{aligned} &S(V_t(p_t)) - S(\mathtt{V}_t|p_t) \\ &= S\left(\sum_j p_t(j)\sum_l \frac{1}{L_{n,t}}\mathtt{V}_t^{\otimes n}(X_{j,l}^{(t)})\right) \\ &- \sum_j p_t(j) S\left(\sum_l \frac{1}{L_{n,t}}\mathtt{V}_t^{\otimes n}(X_{j,l}^{(t)})\right) \\ &= \chi(p_t; Z_t) . \end{aligned}$$

Let $L_{n,t} = \lceil 2^{n(\chi(p_t;Z_t)+2\zeta)}\rceil$, and $n$ be large enough, then by (35) for all $j$ it holds that

$$Pr\left(\|\sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Theta_t\| > \epsilon + 4\sqrt{\epsilon} + 24\sqrt[4]{\epsilon}\right) \leq \exp(-2^{n\zeta}) \tag{36}$$

and

$$\begin{aligned} &Pr\left(\|\sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Theta_t\| \leq \epsilon \ \forall t \ \forall j\right) \\ &= 1 - Pr\left(\bigcup_t \bigcup_j \{\|\sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Theta_t\| > \epsilon + 4\sqrt{\epsilon} + 24\sqrt[4]{\epsilon}\}\right) \\ &\geq 1 - TJ_n \exp(-2^{n\zeta}) \\ &\geq 1 - T2^{n(\min_{t\in\theta}(I(p_t;\mathsf{W}_t)-\frac{1}{n}\log L_{n,t})} \exp(-2^{n\zeta}) \\ &\geq 1 - 2^{-n\upsilon} , \end{aligned} \tag{37}$$

where $\upsilon$ is some suitable positive constant which does not depend on $j$ and $t$.

**Remark 3.5.** *Since* $\exp(-2^{n\zeta})$ *converges to zero double exponentially quickly, the inequality (37) remains true even if* $T$ *depends on* $n$ *and is exponentially large over* $n$*, i.e. we can still achieve an exponentially small error.*

From (28) and (37) it follows: For any $\varepsilon > 0$, if $n$ is large enough then the event

$$\left(\bigcap_t \left\{\max_{j\in\{1,\ldots,J_n\}} \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} \mathsf{W}_t^n(D_j^c(\mathcal{X})|X_{j,l}^{(t)}) \leq \varepsilon\right\}\right)$$

$$\cap \left( \left\{ \| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(X_{j,l}^{(t)}) - \Theta_t \| \leq \varepsilon \ \forall t \ \forall j \right\} \right)$$

has a positive probability. This means that we can find a realization $x_{j,l}^{(t)}$ of $X_{j,l}^{(t)}$ with a positive probability such that for all $t \in \theta$ and $j \in \{1, \ldots, J_n\}$, we have

$$\sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} \mathsf{W}_t^n(D_j^c | x_{j,l}^{(t)}) \leq \epsilon \,, \tag{38}$$

and

$$\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(x_{j,l}^{(t)}) - \Theta_t \| \leq \epsilon \,. \tag{39}$$

For an arbitrary $\gamma > 0$ let

$$R := \min_{t \in \theta} \max_{U \to A \to (BZ)_t} (I(p_U; B_t) - \limsup_{n \to \infty} \frac{1}{n} \chi(p_U; Z_t^n)) - \gamma \,.$$

Choose $\mu < \frac{1}{2}\gamma$, then for every $t \in \theta$, there is an $(n, J_n)$ code $\left( (x_{j,l}^{(t)})_{j=1,\ldots,J_n, l=1,\ldots,L_{n,t}}, \{D_j : j = 1, \ldots, J_n\} \right)$ such that

$$\frac{1}{n} \log J_n \geq R \,, \tag{40}$$

$$\lim_{n \to \infty} \max_{t \in \theta} \max_{j \in \{1,\ldots,J_n\}} \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} \mathsf{W}_t^n(D_j^c | x_{j,l}^{(t)}) = 0 \,. \tag{41}$$

Choose a suitable $\alpha$ in (30) such that for all $j$, it holds $\| \mathtt{V}_t^{\otimes n}(x_{j,l}^{(t)}) - Q_t(x_{j,l}^{(t)}) \| < \epsilon$ . For any given $j' \in \{1, \ldots, J_n\}$, (30) and (39) yield

$$\begin{aligned}
&\| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} \mathtt{V}_t^{\otimes n}(x_{j',l}^{(t)}) - \Theta_t \| \\
&\leq \| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} \mathtt{V}_t^{\otimes n}(x_{j',l}^{(t)}) - \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(x_{j',l}^{(t)}) \| \\
&+ \| \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} Q_t(x_{j',l}^{(t)}) - \Theta_t \| \\
&\leq \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} \| \mathtt{V}_t^{\otimes n}(x_{j',l}^{(t)}) - Q_t(x_{j',l}^{(t)}) \| \\
&+ \| \sum_{l=1}^{L_{n,t}^{(t)}} \frac{1}{L_{n,t}} Q_t(x_{j',l}^{(t)}) - \Theta_t \| \\
&\leq 2\epsilon \,,
\end{aligned} \tag{42}$$

and $\| \sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_{n,t}} \frac{1}{L_{n,t}} \mathtt{V}_t^{\otimes n}(x_{j,l}^{(t)}) - \Theta_t \|_1 \leq \epsilon$.

**Lemma 3.6 (Fannes-Audenaert Ineq., cf. [37], [10]).** *Let $\Phi$ and $\Psi$ be two quantum states in a $d$-dimensional complex Hilbert space and $\|\Phi-\Psi\|_1 \leq \mu < \frac{1}{e}$, then*

$$|S(\Phi)-S(\Psi)| \leq \mu \log(d-1) + h(\mu)\,, \tag{43}$$

*where $h(\nu) := -\nu\log\nu - (1-\nu)\log(1-\nu)$ for $\nu \in [0,1]$.*

The Fannes Inequality was first introduced in [37], where it has been shown that $|S(\mathfrak{X})-S(\mathfrak{Y})| \leq \mu \log d - \mu \log \mu$. In [10] the result of [37] has been improved, and (43) has been proved.

By Lemma 3.6 and the inequality (42), for a uniformly distributed distributed random variable $R_{uni}$ with value in $\{1,\ldots,J_n\}$, we have

$$\begin{aligned}
&\chi(R_{uni};Z_t^n)\\
&= S\left(\sum_{j=1}^{J_n}\frac{1}{J_n}\sum_{l=1}^{L_{n,t}}\frac{1}{L_{n,t}}\mathtt{V}_t^{\otimes n}(x_{j,l}^{(t)})\right)\\
&- \sum_{j=1}^{J_n}\frac{1}{J_n}S\left(\sum_{l=1}^{L_{n,t}}\frac{1}{L_{n,t}}\mathtt{V}_t^{\otimes n}(x_{j,l}^{(t)})\right)\\
&\leq \left|S\left(\sum_{j=1}^{J_n}\frac{1}{J_n}\sum_{l=1}^{L_{n,t}}\frac{1}{L_{n,t}}\mathtt{V}_t^{\otimes n}(x_{j,l}^{(t)})\right) - S\left(\Theta_t\right)\right|\\
&+ \left|S(\Theta_t) - \sum_{j=1}^{J_n}\frac{1}{J_n}S\left(\sum_{l=1}^{L_{n,t}}\frac{1}{L_{n,t}}\mathtt{V}_t^{\otimes n}(x_{j,l}^{(t)})\right)\right|\\
&\leq \epsilon\log(d-1) - \epsilon\log\epsilon - (1-\epsilon)\log(1-\epsilon)\\
&+ \left|\sum_{j=1}^{J_n}\frac{1}{J_n}\left[S(\Theta_t) - S\left(\sum_{l=1}^{L_{n,t}}\frac{1}{L_{n,t}}\mathtt{V}_t^{\otimes n}(x_{j,l}^{(t)})\right)\right]\right|\\
&\leq 3\epsilon\log(d-1) - \epsilon\log\epsilon - (1-\epsilon)\log(1-\epsilon) - 2\epsilon\log 2\epsilon\,.
\end{aligned} \tag{44}$$

By (44), for any positive $\lambda$ if $n$ is sufficiently large, we have

$$\max_{t\in\theta}\chi(R_{uni};Z_t^n) \leq \lambda\,. \tag{45}$$

For every $t \in \theta$ we define an $(n, J_n)$ code $(E_t, \{D_j : j = 1,\ldots,J_n\})$, where $E_t$ is built such that $Pr\left(E_t(j) = x_{j,l}^{(t)}\right) = \frac{1}{L_{n,t}}$ for $l \in \{1,\ldots,L_{n,t}\}$. Combining (41) and (45) we obtain

$$C_{S,CSI}(\{(\mathsf{W}_t,\mathtt{V}_t) : t\in\theta\}) \geq \min_{t\in\theta}\max_{U\to A\to(BZ)_t}\left(I(p_U;B_t) - \limsup_{n\to\infty}\frac{1}{n}\chi(p_U;Z_t^n)\right). \tag{46}$$

Thus, we have shown the "$\geq$" part of (26).

*ii) Upper Bound for Case With CSI*

Let $(\mathcal{C}_n)$ be a sequence of $(n, J_n)$ codes such that

$$\max_{t\in\theta} \max_{j\in\{1,\dots,J_n\}} \sum_{x^n\in\mathbf{A}^n} E(x^n|j)\mathsf{W}_t^n(D_j^c|x^n) =: \epsilon_{1,n} , \tag{47}$$

$$\max_{t\in\theta} \chi(J; Z_t^n) =: \epsilon_{2,n} , \tag{48}$$

where $\lim_{n\to\infty}\epsilon_{1,n} = 0$ and $\lim_{n\to\infty}\epsilon_{2,n} = 0$. $J$ denotes the random variable which is uniformly distributed on the message set $\{1,\dots,J_n\}$.

We denote the security capacity of the wiretap channel $(\mathsf{W}_t, \mathsf{V}_t)$ in the sense of [66] by $C(\mathsf{W}_t, \mathsf{V}_t)$. Choose $t' \in \theta$ such that $C(\mathsf{W}_{t'}, \mathsf{V}_{t'}) = \min_{t\in\theta} C(\mathsf{W}_t, \mathsf{V}_t)$.

We denote a new random variable by $\hat{X}$ with values in $\{1,\dots,J_n\}$ determined by the Markov chain $R_{uni} \to A \to B_{t'} \to \hat{X}$, where the first transition is governed by the sender's encoding strategy, the second by $\mathsf{W}_{t'}$, and the last by the legal receiver's decoding strategy. Then we have from the data processing inequality

$$\begin{aligned}
&\log J_n = H(R_{uni}) \\
&= I(R_{uni}, \hat{X}) + H(R_{uni} \mid \hat{X}) \\
&\le I(R_{uni}, B_{t'}^n) + H(R_{uni} \mid \hat{X}) .
\end{aligned}$$

Using Fano's inequality we have

$$H(R_{uni} \mid \hat{X}) \le 1 + \epsilon_{1,n} \log J_n .$$

Thus $\log J_n \le I(R_{uni}, B_{t'}^n) + 1 + \epsilon_{1,n}\log J_n$. Applying the standard technique for single letter formula in classical information theory we have

$$\log J_n \le nI(R_{uni}, B_{t'}) + 1 + \epsilon_{1,n}\log J_n . \tag{49}$$

Thus for any $\epsilon > 0$, if $n$ is sufficiently large $\frac{1}{n}\log J_n$ cannot be greater than

$$\begin{aligned}
&I(R_{uni}; B_{t'}) + \frac{1}{n} + \frac{1}{n}\epsilon_{1,n}\log J_n \\
&\le [I(R_{uni}; B_{t'}) - \frac{1}{n}\chi(R_{uni}; Z_{t'}^n)] + \frac{\epsilon_{1,n}}{n} + \frac{1}{n}\log J_n + \frac{\epsilon_{2,n}}{n} \\
&\le [I(R_{uni}; B_{t'}) - \frac{1}{n}\chi(R_{uni}; Z_{t'}^n)] + \epsilon .
\end{aligned} \tag{50}$$

We cannot exceed the secrecy capacity of the worst wiretap channel, since we have to guarantee that the legal receiver can decode the message in the worst case (cf. (9)). Thus, we have

$$C_{S,CSI}(\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}) \le \min_{t\in\theta} \max_{U\to A\to(BZ)_t} (I(p_U; B_t) - \limsup_{n\to\infty} \frac{1}{n}\chi(p_U; Z_t^n)) . \tag{51}$$

Combining (51) and (46) we obtain (26).

*iii) Lower Bound for Case Without CSI*

Fix a probability distribution $p$ on $\mathbf{A}^n$. For any $\omega > 0$, we define

$$J_n = \lfloor 2^{n(\min_{t\in\theta}(I(p;\mathsf{W}_t) - \frac{1}{n}\log L_n - \mu)} \rfloor ,$$

where $\mu$ is a positive constant which does not depend on $j$ and $t$, and can be arbitrarily small when $\omega$ goes to 0. Let $p'(x^n) := \begin{cases} \frac{p^n(x^n)}{p^n(\mathcal{T}^n_{p,\delta})} & \text{if } x^n \in \mathcal{T}^n_{p,\delta} \ ; \\ 0 & \text{else} \ . \end{cases}$ and $X^n := \{X_{j,l}\}_{j\in\{1,\ldots,J_n\},l\in\{1,\ldots,L_n\}}$, where $L_n$, a natural number, will be specified later, be a family of random matrices whose components are selected i.i.d. according to $p'$.

There are $\{D_j : j = 1,\ldots,J_n\}$ such that for all $t \in \theta$ and for all $L_n \in \mathbb{N}$

$$\begin{aligned} & Pr\left(\max_{j\in\{1,\ldots,J_n\}} \sum_{l=1}^{L_n} \frac{1}{L_n} \mathsf{W}_t^n(D_j^c|X_{j,l}) > \sqrt{t}2^{-n\omega/2}\right) \\ & \le \sqrt{t}2^{-n\omega/2} \ . \end{aligned} \tag{52}$$

For a positive $\alpha$, we define

$$Q_t(x^n) := \Pi_{p\mathbb{V}_t,\alpha\sqrt{a}}\Pi_{\mathbb{V}_t,\alpha}(x^n)\cdot \mathbb{V}_t^{\otimes n}(x^n)\cdot \Pi_{\mathbb{V}_t,\alpha}(x^n)\Pi_{p\mathbb{V}_t,\alpha\sqrt{a}}$$

and $\Theta_t := \sum_{x^n\in\mathcal{T}^n_{p,\delta}} {p'}^n(x^n)Q_t(x^n)$.

For any positive $\delta$ let $L_n = \lceil 2^{\max_t \chi(p;Z_t^n)+n\delta}\rceil$ and $n$ be large enough, in the same way as our proof of (37) for the case with CSI at the encoder, there is a positive constant $\upsilon$ so that

$$Pr\left(\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_t(X_{j,l}^{(t)}) - \Theta_t\| \le \epsilon \ \forall t \ \forall j\right) \ge 1 - 2^{-n\upsilon} \ . \tag{53}$$

For any positive $\epsilon$ we choose a suitable $\alpha$, by (52) and (53) there is a realization $x_{j,l}$ of $X_{j,l}$ with a positive probability such that: For all $t \in \theta$ and all $j \in \{1,\ldots J_n\}$, we have

$$\sum_{l=1}^{L_n}\frac{1}{L_n}\mathsf{W}_t^n(D_j^c|x_{j,l}) \le \epsilon \ ,$$

$$\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_t(x_{j,l}) - \Theta_t\| \le \epsilon \ .$$

For any $\gamma > 0$ let

$$R = \max_{U\to A\to(BZ)_t}\left(\min_{t\in\theta} I(p_U;B_t) - \max_t \frac{1}{n}\chi(p_U;Z_t^n)\right) - \gamma \ .$$

Then there is an $(n, J_n)$ code $(E, \{D_j : j = 1,\ldots,J_n\})$, where $E$ is so built that $Pr\left(E(j) = x_{j,l}\right) = \frac{1}{L_n}$ for $l \in \{1,\ldots,L_{n,t}\}$, such that $\liminf_{n\to\infty}\frac{1}{n}\log J_n \ge R$, and

$$\lim_{n\to\infty}\max_{t\in\theta}\max_{j\in\{1,\ldots,J_n\}}\sum_{l=1}^{L_n}\frac{1}{L_n}\mathsf{W}_t^n(D_j^c|x_{j,l}) = 0 \ . \tag{54}$$

In the same way as our proof of (45) for the case with CSI at the encoder,

$$\max_{t\in\theta}\chi(R_{uni};Z_t^n) \le \epsilon \ , \tag{55}$$

for any uniformly distributed random variable $R_{uni}$ with value in $\{1, \ldots, J_n\}$.

Combining (54) and (55) we obtain

$$C_S(\{(\mathsf{W}_t, \mathsf{V}_t) : t \in \theta\}) \geq \max_{U \to A \to (BZ)_t} (\min_{t \in \theta} I(p_U; B_t) - \max_{t \in \theta} \limsup_{n \to \infty} \frac{1}{n} \chi(p_U; Z_t^n)) .$$

□

## 3.2 Compound Classical-Quantum Wiretap Channel

In this section, we derive the secrecy capacity of the compound classical-quantum wiretap channel with CSI. In this model, both the receiver and the wiretapper use classical-quantum channels and the set of the channel states may be finite or infinite.

Let $A$, $H$, $H'$, $H''$, $\theta$, and $\{(W_t, V_t) : t \in \theta\}$ be defined as in Section 2.

**Theorem 3.7.** *The secrecy capacity of the compound classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$ in the case with CSI is given by*

$$C_{S,CSI}(\{(W_t, V_t) : t \in \theta\}) = \lim_{n \to \infty} \min_{t \in \theta} \max_{P_{inp}, w_t} \frac{1}{n} (\chi(P_{inp}; B_t^n) - \chi(P_{inp}; Z_t^n)) \tag{56}$$

*where $B_t$ are the resulting random quantum states at the output of legal receiver channels and $Z_t$ are the resulting random quantum states at the output of wiretap channels. The maximum is taken over all probability distributions $P_{inp}$ on the input quantum states $w_t$.*

*Assume that the sender's encoding is restricted to transmitting an indexed finite set of orthogonal quantum states $\{\rho_x : x \in A\} \subset \mathcal{S}(H'^{\otimes n})$, then the secrecy capacity of the compound classical-quantum wiretap channel in the case with no CSI at the encoder is given by*

$$C_S(\{(W_t, V_t) : t \in \theta\}) = \lim_{n \to \infty} \max_{U \to A \to (BZ)_t} \frac{1}{n} \left( \min_{t \in \theta} \chi(p_U; B_t^n) - \max_{t \in \theta} \chi(p_U; Z_t^n) \right) . \tag{57}$$

*Proof.* At first we are going to prove (56). Our idea is to send the information in two parts. First, we send the channel state information with finite blocks of finite bits with a code $C_1$ to the receiver, and then, depending on $t$, we send the message with a code $C_2^{(t)}$ in the second part.

*i.1) Sending Channel State Information with Finite Bits*

We do not require that the first part should be secure against the wiretapper, since we assume that the wiretapper already has the full knowledge of the CSI.

By ignoring the security against the wiretapper, we consider only the compound channel $(W_t)_{t \in \theta}$. Let $W = (W_t)_t$ be a compound classical-quantum channel. Then, by [40] and [51], for each $\lambda \in (0, 1)$, the $\lambda$ capacity $C(W, \lambda)$ equals

$$C(W, \lambda) = \max_{P_{inp} \in P(\mathbf{A})} \min_t \chi(P_{inp}; W_t) . \tag{58}$$

If $\max_{P_{inp}} \min_t \chi(P_{inp}; W_t) > 0$ holds, then the sender can build a code $C_1$ such that the CSI can be sent to the legal receiver with a block with length $l \leq \frac{\log T}{\min_t \max_{P_{inp}} \chi(P_{inp}, W_t)} -$

$\epsilon$. Here $T < \infty$ is the size of $\theta$, as we defined in Section 2. If $\max_{P_{inp}} \min_t \chi(P_{inp}; W_t) = 0$ holds, we cannot build a code $C_1$ such that the CSI can be sent to the legal receiver. But, this does not cause any problem, since when for every $t \in \theta$ there is a $P_{inp}$ such that $\chi(P_{inp}; W_t) > 0$ then for every $t$ there are $x_1^{(t)}$ and $x_a^{(t)} \in A$ such that $W_t(x_1^{(t)}) \neq W_t(x_2^{(t)})$. In this case $\min_t \chi(P_{uni}; W_t) > 0$, where $P_{inp}$ is the uniform distribution over $A$. This means that if $\max_{P_{inp}} \min_t \chi(P_{inp}; W_t) = 0$, the right-hand side of (56) is also zero.

*i.2) Message Transformation When Both the Sender and the Legal Receiver Know CSI*

If both the sender and the legal receiver have the full knowledge of $t$, then we only have to look at the single wiretap channel $(W_t, V_t)$.

In [31] and [34] it was shown that if $n$ is sufficiently large, there exists an $(n, J_n)$ code for the quantum wiretap channel $(W, V)$ with

$$\log J_n = \max_{P_{inp}, w} (\chi(P_{inp}; B^n) - \chi(P_{inp}; Z^n)) - \epsilon \,, \tag{59}$$

for any positive $\epsilon$ and positive $\delta$, where $B$ is the resulting random variable at the output of legal receiver's channel and $Z$ the output of the wiretap channel.

When the sender and the legal receiver both know $t$, they can build an $(n, J_{n,t})$ code $C_2^{(t)}$ where

$$\log J_{n,t} = \max_{P_{inp}, w_t} (\chi(P_{inp}; B_t^n) - \chi(P_{inp}; Z_t^n)) - \epsilon \,. \tag{60}$$

Thus,

$$C_{S,CSI}(\{(W_t, V_t) : t \in \theta\}) \geq \lim_{n\to\infty} \min_{t\in\theta} \max_{P_{inp}, w_t} \frac{1}{n} (\chi(P_{inp}; B_t^n) - \chi(P_{inp}; Z_t^n)) \,. \tag{61}$$

**Remark 3.8.** *For the construction of the second part of our code, we use random coding and request that the randomization can be sent (cf. [31]). However, it was shown in [18] that the randomization could not always be sent if we require that we use one unique code which is secure against the wiretapper and suitable for every channel state, i.e. it does not depend on t. This is not a counterexample to our results above, neither to the construction of $C_1$ nor to the construction of $C_2^{(t)}$, because of the following facts.*

*The first part of our code does not need to be secure. For our second part, the legal transmitters can use the following strategy: At first they build a code $C_1 = (E, \{D_t : t = 1, \ldots, |\theta|\})$ and a code $C_2^{(t)} = (E^{(t)}, \{D_j^{(t)} : j = 1, \ldots, J_n\})$ for every $t \in \theta$. If the sender wants to send the CSI $t' \in \theta$ and the message $j$, he encodes $t'$ with $E$ and $j$ with $E^{(t')}$, then he sends both parts together through the channel. After receiving both parts, the legal receiver decodes the first part with $\{D_t : t\}$, and chooses the right decoders $\{D_j^{(t')} : j\} \in \left\{\{D_j^{(t)} : j\} : t \in \theta\right\}$ to decode the second part. With this strategy, we can avoid using one unique code which is suitable for every channel state.*

*i.3) Upper Bound for the Case CSI at the Encoder*

For any $\epsilon > 0$, we choose $t' \in \theta$ such that $C_{S,CSI}(W_{t'}, V_{t'}) \leq \inf_{t\in\theta} C_{S,CSI}(W_t, V_t) + \epsilon$.

From [31] and [34] we know that the secrecy capacity of the quantum wiretap channel $(W_{t'}, V_{t'})$ cannot be greater than

$$\lim_{n\to\infty} \max_{P_{inp},w_{t'}} \frac{1}{n}(\chi(P_{inp}; B^n_{t'}) - \chi(P_{inp}; Z^n_{t'}))\,.$$

Since we cannot exceed the capacity of the worst wiretap channel, we have

$$C_{S,CSI}(\{(W_t, V_t) : t \in \theta\}) \leq \lim_{n\to\infty} \min_{t\in\theta} \max_{P_{inp},w_t} \frac{1}{n}(\chi(P_{inp}; B^n_t) - \chi(P_{inp}; Z^n_t)). \tag{62}$$

This together with (61) completes the proof of (56).

**Remark 3.9.** *In [64] it was shown that if for a given $t$ and any $n \in \mathbb{N}$,*

$$\chi(P_{inp}; B^n_t) \geq \chi(P_{inp}; Z^n_t)$$

*holds for all $P_{inp} \in P(\mathbf{A})$ and $\{w_t(j) : j = 1, \ldots, J_n\} \subset S(H^{\otimes n})$, then*

$$\begin{aligned}&\lim_{n\to\infty} \max_{P_{inp},w_t} \frac{1}{n}(\chi(P_{inp}; B^n_t) - \chi(P_{inp}; Z^n_t))\\ &= \max_{P_{inp},w_t} (\chi(P_{inp}; B_t) - \chi(P_{inp}; Z_t))\,.\end{aligned}$$

*Thus if for every $t \in \theta$ and $n \in \mathbb{N}$,*

$$\chi(P_{inp}, B^n_t) \geq \chi(P_{inp}; Z^n_t)$$

*holds for all $P_{inp} \in P(\mathbf{A})$ and $\{w_t(j) : j = 1, \ldots, J_n\} \subset S(H^{\otimes n})$, we have*

$$C_{S,CSI} = \min_{t\in\theta} \max_{P_{inp},w_t} (\chi(P_{inp}; B_t) - \chi(P_{inp}; Z_t))\,.$$

Now we are going to prove (57).

*ii.1) Lower Bound for Case Without CSI*

Fix a probability distribution $p$ on $\mathbf{A}^n$. Let

$$J_n = \lfloor 2^{\min_{t\in\theta} \chi(p;B^n_t) - \max_{t\in\theta} \chi(p;Z^n_t) - 2n\mu} \rfloor\,,$$

$$L_n = \lceil 2^{\max_t \chi(p;Z^n_t) + n\mu} \rceil\,,$$

and let $p'$ and $X^n = \{X_{j,l} : j, l\}$ be defined as in Section 3.1. Since $J_n \cdot L_n \leq 2^{\min_t \chi(p;B^n_t) - n\mu}$, in [40] and [51] it was shown that if $n$ is sufficiently large, there exist a collection of quantum states $\{\rho_{x^n} : x^n \in \mathbf{A}^n\} \subset \mathcal{S}(H'^{\otimes n})$, a collection of positive-semidefinite operators $\{D_{x^n} : x^n \in \mathbf{A}^n\}$, and a positive constant $\beta$, such that for any $(j, l) \in \{1, \ldots, J_n\} \times \{1, \ldots, L_n\}$ it holds

$$Pr\left[\mathrm{tr}\left(W^n_t(\rho^n_{X_{j,l}})D_{X_{j,l}}\right) \geq 1 - 2^{-n\beta}\right] > 1 - 2^{-n\beta}\,, \tag{63}$$

and for any realization $\{x_{j,l} : j, l\}$ of $\{X_{j,l} : j, l\}$ it holds that

$$\sum_{t\in\theta} \sum_{j=1}^{J_n} \sum_{l=1}^{L_n} D_{x_{j,l}} \leq \mathrm{id}\,.$$

We define

$$Q_t(\rho_{x^n}) := \Pi_{pV_t,\alpha}\Pi_{V_t,\alpha}(x^n)\cdot V_t^{\otimes n}(\rho_{x^n})\cdot\Pi_{V_t,\alpha}(x^n)\Pi_{pV_t,\alpha}\,,$$

and $\Theta_t := \sum_{x^n\in\mathcal{T}^n_{p,\delta}} p'^n(x^n)Q_t(\rho_{x^n})$.

Choosing $n$ sufficiently large, in the same way as our proof of (37) for the classical compound channel with quantum wiretapper, there is a positive constant $\upsilon$ such that

$$Pr\left(\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_t(\rho_{X^{(t)}_{j,l}})-\Theta_t\|\le\epsilon\ \forall t\ \forall j\right)\ge 1-2^{-n\upsilon}\,. \tag{64}$$

We choose a suitable $\alpha$. If $n$ is sufficiently large, we can find a realization $x_{j,l}$ of $X_{j,l}$ with a positive probability such that for all $j\in\{1,\dots J_n\}$, we have

$$\min_{t\in\theta}\mathrm{tr}\left(W_t^n(\rho^n_{x_{j,l}})D_{x_{j,l}}\right)\ge 1-2^{-n\beta}$$

and

$$\max_{t\in\theta}\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_t(\rho_{x_{j,l}})-\Theta_t\|\le\epsilon\,.$$

We define $D_j := \sum_{t\in\theta}\sum_{l=1}^{L_n}D_{x_{j,l}}$, then $\sum_{j=1}^{J_n}D_j=\sum_{t\in\theta}\sum_{j=1}^{J_n}\sum_{l=1}^{L_n}D_{x_{j,l}}\le$ id. Furthermore, for all $t'\in\theta$ and $l'\in\{1,\dots,L_n\}$ we have

$$\begin{aligned}
&\mathrm{tr}\left(W_{t'}^n(\rho^n_{x_{j,l'}})D_j\right)\\
&=\sum_{t\in\theta}\sum_{l=1}^{L_n}\mathrm{tr}\left(W_{t'}^n(\rho^n_{x_{j,l'}})D_{x_{j,l}}\right)\\
&\ge\mathrm{tr}\left(W_{t'}^n(\rho^n_{x_{j,l'}})D_{x_{j,l'}}\right)\\
&\ge 1-2^{-n\beta}\,,
\end{aligned}$$

the inequality in the third line holds because for two positive-semidefinite matrices $M_1$ and $M_2$, we always have $\mathrm{tr}\,(M_1M_2)=\mathrm{tr}\left(\sqrt{M_1}M_2\sqrt{M_1}\right)\ge 0$.

For any $\gamma>0$ let

$$R:=\max_{U\to A\to(BZ)_t}\frac{1}{n}\left[\min_{t\in\theta}\chi(p;B_t^n)-\max_{t\in\theta}\chi(p;Z_t^n)\right]-\gamma\,.$$

Then for any positive $\lambda$, there is an $(n,J_n,\lambda)$ code $\left(\{w(j):=\sum_{l=1}^{L_n}\frac{1}{L_n}\rho^n_{x_{j,l}}:j=1,\dots,J_n,\},\{D_j:j=1,\dots,J_n\}\right)$, such that $\liminf_{n\to\infty}\frac{1}{n}\log J_n\ge R$,

$$\max_{t\in\theta}\max_{j\in\{1,\dots,J_n\}}\mathrm{tr}\left((\mathrm{id}_{H''^{\otimes n}}-D_j)W_t^{\otimes n}(w(j))\right)\le\lambda\,, \tag{65}$$

and in the same way as our proof of (45) for the classical compound channel with quantum wiretapper,

$$\max_{t\in\theta}\chi(R_{uni};Z_t^n)\le\lambda\,, \tag{66}$$

for any uniformly distributed random variable $R_{uni}$ with value in $\{1,\dots,J_n\}$.

Combining (65) and (66) we obtain

$$C_S \geq \lim_{n\to\infty} \max_{U\to A\to (BZ)_t} \frac{1}{n}\left(\min_{t\in\theta}\chi(p_U;B_t^n)-\max_{t\in\theta}\chi(p_U;Z_t^n)\right). \tag{67}$$

*ii.2) Upper Bound for Case Without CSI*

Let $(\mathcal{C}_n)=(\{\rho_j^{(n)}:j\},\{D_j^{(n)}:j\})$ be a sequence of $(n,J_n,\lambda_n)$ code such that

$$\max_{t\in\theta}\max_{j\in\{1,\dots,J_n\}} \operatorname{tr}\left((\mathrm{id}-D_j^{(n)})W_t^{\otimes n}\left(\rho_j^{(n)}\right)\right)\leq\lambda_n\,, \tag{68}$$

$$\max_{t\in\theta}\chi(R_{uni};Z_t^n)=:\epsilon_{2,n}\,, \tag{69}$$

where $\lim_{n\to\infty}\lambda_n=0$ and $\lim_{n\to\infty}\epsilon_{2,n}=0$. $R_{uni}$ denotes the random variable which is uniformly distributed on the message set $\{1,\dots,J_n\}$.

We denote the classical capacity of the quantum channel $W_t$ in the sense of [66] by $C(W_t)$. Choose $t'\in\theta$ such that $C(W_{t'})=\min_{t\in\theta}C(W_t)$.

It is known (cf. Section 3.1 *ii)*and [52]) that $C(W_{t'})$ cannot exceed $\chi(R_{uni};B_{t'}^n)+\xi$ for any constant $\xi>0$. Since the secrecy capacity of a compound wiretap channel cannot exceed the capacity of the worst channel without wiretapper, for any $\epsilon>0$ choose $\xi=\frac{1}{2}\epsilon$, if $n$ is large enough, the secrecy rate of $(\mathcal{C}_n)$ cannot be greater than

$$\begin{aligned}
&\frac{1}{n}\chi(R_{uni};B_{t'}^n)+\xi\\
&=\min_{t\in\theta}\frac{1}{n}\chi(R_{uni};B_t^n)+\xi\\
&\leq\min_{t\in\theta}\frac{1}{n}\chi(R_{uni};B_t^n)-\max_{t\in\theta}\frac{1}{n}\chi(R_{uni};Z_t^n)+\xi+\frac{1}{n}\epsilon_{2,n}\\
&\leq\frac{1}{n}\left(\min_{t\in\theta}\chi(R_{uni};B_t^n)-\max_{t\in\theta}\chi(R_{uni};Z_t^n)\right)+\epsilon\,.
\end{aligned} \tag{70}$$

Thus

$$C_S\leq\lim_{n\to\infty}\max_{U\to A\to(BZ)_t}\frac{1}{n}\left(\min_{t\in\theta}\chi(p_U;B_t^n)-\max_{t\in\theta}\chi(p_U;Z_t^n)\right). \tag{71}$$

□

Combining (71) and (67) we obtain (57).

So far, we assumed that $|\theta|$, the number of the channels, is finite, therefore we can send the CSI with finite bits to the receiver in the case where the sender has CSI. Now we look at the case where $|\theta|$ can be arbitrary. We of course are not allowed to send the CSI with finite bits if $|\theta|=\infty$, but in this case, we may use a "finite approximation" to obtain the following corollary.

**Corollary 3.10.** *For an arbitrary set $\theta$ we have*

$$C_{S,CSI}(\{(W_t,V_t):t\in\theta\})=\lim_{n\to\infty}\inf_{t\in\theta}\max_{P_{inp},w_t}\frac{1}{n}(\chi(P_{inp};B_t^n)-\chi(P_{inp};Z_t^n))\,. \tag{72}$$

*Proof.* Let $W : \mathcal{S}(H') \to \mathcal{S}(H'')$ be a linear map, then let

$$\|W\|_\diamond := \sup_{n\in\mathbb{N}} \max_{a\in S(\mathbb{C}^n\otimes H'),\|a\|_1=1} \|(\mathrm{id}_n \otimes W)(a)\|_1 \,. \tag{73}$$

It is known [56] that this norm is multiplicative, i.e. $\|W\otimes W'\|_\diamond = \|W\|_\diamond \cdot \|W'\|_\diamond$.

A $\tau$-net in the space of the completely positive trace-preserving maps $\mathcal{S}(H') \to \mathcal{S}(H'')$ is a finite set $\left(W^{(k)}\right)_{k=1}^K$ of completely positive trace-preserving maps $\mathcal{S}(H') \to \mathcal{S}(H'')$ with the property that for each completely positive trace-preserving map $W : \mathcal{S}(H') \to \mathcal{S}(H'')$, there is at least one $k \in \{1,\dots,K\}$ with $\|W - W^{(k)}\|_\diamond < \tau$.

**Lemma 3.11 ($\tau-$net [50]).** *Let $H'$ and $H''$ be finite-dimensional complex Hilbert spaces. For any $\tau \in (0,1]$, there is a $\tau$-net of quantum channels $\left(W^{(k)}\right)_{k=1}^K$ in the space of the completely positive trace-preserving maps $\mathcal{S}(H') \to \mathcal{S}(H'')$ with $K \le (\frac{3}{\tau})^{2d'^4}$, where $d' = \dim H'$.*

If $|\theta|$ is arbitrary, then for any $\xi > 0$ let $\tau = \frac{\xi}{-\log\xi}$. By Lemma 3.11 there exists a finite set $\theta'$ with $|\theta'| \le (\frac{3}{\tau})^{2d'^4}$ and $\tau$-nets $(W_{t'})_{t'\in\theta'}$, $(V_{t'})_{t'\in\theta'}$ such that for every $t \in \theta$ we can find a $t' \in \theta'$ with $\|W_t - W_{t'}\|_\diamond \le \tau$ and $\|V_t - V_{t'}\|_\diamond \le \tau$. For every $t' \in \theta'$, the legal transmitters build a code $C_2^{(t')} = \{w_{t'}, \{D_j : j\}\}$. Since by [31], the error probability of the code $C_2^{(t')}$ decreases exponentially with its length, there is an $N = O(-\log\xi)$ such that for all $t'' \in \theta'$ it holds

$$\frac{1}{J_N}\sum_{j=1}^{J_N} \mathrm{tr}\left(W_{t''}^{\otimes N}\left(w_{t''}(j)\right) D_j\right) \ge 1 - \lambda - \xi \,, \tag{74}$$

$$\chi(R_{uni}; Z_{t'}^N) \le \xi \,. \tag{75}$$

Then, if the sender obtains the channel state information "$t$" , he chooses a "$t'$" $\in \theta'$ such that $\|W_t - W_{t'}\|_\diamond \le \tau$ and $\|V_t - V_{t'}\|_\diamond \le \tau$. He can send "$t'$" to the legal receiver in the first part with finite bits, and then they build a code $C_2^{(t')}$ that fulfills (74) and (75) to transmit the message.

For every $t'$ and $j$ let $|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)| \in \mathcal{S}(H'^{\otimes N}\otimes H'^{\otimes N})$ be an arbitrary purification of the quantum state $w_{t'}(j)$, then $\mathrm{tr}\left[\left(W_t^{\otimes N} - W_{t'}^{\otimes N}\right)\left(w_{t'}(j)\right)\right] = \mathrm{tr}\Bigg(\mathrm{tr}_{H'^{\otimes N}}\Big[\mathrm{id}_{H'}^{\otimes N}\otimes \left(W_t^{\otimes N} - W_{t'}^{\otimes N}\right)\left(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|\right)\Big]\Bigg)$. We have

$$\begin{aligned}
&\mathrm{tr}\left|\left(W_t^{\otimes N} - W_{t'}^{\otimes N}\right)\left(w_{t'}(j)\right)\right| \\
&= \mathrm{tr}\left(\mathrm{tr}_{H'^{\otimes N}}\left|\mathrm{id}_{H'}^{\otimes N}\otimes\left(W_t^{\otimes N} - W_{t'}^{\otimes N}\right)\left(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|\right)\right|\right) \\
&= \mathrm{tr}\left|\mathrm{id}_{H'}^{\otimes N}\otimes\left(W_t^{\otimes n} - W_{t'}^{\otimes N}\right)\left(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|\right)\right| \\
&= \left\|\mathrm{id}_{H'}^{\otimes N}\otimes\left(W_t^{\otimes N} - W_{t'}^{\otimes N}\right)\left(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|\right)\right\|_1 \\
&\le \|W_t^{\otimes N} - W_{t'}^{\otimes N}\|_\diamond \cdot \|\left(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|\right)\|_1 \\
&\le N\tau \,.
\end{aligned}$$

The second equality follows from the definition of trace. The second inequality follows by the definition of $\|\cdot\|_\diamond$. The third inequality follows from the facts that $\|\left(|\psi_{t'}(j)\rangle\langle\psi_{t'}(j)|\right)\|_1 =$

1 and $\left\|W_t^{\otimes N} - W_{t'}^{\otimes N}\right\|_\diamond = \left\|\sum_{k=1}^N W_t^{\otimes k-1} W_{t'}^{\otimes N-k} \left(W_t - W_{t'}\right)\right\|_\diamond = N \cdot \|W_t - W_{t'}\|_\diamond$, since $\|\cdot\|_\diamond$ is multiplicative and $\|W_t\|_\diamond = \|W_{t'}\|_\diamond = 1$.

It follows that

$$\begin{aligned}
&\left| \frac{1}{J_N} \sum_{j=1}^{J_N} \operatorname{tr}\left(W_t^{\otimes N}\left(w_{t'}(j)\right) D_j\right) \right. \\
&\left. - \frac{1}{J_N} \sum_{j=1}^{J_N} \operatorname{tr}\left(W_{t'}^{\otimes N}\left(w_{t'}(j)\right) D_j\right) \right| \\
&\leq \frac{1}{J_N} \sum_{j=1}^{J_N} \left| \operatorname{tr}\left[\left(W_t^{\otimes N} - W_{t'}^{\otimes N}\right)\left(w_{t'}(j)\right) D_j\right] \right| \\
&\leq \frac{1}{J_N} \sum_{j=1}^{J_N} \operatorname{tr}\left| \left(W_t^{\otimes N} - W_{t'}^{\otimes N}\right)\left(w_{t'}(j)\right) D_j \right| \\
&\leq \frac{1}{J_N} \sum_{j=1}^{J_N} \operatorname{tr}\left| \left(W_t^{\otimes N} - W_{t'}^{\otimes N}\right)\left(w_{t'}(j)\right) \right| \\
&\leq \frac{1}{J_N} J_N N\tau \\
&= N\tau \,. \qquad (76)
\end{aligned}$$

$N\tau$ can be arbitrarily small when $\xi$ is close to zero, since $N = O(-\log \xi)$.

Let $R_{uni}$ be a random variable uniformly distributed on $\{1, \ldots, J_N\}$, and $\{\rho(j) : j = 1, \ldots, J_n\}$ be a set of quantum states labeled by elements of $\{1, \ldots, J_n\}$. We have

$$\begin{aligned}
&|\chi(R_{uni}; V_t) - \chi(R_{uni}; V_{t'})| \\
&\leq \left| S\left(\sum_{j=1}^{J_N} \frac{1}{J_N} V_t(\rho(j))\right) - S\left(\sum_{j=1}^{J_N} \frac{1}{J_N} V_{t'}(\rho(j))\right) \right| \\
&+ \left| \sum_{j=1}^{J_N} \frac{1}{J_N} S\left(V_t(\rho(j))\right) - \sum_{j=1}^{J_N} \frac{1}{J_N} S\left(V_{t'}(\rho(j))\right) \right| \\
&\leq \tau \log(d-1) - \tau \log \tau - (1-\tau)\log(1-\tau) \,, \qquad (77)
\end{aligned}$$

where $d = \dim H$. The inequality in the last line holds by Lemma 3.6 and because $\|V_t(\rho) - V_{t'}(\rho)\|_1 \leq \tau$ for all $\rho \in \mathcal{S}(H)$ when $\|V_t - V_{t'}\|_\diamond \leq \tau$.

By (76) and (77) we have

$$\sup_{t \in \theta} \frac{1}{J_N} \sum_{j=1}^{J_N} \operatorname{tr}\left(W_t^{\otimes N}\left(w_{t'}(j)\right) D_j\right) \geq 1 - \lambda - \xi - N\tau \,,$$

$$\chi(R_{uni}; Z_t^N) \leq \xi + \tau \log(d-1) - \tau \log \tau - (1-\tau)\log(1-\tau) \,.$$

Since $\xi + N\tau$ and $\tau \log(d-1)$ can be arbitrarily small, when $\xi$ is close to zero, we have

$$\sup_{t \in \theta} \frac{1}{J_N} \sum_{j=1}^{J_N} \operatorname{tr}\left(W_t^{\otimes N}\left(w_{t'}(j)\right) D_j\right) \geq 1 - \lambda \,,$$

$$\sup_{t\in\theta} \chi(R_{uni}; Z_t^N) \leq \epsilon \, .$$

The bits that the sender uses to transform the CSI are large but constant, so it is still negligible compared to the second part. We obtain

$$C_{S,CSI}(\{(W_t, V_t) : t \in \theta\}) \geq \lim_{n\to\infty} \inf_{t\in\theta} \max_{P_{inp}, w_t} \frac{1}{n}(\chi(P_{inp}; B_t^n) - \chi(P_{inp}; Z_t^n)) \, . \tag{78}$$

The proof of the converse is similar to those given in the proof of Theorem 3.7, where we consider a worst $t'$. □

**Remark 3.12.** *In (56) and Corollary 3.10 we have only required that the legal receiver can decode the correct message with a high probability if $n$ is sufficiently large. We have not specified how fast the error probability tends to zero when the code length goes to infinity. If we analyze the relation between the error probability $\varepsilon$ and the code length, then we have the following facts.*

*In the case of finite $\theta$, let $\varepsilon_1$ denote the error probability of the first part of the code (i.e. the legal receiver does not decode the correct CSI), and let $\varepsilon_2$ denote the error probability of the second part of the code (i.e. the legal receiver decodes the correct CSI, but does not decode the message). Since the length of the first part of the code is $l \cdot \log c \cdot c' = O(\log \varepsilon_1)$, we have $\varepsilon_1^{-1}$ is $O(\exp(l \cdot \log c \cdot c')) = O(\exp(n))$, where $n$ stands for the length of the first part of the code. For the second part of the code, $\varepsilon_2$ decreased exponentially with the length of the second part, as proven in [31]. Thus, the error probability $\varepsilon = \max\{\varepsilon_1, \varepsilon_2\}$ decreases exponentially with the code length in the case of finite $\theta$.*

*If $\theta$ is infinite, let $\varepsilon_1$ denote the error probability of the first part of the code probability. Here we have to build two $\tau$-nets for a suitable $\tau$, each contains $O((\frac{-\log \varepsilon_1}{\varepsilon_1})^{-2d'^4})$ channels. If we want to send the CSI of these $\tau$-nets, the length of first part $l$ will be $O(-2d'^4 \cdot \log(\varepsilon_1 \log \varepsilon_1))$, which means here $\varepsilon_1^{-1}$ will be $O(\exp(\frac{n}{4d'^4})) = O(\exp(n))$. Thus we can still achieve that the error probability decreases exponentially with the code length in case of infinite $\theta$.*

## 3.3 Entanglement Generation over Compound Quantum Channels

The entanglement generating capacity of a given quantum channel describes the maximal amount of entanglement that we can generate or transmit over the channel. A code for the secure message transmission over a classical-quantum wiretap channel can be used to build a code for the entanglement transmission over a quantum channel (cf. [34]). Our technique for entanglement generation over compound quantum channels is similar to the proof of entanglement generating capacity over quantum channels in [34]. The difference between our technique and the proofs in [34] is that we have to consider the channel uncertainty (c.f. the discussion in Section 3.4).

Let $\mathfrak{P}$, $\mathfrak{Q}$, $H^{\mathfrak{P}}$, $H^{\mathfrak{Q}}$, $\theta$, and $\{\left(N_t^{\otimes n}\right) : t \in \theta\}$ be defined as in Section 2 (i.e. we assume that $\theta$ is finite).

We denote $\dim H^{\mathfrak{P}}$ by $a$, and denote $\mathcal{X} := \{1, \ldots, a\}$. Consider the eigen-decomposition of $\rho^{\mathfrak{P}}$ into the orthonormal pure quantum state ensemble $\{p(x), |\phi_x\rangle^{\mathfrak{P}} : x \in \mathcal{X}\}$,

$$\sum_{x\in\mathcal{X}} p(x)|\phi_x\rangle\langle\phi_x|^{\mathfrak{P}} = \rho^{\mathfrak{P}} \, .$$

The distribution $p$ defines a random variable $X$.

**Theorem 3.13.** *The entanglement generating capacity of $\{(N_t) : t \in \theta\}$ is bounded as follows*

$$A(\{(N_t) : t \in \theta\}) \geq \max_p \left( \min_{t\in\theta} \chi(p; Q_t) - \max_{t\in\theta} \chi(p; E_t) \right) , \tag{79}$$

*where $Q_t$ stands for the quantum outputs that the receiver observes at the channel state $t$, and $E_t$ the quantum outputs at the environment.*

(Theorem 3.13 is weaker than the result in [17], the reason is that we use for our proof a different quantum channel representation. For details and the result in [17] cf. Section 3.4.)

*Proof.* Let $\rho^{\mathfrak{P}} \to U_{N_t} \rho^{\mathfrak{P}} U_{N_t}^*$ be a isometric transformation which represents $N_t$ (cf. Section 3.4), where $U_{N_t}$ is a linear operator $\mathcal{S}(H^{\mathfrak{P}}) \to \mathcal{S}(H^{\mathfrak{O}\mathfrak{E}})$, and $\mathfrak{E}$ is the quantum system of the environment. Fix a $\rho^{\mathfrak{P}}$ with eigen-decomposition $\sum_{x\in\mathcal{X}} p(x)|\phi_x\rangle^{\mathfrak{P}}\langle\phi_x|^{\mathfrak{P}}$. If the channel state is $t$, the local output density matrix seen by the receiver is

$$\mathrm{tr}_{\mathfrak{E}} \left( \sum_x p(x) U_{N_t} |\phi_x\rangle\langle\phi_x|^{\mathfrak{P}} U_{N_t}^* \right) ,$$

and the local output density matrix seen by the environment (which we interpret as the wiretapper) is

$$\mathrm{tr}_{\mathfrak{O}} \left( \sum_x p(x) U_{N_t} |\phi_x\rangle\langle\phi_x|^{\mathfrak{P}} U_{N_t}^* \right) .$$

Therefore $(N_t)_{t\in\theta}$ defines a quantum compound wiretap channel $(W_{N_t}, V_{N_t})_{t\in\theta}$, where $W_{N_t} : H^{\mathfrak{P}} \to H^{\mathfrak{O}}$, $\sum_{x\in\mathcal{X}} p(x)|\phi_x\rangle\langle\phi_x|^{\mathfrak{P}} \to \mathrm{tr}_{\mathfrak{E}} \left( \sum_x p(x) U_{N_t} |\phi_x\rangle\langle\phi_x|^{\mathfrak{P}} U_{N_t}^* \right)$, and $V_{N_t} : H^{\mathfrak{P}} \to H^{\mathfrak{O}}$, $\sum_{x\in\mathcal{X}} p(x)|\phi_x\rangle\langle\phi_x|^{\mathfrak{P}} \to \mathrm{tr}_{\mathfrak{E}} \left( \sum_x p(x) U_{N_t} |\phi_x\rangle\langle\phi_x|^{\mathfrak{P}} U_{N_t}^* \right)$.

*i) Building the Encoder and the First Part of the Decoding Operator*

Let

$$J_n = \lceil 2^{n[\min_t \chi(X;Q_t) - \max_t \chi(X;E_t) - 2\delta]} \rceil ,$$

and

$$L_n = \lceil 2^{n(\max_t \chi(X;E_t) + \delta)} \rceil .$$

For the compound classical-quantum wiretap channel $(W_{N_t}, V_{N_t})_{t\in\theta}$, since

$$\begin{aligned} &|\{(j,l) : j = 1, \ldots, J_n, l = 1, \ldots, L_n\}| \\ &= J_n \cdot L_n \leq 2^{n \min_t [\chi(X;Q_t) - \delta]} , \end{aligned}$$

if $n$ is large enough, by Theorem 3.7, [40], and [51], the following holds. There is a collection of quantum states $\{\rho_{x_{j,l}}^{\mathfrak{P}^n} : j = 1, \ldots, J_n, l = 1, \ldots, L_n\} \subset \mathcal{S}(H^{\mathfrak{P}^n})$, a collection of positive-semidefinite operators $\{D_{j,l} := D_{x_{j,l}} : t \in \theta, j = 1, \ldots, J_n, l = 1, \ldots, L_n\}$, a positive constant $\beta$, and a quantum state $\xi_t^{\mathfrak{E}^n}$ on $H^{\mathfrak{E}^n}$, such that

$$\mathrm{tr} \left( (D_{x_{j,l}}^{\mathfrak{O}^n} \otimes \mathrm{id}^{\mathfrak{E}^n}) U_{N_t} \rho_{x_{j,l}}^{\mathfrak{P}^n} U_{N_t}^* \right) \geq 1 - 2^{-n\beta} , \tag{80}$$

and

$$\| \omega_{j,t}^{\mathfrak{E}^n} - \xi_t^{\mathfrak{E}^n} \|_1 < \epsilon , \tag{81}$$

where $\omega_{j,t}^{\mathfrak{E}^n} := \frac{1}{L_n} \sum_{l=1}^{L_n} \mathrm{tr}_{\mathfrak{O}^n} \left( U_{N_t} \rho_{x_{j,l}}^{\mathfrak{P}^n} U_{N_t}^* \right)$.

Now the quantum state $\rho_{x_{j,l}}^{\mathfrak{P}^n}$ may be pure or mixed. Assume $\rho_{x_{j,l}}^{\mathfrak{P}^n}$ is a mixed quantum state $\sum_{i=1}^n {p'}_{j,l}(i)|\aleph_{x_{j,l}}^{(i)}\rangle\langle\aleph_{x_{j,l}}^{(i)}|^{\mathfrak{P}^n}$, then

$$\sum_{i=1}^n {p'}_{j,l}(i)\mathrm{tr}\left((D_{x_{j,l}}^{\mathfrak{Q}^n}\otimes \mathrm{id}^{\mathfrak{E}^n})U_{N_t}|\aleph_{x_{j,l}}^{(i)}\rangle\langle\aleph_{x_{j,l}}^{(i)}|^{\mathfrak{P}^n}U_{N_t}^*\right)$$
$$\mathrm{tr}\left((D_{x_{j,l}}^{\mathfrak{Q}^n}\otimes \mathrm{id}^{\mathfrak{E}^n})U_{N_t}(\sum_{i=1}^n {p'}_{j,l}(i)|\aleph_{x_{j,l}}^{(i)}\rangle\langle\aleph_{x_{j,l}}^{(i)}|^{\mathfrak{P}^n})U_{N_t}^*\right)$$
$$\geq 1-2^{-n\beta}\,.$$

Thus, for all $i$ such that ${p'}_{j,l}(i)\geq\frac{2^{-n\beta}}{1-2^{-n\beta}}$ it must hold

$$\mathrm{tr}\left((D_{x_{j,l}}^{\mathfrak{Q}^n}\otimes \mathrm{id}^{\mathfrak{E}^n})U_{N_t}|\aleph_{x_{j,l}}^{(i)}\rangle\langle\aleph_{x_{j,l}}^{(i)}|^{\mathfrak{P}^n}U_{N_t}^*\right)\geq 1-2^{-n\beta}\,.$$

If $n$ is large enough, then there is at least one $i_{l,j}\in\{1,\ldots,n\}$ such that ${p'}_{j,l}(i_{l,j})\geq\frac{2^{-n\beta}}{1-2^{-n\beta}}$. By Theorem 3.7, there is a $\xi_t^{\mathfrak{E}^n}$ on $H^{\mathfrak{E}^n}$, such that

$$\|\frac{1}{L_n}\sum_{l=1}^{L_n}\mathrm{tr}_{\mathfrak{Q}^n}\left(U_{N_t}|\aleph_{x_{j,l}}^{(i_{l,j})}\rangle\langle\aleph_{x_{j,l}}^{(i_{l,j})}|^{\mathfrak{P}^n}U_{N_t}^*\right)-\xi_t^{\mathfrak{E}^n}\|_1<\epsilon\,.$$

Thus,

$$\left(\{|\aleph_{x_{j,l}}^{(i_{l,j})}\rangle\langle\aleph_{x_{j,l}}^{(i_{l,j})}|^{\mathfrak{P}^n}:j,l\},\{D_{x_{j,l}}^{\mathfrak{Q}^n}:j,l,t\}\right)$$

is a code with the same security rate as

$$\left(\{\rho_{x_{j,l}}^{\mathfrak{P}^n}:j,l\},\{D_{x_{j,l}}^{\mathfrak{Q}^n}:j,l,t\}\right)\,.$$

When some $\rho_{x_{j,l}}^{\mathfrak{P}^n}$ are mixed quantum states we may replace them with $|\aleph_{x_{j,l}}^{(i_{l,j})}\rangle\langle\aleph_{x_{j,l}}^{(i_{l,j})}|^{\mathfrak{P}^n}$. Hence we may assume that every $\rho_{x_{j,l}}^{\mathfrak{P}^n}$ is a pure quantum state.

Assume $\rho_{x_{j,l}}^{\mathfrak{P}^n}=|\aleph_{j,l}\rangle\langle\aleph_{j,l}|^{\mathfrak{P}^n}$. Let $H^{\mathfrak{M}}$ be a $J_n$-dimensional Hilbert space with an orthonormal basis $\{|j\rangle^{\mathfrak{M}}:j=1,\ldots,J_n\}$, $H^{\mathfrak{L}}$ be a $L_n$-dimensional Hilbert space with an orthonormal basis $\{|l\rangle^{\mathfrak{L}}:l=1,\ldots,L_n\}$, and $H^{\theta}$ be a $|\theta|$-dimensional Hilbert space with an orthonormal basis $\{|t\rangle^{\theta}:t\in\theta\}$. Let $|0\rangle^{\mathfrak{M}}|0\rangle^{\mathfrak{L}}|0\rangle^{\theta}$ be the ancillas on $H^{\mathfrak{M}}$, $H^{\mathfrak{L}}$, and $H^{\theta}$, respectively, that the receiver adds. We can (cf. [52]) define a unitary matrix $V^{\mathfrak{Q}^n\mathfrak{M}\mathfrak{L}\theta}$ on $H^{\mathfrak{Q}^n\mathfrak{M}\mathfrak{L}\theta}$ such that for any given quantum state $\rho^{\mathfrak{Q}^n}\in\mathcal{S}(H^{\mathfrak{Q}^n})$ we have

$$V^{\mathfrak{Q}^n\mathfrak{M}\mathfrak{L}\theta}\left(\rho^{\mathfrak{Q}^n}\otimes|0\rangle\langle 0|^{\mathfrak{M}}\otimes|0\rangle\langle 0|^{\mathfrak{L}}\otimes|0\rangle\langle 0|^{\theta}\right)(V^{\mathfrak{Q}^n\mathfrak{M}\mathfrak{L}\theta})^*$$
$$=\sum_t\sum_j\sum_L\left(D_{x_{j,l}}^{\mathfrak{Q}^n}\rho^{\mathfrak{Q}^n}\right)\otimes|j\rangle\langle j|^{\mathfrak{M}}|l\rangle\langle l|^{\mathfrak{L}}|t\rangle\langle t|^{\theta}\,.$$

We denote

$$\psi_{j,l,t}^{\mathfrak{Q}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta}$$
$$:=\left(\mathrm{id}^{\mathfrak{E}^n}\otimes V^{\mathfrak{Q}^n\mathfrak{M}\mathfrak{L}\theta}\right)\left(U_N\otimes\mathrm{id}^{\mathfrak{M}\mathfrak{L}\theta}\right)\left[|\aleph_{j,l}\rangle\langle\aleph_{j,l}|^{\mathfrak{P}^n}\right.$$

$$\otimes |0\rangle\langle 0|^{\mathfrak{M}} \otimes |0\rangle\langle 0|^{\mathfrak{L}} \otimes |0\rangle\langle 0|^{\theta}\Big] \left(U_N \otimes \mathrm{id}^{\mathfrak{M}\mathfrak{L}\theta}\right)^* \left(\mathrm{id}^{\mathfrak{E}^n} \otimes V^{\mathfrak{Q}^n\mathfrak{M}\mathfrak{L}\theta}\right)^* ,$$

in view of (80), we have

$$F\left(\mathrm{tr}_{\mathfrak{Q}^n\mathfrak{E}^n}\left(\psi_{j,l,t}^{\mathfrak{Q}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta}\right), |j\rangle\langle j|^{\mathfrak{M}} \otimes |l\rangle\langle l|^{\mathfrak{L}} \otimes |t\rangle\langle t|^{\theta}\right) \geq 1-\epsilon . \tag{82}$$

By Uhlmann's theorem (cf. e.g. [66]) we can find a $|\zeta_{j,l,t}\rangle^{\mathfrak{Q}^n\mathfrak{E}^n}$ on $H^{\mathfrak{Q}^n\mathfrak{E}^n}$, such that

$$\begin{aligned}
&\langle 0|^{\theta}\langle 0|^{\mathfrak{L}}\langle 0|^{\mathfrak{M}}\langle \aleph_{j,l}|^{\mathfrak{P}^n} \left(U_{N_t} \otimes \mathrm{id}^{\mathfrak{M}\mathfrak{L}\theta}\right)^* \\
&\left(\mathrm{id}^{\mathfrak{E}^n} \otimes V^{\mathfrak{Q}^n\mathfrak{M}\mathfrak{L}\theta}\right)^* |\zeta_{j,l,t}\rangle^{\mathfrak{Q}^n\mathfrak{E}^n}|j\rangle^{\mathfrak{M}}|l\rangle^{\mathfrak{L}}|t\rangle^{\theta} \\
&= F\Bigg(\psi_{j,l,t}^{\mathfrak{Q}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta}, |\zeta_{j,l,t}\rangle\langle\zeta_{j,l,t}|^{\mathfrak{Q}^n\mathfrak{E}^n} \\
&\otimes |j\rangle\langle j|^{\mathfrak{M}} \otimes |l\rangle\langle l|^{\mathfrak{L}} \otimes |t\rangle\langle t|^{\theta}\Bigg) \\
&\geq 1-\epsilon .
\end{aligned} \tag{83}$$

*ii) Building the Second Part of the Decoding Operator*

We define

$$|a_{j,l}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} := |\aleph_{j,l}\rangle^{\mathfrak{P}^n}|0\rangle^{\mathfrak{M}}|0\rangle^{\mathfrak{L}}|0\rangle^{\theta} ,$$

and

$$\begin{aligned}
|b_{j,l,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} &:= \left(U_{N_t} \otimes \mathrm{id}^{\mathfrak{M}\mathfrak{L}\theta}\right)^* \left(\mathrm{id}^{\mathfrak{E}^n} \otimes V^{\mathfrak{Q}^n\mathfrak{M}\mathfrak{L}\theta}\right)^* \\
&|\zeta_{j,l,t}\rangle^{\mathfrak{Q}^n\mathfrak{E}^n}|j\rangle^{\mathfrak{M}}|l\rangle^{\mathfrak{L}}|t\rangle^{\theta} .
\end{aligned}$$

For every $j$, $l$, and $t$, we have $\langle a_{j,l}|b_{j,l,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} \geq 1-\epsilon$.

We define

$$|\hat{a}_{j,k}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} := \frac{1}{\sqrt{L_n}} \sum_{l=1}^{L_n} e^{-2\pi i l \frac{k}{L_n}} |a_{j,l}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} ,$$

$$|\hat{b}_{j,k,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} := \frac{1}{\sqrt{L_n}} \sum_{l=1}^{L_n} e^{-2\pi i l \frac{k}{L_n}} |b_{j,l,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} ,$$

and

$$|\bar{b}_{j,k}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} := \frac{1}{|\theta|} \sum_{t=1}^{|\theta|} |\hat{b}_{j,k,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} .$$

For every $j \in \{1, \ldots, J_n\}$, by (83) it holds

$$\frac{1}{L_n} \sum_{k=1}^{L_n} \langle \hat{a}_{j,k}|\bar{b}_{j,k}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta}$$

$$
\begin{aligned}
&= \frac{1}{|\theta|}\frac{1}{L_n}\sum_{t=1}^{|\theta|}\sum_{k=1}^{L_n}\langle \hat{a}_{j,k}|\hat{b}_{j,k,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} \\
&= \frac{1}{|\theta|}\frac{1}{L_n}\sum_{t=1}^{|\theta|}\sum_{l=1}^{L_n}\langle a_{j,l}|b_{j,l,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} \\
&\geq 1-\epsilon\,. \qquad (84)
\end{aligned}
$$

Hence there is at least one $k_j \in \{1,\dots,L_n\}$ such that for every $j$, we have

$$
\begin{aligned}
&1-\epsilon \\
&\leq e^{-is_{k_j}}\langle \hat{a}_{j,k_j}|\bar{b}_{j,k_j}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} \\
&= \frac{1}{|\theta|}\sum_{t=1}^{|\theta|} e^{-is_{k_j}}\langle \hat{a}_{j,k_j}|\hat{b}_{j,k_j,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta}\,,
\end{aligned}
$$

for a suitable phase $s_{k_j}$. Since for all $t$ it holds $\left|e^{-is_{k_j}}\langle \hat{a}_{j,k_j}|\hat{b}_{j,k_j,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta}\right| \leq 1$, we have

$$
\min_{t\in\theta}\left|e^{-is_{k_j}}\langle \hat{a}_{j,k_j}|\hat{b}_{j,k_j,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta}\right| \geq 1-|\theta|\epsilon\,.
$$

Therefore, there is a suitable phase $r_{k_j}$ such that for all $t\in\theta$,

$$
\begin{aligned}
&1-|\theta|\epsilon \\
&\leq \left|e^{-is_{k_j}}\langle \hat{a}_{j,k_j}|\hat{b}_{j,k_j,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta}\right| \\
&= e^{-ir_{k_j}}\langle \hat{a}_{j,k_j}|\hat{b}_{j,k_j,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta} \\
&= e^{-ir_{k_j}}\frac{1}{L_n}\left(\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}}\langle a_{j,l}|^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta}\right) \\
&\left(\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}}|b_{j,l,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta}\right)\,. \qquad (85)
\end{aligned}
$$

For every $t\in\theta$, we set

$$
|\varpi_{j,t}\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} := \sqrt{\frac{1}{L_n}}\sum_{l=1}^{L_n} e^{-2\pi i(l\frac{k_j}{L_n}+r_{k_j})}|\zeta_{j,l,t}\rangle^{\mathfrak{O}^n\mathfrak{E}^n}\otimes|l\rangle^{\mathfrak{L}}
$$

and

$$
\begin{aligned}
&|\vartheta_{j,t}\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta} := \sqrt{\frac{1}{L_n}}\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}}\left[\mathrm{id}^{\mathfrak{E}^n}\otimes V^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}\theta}\right] \\
&(U_N^n|\aleph_{j,l}\rangle^{\mathfrak{P}^n})|0\rangle^{\mathfrak{M}}|0\rangle^{\mathfrak{L}}|0\rangle^{\theta}\,.
\end{aligned}
$$

For all $t\in\theta$ and $j\in\{1,\dots J_n\}$ it holds by (85)

$$
F\Bigg(|\vartheta_{j,t}\rangle\langle\vartheta_{j,t}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta},
$$

$$
\begin{aligned}
&|\varpi_{j,t}\rangle\langle\varpi_{j,t}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} \otimes |j\rangle\langle j|^{\mathfrak{M}} \otimes |t\rangle\langle t|^{\theta}\Big) \\
&= \left|\langle\vartheta_{j,t}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta}|\varpi_{j,t}\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}}|j\rangle^{M}|t\rangle^{\theta}\right| \\
&= \frac{1}{L_n}\left(\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}}\langle a_{j,l}|^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta}\right) \\
&\left(\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}} e^{-ir_{k_j}}|b_{j,l,t}\rangle^{\mathfrak{P}^n\mathfrak{M}\mathfrak{L}\theta}\right) \\
&\geq 1-|\theta|\epsilon\,. \qquad (86)
\end{aligned}
$$

Furthermore, since (81) holds there is a quantum state $\xi_t^{\mathfrak{E}^n}$, which does not depend on $j$ and $l$, on $H^{\mathfrak{E}^n}$ such that

$$
\left\|\xi_t^{\mathfrak{E}^n} - \mathrm{tr}_{\mathfrak{O}^n}\left(U_{N_t}|\aleph_{j,l}\rangle\langle\aleph_{j,l}|^{\mathfrak{P}^n}U_{N_t}^*\right)\right\|_1 \leq \epsilon\,. \qquad (87)
$$

By monotonicity of fidelity, for any $l\in\{1,\ldots,L_n\}$

$$
\begin{aligned}
&\left\|\mathrm{tr}_{\mathfrak{O}^n}\left(U_{N_t}|\aleph_{j,l}\rangle\langle\aleph_{j,l}|^{\mathfrak{P}^n}U_{N_t}^*\right) - \mathrm{tr}_{\mathfrak{O}^n}\left(|\zeta_{j,l,t}\rangle\langle\zeta_{j,l,t}|^{\mathfrak{O}^n\mathfrak{E}^n}\right)\right\|_1 \\
&\leq 2\Bigg[1-F\Bigg(\mathrm{tr}_{\mathfrak{O}^n}\left(U_{N_t}|\aleph_{j,l}\rangle\langle\aleph_{j,l}|^{\mathfrak{P}^n}U_{N_t}^*\right), \\
&\mathrm{tr}_{\mathfrak{O}^n}\left(|\zeta_{j,l,t}\rangle\langle\zeta_{j,l,t}|^{\mathfrak{O}^n\mathfrak{E}^n}\right)\Bigg)\Bigg]^{\frac{1}{2}} \\
&\leq 2\Bigg[1-F\Bigg(\psi_{j,l,t}^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta}, |\zeta_{j,l,t}\rangle\langle\zeta_{j,l,t}|^{\mathfrak{O}^n\mathfrak{E}^n} \\
&\otimes|j\rangle\langle j|^{\mathfrak{M}}\otimes|l\rangle\langle l|^{\mathfrak{L}}\otimes|t\rangle\langle t|^{\theta}\Bigg)\Bigg]^{\frac{1}{2}} \\
&\leq 2\sqrt{\epsilon}\,, \qquad (88)
\end{aligned}
$$

the first inequality holds because for two quantum states $\varrho$ and $\eta$, we have $\frac{1}{2}\|\varrho-\eta\|_1 \leq \sqrt{1-F(\varrho,\eta)}$.

By (87) and (88)

$$
\begin{aligned}
&\left\|\mathrm{tr}_{\mathfrak{O}^n\mathfrak{L}}\left(|\varpi_{j,t}\rangle\langle\varpi_{j,t}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}}\right) - \xi_t^{\mathfrak{E}^n}\right\|_1 \\
&= \left\|\frac{1}{L_n}\sum_{l=1}^{L_n}\mathrm{tr}_{\mathfrak{O}^n}\left(|\zeta_{j,l,t}\rangle\langle\zeta_{j,l,t}|^{\mathfrak{O}^n\mathfrak{E}^n}\right) - \xi_t^{\mathfrak{E}^n}\right\|_1 \\
&\leq \frac{1}{L_n}\sum_{l=1}^{L_n}\Bigg\|\mathrm{tr}_{\mathfrak{O}^n}\left(U_{N_t}|\aleph_{j,l}\rangle\langle\aleph_{j,l}|^{\mathfrak{P}^n}U_{N_t}^*\right) \\
&- \mathrm{tr}_{\mathfrak{O}^n}\left(|\zeta_{j,l,t}\rangle\langle\zeta_{j,l,t}|^{\mathfrak{O}^n\mathfrak{E}^n}\right)\Bigg\|_1 \\
&+ \left\|\xi_t^{\mathfrak{E}^n} - \mathrm{tr}_{\mathfrak{O}^n}\left(U_{N_t}|\aleph_{j,l}\rangle\langle\aleph_{j,l}|^{\mathfrak{P}^n}U_{N_t}^*\right)\right\|_1
\end{aligned}
$$

$$\leq 2\sqrt{\epsilon} + \epsilon \,, \tag{89}$$

holds for all $t \in \theta$ and $j \in \{1, \ldots, J_n\}$.

In [59] (cf. also [34]) it was shown that when (89) holds, for every $t \in \theta$ we can find a unitary operator $U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}}$ such that if we set

$$\chi_{j,j',t}^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}} := \left(U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes \mathrm{id}^{\mathfrak{E}^n}\right) \left(|\varpi_{j,t}\rangle\langle\varpi_{j,t}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} \otimes |j\rangle\langle j'|^{\mathfrak{M}}\right) \left(U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes \mathrm{id}^{\mathfrak{E}^n}\right)^* \,,$$

then

$$F\left(|\xi_t\rangle\langle\xi_t|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} \otimes |j\rangle\langle j'|^{\mathfrak{M}}, \chi_{j,j',t}^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}}\right) \geq 1 - 4\epsilon - 4\sqrt{\epsilon} \,, \tag{90}$$

where $|\xi_t\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}}$ is chosen so that $|\xi_t\rangle\langle\xi_t|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}}$ is a purification of $\xi_t^{\mathfrak{E}^n}$ on $H^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}}$.

*iii) Defining the Code*

We can now define our entanglement generating code. Let $t'$ be arbitrary in $\theta$. The sender prepares the quantum state

$$\frac{1}{J_n}\frac{1}{L_n}\left(\sum_{j=1}^{J_n}\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}} |\aleph_{j,l}\rangle^{\mathfrak{P}^n} |j\rangle^{\mathfrak{A}}\right) \left(\sum_{j=1}^{J_n}\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}} \langle j|^{\mathfrak{A}} \langle\aleph_{j,l}|^{\mathfrak{P}^n}\right) \,, \tag{91}$$

keeps the system $\mathfrak{A}$, and sends the system $\mathfrak{P}^n$ through the channel $N_{t'}^{\otimes n}$, i.e. the resulting quantum state is

$$\begin{aligned}
&\frac{1}{J_n}\frac{1}{L_n}\left(\mathrm{id}^{\mathfrak{A}} \otimes U_{N_{t'}}^n\right)\left[\left(\sum_{j=1}^{J_n}\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}} |j\rangle^A |\aleph_{j,l}\rangle^{\mathfrak{P}^n}\right)\right. \\
&\left.\left(\sum_{j=1}^{J_n}\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}} \langle\aleph_{j,l}|^{\mathfrak{P}^n}\langle j|^{\mathfrak{A}}\right)\right]\left(\mathrm{id}^{\mathfrak{A}} \otimes U_{N_{t'}}^n\right)^* \\
&= \frac{1}{J_n}\frac{1}{L_n}\left[\sum_{j=1}^{J_n} |j\rangle^{\mathfrak{A}}\left(\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}} U_{N_{t'}}^n |\aleph_{j,l}\rangle^{\mathfrak{P}^n}\right)\right] \\
&\left[\sum_{j=1}^{J_n}\left(\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}} \langle\aleph_{j,l}|^{\mathfrak{P}^n} (U_{N_{t'}}^n)^*\right)\langle j|^{\mathfrak{A}}\right] \,.
\end{aligned}$$

The receiver subsequently applies the decoding operator

$$\begin{aligned}
\tau^{\mathfrak{O}^n} \to \mathrm{tr}_{\mathfrak{O}^n\mathfrak{L}\theta}&\left[\left(\sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes |t\rangle\langle t|^{\theta}\right) V^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}\theta}\right. \\
&\left(\tau^{\mathfrak{O}^n} \otimes |0\rangle\langle 0|^{\mathfrak{M}} \otimes |0\rangle\langle 0|^{\mathfrak{L}} \otimes |0\rangle\langle 0|^{\theta}\right) \\
&\left. V^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}\theta *}\left(\sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes |t\rangle\langle t|^{\theta}\right)^*\right] \,,
\end{aligned} \tag{92}$$

to his outcome.

*iii.1) The Resulting Quantum State after Performing the Decoding Operator*

We define

$$\begin{aligned}
&\iota_{t'}^{\mathfrak{A}\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta} \\
&:= \left(\sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes \mathrm{id}^{\mathfrak{A}\mathfrak{E}^n} \otimes |t\rangle\langle t|^{\theta}\right) (V^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}\theta} \otimes \mathrm{id}^{\mathfrak{A}\mathfrak{E}^n}) \\
&\left(\frac{1}{J_n}\frac{1}{L_n}\left[\sum_{j=1}^{J_n} |j\rangle^{\mathfrak{A}} \left(\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}} U_{N_{t'}}^n |\aleph_{j,l}\rangle^{P^n}\right)\right]\right. \\
&\left.\left[\sum_{j=1}^{J_n} \left(\sum_{l=1}^{L_n} e^{-2\pi i l\frac{k_j}{L_n}} \langle\aleph_{j,l}|^{P^n} (U_{N_{t'}}^n)^*\right) \langle j|^{\mathfrak{A}}\right]\right. \\
&\left.\otimes |0\rangle\langle 0|^{\mathfrak{M}} \otimes |0\rangle\langle 0|^{\mathfrak{L}} \otimes |0\rangle\langle 0|^{\theta}\right) (V^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}\theta} \otimes \mathrm{id}^{\mathfrak{A}\mathfrak{E}^n})^* \\
&\left(\sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes \mathrm{id}^{\mathfrak{A}\mathfrak{E}^n} \otimes |t\rangle\langle t|^{\theta}\right)^* \\
&= \left(\sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes \mathrm{id}^{\mathfrak{A}\mathfrak{E}^n} \otimes |t\rangle\langle t|^{\theta}\right) \\
&\left(\frac{1}{J_n}(\sum_{j=1}^{J_n} |j\rangle^{\mathfrak{A}} |\vartheta_{j,t'}\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta})(\sum_{j=1}^{J_n} \langle\vartheta_{j,t'}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta} \langle j|^{\mathfrak{A}})\right) \\
&\left(\sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes \mathrm{id}^{\mathfrak{A}\mathfrak{E}^n} \otimes |t\rangle\langle t|^{\theta}\right)^*, \qquad (93)
\end{aligned}$$

then the resulting quantum state after performing the decoding operator is $\mathrm{tr}_{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}\theta}(\iota_{t'}^{\mathfrak{A}\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta})$.

*iii.2) The Fidelity of $\frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{j'=1}^{J_n} \chi_{j,j',t'}^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}} \otimes |j\rangle\langle j|^{\mathfrak{A}} \otimes |t'\rangle\langle t'|^{\theta}$ and the Actual Quantum State*

$$\begin{aligned}
&\left(\sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes \mathrm{id}^{\mathfrak{A}\mathfrak{E}^n} \otimes |t\rangle\langle t|^{\theta}\right) \\
&\left(\sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes \mathrm{id}^{\mathfrak{E}^n} \otimes |t\rangle\langle t|^{\theta}\right)^* \\
&= \mathrm{id}^{\mathfrak{A}\mathfrak{E}^n} \otimes \sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} (U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}})^* \otimes |t\rangle\langle t|^{\theta} \\
&= \mathrm{id}^{\mathfrak{A}\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta},
\end{aligned}$$

$\sum_{t\in\theta} U_{(t)}^{\mathfrak{O}^n\mathfrak{M}\mathfrak{L}} \otimes \mathrm{id}^{\mathfrak{E}^n} \otimes |t\rangle\langle t|^{\theta}$ is unitary.

Because of this unitarity and by (86)

$$F\left(\iota_{t'}^{\mathfrak{A}\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta}, \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{j'=1}^{J_n} \chi_{j,j',t'}^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}} \otimes |j\rangle\langle j'|^{\mathfrak{A}} \otimes |t'\rangle\langle t'|^{\theta}\right)$$

$$
\begin{aligned}
&= F\Bigg( \frac{1}{J_n}(\sum_{j=1}^{J_n} |j\rangle^{\mathfrak{A}} |\vartheta_{j,t'}\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta})(\sum_{j=1}^{J_n} \langle\vartheta_{j,t'}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta}\langle j|^{\mathfrak{A}}), \\
&\frac{1}{J_n}(\sum_{j=1}^{J_n} |\varpi_{j,t'}\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} \otimes |j\rangle^{\mathfrak{A}} \otimes |j\rangle^{\mathfrak{M}}) \\
&(\sum_{j=1}^{J_n} \langle j|^{\mathfrak{M}} \otimes \langle j|^{\mathfrak{A}} \otimes \langle\varpi_{j,t'}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}}) \otimes |t'\rangle\langle t'|^{\theta} \Bigg) \\
&= \frac{1}{J_n}\Bigg| \Bigg( \sum_{j=1}^{J_n} \langle\vartheta_{j,t'}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta} \Bigg) \\
&\Bigg( \sum_{j=1}^{J_n} |\varpi_{j,t'}\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} \otimes |j\rangle^{\mathfrak{M}} \otimes |t'\rangle^{\theta} \Bigg) \Bigg| \\
&\geq 1 - |\theta|\epsilon \,. \qquad (94)
\end{aligned}
$$

*iii.3) The Fidelity of* $\frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{j'=1}^{J_n} \chi_{j,j',t'}^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}} \otimes |j\rangle\langle j'|^{\mathfrak{A}} \otimes |t'\rangle\langle t'|^{\theta}$ *and the Standard Maximally Entanglement State*

By (90) we have

$$
\begin{aligned}
&F\Bigg( \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{j'=1}^{J_n} \chi_{j,j',t'}^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}} \otimes |t'\rangle\langle t'|^{\theta} \otimes |j\rangle\langle j'|^{\mathfrak{A}}, \\
&\frac{1}{J_n}(\sum_{j=1}^{J_n} |\xi_{t'}\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} \otimes |j\rangle^{\mathfrak{A}} \otimes |j\rangle^{\mathfrak{M}} \otimes |t'\rangle^{\theta}) \\
&(\sum_{j=1}^{J_n} \langle\xi_{t'}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} \otimes \langle j|^{\mathfrak{A}} \otimes \langle j|^{\mathfrak{M}} \otimes \langle t'|^{\theta}) \Bigg) \\
&\geq 1 - 4\epsilon - 4\sqrt{\epsilon} \,. \qquad (95)
\end{aligned}
$$

*iii.4) The Fidelity of the Actual Quantum State and the Standard Maximally Entanglement State*

Since for two quantum states $\varrho$ and $\eta$, it holds

$$
1 - \sqrt{F(\varrho,\eta)} \leq \frac{1}{2}\|\varrho - \eta\|_1 \leq \sqrt{1 - F(\varrho,\eta)^2} \,,
$$

for three quantum states $\varrho$, $\eta$, and $\upsilon$, we have

$$
\begin{aligned}
&F(\varrho,\eta) \\
&\geq 1 - \frac{1}{2}\|\varrho - \eta\|_1 \\
&\geq 1 - \frac{1}{2}\|\varrho - \upsilon\|_1 - \frac{1}{2}\|\upsilon - \eta\|_1 \\
&\geq 1 - \sqrt{1 - F(\varrho,\upsilon)} - \sqrt{1 - F(\upsilon,\eta)} \,.
\end{aligned}
$$

Combining (94) and (95), for all $t' \in \theta$ we have

$$
\begin{aligned}
& F\Bigg( \mathrm{tr}_{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}\theta}(\iota_{t'}^{\mathfrak{A}\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta}), \\
& (\sum_{j=1}^{J_n} |j\rangle^{\mathfrak{A}} \otimes |j\rangle^{\mathfrak{M}})(\sum_{j=1}^{J_n} \langle j|^{\mathfrak{A}} \otimes \langle j|^{\mathfrak{M}}) \Bigg) \\
& \geq F\Bigg( \iota_{t'}^{\mathfrak{A}\mathfrak{O}^n\mathfrak{E}^n\mathfrak{M}\mathfrak{L}\theta}, \frac{1}{J_n}(\sum_{j=1}^{J_n} |\xi_{t'}\rangle^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} \otimes |j\rangle^{\mathfrak{A}} \otimes |j\rangle^{\mathfrak{M}} \otimes |t'\rangle^{\theta}) \\
& (\sum_{j=1}^{J_n} \langle \xi_{t'}|^{\mathfrak{O}^n\mathfrak{E}^n\mathfrak{L}} \otimes \langle j|^{\mathfrak{A}} \otimes \langle j|^{\mathfrak{M}} \otimes \langle t'|^{\theta}) \Bigg) \\
& \geq 1 - \sqrt{2|\theta|\epsilon - |\theta|^2\epsilon^2} - \sqrt{8\sqrt{\epsilon} - 16\epsilon^2 - 32\epsilon\sqrt{\epsilon} - 8\epsilon} \qquad (96) \\
& \geq 1 - \sqrt{2|\theta|}\sqrt{\epsilon} - \sqrt{8}\sqrt[4]{\epsilon}\,. \qquad (97)
\end{aligned}
$$

This means that if $n$ is large enough, then for any positive $\delta$ and $\epsilon$, there is an $(n, \sqrt{2|\theta|}\sqrt{\epsilon} + \sqrt{8}\sqrt[4]{\epsilon})$ code with rate

$$\min_t \chi(X; Q_t) - \max_t \chi(X; E_t) - 2\delta\,. \qquad \square$$

**Proposition 3.14.** *The entanglement generating capacity of $\{(N_t) : t \in \theta\}$ with CSI at the encoder is*

$$A_{CSI}(\{(N_t) : t \in \theta\}) = \lim_{n\to\infty} \frac{1}{n} \min_{t\in\theta} \max_{\rho\in\mathcal{S}(H)^{\mathfrak{O}^n}} I_C(\rho; {N_t}^{\otimes n})\,. \qquad (98)$$

*Proof.* As the authors of [23] showed, after receiving a dummy code word as the first block, the receiver also can have CSI. Then we have the case where both the sender and the receiver have CSI. But this case is equivalent to the case where we only have one channel $(N_t)$ instead of a family of channels $\{(N_t) : t = 1, \ldots, |\theta|\}$, and we may assume it is the worst channel. The bits that we use to detect the CSI are large but constant, so it is negligible compared to the rest. By [34], the entanglement generating capacity of the quantum channel $N_t$ is

$$\lim_{n\to\infty} \frac{1}{n} \max_{\rho\in\mathcal{S}(H)^{\mathfrak{O}^n}} I_C(\rho; N_t^{\otimes n})\,.$$

The proof of the converse is similar to those given in the proof of Theorem 3.7, where we consider a worst $t'$. $\square$

**Proposition 3.15.** *The entanglement generating capacity of $(N_t)_{t\in\theta}$ with feedback is bounded as follows*

$$A_{feed}(\{(N_t) : t \in \theta\}) \geq \lim_{n\to\infty} \frac{1}{n} \min_{t\in\theta} \max_{\rho\in\mathcal{S}(H)^{\mathfrak{O}^n}} I_C(\rho; {N_t}^{\otimes n})\,. \qquad (99)$$

*Proof.* As the authors of [23] showed, the receiver can detect the channel state $t$ correctly after receiving a dummy word as the first block. Then he can send $t$ back to the sender via feedback. □

**Remark 3.16.** *Feedback can improve the channel capacity of quantum channels in some cases (c.f. [45]). Thus it can be possible that the lower bound in Proposition 3.15 is not tight. For a one-way entanglement distillation protocol using secret key, cf. [35].*

## 3.4 Further Notes

In this section we will discuss the proof of our result of the previous section.

Let $\mathfrak{P}$, $\mathfrak{Q}$, $H^{\mathfrak{P}}$, and $H^{\mathfrak{Q}}$ be defined as in Section 2. Let $N$ be a quantum channel $\mathcal{S}(H^{\mathfrak{P}}) \to \mathcal{S}(H^{\mathfrak{Q}})$. In general, there are two ways to represent a quantum channel, i. e. a completely positive trace-preserving map $H^{\mathfrak{P}} \to H^{\mathfrak{Q}}$, with linear algebraic tools.

*1. Operator Sum Decomposition (Kraus Representation)*

$$N(\rho) = \sum_{i=1}^{K} A_i \rho {A_i}^* \,, \tag{100}$$

where $A_1, \ldots, A_K$ (Kraus operators) are linear operators $\mathcal{S}(H^{\mathfrak{P}}) \to \mathcal{S}(H^{\mathfrak{Q}})$ (cf.[44], [12], and [52]). They satisfy the completeness relation $\sum_{i=1}^{K} {A_i}^* A_i = \mathrm{id}_{H^{\mathfrak{P}}}$. The representation of a quantum channel $N$ according to (100) is not unique. Let $A_1, \ldots, A_K$ and $B_1, \ldots, B_{K'}$ be two sets of Kraus operators (by appending zero operators to the shorter list of operation elements we may ensure that $K' = K$). Suppose $A_1, \ldots, A_K$ represents $N$, then $B_1, \ldots, B_K$ also represents $N$ if and only if there exists a $K \times K$ unitary matrix $(u_{i,j})_{i,j=1,\ldots,K}$ such that for all $i$ we have $A_i = \sum_{j=1}^{K} u_{i,j} B_j$ (cf. [52]).

*2. Isometric Extension (Stinespring Dilation)*

$$N(\rho) = \mathrm{tr}_{\mathfrak{E}} \left( U_N \rho U_N^* \right) \,, \tag{101}$$

where $U_N$ is a linear operator $\mathcal{S}(H^{\mathfrak{P}}) \to \mathcal{S}(H^{\mathfrak{Q}\mathfrak{E}})$ such that $U_N^* U_N = \mathrm{id}_{H^{\mathfrak{P}}}$, and $\mathfrak{E}$ is the quantum system of the environment (cf. [60], [12], and also cf. [63] for a more general Stinespring Dilation Theorem). $H^{\mathfrak{E}}$ can be chosen such that $\dim H^{\mathfrak{E}} \leq (\dim H^{\mathfrak{P}})^2$. The isometric extension of a quantum channel $N$ according to (101) is not unique either. Let $U$ and $U'$ be two linear operators $\mathcal{S}(H^{\mathfrak{P}}) \to \mathcal{S}(H^{\mathfrak{Q}\mathfrak{E}})$. Suppose $U$ represents $N$, then $U'$ also represents $N$ if and only if $U$ and $U'$ are unitarily equivalent.

It is well known that we can reduce each of these two representations of the quantum channel from the other one. Let $A_1, \ldots, A_K$ be a set of Kraus operators which represents $N$. Let $\{|j\rangle^{\mathfrak{E}} : j = 1, \ldots, K\}$ be an orthonormal system on $H^{\mathfrak{E}}$. Then $U_N = \sum_{j=1}^{K} A_j \otimes |j\rangle^{\mathfrak{E}}$ is an isometric extension which represents $N$, since $\left(\sum_{j=1}^{K} A_j \otimes |j\rangle^{\mathfrak{E}}\right) \rho \left(\sum_{k=1}^{K} A_k \otimes |k\rangle^{\mathfrak{E}}\right)^* = \sum_{j=1}^{K} A_j \rho {A_j}^*$ and $\left(\sum_{j=1}^{K} A_j \otimes |j\rangle^{\mathfrak{E}}\right)^* \left(\sum_{k=1}^{K} A_k \otimes |k\rangle^{\mathfrak{E}}\right) = \sum_{j=1}^{K} {A_j}^* A_j$. For the other way around, every isometric extension $U_N$ that represents $N$ can be written in the form $U_N = \sum_{j=1}^{K} A_j \otimes |j\rangle^{\mathfrak{E}}$, i.e. if the sender sends $\rho$, and if the environment's measurement gives $|i\rangle^{\mathfrak{E}}$, the receiver's outcome will be $A_i \rho {A_i}^*$. Here $A_1, \ldots, A_K$ is a set of Kraus operators which represents $N$, and $\{|j\rangle^{\mathfrak{E}} : j = 1, \ldots, K\}$ is an orthonormal system on $H^{\mathfrak{E}}$.

Using either of both methods to represent a quantum channel, one can show that (cf. [34]) the entanglement generating capacity of a quantum channel $N$ is

$$\mathcal{A}(N) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho\in\mathcal{S}(H)^{\Omega^n}} I_C(\rho; N^{\otimes n}) \,. \tag{102}$$

The Kraus representation describes the dynamics of the principal system without having to explicitly consider properties of the environment, whose dynamics are often unimportant. All that we need to know is the system of the receiver alone; this simplifies calculations. In [43], an explicit construction of a quantum error correction code (both perfect and approximate information recovery) with the Kraus operators is given. In the Stinespring dilation, we have a natural interpretation of the system of the environment. From the Stinespring dilation, we can conclude that the receiver can detect almost all quantum information if and only if the channel releases almost no information to the environment. In [59], an alternative way to build a quantum error correction code (both perfect and approximate information recovery) is given using this fact. The disadvantage is that we suppose it is suboptimal for calculating the entanglement generating capacity of a compound quantum channel without CSI at the encoder.

In [17], the entanglement generating capacity for the compound quantum channel is determined, using a quantum error correction code of [43], which is built by Kraus operators. Their result is the following. The entanglement generating capacity of a quantum wiretap channel $N = (N_t)_{t\in\theta}$ is

$$\mathcal{A}(N) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho\in\mathcal{S}(H)^{\Omega^n}} \min_{t\in\theta} I_C(\rho; {N_t}^{\otimes n}) \,. \tag{103}$$

This result is stronger than our result in Theorem 3.13. This is due to the fact that we use for our proof a quantum error correction code of [59], which is based upon the Stinespring dilation. If we use the Kraus operators to represent a compound quantum channel, we have a bipartite system, and for calculating the entanglement generating capacity of a compound quantum channel, we can use the technique which is similar to the case of a single quantum channel. However, if we use the Stinespring dilation to represent a compound quantum channel, we have a tripartite system which includes the sender, the receiver, and in addition, the environment. Unlike in the case of a single quantum channel, for compound quantum channel we have to deal with uncertainty at the environment. If the sender knows the CSI, the transmitters can build an $(n, \epsilon)$ code for entanglement generating with rate $\min_t [\chi(X; Q_t) - \chi(X; E_t)] - \delta = \min_{t\in\theta} I_C(\rho; N_t) - \delta$ (Proposition 3.14) for any positive $\delta$ and $\epsilon$. This result is optimal (cf. [17]). But if the sender does not know the CSI, he has to build an encoding operator by considering every possible channel state for the environment. Therefore the maximal rate that we can achieve is $\min_t \chi(X; Q_t) - \max_t \chi(X; E_t)$, but not $\min_{t\in\theta} I_C(\rho; N_t) = \min_t [\chi(X; Q_t) - \chi(X; E_t)]$. This is only a lower bound of the entanglement generating capacity. It is unknown if we can achieve the stronger result (103) using the Stinespring dilation.

# 4 Classical-Quantum Arbitrarily Varying Wiretap Channel and Resources

The results in this section was published in [25], [26], and [27].

## 4.1 Ahlswede Dichotomy for Arbitrarily Varying Classical-Quantum Wiretap Channels

In this section we analyze the secrecy capacities of various coding schemes with resource assistance. Our goal is to see what the effects are on the secrecy capacities of an arbitrarily varying classical-quantum wiretap channel if we use deterministic code, randomness assisted code, or common randomness assisted code.

**Theorem 4.1 (Ahlswede dichotomy).** *Let $\{(W_t, V_t) : t \in \theta\}$ be an arbitrarily varying classical-quantum wiretap channel.*

1. *(a) If the arbitrarily varying classical-quantum channel $\{W_t : t \in \theta\}$ is not symmetrizable, then*

$$C_s(\{(W_t, V_t) : t \in \theta\}) = C_s(\{(W_t, V_t) : t \in \theta\}; r)\,. \tag{104}$$

*(b) If $\{W_t : t \in \theta\}$ is symmetrizable,*

$$C_s(\{(W_t, V_t) : t \in \theta\}) = 0\,. \tag{105}$$

2.

$$C_s(\{(W_t, V_t) : t \in \theta\}; cr) = C_s(\{(W_t, V_t) : t \in \theta\}; r)\,. \tag{106}$$

*Proof.* Our proof is similar to the proof of Ahlswede Dichotomy for arbitrarily varying classical-quantum channels in [6]. The difference between our proof and the proofs in [6] is that we have to additionally consider the security.

*i) Proof of Theorem 4.1. 2*

At first we use random encoding technique to show the existence of a common randomness assisted code.

Choose arbitrary positive $\epsilon$ and $\zeta$. Assume we have an $(n, J_n)$ randomness assisted code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$ for $\{(W_t, V_t) : t \in \theta\}$ such that

$$\max_{t^n \in \theta^n} \int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < \zeta\,,$$

$$\max_{t^n \in \theta^n} \int_\Lambda \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG(\gamma) < \epsilon\,.$$

Consider now $n^3$ independent and identically distributed random variables $\overline{\mathcal{C}}_1, \overline{\mathcal{C}}_2, \cdots, \overline{\mathcal{C}}_{n^3}$ with values in $\{\mathcal{C}^\gamma : \gamma \in \Lambda\}$ such that $Pr(\overline{\mathcal{C}}_i = \mathcal{C}) = G(\mathcal{C})$ for all $\mathcal{C} \in \{\mathcal{C}^\gamma : \gamma \in \Lambda\}$ and for all $i \in \{1, \cdots, n^3\}$. For a fixed $t^n \in \theta^n$ we have

$$Pr\left(\sum_{i=1}^{n^3} \chi\left(R_{uni}, Z_{\overline{\mathcal{C}}_i, t^n}\right) > n^3 \lambda\right)$$

$$
\begin{aligned}
&= Pr\left(\exp\left(\sum_{i=1}^{n^3}\frac{1}{n}2\chi\left(R_{uni},Z_{\overline{\mathcal{C}}_i,t^n}\right)\right) > \exp(\frac{1}{n}2n^3\lambda)\right)\\
&\leq \exp\left(-2n^2\lambda\right)\prod_{i=1}^{n^3}\mathbb{E}_G\exp\left(\frac{1}{n}2\chi\left(R_{uni},Z_{\overline{\mathcal{C}}_i,t^n}\right)\right)\\
&= \exp\left(-2n^2\lambda\right)\mathbb{E}_G\exp\left(\sum_{i=1}^{n^3}\frac{1}{n}2\chi\left(R_{uni},Z_{\overline{\mathcal{C}}_i,t^n}\right)\right)\\
&\leq \exp\left(-2n^2\lambda\right)\prod_{i=1}^{n^3}\mathbb{E}_G\left[1+\sum_{k=1}^{\infty}\frac{2^k\frac{1}{n}\chi\left(R_{uni},Z_{\overline{\mathcal{C}}_i,t^n}\right)}{k!}\right]\\
&= \exp\left(-2n^2\lambda\right)\left[1+\sum_{k=1}^{\infty}\frac{2^k\frac{1}{n}\mathbb{E}_G\chi\left(R_{uni},Z_{\overline{\mathcal{C}}_i,t^n}\right)}{k!}\right]^{n^3}\\
&\leq \exp\left(-2n^2\lambda\right)\left[1+\sum_{k=1}^{\infty}\frac{2^k\epsilon}{nk!}\right]^{n^3}\\
&= \exp\left(-2n^2\lambda\right)\left[1+\frac{1}{n}\epsilon\exp 2\right]^{n^3}\quad, \qquad (107)
\end{aligned}
$$

the second inequality holds because the right side is part of the Taylor series.

We fix $n\in\mathbb{N}$ and define

$$h_n(x) := n\log(1+\frac{1}{n}e^2x) - x\,.$$

We have $h_n(0)=0$ and

$$
\begin{aligned}
&h_n'(x)\\
&= n\frac{1}{1+\frac{1}{n}e^2x}\frac{1}{n}e^2 - 1\\
&= \frac{ne^2}{e^2x+n} - 1\,.
\end{aligned}
$$

$\frac{ne^2}{e^2x+n}-1$ is positive if $x<\frac{e^2-1}{e}n$, thus if $\hat{c}<\frac{e^2-1}{e}n$, $h_n(x)$ is strictly monotonically increasing in the interval $]0,\hat{c}]$. Thus $h_n(x)$ is positive for $0<x\leq\hat{c}$. For every positive $\hat{c}$, $\hat{c}<\frac{e^2-1}{e}n$ holds if $n>\frac{e}{e^2-1}\hat{c}$. Thus for any positive $\epsilon$, $\epsilon\leq n\log(1+\frac{1}{n}\epsilon\exp 2)$ if $n$ is large enough. Choose $\lambda\geq 2\epsilon$ and let $n$ be sufficiently large, we have $\lambda\geq\log(1+\frac{1}{n}\epsilon\exp 2)$, therefore

$$
\begin{aligned}
&\exp\left(-2\lambda n^2\right)\left[1+\frac{1}{n}\epsilon\exp 2\right]^{n^3}\\
&= \exp\left(-\lambda n^2\right)\exp\left(n^2(-\lambda+n\log(1+\frac{1}{n}\epsilon\exp 2))\right)\\
&\leq \exp\left(-\lambda n^2\right)\,. \qquad (108)
\end{aligned}
$$

By (107) and (108)

$$\begin{aligned} & P\left(\sum_{i=1}^{n^3}\chi\left(R_{uni}, Z_{\overline{\mathcal{C}}_i,t^n}\right) > \lambda n^3 \ \forall t^n \in \theta^n\right) \\ & < |\theta|^n \exp(-\lambda n^2) \\ & = \exp(n\log|\theta| - \lambda n^2) \\ & = \exp(-n\lambda)\,. \end{aligned} \tag{109}$$

In a similar way as $(107)$, choose $\lambda \geq 2\zeta$, we can show that

$$Pr\left(\sum_{i=1}^{n^3} P_e(\overline{\mathcal{C}}_i, t^n) > \lambda n^3 \ \forall t^n \in \theta^n\right) < e^{-\lambda n}\,. \tag{110}$$

Let $\lambda := \max\{2\epsilon, 2\zeta\}$, we have

$$\begin{aligned} & Pr\left(\sum_{i=1}^{n^3} P_e(\overline{\mathcal{C}}_i, t^n) > \lambda n^3 \text{ or } \sum_{i=1}^{n^3}\chi\left(R_{uni}, Z_{\overline{\mathcal{C}}_i,t^n}\right) > \lambda n^3 \ \forall t^n \in \theta^n\right) \\ & \leq 2e^{-\lambda n^3}\,. \end{aligned}$$

We denote the event

$$\begin{aligned} \mathbf{E}_n := \Bigg\{ & \mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_{n^3} \in \mathcal{C}'_\nu : \frac{1}{n^3}\sum_{i=1}^{n^3} P_e(\mathcal{C}_i, t^n) \leq \lambda \\ & \text{and } \frac{1}{n^3}\sum_{i=1}^{n^3}\chi\left(R_{uni}, Z_{\mathcal{C}_i,t^n}\right) \leq \lambda \Bigg\}\,. \end{aligned}$$

If $n$ is large enough, then $P(\mathbf{E}_n)$ is positive. This means $\mathbf{E}_n$ is not the empty set, since $P(\emptyset) = 0$ by definition. Thus there exist codes $\mathcal{C}_i = \left(E_i^n, \left\{D_{j,i}^n : j = 1, \ldots, J_n\right\}\right) \in \mathcal{C}'_\nu$ for $i \in \{1, \ldots, n^3\}$, with a positive probability such that

$$\frac{1}{n^3}\sum_{i=1}^{n^3} P_e(\mathcal{C}_i, t^n) < \lambda \text{ and } \frac{1}{n^3}\sum_{i=1}^{n^3}\chi\left(R_{uni}, Z_{\mathcal{C}_i,t^n}\right) \leq \lambda\,. \tag{111}$$

By (111), for any $n \in \mathbb{N}$ and positive $\lambda$, if there is an $(n, J_n)$ randomness assisted code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$ for $\{(W_t, V_t) : t \in \theta\}$ such that

$$\max_{t^n \in \theta^n}\int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < \lambda\,,$$

$$\max_{t^n \in \theta^n}\int_\Lambda \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma,t^n}\right) dG(\gamma) < \lambda\,,$$

there is also an $(n, J_n)$ common randomness assisted code $\{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_{n^3}\}$ such that

$$\max_{t^n \in \theta^n}\frac{1}{n^3}\sum_{i=1}^{n^3} P_e(\mathcal{C}_i, t^n) < \lambda\,,$$

$$\max_{t^n \in \theta^n} \frac{1}{n^3} \sum_{i=1}^{n^3} \chi\left(R_{uni}, Z_{\mathcal{C}_i, t^n}\right) < \lambda \,.$$

Therefore we have

$$C_s(\{(W_t, V_t) : t \in \theta\}; cr) \geq C_s(\{(W_t, V_t) : t \in \theta\}; r) \,.$$

This and the fact that

$$C_s(\{(W_t, V_t) : t \in \theta\}; cr) \leq C_s(\{(W_t, V_t) : t \in \theta\}; r) \,,$$

prove Theorem 4.1. 2.

*ii) Proof of Theorem 4.1. 1a*

Now we are going to use Theorem 4.1. 2 to prove Theorem 4.1. 1a.

To show the lower bound in Theorem 4.1. 1a, we build a two-part code word, which consists of a non-secure code word and a common randomness assisted secure code word. The non-secure one is used to create the common randomness for the sender and the legal receiver. The common randomness assisted secure code word is used to transmit the message to the legal receiver.

Choose arbitrary positive $\epsilon$ and $\zeta$. Assume we have an $(n, J_n)$ randomness assisted code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$ for $\{(W_t, V_t) : t \in \theta\}$ such that

$$\max_{t^n \in \theta^n} \int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < \epsilon \,,$$

$$\max_{t^n \in \theta^n} \int_\Lambda \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG(\gamma) < \zeta \,,$$

by Theorem 4.1. 2, there is also an $(n, J_n)$ common randomness assisted code $\{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_{n^3}\}$ such that

$$\max_{t^n \in \theta^n} \frac{1}{n^3} \sum_{i=1}^{n^3} P_e(\mathcal{C}_i, t^n) < \lambda \,, \tag{112}$$

$$\max_{t^n \in \theta^n} \frac{1}{n^3} \sum_{i=1}^{n^3} \chi\left(R_{uni}, Z_{\mathcal{C}_i, t^n}\right) < \lambda \,, \tag{113}$$

where $\lambda := \max\{2\epsilon, 2\zeta\}$.

If the arbitrarily varying classical-quantum channel $\{W_t : x \in \mathcal{X}\}$ is not symmetrizable, then by [6], the capacity for message transmission of $\{W_t : x \in \mathcal{X}\}$ is positive. By Remark 2.21 we may assume that the capacity for message transmission of $\{W_t : x \in \mathcal{X}\}$ using deterministic encoder is positive. This means for any positive $\vartheta$, if $n$ is sufficiently large, there is a code $\left(\left(c_i^{\mu(n)}\right)_{i \in \{1, \cdots, n^3\}}, \{D_i^{\mu(n)} : i \in \{1, \cdots, n^3\}\}\right)$ with deterministic encoder of length $\mu(n)$, where $2^{\mu(n)} = o(n)$ such that

$$1 - \frac{1}{n^3} \sum_{i=1}^{n^3} \mathrm{tr}(W_{t^n}(c_i^{\mu(n)}) D_i^{\mu(n)}) \leq \vartheta \,. \tag{114}$$

We now can construct a code $\mathcal{C}^{det} = \left(E^{\mu(n)+n}, \left\{D_j^{\mu(n)+n} : j = 1, \cdots, J_n\right\}\right)$, where for $a^{\mu(n)+n} = (a^{\mu(n)}, a^n) \in \mathbf{A}^{\mu(n)+n}$

$$E^{\mu(n)+n}(a^{\mu(n)+n}|j) = \begin{cases} \frac{1}{n^3} E_i^n(a^n|j) \text{ if } a^{\mu(n)} = c_i^{\mu(n)} \\ 0 \text{ else} \end{cases} ,$$

and

$$D_j^{\mu(n)+n} := \sum_{i=1}^{n^3} D_i^{\mu(n)} \otimes D_{i,j}^n .$$

It is a composition of the code $\left(c_i^{\mu(n)}\right)_{i=1,\cdots,n^3}, \{D_i^{\mu(n)} : i = 1, \cdots, n^3\})$ and the code $\mathcal{C}_i = (E_i^n, \{D_{i,j}^n : j = 1, \cdots, J_n\}$. This is a code of length $\mu(n) + n$.

*iii) This Code Is Secure Against Eavesdropping*

We are going to show that the two-part code word is secure when the common randomness assisted part is secure. Since the two-part code can be seen as a function of its common randomness assisted part the idea is similar to applying the quantum data processing inequality (cf. [66]) when we consider quantum mutual information as security criterion.

For any $i \in \{1, \cdots, n^3\}$ let

$$\begin{aligned} &\mathfrak{Z}_{i,t^{\mu(n)+n}} \\ &:= \left\{V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,1)), \cdots, V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,J_n))\right\} . \end{aligned}$$

For any $t^{\mu(n)+n} = (t^{\mu(n)}, t^n)$ we have

$$\begin{aligned} &\chi\left(R_{uni}, \mathfrak{Z}_{i,t^{\mu(n)+n}}\right) \\ &= S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{a^n\in\mathbf{A}^n} E_i^n(a^n \mid j) V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(a^n)\right) \\ &- \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(\sum_{a^n\in\mathbf{A}^n} E_i^n(a^n \mid j) V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(a^n)\right) \\ &= S\left(V_{t^{\mu(n)}}(c_i^{\mu(n)})\right) + S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{a^n\in\mathbf{A}^n} E_i^n(a^n \mid j) V_{t^n}(a^n)\right) - S\left(V_{t^{\mu(n)}}(c_i^{\mu(n)})\right) \\ &- \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(\sum_{a^n\in\mathbf{A}^n} E_i^n(a^n \mid j) V_{t^n}(a^n)\right) \\ &= S\left(\frac{1}{J_n}\sum_{j=1}^{J_n} V_{t^n}(E_i^n(\,|\,j))\right) - \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(V_{t^n}(E_i^n(\,|\,j))\right) \qquad (115) \\ &= \chi\left(R_{uni}, Z_{\mathcal{C}_i,t^n}\right) . \qquad (116) \end{aligned}$$

By definition we have

$$Z_{\mathcal{C}^{det},t^{\mu(n)+n}}$$

$$
\begin{aligned}
&= \left\{ V_{t^{\mu(n)+n}}(E^{\mu(n)+n}(\,|\,1)), \cdots, V_{t^{\mu(n)+n}}(E^{\mu(n)+n}(\,|\,J_n)) \right\} \\
&= \left\{ \frac{1}{n^3} \sum_{i=1}^{n^3} \sum_{a^n \in \mathbf{A}^n} E_i^n(a^n \,|\, 1) V_{t^{\mu(n)+n}}\left( \left( c_i^{\mu(n)}, a^n \right) \right), \cdots, \right. \\
&\left. \frac{1}{n^3} \sum_{i=1}^{n^3} \sum_{a^n \in \mathbf{A}^n} E_i^n(a^n \,|\, J_n) V_{t^{\mu(n)+n}}\left( \left( c_i^{\mu(n)}, a^n \right) \right) \right\} \\
&= \left\{ \frac{1}{n^3} \sum_{i=1}^{n^3} \sum_{a^n \in \mathbf{A}^n} E_i^n(a^n \,|\, 1) V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(a^n), \cdots, \right. \\
&\left. \frac{1}{n^3} \sum_{i=1}^{n^3} \sum_{a^n \in \mathbf{A}^n} E_i^n(a^n \,|\, J_n) V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(a^n) \right\} \\
&= \left\{ \frac{1}{n^3} \sum_{i=1}^{n^3} V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,1)), \cdots, \right. \\
&\left. \frac{1}{n^3} \sum_{i=1}^{n^3} V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,J_n)) \right\}.
\end{aligned}
$$

By (113) and (116) for any $t^{\mu(n)+n} = (t^{\mu(n)} t^n)$ we have

$$
\begin{aligned}
&\chi\left(R_{uni}, Z_{\mathcal{C}^{det}, t^{\mu(n)+n}}\right) \\
&\leq \chi\left(R_{uni}, Z_{\mathcal{C}^{det}, t^{\mu(n)+n}}\right) - \frac{1}{n^3} \sum_{i=1}^{n^3} \chi\left(R_{uni}, Z_{\mathcal{C}_i, t^n}\right) + \lambda \\
&= \chi\left(R_{uni}, Z_{\mathcal{C}^{det}, t^{\mu(n)+n}}\right) - \frac{1}{n^3} \sum_{i=1}^{n^3} \chi\left(R_{uni}, \mathfrak{Z}_{i, t^{\mu(n)+n}}\right) + \lambda \\
&= S\left( \frac{1}{J_n} \frac{1}{n^3} \sum_{j=1}^{J_n} \sum_{i=1}^{n^3} V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j)) \right) \\
&- \frac{1}{J_n} \sum_{j=1}^{J_n} S\left( \frac{1}{n^3} \sum_{i=1}^{n^3} V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j)) \right) \\
&- \frac{1}{n^3} \sum_{i=1}^{n^3} S\left( \frac{1}{J_n} \sum_{j=1}^{J_n} V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j)) \right) \\
&+ \frac{1}{J_n} \frac{1}{n^3} \sum_{j=1}^{J_n} \sum_{i=1}^{n^3} S\left( V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j)) \right) + \lambda\,. \qquad (117)
\end{aligned}
$$

Let $H^{\mathfrak{H}}$ be a $n^3$-dimensional Hilbert space, spanned by an orthonormal basis $\{|i\rangle : i = 1, \cdots, n^3\}$. Let $H^{\mathfrak{J}}$ be a $J_n$-dimensional Hilbert space, spanned by an orthonormal

basis $\{|j\rangle : j = 1, \cdots, J_n\}$. We define

$$\varphi^{\mathfrak{J}\mathfrak{H}H^{\mu(n)+n}} := \frac{1}{J_n}\frac{1}{n^3}\sum_{j=1}^{J_n}\sum_{i=1}^{n^3} |j\rangle\langle j| \otimes |i\rangle\langle i| \otimes V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\,.$$

We have

$$\varphi^{\mathfrak{J}H^{\mu(n)+n}} = \mathrm{tr}_{\mathfrak{H}}\left(\varphi^{\mathfrak{J}\mathfrak{H}H^{\mu(n)+n}}\right) = \frac{1}{J_n}\frac{1}{n^3}\sum_{j=1}^{J_n}\sum_{i=1}^{n^3} |j\rangle\langle j| \otimes V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\,,$$

$$\varphi^{\mathfrak{H}H^{\mu(n)+n}} = \mathrm{tr}_{\mathfrak{J}}\left(\varphi^{\mathfrak{J}\mathfrak{H}H^{\mu(n)+n}}\right) = \frac{1}{J_n}\frac{1}{n^3}\sum_{j=1}^{J_n}\sum_{i=1}^{n^3} |i\rangle\langle i| \otimes V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\,,$$

$$\varphi^{H^{\mu(n)+n}} = \mathrm{tr}_{\mathfrak{J}\mathfrak{H}}\left(\varphi^{\mathfrak{J}\mathfrak{H}H^{\mu(n)+n}}\right) = \frac{1}{J_n}\frac{1}{n^3}\sum_{j=1}^{J_n}\sum_{i=1}^{n^3} V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\,.$$

Furthermore,

$$\begin{aligned}
&S(\varphi^{\mathfrak{J}H^{\mu(n)+n}})\\
&= S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{i=1}^{n^3} |j\rangle\langle j| \otimes V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\right)\\
&= H(R_{uni}) + \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(\frac{1}{n^3}\sum_{i=1}^{n^3} V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\right)\,,
\end{aligned}$$

$$\begin{aligned}
&S(\varphi^{\mathfrak{H}H^{\mu(n)+n}})\\
&= S\left(\frac{1}{J_n}\frac{1}{n^3}\sum_{j=1}^{J_n}\sum_{i=1}^{n^3} |i\rangle\langle i| \otimes V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\right)\\
&= H(Y_{uni}) + \frac{1}{n^3}\sum_{i=1}^{n^3} S\left(\frac{1}{J_n}\sum_{j=1}^{J_n} V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\right)\,,
\end{aligned}$$

$$\begin{aligned}
&S(\varphi^{\mathfrak{J}\mathfrak{H}H^{\mu(n)+n}})\\
&= S\left(\frac{1}{J_n}\frac{1}{n^3}\sum_{j=1}^{J_n}\sum_{i=1}^{n^3} |j\rangle\langle j| \otimes |i\rangle\langle i| \otimes V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\right)\\
&= H(R_{uni}) + H(Y_{uni}) + \frac{1}{J_n}\frac{1}{n^3}\sum_{j=1}^{J_n}\sum_{i=1}^{n^3} S\left(V_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes V_{t^n}(E_i^n(\,|\,j))\right)\,,
\end{aligned}$$

By strong subadditivity of von Neumann entropy it holds $S(\varphi^{\mathfrak{J}H^{\mu(n)+n}})+S(\varphi^{\mathfrak{H}H^{\mu(n)+n}}) \geq S(\varphi^{H^{\mu(n)+n}}) + S(\varphi^{\mathfrak{J}\mathfrak{H}H^{\mu(n)+n}})$. Thus by (117) we have

$$\chi\left(R_{uni}, Z_{\mathcal{C}^{det},t^{\mu(n)+n}}\right) \leq \lambda\,. \tag{118}$$

*iv) The Legal Receiver Is Able To Decode the Message*

We now use Theorem 4.1. 2 to show that the legal receiver's average error goes to zero.

For any $t^{\mu(n)+n} \in \theta^{\mu(n)+n}$, by (112) and (113),

$$\begin{aligned}
&P_e(\mathcal{C}^{det}, t^{\mu(n)+n}) \\
&= 1 - \frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}\left( \left[\frac{1}{n^3}\sum_{i=1}^{n^3} U_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes U_{t^n}(E_i^n(\,|\,j))\right] \cdot \left[\sum_{k=1}^{n^3} D_k^{\mu(n)} \otimes D_{k,j}^n\right]\right) \\
&\leq 1 - \frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}\left( \frac{1}{n^3}\sum_{i=1}^{n^3} \left[U_{t^{\mu(n)}}(c_i^{\mu(n)}) \otimes U_{t^n}(E_i^n(\,|\,j))\right] \cdot \left[D_k^{\mu(n)} \otimes D_{k,j}^n\right]\right) \\
&= 1 - \frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}\left( \frac{1}{n^3}\sum_{i=1}^{n^3} \left[U_{t^{\mu(n)}}(c_i^{\mu(n)}) D_k^{\mu(n)}\right] \otimes \left[U_{t^n}(E_i^n(\,|\,j)) D_{k,j}^n\right]\right) \\
&= 1 - \frac{1}{n^3}\sum_{i=1}^{n^3} \left( \mathrm{tr}\left[U_{t^{\mu(n)}}(c_i^{\mu(n)}) D_k^{\mu(n)}\right] \cdot \frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}\left[U_{t^n}(E_i^n(\,|\,j)) D_{k,j}^n\right]\right) \\
&= 1 - \frac{1}{n^3}\sum_{i=1}^{n^3} \left( \mathrm{tr}\left[U_{t^{\mu(n)}}(c_i^{\mu(n)}) D_k^{\mu(n)}\right] \cdot (1 - P_e(C_i, t^n))\right) \\
&\leq 1 - (1 - \vartheta - \lambda) \\
&= \lambda + \vartheta\,,
\end{aligned} \tag{119}$$

the second inequality holds because for non-negative numbers $\{\alpha_i, \beta_i : i = 1, \cdots, M\}$ such that $\frac{1}{M}\sum_{i=1}^M \alpha_i \leq \vartheta$ and $\frac{1}{M}\sum_{i=1}^M \beta_i \leq \lambda$ we have $\frac{1}{M}\sum_{i=1}^M (1-\alpha_i)(1-\beta_i) \geq 1 - \vartheta - \lambda$.

For any $n \in \mathbb{N}$ and positive $\lambda$, if there is an $(n, J_n)$ randomness assisted code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$ for $\{(W_t, V_t) : t \in \theta\}$ such that

$$\max_{t^n \in \theta^n} \int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < \epsilon\,,$$

$$\max_{t^n \in \theta^n} \int_\Lambda \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG(\gamma) < \zeta\,,$$

choose $\delta = \max\{2\epsilon, 2\zeta\} + \vartheta$, by (119) and (118), we can find a $(\mu(n)+n, J_n)$ deterministic code $\mathcal{C}^{det} = \left(E^{\mu(n)+n}, \{D_j^{\mu(n)+n} : j = 1, \cdots, J_n\}\right)$ such that such that

$$\max_{t^{\mu(n)+n} \in \theta^{\mu(n)+n}} P_e(\mathcal{C}^{det}, t^{\mu(n)+n}) < \delta\,,$$

$$\max_{t^{\mu(n)+n} \in \theta^{\mu(n)+n}} \chi\left(R_{uni}, Z_{\mathcal{C}^{det}, t^{\mu(n)+n}}\right) < \delta\,.$$

We know that $2^{\mu(n)} = o(n)$. For any positive $\varepsilon$, if $n$ is large enough we have $\frac{1}{n}\log J_n - \frac{1}{\log n+n}\log J_n \leq \varepsilon$. Therefore, if the arbitrarily varying classical-quantum channel $\{W_t : x \in \mathcal{X}\}$ is not symmetrizable, we have

$$C_s(\{(W_t, V_t) : t \in \theta\}; cr) \geq C_s(\{(W_t, V_t) : t \in \theta\}; r) - \varepsilon\,. \tag{120}$$

This and the fact that

$$C_s(\{(W_t, V_t) : t \in \theta\}; cr) \leq C_s(\{(W_t, V_t) : t \in \theta\}; r)$$

prove Theorem 4.1. 1a (c.f. [6] for Ahlswede dichotomy for arbitrarily varying classical-quantum channel Channels).

*v) The Proof of Theorem 4.1. 1b*

If $\{W_t : t \in \theta\}$ is symmetrizable, the deterministic capacity of $\{W_t : t \in \theta\}$ using a deterministic encoder is equal to zero by [6]. Now we have to check whether $C_s(\{(W_t, V_t) : t \in \theta\})$ using stochastic encoder remains equal to zero. The proof is rather standard. Readers with experiences in information theory may pass over this subsection.

For any $n \in \mathbb{N}$ and $J_n \in \mathbb{N} \setminus \{1\}$ let $\mathcal{C} = \Big(E^n, \{D_j^n : j \in \{1, \cdots J_n\}\}\Big)$ be an $(n, J_n)$ deterministic code with a random encoder. We denote the set of all deterministic encoders by $\mathbf{F}_n := \Big\{f_n : \{1, \cdots, J_n\} \to \mathbf{A}^n\Big\}$. Since the deterministic capacity of $\{W_t : t \in \theta\}$ using deterministic encoder is zero, there is a positive $c$ such that for any $n \in \mathbb{N}$ we have

$$\max_{t^n \in \theta^n} \frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}\Big(W_{t^n}(f_n(j))D_j^n\Big) < 1 - c\,. \tag{121}$$

For any $t^n \in \theta^n$, we have

$$\begin{aligned}
&1 - c\\
&= (1-c)\sum_{f_n \in \mathbf{F}_n}\prod_{k=1}^{J_n} E^n(f_n(k) \mid k)\\
&> \sum_{f_n \in \mathbf{F}_n}\prod_{k=1}^{J_n} E^n(f_n(k) \mid k)\frac{1}{J_n}\sum_{j=1}^{J_n}\mathrm{tr}\Big(W_{t^n}(f_n(j))D_j^n\Big)\\
&= \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{a^n \in \mathbf{A}^n} E^n(a^n \mid j)\mathrm{tr}\Big(W_{t^n}(a^n)D_j^n\Big)\\
&= \frac{1}{J_n}\sum_{j=1}^{J_n}\mathrm{tr}\Big(W_{t^n}(E^n(\,\mid j))D_j^n\Big)\,,
\end{aligned} \tag{122}$$

the first equation holds because

$$\sum_{f_n \in \mathbf{F}_n}\prod_{j=1}^{J_n} E^n(f_n(j) \mid j)$$

$$
\begin{aligned}
&= \sum_{a^n} \sum_{f_n(1)=a^n} \left( \sum_{a^n} \sum_{f_n(2)=a^n} \left( \cdots \left( \sum_{a^n} \sum_{f_n(J_n-1)=a^n} \left( \sum_{a^n} \sum_{f_n(J_n)=a^n} \prod_{j=1}^{J_n} E^n(f_n(j) \mid j) \right) \right) \cdots \right) \right) \\
&= \sum_{a^n} \sum_{f_n(1)=a^n} \left( \sum_{a^n} \sum_{f_n(2)=a^n} \left( \cdots \left( \sum_{a^n} \sum_{f_n(J_n-1)=a^n} \left( \sum_{a^n} E^n(a^n \mid J_n) \prod_{j=1}^{J_n-1} E^n(f_n(j) \mid j) \right) \right) \cdots \right) \right) \\
&= \sum_{a^n} \sum_{f_n(1)=a^n} \left( \sum_{a^n} \sum_{f_n(2)=a^n} \left( \cdots \left( \sum_{a^n} \sum_{f_n(J_n-1)=a^n} \prod_{j=1}^{J_n-1} E^n(f_n(j) \mid j) \right) \cdots \right) \right) \\
&= \sum_{a^n} \sum_{f_n(1)=a^n} \left( \sum_{a^n} \sum_{f_n(2)=a^n} \left( \cdots \left( \sum_{a^n} E^n(a^n \mid J_n) \prod_{j=1}^{J_n-1} E^n(f_n(j) \mid j) \right) \cdots \right) \right) \\
&= \cdots \\
&= \sum_{a^n} \sum_{f_n(1)=a^n} E^n(f_n(1) \mid 1) \\
&= \sum_{a^n} E^n(a^n \mid 1) \\
&= 1 \,,
\end{aligned}
$$

the second equation holds because because for any $j \in \{1, \cdots, J_n\}$, we have

$$
\begin{aligned}
&\sum_{f_n \in \mathbf{F}_n} \prod_{k=1}^{J_n} E^n(f_n(k) \mid k) \mathrm{tr}\left( W_{t^n}(f_n(j)) D_j^n \right) \\
&= \sum_{a^n} \sum_{f_n(j)=a^n} E^n(a^n \mid j) \left( \prod_{k \neq j} E^n(f_n(k) \mid k) \right) \mathrm{tr}\left( W_{t^n}(f_n(j)) D_j^n \right) \\
&= \sum_{a^n} \sum_{f_n(j)=a^n} E^n(a^n \mid j) \mathrm{tr}\left( W_{t^n}(f_n(j)) D_j^n \right) \\
&= \sum_{a^n} E^n(a^n \mid j) \mathrm{tr}\left( W_{t^n}(a^n) D_j^n \right) .
\end{aligned}
$$

By (122), for any $n \in \mathbb{N}$, $J_n \in \mathbb{N} \setminus \{1\}$, let $\mathcal{C}$ be any $(n, J_n)$ deterministic code with a random encoder, if $\{W_t : t \in \theta\}$ is symmetrizable, we have

$$
\max_{t \in \theta} P_e(\mathcal{C}, t^n) > c \,.
$$

Thus the only achievable deterministic secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ is $\log 1 = 0$. Thus $C_s(\{(W_t, V_t) : t \in \theta\}) = 0$. (Actually, (122) shows that if $\{W_t : t \in \theta\}$ is symmetrizable, even the deterministic capacity for message transmission of $\{(W_t, V_t) : t \in \theta\}$ with random encoding technique is equal to zero. Since the deterministic secrecy capacity $C_s(\{(W_t, V_t) : t \in \theta\})$ cannot exceed the deterministic capacity for message transmission, we have $C_s(\{(W_t, V_t) : t \in \theta\}) = 0$). This completes the proof of Theorem 4.1. 1b. □

## 4.2 Secrecy Capacity under Common Randomness Assisted Quantum Coding

In this Section we determine the secrecy capacities under common randomness assisted coding of arbitrarily varying classical-quantum wiretap channels.

### 4.2.1 Compound-Arbitrarily Varying Wiretap Classical-Quantum Channel

Let $A$, $B$, $H$, $\overline{\theta}$, $\theta$, and $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$ be defined as in Section 2.

Following the idea of [65], we first prove the following Theorem.

**Theorem 4.2.** *Let $\overline{\theta} := \{1, \cdots, \overline{T}\}$ and $\theta := \{1, \cdots, T\}$ be finite index sets. Let $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$ be a compound-arbitrarily varying wiretap classical-quantum channel. We have*

$$\begin{aligned}&\hat{C}_s(\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}) \\ &= \lim_{n\to\infty} \frac{1}{n} \max_{U \to A^n \to \{B_s^{\otimes n}, Z_{t^n} : s, t_n\}} \left( \min_{s\in\overline{\theta}} \chi(p_U; B_s^{\otimes n}) - \max_{t^n \in \theta^n} \chi(p_U; Z_{t^n}) \right), \quad (123)\end{aligned}$$

*where $B_s$ are the resulting quantum states at the output of the legal receiver's channels. $Z_{t^n}$ are the resulting quantum states at the output of wiretap channels. The maximum is taken over all random variables that satisfy the Markov chain relationships: $U \to A^n \to B_s^{\otimes n} Z_{t^n}$ for every $s \in \overline{\theta}$ and $t^n \in \theta^n$. $A$ is here a random variable taking values on $\mathbf{A}$, $U$ a random variable taking values on some finite set $\mathbf{U}$ with probability distribution $p_U$.*

*Proof.* We fix a probability distribution $p \in P(\mathbf{A})$. Let

$$J_n = \lfloor 2^{n \min_{s\in\overline{\theta}} \chi(p;B_s) - \log L_n - 2n\mu} \rfloor .$$

Let $p'(x^n) := \begin{cases} \frac{p^n(x^n)}{p^n(\mathcal{T}^n_{p,\delta})} ; & \text{if } x^n \in \mathcal{T}^n_{p,\delta} \\ 0 , & \text{else.} \end{cases}$

Let $X^n := \{X_{j,l}\}_{j\in\{1,\ldots,J_n\}, l\in\{1,\ldots,L_n\}}$ be a family of random variables taking value according to $p'$, i.e. with the uniform distribution over $\mathcal{T}^n_{P,\delta}$. Here $L_n$ is a natural number which will be specified later.

We fix a $t^n \in \theta^n$ and define a map $\mathsf{V} : P(\theta) \times P(\mathbf{A}) \to \mathcal{S}(H)$ by

$$\mathsf{V}(t, p) := V_t(p) .$$

For $t \in \theta$ we define $q(t) := \frac{N(t|t^n)}{n}$. $t^n$ is trivially a typical sequence of $q$. For $p \in P(\mathbf{A})$, $\mathsf{V}$ defines a map $\mathsf{V}(\cdot, p) : P(\theta) \to \mathcal{S}(H)$.

Let

$$Q_{t^n}(x^n) := \Pi_{\mathsf{V}(\cdot,p),\alpha}(t^n) \Pi_{\mathsf{V},\alpha}(t^n, x^n) \cdot V_{t^n}(x^n) \cdot \Pi_{\mathsf{V},\alpha}(t^n, x^n) \Pi_{\mathsf{V}(\cdot,p),\alpha}(t^n) .$$

In view of the fact that $\Pi_{\mathsf{V}(\cdot,p),\alpha}(t^n)$ and $\Pi_{\mathsf{V},\alpha}(t^n, x^n)$ are both projection matrices, by (1), (7), and Lemma 3.3 for any $t$ and $x^n$, it holds that

$$\|Q_{t^n}(x^n) - V_{t^n}(x^n)\|_1 \leq \sqrt{2^{-n\beta(\alpha)} + 2^{-n\beta(\alpha)''}} . \quad (124)$$

For our result, we use an alternative Covering Lemma

**Lemma 4.3.** *Let $\mathcal{V}$ be a finite-dimensional Hilbert space. Let $\mathbf{M}$ and $\mathbf{M}' \subset \mathbf{M}$ be finite sets. Suppose we have an ensemble $\{\rho_m : m \in \mathbf{M}\} \subset \mathcal{S}(\mathcal{V})$ of quantum states. Let $p$ be a probability distribution on $\mathbf{M}$.*

*Suppose a total subspace projector $\Pi$ and codeword subspace projectors $\{\Pi_m : m \in \mathbf{M}\}$ exist, which project onto subspaces of the Hilbert space in which the states exist and for all $m \in \mathbf{M}'$ there are positive constants $\epsilon \in ]0,1[$, $D$, $d$ such that the following conditions hold:*

$$\operatorname{tr}(\rho_m \Pi) \geq 1 - \epsilon \,,$$

$$\operatorname{tr}(\rho_m \Pi_m) \geq 1 - \epsilon \,,$$

$$\operatorname{tr}(\Pi) \leq D \,,$$

*and*

$$\Pi_m \rho_m \Pi_m \leq \frac{1}{d} \Pi_m \,.$$

*We denote $\omega := \sum_{m \mathbf{M}'} p(m) \rho_m$. Notice that $\omega$ is not a density operator in general. We define a sequence of i.i.d. random variables $X_1, \ldots, X_L$, taking values in $\{\rho_m : m \in \mathbf{M}\}$. If $L \gg \frac{d}{D}$, then*

$$\begin{aligned} & Pr\Bigg( \| L^{-1} \sum_{i=1}^{L} \Pi \cdot \Pi_{X_i} \cdot X_i \cdot \Pi_{X_i} \cdot \Pi - \omega \|_1 \\ & \leq 1 - p(\mathbf{M}') + 4\sqrt{1 - p(\mathbf{M}')} + 42 \sqrt[8]{\epsilon} \Bigg) \\ & \geq 1 - 2D \exp\left( -p(\mathbf{M}') \frac{\epsilon^3 L d}{2 \ln 2D} \right) \,. \end{aligned} \tag{125}$$

*Proof.* We define a function $\mathbb{1}_{\mathbf{M}'} : \mathbf{M} \to \mathbf{M}' \cup \{0^{\mathcal{V}}\}$ by

$$\mathbb{1}_{\mathbf{M}'}(\rho_m) := \begin{cases} \rho_m \,, & \text{if } m \in \mathbf{M}' \\ 0^{\mathcal{V}} \,, & \text{if } m \notin \mathbf{M}' \,, \end{cases}$$

where $0^{\mathcal{V}}$ is the zero operator on $\mathcal{V}$, i.e. $\langle j | 0^{\mathcal{V}} | j \rangle = 0$ for all $j \in \mathcal{V}$. Notice that $0^{\mathcal{V}}$ is not a density operator.

We have

$$\begin{aligned} & \operatorname{tr}\left( \sum_{m \in \mathbf{M}} p(m) \mathbb{1}_{\mathbf{M}'}(\rho_m) \right) \\ & = \operatorname{tr}\left( \sum_{m \in \mathbf{M}'} p(m) \rho_m \right) \\ & = \sum_{m \in \mathbf{M}'} p(m) \operatorname{tr}(\rho_m) \\ & = p(\mathbf{M}') \,. \end{aligned} \tag{126}$$

Let $\hat{\Pi}$ be the projector onto the subspace spanned by the eigenvectors of $\sum_{m \in \mathbf{M}'} p(m) \Pi \Pi_m \rho_m \Pi_m \Pi$ whose corresponding eigenvalues are greater than $p(\mathbf{M}') \frac{\epsilon}{D}$.

The following three inequalities can be shown by the same arguments as in the proof of Lemma 3.4 in [66]:

$$\sum_{m\in\mathbf{M}} p(m) d\cdot \hat{\Pi}\Pi\Pi_m \mathbb{1}_{\mathbf{M}'}(\rho_m)\Pi_m\Pi\hat{\Pi} \geq p(\mathbf{M}')\frac{d\epsilon}{D}\hat{\Pi}\,. \tag{127}$$

$$\begin{aligned}
&\|\sum_{m\in\mathbf{M}} p(m)\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi - \sum_{m\in\mathbf{M}} p(m)\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\|_1\\
&\leq \sum_{m\in\mathbf{M}'} p(m)\|\Pi\cdot\Pi_m\rho_m\cdot\Pi_m\cdot\Pi-\rho_m\|_1\\
&\leq \sum_{m\in\mathbf{M}'} p(m)\left(2\sqrt{\epsilon}+2\sqrt{\epsilon+2\sqrt{\epsilon}}\right)\\
&= p(\mathbf{M}')\left(2\sqrt{\epsilon}+2\sqrt{\epsilon+2\sqrt{\epsilon}}\right)\\
&\leq 2\sqrt{\epsilon}+2\sqrt{\epsilon+2\sqrt{\epsilon}}\\
&\leq 6\sqrt[4]{\epsilon}\,.
\end{aligned} \tag{128}$$

The last inequality holds because $\sqrt{\epsilon+2\sqrt{\epsilon}} \leq 2\sqrt[4]{\epsilon}$ for $0\leq\epsilon\leq 1$.

When $\{\rho_1,\cdots,\rho_L\}$ fulfills

$$\begin{aligned}
&(1-\epsilon)\sum_{m\in\mathbf{M}} p(m)\hat{\Pi}\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\hat{\Pi}\\
&\leq L^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\cdot\Pi_{\rho_i}\cdot(\mathbb{1}_{\mathbf{M}'}(\rho_i))\cdot\Pi_{\rho_i}\cdot\Pi\hat{\Pi}\\
&\leq (1+\epsilon)\sum_{m\in\mathbf{M}} p(m)\hat{\Pi}\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\hat{\Pi}\,,
\end{aligned}$$

(i.e. we assume the event considered in (131) below),

then

$$\begin{aligned}
&\|L^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\cdot\Pi_{\rho_i}\cdot(\mathbb{1}_{\mathbf{M}'}(\rho_i))\cdot\Pi_{\rho_i}\cdot\Pi\hat{\Pi}\\
&-\sum_{m\in\mathbf{M}} p(m)\hat{\Pi}\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\hat{\Pi}\|_1\\
&\leq\epsilon
\end{aligned} \tag{129}$$

*i) Application of the Operator Chernoff Bound*

For all $m\in\mathbf{M}'$ we have

$$\begin{aligned}
&d\cdot\hat{\Pi}\Pi\Pi_m\mathbb{1}_{\mathbf{M}'}(\rho_m)\Pi_m\Pi\hat{\Pi}\\
&= d\cdot\hat{\Pi}\Pi\Pi_m\rho_m\Pi_m\Pi\hat{\Pi}\\
&\leq\hat{\Pi}
\end{aligned} \tag{130}$$

as a consequence of the inequality $A^\dagger BA \leq A^\dagger A$ which is valid whenever $B \leq id$.

By (130) and the fact that $d \cdot 0^{\mathcal{V}} \leq \hat{\Pi}$ we have for all $m \in \mathbf{M}$

$$0^{\mathcal{V}} \leq d \cdot \hat{\Pi}\Pi\Pi\Pi_m \mathbb{1}_{\mathbf{M}'}(\rho_m)\Pi_m\Pi\Pi\hat{\Pi} \leq \hat{\Pi} .$$

Now we apply the Operator Chernoff Bound (cf. [66]) on the set of operator $\{d\mathbb{1}_{\mathbf{M}'}(\rho_m) : m \in \mathbf{M}\}$ and the subspace spanned by the eigenvectors of $\sum_{m\in\mathbf{M}'} p(m)\Pi\Pi\Pi_m\rho_m\Pi\Pi\Pi_m$ whose corresponding eigenvalues are greater than $p(\mathbf{M}')\frac{\epsilon}{D}$; here $\hat{\Pi}$ acts as the identity on the subspace.

By (127) we obtain

$$\begin{aligned}
&Pr\Bigg((1-\epsilon)\sum_{m\in\mathbf{M}} p(m)\hat{\Pi}\Pi\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\Pi\hat{\Pi}\\
&\quad\leq L^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\Pi\cdot\Pi_{X_i}\cdot(\mathbb{1}_{\mathbf{M}'}(X_i))\cdot\Pi_{X_i}\cdot\Pi\Pi\hat{\Pi}\\
&\quad\leq (1+\epsilon)L^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\Pi\cdot\Pi_{X_i}\cdot(\mathbb{1}_{\mathbf{M}'}(X_i))\cdot\Pi_{X_i}\cdot\Pi\Pi\hat{\Pi}\\
&= Pr\Bigg(d(1-\epsilon)\sum_{m\in\mathbf{M}} p(m)\hat{\Pi}\Pi\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\Pi\hat{\Pi}\\
&\quad\leq dL^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\Pi\cdot\Pi_{X_i}\cdot(\mathbb{1}_{\mathbf{M}'}(X_i))\cdot\Pi_{X_i}\cdot\Pi\Pi\hat{\Pi}\\
&\quad\leq d(1+\epsilon)\sum_{m\in\mathbf{M}} p(m)\hat{\Pi}\Pi\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\Pi\hat{\Pi}\Bigg)\\
&\geq 1-2D\exp\left(-p(\mathbf{M}')\frac{\epsilon^3 Ld}{2\ln 2D}\right) .
\end{aligned} \tag{131}$$

*ii) Upper Bound for* $\|\sum_{m\in\mathbf{M}} p(m)\mathbb{1}_{\mathbf{M}'}(\rho_m) - \sum_{m\in\mathbf{M}} p(m)\hat{\Pi}\Pi\Pi\Pi_m\mathbb{1}_{\mathbf{M}'}(\rho_m)\Pi_m\Pi\Pi\hat{\Pi}\|_1$

Let $\sum_i \lambda_i|i\rangle\langle i|$ be a spectral decomposition of $\sum_{m\in\mathbf{M}'}\frac{p(m)}{p(\mathbf{M}')}\Pi\Pi\Pi_m\rho_m\Pi_m\Pi\Pi$. In view of the fact that $\hat{\Pi}$ is the projector onto the subspace spanned by the eigenvectors of the density operator $\sum_{m\in\mathbf{M}'}\frac{p(m)}{p(\mathbf{M}')}\Pi\Pi\Pi_m\rho_m\Pi_m\Pi\Pi$ whose corresponding eigenvalues are greater than $\frac{\epsilon}{D}$, we have

$$\begin{aligned}
&\mathrm{tr}\left(\sum_{m\in\mathbf{M}}\frac{p(m)}{p(\mathbf{M}')}\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\right)\\
&-\mathrm{tr}\left(\sum_{m\in\mathbf{M}}\frac{p(m)}{p(\mathbf{M}')}\hat{\Pi}\Pi\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\Pi\hat{\Pi}\right)\\
&=\sum_{\lambda_i\geq\frac{\epsilon}{D}}\lambda_i\\
&\leq\epsilon .
\end{aligned}$$

We apply Lemma 3.3 to obtain

$$\begin{aligned}
&\| \sum_{m\in\mathbf{M}} p(m)\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi - \sum_{m\in\mathbf{M}} p(m)\hat{\Pi}\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\hat{\Pi}\|_1\\
&= p(\mathbf{M}')\| \sum_{m\in\mathbf{M}} \frac{p(m)}{p(\mathbf{M}')}\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi - \sum_{m\in\mathbf{M}} \frac{p(m)}{p(\mathbf{M}')}\hat{\Pi}\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\hat{\Pi}\|_1\\
&\leq 2\sqrt{\epsilon+2\sqrt{\epsilon}}\\
&\leq 4\sqrt[4]{\epsilon}\,. \qquad (132)
\end{aligned}$$

When $\{\rho_1,\cdots,\rho_L\}$ fulfills

$$\begin{aligned}
&\|L^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\cdot\Pi_i\cdot(\mathbb{1}_{\mathbf{M}'}(\rho_i))\cdot\Pi_i\cdot\Pi\hat{\Pi}\\
&-\sum_{m\in\mathbf{M}} p(m)\hat{\Pi}\Pi\cdot\Pi_m\cdot\mathbb{1}_{\mathbf{M}'}(\rho_m)\cdot\Pi_m\cdot\Pi\hat{\Pi}\|_1\\
&\leq\epsilon
\end{aligned}$$

(i.e. we assume the event considered in (131) occurs and thus (129) holds), then by (128) and (132) it holds

$$\begin{aligned}
&\|L^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\cdot\Pi_i\cdot(\mathbb{1}_{\mathbf{M}'}(\rho_i))\cdot\Pi_i\cdot\Pi\hat{\Pi} - \sum_{m\in\mathbf{M}} p(m)\mathbb{1}_{\mathbf{M}'}(\rho_m)\|_1\\
&\leq\epsilon+10\sqrt[4]{\epsilon}\\
&\leq 11\sqrt[4]{\epsilon}\,. \qquad (133)
\end{aligned}$$

*iii) Upper Bound for* $\|L^{-1}\sum_{i=1}^{L}\Pi\Pi_i(\mathbb{1}_{\mathbf{M}'}(\rho_i))\Pi_i\Pi - L^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\Pi_i(\mathbb{1}_{\mathbf{M}'}(\rho_i))\Pi_i\Pi\hat{\Pi}\|_1$

When the event considered in (131) is true, i.e. when (133) holds, then by (126)

$$\begin{aligned}
&\mathrm{tr}\left(L^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\cdot\Pi_i\cdot(\mathbb{1}_{\mathbf{M}'}(\rho_i))\cdot\Pi_i\cdot\Pi\hat{\Pi}\right)\\
&\geq p(\mathbf{M}')-11\sqrt[4]{\epsilon}\,.
\end{aligned}$$

We apply Lemma 3.3 to obtain

$$\begin{aligned}
&\|L^{-1}\sum_{i=1}^{L}\hat{\Pi}\Pi\cdot\Pi_i\cdot(\mathbb{1}_{\mathbf{M}'}(\rho_i))\cdot\Pi_i\cdot\Pi\hat{\Pi} - L^{-1}\sum_{i=1}^{L}\Pi\cdot\Pi_i\cdot(\mathbb{1}_{\mathbf{M}'}(\rho_i))\cdot\Pi_i\cdot\Pi\|_1\\
&\leq 2\sqrt{1-p(\mathbf{M}')+11\sqrt[4]{\epsilon}}\\
&\leq 2\sqrt{1-p(\mathbf{M}')}+22\sqrt[8]{\epsilon}\,, \qquad (134)
\end{aligned}$$

the last inequality holds because $\sqrt{a+b}\leq\sqrt{a}+\sqrt{b}$ for positive $a$ and $b$.

*iv) Upper Bound for* $\|L^{-1}\sum_{i=1}^{L}\Pi\Pi_i\rho_i\Pi_i\Pi - L^{-1}\sum_{i=1}^{L}\Pi\Pi_i(\mathbb{1}_{\mathbf{M}'}(\rho_i))\Pi_i\Pi\|_1$

In view of the fact that $\Pi$ and $\Pi_i$ are projection matrices for every $\rho_i \in \{\rho_1, \cdots, \rho_L\}$ it holds

$$\mathrm{tr}(\Pi_i\rho_l\Pi_i) \leq \mathrm{tr}(\rho_l) = 1$$

and

$$\begin{aligned}&\mathrm{tr}(L^{-1}\sum_{i=1}^{L}\Pi\Pi_i\rho_l\Pi_i\Pi)\\ &\leq \mathrm{tr}(L^{-1}\sum_{i=1}^{L}\Pi_i\rho_l\Pi_i)\\ &\leq 1\,.\end{aligned}$$

When $\{\rho_1, \cdots, \rho_L\}$ fulfills

$$\begin{aligned}&\|L^{-1}\sum_{i=1}^{L}\Pi\cdot\Pi_i\cdot(\mathbb{1}_{\mathbf{M}'}(\rho_i))\cdot\Pi_i\cdot\Pi - \sum_{m\in\mathbf{M}}p(m)\mathbb{1}_{\mathbf{M}'}(\rho_m)\|_1\\ &\leq 2\sqrt{1-p(\mathbf{M}')} + 20\sqrt[8]{\epsilon}\,,\end{aligned}$$

i.e. we assume that the event considered in (131) is true, then by (126) and the triangle inequality we have

$$\begin{aligned}&\mathrm{tr}\left(\Pi\cdot\Pi_i\cdot(\mathbb{1}_{\mathbf{M}'}(\rho_i))\cdot\Pi_i\cdot\Pi\right)\\ &\geq p(\mathbf{M}') - 2\sqrt{1-p(\mathbf{M}')} - 20\sqrt[8]{\epsilon}\,.\end{aligned} \tag{135}$$

Since

$$\begin{aligned}&L^{-1}\sum_{i=1}^{L}\Pi\cdot\Pi_i\cdot\rho_i\cdot\Pi_i\cdot\Pi\\ &= L^{-1}\sum_{i=1}^{L}\Pi\cdot\Pi_i\cdot\mathbb{1}_{\mathbf{M}'}(\rho_i)\cdot\Pi_i\cdot\Pi\\ &+ L^{-1}\sum_{i\notin\mathbf{M}'}\Pi\cdot\Pi_i\cdot\rho_i\cdot\Pi_i\cdot\Pi\,,\end{aligned} \tag{136}$$

we have

$$\begin{aligned}&\|L^{-1}\sum_{i\notin\mathbf{M}'}\Pi\cdot\Pi_i\cdot\rho_i\cdot\Pi_i\cdot\Pi\|_1\\ &= \mathrm{tr}\left(L^{-1}\sum_{i\notin\mathbf{M}'}\Pi\cdot\Pi_i\cdot\rho_i\cdot\Pi_i\cdot\Pi\right)\\ &\leq 1 - p(\mathbf{M}') + 2\sqrt{1-p(\mathbf{M}')} + 20\sqrt[8]{\epsilon}\,,\end{aligned} \tag{137}$$

which implies

$$\|L^{-1}\sum_{i=1}^{L}\Pi\cdot\Pi_i\cdot\rho_i\cdot\Pi_i\cdot\Pi - \sum_{m\in\mathbf{M}}p(m)\mathbb{1}_{\mathbf{M}'}(\rho_m)\|_1$$

$$\leq 1 - p(\mathbf{M}') + 4\sqrt{1 - p(\mathbf{M}')} + 42\sqrt[8]{\epsilon}\,. \tag{138}$$

By (138) we have

$$\begin{aligned} & Pr\Bigg( \|L^{-1}\sum_{i=1}^{L} \Pi\cdot\Pi_{X_i}\cdot X_i\cdot\Pi_{X_i}\cdot\Pi - \sum_{m\in\mathbf{M}} p(m)\cdot \mathbb{1}_{\mathbf{M}'}(\rho_m)\|_1 \\ & \quad \leq 1 - p(\mathbf{M}') + 4\sqrt{1 - p(\mathbf{M}')} + 42\sqrt[8]{\epsilon}\Bigg) \\ & \geq 1 - 2D\exp\left(-p(\mathbf{M}')\frac{\epsilon^3 Ld}{2\ln 2D}\right)\,. \end{aligned} \tag{139}$$

□

By (2) we have

$$\begin{aligned} & \mathrm{tr}(\Pi_{\mathsf{V}(\cdot,p),\alpha}(t^n)) \\ & \leq 2^{n(S(\mathsf{V}(\cdot,p)|q)+\delta(\alpha))} \\ & = 2^{n(\sum_t q(t)\mathsf{V}(t,p)+\delta(\alpha))} \\ & = 2^{n(\sum_t q(t)S(V_t(p))+\delta(\alpha))}\,. \end{aligned} \tag{140}$$

Furthermore, for all $x^n$ holds

$$\begin{aligned} & \Pi_{\mathsf{V},\alpha}(t^n,x^n)V_{t^n}(x^n)\Pi_{\mathsf{V},\alpha}(t^n,x^n) \\ & \leq 2^{-n(S(\mathsf{V}|r)+\delta(\alpha)')}\Pi_{\mathsf{V},\alpha}(t^n,x^n) \\ & = 2^{-n(\sum_{t,x} r(t,x)S(\mathsf{V}(t,x))+\delta(\alpha)')}\Pi_{\mathsf{V},\alpha}(t^n,x^n)\,. \end{aligned} \tag{141}$$

We define

$$\theta' := \left\{t\in\theta : nq(t)\geq\sqrt{n}\right\}\,.$$

By properties of classical typical set (cf. [67] ) there is a positive $\hat{\beta}(\alpha)$ such that

$$\underset{p'}{Pr}\left(x^n\in\left\{x^n\in\mathbf{A}^n : (x_{\mathtt{I}_t})\in\mathcal{T}_{p,\delta}^{nq(t)}\ \forall t\in\theta'\right\}\right) \geq \left(1-2^{-\sqrt{n}\hat{\beta}(\alpha)}\right)^{|\theta|} \geq 1-2^{-\sqrt{n}\frac{1}{2}\hat{\beta}(\alpha)}\,, \tag{142}$$

where $\mathtt{I}_t := \{i\in\{1,\cdots,n\} : t_i = t\}$ is an indicator set that selects the indices $i$ in the sequence $t^n = (t_1,\cdots,t_n)$.

We denote the set $\{x^n : (x_{\mathtt{I}_t})\in\mathcal{T}_{p,\delta}^{nq(t)}\ \forall t\in\theta'\}\subset\mathbf{A}^n$ by $\mathbf{M}_{t^n}$. For all $x^n\in\mathbf{M}_{t^n}$, if $n$ is sufficiently large, we have

$$\begin{aligned} & \left|\sum_{t,x} r(t,x)S(\mathsf{V}(t,x)) - \sum_t q(t)S(V_t|p)\right| \\ & \leq \left|\sum_{t\in\theta',x} r(t,x)S(\mathsf{V}(t,x)) - \sum_{t\in\theta'} q(t)S(V_t|p)\right| \\ & + \left|\sum_{t\notin\theta',x} r(t,x)S(\mathsf{V}(t,x)) - \sum_{t\notin\theta'} q(t)S(V_t|p)\right| \end{aligned}$$

$$\leq \sum_{t\in\theta'}\left|\sum_x r(t,x)S(\mathsf{V}(t,x)) - q(t)S(V_t|p)\right| + 2|\theta|\frac{1}{\sqrt{n}}C$$
$$\leq 2|\theta|\frac{\delta}{n}C + 2|\theta|\frac{1}{\sqrt{n}}C\,, \tag{143}$$

where $C := \max_{t\in\theta}\max_{x\in\mathbf{A}}(S(\mathsf{V}(t,x)) + S(V_t|p))$.

We set $\Theta_{t^n} := \sum_{x^n\in\mathbf{M}_{t^n}} p(x^n)Q_{t^n}(x^n)$. For given $z^n \in \mathbf{M}_{t^n}$ and $t^n \in \theta^n$, $\langle z^n|\Theta_{t^n}|z^n\rangle$ is the expected value of $\langle z^n|Q_{t^n}(x^n)|z^n\rangle$ under the condition $x^n \in \mathbf{M}_{t^n}$.

We choose a positive $\bar\beta(\alpha)$ such that $\bar\beta(\alpha) \leq \min(2^{-n\beta(\alpha)}, 2^{-n\beta(\alpha)'})$, and set $\epsilon := 2^{-n\bar\beta(\alpha)}$. In view of (141) we now apply Lemma 4.3, where we consider the set $\mathbf{M}_{t^n} \subset \mathbf{A}^n$: If $n$ is sufficiently large for all $j$ we have

$$\begin{aligned}
&Pr\left(\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_{t^n}(X_{j,l}) - \Theta_{t^n}\|_1 > 2^{-\sqrt{n}\frac{1}{8}\hat\beta(\alpha)} + 40\sqrt[8]{\epsilon}\right)\\
&\leq 2^{n(\sum_{t,x} r(t,x)S(\mathsf{V}(t,x))+\delta(\alpha))}\\
&\cdot\exp\left(-L_n\frac{\epsilon^3}{2\ln 2}(1-2^{-\sqrt{n}\frac{1}{2}\hat\beta(\alpha)})\cdot 2^{n(\sum_t q(t)S(V_t(p))-\sum_t q(t)S(V_t|p))+\delta(\alpha)+\delta(\alpha)'+2|\theta|\frac{\delta}{n}C+2|\theta|\frac{1}{\sqrt{n}}C}\right)\\
&= 2^{n(\sum_{t,x} r(t,x)S(\mathsf{V}(t,x))+\delta(\alpha)}\\
&\cdot\exp\left(-L_n\frac{\epsilon^3}{2\ln 2}\cdot(1-2^{-\sqrt{n}\frac{1}{2}\hat\beta(\alpha)})2^{n(-\sum_t q(t)\chi(p;Z_t)+\delta(\alpha)+\delta(\alpha)'+2|\theta|\frac{\delta}{n}C+2|\theta|\frac{1}{\sqrt{n}}C)}\right)\,.
\end{aligned} \tag{144}$$

The equality holds since $S(V_t(p)) - S(V_t|p) = \chi(p;Z_t)$.

Furthermore,

$$\begin{aligned}
&Pr\left(\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_{t^n}(X_{j,l}) - \Theta_{t^n}\|_1 > 2^{-\sqrt{n}\frac{1}{8}\hat\beta(\alpha)} + 40\sqrt[8]{\epsilon}\ \forall t^n\ \forall j\right)\\
&\leq J_n|\theta|^n 2^{n(\sum_{t,x} r(t,x)S(\mathsf{V}(t,x))+\delta(\alpha)}\\
&\cdot\exp\left(-L_n\frac{\epsilon^3}{2\ln 2}(1-2^{-\sqrt{n}\frac{1}{2}\hat\beta(\alpha)})2^{n(-\sum_t q(t)\chi(p;Z_t)+\delta(\alpha)+\delta(\alpha)'+2|\theta|\frac{\delta}{n}C+2|\theta|\frac{1}{\sqrt{n}}C)}\right)\,.
\end{aligned} \tag{145}$$

Let $\phi_t^j$ be the quantum state at the output of wiretapper's channel when the channel state is $t$ and $j$ has been sent. We have

$$\begin{aligned}
&\sum_{t\in\theta} q(t)\chi\left(p;Z_t\right) - \chi\left(p;\sum_t q(t)Z_t\right)\\
&= \sum_{t\in\theta} q(t)S\left(\sum_{j=1}^{J_n}\frac{1}{J_n}\phi_t^j\right) - \sum_{t\in\theta}\sum_{j=1}^{J_n} q(t)\frac{1}{J_n}S\left(\phi_t^j\right)\\
&- S\left(\sum_{t\in\theta}\sum_{j=1}^{J_n} q(t)\frac{1}{J_n}\phi_t^j\right) + \sum_{j=1}^{J_n}\frac{1}{J_n}S\left(\sum_{t\in\theta} q(t)\phi_t^j\right)\,.
\end{aligned}$$

Let $H^{\mathfrak{T}}$ be a $|\theta|$-dimensional Hilbert space spanned by an orthonormal basis $\{|t\rangle : t = 1, \cdots, |\theta|\}$. Let $H^{\mathfrak{J}}$ be a $J_n$-dimensional Hilbert space, spanned by an orthonormal basis $\{|j\rangle : j = 1, \cdots, J_n\}$. Similar to (118) we define

$$\varphi^{\mathfrak{J}\mathfrak{T}H^n} := \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t)|j\rangle\langle j| \otimes |t\rangle\langle t| \otimes \phi_t^j \ .$$

We have

$$\varphi^{\mathfrak{J}H^n} = \mathrm{tr}_{\mathfrak{T}}\left(\varphi^{\mathfrak{J}\mathfrak{T}H^n}\right) = \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t)|j\rangle\langle j| \otimes \phi_t^j \ ;$$

$$\varphi^{\mathfrak{T}H^n} = \mathrm{tr}_{\mathfrak{J}}\left(\varphi^{\mathfrak{J}\mathfrak{T}H^n}\right) = \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t)|t\rangle\langle t| \otimes \phi_t^j \ ;$$

$$\varphi^{H^n} = \mathrm{tr}_{\mathfrak{J}}\mathfrak{T}\left(\varphi^{\mathfrak{J}\mathfrak{T}H^n}\right) = \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t)\phi_t^j \ .$$

Thus,

$$S(\varphi^{\mathfrak{J}H^n}) = H(R_{uni}) + \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(\sum_{t\in\theta} q(t)\phi_t^j\right) \ ;$$

$$S(\varphi^{\mathfrak{T}H^n}) = H(Y_q) + \sum_{t\in\theta} q(t) S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\phi_t^j\right) \ ;$$

$$S(\varphi^{\mathfrak{J}\mathfrak{T}H^n}) = H(R_{uni}) + H(Y_q) + \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t) S\left(\phi_t^j\right) \ ,$$

where $Y_q$ is a random variable on $\theta$ with distribution $q(t)$.

By strong subadditivity of von Neumann entropy it holds $S(\varphi^{\mathfrak{J}H^n}) + S(\varphi^{\mathfrak{T}H^n}) \geq S(\varphi^{H^n}) + S(\varphi^{\mathfrak{i}\mathfrak{T}H^n})$, therefore

$$\sum_t q(t)\chi\left(p; Z_t\right) - \chi\left(p; \sum_t q(t) Z_t\right) \geq 0 \ . \tag{146}$$

For an arbitrary $\zeta$ we define $L_n = \lceil 2^{\max_{t^n}\chi(p;Z_{t^n})+n\zeta}\rceil$, and choose a suitable $\alpha$, $\bar{\beta}(\alpha)$, and sufficiently large $n$ such that $6\bar{\beta}(\alpha) + 2\delta(\alpha) + 2\delta(\alpha)' + 2|\theta|\frac{\delta}{n}C + 2|\theta|\frac{1}{\sqrt{n}}C \leq \zeta$. By (146), if $n$ is sufficiently large, we have $L_n \geq \lceil 2^{n(\sum_t q(t)\chi(p;Z_t)+\zeta)}\rceil$ and

$$L_n \frac{\epsilon^3}{2\ln 2}(1 - 2^{-\sqrt{n}\frac{1}{2}\hat{\beta}(\alpha)})2^{n(-\sum_t q(t)\chi(p;Z_t)+\delta(\alpha)+\delta(\alpha)'+2|\theta|\frac{\delta}{n}C+2|\theta|\frac{1}{\sqrt{n}}C)} > 2^{\frac{1}{2}n\zeta} \ .$$

When $n$ is sufficiently large for any positive $\vartheta$

$$\begin{aligned} & J_n|\theta|^n 2^{n(\sum_{t,x} r(t,x)S(\mathsf{V}(t,x))+\delta(\alpha)}\exp(-2^{\frac{1}{4}n\zeta}) \\ & \leq 2^{-n\vartheta} \end{aligned}$$

and

$$2^{-\sqrt{n}\frac{1}{8}\hat{\beta}(\alpha)} + 40\sqrt[8]{\epsilon} \leq 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)} \ .$$

Thus for sufficiently large $n$ we have

$$Pr\left(\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_{t^n}(X_{j,l})-\Theta_{t^n}\|_1\leq 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)}\ \forall t^n\ \forall j\right)$$
$$\geq 1-2^{n\vartheta} \tag{147}$$

for any positive $\vartheta$.

Now we have $J_n\cdot L_n<2^{n(\min_s\chi(p;B_s)-\mu)}$.

In [40], [51] and [17], the following was shown (using results of [41]). Let $\{\dot{X}_{j,l}\}_{j\in\{1,\ldots,J_n\},l\in\{1,\ldots,L_n\}}$ be a family of random variables taking value according to $\dot{p}\in P(\mathbf{A}^n)$. If $n$ is sufficiently large, and if $J_n\cdot L_n\leq 2^{\min_s n(\chi(\dot{p};B_s)-\mu)}$ for an arbitrary positive $\mu$ there exists a projection $q_{x^n}$ on $H$ for every $x^n\in\mathbf{A}^n$ and positive constants $\beta$ and $\gamma$ such that for any $(s,j,l)\in\theta\times\{1,\ldots,J_n\}\times\{1,\ldots,L_n\}$ it holds

$$\Pr_{\dot{p}}\left[\mathrm{tr}\left(\overline{W}_s^{\otimes n}(\dot{X}_{j,l})D_{\dot{X}_{j,l}}\right)\geq 1-|\overline{\theta}|2^{-n\beta}\right]>1-2^{-n\gamma}\,, \tag{148}$$

where for $j\in\{1,\ldots,J_n\},l\in\{1,\ldots,L_n\}$ we have

$$D_{\dot{X}_{j,l}}:=\left(\sum_{j',l'}q_{\dot{X}_{j',l'}}\right)^{-\frac{1}{2}}q_{\dot{X}_{j,l}}\left(\sum_{j',l'}q_{\dot{X}_{j',l'}}\right)^{-\frac{1}{2}}.$$

Notice that by this definition for any realization $\{\dot{x}_{j,l}:j,l\}$ of $\{\dot{X}_{j,l}:j,l\}$ it holds that $\sum_{j=1}^{J_n}\sum_{l=1}^{L_n}D_{\dot{x}_{j,l}}\leq\mathrm{id}$.

(actually in [23] it was shown that there exists a collection of positive-semidefinite operators $\{D_{s,\dot{X}_{j,l}}:s\in\overline{\theta},j\in\{1,\ldots,J_n\},l\in\{1,\ldots,L_n\}\}$ such that for any $s$, $j$, and $l$ it holds

$$Pr\left[\mathrm{tr}\left(\overline{W}_s^{\otimes n}(\dot{X}_{j,l})D_{s,\dot{X}_{j,l}}\right)\geq 1-2^{|\overline{\theta}|}2^{-n\beta}\right]>1-2^{-n\gamma}\,,$$

and for any realization $\{\dot{x}_{j,l}:j,l\}$ of $\{\dot{X}_{j,l}:j,l\}$ it holds that $\sum_{s\in\overline{\theta}}\sum_{j=1}^{J_n}\sum_{l=1}^{L_n}D_{s,\dot{x}_{j,l}}\leq$ id).

For any given $s\in\theta$ it holds

$$\overline{W}_s^{\otimes n}(p^n)-\overline{W}_s^{\otimes n}(p'^n)$$
$$=\left(1-\frac{1}{P(\mathcal{T}_{p,\delta}^n)}\right)\sum_{a^n\in\mathcal{T}_{p,\delta}^n}p^n(a^n)\overline{W}_s^{\otimes n}(a^n)+\sum_{a^n\notin\mathcal{T}_{p,\delta}^n}p^n(a^n)\overline{W}_s^{\otimes n}(a^n)\,.$$

Thus we have $\left|\mathrm{tr}\left(\overline{W}_s^{\otimes n}(p^n)-\overline{W}_s^{\otimes n}(p'^n)\right)\right|\leq 2P(\mathcal{T}_{p,\delta}^n)\leq 2^{-n\eta(\delta)}$ for a positive $\eta(\delta)$.

By Lemma 3.6 for any positive $\omega$ if $n$ is sufficiently large we have

$$\left|S\left(\overline{W}_s^{\otimes n}(p^n)\right)-S\left(\overline{W}_s^{\otimes n}(p'^n)\right)\right|$$
$$\leq 2^{-n\eta(\delta)}\log(d^n-1)+h(2^{-n\eta(\delta)})$$

$$\leq \omega \, .$$

Furthermore, we have

$$\begin{aligned}
&\Big| \sum_{a^n \in \mathcal{T}^n_{p,\delta}} {p'}^n(a^n) S\left(\overline{W}_s^{\otimes n}(a^n)\right) - \sum_{a^n \in \mathcal{T}^n_{p,\delta}} {p'}^n(a^n) S\left(\overline{W}_s^{\otimes n}(a^n)\right) \Big| \\
&= \left| \left(1 - \frac{1}{P(\mathcal{T}^n_{p,\delta})}\right) \sum_{a^n \in \mathcal{T}^n_{p,\delta}} p^n(a^n) S\left(\overline{W}_s^{\otimes n}(a^n)\right) + \sum_{a^n \notin \mathcal{T}^n_{p,\delta}} p^n(a^n) S\left(\overline{W}_s^{\otimes n}(a^n)\right) \right| \\
&\leq 2P(\mathcal{T}^n_{p,\delta}) \max_{a^n \in \mathbf{A}^n} S\left(\overline{W}_s^{\otimes n}(a^n)\right) \\
&\leq \omega \, .
\end{aligned}$$

for any positive $\omega$ if $n$ is sufficiently large.

We now have

$$\begin{aligned}
&\left|\chi(p; B_s^{\otimes n}) - \chi(p'; B_s^{\otimes n})\right| \\
&\leq \left| S\left(\overline{W}_s^{\otimes n}(p^n)\right) - S\left(\overline{W}_s^{\otimes n}({p'}^n)\right) \right| \\
&+ \Big| \sum_{a^n \in \mathcal{T}^n_{p,\delta}} {p'}^n(a^n) S\left(\overline{W}_s^{\otimes n}(a^n)\right) - \sum_{a^n \in \mathcal{T}^n_{p,\delta}} {p'}^n(a^n) S\left(\overline{W}_s^{\otimes n}(a^n)\right) \Big| \\
&\leq 2\omega \, .
\end{aligned}$$

for any positive $\omega$ if $n$ is sufficiently large.

Thus, when $J_n \cdot L_n < 2^{\min_s n\chi(p;B_s)-\mu}$ holds, we also have

$$J_n \cdot L_n < 2^{\min_s n\chi(p';B_s)-\mu} \, . \tag{149}$$

if $n$ is sufficiently large.

By (149) we can apply (148) to $X_{j,l}$. We have: If $n$ is sufficiently large, the event

$$\begin{aligned}
&\left( \bigcap_s \left\{ \max_{j\in\{1,\ldots,J_n\}} \max_{l\in\{1,\ldots,L_n\}} \mathrm{tr}\left(\overline{W}_s^{\otimes n}(X_{j,l}) D_{X_{j,l}}\right) \geq 1 - |\overline{\theta}| 2^{-n\beta} \right\} \right) \\
&\cap \left( \| \sum_{l=1}^{L_n} \frac{1}{L_n} Q_{t^n}(X_{j,l}) - \Theta_{t^n} \|_1 \leq 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)} \ \forall t^n \ \forall j \right)
\end{aligned}$$

has a positive probability with respect to $p'$.

This means that for any $\epsilon > 0$ if $n$ is sufficiently large we can find a realization $x_{j,l}$ of $X_{j,l}$ with a positive probability such that for all $s \in \overline{\theta}$, $t^n \in \theta^n$, $\pi \in \mathsf{S}_n$, and $j \in \{1, \ldots, J_n\}$, we have

$$\sum_{l=1}^{L_n} \mathrm{tr}\left(\overline{W}_s^{\otimes n}(x_{j,l}) D_{x_{j,l}}\right) \geq 1 - \epsilon \, , \tag{150}$$

and

$$\| \sum_{l=1}^{L_n} \frac{1}{L_n} Q_{t^n}(x_{j,l}) - \Theta_{t^n} \|_1 \leq 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)} \, . \tag{151}$$

We define for $\pi \in \mathsf{S}_n$ its permutation matrix on $H^{\otimes n}$ by $P_\pi$. We have $V_{t^n}(\pi(x^n)) = P_\pi V_{\pi^{-1}(t^n)}(x^n)P_\pi^\dagger$. For $\pi \in \mathsf{S}_n$, we define $\Theta_{t^n,\pi} := \sum_{x^n \in \mathcal{T}_{p,\delta}} p'(x^n) Q_{t^n}(\pi(x^n))$. We have $\Theta_{t^n,\pi} = P_\pi \left(\sum_{x^n \in \mathcal{T}_{p,\delta}} p'(x^n) Q_{\pi^{-1}(t^n)}(x^n)\right) P_\pi^\dagger = P_\pi \Theta_{\pi^{-1}(t^n)} P_\pi^\dagger$.

We choose suitable a positive $\alpha$. For any given $j' \in \{1, \ldots, J_n\}$ we have

$$\begin{aligned}
&\left\| \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(\pi(x_{j',l})) - \Theta_{t^n,\pi} \right\|_1 \\
&\leq \| \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(\pi(x_{j',l})) - \sum_{l=1}^{L_n} \frac{1}{L_n} Q_{t^n}(\pi(x_{j',l})) \|_1 \\
&\quad + \| \sum_{l=1}^{L_n} \frac{1}{L_n} Q_{t^n}(\pi(x_{j',l})) - \Theta_{t^n,\pi} \|_1 \\
&\leq \sum_{l=1}^{L_n} 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)} + \| P_\pi Q_{\pi^{-1}(t^n)}(x_{j',l}) P_\pi^\dagger - P_\pi \Theta_{\pi^{-1}(t^n)} P_\pi^\dagger \|_1 \\
&= 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)} + \| \sum_{l=1}^{L_n} \frac{1}{L_n} Q_{\pi^{-1}(t^n)}(x_{j',l}) - \Theta_{\pi^{-1}(t^n)} \|_1 \\
&\leq 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)} + \sqrt{2^{-\frac{1}{2}n\beta(\alpha)} + 2^{-\frac{1}{2}n\beta(\alpha)''}} \\
&\leq 2^{-\sqrt{n}\frac{1}{32}\hat{\beta}(\alpha)} ,
\end{aligned} \tag{152}$$

where the first inequality is an application of the triangle inequality, the second is again the triangle inequality combined with (124). The following equality follows because $\|U \cdot A \cdot U^\dagger\|_1 = \|A\|_1$ for all $A \in \mathcal{B}(H^{\otimes n})$ and unitary matrices $U \in \mathcal{B}(H^{\otimes n})$. At last, we use (151).

By (152) we have

$$\begin{aligned}
&\| \frac{1}{J_n \cdot L_n} \sum_{j=1}^{J_n} \sum_{l=1}^{L_n} V_{t^n}(\pi(x_{j,l})) - \Theta_{t^n,\pi} \|_1 \\
&\leq 2^{-\sqrt{n}\frac{1}{32}\hat{\beta}(\alpha)} .
\end{aligned}$$

By Lemma 3.6 and the inequality (152), for a uniformly distributed random variable $R_{uni}$ with values in $\{1, \ldots, J_n\}$ and all $\pi \in \mathsf{S}_n$ and $t^n \in \theta^n$ we have

$$\begin{aligned}
&\chi(R_{uni}; Z_{t^n,\pi}) \\
&= S\left( \sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(\pi(x_{j,l})) \right) \\
&- \sum_{j=1}^{J_n} \frac{1}{J_n} S\left( \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(\pi(x_{j,l})) \right) \\
&\leq \left| S\left( \sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(\pi(x_{j,l})) \right) - S\left(\Theta_{t^n,\pi}\right) \right|
\end{aligned}$$

$$+\left|S(\Theta_{t^n,\pi})-\sum_{j=1}^{J_n}\frac{1}{J_n}S\left(\sum_{l=1}^{L_n}\frac{1}{L_n}V_{t^n}(\pi(x_{j,l}))\right)\right|$$
$$\leq 2\cdot 2^{-\sqrt{n}\frac{1}{32}\hat{\beta}(\alpha)}\log(nd-1)+2h(2^{-\sqrt{n}\frac{1}{32}\hat{\beta}(\alpha)})\,. \tag{153}$$

By (153), for any positive $\lambda$ if $n$ is sufficiently large, we have

$$\max_{t^n\in\theta^n}\chi(R_{uni};Z_{t^n,\pi})\leq\lambda\,. \tag{154}$$

For an arbitrary positive $\delta$ let

$$J_n:=2^{n\min_{s\in\overline{\theta}}\chi(p;B_s)-\max_{t^n\in\theta^n}\chi(p;Z_{t^n})-n\delta}\,.$$

Now we define a code $(E,\{D_j:j=1,\ldots,J_n\})$, by $E(x^n\mid j)=\frac{1}{L_n}$ if $x^n\in\{x_{j,l}:l\in\{1,\ldots,L_n\}\}$, and $E(x^n\mid j)=0$ if $x\notin\{x_{j,l}:l\in\{1,\ldots,L_n\}\}$, and $D_j:=\frac{1}{L_n}\sum_{l=1}^{L_n}D_{x_{j,l}}$. For any positive $\lambda$ and $\epsilon$ if $n$ is sufficiently large by (150) and (154) it holds

$$\max_{s\in\overline{\theta}}\frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{a^n\in\mathbf{A}^n}E^n(a^n|j)\mathrm{tr}\left(\overline{W}_s^{\otimes n}(a^n)D_j\right)\geq 1-\epsilon\,,$$

$$\max_{t^n\in\theta^n}\max_{\pi\in\mathsf{S}_n}\chi\left(R_{uni};Z_{t^n,\pi}\right)\leq\epsilon\,.$$

We obtain

$$\hat{C}_s(\{(\overline{W}_s,V_t):s\in\overline{\theta},t\in\theta\})\geq\min_{s\in\overline{\theta}}\chi(p;B_s)-\lim_{n\to\infty}\frac{1}{n}\max_{t^n\in\theta^n}\chi(p;Z_{t^n})\,. \tag{155}$$

The achievability of $\lim_{n\to\infty}\frac{1}{n}\Big(\min_{s\in\overline{\theta}}\chi(p_U;B_s^{\otimes n})-\max_{t^n\in\theta^n}\chi(p_U;Z_{t^n})\Big)$ is then shown via standard arguments (cf. [34]).

Now we are going to prove the converse.

Let $(\mathcal{C}_n)=(E^{(n)},\{D_j^{(n)}:j\})$ be a sequence of $(n,J_n)$ code such that

$$\max_{s\in\overline{\theta}}P_e(\mathcal{C}_n,s,n)\leq\lambda_n\,,$$

$$\max_{t^n\in\theta^n}\max_{\pi\in\mathsf{S}_n}\chi\left(R_{uni};Z_{t^n,\pi}\right)\leq\epsilon_n\,,$$

where $\lim_{n\to\infty}\lambda_n=0$ and $\lim_{n\to\infty}\epsilon_n=0$, where $R_{uni}$ is the uniform distribution on $\{1,\cdots J_n\}$.

Since the capacity of a compound classical-quantum channel $(\overline{W}_s)_{s\in\overline{\theta}}$ cannot exceed the worst channel in $\{\overline{W}_s:s\in\overline{\theta}\}$, its capacity is bounded by $\frac{1}{n}\min_{s\in\overline{\theta}}\chi(p_U;B_s)$ (cf. [67]). For any $\xi>0$ we choose $\epsilon_n=\frac{1}{2}\xi$. The enhanced achievable secrecy rate for the compound-arbitrarily varying wiretap classical-quantum channel cannot exceed the capacity without wiretapper, thus , if $n$ is sufficiently large, the secrecy rate of $(\mathcal{C}_n)$ cannot be greater than

$$\min_{s\in\overline{\theta}}\chi(R_{uni};B_s)-\frac{1}{n}\max_{t^n\in\theta^n}\chi(R_{uni};Z_{t^n})-\xi+\frac{1}{n}\epsilon_n$$

$$\leq \frac{1}{n} \max_{U \to A^n \to \{B_s^{\otimes n}, Z_{t^n} : s, t_n\}} \left( \min_{s \in \overline{\theta}} \chi(p_U; B_s^n) - \max_{t^n \in \theta^n} \chi(p_U; Z_{t^n}) \right) - \frac{1}{2}\xi \,. \quad (156)$$

□

The inequality holds because $R_{uni} \to A^n \to \{B_s^{\otimes n}, Z_{t^n} : s, t_n\}$ is always a Markov chain.

This and (155) prove Theorem 4.2.

**Corollary 4.4.** *Let $\theta := \{1, \cdots, T\}$ be a finite index set. Let $\overline{\theta}$ be an infinite index set. Let $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$ be a compound-arbitrarily varying wiretap classical-quantum channel. We have*

$$\hat{C}_s(\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}) = \lim_{n \to \infty} \frac{1}{n} \max_{U \to A^n \to \{B_s^{\otimes n}, Z_{t^n} : s, t_n\}} \left( \inf_{s \in \overline{\theta}} \chi(p_U; B_s^{\otimes n}) - \max_{t^n \in \theta^n} \chi(p_U; Z_{t^n}) \right).$$

*Proof.* We now consider a $\overline{\theta}$ such that $|\overline{\theta}|$ is not finite. We assume there is a series of positive constants $\{\tau_n : n \in \mathbb{N}\}$ such that $(\frac{3}{\tau_n})^{2d'^4} \geq 2^{-n\beta/2}$ and $\lim_{n\to\infty} n\tau_n = 0$. By Lemma 3.11 there exists a finite set $\overline{\theta_{\tau_n}}'$ with $|\overline{\theta_{\tau_n}}'| \leq (\frac{3}{\tau_n})^{2d'^4}$ and $\tau_n$-nets $\left(\overline{W}_{s'}\right)_{s' \in \overline{\theta_{\tau_n}}'}$, $(V_{s'})_{s' \in \overline{\theta_{\tau_n}}'}$ such that for every $t \in \overline{\theta}$ we can find a $s' \in \overline{\theta_{\tau_n}}'$ with $\left\|\overline{W}_s - \overline{W}_{s'}\right\|_\diamond \leq \tau_n$.

We assume that the sender's encoding is restricted to transmitting an indexed finite set of quantum states $\{\rho_x : x \in \mathbf{A}\} \subset \mathcal{S}(H'^{\otimes n})$.

By Theorem 4.2 the legal transmitters build now a code $C_2 = \{E, \{D_j : j\}\}$. such that for all $s'' \in \overline{\theta_{\tau_n}}'$, $t \in \theta$, and $\pi \in \mathsf{S}_n$ it holds that

$$\frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{x^n \in \mathbf{A}^n} E(x^n \mid j) \mathrm{tr}\left(\overline{W}_{s''}^{\otimes n}(\rho_{x^n}) D_j^n\right) \geq 1 - (\frac{3}{\tau_n})^{2d'^4} 2^{-n\beta} \geq 1 - 2^{-n\beta/2}\,, \quad (157)$$

$$\chi(R_{uni}; Z_{t,\pi}^n) \leq 2^{-n\upsilon}\,. \quad (158)$$

Let $|\psi_{x^n}\rangle\langle\psi_{x^n}| \in \mathcal{S}(H'^{\otimes n} \otimes H'^{\otimes n})$ be an arbitrary purification of the quantum state $\rho_{x^n}$, then $\mathrm{tr}\left[\left(\overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n}\right)(\rho_{x^n})\right] = \mathrm{tr}\left(\mathrm{tr}_{H'^{\otimes n}}\left[\mathrm{id}_{H'}^{\otimes n} \otimes (\overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n})(|\psi_{x^n}\rangle\langle\psi_{x^n}|)\right]\right)$. We have

$$\begin{aligned}
&\mathrm{tr}\left|\sum_{x^n \in \mathbf{A}^n} E(x^n \mid j)\left(\overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n}\right)(\rho_{x^n})\right| \\
&= \mathrm{tr}\left(\sum_{x^n \in \mathbf{A}^n} E(x^n \mid j)\mathrm{tr}_{H'^{\otimes n}}\left|\mathrm{id}_{H'}^{\otimes N} \otimes (\overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n})(|\psi_{x^n}\rangle\langle\psi_{x^n}|)\right|\right) \\
&= \mathrm{tr}\left|\sum_{x^n \in \mathbf{A}^n} E(x^n \mid j)\mathrm{id}_{H'}^{\otimes n} \otimes (\overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n})(|\psi_{x^n}\rangle\langle\psi_{x^n}|)\right| \\
&= \sum_{x^n \in \mathbf{A}^n} E(x^n \mid j)\left\|\mathrm{id}_{H'}^{\otimes n} \otimes (\overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n})(|\psi_{x^n}\rangle\langle\psi_{x^n}|)\right\|_1 \\
&\leq \sum_{x^n \in \mathbf{A}^n} E(x^n \mid j)\|\overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n}\|_\diamond \cdot \|(|\psi_{x^n}\rangle\langle\psi_{x^n}|)\|_1
\end{aligned}$$

$$\leq n\tau_n \ .$$

The first inequality follows by the definition of $\| \cdot \|_\diamond$. The second inequality follows from the facts that $\| \left( |\psi_{x^n}\rangle\langle\psi_{x^n}| \right) \|_1 = 1$ and $\left\| \overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n} \right\|_\diamond = \left\| \sum_{k=1}^n \overline{W}_s^{\otimes k-1} \overline{W}_{s'}^{\otimes n-k} \left( \overline{W}_s - \overline{W}_{s'} \right) \right\|_\diamond$ $= n \cdot \left\| \overline{W}_s - \overline{W}_{s'} \right\|_\diamond$, since $\| \cdot \|_\diamond$ is multiplicative and $\left\| \overline{W}_s \right\|_\diamond = \left\| \overline{W}_{s'} \right\|_\diamond = 1$.

It follows that

$$\begin{aligned}
&\left| \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{x^n \in \mathbf{A}^n} E(x^n \mid j) \mathrm{tr} \left( \overline{W}_s^{\otimes n} \left( \rho_{x^n} \right) D_j^n \right) \right. \\
&\left. - \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{x^n \in \mathbf{A}^n} E(x^n \mid j) \mathrm{tr} \left( \overline{W}_{s'}^{\otimes n} \left( \rho_{x^n} \right) D_j^n \right) \right| \\
&\leq \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{x^n \in \mathbf{A}^n} E(x^n \mid j) \left| \mathrm{tr} \left[ \left( \overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n} \right) \left( \rho_{x^n} \right) D_j^n \right] \right| \\
&\leq \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{x^n \in \mathbf{A}^n} E(x^n \mid j) \mathrm{tr} \left[ \left( \overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n} \right) \left( \rho_{x^n} \right) D_j^n \right] \\
&\leq \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{x^n \in \mathbf{A}^n} E(x^n \mid j) \mathrm{tr} \left[ \left( \overline{W}_s^{\otimes n} - \overline{W}_{s'}^{\otimes n} \right) \left( \rho_{x^n} \right) \right] \\
&\leq \frac{1}{J_n} J_n n\tau_n \\
&= n\tau_n \ .
\end{aligned} \tag{159}$$

By (159) we have

$$\sup_{s \in \overline{\theta}} \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{x^n \in \mathbf{A}^n} E(x^n \mid j) \mathrm{tr} \left( \overline{W}_s^{\otimes n} \left( \rho_{x^n} \right) D_j^n \right) \geq 1 - \lambda_{\tau_n} - n\tau_n \ .$$

Thus,

$$\hat{C}_s(\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}) \geq \lim_{n \to \infty} \frac{1}{n} (\inf_{s \in \overline{\theta}} \chi(p; B_s^{\otimes n}) - \max_{t^n \in \theta^n} \chi(p; Z_{t^n})) \ . \tag{160}$$

The achievability of $\lim_{n \to \infty} \frac{1}{n} \left( \min_{s \in \overline{\theta}} \chi(p_U; B_s) - \max_{t^n \in \theta^n} \chi(p_U; Z_{t^n}) \right)$ is then shown via standard arguments.

The proof of the converse is similar to those given in the proof of Theorem 4.2. □

**Corollary 4.5.** *Let $\overline{\theta}$ and $\theta$ be finite index sets. Let $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$ be a compound-arbitrarily varying wiretap classical-quantum channel. The secrecy capacity of $\{(\overline{W}_s, V_t) : s \in \overline{\theta}, t \in \theta\}$ is equal to*

$$\lim_{n \to \infty} \frac{1}{n} \max_{U \to A^n \to \{B_s^{\otimes n}, Z_{t^n : s, t_n}\}} \left( \min_{s \in \overline{\theta}} \chi(p_U; B_s^{\otimes n}) - \max_{t^n \in \theta^n} \chi(p_U; Z_{t^n}) \right) .$$

*Proof.* The corollary follows immediately from the fact that the enhanced secrecy capacity of a compound-arbitrarily varying wiretap classical-quantum channel is less or equal to its secrecy capacity. □

### 4.2.2 Secrecy Capacity Formula of Arbitrarily Varying Classical-Quantum Wiretap Channel under Common Randomness Assisted Quantum Coding

In this section we use the results of Section 4.2.1 to determine the formula for the secrecy capacities under common randomness assisted coding of arbitrarily varying classical-quantum wiretap channels.

**Theorem 4.6.** *Let $\theta := \{1, \cdots, T\}$ be a finite index set. Let $(W_t, V_t)_{t\in\theta}$ be an arbitrarily varying classical-quantum wiretap channel. We have*

$$C_s(\{(W_t, V_t) : t \in \theta\}; cr)$$
$$= \lim_{n\to\infty} \frac{1}{n} \max_{U\to A^n\to\{B_q^{\otimes n}, Z_{t^n}:q,t_n\}} \Big( \inf_{B_q\in Conv((B_t)_{t\in\theta})} \chi(p_U; B_q^{\otimes n}) - \max_{t^n\in\theta^n} \chi(p_U; Z_{t^n}) \Big) . \tag{161}$$

*Here $Conv((B_t)_{t\in\theta})$ is the convex hull of $\{B_t : t \in \theta\}$.*

*Proof. i) Achievement*

Our idea is similar to the results for classical arbitrarily varying wiretap channel in [65]: Applying Ahlswede's robustification technique (cf. [16]), we use the results of Section 4.2.1 to show the existence of a common randomness assisted quantum code. Additionally, we have to consider the security.

We denote the set of distribution function on $\theta$ by $P(\theta)$. For every $q \in P(\theta)$ we define a classical-quantum channel $\overline{W}_q := \sum_{s\in\theta} q(s)W_s$. We now define a compound-arbitrarily varying wiretap classical-quantum channel by

$$\{(\overline{W}_q, V_t); q \in P(\theta), t \in \theta\} .$$

We fix a probability distribution $p \in \mathbf{A}$. We choose arbitrarily $\epsilon > 0$, $\delta > 0$, and $\zeta > 0$. Let

$$J_n = \lfloor 2^{\inf_{B_q\in Conv((B_s)_{s\in\theta})} \chi(p;B_q^{\otimes n}) - \max_{t^n\in\theta^n} \chi(p;Z_{t^n}) - n\delta} \rfloor .$$

By Corollary 4.4 if $n$ is sufficiently large there exists an $(n, J_n)$ code $C = \big(E^n, \{D_j^n : j = 1, \cdots J_n\}\big)$ such that

$$\max_{q\in P(\theta)} 1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \mathrm{tr}(\overline{W}_q(E^n(\ |j))D_j^n) < \epsilon ,$$

$$\max_{t^n\in\theta^n} \max_{\pi\in\mathsf{S}_n} \chi\left(R_{uni}; Z_{t^n,\pi}\right) < \zeta .$$

Similar to the proofs in [16] we now apply Ahlswede's robustification technique.

**Lemma 4.7 (cf. [3], [4], and [5]).** *Let $S$ be a finite set and $n \in \mathbb{N}$. If a function $f : S^n \to [0, 1]$ satisfies*

$$\sum_{s^n\in S^n} f(s^n)q(s_1)q(s_2)\cdots q(s_n) \geq 1 - \epsilon ,$$

*for all $q \in P(\theta)$ and a positive $\epsilon \in [0, 1]$, then*

$$\frac{1}{n!} \sum_{\pi\in\mathsf{S}_n} f(\pi(s^n)) \geq 1 - 3(n + 1)^{|S|}\epsilon . \tag{162}$$

We define a function $f : \theta^n \to [0,1]$ by

$$f(t^n) := \frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}(W_{t^n}(E^n(\ |j))D_j^n)\ .$$

For every $q \in P(\theta)$ we have

$$\begin{aligned}
&\sum_{t^n\in\theta^n} f(t^n)q(t_1)\cdots q(t_n)\\
&= \sum_{t^n\in\theta^n}\frac{1}{J_n}\sum_{j=1}^{J_n}\mathrm{tr}(W_{t^n}(E^n(\ |j))D_j^n)q(t_1)\cdots q(t_n)\\
&= \frac{1}{J_n}\sum_{j=1}^{J_n}\mathrm{tr}\left(\sum_{t^n\in\theta^n} q(t_1)\cdots q(t_n)W_{t^n}(E^n(\ |j))D_j^n\right)\\
&= \frac{1}{J_n}\sum_{j=1}^{J_n}\mathrm{tr}(\overline{W}_q(E^n(\ |j))D_j^n)\\
&> 1-2^{-n\beta/2}\ .
\end{aligned}$$

Applying Lemma 4.7 we have

$$\begin{aligned}
&1-3(n+1)^{|\theta|}2^{-n\beta/2}\\
&\le \frac{1}{n!}\sum_{\pi\in\mathsf{S}_n} f(\pi(t^n))\\
&= \frac{1}{n!}\sum_{\pi\in\mathsf{S}_n}\frac{1}{J_n}\sum_{j=1}^{J_n}\mathrm{tr}(W_{\pi(t^n)}(E^n(\ |j))D_j^n)\\
&= \frac{1}{n!}\sum_{\pi\in\mathsf{S}_n}\frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{a^n\in\mathbf{A}^n} E^n(a^n|j)\mathrm{tr}(W_{\pi(t^n)}(a^n)D_j^n)\\
&= \frac{1}{n!}\sum_{\pi\in\mathsf{S}_n}\frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{a^n\in\mathbf{A}^n} E^n(a^n|j)\mathrm{tr}(W_{t^n}(\pi^{-1}(a^n))P_\pi^\dagger D_j^n P_\pi)\ , \qquad (163)
\end{aligned}$$

where for $\pi \in \mathsf{S}_n$, $P_\pi$ is its permutation matrix on $H^{\otimes n}$.

We now define our common randomness assisted quantum code by

$$\left\{\left(\pi\circ E^n, \{P_\pi D_j^n P_\pi^\dagger, j\in\{1,\cdots,J_n\}\}\right) : \pi\in\mathsf{S}_n\right\}\ .$$

$P_\pi D_j^n P_\pi^\dagger$ is Hermitian and positive-semidefinite. Furthermore, it holds $\sum_{j=1}^{J_n} P_\pi D_j^n P_\pi^\dagger$ $= \sum_{j=1}^{J_n} P_\pi \mathrm{id}_{H^{\otimes n}} P_\pi^\dagger = \mathrm{id}_{H^{\otimes n}}$.

By (163) and by the fact that

$$\begin{aligned}
&\frac{1}{n!}\sum_{\pi\in\mathsf{S}_n}\max_{t^n\in\theta^n}\chi\left(R_{uni};Z_{t^n,\pi}\right)\\
&\le \max_{t^n\in\theta^n}\max_{\pi\in\mathsf{S}_n}\chi\left(R_{uni};Z_{t^n,\pi}\right)
\end{aligned}$$

$$< \zeta \, ,$$

for any positive $\varepsilon$ when $n$ is sufficiently large it holds that:

$$C_s(\{(W_t, V_t) : t \in \theta\}; cr) \geq \inf_{B_q \in Conv((B_s)_{s\in\theta})} \chi(p; B_q) - \lim_{n\to\infty} \frac{1}{n} \max_{t^n \in \theta^n} \chi(p; Z_{t^n}) - \varepsilon \, . \tag{164}$$

The achievability of $\lim_{n\to\infty} \frac{1}{n}\Big(\min_{B_q \in Conv((B_s)_{s\in\theta})} \chi(p_U; B_q^{\otimes n}) - \max_{t^n\in\theta^n} \chi(p_U; Z_{t^n})\Big)$ is then shown via standard arguments (cf. [34]).

*ii) Converse*

Now we are going to prove the converse. Similar to the results for classical arbitrarily varying wiretap channel in [65] we limit the amount of common randomness.

Let $(\{\mathcal{C}_n^\gamma : \gamma \in \Gamma\})$ be a sequence of $(n, J_n)$ common randomness assisted codes such that

$$\max_{s\in\theta} \frac{1}{|\Gamma|} \sum_{\gamma=1}^{|\Gamma|} P_e(\mathcal{C}_n^\gamma, t^n) \leq \lambda_n \, , \tag{165}$$

$$\max_{t^n\in\theta^n} \frac{1}{|\Gamma|} \sum_{\gamma=1}^{|\Gamma|} \chi\left(R_{uni}; Z_{\mathcal{C}^\gamma, t^n}\right) \leq \epsilon_n \, , \tag{166}$$

where $\lim_{n\to\infty} \lambda_n = 0$ and $\lim_{n\to\infty} \epsilon_n = 0$.

We consider a $|\Gamma|$-long sequence of outputs $(1, \cdots, |\Gamma|)$ has been given by the common randomness and a $n\,|\Gamma|$-long block has been sent. The legal receiver obtains the quantum states $\{B_q^\gamma : \gamma \in \Gamma\}$. By (165) he is able to decode $2^{n|\Gamma| \log J_n}$ messages. By [16], for every $B_q \in Conv((B_s)_{s\in\theta})$ we have

$$\log J_n \leq \frac{1}{|\Gamma|} \frac{1}{n} \sum_{\gamma=1}^{|\Gamma|} \chi(R_{uni}; B_q^{\gamma\otimes n}) \, ,$$

and by (166), we have for and every $t^n \in \theta^n$

$$\frac{1}{n} \log J_n \leq \frac{1}{|\Gamma|} \frac{1}{n} \sum_{\gamma=1}^{|\Gamma|} (\chi(R_{uni}; B_q^{\gamma\otimes n}) - \chi(R_{uni}; Z_{t^n}^\gamma)) + \epsilon_n \, .$$

**Lemma 4.8 (cf. [26]).** *Let $c > 0$. For every $q \in P(\theta)$ and $s^n \in \theta^n$, let a function $I_{q,s^n} : \Gamma \to [0, c]$ be given. Assume these functions satisfy the following: for every $\gamma \in \Gamma$ and $s^n \in \theta^n$*

$$|I_{q,s^n}(\gamma) - I_{q',s^n}(\gamma)| \leq f(\delta) \, ,$$

*if $q, q' \in P(\theta)$ satisfy $\|q - q'\|_1 \leq \delta$, for some $f(\delta)$ which tends to $0$ as $\delta$ tends to $0$. We write $\mu(I_{q,s^n}) := \sum_{\gamma\in\Gamma} \mu(\gamma) I_{q,s^n}(\gamma)$, where $\mu(\gamma)$ is the probability of $\gamma$. Then for every $\varepsilon > 0$ and sufficiently large $n$, there are $L = n^2$ realizations $\gamma_1, \cdots, \gamma_L$ such that*

$$\frac{1}{L} \sum_{l=1}^{L} I_{q,s^n}(\gamma_l) \geq (1 - \varepsilon)\mu(I_{q,s^n}) - \varepsilon$$

*for every $q \in P(\theta)$ and $s^n \in \theta^n$.*

For $q \in Conv(\{s : s \in \theta\})$ we define

$$I_{q,s^n}(\gamma) := \frac{1}{n}\left(\chi(R_{uni}; B_q^{\gamma\otimes n}) - \chi(R_{uni}; Z_{t^n}^{\gamma})\right) .$$

In [29] the continuity of $q \to \frac{1}{n}\chi(R_{uni}; B_q^{\gamma\otimes n})$ has been shown; thus there is a $f(\delta)$ such that $|I_{q,s^n}(\gamma) - I_{q',s^n}(\gamma)| \frac{1}{n}\frac{1}{|\Gamma|}\sum_{\gamma=1}^{|\Gamma|}(\chi(R_{uni}; B_q^{\gamma\otimes n}) - \frac{1}{n}\frac{1}{|\Gamma|}\sum_{\gamma=1}^{|\Gamma|}(\chi(R_{uni}; B_{q'}^{\gamma\otimes n})$ $\leq f(\delta)$ for a $f(\delta)$ that fulfills $f(\delta) \to 0$ when $\|q - q'\|_1 = \delta \to 0$. By Lemma 4.8 there is a set $\Gamma' \subset \Gamma$ such that $|\Gamma'| = n^2$ and

$$\begin{aligned}
&\frac{1}{|\Gamma'|}\frac{1}{n}\sum_{\gamma'\in\Gamma'}\left(\chi(R_{uni}; B_q^{\gamma'\otimes n}) - \chi(R_{uni}; Z_{t^n}^{\gamma'})\right) \\
&\geq (1-\varepsilon)\frac{1}{n}\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\left(\chi(R_{uni}; B_q^{\gamma\otimes n}) - \chi(R_{uni}; Z_{t^n}^{\gamma})\right) ,
\end{aligned}$$

where $B_q^{\gamma'}$ and $Z_{t^n}^{\gamma}$ are the quantum states at the output of legal receiver channel and the wiretapper's channel, respectively, when the output of the common randomness is $\gamma'$.

Thus

$$\frac{1}{n}\log J_n \leq \frac{1}{1-\varepsilon}\frac{1}{n}\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\left(\chi(R_{uni}; B_q^{\gamma\otimes n}) - \chi(R_{uni}; Z_{t^n}^{\gamma}) + \epsilon_n\right) . \qquad (167)$$

To prove the converse we now consider

$$\begin{aligned}
&\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\frac{1}{n}\left(\chi(R_{uni}; B_q^{\gamma\otimes n}) - \chi(R_{uni}; Z_{t^n}^{\gamma})\right) - \frac{1}{n}\left(\chi(R_{uni}; B_q^{\otimes n}) - \chi(R_{uni}; Z_{t^n})\right) \\
&= \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\frac{1}{n}\left(\chi(R_{uni}; B_q^{\gamma\otimes n}) - \chi(R_{uni}; Z_{t^n}^{\gamma})\right) \\
&- \frac{1}{n}\left(\chi(R_{uni}; \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'} B_q^{\gamma\otimes n}) - \chi(R_{uni}; \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'} Z_{t^n}^{\gamma})\right) \\
&= \frac{1}{n}\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\left(\chi(R_{uni}; B_q^{\gamma\otimes n}) - \chi(R_{uni}; \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'} B_q^{\gamma\otimes n})\right) \\
&- \frac{1}{n}\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\left(\chi(R_{uni}; Z_{t^n}^{\gamma}) + \frac{1}{n}\chi(R_{uni}; \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'} Z_{t^n}^{\gamma})\right) .
\end{aligned}$$

Let $\psi_{t^n}^{j,\gamma}$ be the quantum state at the output of legal receiver channel when the channel state is $t^n$, the output of the common randomness is $\gamma$, and $j$ has been sent. We denote $\tilde{B}_{q^n}^{j} := \{\psi_{q^n}^{j,\gamma} : \gamma \in \Gamma'\}$ and $\tilde{B}_{q^n} := \{\frac{1}{J_n}\psi_{q^n}^{j,\gamma} : \gamma \in \Gamma'\}$. Let $G_{uni}$ be the uniformly distributed random variable with value in $\Gamma'$.

We have

$$\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\chi\left(R_{uni}; B_q^{\gamma}\right) - \chi\left(R_{uni}; \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'} B_q^{\gamma}\right)$$

$$
\begin{aligned}
&= \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'} S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\psi_{q^n}^{j,\gamma}\right) - \frac{1}{|\Gamma'|}\frac{1}{J_n}\sum_{\gamma\in\Gamma'}\sum_{j=1}^{J_n} S\left(\psi_{q^n}^{j,\gamma}\right) \\
&- \left[S\left(\frac{1}{|\Gamma'|}\frac{1}{J_n}\sum_{\gamma\in\Gamma'}\sum_{j=1}^{J_n}\psi_{q^n}^{j,\gamma}\right) - \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\psi_{q^n}^{j,\gamma}\right)\right] \\
&= \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'} S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\psi_{q^n}^{j,\gamma}\right) - S\left(\frac{1}{|\Gamma'|}\frac{1}{J_n}\sum_{\gamma\in\Gamma'}\sum_{j=1}^{J_n}\psi_{q^n}^{j,\gamma}\right) \\
&- \left[\frac{1}{|\Gamma'|}\frac{1}{J_n}\sum_{\gamma\in\Gamma'}\sum_{j=1}^{J_n} S\left(\psi_{q^n}^{j,\gamma}\right) - \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\psi_{q^n}^{j,\gamma}\right)\right] \\
&= \frac{1}{J_n}\sum_{j=1}^{J_n}\chi\left(G_{uni}, \tilde{B}_{q^n}^{j}\right) - \chi\left(G_{uni}, \tilde{B}_{q^n}\right) \\
&\leq \frac{1}{J_n}\sum_{j=1}^{J_n}\chi\left(G_{uni}, \tilde{B}_{q^n}^{j}\right) \\
&\leq \frac{1}{J_n}\sum_{j=1}^{J_n} H\left(G_{uni}\right) \\
&= H\left(G_{uni}\right) \\
&= 2\log n\,. \qquad (168)
\end{aligned}
$$

Let $\phi_{t^n}^{j,\gamma}$ be the quantum state at the output of the wiretapper's channel when the channel state is $t^n$, the output of the common randomness is $\gamma$, and $j$ has been sent.

We have

$$
\begin{aligned}
&\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\chi\left(R_{uni}; Z_{t^n}^{\gamma}\right) - \chi\left(R_{uni}; \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'} Z_{t^n}^{\gamma}\right) \\
&= \frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'} S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\phi_{t^n}^{j,\gamma}\right) - \frac{1}{|\Gamma'|}\frac{1}{J_n}\sum_{\gamma\in\Gamma'}\sum_{j=1}^{J_n} S\left(\phi_{t^n}^{j,\gamma}\right) \\
&- S\left(\frac{1}{|\Gamma'|}\frac{1}{J_n}\sum_{\gamma\in\Gamma'}\sum_{j=1}^{J_n}\phi_{t^n}^{j,\gamma}\right) + \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\phi_{t^n}^{j,\gamma}\right)\,. \qquad (169)
\end{aligned}
$$

Let $H^{\mathfrak{G}}$ be a $|\Gamma'|$-dimensional Hilbert space, spanned by an orthonormal basis $\{|i\rangle : i = 1, \cdots, |\Gamma'|\}$. Let $H^{\mathfrak{J}}$ be a $J_n$-dimensional Hilbert space, spanned by an orthonormal basis $\{|j\rangle : j = 1, \cdots, J_n\}$. Similar to (118) we define

$$
\varphi^{\mathfrak{J}\mathfrak{G}H^n} := \frac{1}{J_n}\frac{1}{|\Gamma'|}\sum_{j=1}^{J_n}\sum_{\gamma\in\Gamma'} |j\rangle\langle j| \otimes |i\rangle\langle i| \otimes \phi_{t^n}^{j,\gamma}\,,
$$

By strong subadditivity of von Neumann entropy it holds $S(\varphi^{\mathfrak{J}H^n}) + S(\varphi^{\mathfrak{G}H^n}) \geq$

$S(\varphi^{H^n}) + S(\varphi^{\mathfrak{J}\mathfrak{G}H^n})$, therefore

$$\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\chi\left(R_{uni};Z_{t^n}^{\gamma}\right)-\chi\left(R_{uni};\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}Z_{t^n}^{\gamma}\right)\geq 0\,. \tag{170}$$

By (168) and (170) we have

$$\chi(R_{uni};B_q)-\frac{1}{n}\chi(R_{uni};Z_{t^n})+2\log n\geq\frac{1}{|\Gamma'|}\sum_{\gamma\in\Gamma'}\frac{1}{n}\left(\chi(R_{uni};B_q^{\gamma\otimes n})-\chi(R_{uni};Z_{t^n}^{\gamma})\right)\,.$$

Thus for every $B_q\in Conv((B_s)_{s\in\theta})$ and every $t^n\in\theta^n$ we have

$$\frac{1}{n}\log J_n\leq\frac{1}{1-\varepsilon}\frac{1}{n}\left(\chi(R_{uni};B_q^{\otimes n})-\chi(R_{uni};Z_{t^n})+\epsilon_n+2\frac{1}{n}\log n\right)\,. \tag{171}$$

□

Similar to the proof of Theorem 4.2 we have $\frac{1}{n}\Big(\inf_{B_q\in Conv((B_t)_{t\in\theta})}\chi(R_{uni};B_q^{\otimes n})$ $-\max_{t^n\in\theta^n}\chi(R_{uni};Z_{t^n})\Big)\leq\frac{1}{n}\max_{U\to A^n\to\{B_q^{\otimes n},Z_{t^n}:q,t_n\}}\Big(\inf_{B_q\in Conv((B_t)_{t\in\theta})}\chi(p_U;B_q^{\otimes n})$ $-\max_{t^n\in\theta^n}\chi(p_U;Z_{t^n})\Big)$. The converse has been shown. (164) and (171) prove Theorem 4.6.

**Corollary 4.9.** *Let $\{(W_t,V_t):t\in\theta\}$ be an arbitrarily varying classical-quantum wiretap channel.*

*1) Let $\mathbf{X}$ and $\mathbf{Y}$ be finite sets. If $I(X,Y)>0$ holds for a random variable $(X,Y)$ which is distributed to a joint probability distribution $p\in P(\mathbf{X},\mathbf{Y})$, then the $(X,Y)$ correlation assisted secrecy capacity of $\{(W_t,V_t):t\in\theta\}$ is equal to*

$$\lim_{n\to\infty}\frac{1}{n}\max_{U\to A^n\to\{B_q^{\otimes n},Z_{t^n}:q,t_n\}}\Big(\inf_{B_q\in Conv((B_t)_{t\in\theta})}\chi(p_U;B_q^{\otimes n})-\max_{t^n\in\theta^n}\chi(p_U;Z_{t^n})\Big)\,.$$

*2)If the arbitrarily varying classical-quantum channel $\{W_t:t\in\theta\}$ is not symmetrizable, then the deterministic secrecy capacity of $\{(W_t,V_t):t\in\theta\}$ under weak code concept is equal to*

$$\lim_{n\to\infty}\frac{1}{n}\max_{U\to^n A\to\{B_q^{\otimes n},Z_{t^n}:q,t_n\}}\Big(\inf_{B_q\in Conv((B_t)_{t\in\theta})}\chi(p_U;B_q^{\otimes n})-\max_{t^n\in\theta^n}\chi(p_U;Z_{t^n})\Big)\,.$$

*Proof.* 1) follows immediately from Theorem 4.6 and Theorem 4.14.

To show 2) we use a technique similar to the proof of Theorem 4.1: We build a two-part code word which consists of a non-secure code word and a common randomness assisted secure code word. The first part is used to create the common randomness for the sender and the legal receiver. The second part is a common randomness assisted secure code word transmitting the message to the legal receiver.

We consider the Markov chain $U\to A^n\to\{B_q^{\otimes n},Z_{t^n}:q,t_n\}$, where we define the classical channel $U\to A$ by $T_U$. Let

$$J_n=\lfloor 2^{n\inf_{B_q\in Conv((B_s)_{s\in\theta})}\chi(p_U;B_q)-\max_{t^n\in\theta^n}\chi(p_U;Z_{t^n})-n\delta}\rfloor\,.$$

By Theorem 4.6 for any positive $\epsilon$ if $n$ is sufficiently large there is an $(n, J_n)$ code $\left(E^n, \{D_j^n : j = 1, \cdots J_n\}\right)$ for the arbitrarily varying classical-quantum wiretap channel $\{(W_t \circ T_U, V_t \circ T_U) : t \in \theta\}$ such that

$$\frac{1}{n!}\sum_{\pi\in\mathsf{S}_n}\frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{a^n\in\mathbf{A}^n} E^n(a^n|j)\mathrm{tr}(W_{t^n}(\pi^{-1}(a^n))P_\pi^\dagger D_j^n P_\pi) \geq 1-\epsilon$$

and

$$\frac{1}{n!}\sum_{\pi\in\mathsf{S}_n}\max_{t^n\in\theta^n}\chi\left(R_{uni}; Z_{t^n,\pi}\right) \leq \epsilon\,.$$

By Theorem 4.1 for any positive $\lambda$ if $n$ is sufficiently large there is an $(n, J_n)$ common randomness assisted code $\{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_{n^3}\}$ for the arbitrarily varying classical-quantum wiretap channel $\{(W_t \circ T_U, V_t \circ T_U) : t \in \theta\}$ such that

$$\max_{t^n\in\theta^n}\frac{1}{n^3}\sum_{i=1}^{n^3} P_e(\mathcal{C}_i, t^n) < \lambda\,,$$

and

$$\max_{t^n\in\theta^n}\frac{1}{n^3}\sum_{i=1}^{n^3}\chi\left(R_{uni}, Z_{\mathcal{C}_i,t^n}\right) < \lambda\,.$$

Similar to the proof of Theorem 4.1, for any positive $\vartheta$ if $\{W_t : t \in \theta\}$ is not symmetrizable and $n$ is sufficiently large there is a code $\left(\left(c_i^{\mu(n)}\right)_{i\in\{1,\cdots,n^3\}}, \{D_i^{\mu(n)} : i \in \{1, \cdots, n^3\}\}\right)$ with deterministic encoder of length $\mu(n)$, where $2^{\mu(n)} = o(n)$ for the arbitrarily varying classical-quantum wiretap channel $\{(W_t, V_t) : t \in \theta\}$ such that

$$1 - \frac{1}{n^3}\sum_{i=1}^{n^3}\mathrm{tr}(W_{t^n}(c_i^{\mu(n)})D_i^{\mu(n)}) \leq \vartheta\,.$$

We now can construct a code $\mathcal{C}^{det} = \left(E^{\mu(n)+n}, \left\{D_j^{\mu(n)+n} : j = 1, \cdots, J_n\right\}\right)$, where for $a^{\mu(n)+n} = (a^{\mu(n)}, a^n) \in \mathbf{A}^{\mu(n)+n}$

$$E^{\mu(n)+n}(a^{\mu(n)+n}|j) = \begin{cases} \frac{1}{n^3}E_i^n(a^n|j) \text{ if } a^{\mu(n)} = c_i^{\mu(n)} \\ 0 \text{ else} \end{cases},$$

and

$$D_j^{\mu(n)+n} := \sum_{i=1}^{n^3} D_i^{\mu(n)} \otimes D_{i,j}^n\,.$$

Similar to the proof of Theorem 4.1 for any positive $\lambda$ if $n$ is sufficiently large we have

$$\max_{t^{\mu(n)+n}\in\theta^{\mu(n)+n}} P_e(\mathcal{C}^{det}, t^{\mu(n)+n}) < \lambda\,,$$

$$\max_{t^{\mu(n)+n}\in\theta^{\mu(n)+n}}\chi\left(R_{uni}, Z_{\mathcal{C}^{det},t^{\mu(n)+n}}\right) < \lambda\,.$$

□

**Remark 4.10.** *For the proof of Corollary 4.9. 2) it is important to assume that* $\left(\left(c_i^{\mu(n)}\right)_{i\in\{1,\cdots,n^3\}}, \{D_i^{\mu(n)} : i \in \{1,\cdots,n^3\}\}\right)$ *is a code for the channel* $\{(W_t, V_t) : t \in \theta\}$ *and not for* $\{(W_t \circ T_U, V_t \circ T_U) : t \in \theta\}$*, since it may happen that* $\{W_t \circ T_U : t \in \theta\}$ *is symmetrizable although* $\{W_t : t \in \theta\}$ *is not symmetrizable, as the following example shows:*

*We assume that* $\{W_t : t \in \theta\}$ *:*$P(\mathbf{A}) \to \mathcal{S}(H)$ *is not symmetrizable, but there is a subset* $\mathbf{A}' \subset \mathbf{A}$ *such that* $\{W_t : t \in \theta\}$ *limited on* $\mathbf{A}'$ *is symmetrizable. We choose a* $T_U$ *such that for every* $u \in \mathsf{U}$ *there is* $a \in \mathbf{A}'$ *such that* $T_U(a \mid u) = 1$*, and* $T_U(a \mid u) = 0$ *for all* $a \in \mathbf{A} \setminus \mathbf{A}'$ *and* $u \in \mathsf{U}$*. It is clear that* $\{W_t \circ T_U : t \in \theta\}$ *is symmetrizable (cf. also [53] for an example for classical channels).*

## 4.3 Secrecy Capacity under Strong Code Concept

The code concept of Corollary 4.9 still leaves something to be desired because we had to reduce the generality of the code concept when we explicitly allowed a small part of the code word to be non-secure. In light of the importance of shared randomness for robustness the shared randomness is not allowed to be known by the jammer (cf. Corollary 4.15). Thus the code concept of Corollary 4.9 does not work when there is a two-way communication between the jammer and the eavesdropper. Hence, in this Section we analyze arbitrarily varying classical-quantum wiretap channels with strong code concept.

For classical arbitrarily varying wiretap channels the authors of [53] developed a new method to overcome this problem: Applying a technique introduced in [33] they made the first part secure and used it to send the message instead just the common randomness. The code they constructed is thus a one-part deterministic secure code. However, it is technically difficult to extend the random classical code technique introduced in [33] to classical-quantum channels, thus we have to find another way.

### 4.3.1 Classical Arbitrarily Varying Quantum Wiretap Channel

At first we determine a capacity formula for a mixed channel model, i.e. the secrecy capacity of the classical arbitrarily varying quantum wiretap channel. This formula will be used for our result for secrecy capacity of arbitrarily varying classical-quantum wiretap channels using secretly sent common randomness.

**Theorem 4.11.** *Let* $t$ *be a finite set and* $\{(\grave{W}_t, V_t) : t \in \theta\}$ *be a classical arbitrarily varying quantum wiretap channel. If* $\{\grave{W}_t : t \in \theta\}$ *is not symmetrizable, then*

$$C_s(\{(\grave{W}_t, V_t) : t \in \theta\}) = \lim_{n\to\infty} \frac{1}{n} \left( \max_{U \to A^n \to \{B_q^{\otimes n}, Z_{t^n} : q, t^n\}} \min_{q \in P(\theta)} I(p_U, \grave{B}_q^n) - \max_{t^n \in \theta^n} \chi(p_U; Z_{t^n}) \right).$$

*Here* $\grave{B}_t$ *are the resulting classical random variables at the output of the legitimate receiver's channels and* $Z_{t^n}$ *are the resulting quantum states at the output of wiretap channels. The maximum is taken over all random variables that satisfy the Markov chain relationships:* $U \to A^n \to \{B_q^{\otimes n}, Z_{t^n} : q, t^n\}$ *for every* $\grave{B}_q \in Conv((\grave{B}_t)_{t\in\theta})$ *and* $t \in \theta$*.* $A$ *is here a random variable taking values on* $\mathbf{A}$*,* $U$ *a random variable taking values on some finite set* $\mathbf{U}$ *with probability distribution* $p_U$*.*

*Proof.* We fix a probability distribution $p \in P(\mathbf{A})$ and choose an arbitrarily positive $\delta$. Let

$$J_n = \lfloor 2^{\inf_{\grave{B}_q \in Conv((\grave{B}_{t'})_{t' \in \theta})} I(p;\grave{B}_q^n) - \max_{t^n \in \theta^n} \chi(p;Z_{t^n}) - n\delta} \rfloor ,$$

and

$$L_n = \lceil 2^{\max_{t^n \in \theta^n} (\chi(p;Z_{t^n}) + n\delta)} \rceil .$$

Let $p'(x^n) := \begin{cases} \frac{p^n(x^n)}{p^n(\mathcal{T}^n_{p,\delta})} & \text{if } x^n \in \mathcal{T}^n_{p,\delta} ; \\ 0 & \text{else} \end{cases}$

and $X^n := \{X_{j,l}\}_{j \in \{1,\cdots,J_n\}, l \in \{1,\cdots,L_n\}}$ be a family of random matrices whose components are i.i.d. according to $p'$.

We fix a $t^n \in \theta^n$ and define a map $\mathsf{V} : P(\theta) \times P(\mathbf{A}) \to \mathcal{S}(H)$ by

$$\mathsf{V}(t,p) := V_t(p) .$$

For $t \in \theta$ we define $q(t) := \frac{N(t|t^n)}{n}$. $t^n$ is trivially a typical sequence of $q$. For $p \in P(\mathbf{A})$, $\mathsf{V}$ defines a map $\mathsf{V}(\cdot,p) : P(\theta) \to \mathcal{S}(H)$.

Let

$$Q_{t^n}(x^n) := \Pi_{\mathsf{V}(\cdot,p),\alpha}(t^n) \Pi_{\mathsf{V},\alpha}(t^n,x^n) \cdot V_{t^n}(x^n) \cdot \Pi_{\mathsf{V},\alpha}(t^n,x^n) \Pi_{\mathsf{V}(\cdot,p),\alpha}(t^n) .$$

In view of the fact that $\Pi_{\mathsf{V}(\cdot,p),\alpha}(t^n)$ and $\Pi_{\mathsf{V},\alpha}(t^n,x^n)$ are both projection matrices, by (1), (7), and Lemma 3.3 for any $t$ and $x^n$, it holds that

$$\|Q_{t^n}(x^n) - V_{t^n}(x^n)\|_1 \leq \sqrt{2^{-n\beta(\alpha)} + 2^{-n\beta(\alpha)''}} . \tag{172}$$

Now we are going to apply Lemma 4.3 on $\mathsf{V}$.

By (2) we have

$$\begin{aligned} &\mathrm{tr}(\Pi_{\mathsf{V}(\cdot,p),\alpha}(t^n)) \\ &\leq 2^{n(S(\mathsf{V}(\cdot,p)|q) + \delta(\alpha))} \\ &= 2^{n(\sum_t q(t) \mathsf{V}(t,p) + \delta(\alpha))} \\ &= 2^{n(\sum_t q(t) S(V_t(p)) + \delta(\alpha))} . \end{aligned} \tag{173}$$

Furthermore, by (5) for all $x^n$ it holds that

$$\begin{aligned} &\Pi_{\mathsf{V},\alpha}(t^n,x^n) V_{t^n}(x^n) \Pi_{\mathsf{V},\alpha}(t^n,x^n) \\ &\leq 2^{-n(S(\mathsf{V}|r) + \delta(\alpha)')} \Pi_{\mathsf{V},\alpha}(t^n,x^n) \\ &= 2^{-n(\sum_{t,x} r(t,x) S(\mathsf{V}(t,x)) + \delta(\alpha)')} \Pi_{\mathsf{V},\alpha}(t^n,x^n) \end{aligned} \tag{174}$$

where $r$ is a probability distribution on $\theta \times \mathbf{A}$ such that $r(t,x) = q(t) \cdot p(x)$.

We define

$$\theta' := \left\{ t \in \theta : nq(t) \geq \sqrt{n} \right\} .$$

By properties of classical typical set (cf. [67] ) there is a positive $\hat{\beta}(\alpha)$ such that

$$\underset{p'}{Pr}\left(x^n \in \left\{x^n \in \mathbf{A}^n : (x_{\mathtt{I}_t}) \in \mathcal{T}_{p,\delta}^{nq(t)}\ \forall t \in \theta'\right\}\right) \geq \left(1-2^{-\sqrt{n}\hat{\beta}(\alpha)}\right)^{|\theta|} \geq 1-2^{-\sqrt{n}\frac{1}{2}\hat{\beta}(\alpha)}\,, \tag{175}$$

where $\mathtt{I}_t := \{i \in \{1,\cdots,n\} : t_i = t\}$ is an indicator set that selects the indices $i$ in the sequence $t^n = (t_1,\cdots,t_n)$. Here $\underset{p'}{Pr}$ is the probability according to $p'$.

We denote the set $\{x^n : (x_{\mathtt{I}_t}) \in \mathcal{T}_{p,\delta}^{nq(t)}\ \forall t \in \theta'\} \subset \mathbf{A}^n$ by $\mathsf{M}_{t^n}$. For all $x^n \in \mathsf{M}_{t^n}$, if $n$ is sufficiently large, we have

$$\begin{aligned}
&\left|\sum_{t,x} r(t,x)S(\mathsf{V}(t,x)) - \sum_t q(t)S(V_t|p)\right| \\
&\leq \left|\sum_{t\in\theta',x} r(t,x)S(\mathsf{V}(t,x)) - \sum_{t\in\theta'} q(t)S(V_t|p)\right| \\
&+ \left|\sum_{t\notin\theta',x} r(t,x)S(\mathsf{V}(t,x)) - \sum_{t\notin\theta'} q(t)S(V_t|p)\right| \\
&\leq \sum_{t\in\theta'}\left|\sum_x r(t,x)S(\mathsf{V}(t,x)) - q(t)S(V_t|p)\right| + 2|\theta|\frac{1}{\sqrt{n}}C \\
&\leq 2|\theta|\frac{\delta}{n}C + 2|\theta|\frac{1}{\sqrt{n}}C\,,
\end{aligned} \tag{176}$$

where $C := \max_{t\in\theta}\max_{x\in\mathbf{A}}(S(\mathsf{V}(t,x))+S(V_t|p))$. We set $\Xi_{t^n} := \sum_{x^n\in\mathsf{M}_{t^n}} p(x^n)Q_{t^n}(x^n)$. For any $z^n \in \mathsf{M}_{t^n}$ and $t^n \in \theta^n$, $\langle z^n|\Xi_{t^n}|z^n\rangle$ is the expected value of $\langle z^n|Q_{t^n}(x^n)|z^n\rangle$ under the condition $x^n \in \mathsf{M}_{t^n}$.

We choose a positive $\bar{\beta}(\alpha)$ such that $\bar{\beta}(\alpha) \leq \min(\beta(\alpha),\beta(\alpha)')$, and set $\epsilon := 2^{-n\bar{\beta}(\alpha)}$. In view of (174) we now apply Lemma 4.3, where we consider the set $\mathsf{M}_{t^n} \subset \mathbf{A}^n$. If $n$ is sufficiently large, for all $j$ we have

$$\begin{aligned}
&Pr\left(\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_{t^n}(X_{j,l}) - \Xi_{t^n}\|_1 > 2^{-\sqrt{n}\frac{1}{8}\hat{\beta}(\alpha)} + 40\sqrt[8]{\epsilon}\right) \\
&\leq 2^{n(\sum_{t,x} r(t,x)S(\mathsf{V}(t,x))+\delta(\alpha))} \\
&\cdot \exp\left(-L_n\frac{\epsilon^3}{2\ln 2}(1-2^{-\sqrt{n}\frac{1}{2}\hat{\beta}(\alpha)})\cdot 2^{n(\sum_t q(t)S(V_t(p))-\sum_t q(t)S(V_t|p))+\delta(\alpha)+\delta(\alpha)'+2|\theta|\frac{\delta}{n}C+2|\theta|\frac{1}{\sqrt{n}}C}\right) \\
&= 2^{n(\sum_{t,x} r(t,x)S(\mathsf{V}(t,x))+\delta(\alpha)} \\
&\cdot \exp\left(-L_n\frac{\epsilon^3}{2\ln 2}\cdot(1-2^{-\sqrt{n}\frac{1}{2}\hat{\beta}(\alpha)})2^{n(-\sum_t q(t)\chi(p;Z_t)+\delta(\alpha)+\delta(\alpha)'+2|\theta|\frac{\delta}{n}C+2|\theta|\frac{1}{\sqrt{n}}C)}\right)\,.
\end{aligned} \tag{177}$$

The equality holds since $S(V_t(p)) - S(V_t|p) = \chi(p;Z_t)$.

Furthermore, application of a union bound gives

$$Pr\left(\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_{t^n}(X_{j,l}) - \Xi_{t^n}\|_1 > 2^{-\sqrt{n}\frac{1}{8}\hat{\beta}(\alpha)} + 40\sqrt[8]{\epsilon}\ \forall t^n\ \forall j\right)$$

$$\leq J_n|\theta|^n 2^{n(\sum_{t,x} r(t,x)S(\mathsf{V}(t,x))+\delta(\alpha)}$$
$$\cdot \exp\left(-L_n \frac{\epsilon^3}{2\ln 2}(1-2^{-\sqrt{n}\frac{1}{2}\hat{\beta}(\alpha)})2^{n(-\sum_t q(t)\chi(p;Z_t)+\delta(\alpha)+\delta(\alpha)'+2|\theta|\frac{\delta}{n}C+2|\theta|\frac{1}{\sqrt{n}}C)}\right) . \tag{178}$$

We denote the quantum state at the output of the wiretapper's channel when the channel state is $t$ and $j$ has been sent by $\phi_t^j$. We have

$$\begin{aligned}
&\sum_{t\in\theta} q(\theta)\chi\left(p;Z_t\right) - \chi\left(p;\sum_t q(t)Z_t\right) \\
&= \sum_{t\in\theta} q(t) S\left(\sum_{j=1}^{J_n}\frac{1}{J_n}\phi_t^j\right) - \sum_{t\in\theta}\sum_{j=1}^{J_n} q(t)\frac{1}{J_n}S\left(\phi_t^j\right) \\
&- S\left(\frac{1}{J_n}\sum_{t\in\theta}\sum_{j=1}^{J_n} q(t)\phi_t^j\right) + \sum_{j=1}^{J_n}\frac{1}{J_n}S\left(\sum_{t\in\theta} q(t)\phi_t^j\right) .
\end{aligned}$$

Let $H^{\mathfrak{T}}$ be a $|\theta|$-dimensional Hilbert space spanned by an orthonormal basis $\{|t\rangle : t = 1,\cdots,|\theta|\}$. Let $H^{\mathfrak{J}}$ be a $J_n$-dimensional Hilbert space spanned by an orthonormal basis $\{|j\rangle : j = 1,\cdots,J_n\}$. We define

$$\varphi^{\mathfrak{J}\mathfrak{T}H^n} := \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t)|j\rangle\langle j|\otimes|t\rangle\langle t|\otimes\phi_t^j .$$

We have

$$\varphi^{\mathfrak{J}H^n} = \mathrm{tr}_{\mathfrak{T}}\left(\varphi^{\mathfrak{J}\mathfrak{T}H^n}\right) = \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t)|j\rangle\langle j|\otimes\phi_t^j ;$$

$$\varphi^{\mathfrak{T}H^n} = \mathrm{tr}_{\mathfrak{J}}\left(\varphi^{\mathfrak{J}\mathfrak{T}H^n}\right) = \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t)|t\rangle\langle t|\otimes\phi_t^j ;$$

$$\varphi^{H^n} = \mathrm{tr}_{\mathfrak{J}\mathfrak{T}}\left(\varphi^{\mathfrak{J}\mathfrak{T}H^n}\right) = \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t)\phi_t^j .$$

Thus,

$$S(\varphi^{\mathfrak{J}H^n}) = H(R_{uni}) + \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(\sum_{t\in\theta} q(t)\phi_t^j\right) ;$$

$$S(\varphi^{\mathfrak{T}H^n}) = H(Y_q) + \sum_{t\in\theta} q(t) S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\phi_t^j\right) ;$$

$$S(\varphi^{\mathfrak{J}\mathfrak{T}H^n}) = H(R_{uni}) + H(Y_q) + \frac{1}{J_n}\sum_{j=1}^{J_n}\sum_{t\in\theta} q(t) S\left(\phi_t^j\right) ,$$

where $Y_q$ is a random variable on $\theta$ with distribution $q(t)$.

By strong subadditivity of von Neumann entropy, it holds that $S(\varphi^{\Im H^n})+S(\varphi^{\Im H^n}) \geq S(\varphi^{H^n}) + S(\varphi^{\mathrm{i}\Im H^n})$, therefore

$$\sum_t q(t)\chi\left(p; Z_t\right) - \chi\left(p; \sum_t q(t) Z_t\right) \geq 0\,. \tag{179}$$

For an arbitrary $\zeta$, we define $L_n = \lceil 2^{\max_{t^n} \chi(p;Z_{t^n})+n\zeta}\rceil$, and choose a suitable $\alpha$, $\bar{\beta}(\alpha)$, and sufficiently large $n$ such that $6\bar{\beta}(\alpha) + 2\delta(\alpha) + 2\delta(\alpha)' + 2|\theta|\frac{\delta}{n}C + 2|\theta|\frac{1}{\sqrt{n}}C \leq \zeta$. By (179), if $n$ is sufficiently large, we have $L_n \geq \lceil 2^{n(\sum_t q(t)\chi(p;Z_t)+\zeta)}\rceil$ and

$$L_n \frac{\epsilon^3}{2\ln 2}(1 - 2^{-\sqrt{n}\frac{1}{2}\hat{\beta}(\alpha)})2^{n(-\sum_t q(t)\chi(p;Z_t)+\delta(\alpha)+\delta(\alpha)'+2|\theta|\frac{\delta}{n}C+2|\theta|\frac{1}{\sqrt{n}}C)} > 2^{\frac{1}{2}n\zeta}\,.$$

When $n$ is sufficiently large for any positive $\vartheta$ it holds that

$$\begin{aligned} & J_n|\theta|^n 2^{n(\sum_{t,x} r(t,x)S(\mathsf{V}(t,x))+\delta(\alpha))}\exp(-2^{\frac{1}{4}n\zeta}) \\ & \leq 2^{-n\vartheta} \end{aligned}$$

and

$$2^{-\sqrt{n}\frac{1}{8}\hat{\beta}(\alpha)} + 40\sqrt[8]{\epsilon} \leq 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)}\,.$$

Thus for sufficiently large $n$ we have

$$\begin{aligned} & Pr\left(\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_{t^n}(X_{j,l}) - \Xi_{t^n}\|_1 \leq 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)}\ \forall t^n\ \forall j\right) \\ & \geq 1 - 2^{n\vartheta} \end{aligned} \tag{180}$$

for any positive $\varrho$.

In [33] , the following was shown: Let $\{X_{j,l}\}_{j\in\{1,\ldots,J_n\},l\in\{1,\ldots,L_n\}}$ be a family of random variables that are distributed according to according to $p'$. We assume $\{\grave{W}_t : t \in \theta\}$ is not symmetrizable. If $n$ is sufficiently large, and if $J_n \cdot L_n \leq 2^{n(\inf_{\grave{B}_q\in Conv((\grave{B}_t)_{t\in\theta})} I(p;\grave{B}_q)-\mu)}$ for an arbitrary positive $\mu$, there exists a set of mutually disjoint sets $\{D_{j,l} : j \in \{1,\cdots,J_n\}, l \in \{1,\cdots,L_n\}\}$ on $\mathbf{B}^n$ such that for all positive $\epsilon$, $t^n \in \theta^n$, and $j \in \{1,\ldots,J_n\}$,

$$\Pr_{p'}\left[\mathrm{tr}\left(\grave{W}_{t^n}(X_{j,l})D_{j,l}\right) \geq 1 - 2^{-n\beta}\right] > 1 - 2^{-n\gamma}\,. \tag{181}$$

By (180) and (181), when $\{\grave{W}_t : t \in \theta\}$ is not symmetrizable we can find with positive probability a realization $x_{j,l}$ of $X_{j,l}$ and a set of mutually disjoint sets $\{D_{j,l} : j \in \{1,\cdots,J_n\}, l \in \{1,\cdots,L_n\}\}$ such that for all positive $\epsilon$, $t^n \in \theta^n$, and $j \in \{1,\ldots,J_n\}$

$$\max_{t\in\theta}\frac{1}{J_n}\sum_{j=1}^{J_n}\grave{W}_{t^n}(D_{j,l}^c|x_{j,l}) \leq \epsilon\,, \tag{182}$$

and

$$\|\sum_{l=1}^{L_n}\frac{1}{L_n}Q_{t^n}(x_{j,l}) - \Xi_{t^n}\|_1 \leq \epsilon\,. \tag{183}$$

Here we define $E(x^n \mid j) = \frac{1}{L_n}$ if $x^n \in \{x_{j,l} : l \in \{1, \ldots, L_n\}\}$.

We choose a suitable positive $\alpha$. For any given $j' \in \{1, \ldots, J_n\}$, by (172) and (183) we have

$$\begin{aligned}
&\left\| \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(x_{j',l}) - \Xi_{t^n} \right\|_1 \\
&\le \| \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(x_{j',l}) - \sum_{l=1}^{L_n} \frac{1}{L_n} Q_{t^n}(x_{j',l}) \|_1 \\
&\quad + \| \sum_{l=1}^{L_n} \frac{1}{L_n} Q_{t^n}(x_{j',l}) - \Xi_{t^n} \|_1 \\
&\le 2^{-\sqrt{n}\frac{1}{16}\hat{\beta}(\alpha)} + \sqrt{2^{-\frac{1}{2}n\beta(\alpha)} + 2^{-\frac{1}{2}n\beta(\alpha)''}} \\
&\le 2^{-\sqrt{n}\frac{1}{32}\hat{\beta}(\alpha)} . \qquad (184)
\end{aligned}$$

Notice that by (184) we have $\| \frac{1}{J_n \cdot L_n} \sum_{j=1}^{J_n} \sum_{l=1}^{L_n} V_{t^n}(x_{j,l}) - \Xi_{t^n} \|_1 \le 2^{-\sqrt{n}\frac{1}{32}\hat{\beta}(\alpha)}$.

By Lemma 3.6 and the inequality (184), for a uniformly distributed random variable $R_{uni}$ with values in $\{1, \ldots, J_n\}$ a and $t^n \in \theta^n$, we have

$$\begin{aligned}
&\chi(R_{uni}; Z_{t^n}) \\
&= S\left( \sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(x_{j,l}) \right) \\
&- \sum_{j=1}^{J_n} \frac{1}{J_n} S\left( \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(\pi(x_{j,l})) \right) \\
&\le \left| S\left( \sum_{j=1}^{J_n} \frac{1}{J_n} \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(x_{j,l}) \right) - S\left(\Xi_{t^n}\right) \right| \\
&+ \left| S(\Xi_{t^n}) - \sum_{j=1}^{J_n} \frac{1}{J_n} S\left( \sum_{l=1}^{L_n} \frac{1}{L_n} V_{t^n}(x_{j,l}) \right) \right| \\
&\le 2 \cdot 2^{-\sqrt{n}\frac{1}{32}\hat{\beta}(\alpha)} \log(nd-1) + 2h(2^{-\sqrt{n}\frac{1}{32}\hat{\beta}(\alpha)}) . \qquad (185)
\end{aligned}$$

By (185), for any positive $\lambda$ if $n$ is sufficiently large, we have

$$\chi\left(R_{uni}; Z_{t^n}\right) \le \lambda . \qquad (186)$$

We define $E(x^n \mid j) = \begin{cases} \frac{1}{L_n} \text{ if } x^n \in \{x_{j,l} : l \in \{1, \ldots, L_n\}\} ; \\ 0 \text{ if } x \notin \{x_{j,l} : l \in \{1, \ldots, L_n\}\} . \end{cases}$ and $D_j := \bigcup_l D_{j,l}$. By (183) and (186), when $\{\grave{W}_t : t \in \theta\}$ is not symmetrizable the deterministic secrecy capacity of $\{(\grave{W}_t, V_t) : t \in \theta\}$ is larger or equal to

$$\lim_{n\to\infty} \frac{1}{n} \left( \inf_{B_q \in Conv((B_{t'})_{t' \in \theta})} \chi(p; \grave{B}_q^n) - \max_{t^n \in \theta^n} \chi(p; Z_{t^n}) \right) - \varepsilon . \qquad (187)$$

The achievability of $\lim_{n\to\infty} \frac{1}{n} \max_{U \to A^n \to \{B_q^{\otimes n}, Z_{t^n} : q, t^n\}} (\inf_{\grave{B}_q \in Conv((\grave{B}_{t'})_{t' \in \theta})} I(p_U; \grave{B}_q^n)$ $- \max_{t^n \in \theta^n} \chi(p_U; Z_{t^n}))$ and the converse are shown by the standard arguments (cf. [34] and [15]). □

### 4.3.2 The Secure Message Transmission With Strong Code Concept

Now we are going to prove the secrecy capacity formula for arbitrarily varying classical-quantum wiretap channels using secretly sent common randomness. In Corollary 4.9 we determined the secrecy capacity formula for arbitrarily varying classical-quantum wiretap channels. Our strategy is to build a two-part code word, which consists of a non-secure code word and a common randomness-assisted secure code word. The non-secure one is used to create the common randomness for the sender and the legal receiver. The common randomness-assisted secure code word is used to transmit the message to the legal receiver.

We build a code in such a way that the transmission of both the message and the randomization is secure. Since the technique introduced in [33] for classical channels cannot be easily transferred into quantum channels, our idea is to construct a classical arbitrarily varying quantum wiretap channel and apply Theorem 4.11. In [6], a technique has been introduced to construct a classical arbitrarily varying channel by means of an arbitrarily varying classical-quantum channel. However this technique does not work for the classical arbitrarily varying quantum wiretap channel since it cannot provide security. We have to find a more sophisticated way.

**Theorem 4.12.** *If the arbitrarily varying classical-quantum channel $\{W_t : t \in \theta\}$ is not symmetrizable, then*

$$C_s(\{(W_t, V_t) : t \in \theta\}) = \lim_{n\to\infty} \frac{1}{n} \max_{U\to A^n\to\{B_q^{\otimes n}, Z_{t^n}:q,t^n\}} \left( \inf_{B_q\in Conv((B_t)_{t\in\theta})} \chi(p_U; B_q^{\otimes n}) - \max_{t^n\in\theta^n} \chi(p_U; Z_{t^n}) \right), \tag{188}$$

*when we use a two-part code word where both parts are secure.*

*Here $B_t$ are the resulting quantum states at the output of the legitimate receiver's channels. $Z_{t^n}$ are the resulting quantum states at the output of wiretap channels. The maximum is taken over all random variables that satisfy the Markov chain relationships: $U \to A^n \to B_q Z_t$ for every $B_q \in Conv((B_t)_{t\in\theta})$ and $t \in \theta$. Here $A$ is a random variable taking values on $\mathbf{A}$, $U$ a random variable taking values on some finite set $\mathbf{U}$ with probability distribution $p_U$.*

*Proof.* Since the security of both the message and the randomization implies the security of only the message, the secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ for the message and the randomization transmission cannot exceed $\lim_{n\to\infty} \frac{1}{n} \left( \max_{U\to A^n\to\{B_q^{\otimes n}, Z_{t^n}:q,t^n\}} \inf_{B_q\in Conv((B_t)_{t\in\theta})} \chi(p_U; B_q^{\otimes n}) - \max_{t^n\in\theta^n} \chi(p_U; Z_{t^n}) \right)$, which is the secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ for only the message transmission (cf. Theorem 4.6). Thus the converse is trivial.

For the achievability, we at first assume that for all $p \in P(\mathbf{U})$ we have $\lim_{n\to\infty} \frac{1}{n} \max_{U\to A^n\to\{B_q^{\otimes n}, Z_{t^n}:q,t^n\}} \left( \inf_{B_q\in Conv((B_t)_{t\in\theta})} \chi(p_U; B_q^{\otimes n}) - \max_{t^n\in\theta^n} \chi(p_U; Z_{t^n}) \right) \leq 0$. In this case the secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ for only the message transmission is also zero and there is nothing to prove. Now let us assume that $\{W_t : t \in \theta\}$ is not symmetrizable and for all sufficiently large $n$ and a positive $\epsilon$

$$\frac{1}{n} \max_{U\to A^n\to\{B_q^{\otimes n}, Z_{t^n}:q,t^n\}} \left( \inf_{B_q\in Conv((B_t)_{t\in\theta})} \chi(p_U; B_q^{\otimes n}) - \max_{t^n\in\theta^n} \chi(p; Z_{t^n}) \right) > 2\epsilon \tag{189}$$

holds.

*i) Construction of a Channel with Random Pre-coding That is Not Symmetrizable*

We consider the Markov chain $U \to A^n \to \{B_q^{\otimes n}, Z_{t^n} : q, t^n\}$, where we define the classical channel $P(\mathbf{U}) \to P(\mathbf{A})$ by $T_U$. It may happen that $\{W_t \circ T_U : t \in \theta\}$ is symmetrizable although $\{W_t : t \in \theta\}$ is not symmetrizable, as the following example shows:

We assume that $\{W_t : t \in \theta\} : P(\mathbf{A}) \to \mathcal{S}(H)$ is not symmetrizable, but that there is a subset $\mathbf{A}' \subset \mathbf{A}$ such that $\{W_t : t \in \theta\}$ limited on $A'$ is symmetrizable. We choose a $T_U$ such that for every $u \in \mathbf{U}$ there is $a \in \mathbf{A}'$ such that $T_U(a \mid u) = 1$, and $T_U(a \mid u) = 0$ for all $a \in \mathbf{A} \setminus \mathbf{A}'$ and $u \in \mathbf{U}$. It is clear that $\{W_t \circ T_U : t \in \theta\}$ is symmetrizable (cf. also [53] for an example for classical channels).

We now use a technique introduced in [53] to overcome this: Without loss of generality we may assume that $|\mathbf{A}^n| = |\mathbf{U}|$ in the optimizations carried out in (188). Furthermore, without loss of generality we may assume that $\mathbf{A} = \mathbf{U}$ by relabeling the symbols. For every $n \in \mathbb{N} \setminus \{1\}$ we define a new classical channel $\tilde{T}_U^n : P(\mathbf{A}^n) \to P(\mathbf{A}^n)$ by setting $\tilde{T}_U^n := T_U^{n-1} \times id_{\mathbf{A}}$, i.e.,

$$\tilde{T}_U^n(a_1, \cdots, a_{n-1}, a_n) := T_U^n(a_1, \cdots, a_{n-1}) \cdot \delta_{a_n} .$$

We have

$$W_{t^n} \circ \tilde{T}_U^n(a_1, \cdots, a_{n-1}, a_n) = W_{t^n n-1} \circ T_U^n(a_1, \cdots, a_{n-1}) W_{t_n}(a_n) , \qquad (190)$$

where for $t^n = (t_1, \cdots, t_{n-1}, t_n)$ we denote $t^{n-1} := (t_1, \cdots, t_{n-1})$. Since $\{W_t : t \in \theta\}$ is not symmetrizable, $\{W_{t^n} \circ \tilde{T}_U^n : t \in \theta\}$ is not symmetrizable. Furthermore, for any positive $\delta$ sufficiently large $n$ we have

$$C(\{W_{t^n} \circ T_U^n : t \in \theta\}; r) \leq C(\{W_{t^{n-1}} \circ T_U^{n-1} : t \in \theta\}; r) + \delta \leq C(\{W_{t^n} \circ \tilde{T}_U^{n-1} : t \in \theta\}; r) + \delta .$$

For every $n > 1$ and $t^n \in t^n n$ we define $\check{W}_{t^n} : P(\mathbf{U}^n) \to P(\mathbf{A}^n)$ by

$$\check{W}_{t^n} := W_{t^n} \circ \tilde{T}_U^n . \qquad (191)$$

This shows that, if for a channel $\{W_t : t \in \theta\}$ that is not symmetrizable we have (189), then

$$\frac{1}{n} \max_{U \to A^n \to \{\check{B}_q^{\otimes n}, Z_{t^n} : q, t^n\}} \left( \inf_{\check{B}_q \in Conv((\check{B}_t)_{t \in \theta})} \chi(p_U; \check{B}_q^{\otimes n}) - \max_{t^n \in t^n} \chi(p; Z_{t^n}) \right) > \epsilon \qquad (192)$$

holds, where $\check{B}_{t^n}$ are the resulting quantum states at the output of $\check{W}_{t^n}$ as defined in (191).

*ii) Definition of a Classical Arbitrarily Varying Channel which is not Symmetrizable*

We denote $m := \lfloor \log n \rfloor$ and define $\check{V}_t := V_t \circ T_U$ for all $t \in \theta$. Now we consider the arbitrarily varying wiretap classical-quantum channel $\{(\check{W}_t, \check{V}_t); t \in \theta\}$. We choose an arbitrary $\delta > 0$. By (192), if $m$ is sufficiently large we may assume that for at least one $p \in P(\mathbf{U})$

$$2 \leq 2^{m\epsilon} \leq \lfloor 2^{m \inf_{\check{B}_q \in Conv((\check{B}_{t'})_{t' \in \theta})} \chi(p; \check{B}_q) - \max_{t^m \in \theta^m} \chi(p; Z_{t^m}) - m\delta} \rfloor .$$

By Theorem 4.2, if $m$ is sufficiently large we can find a $(m, 2)$ code $\big(E^m, \{D_j^m : j \in \{1,2\}\}\big)$ and positive $\lambda$, $\zeta$ such that for all $q \in P(\theta)$

$$\frac{1}{2}\sum_{j=1}^{2} \mathrm{tr}\left(\check{W}_q^{\otimes m}(E^m(\,|j))D_j^m\right) \geq 1 - 2^{-m^{1/16}\lambda} \tag{193}$$

and for all $t^m \in \theta^m$ and $\pi \in \Pi_m$

$$\|\check{V}_{t^m}(\pi(E^m(\,|j))) - \Xi_{t^m}\|_1 < 2^{-\sqrt{m}\zeta} \tag{194}$$

for a $\Xi_{t^m} \in \mathcal{S}(H^m)$ which is independent of $j$. Here, for $\pi \in \mathsf{S}_m$, we define its matrix representation on on $H^{\otimes m}$ by $P_\pi$.

Notice that $\big(E^m, \{D_j^m : j \in \{1,2\}\}\big)$ is a deterministic code for a mixed channel model called compound-arbitrarily varying wiretap classical-quantum channel which we introduced in Section 4.2.1.

We now combine a technique introduced in [6] with the concept of the superposition code to define a set of classical channels.

We choose $d^{m^2}+1$ Hermitian operators $L_i \geq 0$, $i = 1, \cdots, d^{m^2}+1$ which span the space of Hermitian operators on $H^m$ and fulfill $\sum_{i=1}^{d^{m^2}+1} L_i = id_{H^m}$ by the technique introduced in [6]: We choose arbitrarily $d^{m^2}$ Hermitian operators $\bar{L}_i \geq 0$, $i = 1, \cdots, d^{m^2}$ which span the space of Hermitian operators on $H^m$ and denote the trace of $\sum_{i=1}^{d^{m^2}} \bar{L}_i$ by $\lambda'$. Now we define $L_i := \frac{1}{\lambda'}\bar{L}_i$ for $i \in \{1, \cdots, d^{m^2}\}$ and $L_{d^{m^2}+1} := id_{H^m} - \sum_{i=1}^{d^{m^2}} L_i$.

Now we defined the classical arbitrarily varying channel $\{\grave{W}_{t^m} : t^m \in \theta^m\} : P(\mathbf{U}^m) \to P(\{1, \cdots, d^{m^2}+3\})$ by

$$\grave{W}_{t^m}(i \mid p^m) := \begin{cases} \frac{1}{2}\mathrm{tr}\left(\check{W}_{t^m}(p^m)D_i^m\right) & \text{for } i \in \{1,2\}\,; \\ \frac{1}{2}\mathrm{tr}\left(\check{W}_{t^m}(p^m)L_{i-2}\right) & \text{for } i = 3, \cdots, d^{m^2}+3\,. \end{cases} \tag{195}$$

Since $\frac{1}{2}\sum_{j=1}^{2} D_j^m + \frac{1}{2}\sum_{i=1}^{d^{m^2}+1} L_i = id_{H^m}$ we have

$$\sum_{i=1}^{d^{m^2}+3} \grave{W}_{t^m}(i \mid p^m) = 1$$

for all $p^m \in P(\mathbf{U}^m)$. Thus the definition in (195) is valid.

When $\{\grave{W}_{t^m} : t^m \in \theta^m\}$ is symmetrizable then there is a $\{\tau(\cdot \mid a^m) : a^m \in \mathbf{U}^m\}$ on $\theta^m$ such that

$$\sum_{t^m \in \theta^m} \tau(t^m \mid a^m)\grave{W}_{t^m}(i \mid a'^m) = \sum_{t^m \in \theta^m} \tau(t^m \mid a'^m)\grave{W}_{t^m}(i \mid a^m)$$

for all $i \in \{1, \cdots, d^{m^2}+3\}$ and all $a^m, a'^m \in \mathbf{U}^m$. This implies that

$$\frac{1}{2}\sum_{t^m \in \theta^m} \tau(t^m \mid a^m)\mathrm{tr}\left(\check{W}_{t^m}(a'^m)L_i\right) = \frac{1}{2}\sum_{t^m \in \theta^m} \tau(t^m \mid a'^m)\mathrm{tr}\left(\check{W}_{t^m}(a^m)L_i\right)$$

for all $i \in \{1, \cdots, d^{m^2}+1\}$ and all $a^m, a'^m \in \mathbf{U}^m$.

Since $\{L_i : i = 1, \cdots, {d^m}^2\}$ span the space of Hermitian operators on $H^m$, and all $a^m, {a'}^m \in \mathbf{U}^m$ we have

$$\sum_{t^m \in \theta^m} \tau(t^m \mid a^m) \check{W}_{t^m}({a'}^m) = \sum_{t^m \in \theta^m} \tau(t^m \mid {a'}^m) \check{W}_{t^m}(a^m) \, .$$

This is a contradiction to our assumption that $\{\check{W}_{t^m} : t^m \in \theta^m\}$ is not symmetrizable, therefore $\{\grave{W}_{t^m} : t^m \in \theta^m\}$ is not symmetrizable.

*iii) The Deterministic Secrecy Capacity of* $\{(\grave{W}_{t^m}, \check{V}_{t^m}) : t^m \in \theta^m\}$ *is Positive*

By (193) for all $q \in P(\theta)$ and $j \in \{1, 2\}$ we have

$$\grave{W}_q(j \mid E^m(\cdot \mid j)) \geq \frac{1}{2} - \frac{1}{2} 2^{-m^{1/16}\lambda} \, , \tag{196}$$

and for all $q \in P(\theta)$ and $j \neq i \in \{1, 2\}$

$$\grave{W}_q(j \mid E^m(\cdot \mid i)) \leq \frac{1}{2} 2^{-m^{1/16}\lambda} \, . \tag{197}$$

We denote the uniform distribution on $\{1, 2\}$ by $R'$. For any positive $\zeta'$, if $m$ is sufficiently large by (196) and (197), for all $q \in P(\theta)$ we have

$$\begin{aligned}
&\min_{q \in P(t^m)} I\left(E^m(\cdot \mid R'), \grave{B}_q\right) \\
&> (\frac{1}{2} - \frac{1}{2} 2^{-m^{1/16}\lambda}) \log (\frac{1}{2} - \frac{1}{2} 2^{-m^{1/16}\lambda}) - (\frac{1}{2} + \frac{1}{2} 2^{-m^{1/16}\lambda}) \log (\frac{1}{4} + \frac{1}{4} 2^{-m^{1/16}\lambda}) - \zeta' \\
&\geq \frac{1}{2} - 2\zeta' \, ,
\end{aligned} \tag{198}$$

where $\grave{B}_q$ is the resulting distribution at the output of $\grave{W}_q$.

Applying the Lemma 3.6 and (194), if $m$ is sufficiently large for any $n' \in \mathbb{N}$, positive $\zeta'$, and for all $t^{mn'} = (t_1^m, \cdots, t_{n'}^m) = (t_1, \cdots, t_m, t_{m+1}, \cdots, t_{2m}, t_{2m+1}, \cdots, t_{mn'}) \in \theta^{mn'}$ we have

$$\begin{aligned}
&\frac{1}{n'} \chi\left({R'}^{\otimes n'}; \check{Z}_{t^{mn'}}\right) \\
&= \frac{1}{n'} \left( S\left( \frac{1}{2^{n'}} \sum_{j \in \{1,2\}^{n'}} \check{V}_{t^{mn'}}((E^m)^{\otimes n'}(\cdot \mid j)) \right) - \frac{1}{2^{n'}} \sum_{j \in \{1,2\}^{n'}} S\left( \check{V}_{t^{mn'}}((E^m)^{\otimes n'}(\cdot \mid j)) \right) \right) \\
&\leq \frac{1}{n'} \left| S\left( \frac{1}{2^{n'}} \sum_{j \in \{1,2\}^{n'}} \check{V}_{t^{mn'}}((E^m)^{\otimes n'}(\cdot \mid j)) \right) - S\left(\Xi_{t^{mn'}}\right) \right| \\
&+ \frac{1}{n'} \left| S\left(\Xi_{t^{mn'}}\right) - \frac{1}{2^{n'}} \sum_{j \in \{1,2\}^{n'}} S\left( \check{V}_{t^{mn'}}((E^m)^{\otimes n'}(\cdot \mid j)) \right) \right| \\
&= \frac{1}{n'} \left| \sum_{i=1}^{n'} \left( S\left( \frac{1}{2} \sum_{j \in \{1,2\}} \check{V}_{t_i^m}((E^m)(\cdot \mid j)) \right) - S\left(\Xi_{t_i^m}\right) \right) \right| \\
&+ \frac{1}{n'} \left| \sum_{i=1}^{n'} \left( S\left(\Xi_{t_i^m}\right) - \frac{1}{2} \sum_{j \in \{1,2\}} S\left( \check{V}_{t_i^m}((E^m)(\cdot \mid j)) \right) \right) \right|
\end{aligned}$$

$$\leq 2 \cdot 2^{-\sqrt{m}\zeta} \log(d^m - 1) + 2 \cdot h(2^{-\sqrt{m}\zeta})$$
$$\leq \zeta' , \tag{199}$$

where $\check{Z}_{t^{mn'}}$ is the resulting quantum state at the output of $\check{V}_{t^{mn'}}$.

We choose $\zeta' < \frac{1}{18}$ and a sufficiently large $m$ such that (198) and (199) hold. Since $\{\grave{W}_{t^m} : t^m \in \theta^m\}$ is not symmetrizable, by Theorem 4.11 the deterministic secrecy capacity of $\{(\grave{W}_{t^m}, \check{V}_{t^m}) : t^m \in \theta^m\}$ is equal to

$$\limsup_{n' \to \infty} \frac{1}{n'} \max_{p \in P(\mathbf{U}^m)} \min_{q \in P(t^n m)} I(p, \grave{B}_q^{n'}) - \max_{t^{mn'} \in \theta^{mn'}} \chi(p; \check{Z}_{t^{mn'}}) \geq \frac{1}{2} - 3\zeta' > \frac{1}{3} .$$

*iv) The Secure Transmission of the Randomization Index with a Deterministic Code*

We define $r(n) := \lfloor (\log n)^3 \rfloor$. Since $(\log n)^2 > 3 \log(n^3)$ for sufficiently large $n$, we can build a $(\lfloor (\log n)^2 \rfloor, n^3)$ code $(\tilde{E}^{r(n)}, \{\tilde{S}_i^{r(n)} : i \in \{1, \cdots, n^3\}\})$ such that

$$1 - \min_{t^{r(n)} \in \theta^{r(n)}} \min_{i \in \{1, \cdots, n^3\}} \grave{W}_{t^{r(n)}} \left( \tilde{S}_i^{r(n)} \mid \tilde{E}^{r(n)} (\cdot \mid i) \right) \leq \varepsilon \tag{200}$$

and

$$\max_{t^{r(n)} \in \theta^{r(n)}} \left\| \check{V}_{t^{r(n)}} \left( \tilde{E}^{r(n)} (\cdot \mid i) \right) - \Xi_{t^{r(n)}} \right\| \leq \varepsilon \tag{201}$$

for a $\Xi_{t^{r(n)}} \in \mathcal{S}(H^{r(n)})$ which is independent of $j$.

We define $D_j := L_{j-2}$ for $j \in \{3, \cdots, d^{m^2} + 3\}$. For $i \in \{1, \cdots, n^3\}$ we define

$$\tilde{D}_i^{r(n)} := \frac{1}{2} \sum_{j^m \in \tilde{S}_i^{r(n)}} D_{j^m} .$$

Here for $j^m = (j_1, \cdots, j_m)$ we set $D_{j^m} = D_{j_1} \otimes \cdots \otimes D_{j_m}$. Since $\sum_{i=1}^{n^3} \tilde{D}_i^{r(n)}$ $= \frac{1}{2} \sum_{j^m \in \{1, \cdots, d^{m^2}+3\}^m} D_{j^m} = id_{H^m}$, $\{\tilde{D}_i^{r(n)} : i \in \{1, \cdots, n^3\}\}$ is a valid set of decoding operators.

$(\tilde{E}^{r(n)}, \{\tilde{D}_i^{r(n)} : i \in \{1, \cdots, n^3\}\})$ is a $(r(n), n^3)$ code which fulfills

$$\min_{t^{r(n)} \in \theta^{r(n)}} \frac{1}{n^3} \sum_{i=1}^{n^3} \mathrm{tr}(\check{W}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i)) \tilde{D}_i^{r(n)}) \geq 1 - 2^{-n^{1/16}\lambda} . \tag{202}$$

*v) The Secure Transmission of Both the Message and the Randomization Index*

We choose an arbitrary positive $\delta$. Let

$$J_n = \frac{1}{n} \left( \max_{U \to A^n \to \{B_q^n, Z_{t^n} : q, t^n\}} \left( \inf_{B_q \in Conv((B_t)_{t \in \theta})} \chi(p_U; B_q^{\otimes n}) - \max_{t^n \in \theta^n} \chi(p_U; Z_{t^n}) \right) \right) - \delta .$$

By the results of Section 4.2.2 if $n$ is sufficiently large there is a $(n, J_n)$ common randomness assisted quantum code $\left\{ \left( \pi \circ E^n, \{P_\pi D_j^n P_\pi^\dagger : j \in \{1, \cdots, J_n\}\} \right) : \pi \in \mathsf{S}_n \right\}$ such that for all $t^n \in \theta^n$

$$\frac{1}{n!} \frac{1}{J_n} \sum_{\pi \in \mathsf{S}_n} \sum_{j=1}^{J_n} \mathrm{tr} \left( \check{W}_{t^n} \left( E^n(\pi(\cdot|j)) \right) P_\pi D_j^n P_\pi^\dagger \right) \geq 1 - 2^{-n^{1/16}\lambda} , \tag{203}$$

and for all $t^n \in \theta^n$, $j \in \{1, \cdots, J_n\}$ and all $\pi \in \mathsf{S}_n$

$$\|\check{V}_{t^n}(\pi(E^n(\cdot|j))) - P_\pi \Xi_{\pi^{-1}(t^n)} P_\pi^\dagger\|_1 < 2^{-\sqrt{n}\zeta} \tag{204}$$

where $\Xi_{t^n}$ is defined as in Section 4.3.1.

Using the technique in Section 4.1 to reduce the amount of common randomness if $n$ is sufficiently large, we can find a set $\{\pi_1, \cdots, \pi_{n^3}\} \subset \mathsf{S}_n$ such that

$$\max_{t^n \in \theta^n} \frac{1}{n^3}\frac{1}{J_n}\sum_{i=1}^{n^3}\sum_{j=1}^{J_n} \mathrm{tr}\left(\check{W}_{t^n}(\pi_i(E^n(\cdot|j)))P_{\pi_i} D_j^n P_{\pi_i}^\dagger\right) \geq 1 - 2\cdot 2^{-n^{1/16}\lambda}, \tag{205}$$

and

$$\|\check{V}_{t^n}(\pi_i(E^n(\cdot|j))) - P_{\pi_i}\Xi_{\pi_i^{-1}(t^n)} P_{\pi_i}^\dagger\|_1 < 2\cdot 2^{-\sqrt{n}\zeta}, \tag{206}$$

Furthermore, for any $\pi \in \mathsf{S}_n$ we have $\check{V}_{t^n}(\pi(x^n)) = P_\pi \check{V}_{\pi^{-1}(t^n)}(x^n) P_\pi^\dagger$. Thus

$$\begin{aligned}
&P_\pi \Xi_{\pi^{-1}(t^n)} P_\pi^\dagger \\
&= P_\pi \left(\sum_{x^n \in \mathcal{T}_{p,\delta}} p'(x^n) Q_{\pi^{-1}(t^n)}(x^n)\right) P_\pi^\dagger \\
&= \sum_{x^n \in \mathcal{T}_{p,\delta}} p'(x^n) Q_{t^n}(\pi(x^n)) \\
&= \sum_{\pi(x^n) \in \mathcal{T}_{p,\delta}} p'(x^n) Q_{t^n}(\pi(x^n)) .
\end{aligned}$$

Since $\{\pi(x^n) \in \mathcal{T}_{p,\delta}\} = \{x^n \in \mathcal{T}_{p,\delta}\}$ we have

$$\Xi_{t^n} = P_\pi \Xi_{\pi(t^n)} P_\pi^\dagger \tag{207}$$

for all $\pi \in \mathsf{S}_n$.

Now we can construct a $(r(n)+n, n^3 J_n)$ code $\left(E^{r(n)+n}, \{D_{i,j}^{r(n)+n} : i = 1, \cdots, n^3,\ j = 1, \cdots J_n\}\right)$ by

$$E^{r(n)+n}(a^{r(n)+n} \mid i, j) := \tilde{E}^{r(n)}(a^{r(n)}|i) \cdot E^n(\pi_i(a^n)|j), \tag{208}$$

for every $a^{r(n)+n} = (a^{r(n)}, a^n) \in \mathbf{U}^{r(n)+n}$ and

$$D_{i,j}^{r(n)+n} := \tilde{D}_i^{r(n)} \otimes (P_{\pi_i} D_j^n P_{\pi_i}^\dagger). \tag{209}$$

By (202) and (205), for every $t^{r(n)+n} = (t^{r(n)}, t^n) \in \theta^{r(n)+n}$, we have

$$\begin{aligned}
&\frac{1}{n^3}\frac{1}{J_n}\sum_{i=1}^{n^3}\sum_{j=1}^{J_n} \mathrm{tr}\left(\check{W}_{t^{r(n)+n}}(E^{r(n)+n}(\cdot \mid i,j)) D_{i,j}^{r(n)+n}\right) \\
&= \frac{1}{n^3}\frac{1}{J_n}\sum_{i=1}^{n^3}\sum_{j=1}^{J_n} \mathrm{tr}\left(\left[\check{W}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i)) \otimes \left(\check{W}_{t^n}(\pi_i(E^n(\cdot|j)))\right)\right]\left[\tilde{D}_i^{r(n)} \otimes (P_{\pi_i} D_j^n P_{\pi_i}^\dagger)\right]\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n^3}\sum_{i=1}^{n^3} \mathrm{tr}\left(\left[\check{W}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i))\tilde{D}_i^{r(n)}\right] \otimes \left[\frac{1}{J_n}\sum_{j=1}^{J_n}\left(\check{W}_{t^n}(\pi_i(E^n(\cdot|j)))\right) P_{\pi_i} D_j^n P_{\pi_i}^{\dagger}\right]\right) \\
&= \frac{1}{n^3}\sum_{i=1}^{n^3}\left(\mathrm{tr}\left(\check{W}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i))\tilde{D}_i^{r(n)}\right)\cdot \mathrm{tr}\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\left(\check{W}_{t^n}(\pi_i(E^n(\cdot|j)))\right) P_{\pi_i} D_j^n P_{\pi_i}^{\dagger}\right)\right) \\
&\geq 1 - \frac{1}{n^{1/16}}2^{\lambda} - 2\cdot 2^{-n^{1/16}\lambda} \\
&\geq 1-\varepsilon
\end{aligned} \tag{210}
$$

for any positive $\varepsilon$ when $n$ is sufficiently large.

By (201) and (206), for every $t^{r(n)+n} = (t^{r(n)}, t^n) \in \theta^{r(n)+n}$ and $i \in \{1, \cdots, n^3\}$ and $j \in \{1, \cdots, J_n\}$, we have

$$
\begin{aligned}
&\|\check{V}_{t^{r(n)+n}}(E^{r(n)+n}(\cdot \mid i,j)) - \Xi_{t^{r(n)}} \otimes \Xi_{t^n}\|_1 \\
&= \|\check{V}_{t^{r(n)}}\left(\tilde{E}^{r(n)}(\cdot \mid i)\right) \otimes \check{V}_{t^n}(\pi(E^n(\cdot|j))) - \Xi_{t^{r(n)}} \otimes \Xi_{t^n}\|_1 \\
&< \frac{1}{\sqrt{n}}2^{\zeta} + 2\cdot 2^{-\sqrt{n}\zeta}\,.
\end{aligned} \tag{211}
$$

Let $R_{n^3}$ be the uniform distribution on $\{1, \cdots, n^3\}$. We define a random variable $R_{n^3,uni}$ on the set $\{1, \cdots, n^3\} \times \{1, \cdots, R_n\}$ by $R_{n^3,uni} := R_{n^3} \times R_{uni}$. Applying Lemma 3.6 we obtain

$$
\begin{aligned}
&\max_{t^{r(n)+n}\in\theta^{r(n)+n}} \chi\left(R_{n^3,J_n}; Z_{t^{r(n)+n}}\right) \\
&\leq \max_{t^{r(n)}\in\theta^{r(n)}} \chi\left(R_{n^3}; Z_{t^{r(n)}+n}\right) \\
&+ \frac{1}{n^3}\sum_{i=1}^{n^3}\max_{t^n\in\theta^n}\chi\left(R_{uni}; \check{V}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i)) \otimes Z_{t^n,\pi_i}\right) \\
&= \max_{t^{r(n)}\in\theta^{r(n)}}\left(S\left(\frac{1}{n^3}\frac{1}{J_n}\sum_{i=1}^{n^3}\sum_{j=1}^{J_n}\check{V}_{t^{r(n)+n}}(E^{r(n)+n}(\cdot \mid i,j))\right)\right. \\
&\left. - \frac{1}{n^3}\sum_{i=1}^{n^3} S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\check{V}_{t^{r(n)+n}}(E^{r(n)+n}(\cdot \mid i,j))\right)\right) \\
&+ \max_{t^n\in\theta^n}\frac{1}{n^3}\sum_{i=1}^{n^3}\left(S\left(\frac{1}{n^3}\frac{1}{J_n}\sum_{i=1}^{n^3}\sum_{j=1}^{J_n}\check{V}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i)) \otimes \left(\check{V}_{t^n}(\pi_i(E^n(\cdot|j)))\right)\right)\right. \\
&\left. - \frac{1}{J_n}\sum_{j=1}^{J_n} S\left(\check{V}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i)) \otimes \left(\check{V}_{t^n}(\pi_i(E^n(\cdot|j)))\right)\right)\right) \\
&\leq \max_{t^{r(n)}\in\theta^{r(n)}}\left(\left|S\left(\frac{1}{n^3}\frac{1}{J_n}\sum_{i=1}^{n^3}\sum_{j=1}^{J_n}\check{V}_{t^{r(n)+n}}(E^{r(n)+n}(\cdot \mid i,j))\right) - \Xi_{t^{r(n)}} \otimes \Xi_{t^n}\right|\right.
\end{aligned}
$$

$$
\begin{aligned}
&+\left|\Xi_{t^{r(n)}}\otimes\Xi_{t^n}-\frac{1}{n^3}\sum_{i=1}^{n^3}S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\check{V}_{t^{r(n)+n}}(E^{r(n)+n}(\cdot\mid i,j))\right)\right|\Bigg)\\
&+\max_{t^n\in\theta^n}\frac{1}{n^3}\Bigg(\left|S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\check{V}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i))\otimes\left(\check{V}_{t^n}(\pi_i(E^n(\cdot|j)))\right)\right)-\check{V}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i))\otimes\Xi_{t^n}\right|\\
&+\left|\check{V}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i))\otimes\Xi_{t^n}-\frac{1}{J_n}\sum_{j=1}^{J_n}S\left(\check{V}_{t^{r(n)}}(\tilde{E}^{r(n)}(\cdot|i))\otimes\left(\check{V}_{t^n}(\pi_i(E^n(\cdot|j)))\right)\right)\right|\Bigg)\\
&\le(\frac{1}{\sqrt{n}}2^{\zeta}+2\cdot2^{-\sqrt{n}\zeta})\log(d^{r(n)}-1)+h(\frac{1}{\sqrt{n}}2^{\zeta}+2\cdot2^{-\sqrt{n}\zeta})\\
&+2\cdot2^{-\sqrt{n}\zeta}\log(d^n-1)+h(2\cdot2^{-\sqrt{n}\zeta})\\
&\le\varepsilon
\end{aligned}
\tag{212}
$$

for any positive $\varepsilon$ when $n$ is sufficiently large. Here $Z_{i,t^n}$ is the resulting quantum state at $\check{V}_{t^n}$ after $i\in\{1,\cdots,n^3\}$ has been sent with $E^{r(n)}$.

For any positive $\delta$, if $n$ is sufficiently large, we have $\frac{1}{n}\log J_n-\frac{1}{(\log n)^3+n}\log J_n\le\delta$. Thus the secrecy rate of $\{(W_t,V_t):t\in\theta\}$ to transmit both the message and the randomization index is larger than

$$
\frac{1}{n}\max_{U\to A^n\to\{B_q^{\otimes n},Z_{t^n}:q,t^n\}}\Big(\inf_{B_q\in Conv((B_t)_{t\in\theta})}\chi(p_U;B_q^{\otimes n})-\max_{t^n\in\theta^n}\chi(p_U;Z_{t^n})\Big)-2\delta\,.
$$

□

**Corollary 4.13 (Ahlswede Dichotomy under strong code concept).** *Let $\theta$ be a finite set and $\{(W_t,V_t):t\in\theta\}$ be an arbitrarily varying classical-quantum wiretap channel.*

*1) If the arbitrarily varying classical-quantum channel $\{W_t:t\in\theta\}$ is not symmetrizable, then*

$$
C_s(\{(W_t,V_t):t\in\theta\})=C_s(\{(W_t,V_t):t\in\theta\};r)\,. \tag{213}
$$

*2) If $\{W_t:t\in\theta\}$ is symmetrizable,*

$$
C_s(\{(W_t,V_t):t\in\theta\})=0\,. \tag{214}
$$

*Proof.* 1) follows immediately from Theorem 4.12 and Theorem 4.6.

Let us assume that $\{W_t:t\in\theta\}$ is symmetrizable. By the Theorem 4.1, the secrecy capacity of $\{(W_t,V_t):t\in\theta\}$ for only the message transmission is zero. The secrecy capacity for the message and the randomization transmission cannot exceed the secrecy capacity for only the message transmission, thus 2) holds. □

## 4.4 Communication with Resources

As we learn from Example 4.23, there are indeed arbitrarily varying classical-quantum wiretap channels which have zero deterministic secrecy capacity and positive random secrecy capacity. Therefore, as Theorem 1 shows, randomness is indeed a very helpful resource for the secure message transmission through an arbitrarily varying classical-quantum wiretap channel. But the problem is: how should the sender and the receiver know which code is used in the particular transmission?

Theorem 2 shows that common randomness assisted secrecy capacity is always equal to the random secrecy capacity, even for the arbitrarily varying classical-quantum wiretap channels of Example 4.23. Therefore, common randomness is an equally helpful resource for the secure message transmission through an arbitrarily varying classical-quantum wiretap channel. However, as [28] showed, common randomness is a very "costly" resource. As Theorem 4.1 shows, for the transmission of common randomness we have to require that the deterministic capacity for message transmission of the sender's and legal receiver's channel is positive. In Section 4.4.1, we will see that the much "cheaper" resource, the $m-a-(X,Y)$ correlation, is also an equally helpful resource for the message transmission through an arbitrarily varying classical-quantum channel. The advantage here is that we do not have to require that the deterministic capacity for message transmission of the sender's and legal receiver's channel is positive.

### 4.4.1 Arbitrarily Varying Classical-Quantum Wiretap Channel with Correlation Assistance

In this section we consider the $m-a-(X,Y)$ correlation assisted secrecy capacity of an arbitrarily varying classical-quantum wiretap channel.

Theorem 2 shows that common randomness is a helpful resource for the secure message transmission through an arbitrarily varying classical-quantum wiretap channel. The $m-a-(X,Y)$ correlation is a weaker resource than common randomness (cf. [28]). We can simulate any $m-a-(X,Y)$ correlation by common randomness asymptotically, but there exists a class of sequences of bipartite distributions which cannot model common randomness (cf. Lemma 1 of [28]). However, the results of [28] show that the "cheaper" $m-a-(X,Y)$ correlation is nevertheless a helpful resource for message transmission through an arbitrarily varying classical-quantum channel. Our following Theorem 4.14 shows that also in case of secure message transmission through an arbitrarily varying classical-quantum wiretap channel, the $m-a-(X,Y)$ correlation assistance is an equally helpful resource as common randomness.

**Theorem 4.14.** *Let $\{(W_t,V_t) : t \in \theta\}$ be an arbitrarily varying classical-quantum wiretap channel. Let $\mathbf{X}$ and $\mathbf{Y}$ be finite sets. If $I(X,Y) > 0$ holds for a random variable $(X,Y)$ which is distributed according to a joint probability distribution $p \in P(\mathbf{X}\times\mathbf{Y})$, then the randomness assisted secrecy capacity is equal to the $m-a-(X,Y)$ correlation assisted secrecy capacity.*

*Proof.* Our proof is similar to the capacity results of arbitrarily varying channels with correlation assistance in [7] and [28].

*i) When the Randomness Assisted Code Has Positive Secrecy Capacity*

If the randomness assisted secrecy capacity of $(W_t,V_t)_{t\in\theta}$ is positive, we can build a new arbitrarily varying classical-quantum channel $\{\tilde{U}_t : t \in \theta\}$ to create common randomness for the sender and the legal receiver. We show that this channel does

not have to be secure to be useful for a secure code for the original arbitrarily varying classical-quantum wiretap channel. Then, similar to our proof of Theorem 1, the sender and the legal receiver can build two-part code words, which consist of a non-secure code words for $\{\tilde{U}_t : t \in \theta\}$ to pass the index and common randomness assisted secure code words to transmit the message.

At first we assume that the $m-a-(X,Y)$ secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ is positive, then the $m-a-(X,Y)$ capacity of the arbitrarily varying classical-quantum channel $\{W_t : t \in \theta\}$ is positive. For the definition of the capacity of an arbitrarily varying classical-quantum channel please see [28].

By Theorem 2, the randomness assisted secrecy capacity is equal to the common randomness assisted secrecy capacity. Let $\delta > 0$, $\zeta > 0$, and $\epsilon > 0$, and $\left\{\mathcal{C}^\gamma = \left(E_\gamma^n, \{D_{\gamma,j}^n : j \in \{1, \cdots J_n\}\}\right) : \gamma \in \Gamma\right\}$ be an $(n, J_n)$ common randomness assisted quantum code such that $\frac{\log J_n}{n} > C_s((W_t, V_t)_{t\in\theta}, r) - \delta$, and

$$\max_{t^n \in \theta^n} \frac{1}{|\Gamma|} \sum_{\gamma=1}^{|\Gamma|} P_e(\mathcal{C}^\gamma, t^n) < \epsilon \,,$$

$$\max_{t^n \in \theta^n} \frac{1}{|\Gamma|} \sum_{\gamma=1}^{|\Gamma|} \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) < \zeta \,.$$

We denote $\mathfrak{F} := \{f : f \text{ is a function } \mathbf{X} \to \mathbf{A}\}$. Let $H_{\mathbf{Y}}$ be a Hilbert space of dimension $|\mathbf{Y}|$ and $\{\breve{\kappa}_y : y \in \mathbf{Y}\}$ be a set of pairwise orthogonal and pure states on $H_{\mathbf{Y}}$. For every $t \in \theta$,

$$\tilde{U}_t(f) := \sum_{\mathbf{x}} \sum_{\mathbf{y}} p(\mathbf{x}, \mathbf{y}) \breve{\kappa}_y \otimes W_t\left(f(\mathbf{x})\right) \tag{215}$$

defines a classical-quantum channel

$$\tilde{U}_t : \mathfrak{F} \to \mathcal{S}(H \otimes H_{\mathbf{Y}}) \,.$$

$\{\tilde{U}_t : t \in \theta\}$ defines an arbitrarily varying classical-quantum channel $\mathfrak{F} \to \mathcal{S}(H) \otimes H_{\mathbf{Y}}$.

In [28] (see also [7] for a classical version), it was shown that if $I(X,Y)$ is positive, the deterministic capacity of $(\tilde{U}_t)_{t\in\theta}$ is equal to the $m-a-(X,Y)$ capacity of $\{W_t : t \in \theta\}$. By Remark 2.21, we may assume that the deterministic capacity of $(\tilde{U}_t)_{t\in\theta}$ using deterministic encoder is positive. This means that the sender and the receiver can build a code $\left(\left(f_\gamma^{\nu(n)}\right)_{\gamma=1,\cdots,|\Gamma|}, \{D_\gamma^{\nu(n)} : \gamma = 1, \cdots, |\Gamma|\}\right)$ with deterministic encoder for $(\tilde{U}_t)_{t\in\theta}$ of length $\nu(n)$, where $2^{\nu(n)}$ is in polynomial order of $n$ and $f_\gamma^{\nu(n)}(\mathbf{x}^{\nu(n)}) = \left(f_{\gamma,1}(\mathbf{x}_1), \cdots, f_{\gamma,\nu(n)}(\mathbf{x}_{\nu(n)})\right)$ for $\mathbf{x}^{\nu(n)} = (\mathbf{x}_1, \cdots, \mathbf{x}_{\nu(n)})$, such that the following statement is valid. For any positive $\vartheta$, if $n$ is large enough, we have

$$1 - \vartheta$$

$$\leq \min_{t^{\nu(n)}\in\theta^{\nu(n)}} \frac{1}{|\Gamma|}\sum_{\gamma=1}^{|\Gamma|}\mathrm{tr}\Big(\tilde{U}_{t^{\nu(n)}}(f_\gamma^{\nu(n)})D_\gamma^{\nu(n)}\Big)$$

$$= \min_{t^{\nu(n)}\in\theta^{\nu(n)}} \frac{1}{|\Gamma|}\mathrm{tr}\Big(\sum_{\gamma=1}^{|\Gamma|}\sum_{\mathbf{x}^{\nu(n)}\in\mathbf{X}^{\nu(n)}}\sum_{\mathbf{y}^{\nu(n)}\in\mathbf{Y}^{\nu(n)}} p(\mathbf{x}^{\nu(n)},\mathbf{y}^{\nu(n)})$$

$$\cdot\Big[\breve{\kappa}_{y^n}\otimes W_{t^{\nu(n)}}(f_\gamma^{\nu(n)}(\mathbf{x}^{\nu(n)}))\Big]D_\gamma^{\nu(n)}\Big)$$

$$= \min_{t^{\nu(n)}\in\theta^{\nu(n)}} \frac{1}{|\Gamma|}\sum_{\gamma=1}^{|\Gamma|}\sum_{\mathbf{x}^{\nu(n)}\in\mathbf{X}^{\nu(n)}}\sum_{\mathbf{y}^{\nu(n)}\in\mathbf{Y}^{\nu(n)}} p(\mathbf{x}^{\nu(n)},\mathbf{y}^{\nu(n)}) \qquad \cdot\mathrm{tr}\Big(W_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma})$$

$$D^{\nu(n)}_{(\mathbf{y}^{\nu(n)}),\gamma}\Big), \tag{216}$$

where for every $\gamma\in\Gamma$, $\mathbf{x}^{\nu(n)}=(\mathbf{x}_1,\cdots,\mathbf{x}_{\nu(n)})\in\mathbf{X}^{\nu(n)}$, and $\mathbf{y}^{\nu(n)}=(\mathbf{y}_1,\cdots,\mathbf{y}_{\nu(n)})\in\mathbf{Y}^{\nu(n)}$, we set $p(\mathbf{x}^{\nu(n)},\mathbf{y}^{\nu(n)})=\prod_{i,j}p(\mathbf{x}_i,\mathbf{y}_j)$,

$$c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}:=f_\gamma^{\nu(n)}(\mathbf{x}^{\nu(n)})\in\mathbf{A}^{\nu(n)},$$

and

$$D^{\nu(n)}_{(\mathbf{y}^{\nu(n)}),\gamma}:=\mathrm{tr}_{H_{\mathbf{Y}^{\nu(n)}}}\Big((\breve{\kappa}_{y^{\nu(n)}}\otimes\mathrm{id}_{H^{\otimes\nu(n)}})D_\gamma^{\nu(n)}\Big).$$

The last equation of (216) holds because

$$\mathrm{tr}\Big(W_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma})\mathrm{tr}_{H_{\mathbf{Y}^{\nu(n)}}}\Big((\breve{\kappa}_{y^{\nu(n)}}\otimes\mathrm{id}_{H^{\otimes\nu(n)}})D_\gamma^{\nu(n)}\Big)\Big)$$
$$=\mathrm{tr}\Big(\Big[\mathrm{id}_{H_{\mathbf{Y}^{\nu(n)}}}\otimes W_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma})\Big]\Big[\breve{\kappa}_{y^{\nu(n)}}\otimes\mathrm{id}_{H^{\otimes\nu(n)}}\Big]D_\gamma^{\nu(n)}\Big)$$
$$=\mathrm{tr}\Big(\Big[\breve{\kappa}_{y^{\nu(n)}}\otimes W_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma})\Big]D_\gamma^{\nu(n)}\Big).$$

Since $\sum_{\gamma=1}^{|\Gamma|}D^{\nu(n)}_{(\mathbf{y}^{\nu(n)}),\gamma}=\sum_{\gamma=1}^{|\Gamma|}\mathrm{tr}_{H_{\mathbf{Y}^{\nu(n)}}}\Big((\breve{\kappa}_{y^{\nu(n)}}\otimes\mathrm{id}_{H^{\otimes\nu(n)}})D_\gamma^{\nu(n)}\Big)=\mathrm{tr}_{H_{\mathbf{Y}^{\nu(n)}}}\Big((\breve{\kappa}_{y^{\nu(n)}}\otimes\mathrm{id}_{H^{\otimes\nu(n)}})\sum_{\gamma=1}^{|\Gamma|}D_\gamma^{\nu(n)}\Big)=\mathrm{id}_{H^{\otimes\nu(n)}}$, we can define an $(X,Y)$-correlation assisted $(\nu(n),|\Gamma|)$ code (this is a code with deterministic encoder) by $\Big(\Big(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}\Big)_{\gamma\in\{1,\cdots,|\Gamma|\}}, \{D^{\nu(n)}_{(\mathbf{y}^{\nu(n)}),\gamma}:\gamma\in\{1,\cdots,|\Gamma|\}\}\Big)$.

Now we can construct an $(X,Y)$-correlation assisted $(\nu(n)+n,J_n)$ code $\mathcal{C}(X,Y)=\Big\{\Big(E_{\mathbf{x}^{\nu(n)+n}},\{D_j^{\mathbf{y}^{\nu(n)+n}}:j\in\{1,\cdots,J_n\}\}\Big):\mathbf{x}^{\nu(n)+n}\in\mathbf{X}^{\nu(n)+n},\mathbf{y}^{\nu(n)+n}\in\mathbf{Y}^{\nu(n)+n}\Big\}$, where for $\mathbf{x}^{\nu(n)+n}=(\mathbf{x}^{\nu(n)},\mathbf{x}^n)$, $\mathbf{y}^{\nu(n)+n}=(\mathbf{y}^{\nu(n)},\mathbf{y}^n)$ and $a^{\nu(n)+n}=(a^{\nu(n)},a^n)\in\mathbf{A}^{\nu(n)+n}$

$$E_{\mathbf{x}^{\nu(n)+n}}(a^{\nu(n)+n}|j)=\begin{cases}\frac{1}{|\Gamma|}E_\gamma(a^n|j)\text{ if }a^{\nu(n)}=c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma};\\ 0\text{ else}\end{cases},$$

and

$$D_j^{\mathbf{y}^{\nu(n)+n}} := \sum_{\gamma=1}^{|\Gamma|} D^{\nu(n)}_{(\mathbf{y}^{\nu(n)}),\gamma} \otimes D^n_{\gamma,j} \, .$$

For any $\gamma \in \{1, \cdots, |\Gamma|\}$ let

$$\begin{aligned} &\mathfrak{Z}_{\gamma, t^{\nu(n)+n}, \mathbf{x}^{\nu(n)+n}} \\ &:= \left\{ V_{t^{\nu(n)}} \left( c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma} \right) \otimes V_{t^n} \left( E_\gamma(a^n|1) \right), \cdots , \right. \\ &\left. V_{t^{\nu(n)}} \left( c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma} \right) \otimes V_{t^n} \left( E_\gamma(a^n|J_n) \right) \right\} . \end{aligned}$$

Similar to (116), for any $\mathbf{x}^{\nu(n)+n} \in \mathbf{X}^{\nu(n)+n}$, $\gamma \in \Gamma$, and $t^{\nu(n)+n} = (t^{\nu(n)} t^n)$ we have

$$\chi \left( R_{uni}, \mathfrak{Z}_{\gamma, t^{\nu(n)+n}, \mathbf{x}^{\nu(n)+n}} \right) = \chi \left( R_{uni}, Z_{\mathcal{C}^\gamma, t^n} \right) \, . \tag{217}$$

By definition we have

$$\begin{aligned} &Z_{t^{\nu(n)+n}, \mathbf{x}^{\nu(n)+n}} := \\ &\left\{ \frac{1}{|\Gamma|} \sum_{\gamma=1}^{|\Gamma|} V_{t^{\nu(n)}} (c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) \otimes V_{t^n} (E_\gamma( \, |1)), \cdots , \right. \\ &\left. \frac{1}{|\Gamma|} \sum_{i=1}^{|\Gamma|} V_{t^{\nu(n)}} (c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) \otimes V_{t^n} (E_\gamma( \, |J_n)) \right\} . \end{aligned}$$

Similar to (117) let $\lambda := \max\{2\epsilon, 2\zeta\}$, for any $t^{\nu(n)+n} = (t^{\nu(n)}, t^n)$, $\mathbf{x}^{\nu(n)+n} = (\mathbf{x}^{\nu(n)}, \mathbf{x}^n)$ and $\mathbf{y}^{\nu(n)+n} = (\mathbf{y}^{\nu(n)}, \mathbf{y}^n)$ we have

$$\begin{aligned} &\sum_{\mathbf{x}^{\nu(n)+n} \in \mathbf{X}^{\nu(n)+n}} \sum_{\mathbf{y}^{\nu(n)+n} \in \mathbf{Y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}) \chi \left( R_{uni}, Z_{t^{\nu(n)+n}, \mathbf{x}^{\nu(n)+n}} \right) \\ &\leq \sum_{\mathbf{x}^{\nu(n)+n}} \sum_{\mathbf{y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}) \chi \left( R_{uni}, Z_{t^{\nu(n)+n}, \mathbf{x}^{\nu(n)+n}} \right) \\ &- \sum_{\mathbf{x}^{\nu(n)+n}} \sum_{\mathbf{y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}) \frac{1}{|\Gamma|} \sum_{\gamma=1}^{|\Gamma|} \chi \left( R_{uni}, Z_{\mathcal{C}^\gamma, t^n} \right) + \lambda \\ &= \sum_{\mathbf{x}^{\nu(n)+n}} \sum_{\mathbf{y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}) \chi \left( R_{uni}, Z_{t^{\nu(n)+n}, \mathbf{x}^{\nu(n)+n}} \right) \\ &- \sum_{\mathbf{x}^{\nu(n)+n}} \sum_{\mathbf{y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}) \frac{1}{|\Gamma|} \sum_{\gamma=1}^{|\Gamma|} \chi \left( R_{uni}, \mathfrak{Z}_{\mathcal{C}^\gamma, t^{\nu(n)+n}, \mathbf{x}^{\nu(n)+n}} \right) + \lambda \\ &= \sum_{\mathbf{x}^{\nu(n)+n}} \sum_{\mathbf{y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}) \left[ S \left( \frac{1}{J_n} \frac{1}{|\Gamma|} \sum_{j=1}^{J_n} \sum_{i=1}^{|\Gamma|} V_{t^{\nu(n)}} (c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) \otimes V_{t^n} (E_\gamma( \, |j)) \right) \right. \\ &\left. - \frac{1}{J_n} \sum_{j=1}^{J_n} S \left( \frac{1}{|\Gamma|} \sum_{i=1}^{|\Gamma|} V_{t^{\nu(n)}} (c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) \otimes V_{t^n} (E_\gamma( \, |j)) \right) \right. \end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{|\Gamma|}\sum_{i=1}^{|\Gamma|} S\left(\frac{1}{J_n}\sum_{j=1}^{J_n} V_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) \otimes V_{t^n}(E_\gamma(\ |j))\right) \\
&+ \frac{1}{J_n}\frac{1}{|\Gamma|}\sum_{j=1}^{J_n}\sum_{i=1}^{|\Gamma|} S\left(V_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) \otimes V_{t^n}(E_\gamma(\ |j))\right)\Bigg] + \lambda \\
&\leq \lambda\,. \qquad (218)
\end{aligned}
$$

By (216), for any $t^{\nu(n)+n} \in \theta^{\nu(n)+n}$,

$$
\begin{aligned}
&\sum_{\mathbf{x}^{\nu(n)+n}}\sum_{\mathbf{y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n},\mathbf{y}^{\nu(n)+n}) P_e(\mathcal{C}(\mathbf{x}^{\nu(n)+n},\mathbf{y}^{\nu(n)+n}), t^{\nu(n)+n}) \\
&= 1 - \sum_{\mathbf{x}^{\nu(n)+n}}\sum_{\mathbf{y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n},\mathbf{y}^{\nu(n)+n}) \frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}\Bigg(\Bigg[\frac{1}{|\Gamma|}\sum_{\gamma=1}^{|\Gamma|} \\
&V_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) \otimes V_{t^n}(E_\gamma(\ |j))\Bigg] \cdot \left[\sum_{\gamma=1}^{|\Gamma|} D^{\nu(n)}_{(\mathbf{y}^{\nu(n)}),\gamma} \otimes D^n_{\gamma,j}\right]\Bigg) \\
&\leq 1 - \sum_{\mathbf{x}^{\nu(n)+n}}\sum_{\mathbf{y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n},\mathbf{y}^{\nu(n)+n}) \frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}\Bigg(\frac{1}{|\Gamma|}\sum_{\gamma=1}^{|\Gamma|} \\
&\left[V_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) \otimes V_{t^n}(E_\gamma(\ |j))\right] \cdot \left[D^{\nu(n)}_{(\mathbf{y}^{\nu(n)}),\gamma} \otimes D^n_{\gamma,j}\right]\Bigg) \\
&= 1 - \sum_{\mathbf{x}^{\nu(n)}}\sum_{\mathbf{y}^{\nu(n)}} p(\mathbf{x}^{\nu(n)},\mathbf{y}^{\nu(n)}) \frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}\Bigg(\frac{1}{|\Gamma|}\sum_{\gamma=1}^{|\Gamma|} \\
&\left[V_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) D^{\nu(n)}_{(\mathbf{y}^{\nu(n)}),\gamma}\right] \otimes \left[V_{t^n}(E_\gamma(\ |j)) D^n_{\gamma,j}\right]\Bigg) \\
&= 1 - \sum_{\mathbf{x}^{\nu(n)}}\sum_{\mathbf{y}^{\nu(n)}} p(\mathbf{x}^{\nu(n)},\mathbf{y}^{\nu(n)}) \frac{1}{|\Gamma|}\sum_{\gamma=1}^{|\Gamma|} \mathrm{tr}\left(V_{t^{\nu(n)}}(c^{\nu(n)}_{\mathbf{x}^{\nu(n)},\gamma}) D^{\nu(n)}_{(\mathbf{y}^{\nu(n)}),\gamma}\right) \\
&\cdot \left(\frac{1}{J_n}\sum_{j=1}^{J_n} \mathrm{tr}(V_{t^n}(E_\gamma(\ |j)) D^n_{\gamma,j})\right) \\
&\leq \lambda + \vartheta\,. \qquad (219)
\end{aligned}
$$

We now combine (219) and (218) and obtain the following result.

If $I(X,Y)$ and the $m-a-(X,Y)$ secrecy capacity of $\{(W_t,V_t) : t \in \theta\}$ are positive, we define $\lambda := \max\{2\epsilon, 2\zeta\} + \vartheta$ and the following statement is valid. For any $n \in \mathbb{N}$ and positive $\lambda$, if there is an $(n, J_n)$ randomness assisted code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$ for $\{(W_t,V_t) : t \in \theta\}$ such that

$$
\max_{t^n \in \theta^n} \int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < \epsilon\,,
$$

and

$$
\max_{t^n \in \theta^n} \int_\Lambda \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG(\gamma) < \zeta\,,
$$

then there is also a $(\nu(n)+n, J_n)$ common randomness assisted code $\mathcal{C}(X,Y) = \left\{ \left( E_{\mathbf{x}^{\nu(n)+n}}, D_j^{\mathbf{y}^{\nu(n)+n}} \right) : j \in \{1,\cdots,J_n\}, \mathbf{x}^{\nu(n)+n} \in \mathbf{X}^{\nu(n)+n} \mathbf{y}^{\nu(n)+n} \in \mathbf{Y}^{\nu(n)+n} \right\}$ such that

$$\max_{t^{\nu(n)+n} \in \theta^{\nu(n)+n}} \sum_{\mathbf{x}^{\nu(n)+n} \in \mathbf{X}^{\nu(n)+n}} \sum_{\mathbf{y}^{\nu(n)+n} \in \mathbf{Y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}) \cdot P_e(C(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}), t^{\nu(n)+n}) < \lambda \,, \tag{220}$$

and

$$\max_{t^{\nu(n)+n} \in \theta^{\nu(n)+n}} \chi\left(R_{uni}; Z_{t^{\nu(n)+n}, \mathbf{x}^{\nu(n)+n}} \mid X\right) < \lambda \,. \tag{221}$$

(220) and (221) mean that

$$C_s(\{(W_t,V_t) : t \in \theta\}; corr(X,Y)) \geq C_s(\{(W_t,V_t) : t \in \theta\}; r) - \frac{1}{n} \cdot \log J_n + \frac{1}{\nu(n)+n} \log J_n \,.$$

We know that $2^{\nu(n)}$ is in polynomial order of $n$. For any positive $\varepsilon$, if $n$ is large enough we have $\frac{1}{n} \log J_n - \frac{1}{\log n + n} \log J_n \leq \varepsilon$. Therefore, if $I(X,Y)$ and $C_s(\{(W_t,V_t) : t \in \theta\}; corr(X,Y))$ are both positive, we have

$$C_s(\{(W_t,V_t) : t \in \theta\}; corr(X,Y)) \geq C_s(\{(W_t,V_t) : t \in \theta\}; r) - \varepsilon \,.$$

This and the fact that

$$C_s(\{(W_t,V_t) : t \in \theta\}; corr(X,Y)) \leq C_s(\{(W_t,V_t) : t \in \theta\}; r) \,, \tag{222}$$

prove Theorem 4.14 for the case that $C_s((W_t,V_t)_{t\in\theta}; corr(X,Y))$ is positive.

*ii) When the Randomness Assisted Code Has Zero Secrecy Capacity*

If the randomness assisted secrecy capacity of $(W_t,V_t)_{t\in\theta}$ is equal to zero, with a similar technique as the techniques in [7] and [28] we show that the $(X,Y)$ correlation assisted secrecy capacity of $(W_t,V_t)_{t\in\theta}$ is also equal to zero.

Now we assume that the $m-a-(X,Y)$ secrecy capacity of $\{(W_t,V_t) : t \in \theta\}$ is equal to zero. If $C_s(\{(W_t,V_t) : t \in \theta\}; r)$ is also equal to zero, then there is nothing to prove. Thus let us assume that $C_s(\{(W_t,V_t) : t \in \theta\}; r)$ is positive.

Assume that there is an $(n, J_n)$ randomness assisted code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$ for $\{(W_t,V_t) : t \in \theta\}$ such that

$$\max_{t^n \in \theta^n} \int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < \lambda \,,$$

$$\max_{t^n \in \theta^n} \int_\Lambda \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG(\gamma) < \lambda \,.$$

We denote $\mathfrak{F}$ and the arbitrarily varying classical-quantum channel $(\tilde{U}_t)_{t\in\theta} : \mathfrak{F} \to \mathcal{S}(H^{n|\mathbf{Y}|})$ as above. If the deterministic capacity of $(\tilde{U}_t)_{t\in\theta}$ is positive, we can build, as above, a $(\nu(n)+n, J_n)$ common randomness assisted code $\mathcal{C}(X,Y) = \left\{ \left( E_{\mathbf{x}^{\nu(n)+n}}, \{D_j^{\mathbf{y}^{\nu(n)+n}} : j \in \{1,\cdots,J_n\}\} \right) : \mathbf{x}^{\nu(n)+n} \in \mathbf{X}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n} \in \mathbf{Y}^{\nu(n)+n} \right\}$ such that

$$\max_{t^{\nu(n)+n} \in \theta^{\nu(n)+n}} \sum_{\mathbf{x}^{\nu(n)+n} \in \mathbf{X}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n} \in \mathbf{Y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n})$$

$$\cdot P_e(C(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}), t^{\nu(n)+n}) < \epsilon \,,$$

$$\max_{t^{\nu(n)+n} \in \theta^{\nu(n)+n}} \sum_{\mathbf{y}^{\nu(n)+n} \in \mathbf{Y}^{\nu(n)+n}} p(\mathbf{x}^{\nu(n)+n}, \mathbf{y}^{\nu(n)+n}) \chi\left(R_{uni}; Z_{t^{\nu(n)+n}, \mathbf{x}^{\nu(n)+n}}\right) < \zeta \,.$$

But this would mean

$$C_s(\{(W_t, V_t) : t \in \theta\}; corr(X, Y)) = C_s(\{(W_t, V_t) : t \in \theta\}; r) \,,$$

and there is nothing to prove.

Thus we may assume that the the deterministic capacity of $(\tilde{U}_t)_{t\in\theta}$ is equal to zero. This implies that $(\tilde{U}_t)_{t\in\theta}$ is symmetrizable (cf. [6]), i.e. there is a parametrized set of distributions $\{\tau(\cdot \mid f) : f \in \mathfrak{F}\}$ on $\theta$ such that for all $f$, $f' \in \mathfrak{F}$ we have

$$\begin{aligned} &\sum_{t\in\theta} \tau(t \mid f') \sum_{\mathbf{x}} \sum_{\mathbf{y}} P(\mathbf{X} \times \mathbf{Y}) \breve{\kappa}_y \otimes W_t\left(f(\mathbf{x})\right) = \sum_{t\in\theta} \tau(t \mid f) \sum_{\mathbf{x}} \sum_{\mathbf{y}} P(\mathbf{X} \times \mathbf{Y}) \breve{\kappa}_y \otimes W_t\left(f'(\mathbf{x})\right) \\ &\Rightarrow \sum_{t\in\theta} \tau(t \mid f') \sum_{\mathbf{x}} P(\mathbf{X} \times \mathbf{Y}) W_t\left(f(\mathbf{x})\right) = \sum_{t\in\theta} \tau(t \mid f) \sum_{\mathbf{x}} P(\mathbf{X} \times \mathbf{Y}) W_t\left(f'(\mathbf{x})\right) \end{aligned} \tag{223}$$

for all $\mathbf{y} \in \mathbf{Y}$.

Our approach is similar to the technique of [7]. Let $\mathbf{A} = \{0, 1, \cdots, |\mathbf{A}| - 1\}$, $\mathbf{X} = \mathbf{Y} = \{0, 1\}$. We define functions $g^*$ and $g_i \in \mathfrak{F}$ for $i = 1, \cdots, a-1$ such that $g^*(0) = g^*(1) = 0$ and $g_i(u) := i + u \bmod |\mathbf{A}|$ for $u \in \{0, 1\}$. Since $(\tilde{U}_t)_{t\in\theta}$ is symmetrizable, by (223) there is a parametrized set of distributions $\{\tau(t \mid f) : f \in \mathfrak{F}\}$ on $\theta$ such that for all $a \in \mathbf{A}$, the following two equalities are valid

$$\begin{aligned} &\sum_{t\in\theta} p(0,0)\tau(t \mid g^*) W_t(a) \\ &+ \sum_{t\in\theta} p(1,0)\tau(t \mid g^*) W_t(a + 1 \bmod |\mathbf{A}|) \\ &= \sum_{t\in\theta} p(0,0)\tau(t \mid g_i) W_t(a) \\ &+ \sum_{t\in\theta} p(1,0)\tau(t \mid g_i) W_t(a) \\ &= \sum_{t\in\theta} \tau(t \mid g_i) W_t(a) \,; \end{aligned}$$

$$\begin{aligned} &\sum_{t\in\theta} p(0,1)\tau(t \mid g^*) W_t(a) \\ &+ \sum_{t\in\theta} p(1,1)\tau(t \mid g^*) W_t(a + 1 \bmod |\mathbf{A}|) \\ &= \sum_{t\in\theta} p(0,1)\tau(t \mid g_i) W_t(a) \\ &+ \sum_{t\in\theta} p(1,1)\tau(t \mid g_i) W_t(a) \end{aligned}$$

$$= \sum_{t\in\theta} \tau(t \mid g_i) W_t(a) \, .$$

If we choose an arbitrary orthonormal basis on $H$ to write the following quantum states in form of matrices

$$(m_{k,l})_{k,l=1,\cdots,\dim H} = \sum_{t\in\theta} \tau(t \mid g^*) W_t(a) \, ,$$

$$(m'_{k,l})_{k,l=1,\cdots,\dim H} = \sum_{t\in\theta} \tau(t \mid g^*) W_t(a+1 \bmod |\mathbf{A}|) \, ,$$

$$(m*_{k,l})_{k,l=1,\cdots,\dim H} = \sum_{t\in\theta} \tau(t \mid g_i) W_t(a) \, ,$$

for all $k, l \in \{1, \cdots, \dim H\}$ we have

$$p(0,0) m_{k,l} + p(1,0) m'_{k,l} = m*_{k,l} \, ,$$

$$p(0,1) m_{k,l} + p(1,1) m'_{k,l} = m*_{k,l} \, .$$

Since $I(X,Y)$ is positive, $p(0,0) \neq p(1,0)$ and $p(0,1) \neq p(1,1)$, therefore $\det \begin{pmatrix} p(0,0) & p(1,0) \\ p(0,1) & p(1,1) \end{pmatrix} \neq 0$. Thus $m_{k,l} = m'_{k,l} = m*_{k,l}$ for all $k, l \in \{1, \cdots, \dim H\}$, this means

$$\sum_{t\in\theta} \tau(t \mid g^*) W_t(a) = \sum_{t\in\theta} \tau(t \mid g^*) W_t(a+1 \bmod |\mathbf{A}|)$$

for all $a \in \mathbf{A}$.

Therefore, for any $n \in \mathbb{N}$ and any given $(n, J_n)$ code $\mathcal{C}^\gamma = \left(E^\gamma, \{D_j^\gamma : j = 1, \cdots, J_n\}\right)$, the following statement is valid. Let $a'^n$ be an arbitrary sequence in $\mathbf{A}^n$, we have

$$\begin{aligned}
&\sum_{t^n\in\theta^n} \tau(t^n \mid g^*) P_e(\mathcal{C}^\gamma, t^n) \\
&= \sum_{t^n\in\theta^n} \tau(t^n \mid g^*) \left[1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \mathrm{tr}(W_{t^n}(E_\gamma(\,|j)) D_j^\gamma)\right] \\
&= \sum_{t^n\in\theta^n} \tau(t^n \mid g^*) \left[1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{a^n\in\mathbf{A}^n} E^\gamma(a^n|j) \mathrm{tr}(W_{t^n}(a^n) D_j^\gamma)\right] \\
&= 1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{a^n\in\mathbf{A}^n} E^\gamma(a^n|j) \mathrm{tr}\left(\sum_{t^n\in\theta^n} \tau(t^n \mid g^*) W_{t^n}(a^n) D_j^\gamma\right) \\
&= 1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \sum_{a^n\in\mathbf{A}^n} E^\gamma(a^n|j) \mathrm{tr}\left(\sum_{t^n\in\theta^n} \tau(t^n \mid g^*) W_{t^n}(a'^n) D_j^\gamma\right) \\
&= 1 - \frac{1}{J_n} \sum_{j=1}^{J_n} \mathrm{tr}\left(\sum_{t^n\in\theta^n} \tau(t^n \mid g^*) W_{t^n}(a'^n) D_j^\gamma\right)
\end{aligned}$$

$$= 1 - \frac{1}{J_n} \sum_{t^n \in \theta^n} \tau(t^n \mid g^*) \mathrm{tr} \left( W_{t^n}(a'^n) \sum_{j=1}^{J_n} D_j^\gamma \right)$$
$$= 1 - \frac{1}{J_n} \sum_{t^n \in \theta^n} \tau(t^n \mid g^*) \mathrm{tr} \left( W_{t^n}(a'^n) \right)$$
$$= 1 - \frac{1}{J_n} , \tag{224}$$

where $\tau(t^n \mid g^*) := \tau(t_1 \mid g^*)\tau(t_2 \mid g^*) \cdots \tau(t_n \mid g^*)$ for $t^n = (t_1, t_2, \cdots, t_n)$. The second and the fifth equations hold because the trace function and matrices' multiplication are linear. The first, the fourth, and the last equations hold because $\sum_{t^n \in \theta^n} \tau(t^n \mid g^*) = \sum_{a^n \in \mathbf{A}^n} E^\gamma(a^n|j) = \mathrm{tr}\left(W_{t^n}(a'^n)\right) = 1$ for all $g^*$, $j$, and $a'^n$. The sixth equation holds because $\sum_{j=1}^{J_n} D_j^\gamma = id$.

Thus for any $n \in \mathbb{N}$, any $J_n \in \mathbb{N} \setminus \{1\}$, and any $(n, J_n)$ randomness assisted quantum code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$ we have

$$\frac{J_n - 1}{J_n}$$
$$= \int_\Lambda \sum_{t^n \in \theta^n} \tau(t^n \mid g^*) P_e(\mathcal{C}^\gamma, t^n) dG(\gamma)$$
$$= \sum_{t^n \in \theta^n} \tau(t^n \mid g^*) \int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma)$$
$$= \mathbb{E}\left( \int_\Lambda P_e(\mathcal{C}^\gamma, \mathfrak{T}^n) dG(\gamma) \right) , \tag{225}$$

where $\mathfrak{T}^n$ is a random variable on $\theta^n$ such that $Pr(\mathfrak{T}^n = t^n) = \tau(t^n \mid g^*)$ for all $t^n \in \theta^n$.

By (225) for any $n \in \mathbb{N}$, any $J_n \in \mathbb{N} \setminus \{1\}$ and any $(n, J_n)$ random assisted quantum code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$, there exists at least one $t^n \in \theta^n$ such that

$$\int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) \geq \frac{J_n - 1}{J_n} . \tag{226}$$

By (226) for any $n \in \mathbb{N}$, any $J_n > 1$, there is no $(n, J_n)$ randomness assisted code $(\{\mathcal{C}^\gamma : \gamma \in \Lambda\}, G)$ for $\{(W_t, V_t) : t \in \theta\}$ such that

$$\max_{t^n \in \theta^n} \int_\Lambda P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < \frac{1}{2} ,$$

therefore if the $m - a - (X, Y)$ secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ is equal to zero and $I(X, Y)$ is positive, the randomness assisted secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ is equal to $\log 1 = 0$. But this is a contradiction to our assumption that $C_s(\{(W_t, V_t) : t \in \theta\}; r)$ is positive.

This result and the result for the case when the $m - a - (X, Y)$ secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ is positive complete our proof for Theorem 4.14. □

Theorem 4.14 shows that the correlation is a very helpful resource for the secure message transmission through an arbitrarily varying classical-quantum wiretap channel. As Example 4.23 shows, there are indeed arbitrarily varying classical-quantum

wiretap channels which have zero deterministic secrecy capacity, but at the same time positive random secrecy capacity. Theorem 4.14 shows that if we have a $m-a-(X,Y)$ correlation as a resource, even when it is insecure and very weak (i.e. $I(X,Y)$ needs only to be slightly larger than zero), these channels will have a positive $m-a-(X,Y)$ secrecy capacity.

### 4.4.2 Further Notes on Resources

In Section 4.6.1 we gave an example when the deterministic capacity of an arbitrarily varying classical-quantum wiretap channel is not equal to its randomness-assisted capacity. Thus having resources is very helpful for achieving a positive secrecy capacity. For the proofs in Section 4.2 and Section 4.4.1 we did not allow the jammer to have access to the shared randomness.

Now we consider the case when the shared randomness is not secure, i.e. when the jammer can have access to the shared randomness (cf. Figure 3).

**Corollary 4.15.** *Let $\{(W_t,V_t) : t \in \theta\}$ be an arbitrarily varying classical-quantum wiretap channel. We have*

$$C_s(\{(W_t,V_t) : t \in \theta\}) = C_s(\{(W_t,V_t) : t \in \theta\}; r_{ns}) \tag{227}$$

*Proof.* Let $\mathcal{C} = \left(E, \{D_j^n : j = 1, \cdots J_n\}\right)$ be an $(n, J_n)$ code such

$$\max_{t^n \in \theta^n} P_e(\mathcal{C}, t^n) < \epsilon \,,$$

and

$$\max_{t^n \in \theta^n} \chi\left(R_{uni}; Z_{t^n}\right) < \zeta \,.$$

We define a $G'$ such that $G'(\{\mathcal{C}\}) = 1$, it holds

$$\int_\Lambda \max_{t^n \in \theta^n} P_e(\mathcal{C}^\gamma, t^n) dG'(\gamma) < \epsilon \,,$$

$$\int_\Lambda \max_{t^n \in \theta^n} \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG'(\gamma) < \zeta \,.$$

Thus every achievable secrecy rate for $\{(W_t,V_t) : t \in \theta\}$ is also an achievable secrecy rate for $\{(W_t,V_t) : t \in \theta\}$ under non-secure randomness assisted coding.

Now we assume that there is a $G''$ such that

$$\int_\Lambda \max_{t^n \in \theta^n} P_e(\mathcal{C}^\gamma, t^n) dG''(\gamma) < \epsilon \,,$$

$$\int_\Lambda \max_{t^n \in \theta^n} \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG''(\gamma) < \zeta \,.$$

Then for any $s^n \in \theta^n$ we have

$$\int_\Lambda P_e(\mathcal{C}^\gamma, s^n) dG''(\gamma) \leq \int_\Lambda \max_{t^n \in \theta^n} P_e(\mathcal{C}^\gamma, t^n) dG''(\gamma) < \epsilon \,,$$

$$\int_\Lambda \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, s^n}\right) dG''(\gamma) \leq \int_\Lambda \max_{t^n \in \theta^n} \chi\left(R_{uni}, Z_{\mathcal{C}^\gamma, t^n}\right) dG''(\gamma) < \zeta \,.$$

Thus every achievable secrecy rate for $\{(W_t, V_t) : t \in \theta\}$ under non-secure randomness assisted coding is also an achievable secrecy rate for $\{(W_t, V_t) : t \in \theta\}$ under randomness assisted coding.

Therefore,

$$C_s(\{(W_t, V_t) : t \in \theta\}) \leq C_s(\{(W_t, V_t) : t \in \theta\}; r_{ns}) \leq C_s(\{(W_t, V_t) : t \in \theta\}; r) \,. \tag{228}$$

At first let us assume that $\{W_t : t \in \theta\}$ is not symmetrizable. By Theorem 4.1 when $\{W_t : t \in \theta\}$ is not symmetrizable it holds $C_s(\{(W_t, V_t) : t \in \theta\}) = C_s(\{(W_t, V_t) : t \in \theta\}; r)$. Thus when $\{W_t : t \in \theta\}$ is not symmetrizable we have

$$C_s(\{(W_t, V_t) : t \in \theta\}) = C_s(\{(W_t, V_t) : t \in \theta\}; r_{ns}) \,.$$

Now let us assume that $\{W_t : t \in \theta\}$ is symmetrizable. When $\{W_t : t \in \theta\}$ is symmetrizable and $J_n > 1$ holds then by Section 4.1, for any $(n, J_n)$ code $\mathcal{C}$ there is a $t^n \in \theta^n$ and a positive $c$ such that

$$P_e(\mathcal{C}, t^n) > c \,.$$

Thus, when $\{W_t : t \in \theta\}$ is symmetrizable, for any $G$ we have

$$\int_\Lambda \max_{t^n \in \theta^n} P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) > c \,,$$

which implies we can only have $\int_\Lambda \max_{t^n \in \theta^n} P_e(\mathcal{C}^\gamma, t^n) dG(\gamma) < c$ when $J_n$ is less or equal to 1. This means

$$C_s(\{(W_t, V_t) : t \in \theta\}; r_{ns}) = \log 1 = 0 \,.$$

By Theorem 4.1 when $\{W_t : t \in \theta\}$ is symmetrizable it holds $C_s(\{(W_t, V_t) : t \in \theta\}) = 0$ and therefore when $\{W_t : t \in \theta\}$ is symmetrizable we have

$$C_s(\{(W_t, V_t) : t \in \theta\}) = C_s(\{(W_t, V_t) : t \in \theta\}; r_{ns}) \,. \qquad \square$$

Theorem 4.1 shows that an arbitrarily varying classical-quantum channel with zero deterministic secrecy capacity allows secure transmission if the sender and the legal receiver has the possibility to use shared randomness, as long as the shared randomness is kept secret against the jammer. Corollary 4.15 shows that when the jammer is able have access to the outcomes of the shared random experiment we can only achieve the rate as when we do not use any shared randomness at all. This means the shared randomness will be completely useless when it is known by the jammer.

Applying Theorem 4.12 we can now determine the random assisted secrecy capacity with the strongest code concept for shared randomness, i.e., the randomness which is secure against both the jammer and eavesdropping (cf. Figure 4).

**Corollary 4.16.** *Let $\theta := \{1, \cdots, T\}$ be a finite index set. Let $\{(W_t, V_t) : t \in \theta\}$ be an arbitrarily varying classical-quantum wiretap channel.*

*When $\{W_t : t \in \theta\}$ is not symmetrizable, we have*

$C_{key}(\{(W_t, V_t) : t \in \theta\}; g_n)$

$$= \min\Bigg( \lim_{n\to\infty} \frac{1}{n} \max_{U\to A^n \to \{B_q^{\otimes n}, Z_{t^n}: q, t_n\}} \left( \inf_{B_q \in Conv((B_t)_{t\in\theta})} \chi(p_U; B_q^{\otimes n}) - \max_{t^n \in \theta^n} \chi(p_U; Z_{t^n}) \right) + g_n,$$

$$\max_{U\to A^n \to \{B_q^{\otimes n}: q\}} \inf_{B_q \in Conv((B_t)_{t\in\theta})} \chi(p_U; B_q^{\otimes n}) \Bigg). \tag{229}$$

*Here we use the strong code concept.*

**Remark 4.17.** *When $g_n$ is positive and independent of $n$, (229) always holds and we do not have to assume that $\{W_t : t \in \theta\}$ is not symmetrizable.*

*Proof.* We define $\Gamma'_n := \{1, \cdots, \lceil \frac{n^3}{|\Gamma_n|} \rceil\}$. It holds $n! > |\Gamma'_n \times \Gamma_n| \geq n^3$. Notice that when $g_n$ is positive and independent of $n$ we always have have $n! \geq 2^{ng_n} \geq n^3$ for sufficiently large $n$ and thus $\Gamma'_n := \{1\}$.

We fix a probability distribution $p \in P(\mathbf{A})$. Let

$$J_n = \min\left( \lfloor 2^{n \min_{s\in\overline{\theta}} \chi(p;B_s) - \log L_n + n g_n - 2n\mu} \rfloor, \lfloor 2^{n \min_{s\in\overline{\theta}} \chi(p;B_s) - 2n\mu} \rfloor \right),$$

$$L_n = \max\left( \lceil 2^{\max_{t^n} \chi(p;Z_{t^n}) - n g_n + 2n\zeta} \rceil, 1 \right),$$

and $p'(x^n) := \begin{cases} \frac{p^n(x^n)}{p^n(\mathcal{T}^n_{p,\delta})}, & \text{if } x^n \in \mathcal{T}^n_{p,\delta}; \\ 0, & \text{else}. \end{cases}$ Let $X^n := \{X_{j,l} : j \in \{1, \ldots, J_n\}, l \in \{1, \ldots, L_n\}\}$ be a family of random variables taking value according to $p'$.

It holds $J_n L_n < 2^{n \min_{s\in\overline{\theta}} \chi(p;B_s)}$ and $L^n 2^{g_n} > 2^{\max_{t^n} \chi(p;Z_{t^n})}$. Similar to the proof of Theorem 4.2 and the proof of Theorem 4.14, with a positive probability there is a realization $\{x_{j,l} : j, l\}$ of $\{X_{j,l,} : j, l\}$ and a set $\{\pi_\gamma : \gamma \in \Gamma'_n \times \Gamma_n\} \subset \mathsf{S}_n$ with the following properties:

There exits a set of decoding operators $\{D_{j,l} : j = 1, \cdots, J_n, l = 1, \cdots, L_n, \}$ such that for every $t^n \in \theta^n$ $\epsilon > 0$, $\zeta > 0$, and sufficiently large $n$,

$$1 - \frac{1}{J_n} \frac{1}{L_n} \frac{1}{|\Gamma'_n \times \Gamma_n|} \sum_{j=1}^{J_n} \sum_{j=1}^{L_n} \sum_{\gamma=1}^{|\Gamma'_n \times \Gamma_n|} \mathrm{tr}\left( W_{t^n}(\pi_\gamma^{-1}(x_{j,l})) P^\dagger_{\pi_\gamma} D_{j,l} P_{\pi_\gamma} \right) < \epsilon$$

and

$$\chi\left( R_{uni}, \frac{1}{J_n} \frac{1}{L_n} \frac{1}{|\Gamma'_n \times \Gamma_n|} \sum_{j=1}^{J_n} \sum_{j=1}^{L_n} \sum_{\gamma=1}^{|\Gamma'_n \times \Gamma_n|} V_{t^n}(\pi_\gamma^{-1}(x_{j,l})) \right) < \zeta.$$

When $|\Gamma'_n| > 1$ holds, we use the strategy of Theorem 4.12 by building a two-part secure code word, the first part is used to send $\gamma' \in \Gamma'_n$, the second is used to transmit the message to the legal receiver.

Thus,

$$C_{key}(\{(W_t, V_t) : t \in \theta\}; g_n)$$
$$\geq \min\Bigg( \lim_{n\to\infty} \frac{1}{n} \max_p \left( \inf_{B_q \in Conv((B_t)_{t\in\theta})} \chi(p; B_q^{\otimes n}) - \max_{t^n \in \theta^n} \chi(p; Z_{t^n}) \right) + g_n,$$
$$\max_p \inf_{B_q \in Conv((B_t)_{t\in\theta})} \chi(p; B_q^{\otimes n}) \Bigg).$$

The achievability of $\lim_{n\to\infty}\frac{1}{n}\Big(\min_q \chi(p_U;B_q)-\max_{t^n}\chi(p_U;Z_{t^n})\Big)+g_n$ and $\inf_{B_q\in Conv((B_t)_{t\in\theta})}\chi(p_U;B_q^{\otimes n})$ is then shown via standard arguments.

Now we are going to prove the converse.

$$C_{key}(\{(W_t,V_t):t\in\theta\};g_n)\le \max_{U\to A^n\to\{B_q^{\otimes n}:q\}}\inf_{B_q\in Conv((B_t)_{t\in\theta})}\chi(p_U;B_q^{\otimes n}) \tag{230}$$

holds trivially.

Let $(E^{\gamma,(n)},\{D_j^{\gamma,(n)}:j\})$ be a sequence of $(n,J_n)$ code such that for every $t^n\in\theta^n$

$$1-\frac{1}{J_n}\frac{1}{2^{ng_n}}\sum_{j=1}^{J_n}\sum_{\gamma=1}^{2^{ng_n}}\mathrm{tr}\left(W_{t^n}(E^{\gamma,(n)}(j))D_j^{\gamma,(n)}\right)<\epsilon_n$$

and

$$\chi\left(R_{uni},\frac{1}{J_n}\frac{1}{2^{ng_n}}\sum_{j=1}^{J_n}\sum_{\gamma=1}^{2^{ng_n}}V_{t^n}(E^{\gamma,(n)}(j))\right)<\zeta_n\ ,$$

where $\lim_{n\to\infty}\epsilon_n=0$ and $\lim_{n\to\infty}\zeta_n=0$. It is known that for sufficiently large $n$ we have

$$\log J_n\le\frac{1}{2^{ng_n}}\sum_{\gamma=1}^{2^{ng_n}}\chi\left(R_{uni},B_q^{\gamma\otimes n}\right)-\chi\left(R_{uni},Z_{t^n}\right)\ . \tag{231}$$

Let $\psi_q^{j,\gamma\otimes n}:=W_q^{\otimes n}(E^{\gamma,(n)}(j))$. We denote $\tilde{B}_q^{j\otimes n}:=\{W_q^{\otimes n}(E^{\gamma,(n)}(j)):\gamma\in\Gamma_n\}$ and $\tilde{B}_q^{\otimes n}:=\{\frac{1}{J_n}W_q^{\otimes n}(E^{\gamma,(n)}(j)):\gamma\in\Gamma_n\}$. Let $G_{uni}$ be the uniformly distributed random variable with value in $\Gamma_n$.

We have

$$\begin{aligned}
&\frac{1}{2^{ng_n}}\sum_{\gamma=1}^{2^{ng_n}}\chi\left(R_{uni};B_q^{\gamma\otimes n}\right)-\chi\left(R_{uni};\frac{1}{2^{ng_n}}\sum_{\gamma=1}^{2^{ng_n}}B_q^{\gamma\otimes n}\right)\\
&=\frac{1}{2^{ng_n}}\sum_{\gamma=1}^{2^{ng_n}}S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\psi_q^{j,\gamma\otimes n}\right)-\frac{1}{2^{ng_n}}\frac{1}{J_n}\sum_{\gamma=1}^{2^{ng_n}}\sum_{j=1}^{J_n}S\left(\psi_q^{j,\gamma\otimes n}\right)\\
&-\left[S\left(\frac{1}{2^{ng_n}}\frac{1}{J_n}\sum_{\gamma=1}^{2^{ng_n}}\sum_{j=1}^{J_n}\psi_q^{j,\gamma\otimes n}\right)-\frac{1}{J_n}\sum_{j=1}^{J_n}S\left(\frac{1}{2^{ng_n}}\sum_{\gamma=1}^{2^{ng_n}}\psi_q^{j,\gamma\otimes n}\right)\right]\\
&=\frac{1}{2^{ng_n}}\sum_{\gamma=1}^{2^{ng_n}}S\left(\frac{1}{J_n}\sum_{j=1}^{J_n}\psi_q^{j,\gamma\otimes n}\right)-S\left(\frac{1}{2^{ng_n}}\frac{1}{J_n}\sum_{\gamma=1}^{2^{ng_n}}\sum_{j=1}^{J_n}\psi_q^{j,\gamma\otimes n}\right)\\
&-\left[\frac{1}{2^{ng_n}}\frac{1}{J_n}\sum_{\gamma=1}^{2^{ng_n}}\sum_{j=1}^{J_n}S\left(\psi_q^{j,\gamma\otimes n}\right)-\frac{1}{J_n}\sum_{j=1}^{J_n}S\left(\frac{1}{2^{ng_n}}\sum_{\gamma=1}^{2^{ng_n}}\psi_q^{j,\gamma\otimes n}\right)\right]\\
&=\frac{1}{J_n}\sum_{j=1}^{J_n}S\left(G_{uni},\tilde{B}_q^{j\otimes n}\right)-S\left(G_{uni},\tilde{B}_q^{\otimes n}\right)\\
&\le\frac{1}{J_n}\sum_{j=1}^{J_n}S\left(G_{uni},\tilde{B}_q^{j\otimes n}\right)
\end{aligned}$$

$$\leq \frac{1}{J_n}\sum_{j=1}^{J_n} H\left(G_{uni}\right)$$
$$= H\left(G_{uni}\right)$$
$$= ng_n\,. \tag{232}$$

By (230), (231), and (232) we have

$$C_{key}(\{(W_t,V_t):t\in\theta\};g_n)$$
$$\leq \lim_{n\to\infty}\frac{1}{n}\max_{U\to A^n\to\{B_q^{\otimes n},Z_{t^n}:q,t_n\}}\Bigg(\inf_{B_q\in Conv((B_t)_{t\in\theta})}\chi(p_U;B_q^{\otimes n})$$
$$-\max_{t^n\in\theta^n}\chi(p_U;Z_{t^n})\Bigg)+g_n\,. \qquad \square$$

## 4.5 Investigation of Secrecy Capacity's Continuity

In this section we show that the secrecy capacity of an arbitrarily varying classical-quantum wiretap channel under common randomness assisted quantum coding is continuous in the following sense:

**Corollary 4.18.** *For an arbitrarily varying classical-quantum wiretap channel $\{(W_t,V_t):t\in\theta\}$, where $W_t:P(\mathbf{A})\to\mathcal{S}(H)$ and $V_t:P(\mathbf{A})\to\mathcal{S}(H')$, and a positive $\delta$ let $\mathbf{C}_\delta$ be the set of all arbitrarily varying classical-quantum wiretap channels $\{(W'_t,V'_t):t\in\theta\}$, where $W'_t:P(\mathbf{A})\to\mathcal{S}(H)$ and $V'_t:P(\mathbf{A})\to\mathcal{S}(H')$, such that*

$$\max_{a\in\mathbf{A}}\|W_t(a)-W'_t(a)\|_1<\delta$$

*and*

$$\max_{a\in\mathbf{A}}\|V_t(a)-V'_t(a)\|_1<\delta$$

*for all $t\in\theta$.*

*For any positive $\epsilon$ there is a positive $\delta$ such that for all $\{(W'_t,V'_t):t\in\theta\}\in\mathbf{C}_\delta$ we have*

$$|C_s(\{(W_t,V_t):t\in\theta\};cr)-C_s(\{((W'_t,V'_t):t\in\theta\};cr)|\leq\epsilon\,. \tag{233}$$

*Proof.* By Theorem 4.11 the secrecy capacity of $\{(W_t,V_t):t\in\theta\}$ is

$$\lim_{n\to\infty}\frac{1}{n}\max_{U\to A^n\to\{B_q^{\otimes n},Z_{t^n}:q,t_n\}}\Big(\inf_{B_q\in Conv((B_t)_{t\in\theta})}\chi(p_U;B_q^{\otimes n})-\max_{t^n\in\theta^n}\chi(p_U;Z_{t^n})\Big)\,,$$

and for every $\{(W'_t,V'_t):t\in\theta\}\in\mathbf{C}_\delta$ the secrecy capacity of $\{(W'_t,V'_t):t\in\theta\}$ is

$$\lim_{n\to\infty}\frac{1}{n}\max_{U\to A^n\to\{B_q^{\otimes n},Z_{t^n}:q,t_n\}}\Big(\inf_{B'_q\in Conv((B'_t)_{t\in\theta})}\chi(p_U;B'^{\otimes n}_q)-\max_{t^n\in\theta^n}\chi(p_U;Z'_{t^n})\Big)\,,$$

where $B'_t$ is the resulting quantum state at the output of $W'_t$ and $Z'_t$ is the resulting quantum state at the output of $V'_t$.

To analyze $|\chi(p;Z_{t^n})-\chi(p;Z'_{t^n})|$ we use the technique introduced in [46] and apply the following lemma given in [9].

**Lemma 4.19 (Alicki-Fannes Inequality).** *Suppose we have a composite system $\mathfrak{PQ}$ with components $\mathfrak{P}$ and $\mathfrak{Q}$. Let $G^{\mathfrak{P}}$ and $G^{\mathfrak{Q}}$ be Hilbert space of $\mathfrak{P}$ and $\mathfrak{Q}$, respectively. Suppose we have two bipartite quantum states $\phi^{\mathfrak{PQ}}$ and $\sigma^{\mathfrak{PQ}}$ in $\mathcal{S}(G^{\mathfrak{PQ}})$ such that $\|\phi^{\mathfrak{PQ}} - \sigma^{\mathfrak{PQ}}\|_1 = \epsilon < 1$, it holds*

$$S(\mathfrak{P} \mid \mathfrak{Q})_\rho - S(\mathfrak{P} \mid \mathfrak{Q})_\sigma \leq 4\epsilon \log(d-1) - 2h(\epsilon)\,, \tag{234}$$

*where $d$ is the dimension of $G^{\mathfrak{P}}$ and $h(\epsilon)$ is defined as in Lemma 3.6.*

The difference to [9] is we consider here classical-quantum channels instead of quantum-quantum channels.

We fix an $n \in \mathbb{N}$ and a $t^n = (t_1, \cdots t_n) \in \theta^n$. For any $a^n \in \mathbf{A}^n$ we have

$$\begin{aligned}
&|S\left(V_{t^n}(a^n)\right) - S\left({V'}_{t^n}(a^n)\right)| \\
&= \left|\sum_{k=1}^{n} S\left(V_{(t_1,\cdots t_{k-1})} \otimes {V'}_{(t_k,\cdots t_n)}(a^n)\right) - S\left(V_{(t_1,\cdots t_k)} \otimes {V'}_{(t_{k+1},\cdots t_n)}(a^n)\right)\right| \\
&\leq \sum_{k=1}^{n} \left| S\left(V_{(t_1,\cdots t_{k-1})} \otimes {V'}_{(t_k,\cdots t_n)}(a^n)\right) - S\left(V_{(t_1,\cdots t_k)} \otimes {V'}_{(t_{k+1},\cdots t_n)}(a^n)\right)\right| .
\end{aligned}$$

For a $k \in \{1, \cdots, n\}$ and $a^n = (a_1, \cdots a_n) \in \mathbf{A}^n$ by Lemma 4.19 we have

$$\begin{aligned}
&\left| S\left(V_{(t_1,\cdots t_{k+1})} \otimes {V'}_{(t_k,\cdots t_n)}(a^n)\right) - S\left(V_{(t_1,\cdots t_{k+1})} \otimes {V'}_{(t_{k+1},\cdots t_n)}(a^n)\right)\right| \\
&= \Big| S\left(V_{(t_1,\cdots t_k)} \otimes {V'}_{(t_k,\cdots t_n)}(a^n)\right) - S\left(V_{(t_1,\cdots t_{k-1})} \otimes {V'}_{(t_{k+1},\cdots t_n)}((a_1,\cdots a_{k-1}, a_{k+1}, \cdots a_n))\right) \\
&- S\left(V_{(t_1,\cdots t_k)} \otimes {V'}_{(t_{k+1},\cdots t_n)}(a^n)\right) + S\left(V_{(t_1,\cdots t_{k-1})} \otimes {V'}_{(t_{k+1},\cdots t_n)}((a_1,\cdots a_{k-1}, a_{k+1}, \cdots a_n))\right)\Big| \\
&= \Big| S\left({V'}_{t_k}(a_k) \mid V_{(t_1,\cdots t_{k-1})} \otimes {V'}_{(t_{k+1},\cdots t_n)}((a_1,\cdots a_{k-1}, a_{k+1}, \cdots a_n))\right) \\
&- S\left(V_{t_k}(a_k) \mid V_{(t_1,\cdots t_{k-1})} \otimes {V'}_{(t_{k+1},\cdots t_n)}((a_1,\cdots a_{k-1}, a_{k+1}, \cdots a_n))\right)\Big| \\
&\leq 4\delta \log(d_E - 1) - 2 \cdot h(\delta)\,,
\end{aligned}$$

where $d_E$ is the dimension of $H^{\mathfrak{E}}$.

Thus

$$|S\left(V_{t^n}(a^n)\right) - S\left({V'}_{t^n}(a^n)\right)| \leq 4n\delta \log(d_E - 1) - 2n \cdot h(\delta)\,. \tag{235}$$

For any probability distribution $p \in P(\mathbf{A})$, $n \in \mathbb{N}$, and $t^n \in \theta^n$ we have

$$\begin{aligned}
&|\chi(p; Z_{t^n}) - \chi(p; {Z'}_{t^n})| \\
&= \Big| S(\sum_a p(a) V_{t^n}(a)) - \sum_a p(a) S(V_{t^n}(a)) \\
&- S(\sum_a p(a) {V'}_{t^n}(a)) + S(\sum_a p(a) {V'}_{t^n}(a))\Big| \\
&\leq \left| S(\sum_a p(a) V_{t^n}(a)) - S(\sum_a p(a) {V'}_{t^n}(a))\right|
\end{aligned}$$

$$+\Big|\sum_a p(a)S(V'^{t^n}(a)) - \sum_a p(a)S(V'^{t^n}(a))\Big|$$
$$\le 8n\delta\log(d_E-1) - 4n\cdot h(\delta)\,. \tag{236}$$

We fix a probability distribution $q$ on $\theta$, a probability distribution $p \in P(\mathbf{A})$, and an $n \in \mathbb{N}$. By Lemma 3.6 we have

$$\begin{aligned}
&|\chi(p;B_q) - \chi(p;B'_q)| \\
&= \Big|\sum_t q(t)S(\sum_a p(a)W_t(a)) - \sum_t\sum_a q(t)p(a)S(W_t(a)) \\
&- \sum_t q(t)S(\sum_a p(a)W'_t(a)) + S(\sum_t\sum_a q(t)p(a)W'_t(a))\Big| \\
&\le \Big|\sum_t q(t)S(\sum_a p(a)W_t(a)) - \sum_t q(t)S(\sum_a p(a)W'_t(a))\Big| \\
&+ \Big|\sum_t\sum_a q(t)p(a)S(W_t(a)) - S(\sum_t\sum_a q(t)p(a)W'_t(a))\Big| \\
&\le 8\delta\log(d_B-1) - 4\cdot h(\delta)\,,
\end{aligned} \tag{237}$$

where $d_B$ is the dimension of $H^{\mathfrak{B}}$.

Thus for any probability distribution $q$ on $\theta$, $n \in \mathbb{N}$, $p \in P(\mathbf{A})$, $t^n \in \theta^n$ we have for all $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$

$$\begin{aligned}
&\Big|(\chi(p;B_q) - \frac{1}{n}\chi(p;Z_{t^n})) - (\chi(p;B'_q) - \frac{1}{n}\chi(p;Z'^{t^n}))\Big| \\
&\le 8\delta\log(d_B-1) + 8\delta\log(d_E-1) - 8\cdot h(\delta)\,.
\end{aligned} \tag{238}$$

For any positive $\epsilon$ we can find a positive $\delta$ such that $8\delta\log(d_B-1) + 8\delta\log(d_E-1) - 8\cdot h(\delta) \le \epsilon$.

Thus for all $n \in \mathbb{N}$ and any positive $\epsilon$ we can find a positive $\delta$ such that for all $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$

$$\begin{aligned}
&\Big|(\max_p \inf_{B_q \in Conv((B_t)_{t\in\theta})} \chi(p;B_q) - \max_{t^n\in\theta^n}\chi(p;Z_{t^n})) \\
&- (\max_p \inf_{B'_q \in Conv((B'_t)_{t\in\theta})} \chi(p;B'_q) - \frac{1}{n}\max_{t^n\in\theta^n}\chi(p;Z'^{t^n}))\Big| \\
&\le \epsilon\,.
\end{aligned} \tag{239}$$

When $\inf_{q\in P(\theta)}\chi(p;B_q) - \max_{t^n\in\theta^n}\chi(p;Z_{t^n})$ achieves its maximum in $p$, and when $\inf_{q'\in P(\theta)}\chi(\acute{p};B'_{q'}) - \frac{1}{n}\max_{t^{n\prime}\in\theta^n}\chi(\acute{p};Z'^{t^{n\prime}})$ achieves its maximum in $\acute{p}$ a $\acute{p} \in P(A)$, the following inequality holds.

For all $n \in \mathbb{N}$ and any positive $\epsilon$ we can find a positive $\delta$ such that for all $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$

$$\begin{aligned}
&\Big|(\max_p \inf_{q\in P(\theta)} \chi(p;B_q) - \max_{\theta^n\in\Theta^n}\chi(p;Z_{t^n})) \\
&- (\max_{\acute{p}} \inf_{q'\in P(\theta)} \chi(\acute{p};B'_{q'}) - \frac{1}{n}\max_{t^{n\prime}\in\theta^n}\chi(\acute{p};Z'^{t^{n\prime}}))\Big|
\end{aligned}$$

$$\leq 3\epsilon\,, \tag{240}$$

since else we would have

$$\inf_{q\in P(\theta)} \chi(p;{B'}_q) - \frac{1}{n}\max_{t^n\in\theta^n}\chi(p;{Z'}_{t^n}) \quad > \inf_{q'\in P(\theta)} \chi(\acute{p};{B'}_{q'}) - \frac{1}{n}\max_{t^{n\prime}\in\theta^n}\chi(\acute{p};{Z'}_{t^{n\prime}})\,.$$

(240) shows Corollary 4.18. □

**Corollary 4.20.** *The deterministic secrecy capacity of an arbitrarily varying classical-quantum wiretap channel is in general not continuous.*

*Proof.* We show Corollary 4.20 by giving an example.

Let $\theta := \{1,2\}$. Let $\mathbf{A} = \{0,1\}$. Let $H^{\mathfrak{B}} = \mathbb{C}^5$. Let $\{|0\rangle^{\mathfrak{B}}, |1\rangle^{\mathfrak{B}}, |2\rangle^{\mathfrak{B}}, |3\rangle^{\mathfrak{B}}, |4\rangle^{\mathfrak{B}}\}$ be a set of orthonormal vectors on $H^{\mathfrak{B}}$. Let $\lambda$ be $\in [0,1]$.

For $r \in [0,1]$ let $P_r$ be the probability distribution on $\mathbf{A}$ such that $P_r(0) = r$ and $P_r(1) = 1-r$. We define a channel $W_1^\lambda : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{B}})$ by

$$W_1^\lambda(P_r) = (1-\lambda)r|0\rangle\langle 0|^{\mathfrak{B}} + (1-\lambda)(1-r)|1\rangle\langle 1|^{\mathfrak{B}} + \lambda|3\rangle\langle 3|^{\mathfrak{B}}\,,$$

and a channel $W_2^\lambda : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{B}})$ by

$$W_2^\lambda(P_r) = (1-\lambda)r|1\rangle\langle 1|^{\mathfrak{B}} + (1-\lambda)(1-r)|2\rangle\langle 2|^{\mathfrak{B}} + \lambda|4\rangle\langle 4|^{\mathfrak{B}}\,.$$

In other words:

$$W_1^\lambda(0) = (1-\lambda)|0\rangle\langle 0|^{\mathfrak{B}} + \lambda|3\rangle\langle 3|^{\mathfrak{B}}\,,$$
$$W_1^\lambda(1) = (1-\lambda)|1\rangle\langle 1|^{\mathfrak{B}} + \lambda|3\rangle\langle 3|^{\mathfrak{B}}\,,$$

$$W_2^\lambda(0) = (1-\lambda)|1\rangle\langle 1|^{\mathfrak{B}} + \lambda|4\rangle\langle 4|^{\mathfrak{B}}\,,$$
$$W_2^\lambda(1) = (1-\lambda)|2\rangle\langle 2|^{\mathfrak{B}} + \lambda|4\rangle\langle 4|^{\mathfrak{B}}\,.$$

Let $H^{\mathfrak{E}} = \mathbb{C}^5$. Let $\{|0\rangle^{\mathfrak{E}}, |1\rangle^{\mathfrak{E}}, |2\rangle^{\mathfrak{E}}, |3\rangle^{\mathfrak{E}}, |4\rangle^{\mathfrak{E}}$ be a set of orthonormal vectors on $H^{\mathfrak{E}}$.

We define a channel $V_1^\lambda : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{E}})$ by

$$V_1^\lambda(P_r) = \lambda r|0\rangle\langle 0|^{\mathfrak{E}} + \lambda(1-r)|1\rangle\langle 1|^{\mathfrak{E}} + (1-\lambda)|3\rangle\langle 3|^{\mathfrak{E}}\,,$$

and a channel $V_2^\lambda : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{E}})$ by

$$V_2^\lambda(P_r) = \lambda r|1\rangle\langle 1|^{\mathfrak{E}} + \lambda(1-r)|2\rangle\langle 2|^{\mathfrak{E}} + (1-\lambda)|4\rangle\langle 4|^{\mathfrak{E}}\,.$$

In other words:

$$V_1^\lambda(0) = \lambda|0\rangle\langle 0|^{\mathfrak{E}} + (1-\lambda)|3\rangle\langle 3|^{\mathfrak{E}}\,,$$
$$V_1^\lambda(1) = \lambda|1\rangle\langle 1|^{\mathfrak{E}} + (1-\lambda)|3\rangle\langle 3|^{\mathfrak{E}}\,,$$

$$V_2^\lambda(0) = \lambda|1\rangle\langle 1|^{\mathfrak{E}} + (1-\lambda)|4\rangle\langle 4|^{\mathfrak{E}}\,,$$

$$V_2^\lambda(1) = \lambda|2\rangle\langle 2|^{\mathfrak{E}} + (1-\lambda)|4\rangle\langle 4|^{\mathfrak{E}} .$$

For every $a \in \mathbf{A}$ and $t \in \theta$ we have

$$\begin{aligned}&\|W_t^0(a) - W_t^\lambda(a)\|_1\\ &= \|\lambda|t+a-1\rangle\langle t+a-1|^{\mathfrak{B}} - \lambda|t+2\rangle\langle t+2|^{\mathfrak{B}}\|_1\\ &= 2\lambda\end{aligned}$$

and

$$\begin{aligned}&\|V_t^0(a) - V_t^\lambda(a)\|_1\\ &= \| - \lambda|t+a-1\rangle\langle t+a-1|^{\mathfrak{E}} + \lambda|t+2\rangle\langle t+2|^{\mathfrak{E}}\|_1\\ &= 2\lambda .\end{aligned}$$

$\{(W_t^\lambda, V_t^\lambda) : t \in \theta\}$ defines an arbitrarily varying classical-quantum wiretap channel for every $\lambda \in [0,1]$.

At first, we consider $\{(W_t^0, V_t^0) : t \in \theta\}$.

*i) The Deterministic Secrecy Capacity of $\{(W_t^0, V_t^0) : t \in \theta\}$ Is Equal to Zero*

We set

$$\begin{aligned}&\tau(1 \mid 0) = 0 ; \;\; \tau(2 \mid 0) = 1 ;\\ &\tau(1 \mid 1) = 1 ; \;\; \tau(2 \mid 1) = 0 .\end{aligned}$$

It holds

$$\sum_{t\in\theta} \tau(t \mid 0) W_t^0(1) = |1\rangle\langle 1|^{\mathfrak{B}} = \sum_{t\in\theta} \tau(t \mid 1) W_t^0(0) ,$$

and of course for every $a \in \mathbf{A}$

$$\sum_{t\in\theta} \tau(t \mid a) W_t^0(a) = \sum_{t\in\theta} \tau(t \mid a) W_t^0(a) .$$

$\{(W_t^0) : t \in \theta\}$ is therefore symmetrizable. By Corollary 4.13 we have

$$C_s(\{(W_t^0, V_t^0) : t \in \theta\}) = 0 . \tag{241}$$

*ii) The Secrecy Capacity of $\{(W_t^0, V_t^0) : t \in \theta\}$ Under Common Randomness Assisted Quantum Coding Is Positive*

We denote by $p' \in P(\mathbf{A})$ the distribution on $\mathbf{A}$ such that $p'(1) = p'(2) = \frac{1}{2}$. Let $q \in [0,1]$. We define $Q(1) = q$, $Q(2) = 1-q$. We have

$$\begin{aligned}&\chi\left(p', \{W_Q^0(a) : a \in \mathbf{A}\}\right)\\ &= -\frac{1}{2} q \log \frac{1}{2} q + \frac{1}{2}(1-q)\log\frac{1}{2}(1-q) - \frac{1}{2}\log\frac{1}{2}\\ &+ q\log q + (1-q)\log(1-q) .\end{aligned}$$

When we differentiate this term by $q$ we obtain

$$\frac{1}{\log e}\left(-\frac{1}{2}\log\frac{1}{2}q-\frac{1}{2}+\frac{1}{2}\log\frac{1}{2}(1-q)+\frac{1}{2}+\log q+1-\log(1-q)-1\right)$$
$$=\frac{1}{2\log e}\left(\log q-\log(1-q)\right)\ .$$

$\log q-\log(1-q)$ is equal to zero if and only if $q=\frac{1}{2}$. By further calculation, one can show that $\chi\left(p',\{W_Q^0(a):a\in\mathbf{A}\}\right)$ achieves its minimum when $q=\frac{1}{2}$. This minimum is equal to $-\frac{1}{2}\log\frac{1}{4}+\frac{1}{2}\log\frac{1}{2}=\frac{1}{2}>0$. Thus

$$\max_p\min_q\chi\left(p,B_q^0\right)\geq\frac{1}{2}\ .$$

For all $t\in\theta$ it holds $V_t^0(0)=V_t^0(1)$; therefore for all $t^n\in\theta^n$ and any $p^n\in P(\mathbf{A}^n)$ we have

$$\begin{aligned}&\chi(p;Z_{t^n}^0)\\&=S(V_{t^n}^0(p^n))-\sum_{a^n\in\mathbf{A}^n}p^n(a^n)S(V_{t^n}^0(a^n))\\&=S(V_{t^n}^0(0^n))-\sum_{a^n\in\mathbf{A}^n}p^n(a^n)S(V_{t^n}^0(0^n))\\&=0\ .\end{aligned}$$

Thus

$$C_s(\{(W_t^0,V_t^0):t\in\theta\},cr)\geq\frac{1}{2}-0>0\ .\tag{242}$$

Now we consider $\{(W_t^\lambda,V_t^\lambda):t\in\theta\}$ when $\lambda\neq0$.

*iii) When $\lambda\neq0$ the Deterministic Secrecy Capacity of $\{(W_t^\lambda,V_t^\lambda):t\in\theta\}$ Is Equal to Its Secrecy Capacity Under Common Randomness Assisted Quantum Coding*

We suppose that for any $a,a'\in\mathbf{A}$ there are two distributions $\tau(\cdot\mid a)$ and $\tau(\cdot\mid a')$ on $\theta$ such that

$$\begin{aligned}&\sum_{t\in\theta}\tau(t\mid a')\cdot W_t^\lambda(a)=\sum_{t\in\theta}\tau(t\mid a)\cdot W_t^\lambda(a')\\&\Rightarrow(1-\lambda)\sum_{t\in\theta}\tau(t\mid a')|t+a-1\rangle\langle t+a-1|^{\mathfrak{B}}+\lambda\tau(1\mid a')|3\rangle\langle3|^{\mathfrak{B}}+\lambda\tau(2\mid a')|4\rangle\langle4|^{\mathfrak{B}}\\&=(1-\lambda)\sum_{t\in\theta}\tau(t\mid a)|t+a'-1\rangle\langle t+a'-1|^{\mathfrak{B}}+\lambda\tau(1\mid a)|3\rangle\langle3|^{\mathfrak{B}}+\lambda\tau(2\mid a)|4\rangle\langle4|^{\mathfrak{B}}\ .\end{aligned}\tag{243}$$

Since $|t+a-1\rangle\langle t+a-1|^{\mathfrak{B}}\in\left\{|0\rangle\langle0|^{\mathfrak{B}},|1\rangle\langle1|^{\mathfrak{B}},|2\rangle\langle2|^{\mathfrak{B}}\right\}$ for all $t$ and $a$, if $\lambda\neq0$ (243) implies that

$$\tau(t\mid a')=\tau(t\mid a)$$

for all $t\in\theta$. This means we have a distribution $\acute{p}$ on $\theta$ such that $\acute{p}(t)=\tau(t\mid a)$ for all $a\in\mathbf{A}$.

But there is clearly no such distribution $\acute{p}$ such that $\sum_{t\in\theta}\acute{p}(t)W_t^\lambda(0)=\sum_{t\in\theta}\acute{p}(t)W_t^\lambda(1)$, because then we would have

$$\begin{aligned}&\acute{p}(1)|0\rangle\langle 0|^{\mathfrak{B}}+\acute{p}(2)|1\rangle\langle 1|^{\mathfrak{B}}\\&=\acute{p}(1)|1\rangle\langle 1|^{\mathfrak{B}}+\acute{p}(2)|2\rangle\langle 2|^{\mathfrak{B}}\ .\end{aligned}$$

This would mean $\acute{p}(1)=\acute{p}(2)=0$, which obviously cannot be true. Thus $(W_t^\lambda)_{t\in\theta}$ is not symmetrizable.

By Corollary 4.13, if $\lambda\neq 0$

$$C_s(\{(W_t^\lambda,V_t^\lambda):t\in\theta\})=C_s(\{(W_t^\lambda,V_t^\lambda):t\in\theta\},cr)\ . \tag{244}$$

When $\lambda\searrow 0$ for every $a\in\mathbf{A}$ and $t\in\theta$ we have $\|W_t^0(a)-W_t^\lambda(a)\|_1=\|V_t^0(a)-V_t^\lambda(a)\|_1=2\lambda\searrow 0$.

By Corollary 4.18 the secrecy capacity of $\{(W_t^\lambda,V_t^\lambda):t\in\theta\}$ under common randomness assisted quantum coding is continues. Thus for any positive $\varepsilon$ there is a $\delta$ such that for all $\lambda\in]0,\delta[$ we have

$$C_s(\{(W_t^\lambda,V_t^\lambda):t\in\theta\})\geq C_s(\{(W_t^0,V_t^0):t\in\theta\},cr)-\varepsilon\geq\frac{1}{2}-\varepsilon\ . \tag{245}$$

In other words, when $\lambda\neq 0$ tends to zero, the deterministic secrecy capacity of $\{(W_t^\lambda,V_t^\lambda):t\in\theta\}$ tends to the secrecy capacity of $\{(W_t^0,V_t^0):t\in\theta\}$ under common randomness assisted quantum coding, which is positive, but the deterministic secrecy capacity of $\{(W_t^0,V_t^0):t\in\theta\}$ is equal to zero. Hence the deterministic secrecy capacity of $\{(W_t^\lambda,V_t^\lambda):t\in\theta\}$ is not continues at zero. □

Corollary 4.20 shows that small errors in the description of an arbitrarily varying classical-quantum wiretap channel may have severe consequences on the secrecy capacity. Corollary 4.18 shows that resources are very helpful to protect these consequences.

Now we are going to deliver the sufficient and necessary conditions for the continuity of the capacity function of arbitrarily varying classical-quantum wiretap channels.

**Corollary 4.21.** *For an arbitrarily varying classical-quantum channel $\{W_t:t\in\theta\}$ we define*

$$F(\{W_t:t\})\coloneqq\min_{\tau\in C(\theta|\mathbf{A})}\max_{a,a'}\left\|\sum_{t\in\theta}\tau(t\mid a)W_t(a')-\sum_{t\in\theta}\tau(t\mid a')W_t(a)\right\|_1\ ,$$

*where $C(\theta\mid\mathbf{A})$ the set of parametrized distributions sets $\{\tau(\cdot\mid a):a\in\mathbf{A}\}$ on $\theta$. The statement $F(\{W_t:t\})=0$ is equivalent to $\{W_t:t\in\theta\}$ being symmetrizable.*

*For an arbitrarily varying classical-quantum wiretap channel $\{(W_t,V_t):t\in\theta\}$, where $W_t:P(\mathbf{A})\to\mathcal{S}(H)$ and $V_t:P(\mathbf{A})\to\mathcal{S}(H')$, and a positive $\delta$ let $\mathbf{C}_\delta$ be defined as in Corollary 4.18.*

*$C_s(\{(W_t,V_t):t\})$, the deterministic secrecy capacity of arbitrarily varying classical-quantum wiretap channel is discontinuous at $\{(W_t,V_t):t\in\theta\}$ if and only if the following hold:*

*1) the secrecy capacity of $\{(W_t,V_t):t\in\theta\}$ under common randomness assisted*

*quantum coding is positive;*
*2)* $F(\{W_t : t\}) = 0$ *but for every positive* $\delta$ *there is a* $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$ *such that* $F(\{W'_t : t\}) > 0$.

*Proof.* At first we assume that the secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ under common randomness assisted quantum coding is positive and $F(\{W_t : t\}) = 0$. We choose a positive $\epsilon$ such that $C_s(\{(W_t, V_t) : t\}; cr) - \epsilon := C > 0$. By Corollary 4.18 the secrecy capacity under common randomness assisted quantum coding is continuous. Thus there exists a positive $\delta$ such that the for all $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$ we have

$$C_s\left(\{(W'_t, V'_t) : t \in \theta\}; cr\right) \geq C_s\left(\{(W_t, V_t) : t\}; cr\right) - \epsilon\,.$$

Now we assume that there is a $\{(W''_t, V''_t) : t \in \theta\} \in \mathbf{C}_\delta$ such that $F(\{W''_t : t\}) > 0$. This means that $\{W''_t : t\}$ is not symmetrizable. By Corollary 4.13 it holds

$$C_s\left(\{(W''_t, V''_t) : t \in \theta\}\right) = C_s\left(\{(W''_t, V''_t) : t\}; cr\right) \geq C > 0\,.$$

Since $F(\{W_t : t\}) = 0$, $\{W_t : t\}$ is symmetrizable. By Corollary 4.13,

$$C_s(\{(W_t, V_t) : t \in \theta\}) = 0\,.$$

Therefore the deterministic secrecy capacity is discontinuous at $\{(W_t, V_t) : t \in \theta\}$ when 1) and 2) hold.

Now let us consider the case when the deterministic secrecy capacity is discontinuous at $\{(W_t, V_t) : t \in \theta\}$.

We fix a $\tau \in C(\theta \mid \mathbf{A})$ and $a, a' \in \mathbf{A}$. The map

$$\{(W_t, V_t) : t \in \theta\} \to \left\| \sum_{t\in\theta} \tau(t \mid a) W_t(a') - \sum_{t\in\theta} \tau(t \mid a') W_t(a) \right\|_1$$

is continuous in the following sense: When $\left\|\sum_{t\in\theta} \tau(t \mid a) W_t(a') - \sum_{t\in\theta} \tau(t \mid a') W_t(a)\right\|_1 = C$ holds, then for every positive $\delta$ and any $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$ we have

$$\left| \| \sum_{t\in\theta} \tau(t \mid a) W'_t(a') - \sum_{t\in\theta} \tau(t \mid a') W'_t(a) \|_1 - C \right| \leq 2\delta\,.$$

Thus if for a $\tau \in C(\theta \mid \mathbf{A})$ we have $\left\|\sum_{t\in\theta} \tau(t \mid a) W_t(a') - \sum_{t\in\theta} \tau(t \mid a') W_t(a)\right\|_1 = C > 0$ for all $a, a' \in \mathbf{A}$, we also have

$$\left\| \sum_{t\in\theta} \tau(t \mid a) W'_t(a') - \sum_{t\in\theta} \tau(t \mid a') W'_t(a) \right\|_1 \geq C - 2\delta\,.$$

This means that when $F(\{W_t : t\}) > 0$ holds we can find a positive $\delta$ such that $F(\{W'_t : t\}) > 0$ holds for all $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$. By Corollary 4.13 it holds

$$C_s\left(\{(W'_t, V'_t) : t \in \theta\}\right) = C_s\left(\{(W'_t, V'_t) : t\}; cr\right) \geq C > 0\,.$$

By Corollary 4.18, $C_s\left(\{(W'_t, V'_t) : t\}; cr\right)$ is continuous.

Therefore, when the deterministic secrecy capacity is discontinuous at $\{(W_t, V_t) : t \in \theta\}$, $F(\{W_t : t\})$ cannot be positive.

We consider now that $F(\{W_t : t\}) = 0$ holds. By Corollary 4.13,

$$C_s(\{(W_t, V_t) : t \in \theta\} = 0\,.$$

When for every $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$ we have $F(\{W'_t : t\}) = 0$, then by Corollary 4.13

$$C_s(\{(W'_t, V'_t) : t \in \theta\}) = 0$$

and the deterministic secrecy capacity is thus continuous at $\{(W_t, V_t) : t \in \theta\}$.

Therefore, when the deterministic secrecy capacity is discontinuous at $\{(W_t, V_t) : t \in \theta\}$, for every positive $\delta$ there is a $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$ such that $F(\{W'_t : t\}) > 0$.

When for every positive $\delta$ there is a $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$ such that $F(\{W'_t : t\}) > 0$ and $C_s(\{(W_t, V_t) : t \in \theta\}, cr) = 0$ holds, then by Corollary 4.13 we have

$$C_s(\{(W'_t, V'_t) : t \in \theta\}) = C_s(\{(W'_t, V'_t) : t \in \theta\}, cr)\,,$$

and the deterministic secrecy capacity is continuous at $\{(W_t, V_t) : t \in \theta\}$.

Therefore, when the deterministic secrecy capacity is discontinuous at $\{(W_t, V_t) : t \in \theta\}$, $C_s(\{(W_t, V_t) : t \in \theta\}, cr)$ must be positive. □

**Corollary 4.22.** *Let $\{(W_t, V_t) : t \in \theta\}$ be an arbitrarily varying classical-quantum wiretap channel. When the secrecy capacity of $\{(W_t, V_t) : t \in \theta\}$ is positive, then there is a $\delta$ such that for all $\{(W'_t, V'_t) : t \in \theta\} \in \mathbf{C}_\delta$ we have*

$$C_s\left(\{(W'_t, V'_t) : t \in \theta\}\right) > 0\,.$$

*Proof.* Suppose we have $C_s(\{(W_t, V_t) : t \in \theta\}) > 0$. Then $\{W_t : t \in \theta\}$ is not symmetrizable, which means that $F(\{W_t : t\})$ is positive. In the proof of Corollary 4.21 we show that $F$ is continuous. Thus there is a positive $\delta'$ such that $F(\{W'_t : t\}) > 0$ for all $\{(W'_t, V'_t) : t \in \theta\} \in \mathsf{C}_{\delta'}$ . When $\{W_t : t \in \theta\}$ is not symmetrizable then we have $C_s(\{(W_t, V_t) : t \in \theta\}, cr) = C_s(\{(W_t, V_t) : t \in \theta\}) > 0$. By Corollary 4.18, the secrecy capacity under common randomness assisted quantum coding is continuous. Thus there is a positive $\delta''$ such that $C_s(\{(W'_t, V'_t) : t \in \theta\}, cr) > 0$ for all $\{(W'_t, V'_t) : t \in \theta\} \in \mathsf{C}_{\delta''}$ . We define $\delta := \min(\delta', \delta'')$ and the Corollary is shown.□

## 4.6 Applications and Further Notes

In Subsection 4.6.1 we will discuss the importance of the Ahlswede Dichotomy for arbitrarily varying classical-quantum wiretap channels. We will show that it can occur that the deterministic capacity of an arbitrarily varying classical-quantum wiretap channel is not equal to its randomness assisted capacity.

In Subsection 4.6.2 we will show that the research in quantum channels not only sets limitations, but also offers new fascinating possibilities. Applying the Ahlswede Dichotomy, we can prove that two arbitrarily varying classical-quantum wiretap channels, both with zero security capacity, allow perfect secure transmission, if we use them together. This is a phenomenon called "super-activation" which appears in quantum information theory (cf. [48]).

### 4.6.1 Further Notes on Ahlswede Dichotomy

In this subsection we give some notes on resource theory and the Ahlswede Dichotomy.

The Ahlswede Dichotomy states that either the deterministic security capacity of an arbitrarily varying classical-quantum wiretap channel is zero or it equals its randomness assisted security capacity. There are actually arbitrarily varying classical-quantum wiretap channels which have zero deterministic security capacity, but achieve a positive security capacity if the sender and the legal receiver can use a resource, as the following example shows. This shows that the Ahlswede Dichotomy is indeed a "dichotomy", and how helpful it a resource can be for the robust and secure message transmission.

**Example 4.23.** *Let $\{(W_t, V_t) : t \in \theta\}$ be an arbitrarily varying classical-quantum wiretap channel. By Theorem 1, $C_s(\{(W_t, V_t) : t \in \theta\})$ is equal to $C_s(\{(W_t, V_t) : t \in \theta\}; r)$ if $\{W_t : t \in \theta\}$ is not symmetrizable, and equal to zero if $\{W_t : t \in \theta\}$ is symmetrizable. If $\{W_t : t \in \theta\}$ is symmetrizable, it can actually occur that $C_s(\{(W_t, V_t) : t \in \theta\})$ is zero, but $C_s(\{(W_t, V_t) : t \in \theta\}; r)$ is positive, as following example shows (c.f. [6] for the case arbitrarily varying classical-quantum channel without wiretap).*

*Let $\theta := \{1, 2\}$. Let $\mathbf{A} = \{0, 1\}$. Let $H^{\mathfrak{B}} = \mathbb{C}^3$. Let $\{|0\rangle^{\mathfrak{B}}, |1\rangle^{\mathfrak{B}}, |2\rangle^{\mathfrak{B}}$ be a set of orthonormal vectors on $H^{\mathfrak{B}}$.*

*For $r \in [0, 1]$ let $P_r$ be the probability distribution on $\mathbf{A}$ such that $P_r(0) = r$ and $P_r(1) = 1 - r$. We define a channel $W_1 : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{B}})$ by*

$$W_1(P_r) = r|0\rangle\langle 0|^{\mathfrak{B}} + (1 - r)|1\rangle\langle 1|^{\mathfrak{B}} ,$$

*and a channel $W_2 : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{B}})$ by*

$$W_1(P_r) = r|1\rangle\langle 1|^{\mathfrak{B}} + (1 - r)|2\rangle\langle 2|^{\mathfrak{B}} .$$

*In other word*

$$W_1(0) = |0\rangle\langle 0|^{\mathfrak{B}} ,$$
$$W_1(1) = |1\rangle\langle 1|^{\mathfrak{B}} ,$$

$$W_2(0) = |1\rangle\langle 1|^{\mathfrak{B}} ,$$
$$W_2(1) = |2\rangle\langle 2|^{\mathfrak{B}} .$$

*Let $H^{\mathfrak{E}} = \mathbb{C}^2$. Let $\{|3\rangle^{\mathfrak{E}}, |4\rangle^{\mathfrak{E}}$ be a set of orthonormal vectors on $H^{\mathfrak{E}}$.*
*We define a channel $V_1 : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{E}})$ by*

$$V_1(P_r) = |3\rangle\langle 3|^{\mathfrak{E}} ,$$

*and a channel $V_2 : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{E}})$ by*

$$V_1(P_r) = |4\rangle\langle 4|^{\mathfrak{E}} .$$

*$\{(W_t, V_t) : t \in \theta\}$ defines an arbitrarily varying classical-quantum wiretap channel.*

*We set*

$$\tau(1 \mid 0) = 0\,; \quad \tau(2 \mid 0) = 1\,;$$
$$\tau(1 \mid 1) = 1\,; \quad \tau(2 \mid 1) = 0\,.$$

*It holds*

$$\sum_{t\in\theta} \tau(t \mid 0) W_t(0) = \sum_{t\in\theta} \tau(t \mid 0) W_t(0)\,,\ \sum_{t\in\theta} \tau(t \mid 1) W_t(1) = \sum_{t\in\theta} \tau(t \mid 1) W_t(1)\,,$$

*and*

$$\sum_{t\in\theta} \tau(t \mid 0) W_t(1) = |1\rangle\langle 1|^{\mathfrak{E}} = \sum_{t\in\theta} \tau(t \mid 1) W_t(0)\,.$$

$\{(W_t) : t \in \theta\}$ *is therefore symmetrizable. By Theorem 1 we have*

$$C_s(\{(W_t, V_t) : t \in \theta\}) = 0\,. \tag{246}$$

*By [21], for any arbitrarily varying classical-quantum wiretap channel* $\{(W_t, V_t) : t \in \theta\}$*, we have*

$$\begin{aligned} &C_s(\{(W_t, V_t) : t \in \theta\}; r) \\ &\geq \max_{P\in\mathcal{P}} \Bigg( \min_{Q\in\mathcal{Q}} \chi\left(P, \{U^Q(a) : a \in \mathbf{A}\}\right) \\ &- \lim_{n\to\infty} \max_{t^n\in\theta^n} \frac{1}{n} \chi(P^n, \{V_{t^n}(a^n) : a^n \in \mathbf{A}^n\}) \Bigg)\,, \end{aligned} \tag{247}$$

*where* $\mathcal{P}$ *is the set of distributions on* $\mathbf{A}$*,* $\mathcal{Q}$ *is the set of distributions on* $\theta$*, and* $U^Q(a) = \sum_{t\in\theta} Q(t) W_t(a)$ *for* $Q \in \mathcal{Q}$*.*

*For all* $n \in \mathbb{N}$*,* $t^n \in \theta^n$*, and* $P^n \in \mathcal{P}^n$ *we have* $\chi(P^n, \{V_{t^n}(a^n) : a^n \in \mathbf{A}^n\}) = 1\log 1 - 1\log 1 = 0$*, therefore*

$$C_s(\{(W_t, V_t) : t \in \theta\}; r) \geq \max_{P\in\mathcal{P}} \min_{Q\in\mathcal{Q}} \chi\left(P, \{U^Q(a) : a \in \mathbf{A}\}\right)\,.$$

*We denote by* $p' \in P(\mathbf{A})$ *the distribution on* $\mathbf{A}$ *such that* $p'(1) = p'(2) = \frac{1}{2}$*. Let* $q \in [0, 1]$*. We define* $Q(1) = q$*,* $Q(2) = 1 - q$*. We have*

$$\begin{aligned} &\chi\left(p', \{W_Q^0(a) : a \in \mathbf{A}\}\right) \\ &= -\frac{1}{2} q \log \frac{1}{2} q + \frac{1}{2}(1-q) \log \frac{1}{2}(1-q) - \frac{1}{2} \log \frac{1}{2} \\ &+ q \log q + (1-q) \log(1-q)\,. \end{aligned}$$

*By the differentiation by* $q$*, we obtain*

$$\begin{aligned} &\frac{1}{\log e} \left( -\frac{1}{2} \log \frac{1}{2} q - \frac{1}{2} + \frac{1}{2} \log \frac{1}{2}(1-q) + \frac{1}{2} + \log q + 1 - \log(1-q) - 1 \right) \\ &= \frac{1}{2 \log e} \left( \log q - \log(1-q) \right)\,. \end{aligned}$$

*This term is equal to zero if and only if* $q = \frac{1}{2}$. *By further calculation, one can show that* $\chi\left(p', \{W_Q^0(a) : a \in \mathbf{A}\}\right)$ *achieves its minimum when* $q = \frac{1}{2}$. *This minimum is equal to* $-\frac{1}{2}\log\frac{1}{4} + \frac{1}{2}\log\frac{1}{2} = \frac{1}{2} > 0$. *Thus*

$$\max_p \min_q \chi\left(p, B_q^0\right) \geq \frac{1}{2}\,.$$

*For all* $t \in \theta$ *it holds* $V_t^0(0) = V_t^0(1)$, *therefore for all* $t^n \in \theta^n$ *and any* $p^n \in P(\mathbf{A}^n)$ *we have*

$$\begin{aligned}
&\chi(p; Z_{t^n}^0) \\
&= S(V_{t^n}^0(p^n)) - \sum_{a^n \in A^n} p^n(a^n) S(V_{t^n}^0(a^n)) \\
&= 0\,.
\end{aligned}$$

*By (247),*

$$C_s(\{(W_t^0, V_t^0) : t \in \theta\}, cr) \geq \frac{1}{2} - 0 > 0\,. \tag{248}$$

*This shows an example of an arbitrarily varying classical-quantum channel such that its deterministic capacity is zero, but its random capacity is positive.*

Thus, a "useless" arbitrarily varying classical-quantum channel, i.e. with zero deterministic secrecy capacity, allows secure transmission if the sender and the legal receiver have the possibility to use a resource, either randomness, common randomness, or even a "cheap", insecure, and weak correlation. Here we say "cheap" and "weak" in the sense of the discussion in Section 4.4.1.

### 4.6.2 Super-Activation

One of the properties of classical channels is that in the majority of cases, if we have a channel system where two sub-channels are used together, the capacity of this channel system is the sum of the two sub-channels' capacities. Particularly, a system consisting of two orthogonal classical channels, where both are "useless" in the sense that they both have zero capacity for message transmission, the capacity for message transmission of the whole system is zero as well ("$0 + 0 = 0$"). For the definition of "two orthogonal channels" in classical systems please see [38].

In contrast to the classical information theory, it is known that the capacities of quantum channels can be super-additive, i.e., there are cases in which the capacity of the product $W_1 \otimes W_2$ of two quantum channels $W_1$ and $W_2$ are larger than the sum of the capacity of $W_1$ and the capacity of $W_2$ (cf. [48] and [39]). "The whole is greater than the sum of its parts" - Aristotle.

Particularly in quantum information theory, there are examples of two quantum channels, $W_1$ and $W_2$, with zero capacity, which allow perfect transmission if they are used together, i.e., the capacity of their product $W_1 \otimes W_2$ is positive, (cf. [62], [61], [55] and also [30] for a rare case result when this phenomenon occurs using two classical arbitrarily varying wiretap channels). This is due to the fact that there are different reasons why a quantum channel can have zero capacity. If we have two channels which have zero capacity for different reasons, they can "remove" their weaknesses from

each other, or in other words, "activate" each other. We call this phenomenon "super-activation" ("$0 + 0 > 0$").

It is known that arbitrarily varying classical-quantum wiretap channels with positive secrecy capacities are super-additive. This means that the product $W_1 \otimes W_2$ of two arbitrarily varying classical-quantum wiretap channels $W_1$ and $W_2$, both with positive secrecy capacities, can have a capacity which is larger than the sum of the capacity of $W_1$ and the capacity of $W_2$ (cf. [48]).

Using Theorem 4.1, we can demonstrate the following Theorem,

**Theorem 4.24.** *Super-activation occurs for arbitrarily varying classical-quantum wiretap channels.*

Note that the results of [48] (super-additivity of arbitrarily varying classical-quantum wiretap channels with positive secrecy capacities) do not imply super-activation of arbitrarily varying classical-quantum wiretap channels, since here we consider channels with zero secrecy capacity.

We will prove Theorem 4.24 by giving an example (Example 4.25) in which two arbitrarily varying classical-quantum wiretap channels, which are themselves "useless" in the sense that they have both zero secrecy capacity, acquire positive secrecy capacity when used together. This is due the following.

Suppose we have an arbitrarily varying classical-quantum wiretap channel with positive randomness assisted secrecy capacity. By Theorem 2, the randomness assisted secrecy capacity is equal to the common randomness assisted secrecy capacity. But the problem for the sender and the legal receiver is that each party does not know which code is used in the particular transmission if the channel that connects them has zero deterministic capacity for message transmission. However, suppose we have another arbitrarily varying classical-quantum wiretap channel which has a positive deterministic capacity for message transmission. Then the sender and the legal receiver can use it to transmit which code is used in the particular transmission. This is possible even when the second arbitrarily varying classical-quantum wiretap channel has zero randomness assisted secrecy capacity, since we allow the wiretapper to know which specific code is used.

We may see it in the following way. If we have two arbitrarily varying classical-quantum wiretap channels, one of them is relatively secure, but not very robust against jamming, while the other one is relatively robust, but not very secure against eavesdropping. We can achieve that they "remove" their weaknesses from each other, or in other words, "activate" each other.

We now give an example of super-activation for arbitrarily varying classical-quantum wiretap channels.

**Example 4.25.** *Let $\theta$, $\mathbf{A}$, $H^{\mathfrak{B}}$ and $H^{\mathfrak{E}}$ be defined as in Example 4.23. We define $\{(W_t, V_t) : t \in \theta\}$ as in Example 4.23.*

*For $r \in [0,1]$ let $P_r$ be the probability distribution on $\mathbf{A}$ such that $P_r(0) = r$ and $P_r(1) = 1 - r$. We define a channel $W'_1 : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{B}})$ by*

$$W'_1(P_r) = \frac{3}{4} r |0\rangle\langle 0|^{\mathfrak{B}} + \frac{3}{4}(1-r)|1\rangle\langle 1|^{\mathfrak{B}} + \frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}} ,$$

*and a channel $W'_2 : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{B}})$ by*

$$W'_2(P_r) = \frac{3}{4} r |0\rangle\langle 0|^{\mathfrak{B}} + \frac{3}{4}(1-r)|1\rangle\langle 1|^{\mathfrak{B}} + \frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}} .$$

*In other words*

$$
\begin{aligned}
{W'}_1(0) &= \frac{3}{4}|0\rangle\langle 0|^{\mathfrak{B}} + \frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}} \\
{W'}_1(1) &= \frac{3}{4}|1\rangle\langle 1|^{\mathfrak{B}} + \frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}} \\
{W'}_2(0) &= \frac{3}{4}|0\rangle\langle 0|^{\mathfrak{B}} + \frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}} \\
{W'}_2(1) &= \frac{3}{4}|1\rangle\langle 1|^{\mathfrak{B}} + \frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}} \,. \qquad (249)
\end{aligned}
$$

*We define a channel* ${V'}_1 : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{E}})$ *by*

$$
{V'}_1(P_r) = r|0\rangle\langle 0|^{\mathfrak{E}} + (1-r)|1\rangle\langle 1|^{\mathfrak{E}} \,,
$$

*and a channel* ${V'}_2 : P(\mathbf{A}) \to \mathcal{S}(H^{\mathfrak{E}})$ *by*

$$
{V'}_2(P_r) = r|0\rangle\langle 0|^{\mathfrak{E}} + (1-r)|1\rangle\langle 1|^{\mathfrak{E}} \,.
$$

*In other words*

$$
\begin{aligned}
{V'}_1(0) &= |0\rangle\langle 0|^{\mathfrak{E}} \\
{V'}_1(1) &= |1\rangle\langle 1|^{\mathfrak{E}} \\
{V'}_2(0) &= |0\rangle\langle 0|^{\mathfrak{E}} \\
{V'}_2(1) &= |1\rangle\langle 1|^{\mathfrak{E}} \,. \qquad (250)
\end{aligned}
$$

$\{(W'_t, V'_t) : t \in \theta\}$ *defines an arbitrarily varying classical-quantum wiretap channel.*

*We denote the uniform distribution on* $\mathbf{A}$ *by* $P$. *We have* $P(0) = P(1) = \frac{1}{2}$. *The capacity of* $\{W'_t : t \in \theta\}$ *is larger or equal to* $\min_{Q\in\mathcal{Q}} \chi\left(P, \{U^Q(a) : a \in \mathbf{A}\}\right) = \log 3 - 1 > 0$.

*However, for all* $(n, J_n)$ *code* $\left(E^n, \{D^n_j : j = 1, \cdots J_n\}\right)$ *the wiretapper can define a set of decoding operators* $\{D^n_{j,wiretap} : j = 1, \cdots J_n\}$ *by* $D^n_{j,wiretap} := \sum_{a^n} E^n(a^n \mid j) \left(\bigotimes_i |a_i\rangle^{\mathfrak{B}}\right)\left(\bigotimes_i \langle a_i|^{\mathfrak{B}}\right)$. *For any probability distribution* $Q^n$ *on* $\mathbf{A}^n$, *denote the wiretapper's random output using* $\{D^n_{j,wiretap} : j = 1, \cdots J_n\}$ *at channel state* $t^n$ *by* $C_{t^n}$, *then* $\chi(Q^n, Z_{t^n}) \geq I(Q^n, C_{t^n}) = H(Q^n)$, *where* $I(\cdot,\cdot)$ *is the mutual information, and* $H(\cdot)$ *is the Shannon entropy (please see [32] for the definitions of the mutual information and the Shannon entropy for classical random variables). If* $\chi(R_{uni}, Z_{t^n}) < \frac{1}{2}$ *holds, we also have* $\log J_n = H(R_{uni}) < \frac{1}{2}$, *but this implies* $J_n = 1$. *Thus*

$$
C_{s*}(\{(W'_t, V'_t) : t \in \theta\}) = 0 \,. \qquad (251)
$$

*Let us now consider the arbitrarily varying classical-quantum wiretap channel* $\left\{\left(W_{t_1} \otimes {W'}_{t_2}, V_{t_1} \otimes {V'}_{t_2}\right) : (t_1, t_2) \in \theta^2\right\}$, *where* $\left\{\left(W_{t_1} \otimes {W'}_{t_2}\right) : (t_1, t_2) \in \theta^2\right\}$ *is an arbitrarily varying classical-quantum channel* $\{(00), (01), (10), (11)\} \to H^{\otimes 2}$, $(a, a') \to W_{t_1}(a) \otimes {W'}_{t_2}(a')$, *and* $\left\{\left(V_{t_1} \otimes {V'}_{t_2}\right) : (t_1, t_2) \in \theta^2\right\}$ *is an arbitrarily varying classical-quantum channel* $\{(00), (01), (10), (11)\} \to H^{\mathfrak{B}^2}$, $(a, a') \to V_{t_1}(a) \otimes {V'}_{t_2}(a')$, *if the channel state is* $(t_1, t_2)$.

*We have*

$$C_s\left(\left\{\left(W_{t_1}\otimes {W'}_{t_2}, V_{t_1}\otimes {V'}_{t_2}\right):(t_1,t_2)\in\theta^2\right\};r\right)\geq\frac{1}{2}>0\,. \tag{252}$$

*Assume* $\left\{\left(W_{t_1}\otimes {W'}_{t_2}\right):\ (t_1,t_2)\ \in\ \theta^2\right\}$ *is symmetrizable, then there exists a parametrized set of distributions* $\{\tau(\cdot\mid(a,a')):(a,a')\in\{(00),(01),(10),(11)\}\}$ *on* $\theta^2$ *such that for all* $(a,a')$, $(b,b')\in\{(00),(01),(10),(11)\}$ *it holds*

$$\sum_{(t_1,t_2)\in\theta^2}\tau((t_1,t_2)\mid(b,b'))W_{t_1}(a)\otimes{W'}_{t_2}(a')=\sum_{(t_1,t_2)\in\theta^2}\tau((t_1,t_2)\mid(a,a'))W_{t_1}(b)\otimes{W'}_{t_2}(b')\,. \tag{253}$$

*(253) implies that*

$$\begin{aligned}
&\sum_{(t_1,t_2)\in\theta^2}\tau((t_1,t_2)\mid(0,0))W_{t_1}(0)\otimes{W'}_{t_2}(1)=\sum_{(t_1,t_2)\in\theta^2}\tau((t_1,t_2)\mid(0,1))W_{t_1}(0)\otimes{W'}_{t_2}(0)\\
&\Rightarrow\left(\tau((1,1)\mid(0,0))+\tau((1,2)\mid(0,0))\right)|0\rangle\langle 0|^{\mathfrak{B}}\otimes\left(\frac{3}{4}|1\rangle\langle 1|^{\mathfrak{B}}+\frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}}\right)\\
&+\left(\tau((2,1)\mid(0,0))+\tau((2,2)\mid(0,0))\right)|1\rangle\langle 1|^{\mathfrak{B}}\otimes\left(\frac{3}{4}|1\rangle\langle 1|^{\mathfrak{B}}+\frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}}\right)\\
&=\left(\tau((1,1)\mid(0,1))+\tau((1,2)\mid(0,1))\right)|0\rangle\langle 0|^{\mathfrak{B}}\otimes\left(\frac{3}{4}|0\rangle\langle 0|^{\mathfrak{B}}+\frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}}\right)\\
&+\left(\tau((2,1)\mid(0,1))+\tau((2,2)\mid(0,1))\right)|1\rangle\langle 1|^{\mathfrak{B}}\otimes\left(\frac{3}{4}|0\rangle\langle 0|^{\mathfrak{B}}+\frac{1}{4}|2\rangle\langle 2|^{\mathfrak{B}}\right)\\
&\Rightarrow\frac{3}{4}\left(\tau((1,1)\mid(0,0))+\tau((1,2)\mid(0,0))\right)|0\rangle\langle 0|^{\mathfrak{B}}\otimes|1\rangle\langle 1|^{\mathfrak{B}}\\
&+\frac{1}{4}\left(\tau((1,1)\mid(0,0))+\tau((1,2)\mid(0,0))\right)|0\rangle\langle 0|^{\mathfrak{B}}\otimes|2\rangle\langle 2|^{\mathfrak{B}}\\
&+\frac{3}{4}\left(\tau((2,1)\mid(0,0))+\tau((2,2)\mid(0,0))\right)|1\rangle\langle 1|^{\mathfrak{B}}\otimes|1\rangle\langle 1|^{\mathfrak{B}}\\
&+\frac{1}{4}\left(\tau((2,1)\mid(0,0))+\tau((2,2)\mid(0,0))\right)|1\rangle\langle 1|^{\mathfrak{B}}\otimes|2\rangle\langle 2|^{\mathfrak{B}}\\
&=\frac{3}{4}\left(\tau((1,1)\mid(0,1))+\tau((1,2)\mid(0,1))\right)|0\rangle\langle 0|^{\mathfrak{B}}\otimes|0\rangle\langle 0|^{\mathfrak{B}}\\
&+\frac{1}{4}\left(\tau((1,1)\mid(0,1))+\tau((1,2)\mid(0,1))\right)|0\rangle\langle 0|^{\mathfrak{B}}\otimes|2\rangle\langle 2|^{\mathfrak{B}}\\
&+\frac{3}{4}\left(\tau((2,1)\mid(0,1))+\tau((2,2)\mid(0,1))\right)|1\rangle\langle 1|^{\mathfrak{B}}\otimes|0\rangle\langle 0|^{\mathfrak{B}}\\
&+\frac{1}{4}\left(\tau((2,1)\mid(0,1))+\tau((2,2)\mid(0,1))\right)|1\rangle\langle 1|^{\mathfrak{B}}\otimes|2\rangle\langle 2|^{\mathfrak{B}}\\
&\Rightarrow\tau((1,1)\mid(0,0))+\tau((1,2)\mid(0,0))\\
&=\tau((2,1)\mid(0,0))+\tau((2,2)\mid(0,0))\\
&=\tau((1,1)\mid(0,1))+\tau((1,2)\mid(0,1))\\
&=\tau((2,1)\mid(0,1))+\tau((2,2)\mid(0,1))\\
&=0\\
&\Rightarrow\lightning\,.
\end{aligned} \tag{254}$$

*Therefore* $\left\{\left(W_{t_1}\otimes {W'}_{t_2}\right)\colon (t_1,t_2)\in\theta^2\right\}$ *is not symmetrizable, and by Theorem 1,*

$$\begin{aligned} & C_{s*}\left(\left\{\left(W_{t_1}\otimes {W'}_{t_2}, V_{t_1}\otimes {V'}_{t_2}\right) : (t_1,t_2)\in\theta^2\right\}\right) \\ & = C_s\left(\left\{\left(W_{t_1}\otimes {W'}_{t_2}, V_{t_1}\otimes {V'}_{t_2}\right) : (t_1,t_2)\in\theta^2\right\}; r\right) \\ & > 0\,. \end{aligned} \tag{255}$$

This example shows that although both $\{(W_t,V_t) : t\in\theta\}$ and $\{({W'}_t,{V'}_t) : t\in\theta\}$ are themselves useless, they allow secure transmission using together ("$0+0>0$"). Thus Theorem 4.24 is proven. This shows that the research in quantum channels with channel uncertainty and eavesdropping can lead to some promising applications.

**Corollary 4.26.** *Let* $\{(W_t,V_t) : t\in\theta\}$ *and* $\{({W'}_t,{V'}_t) : t\in\theta\}$ *be two arbitrarily varying classical-quantum wiretap channels.*

*1) If* $C_{s*}(\{(W_t,V_t) : t\in\theta\}) = C_{s*}(\{({W'}_t,{V'}_t) : t\in\theta\}) = 0$ *then* $C_{s*}(\{W_t\otimes {W'}_{t'}, V_t\otimes {V'}_{t'} : t,t'\in\theta\})$ *is positive if and only if* $\{W_t\otimes {W'}_{t'} : t,t'\in\theta\}$ *is not symmetrizable and* $C_s(\{W_t\otimes {W'}_{t'}, V_t\otimes {V'}_{t'} : t,t'\in\theta\}, cr)$ *is positive.*

*2) If the secrecy capacity under common randomness assisted quantum coding shows no super-activation for* $\{(W_t,V_t) : t\in\theta\}$ *and* $\{({W'}_t,{V'}_t) : t\in\theta\}$ *then the secrecy capacity can only show super-activation for* $\{(W_t,V_t) : t\in\theta\}$ *and* $\{({W'}_t,{V'}_t) : t\in\theta\}$ *if one of* $\{(W_t,V_t) : t\in\theta\}$ *and* $\{({W'}_t,{V'}_t) : t\in\theta\}$ *has positive secrecy capacity under common randomness assisted quantum coding and a symmetrizable legal channel while the other one has zero secrecy capacity under common randomness assisted quantum coding and a non-symmetrizable legal channel.*

*Proof.* By Theorem 4.1, $C_{s*}(\{W_t\otimes {W'}_{t'}, V_t\otimes {V'}_{t'} : t,t'\in\theta\})$ is equal to $C_s(\{W_t\otimes {W'}_{t'}, V_t\otimes {V'}_{t'} : t,t'\in\theta\}, cr)$ when $\{W_t\otimes {W'}_{t'} : t,t'\in\theta\}$ is not symmetrizable and to zero when $\{W_t\otimes {W'}_{t'} : t,t'\in\theta\}$ is symmetrizable. Thus 1) holds.

When $\{W_t : t\in\theta\}$ and $\{{W'}_t : t\in\theta\}$ are both symmetrizable then there exists two parametrized set of distributions $\{\tau(\cdot\mid a) : a\in\mathbf{A}\}$, $\{\tau'(\cdot\mid a) : a\in\mathbf{A}\}$ on $\theta$ such that for all $a$, $a'\in\mathbf{A}$, we have $\sum_{t\in\theta}\tau(t\mid a)W_t(a') = \sum_{t\in\theta}\tau(t\mid a')W_t(a)$, $\sum_{t\in\theta}\tau'(t\mid a){W'}_t(a') = \sum_{t\in\theta}\tau;(t\mid a'){W'}_t(a)$. We can set $\tau((t,t')\mid(a,a')) := \tau(t\mid a)\tau'(t'\mid a')$ and obtain

$$\sum_{(t,t')\in\theta\times\theta}\tau((t,t')\mid(a_1,a_1'))W_t(a_2)\otimes {W'}_{t'}(a_2') = \sum_{(t,t')\in\theta\times\theta}\tau((t,t')\mid(a_2,a_2'))W_t(a_1)\otimes {W'}_{t'}(a_1')$$

for all $(a_1,a_1')$, $(a_2,a_2')\in\mathbf{A}\times\mathbf{A}$, which means that $\{W_t\otimes {W'}_t : t,t'\in\theta\}$ is symmetrizable and super-activation does not occur because of 1).

When $\{W_t : t\in\theta\}$ and $\{{W'}_t : t\in\theta\}$ are both not symmetrizable then their secrecy capacities are equal to their secrecy capacities under common randomness assisted quantum coding. When $C_s(\{(W_t,V_t) : t\in\theta\}, cr) = C_s(\{({W'}_t,{V'}_t) : t\in\theta\}, cr) = 0$. Because of our assumption, $C_s(\{W_t\otimes {W'}_{t'}, V_t\otimes {V'}_{t'} : t,t'\in\theta\}, cr) = 0$. By 1), super-activation cannot occur.

When one of $\{W_t : t\in\theta\}$ and $\{{W'}_t : t\in\theta\}$, say $\{W_t : t\in\theta\}$, is not symmetrizable while the other one is symmetrizable, then the assumption that $C_{s*}(\{(W_t,V_t) :$

$t \in \theta\}) = 0$ indicating that $C_s(\{(W_t, V_t) : t \in \theta\}, cr) = 0$. When $C_s(\{(W'_t, V'_t) : t \in \theta\}, cr)$ is also zero, then by our assumption, super-activation cannot occur. Thus 2) holds. □

# 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, Ann. Math. Stat., Vol. 41, No. 3, 1970.

[2] R. Ahlswede, Elimination of correlation in random codes for arbitrarily varying channels, Z. Wahrscheinlichkeitstheorie verw. Gebiete, Vol. 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. Inform. Theory, Vol. 32, 621-629, 1986.

[5] 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, 103-156, 2013.

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

[7] R. Ahlswede and N. Cai, Correlation sources help transmission over an arbitrarily varying channel, IEEE Trans. Inform. Theory, Vol. 43, No. 4, 1254-1255, 1997.

[8] R. Ahlswede and A. Winter, Strong converse for identification via quantum channels, IEEE Trans. Inform. Theory, Vol. 48, No. 3, 569-579, 2002. Addendum: IEEE Trans. Inform. Theory, Vol. 49, No. 1, 346, 2003.

[9] R. Alicki and M. Fannes, Continuity of quantum conditional information, J. Phys. A: Math. Gen., Vol. 37, L55, 2004.

[10] K. M. R. Audenaert, A sharp continuity estimate for the von Neumann entropy, J. Phys. A: Math. Theor., Vol. 40, 8127-8136, 2007.

[11] H. Barnum, E. Knill, M. A. Nielsen, On Quantum Fidelities and Channel Capacities, IEEE Trans. Inform. Theory, Vol. 46, 1317-1329, 2000.

[12] H. Barnum, M. A. Nielsen, and B. Schumacher, Information transmission through a noisy quantum channel, Phys. Rev. A, Vol. 57, 4153 , 1998.

[13] C. H. Bennet, Quantum cryptography using any two non-orthogonal states, Phys. Rev. Lett., Vol. 68, 3121-3124, 1992.

[14] C. H. Bennett and G. Brassard, Quantum cryptography: public key distribution and coin tossing, Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, 175, 1984.

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

[16] 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, 247-283, arXiv:1209.6325, 2012.

[17] I. Bjelaković, H. Boche, and J. Nötzel, Entanglement transmission and generation under channel uncertainty: universal quantum channel coding, Commun. Math. Phys, Vol. 292, No. 1, 55-97, 2009.

[18] I. Bjelaković, H. Boche, and J. Sommerfeld, Secrecy results for compound wiretap channels, Problems of Information Transmission, Vol. 49, No. 1, 73-98 2013.

[19] D. Blackwell, L. Breiman, and A. J. Thomasian, The capacity of a class of channels, Ann. Math. Stat., Vol. 30, No. 4, 1229-1241, 1959.

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

[21] V. Blinovsky and M. Cai, Classical-quantum arbitrarily varying wiretap channel, Information Theory, Combinatorics, and Search Theory, in Memory of Rudolf Ahlswede, H. Aydinian, F. Cicalese, and C. Deppe eds., LNCS Vol.7777, 197-206, arXiv:1208.1151, 2012.

[22] M. Bloch and J. N. Laneman, On the secrecy capacity of arbitrary wiretap channels, Communication, Control, and Computing, Forty-Sixth Annual Allerton Conference Allerton House, UIUC, USA, 818-825, 2008.

[23] H. Boche, M. Cai and N. Cai, Channel state detecting code for compound quantum channel, preprint.

[24] H. Boche, M. Cai, N. Cai, and C. Deppe, Secrecy capacities of compound quantum wiretap channels and applications, Phys. Rev. A, Vol.89, No.5, 052320, arXiv:1302.3412, 2014.

[25] H. Boche, M. Cai, C. Deppe, and J. Nötzel, Classical-quantum arbitrarily varying wiretap channel - Ahlswede Dichotomy - positivity - resources - super activation, Quant. Inf. Proc., Vol. 15, No. 11, 4853-489, arXiv:1307.8007, 2016.

[26] H. Boche, M. Cai, C. Deppe, and J. Nötzel, Classical-quantum arbitrarily varying wiretap channel: Common randomness assisted code and continuity, Quant. Inf. Proc., Vol. 16, No. 1, 1-48, 2016.

[27] H. Boche, M. Cai, C. Deppe, and J. Nötzel, Classical-quantum arbitrarily varying wiretap channel: secret message transmission under jamming attacks, *arXiv:1702.03483*, 2017

[28] H. Boche and J. Nötzel, Arbitrarily small amounts of correlation for arbitrarily varying quantum channel, J. Math. Phys., Vol. 54, Issue 11, arXiv 1301.6063, 2013.

[29] H. Boche and J. Nötzel, Positivity, discontinuity, finite resources, and nonzero error for arbitrarily varying quantum channels, J. Math. Phys., Vol. 55, 122201, 2014.

[30] H. Boche and R. F. Schaefer (Wyrembelski), Capacity results and super-activation for wiretap channels with active wiretappers, Trans. on Inform. Forensics and Security, Vol. 8, No. 9, 2013.

[31] N. Cai, A. Winter, and R. W. Yeung, Quantum privacy and quantum wiretap channels, Problems of Information Transmission, Vol. 40, No. 4, 318-336, 2004.

[32] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, Academic Press/Akademiai Kiao, 1981.

[33] I. Csiszár and P. Narayan, The capacity of the arbitrarily varying channel revisited: positivity, constraints, IEEE Trans. Inform. Theory, Vol. 34, No. 2, 181-193, 1988.

[34] I. Devetak, The private classical information capacity and quantum information capacity of a quantum channel, IEEE Trans. Inform. Theory, Vol. 51, No. 1, 44-55, 2005.

[35] I. Devetak and A. Winter, Distillation of secret key and entanglement from quantum states, Proc. R. Soc. A, Vol. 461, 207-235, 2005.

[36] T. Ericson, Exponential error bounds for random codes in the arbitrarily varying channel, IEEE Trans. Inform. Theory, Vol. 31, No. 1, 42-48, 1985.

[37] M. Fannes, A continuity property of the entropy density for spin lattice systems, Commun. Math. Phys, Vol. 31. 291-294, 1973.

[38] K. Fazel and S. Kaiser, Multi-Carrier and Spread Spectrum Systems. From OFDM and MC-CDMA to LTE and WiMAX, 2. edition, ISBN 978-0-470-99821-2, John Wiley & Sons, New York, 2008.

[39] G. Giedke and M. M. Wolf, Quantum communication: super-activated channels, Nature Photonics, Vol. 5, No. 10, 578-580, 2011.

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

[41] M. Hayashi, H. Nagaoka, General formulas for capacity of classical-quantum channels, IEEE Trans. Inform. Theory, Vol. 49. No. 7, 1753-1768, 2003.

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

[43] R. Klesse, Approximate quantum error correction, random codes, and quantum channel capacity, Phys. Rev. A, Vol. 75, 062315, 2007.

[44] K. Kraus, States, Effects, and Operations, Springer, Berlin, 1983.

[45] D. Leung, J. Lim, and P. Shor, On quantum capacity of erasure channel assisted by back classical communication, Phys. Rev. Lett., Vol. 103, No. 24, 240505, 2009.

[46] D. Leung and G. Smith, Continuity of quantum channel capacities, Commun. Math. Phys, Vol. 292, No. 1, 201-215, 2009.

[47] Y. Liang, G. Kramer, H. Poor, and S. Shamai, Compound wiretap channels, EURASIP Journal on Wireless Communications and Networking - Special issue on wireless physical layer security archive, Vol. 2009, Article No. 5, 2009.

[48] K. Li, A. Winter, X. B. Zou, G. C. Guo, Private capacity of quantum channels is not additive, Phys. Rev. Lett., Vol. 103, No. 12, 120501, 2009.

[49] S. Lloyd, Capacity of the noisy quantum channel, Phys. Rev. A, Vol. 55, No. 3, 1613-1622, 1997.

[50] V. D. Milman and G. Schechtman, Asymptotic Theory of Finite Dimensional Normed Spaces. Lecture Notes in Mathematics 1200, Springer-Verlag, corrected second printing, Berlin, UK, 2001.

[51] M. Mosonyi, Coding theorems for compound problems via quantum Rényi divergences, IEEE Trans. Inform. Theory, Vol. 61, No. 6, 2997-3012, 2015.

[52] M. Nielsen and I. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, 2000.

[53] J. Nötzel, M. Wiese, and H. Boche, The Arbitrarily Varying Wiretap Channel — Secret Randomness, Stability and Super-Activation, IEEE Trans. Inform. Theory, Vol. 62, No. 6, 3504-3531, `arXiv:1501.07439`, 2016.

[54] T. Ogawa and H. Nagaoka, Making good codes for classical-quantum channel coding via quantum hypothesis testing, IEEE Trans. Inform. Theory, Vol. 53, No. 6, 2261-2266, 2007.

[55] J. Oppenheim, For quantum information, two wrongs can make a right, Science Magazine, Vol. 321, 1783, 2008.

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

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

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

[59] B. Schumacher and M. D. Westmoreland, Approximate quantum error correction, Quant. Inf. Proc., Vol. 1, No. 8, 5-12, 2002.

[60] P. W. Shor, The quantum channel capacity and coherent information, Lecture Notes, MSRI Workshop on Quantum Computation, 2002.

[61] G. Smith, J. A. Smolin, and J. Yard, Quantum communication with Gaussian channels of zero quantum capacity, Nature Photonics, Vol. 5, 624-627, 2011.

[62] G. Smith and J. Yard, Quantum communication with zero-capacity channels, Science Magazine, Vol. 321, No. 5897, 1812-1815, 2008.

[63] W. F. Stinespring, Positive functions on C*-algebras, Proc. Amer. Math. Soc., Vol. 6, 211, 1955.

[64] S. Watanabe, Private and quantum capacities of more capable and less noisy quantum channels, Phys. Rev. A, Vol. 85, 012326, 2012.

[65] M. Wiese, J. Nötzel, 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, 3844-3862, arXiv:1410.8078, 2016.

[66] M. Wilde, Quantum Information Theory, Cambridge University Press, 2013.

[67] A. Winter, Coding theorem and strong converse for quantum channels, IEEE Trans. Inform. Theory, Vol. 45, No. 7, 2481-2485, 1999.

[68] A. D. Wyner, The wire-tap channel, Bell System Technical Journal, Vol. 54, No. 8, 1355-1387, 1975.

[69] Quantum information problem page of the ITP Hannover, http://qig.itp.uni-hannover.de/qiproblems/11