Quantum Action-Dependent Channels

Michael Korenberg and Uzi Pereg

*Faculty of Electrical and Computer Engineering, Technion
†Helen Diller Quantum Center, Technion
Email: mkorenberg@campus.technion.ac.il, uzipereg@technion.ac.il

Abstract

We study the quantum action-dependent channel. The model can be viewed as a quantum analog of the classical action-dependent channel model. In this setting, the communication channel has two inputs: Alice's transmission and the input environment. The action-dependent mechanism enables the transmitter to influence the channel's environment through an action channel. Specifically, Alice encodes her message into a quantum action, which subsequently affects the environment state. For example, a quantum measurement at the encoder can induce a state collapse of the environment. In addition, Alice has access to side information. Unlike the classical model, she cannot have a copy of the environment state due to the no-cloning theorem. Instead, she shares entanglement with this environment. We establish an achievable communication rate for reliable message transmission via the quantum action-dependent channel, thereby extending the classical action-dependent framework to the quantum domain.

I. Introduction

A fundamental problem in information theory is the characterization of reliable communication over channels affected by *random parameters*, often referred to as channel states [1–8]. Beginning with Shannon's seminal work on channels with side information [9], the study of channels with random parameters has revealed the crucial role of side information at the encoder or decoder. Gel'fand and Pinsker [10] established the capacity for channels with non-causal side information at the encoder. Costa's "writing on dirty paper" result [11] further extended it to Gaussian channels. These works lay the foundation for a rich literature on random-parameter models in both point-to-point and multi-user settings [12–15].

In the random-parameter paradigm, the parameters are typically drawn from nature and cannot be controlled by the communicating parties [16]. The parameters influence the channel by altering its transition law. In the classical setting, the channel is described in terms of a probability function $P_{Y|X,S}(\cdot|x,s)$, where X represents Alice's transmission and Y is Bob's observation at the channel output. The random parameter S has a specified distribution and its variation can significantly affect the channel output. For this reason, side information, i.e., the knowledge of S, has a profound effect on capacity.

Weissman [17] introduced the *action-dependent channel*. In this model, the encoder first selects an action sequence, which in turn, generates the channel parameters in a noisy fashion. The overall channel input then depends on both the message and the induced parameters. This two-stage procedure captures a broad class of practical problems, such as memories with defects [18], magnetic recording with rewriting [19], and other scenarios in which the transmitter can probe or partially control the channel before communication, as was demonstrated in [20]. For example, a two-stage coding strategy can steer a defective memory to improve reliability: first, the transmitter writes to the memory and immediately tries to read it back to learn about defects. Then, the transmitter rewrites the defective bits based on that information. The flexibility of this model led to its extension in several directions. For instance, the concept of probing capacity was introduced to quantify the maximum rate at which the channel state can be learned through actions [21], and other works incorporated cost constraints on these actions [22]. The framework has also proven valuable in multi-user communication, with generalizations to broadcast channels [23, 24] and multiple-access channels [25]. Furthermore, its implications for security have been explored in the context of wiretap channels [26, 27] and other secure communication settings [28], where the action channel acts as a broadcast channel influenced by the transmitter's actions.

The ongoing development of quantum information theory is foundational for engineering next-generation communication and computation systems [29–33]. By leveraging the principles of quantum mechanics, this field aims to overcome the limitations of classical technologies. Quantum technology also unlocks entirely new phenomena with no classical parallel, such as entanglement, i.e., the strongest resource of quantum correlation, as well as the no-cloning theorem, which forbids the perfect duplication of quantum information, motivating a deeper study of fundamental communication limits. Action dependence is relevant in quantum communication as well. For example, a quantum measurement by the encoder on the transmission system could result in a state collapse of the channel input environment.

Quantum environment-dependent channels are crucial in scenarios that involve not only the transmission of a message, but also parameter estimation, a central task in fields such as quantum metrology. These types of quantum Gel'fand-Pinsker channels have been studied, both with and without entanglement assistance [34] (see also [35]). The security implications have also been explored in the quantum setting through wiretap channels [36] and covert communication [37]. Other variations include scenarios in which the decoder performs parameter estimation [38]. The action-dependent framework has not yet been studied in the quantum literature thus far.

In this paper, we study the quantum action-dependent channel. Our action dependence model modifies the standard environment-dependent paradigm by allowing the transmitter's actions to influence a quantum environment, which in turn

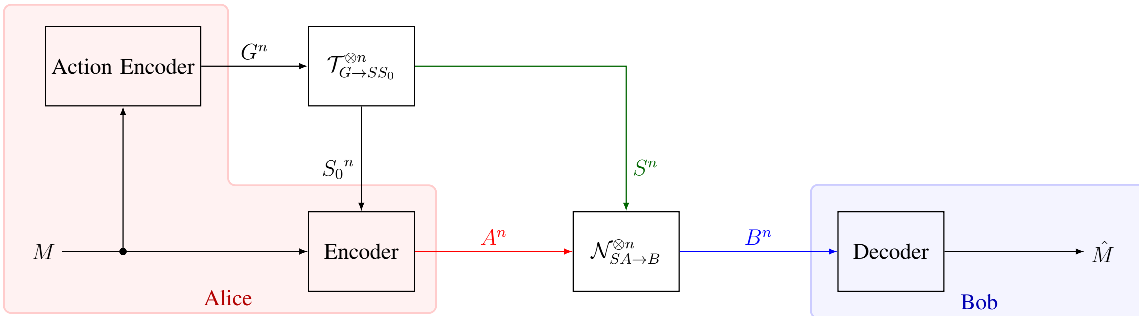

Fig. 1. Coding over a quantum action-dependent channel. Here Alice acts as the Action encoder, encoding the message M into an action sequence G^n, and the main encoder, encoding the message and side information S_0^n into the channel input A^n. The action sequence G^n is fed into the action channel $\mathcal{T}^{\otimes n}_{G\to SS_0}$, which produces the environment state S^n and side-information S_0^n for Alice. The quantum communication channel $\mathcal{N}^{\otimes n}_{SA\to B}$ takes the environment state S^n and input A^n, producing the output B^n, which is measured by Bob to decode the message.

governs the channel transformation. Specifically, an encoder (Alice) first encodes a classical message M into an action sequence G^n, which is fed into a quantum action channel $\mathcal{T}^{\otimes n}_{G\to SS_0}$. Figure 1. The action channel produces side information S_0^n, for Alice, and an environment system S^n. Alice then encodes the message and the side information into her transmission A^n, which is sent through the quantum communication channel $\mathcal{N}^{\otimes n}_{SA\to B}$. The receiver (Bob) obtains the output sequence B^n, and performs a measurement in order to estimate Alice's message.

Our framework can be viewed as the quantum counterpart of the classical action-dependent channel introduced by Weissman [17]. However, the generalization is nontrivial due to fundamental quantum principles such as the no-cloning theorem. While a classical channel parameter can be perfectly copied and then sent back to the transmitter, an unknown quantum state cannot. Thereby, side information is modeled through quantum entanglement shared between two distinct systems, S and S_0, where S represents the environment affecting the channel, and S_0 is the side information available to Alice.

We derive an achievable rate using quantum one-shot information-theoretic methods. These techniques establish non-asymptotic performance bounds by directly analyzing the error probability for a finite number of channel uses, rather than relying on asymptotic arguments. The introduction of action dependence makes our analysis more challenging than in previous environment-dependent channel models [36, 38]. In previous settings, the channel environment and side information are set by an outside source. Whereas here, the transmitter's action *induces* the shared entangled state. Consequently, our analysis must account for this additional layer of control. The formula of our capacity bound thus includes optimization over not only the input state, but also the a quantum ensemble for the sender's action.

The rest of the paper is organized as follows. In Section II, we introduce the notation and definitions used throughout the paper. In Section III, we formally define the quantum action-dependent channel model and then present our main result, an achievable rate for this channel, in Section IV. In Section V, we describe the one-shot coding scheme used to prove the achievability result. The detailed proof is provided in Section VI, with key lemmas proved in Appendices C–A.

II. Notation and Basic Definitions

We use a standard notation in quantum information theory. Quantum systems are denoted by uppercase letters (e.g., A, B) and their corresponding finite-dimensional Hilbert spaces by $\mathcal{H}_A, \mathcal{H}_B$. The set of density operators on $\mathcal{H}_A$ is $\mathscr{D}(\mathcal{H}_A)$. Quantum states (density operators) are denoted by Greek letters, e.g., ρ, σ. A POVM is a set of positive semi-definite operators $\{D_m\}$ that satisfy $\sum_m D_m = \mathbb{1}$, where $\mathbb{1}$ denotes the identity operator. If the quantum state before the measurement is ρ, then the probability of an outcome m is $\Pr(m) = \mathrm{Tr}(D_m\rho)$.

A quantum channel $\mathcal{N}_{A\to B}$ is a completely positive trace-preserving (CPTP) map. We write id_A for the identity channel on system A. Here, we consider a quantum channel $\mathcal{N}_{SA\to B}$, where A, S, and B are associated with the transmitter (Alice), the channel environment ("channel state"), and the receiver (Bob). The channel can be represented through its Stinespring dilation, in terms of an isometry $V_{SA\to BE}$ that couples the output system to an auxiliary environment E. Namely,

$$\mathcal{N}_{SA\to B}(\rho_{SA}) = \mathrm{Tr}_E\left[V\rho_{SA}V^\dagger\right]. \tag{1}$$

We assume that the channel is memoryless. That is, if an input sequence $(S^n, A^n) \equiv (S_1, A_1), \dots, (S_n, A_n)$ is transmitted through the channel, then the input state $\rho_{S^nA^n}$ undergoes the tensor-product map $\mathcal{N}^{\otimes n}_{SA\to B}$. We will see that in the action-dependent model, the channel environment S^n is affected by the encoding operation. See Section V.

For a quantum state $\rho \in \mathscr{D}(\mathcal{H})$, the von Neumann entropy is

$$H(\rho) = -\mathrm{Tr}(\rho\log\rho). \tag{2}$$

For a bipartite state $\rho_{AB} \in \mathscr{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the quantum mutual information is defined as:

$$I(A;B)_\rho = H(\rho_A) + H(\rho_B) - H(\rho_{AB}), \tag{3}$$

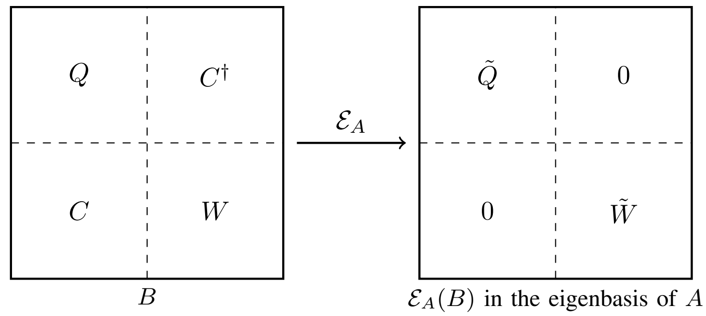

Fig. 2. Effect of pinching on a quantum state. *Left:* Matrix representation of B. Off-diagonal blocks (regions labeled C and $C^\dagger$) indicate non-commutativity with A. *Right:* After applying the pinching map $\mathcal{E}_A$, the modified state $\mathcal{E}_A(B)$ is block-diagonal in the eigenbasis of A (off-diagonal blocks are zero). Now $\mathcal{E}_A(B)$ commutes with A, enabling a measurement comparison of their spectra.

and the conditional entropy as $H(A|B)_\rho = H(\rho_{AB}) - H(\rho_B)$. Unlike its classical counterpart, the quantum conditional entropy can be negative.

The quantum relative entropy between two states ρ and σ in $\mathscr{D}(\mathcal{H})$ is defined as $D(\rho\|\sigma) = \operatorname{Tr}(\rho(\log\rho - \log\sigma))$ if $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$, and $D(\rho\|\sigma) = +\infty$ otherwise. The Sandwiched Rényi Divergence [39] is defined as

$$\tilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\operatorname{Tr}\left[\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right]^\alpha. \tag{4}$$

Other key definitions are given below [40]:

1) **Pinching**: For a Hermitian operator $A = \sum_i a_i \Pi_i$ with projectors Π_i in its eigenspaces, the pinching map is

$$\mathcal{E}_A(B) := \sum_i \Pi_i B \Pi_i. \tag{5}$$

This operation has the properties of a quantum channel and projects B onto a block-diagonal structure dictated by the eigenspaces of A (see Figure 2). One of the properties of the pinching map is that the resulting operator, $\mathcal{E}_A(B)$, always commutes with A, i.e., $[A, \mathcal{E}_A(B)] = 0$. Another key property of pinching map is the pinching inequality. Let ν_A be the number of nonnegative distinct eigenvalue of A, then:

$$B \leq \nu_A \mathcal{E}_A(B). \tag{6}$$

2) **Fidelity**: For $\rho, \sigma \in \mathscr{D}(\mathcal{H})$,

$$F(\rho,\sigma) := \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1. \tag{7}$$

3) **Purified Distance**: For $\rho, \sigma \in \mathscr{D}(\mathcal{H})$,

$$P(\rho,\sigma) := \sqrt{1 - F^2(\rho,\sigma)}. \tag{8}$$

These quantities provide a geometric measure of distance in the space of density matrices, as illustrated in Figure 2.

III. Action-Dependent Coding

Before presenting our main results, we introduce a code for the transmission of messages via a quantum action-dependent channel, where the encoder selects an action that affects the channel environment. Specifically, Alice has two roles: she encodes both the action G and the transmission A through the channel. Her action encoder sends G through an action channel $\mathcal{T}_{G\to SS_0}$ that produces Alice's side information S_0 and the channel environment S.

Definition 1 (Action-Dependent Code). An (M, n) code for communication over a quantum action-dependent channel $\mathcal{N}_{SA\to B}$, that is governed by an action channel $\mathcal{T}_{G\to SS_0}$, consists of:

1) An *encoder* that comprises two stages:
   - action encoder that prepares a quantum action state $\rho^{(m)}_{G^n} \in \mathscr{D}(\mathcal{H}_G^{\otimes n})$, for $m \in \{1,\dots,M\} = \mathsf{M}$.
   - transmission encoder $\mathcal{E}^{(m)}_{S_0^n\to A^n}$ that receives the side-information S_0^n and prepares the channel input A^n.
2) A *decoding measurement*, i.e., a POVM $\{D_m\}_{m=1}^M$ on the output Hilbert space $\mathcal{H}_B^{\otimes n}$.

The coding scheme works as shown in Figure 1. Alice selects a uniform message $m \in \mathsf{M}$. She first prepares the action state $\rho^{(m)}_{G^n}$. The action system G^n then sent through the action channel, producing

$$\rho^{(m)}_{S^n S_0^n} = \mathcal{T}^{\otimes n}_{G\to SS_0}(\rho^{(m)}_{G^n}). \tag{9}$$

Given the side information S_0^n, Alice applies the transmission encoder:

$$\rho^{(m)}_{S^nA^n} = \mathrm{id}_{S^n} \otimes \mathcal{E}^{(m)}_{S_0^n \to A^n}(\rho^{(m)}_{S^nS_0^n}). \tag{10}$$

Both the input environment S^n and Alice's transmission A^n are fed into the channel $\mathcal{N}^{\otimes n}_{SA\to B}$, hence

$$\rho^{(m)}_{B^n} = \mathcal{N}^{\otimes n}_{SA\to B}(\rho^{(m)}_{S^nA^n}). \tag{11}$$

Bob receives B^n. He performs the measurement $\{D_m\}$ to obtain an estimate of Alice's message.

For an $(\mathsf{M}, n, \varepsilon)$ code, *average probability of error* is bounded by ε, i.e.,

$$\bar{p}_e^{(n)} := 1 - \frac{1}{M}\sum_{m\in\mathsf{M}} \mathrm{Tr}\Big[D_m\, \rho^{(m)}_{B^n}\Big] \le \varepsilon. \tag{12}$$

Definition 2 (Achievable Rate). A communication rate R is said to be achievable for the quantum action-dependent channel $\mathcal{N}_{SA\to B}$, with respect to the action channel $\mathcal{T}_{G\to SS_0}$, if for every $\varepsilon, \delta > 0$ and sufficiently large n, there exists a $(2^{n(R-\delta)}, n, \varepsilon)$ code. The channel capacity C_{QAD} is defined as the supremum of all achievable rates, where the subscript 'QAD' indicates the quantum action dependence.

IV. Main Result

We now state our main result, an achievable rate for the quantum action-dependent channel.

Theorem 1 (Achievable Rate). The following rate is achievable for the quantum action-dependent channel:

$$\mathsf{R}_{\text{low}} = I(VU;B)_\rho - I(V;S|U)_\rho, \tag{13}$$

with respect to a classical auxiliary pair $(V,U) \sim p_{VU}$, a state collection $\{\sigma^u_G\}$, and an encoding channel $\mathcal{F}^v_{S_0\to A}$, such that

$$\sigma^u_{SS_0} = \mathcal{T}_{G\to SS_0}(\sigma^u_G), \tag{14}$$

$$\rho^{v,u}_{SA} = \mathrm{id}_S \otimes \mathcal{F}^v_{S_0\to A}(\sigma^u_{SS_0}), \tag{15}$$

$$\rho_{VUSA} = \sum_{u,v} p_{V,U}(v,u)\, |v\rangle\langle v|_V \otimes |u\rangle\langle u|_U \otimes \rho^{v,u}_{SA}, \tag{16}$$

where $\rho^u_V = \sum_{v\in\mathcal{V}} p_{V|U}(v|u)\,|v\rangle\langle v|$. Hence, $\rho_{VUB} = \mathrm{id}_{VU} \otimes \mathcal{N}_{SA\to B}(\rho_{VUSA})$. Equivalently, the capacity of the quantum action-dependent channel $\mathcal{N}_{SA\to B}$ satisfies

$$C_{\text{QAD}} \ge \max_{p_{VU}\,,\, \sigma^u_G\,,\, \mathcal{F}^v_{S_0\to A}} [I(VU;B)_\rho - I(V;S|U)_\rho]. \tag{17}$$

Remark 1. Previous work has considered side information when the quantum state σ_{SS_0} is fixed and dictated by the model [36–38]. This fixed state represents the entanglement between the environment subsystem S and the side-information subsystem S_0, which is accessible to Alice. By utilizing the side information S_0, Alice's encoder can generate entanglement between the channel input and its environment S. In this sense, we can regard Alice as being entangled with the channel, and this is a key feature of quantum side information at the transmitter. In our model, however, the state $\sigma^u_{SS_0}$ depends on the action encoding u chosen by Alice. This added degree of freedom allows Alice to influence the channel environment S by selecting different actions. This is analogous to the classical action-dependent channel model in [17], where the channel parameter is a noisy version of Alice's action.

V. One-Shot Code Construction

In this section, we introduce a coding scheme for the one-shot setting, of $n = 1$, over the quantum action-dependent channel (see Figure 1). Let $\mathcal{N}_{SA\to B}$ be a quantum action-dependent channel. Let $\mathcal{T}_{G\to SS_0}$ be the action channel, that generates Alice's side information subsystem S_0 and the environment subsystem S. Consider the quantum states defined in Theorem 1: $\sigma^u_{SS_0}$ is the entangled state produced by the action channel $\mathcal{T}_{G\to SS_0}$, hence, ρ_{VUSA} is the channel input, and the corresponding output is $\rho_{UVB} = \mathrm{id}_{UV} \otimes \mathcal{N}_{SA\to B}(\rho_{UVSA})$.

We now describe the one-shot coding scheme.

A. Codebook

Let $R,\ R_S > 0$ denote the coding rates corresponding to the information and action encodings. First, the action-encoding codebook $\mathcal{C}_U := \{u(m)\}_{m\in\{1,\dots,2^R\}}$ is sampled from a random set of i.i.d. codebooks $\mathbf{C}_U$, distributed according to p_U. Then, let $\{\mathcal{C}_V(m)\}_{\{m=1,\dots,2^R\}}$ be 2^R subcodebooks, such that $\mathcal{C}_V(m) := \{v(m,1), v(m,2), \dots, v(m, 2^{R_S})\}$, and $\{v(m,\ell)\}_{l\in\{1,\dots,2^{R_S}\}}$ are drawn independently according to $p_{V|U}(\cdot|u(m))$. Both are revealed to Alice and Bob. The overall codebook is $\mathcal{C} := \mathcal{C}_U \cup \{\mathcal{C}_V(m)\}$. We use the notation $\mathcal{C}$ for a deterministic codebook and $\mathbf{C}$ for a random codebook.

B. Encoder

Our encoding scheme consists of two parts, action encoding and message encoding:

1) **Action Encoding:** For each value u, let $\left|\sigma^u_{GK_0}\right\rangle$ be a purification of the state σ^u_G, with K_0 as a reference.
Given a message m, prepare $\left|\sigma^{u(m)}_{GK_0}\right\rangle$, and transmit G through the action channel.
Consider a Stinespring dilation of the action channel, with an isometry $T_{G\to SS_0K_1}$, where K_1 is appended the channel environment S and the side information S_0. Upon the action encoding above, this channel acts on $\left|\sigma^{u(m)}_{GK_0}\right\rangle$ to produce the joint state $\left|\psi^{u(m)}_{SS_0K_1K_0}\right\rangle$:
$$\left|\psi^{u(m)}_{SS_0K_1K_0}\right\rangle = T_{G\to SS_0K_1} \otimes \mathbb{1}_{K_0} \left|\sigma^{u(m)}_{GK_0}\right\rangle . \tag{18}$$

2) **Message Encoding:** Alice implements the encoding map $\mathcal{E}^{v(m,\ell)}_{S_0\to A}$ on the side-information subsystem S_0. Consider a Stinespring dilation with an isometry $E^{v(m,\ell)}_{S_0\to AT}$. This produces the channel input state
$$\left|\rho^{v(m,\ell),u(m)}_{SATK_1K_0}\right\rangle = (\mathbb{1}_S \otimes E^{v(m,\ell)}_{S_0\to AT} \otimes \mathbb{1}_{K_1K_0}) \left|\psi^{u(m)}_{SS_0K_1K_0}\right\rangle \tag{19}$$
Overall, Alice's state is
$$\left|\phi^m_{SATK_1K_0L}\right\rangle = \frac{1}{\sqrt{2^{R_S}}} \sum_{\ell=1}^{2^{R_S}} \left|\rho^{v(m,\ell),u(m)}_{SATK_1K_0}\right\rangle \otimes |\ell\rangle . \tag{20}$$
According to Uhlmann's Theorem [40], for every pair of purifications $|\psi\rangle_{AB}$ and $|\phi\rangle_{AC}$ of ρ_A and σ_A, respectively, there exists an isometry $W_{C\to B}$ such that $F(\rho_A, \sigma_A) = F(|\psi\rangle\langle\psi|_{AB}, W(|\phi\rangle\langle\phi|_{AC})W^\dagger)$. Then, in our case, it follows that there exists a set of isometries,
$$\{W^m_{S_0\to ATL}\}_{m\in\mathsf{M}} \in \mathscr{L}\left(\mathcal{H}_{S_0} \to \mathcal{H}_A \otimes \mathcal{H}_T \otimes \mathcal{H}_L\right) \tag{21}$$
that map from $\left|\psi^{u(m)}_{SS_0K_1K_0}\right\rangle$ to $\left|\phi^m_{SATK_1K_0L}\right\rangle$, or equivalently, from $\left|\psi^{u(m)}_{SS_0K_1K_0}\right\rangle$ to $\left|\rho^{v(m,\ell),u(m)}_{SATK_1K_0}\right\rangle$.
Using the short notation $\tilde{W}^m = \mathbb{1}_S \otimes W^m_{S_0\to ATL} \otimes \mathbb{1}_{K_1K_0}$,
$$P\left(\phi^m_{SATK_1K_0L}, \tilde{W}^m(\psi^{u(m)}_{SS_0K_1K_0})\tilde{W}^{\dagger m}\right) = P\left(\frac{1}{2^{R_S}} \sum_{\ell\in\mathsf{L}} \rho^{v(m,\ell),u(m)}_S, \sigma^{u(m)}_S\right). \tag{22}$$
Given a message m, Alice applies the isometry $W^m_{S_0\to ATL}$ on $\left|\psi^{u(m)}_{SS_0K_1K_0}\right\rangle$ and transmits A over the action-dependent channel.

C. Decoding

We would like to design a POVM measurement $\{D_m\}$ on the system B that distinguishes between the states $\{\rho^{(m)}_B\}_{m\in\mathsf{M}}$ with high probability. Let $\mathcal{E}_1$ be the pinching map associated with $\rho_{VU}\otimes\rho_B$ such that: $\rho_{VU}\otimes\rho_B = \sum_\lambda \lambda\Pi_\lambda$ is the spectral decomposition of $\rho_{VU}\otimes\rho_B$. The pinching map $\mathcal{E}_1$ is defined as: $\mathcal{E}_1(\rho) = \sum_\lambda \Pi_\lambda\rho\Pi_\lambda$, where $\{\Pi_\lambda\}$ are the orthogonal projectors onto the eigenspace of $\rho_{VU}\otimes\rho_B$. We define ν_1 as the number of distinct nonnegative eigenvalues of $\rho_{VU}\otimes\rho_B$. Now $\mathcal{E}_1(\rho_{VUB})$ is block-diagonal in the eigenbasis of $\rho_{VU}\otimes\rho_B$, thus it commutes with $\rho_{VU}\otimes\rho_B$. The pinching of ρ_{VUB} with respect to $\rho_{VU}\otimes\rho_B$ is defined as:
$$\mathcal{E}_1(\rho_{VUB}) = \sum_{\lambda=1}^{\nu_1} \Pi_\lambda \left(\sum_{v,u} p_{VU}(v,u)\, |v\rangle\langle v|_V \otimes |u\rangle\langle u|_U \otimes \rho^{v,u}_B\right) \Pi_\lambda \tag{23}$$

For two Hermitian matrices A and B, we define the projection $\{A \geq B\}$ as $\sum_{\lambda\geq 0} P_\lambda$, where the spectral decomposition of $A - B$ is given as $\sum_\lambda \lambda P_\lambda$. In this notation, P_λ is the projection to the eigenspace corresponding to the eigenvalue λ. Then, let
$$\Pi_{VUB} = \{\mathcal{E}_1\left(\rho_{VUB}\right) \geq 2^{R+R_S}\rho_{VU}\otimes\rho_B\}. \tag{24}$$
For every $m\in\mathsf{M}, \ell\in\mathsf{L}$, we define:
$$\gamma(m,\ell) = \mathrm{Tr}_{VU}\left[\Pi_{VUB}\left(|v(m,\ell)\rangle\langle v(m,\ell)| \otimes |u(m)\rangle\langle u(m)| \otimes \mathbb{1}_B\right)\right]. \tag{25}$$
Our set of POVM operators are then normalized as
$$\beta(m,\ell) = \left(\sum_{m',\ell'} \gamma(m',\ell')\right)^{-\frac{1}{2}} \gamma(m,\ell) \left(\sum_{m',\ell'} \gamma(m',\ell')\right)^{-\frac{1}{2}} \tag{26}$$
for $(m,\ell) \in \mathsf{M}\times\mathsf{L}$.

D. Error Bound

We are now ready to state our one-shot result.

Proposition 2 (One-shot error probability). Let $\alpha \in \left(0, \frac{1}{2}\right)$. Then, the average error probability is bounded by:

$$\mathbb{E}_{\mathbf{C}}[\bar{p}_e^{(1)}] \leq 12 \cdot \nu_1^{\alpha} 2^{\alpha\left[R+R_S-\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]} + \frac{2}{\alpha}\frac{\nu_2^{\alpha}}{2^{\alpha R_S}} 2^{\alpha\tilde{D}_{1+\alpha}(\rho_{VUS}\|\rho_{V-U-S})} \tag{27}$$

with

$$\rho_{V-U-S} = \sum_u p_U(u)\rho_V^u \otimes |u\rangle\langle u| \otimes \sigma_S^u \tag{28}$$

where ν_1 is the number of distinct eigenvalues of $\rho_{VU}\otimes\rho_B$, and ν_2 is the maximum number of distinct eigenvalues of $\{\sigma_S^{u(m)}\}_{\forall m\in\mathsf{M}}$.

The proof of Proposition 2 is given in Appendix C. The outline is given below.

E. Proof Outline

First, we show that

$$\mathbb{E}_{\mathbf{C}}[\bar{p}_e^{(1)}]\leq 2\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\sum_{m'\neq 1,\ell}\beta(m',\ell)\right)\hat{\Theta}_B(1)\right]\right\} + 2\mathbb{E}_{\mathbf{C}}\left\{P\left(\Theta_B(1),\hat{\Theta}_B(1)\right)^2\right\} \tag{29}$$

where $\Theta_B(m)$ Bob's output state given that the message m was sent, and $\hat{\Theta}_B(m) = \frac{1}{2^{R_S}}\sum_{\ell\in\mathsf{L}}\rho_B^{v(m,\ell),u(m)}$ is the average output state, when averaged over the subcodebook.

Intuitively, bounding the second error term requires showing the codebook average state $\Theta_B(1)$ is close to the probabilistic average state $\hat{\Theta}_B(1)$. To this end, we use the subcodebook property below:

$$\mathbb{E}_{\mathbf{C}}\left\{\tilde{D}_{1+\alpha}\left(\frac{1}{2^{R_S}}\sum_{\ell\in\mathsf{L}}\rho_S^{v(m,\ell),u(m)}\|\sigma_S^{u(m)}\right)\right\} \leq \frac{1}{\alpha\ln 2}\frac{\nu_2^{\alpha}}{2^{\alpha R_S}}2^{\alpha\tilde{D}_{1+\alpha}(\rho_{VUS}\|\rho_{V-U-S})} \tag{30}$$

where $\rho_{V-U-S} = \sum_u p_U(u)\rho_V^u \otimes |u\rangle\langle u| \otimes \sigma_S^u$ and $\alpha \in (0, \frac{1}{2})$. This property is shown in Appendix A.

As for the first error term on the right-hand side of (29), we use the Hayashi-Nagaoka inequality in order to show to bound this error term by

$$\begin{aligned}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\sum_{m'\neq 1,\ell}\beta(m',\ell)\right)\hat{\Theta}_B(1)\right]\right\} \leq{}& 2\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\mathbb{1}-\gamma(1,1)\right)\rho_B^{v(1,1),u(1)}\right]\right\} \\ &+ 4\sum_{(m',\ell)\neq(1,1)}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\gamma(m',\ell)\right)\rho_B^{v(1,1),u(1)}\right]\right\}.\end{aligned} \tag{31}$$

The bound on the first term is then obtained as a consequence of the projector properties in Appendix B:

$$\mathrm{Tr}[(\mathbb{1}-\Pi_{VUB})\,\rho_{VUB}] \leq \nu_1^{\alpha}2^{\alpha\left[R+R_S-\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]}, \tag{32}$$

$$2^{R+R_S}\,\mathrm{Tr}\left[\Pi_{VUB}\left(\rho_{VU}\otimes\rho_B\right)\right] \leq \nu_1^{\alpha}2^{\alpha\left[R+R_S-\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]}. \tag{33}$$

VI. Proof of Theorem 1

We consider the average error probability as the number of channel uses n grows to infinity. Let $\varepsilon > 0$. Let d_{VUB} be the dimension of $\mathcal{H}_V\otimes\mathcal{H}_U\otimes\mathcal{H}_B$, and d_S be the dimension of $\mathcal{H}_S$. As shown in [41, Lemma 3.9], we can bound ν_1 and ν_2 as follows:

$$\begin{aligned}\nu_1 &\leq (n+1)^{d_{VUB}-1}, \\ \nu_2 &\leq (n+1)^{d_S-1}.\end{aligned} \tag{34}$$

Based on Proposition 2 for n uses of the channel, there exists a (deterministic) codebook $\mathcal{C}$ such that:

$$\begin{aligned}\bar{p}_e^{(n)} &\leq 12\,(n+1)^{\alpha(d_{VUB}-1)}\,2^{\alpha\left[n(R+R_S)-\tilde{D}_{1-\alpha}\left(\rho_{VUB}^{\otimes n}\|\rho_{VU}^{\otimes n}\otimes\rho_B^{\otimes n}\right)\right]} \\ &\quad+\frac{2}{\alpha}(n+1)^{\alpha(d_S-1)}2^{\alpha\left[-nR_S+\tilde{D}_{1+\alpha}(\rho_{VUS}^{\otimes n}\|\rho_{V-U-S}^{\otimes n})\right]} \\ &= 2^{-n\left[-\alpha(R+R_S)-\frac{\log(12)}{n}-\frac{\alpha(d_{VUB}-1)}{n}\log(n+1)+\frac{\alpha}{n}\tilde{D}_{1-\alpha}\left(\rho_{VUB}^{\otimes n}\|\rho_{VU}^{\otimes n}\otimes\rho_B^{\otimes n}\right)\right]}\end{aligned}$$

$$+ 2^{-n\left[\alpha R_S - \frac{1}{n}\log\left(\frac{2}{\alpha}\right) - \frac{\alpha(d_S-1)}{n}\log(n+1) - \frac{\alpha}{n}\tilde{D}_{1+\alpha}(\rho_{VUS}^{\otimes n}\|\rho_{V-U-S}^{\otimes n})\right]} \tag{35}$$

Hence, the error probability tends to zero as $n \to \infty$, provided that

$$\alpha(R+R_S) < \frac{\alpha}{n}\tilde{D}_{1-\alpha}\left(\rho_{VUB}^{\otimes n}\|\rho_{VU}^{\otimes n}\otimes\rho_B^{\otimes n}\right) - \frac{\alpha(d_{VUB}-1)}{n}\log(n+1) - \frac{\log(12)}{n}, \tag{36}$$

and

$$\alpha R_S > \frac{\alpha}{n}\tilde{D}_{1+\alpha}\left(\rho_{VUS}^{\otimes n}\|\rho_{V-U-S}^{\otimes n}\right) + \frac{\alpha(d_S-1)}{n}\log(n+1) + \frac{1}{n}\log\left(\frac{2}{\alpha}\right). \tag{37}$$

In the limit of $n \to \infty$ and $\alpha \to 0$, the bounds from (36) and (37) can be simplified. The last two terms vanish as $n \to \infty$. We then apply the additivity of the sandwiched Rényi divergence, $\tilde{D}_\alpha(\rho^{\otimes n}\|\sigma^{\otimes n}) = n\tilde{D}_\alpha(\rho\|\sigma)$, and its convergence to the quantum divergence as $\alpha \to 0$ [39, 40]:

$$\begin{aligned}\lim_{\alpha\to 0}\left[\lim_{n\to\infty}\frac{1}{n}\tilde{D}_{1-\alpha}(\rho_{VUB}^{\otimes n}\|\rho_{VU}^{\otimes n}\otimes\rho_B^{\otimes n})\right] &= \lim_{\alpha\to 0}\left[\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]\\ &= D(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\\ &= I(VU;B)_\rho,\end{aligned} \tag{38}$$

and similarly,

$$\lim_{\alpha\to 0}\left[\lim_{n\to\infty}\frac{1}{n}\tilde{D}_{1+\alpha}(\rho_{VUS}^{\otimes n}\|\rho_{V-U-S}^{\otimes n})\right] = I(V;S|U)_\rho. \tag{39}$$

We deduce that the error probability satisfies $\bar{p}_e^{(n)} \le \varepsilon$ provided that

$$R + R_S < I(VU;B)_\rho - \frac{\delta}{3}, \tag{40}$$

$$R_S > I(V;S|U)_\rho + \frac{\delta}{3} \tag{41}$$

for some $\delta > 0$ and for sufficiently large n. Hence, a transmission rate R such that

$$R < I(VU;B)_\rho - I(V;S|U)_\rho - \delta, \tag{42}$$

is achievable. $\square$

Appendix A
Subcodebook Average

We begin our one-shot analysis with the subcodebook average result below (see proof outline in Subsection V-E).

Lemma 3 (see [36, Lemma 7]). Let $\rho_{VUS} = \mathrm{Tr}\left[\rho_{VUSA}\right]$ be a classical-quantum state, and $\alpha \in (0, \frac{1}{2})$. Furthermore, let $\mathcal{C}_m = \{v(m,1), \ldots v(2^{m,R_S}), u(m)\}$ be a collection of random variables such that the sequence $v(m,1), \ldots, v(m, 2^{R_S})$ is conditionally i.i.d. $\sim p_{V|U}(\cdot|u(m))$, for every given $u(m)$. We consider the following state:

$$\tau_{S|\mathcal{C}_m} \triangleq \frac{1}{2^{R_S}}\sum_{\ell\in\mathsf{L}}\rho_S^{v(m,\ell),u(m)}. \tag{43}$$

Then, there exists a constant $\nu_2 \ge 0$ such that:

$$\mathbb{E}_{\mathbf{C}}\left\{\tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m}\|\sigma_S^{u(m)})\right\} \le \frac{1}{\alpha\ln 2}\frac{\nu_2^\alpha}{2^{\alpha R_S}}2^{\alpha\tilde{D}_{1+\alpha}(\rho_{VUS}\|\rho_{V-U-S})} \tag{44}$$

for all $m \in \{1, \ldots, 2^R\}$, where $\rho_{V-U-S} = \sum_u p_U(u)\rho_V^u \otimes |u\rangle\langle u| \otimes \sigma_S^u$, and ν_2 is the maximum number of distinct eigenvalues of the states $\{\sigma_{S|u}\}_u$.

The lemma can be obtained as a consequence of [36, Lemma 7]. For completeness, we provide the full proof below.

Proof. For $\alpha \in (0,1]$ and due to concavity of $\log(\cdot)$, we obtain:

$$\begin{aligned}\mathbb{E}_{\mathbf{C}}\left\{\tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m}\|\sigma_S^{u(m)})\right\} &= \frac{1}{\alpha}\mathbb{E}_{\mathbf{C}}\left\{\alpha\tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m}\|\sigma_S^{u(m)})\right\}\\ &= \frac{1}{\alpha}\mathbb{E}_{\mathbf{C}}\left\{2^{\log_2\left(\alpha\tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m}\|\sigma_S^{u(m)})\right)}\right\}\\ &\le \frac{1}{\alpha}\log_2\left(\mathbb{E}_{\mathbf{C}}\left\{2^{\alpha\tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m}\|\sigma_S^{u(m)})}\right\}\right).\end{aligned} \tag{45}$$

Then, by substituting the definition of $\tilde{D}_{1+\alpha}(\cdot\|\cdot)$: [1]

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{C}}\left\{2^{\alpha\tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m}\|\sigma_{S|u(m)})}\right\}\\
&=\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\tau_{S|\mathcal{C}_m}\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)^{1+\alpha}\right]\right\}\\
&=\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\left(\frac{1}{2^{R_S}}\sum_{\ell\in\mathsf{L}}\rho_{S|v(m,\ell),u(m)}\right)\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)^{1+\alpha}\right]\right\}\\
&=\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\left(\frac{1}{2^{R_S}}\sum_{\ell\in\mathsf{L}}\rho_{S|v(m,\ell),u(m)}\right)\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)\right.\right.\\
&\quad\left.\left.\times\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\left(\frac{1}{2^{R_S}}\sum_{\ell'\in\mathsf{L}}\rho_{S|v(m,\ell'),u(m)}\right)\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)^{\alpha}\right]\right\}\\
&=\frac{1}{2^{R_S}}\sum_{\ell\in\mathsf{L}}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\left(\rho_{S|v(m,\ell),u(m)}\right)\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)\right.\right.\\
&\quad\left.\left.\times\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\frac{1}{2^{R_S}}\left(\rho_{S|v(m,\ell),u(m)}+\sum_{\ell'\neq\ell\in\mathsf{L}}\rho_{S|v(m,\ell'),u(m)}\right)\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)^{\alpha}\right]\right\}\\
&\leq\frac{1}{2^{R_S}}\sum_{\ell\in\mathsf{L}}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\left(\rho_{S|v(m,\ell),u(m)}\right)\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)\right.\right.\\
&\quad\left.\left.\times\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\frac{1}{2^{R_S}}\left(\rho_{S|v(m,\ell),u(m)}+\sum_{\ell'\in\mathsf{L}}\rho_{S|v(m,\ell'),u(m)}\right)\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)^{\alpha}\right]\right\},
\end{aligned}
\tag{46}
$$

where the second equality is obtained by substituting $\tau_{S|\mathcal{C}_m}$ as in (43). Recall that the sequence $v(m,\ell')$, $\ell'\in\{1,\ldots,2^{R_S}\}$, is conditionally i.i.d. $\sim p_{V|U}(\cdot|u(m))$, for every given $u(m)$.

Therefore,

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{C}}\left\{2^{\alpha\tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m}\|\sigma_{S|u(m)})}\right\}\\
&\leq\frac{1}{2^{R_S}}\sum_{\ell\in\mathsf{L}}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\left(\rho_{S|v(m,\ell),u(m)}\right)\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)\right.\right.\\
&\quad\left.\left.\times\left(\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\frac{1}{2^{R_S}}\left(\rho_{S|v(m,\ell),u(m)}+2^{R_S}\mathbb{E}_{V|u(m)}\left[\rho_{S|V,u(m)}\right]\right)\sigma_{S|u(m)}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)^{\alpha}\right]\right\}
\end{aligned}
\tag{47}
$$

We note that the inner expectation term satisfies $\mathbb{E}_{V|u(m)}\left[\rho_{S|V,u(m)}\right]=\sigma_{S|u(m)}$. Taking the expectation over $u(m)$ and $v(m,\ell)$ as well, we obtain

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{C}}\left\{2^{\alpha\tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m}\|\sigma_{S|u(m)})}\right\}\\
&\quad\leq\mathbb{E}_{V,U}\left\{\mathrm{Tr}\left[\left(\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\left(\rho_{S|V,U}\right)\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)\left(\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\frac{1}{2^{R_S}}\left(\rho_{S|V,U}+2^{R_S}\sigma_{S|U}\right)\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)^{\alpha}\right]\right\}.
\end{aligned}
\tag{48}
$$

Next, we introduce the pinching map $\mathcal{E}_2$ with respect to $\sigma_{S|u}$. Let ν_2 denote the maximum number of distinct eigenvalues of $\sigma_{S|u}$. Then, by the pinching inequality in (6),

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{C}}\left\{2^{\alpha\tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m}\|\sigma_{S|u(m)})}\right\}\\
&\leq\mathbb{E}_{V,U}\left\{\mathrm{Tr}\left[\left(\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\left(\rho_{S|V,U}\right)\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)\left(\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\frac{1}{2^{R_S}}\left(\nu_2\mathcal{E}_2(\rho_{S|V,U})+2^{R_S}\sigma_{S|U}\right)\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)^{\alpha}\right]\right\}\\
&\leq\mathbb{E}_{V,U}\left\{\mathrm{Tr}\left[\left(\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\left(\rho_{S|V,U}\right)\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\right)\left(\sigma_{S|U}{}^{\frac{-\alpha^2}{2(1+\alpha)}}\frac{1}{2^{\alpha R_S}}\left(\nu_2^{\alpha}\mathcal{E}_2^{\alpha}(\rho_{S|V,U})+2^{\alpha R_S}\sigma_{S|U}^{\alpha}\right)\sigma_{S|U}{}^{\frac{-\alpha^2}{2(1+\alpha)}}\right)\right]\right\}
\end{aligned}
\tag{49}
$$

where the second inequality holds since the operators $\mathbb{E}_2(\rho_{S|V,U})$ and $\sigma_{S|U}$ commute.

By trace cyclicity, we can write the last expression as

$$
\mathbb{E}_{V,U}\left\{\mathrm{Tr}\left[\sigma_{S|U}{}^{\alpha-\alpha}\rho_{S|V,U}\right]\right\}+\mathbb{E}_{V,U}\left\{\mathrm{Tr}\left[\sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}}\rho_{S|V,U}\;\sigma_{S|U}{}^{\frac{-(\alpha+\alpha^2)}{2(1+\alpha)}}\frac{\nu_2^{\alpha}}{2^{\alpha R_S}}\left(\mathcal{E}_2^{\alpha}(\rho_{S|V,U})\right)\sigma_{S|U}{}^{\frac{-\alpha^2}{2(1+\alpha)}}\right]\right\}
$$

[1]To avoid confusion, from now on, we are using the notation $\sigma_{S|u(m)}$ and $\rho_{S|v(m,\ell),u(m)}$ to represent the states $\sigma_S^{u(m)}$ and $\rho_S^{v(m,\ell),u(m)}$, respectively.

$$
\begin{aligned}
&= 1 + \frac{\nu_2^\alpha}{2^{\alpha R_S}} \mathbb{E}_{V,U} \left\{ \mathrm{Tr}\left[\rho_{S|V,U}\ \mathcal{E}_2^\alpha(\rho_{S|V,U})\ \sigma_{S|U}{}^{-\alpha} \right] \right\} \\
&\overset{(a)}{=} 1 + \frac{\nu_2^\alpha}{2^{\alpha R_S}} \mathbb{E}_{V,U} \left\{ \ \mathrm{Tr}\left[\rho_{S|V,U}\ \mathcal{E}_2\left(\mathcal{E}_2^\alpha(\rho_{S|V,U})\ \sigma_{S|U}{}^{-\alpha} \right) \right] \right\} \\
&\overset{(b)}{=} 1 + \frac{\nu_2^\alpha}{2^{\alpha R_S}} \mathbb{E}_{V,U} \left\{ \ \mathrm{Tr}\left[\mathcal{E}_2^{1+\alpha}(\rho_{S|V,U})\ \sigma_{S|U}{}^{-\alpha} \right] \right\} \\
&\overset{(c)}{=} 1 + \frac{\nu_2^\alpha}{2^{\alpha R_S}} \mathbb{E}_{V,U} \left\{ 2^{\alpha \tilde{D}_{1+\alpha}\left(\mathcal{E}_2(\rho_{S|V,U}) \| \mathcal{E}_2(\sigma_{S|U}) \right)} \right\} \\
&\overset{(d)}{\leq} 1 + \frac{\nu_2^\alpha}{2^{\alpha R_S}} \mathbb{E}_{V,U} \left\{ 2^{\alpha \tilde{D}_{1+\alpha}\left(\rho_{S|V,U} \| \sigma_{S|U} \right)} \right\} \\
&\overset{(e)}{=} 1 + \frac{\nu_2^\alpha}{2^{\alpha R_S}} 2^{\alpha \tilde{D}_{1+\alpha}(\rho_{VUS} \| \rho_{V-U-S})},
\end{aligned} \tag{50}
$$

where

(a) holds since the state is invariant to another application of the pinching map,

(b) since

$$
\mathrm{Tr}[\mathcal{E}_\sigma(\rho)\omega] = \mathrm{Tr}\left[\sum_\lambda \Pi_\lambda \rho \Pi_\lambda \omega \right] = \sum_\lambda \mathrm{Tr}\left[\Pi_\lambda \rho \Pi_\lambda \omega \right] = \sum_\lambda \mathrm{Tr}\left[\rho \Pi_\lambda \omega \Pi_\lambda \right] = \mathrm{Tr}\left[\rho \mathcal{E}_\sigma(\omega) \right],
$$

(c) is obtained by substituting the definition of $\tilde{D}_{1+\alpha}(\cdot\|\cdot)$, and since $\mathcal{E}_2(\sigma_{S|u}) = \sigma_{S|u}$.

(d) is obtained by the data processing inequality of $\tilde{D}_{1+\alpha}(\cdot\|\cdot)$, for $\alpha \in (-\frac{1}{2}, 0) \cup (0, \infty)$ [39, 40].

(e) is obtained by taking the expectation over V, U:

$$
\begin{aligned}
\mathbb{E}_{V,U} \left\{ 2^{\alpha \tilde{D}_{1+\alpha}\left(\rho_{S|V,U} \| \sigma_{S|U} \right)} \right\} &= \mathbb{E}_{V,U} \left\{ \mathrm{Tr}\left[\left(\sigma_{S|u}{}^{\frac{-\alpha}{2(1+\alpha)}} \rho_{S|V,U} \sigma_{S|U}{}^{\frac{-\alpha}{2(1+\alpha)}} \right)^{1+\alpha} \right] \right\} \\
&= \mathrm{Tr}\left[\sum_{v,u} p_{V,U}(v,u)\, |v\rangle\langle v| \otimes |u\rangle\langle u| \otimes \left(\sigma_{S|u}{}^{\frac{-\alpha}{2(1+\alpha)}} \rho_{S|v,u}\ \sigma_{S|u}{}^{\frac{-\alpha}{2(1+\alpha)}} \right)^{1+\alpha} \right] \\
&= \mathrm{Tr}\left[\left(\rho_{V-U-S}{}^{\frac{-\alpha}{2(1+\alpha)}} \rho_{VUS}\ \rho_{V-U-S}{}^{\frac{-\alpha}{2(1+\alpha)}} \right)^{1+\alpha} \right] \\
&= 2^{\alpha \tilde{D}_{1+\alpha}(\rho_{VUS} \| \rho_{V-U-S})}.
\end{aligned} \tag{51}
$$

We conclude that for $\alpha \in (0, \frac{1}{2})$:

$$
\begin{aligned}
\mathbb{E}_{\mathbf{C}} \left\{ \tilde{D}_{1+\alpha}(\tau_{S|\mathcal{C}_m} \| \sigma_{S|u(m)}) \right\} &\leq \frac{1}{\alpha} \log_2 \left(1 + \frac{\nu_2^\alpha}{2^{\alpha R_S}} 2^{\alpha \tilde{D}_{1+\alpha}(\rho_{VUS} \| \rho_{V-U-S})} \right) \\
&\leq \frac{1}{\alpha \ln 2} \frac{\nu_2^\alpha}{2^{\alpha R_S}} 2^{\alpha \tilde{D}_{1+\alpha}(\rho_{VUS} \| \rho_{V-U-S})},
\end{aligned} \tag{52}
$$

where the last line follows from $\log_2(1+x) \leq \frac{x}{\ln 2}$ for $x \in (-1, \infty)$. This completes the proof of the subcodebook property in Lemma 3. ☐

Appendix B
Projector Properties

Our one-shot error analysis makes use of the projector properties below (see proof outline in Subsection V-E).

Lemma 4. For every $\alpha \in \left(0, \frac{1}{2}\right)$, and $R, R_S > 0$ we have:

$$
\mathrm{Tr}[(\mathbb{1} - \Pi_{VUB})\, \rho_{VUB}] \leq \nu_1^\alpha 2^{\alpha\left[R + R_S - \tilde{D}_{1-\alpha}(\rho_{VUB} \| \rho_{VU} \otimes \rho_B) \right]}, \tag{53}
$$

$$
2^{R+R_S}\, \mathrm{Tr}\left[\Pi_{VUB}\, (\rho_{VU} \otimes \rho_B) \right] \leq \nu_1^\alpha 2^{\alpha\left[R + R_S - \tilde{D}_{1-\alpha}(\rho_{VUB} \| \rho_{VU} \otimes \rho_B) \right]}. \tag{54}
$$

Proof. We start by showing the upper bound in (53):

$$
\begin{aligned}
\mathrm{Tr}[(\mathbb{1} - \Pi_{VUB})\, \rho_{VUB}] &\overset{(a)}{=} \mathrm{Tr}[\mathcal{E}_1\left((\mathbb{1} - \Pi_{VUB})\, \rho_{VUB} \right)] \\
&\overset{(b)}{=} \mathrm{Tr}[(\mathbb{1} - \Pi_{VUB})\, \mathcal{E}_1(\rho_{VUB})] \\
&= \mathrm{Tr}\left[(\mathbb{1} - \Pi_{VUB})\, \mathcal{E}_1(\rho_{VUB})^{1-\alpha} \mathcal{E}_1(\rho_{VUB})^\alpha \right] \\
&\overset{(c)}{\leq} 2^{\alpha(R+R_S)}\, \mathrm{Tr}\left[(\mathbb{1} - \Pi_{VUB})\, \mathcal{E}_1(\rho_{VUB})^{1-\alpha} (\rho_{VU} \otimes \rho_B)^\alpha \right] \\
&\leq 2^{\alpha(R+R_S)}\, \mathrm{Tr}\left[\mathcal{E}_1(\rho_{VUB})^{1-\alpha} (\rho_{VU} \otimes \rho_B)^\alpha \right]
\end{aligned}
$$

$$
\begin{aligned}
&= 2^{\alpha(R+R_S)} \operatorname{Tr}\Big[\Big((\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\,\mathcal{E}_1(\rho_{VUB})\,(\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\Big)\\
&\qquad\times\Big((\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\,\mathcal{E}_1(\rho_{VUB})\,(\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\Big)^{-\alpha}\Big]\\
&\overset{(d)}{=} 2^{\alpha(R+R_S)} \operatorname{Tr}\Big[\Big((\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\,\rho_{VUB}\,(\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\Big)\\
&\qquad\times\Big((\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\,\mathcal{E}_1(\rho_{VUB})\,(\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\Big)^{-\alpha}\Big]
\end{aligned}
\tag{55}
$$

where

(a) holds due to the trace-preserving property of pinching,

(b) follows from the definition of Π_{VUB} in (24),

(c) from the definition of Π_{VUB} and the operator monotonicity of the function $f(x)=x^\alpha$ for $\alpha\in(0,1]$,

(d) holds since the pinching map $\mathcal{E}_1$ is with respect to the product state $\rho_{VU}\otimes\rho_B$.

Based on the pinching inequality (6), it follows that

$$
\begin{aligned}
\operatorname{Tr}[(\mathbb{1}-\Pi_{VUB})\,\rho_{VUB}] &\le 2^{\alpha(R+R_S)}\nu_1^\alpha \operatorname{Tr}\Big[\Big((\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\,\rho_{VUB}\,(\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\Big)\\
&\qquad\times\Big((\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\,\rho_{VUB}\,(\rho_{VU}\otimes\rho_B)^{\frac{\alpha}{2(1-\alpha)}}\Big)^{-\alpha}\Big]\\
&= \nu_1^\alpha 2^{\alpha\left[R+R_S-\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]}.
\end{aligned}
\tag{56}
$$

The derivation for (54) follows similar steps:

$$
\begin{aligned}
2^{R+R_S}\operatorname{Tr}[\Pi_{VUB}(\rho_{VU}\otimes\rho_B)] &= 2^{R+R_S}\operatorname{Tr}\Big[\Pi_{VUB}(\rho_{VU}\otimes\rho_B)^{1-\alpha}(\rho_{VU}\otimes\rho_B)^{\alpha}\Big]\\
&\le 2^{R+R_S}2^{-(R+R_S)(1-\alpha)}\operatorname{Tr}\Big[\Pi_{VUB}\mathcal{E}_1(\rho_{VUB})^{1-\alpha}(\rho_{VU}\otimes\rho_B)^{\alpha}\Big]\\
&\le 2^{\alpha(R+R_S)}\operatorname{Tr}\Big[\mathcal{E}_1(\rho_{VUB})^{1-\alpha}(\rho_{VU}\otimes\rho_B)^{\alpha}\Big]\\
&\le \nu_1^\alpha 2^{\alpha\left[R+R_S-\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]}.
\end{aligned}
\tag{57}
$$

□

Appendix C
Proof of Proposition 2

Let $\Theta_B(m)$ be the state Bob receives:

$$
\Theta_B(m) = \operatorname{Tr}_{TK_1K_0L}\Big[\mathcal{N}_{SA\to B}\Big(\tilde{W}^m(\psi^{u(m)}_{SS_0K_1K_0})\tilde{W}^{\dagger m}\Big)\Big]
\tag{58}
$$

given that the message m was transmitted. Furthermore, let $\hat{\Theta}_B(m)$ be the average state that Bob receives, when averaged over the subcodebook of $\mathcal{C}_V(m)$:

$$
\hat{\Theta}_B(m) = \frac{1}{2^{R_S}}\sum_{\ell\in\mathsf{L}}\rho_B^{v(m,\ell),u(m)}.
\tag{59}
$$

By the symmetry of encoding and decoding, we may assume without loss of generality that Alice sent $m=1$. Consider the pinching-based decoder $\{\beta(m,\ell)\}_{(m,\ell)\in\mathsf{M}\times\mathsf{L}}$ that has been constructed in Subsection V-C. We now bound the average error probability as follows:

$$
\begin{aligned}
\mathbb{E}_{\mathbf{C}}[\bar{p}_e^{(1)}] &= \Pr\Big(\hat{M}\neq 1\mid M=1\Big)\\
&\overset{(a)}{=} \mathbb{E}_{\mathbf{C}}\left\{\operatorname{Tr}\left[\left(\sum_{m'\neq 1,\ell}\beta(m',\ell)\right)\Theta_B(1)\right]\right\}\\
&\overset{(b)}{\le} 2\mathbb{E}_{\mathbf{C}}\left\{\operatorname{Tr}\left[\left(\sum_{m'\neq 1,\ell}\beta(m',\ell)\right)\hat{\Theta}_B(1)\right]\right\}\\
&\quad + 2\mathbb{E}_{\mathbf{C}}\left\{\left|\sqrt{\operatorname{Tr}\left[\left(\sum_{m'\neq 1,\ell}\beta(m',\ell)\right)\Theta_B(1)\right]}-\sqrt{\operatorname{Tr}\left[\left(\sum_{m'\neq 1,\ell}\beta(m',\ell)\right)\hat{\Theta}_B(1)\right]}\right|^2\right\}
\end{aligned}
$$

$$\overset{(c)}{\leq} 2\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\sum_{m'\neq 1,\ell}\beta(m',\ell)\right)\hat{\Theta}_B(1)\right]\right\} + 2\mathbb{E}_{\mathbf{C}}\left\{P\left(\Theta_B(1),\hat{\Theta}_B(1)\right)^2\right\} \tag{60}$$

where (a) follows since the error events are disjoint, (b) holds due to the following: Note that $(x-y)^2 \geq 0$ implies $(x+y)^2 \leq 2(x^2+y^2)$. Therefore, $z = (\sqrt{w}+\sqrt{z}-\sqrt{w})^2 \leq 2w + 2|\sqrt{z}-\sqrt{w}|^2$. The inequality follows by taking $z = \mathrm{Tr}\left[\left(\sum_{m'\neq 1,\ell}\beta(m',\ell)\right)\Theta_B(1)\right]$ and $w = \mathrm{Tr}\left[\left(\sum_{m'\neq 1,\ell}\beta(m',\ell)\right)\hat{\Theta}_B(1)\right]$. Then, (c) is obtained by $\sum_{m'\neq 1,\ell}\beta(m',\ell) \leq \mathbb{1}$ and $\left|\sqrt{\mathrm{Tr}[\Delta\sigma]}-\sqrt{\mathrm{Tr}[\Delta\rho]}\right| \leq P(\sigma,\rho)$ for every pair of quantum states $\rho,\sigma \in \mathscr{D}(\mathcal{H})$ and $0 \leq \Delta \leq \mathbb{1}$, based on [36, Fact 7].

Consider the first term on the right-hand side of (60):

$$\begin{aligned}
\sum_{m'\neq 1,\ell}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[(\beta(m',\ell))\,\hat{\Theta}_B(1)\right]\right\} &= \frac{1}{R_S}\sum_{\ell'}\sum_{m'\neq 1,\ell}\mathbb{E}_C\left\{\mathrm{Tr}\left[(\beta(m',\ell))\,\rho_B^{v(1,\ell'),u(1)}\right]\right\}\\
&= \sum_{m'\neq 1,\ell}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[(\beta(m',\ell))\,\rho_B^{v(1,1),u(1)}\right]\right\}\\
&\leq \mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\Big[(\mathbb{1}-\beta(1,1))\,\rho_B^{v(1,1),u(1)}\Big]\right\}\\
&\leq \mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\left[\left(\mathbb{1}-\left(\sum_{m',\ell}\gamma(m',\ell)\right)^{-\frac{1}{2}}\gamma(1,1)\left(\sum_{m',\ell}\gamma(m',\ell)\right)^{-\frac{1}{2}}\right)\rho_B^{v(1,1),u(1)}\right]\right\}\\
&\leq 2\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\Big[(\mathbb{1}-\gamma(1,1))\,\rho_B^{v(1,1),u(1)}\Big]\right\}\\
&\quad + 4\sum_{(m',\ell)\neq(1,1)}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\Big[(\gamma(m',\ell))\,\rho_B^{v(1,1),u(1)}\Big]\right\},
\end{aligned} \tag{61}$$

where the last inequality is based on the Hayashi-Nagaoka operator inequality [42]: $\mathbb{1}-(S+T)^{-\frac{1}{2}}\,S\,(S+T)^{-\frac{1}{2}} \leq 2\,(\mathbb{1}-S)+4T$, for every $0 \leq S \leq \mathbb{1}$ and $T \geq 0$ on a Hilbert space $\mathcal{H}$. In our case, we set $S = \gamma(1,1)$ and $T = \sum_{(m',\ell)\neq(1,1)}\gamma(m',\ell)$.

Our next steps follow similar considerations as in [36, 37]. Specifically, we obtain an upper bound on each term on the right-hand side of (61) by using the properties established in Appendix A and Appendix B. We bound the first term on the right-hand side of (61) as follows:

$$\begin{aligned}
&2\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\Big[(\mathbb{1}-\gamma(1,1))\,\rho_B^{v(1,1),u(1)}\Big]\right\}\\
&\overset{(a)}{=} 2\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\Big[(\mathbb{1}-\mathrm{Tr}_{VU}[\Pi_{VUB}(|v(1,1)\rangle\langle v(1,1)|\otimes|u(1)\rangle\langle u(1)|\otimes\mathbb{1}_B)])\,\rho_B^{v(1,1),u(1)}\Big]\right\}\\
&= 2\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\Big[(\mathbb{1}-\mathrm{Tr}_{VU}[\Pi_{VUB}(|v(1,1)\rangle\langle v(1,1)|\otimes|u(1)\rangle\langle u(1)|\otimes\mathbb{1}_B)])\left(\mathrm{id}_{VU}\otimes\rho_B^{v(1,1),u(1)}\right)\Big]\right\}\\
&\overset{(b)}{=} 2\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\Big[(\mathbb{1}-\Pi_{VUB}(|v(1,1)\rangle\langle v(1,1)|\otimes|u(1)\rangle\langle u(1)|\otimes\mathbb{1}_B))\left(\mathrm{id}_{VU}\otimes\rho_B^{v(1,1),u(1)}\right)\Big]\right\}\\
&= 2\left[\mathrm{Tr}\Big[(\mathbb{1}-\Pi_{VUB})\,\mathbb{E}_{\mathbf{C}}\left\{|v(1,1)\rangle\langle v(1,1)|\otimes|u(1)\rangle\langle u(1)|\otimes\rho_B^{v(1,1),u(1)}\right\}\Big]\right]\\
&= 2\left[\mathrm{Tr}\left[(\mathbb{1}-\Pi_{VUB})\left(\sum_{v,u}p_{VU}(v,u)\,|v\rangle\langle v|\otimes|u\rangle\langle u|\otimes\rho_B^{v,u}\right)\right]\right]\\
&\overset{(c)}{=} 2\,\mathrm{Tr}[(\mathbb{1}-\Pi_{VUB})\,\rho_{VUB}]\\
&\overset{(d)}{\leq} 2\cdot\nu_1^{\alpha}2^{\alpha\left[R+R_S-\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]}.
\end{aligned} \tag{62}$$

where

- (a) is obtained by the definition of $\gamma(1,1)$ as in (25),
- (b) follows from trace linearity,
- (c) from taking the expectation with respect to the random codebook $\mathbf{C}$, and
- (d) from Lemma 4.

As for the second term in (61):

$$\begin{aligned}
&4\sum_{(m',\ell)\neq(1,1)}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\Big[(\gamma(m',\ell))\,\rho_B^{v(1,1),u(1)}\Big]\right\}\\
&\overset{(a)}{=} 4\sum_{(m',\ell)\neq(1,1)}\mathbb{E}_{\mathbf{C}}\left\{\mathrm{Tr}\Big[(\mathrm{Tr}_{VU}[\Pi_{VUB}(|v(m',\ell)\rangle\langle v(m',\ell)|\otimes|u(m')\rangle\langle u(m')|\otimes\mathbb{1}_B)])\,\rho_B^{v(1,1),u(1)}\Big]\right.
\end{aligned}$$

$$
\begin{aligned}
&= 4 \sum_{(m',\ell)\neq(1,1)} \mathbb{E}_{\mathbf{C}} \left\{ \mathrm{Tr}\Big[\Big(\Pi_{VUB}(|v(m',\ell)\rangle\langle v(m',\ell)| \otimes |u(m')\rangle\langle u(m')| \otimes \rho_B^{v(1,1),u(1)}\Big)\Big]\right\} \\
&\overset{(b)}{=} 4 \sum_{(m',\ell)\neq(1,1)} \mathrm{Tr}\left[\mathbb{E}_{\mathbf{C}} \left\{ \Big(\Pi_{VUB}(|v(m',\ell)\rangle\langle v(m',\ell)| \otimes |u(m')\rangle\langle u(m')| \otimes \rho_B^{v(1,1),u(1)}\Big)\right\}\right] \\
&= 4 \sum_{(m',\ell)\neq(1,1)} \mathrm{Tr}\ \Pi_{VUB} \left(\sum_{v,v',u,u'} \Pr\left(v(1,1)=v, v(m',\ell)=v', u(1)=u, u(m')=u'\right) |v'\rangle\langle v'| \otimes |u'\rangle\langle u'| \otimes \rho_B^{v,u} \right) \\
&= 4 \sum_{(m',\ell)\neq(1,1)} \mathrm{Tr}\left[\Pi_{VUB}\left(\sum_{v,v',u,u'} p_{VU}(v,u) p_{VU}(v',u')\, |v'\rangle\langle v'| \otimes |u'\rangle\langle u'| \otimes \rho_B^{v,u} \right)\right] \\
&= 4 \sum_{(m',\ell)\neq(1,1)} \mathrm{Tr}\left[\Pi_{VUB}\left(\sum_{v',u'} p_{VU}(v',u')\, |v'\rangle\langle v'| \otimes |u'\rangle\langle u'| \right) \otimes \left(\sum_{v,u} p_{VU}(v,u) \rho_B^{v,u} \right)\right] \\
&\overset{(c)}{=} 4 \sum_{(m',\ell)\neq(1,1)} \mathrm{Tr}\left[\Pi_{VUB}\left(\rho_{VU} \otimes \rho_B\right)\right] \\
&\overset{(d)}{\leq} 4 \cdot 2^{R+R_S}\, \mathrm{Tr}\left[\Pi_{VUB}\left(\rho_{VU} \otimes \rho_B\right)\right] \\
&\overset{(e)}{\leq} 4 \cdot \nu_1^{\alpha} 2^{\alpha\left[R+R_S-\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]}.
\end{aligned} \tag{63}
$$

where

- (a) holds by the definition of $\gamma(m,\ell)$ in (25),
- (b) by linearity,
- (c) follows by taking the expectation with respect to the random codebook,
- (d) is obtained due to the fact that $v(m,\ell)$ is conditionally independent of $v(m',\ell')$, given $u(m)$, for $(m,\ell)\neq(m',\ell')$,
- (e) follows from Lemma 4.

By plugging (62)-(63) into (61), we obtain the following bound on the confusion error term:

$$
\sum_{m'\neq 1,\ell} \mathbb{E}_{\mathbf{C}} \left\{ \mathrm{Tr}\left[\left(\beta(m',\ell)\right)\hat{\Theta}_B(1)\right]\right\} \leq 6 \cdot \nu_1^{\alpha} 2^{\alpha\left[R+R_S-\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]}. \tag{64}
$$

Hence, the expected error probability is bounded by

$$
\mathbb{E}_{\mathbf{C}}[\bar{p}_e^{(1)}] \leq 12 \cdot \nu_1^{\alpha} 2^{\alpha\left[R+R_S-\tilde{D}_{1-\alpha}(\rho_{VUB}\|\rho_{VU}\otimes\rho_B)\right]} + 2\mathbb{E}_{\mathbf{C}}\left\{ P\left(\Theta_B(1), \hat{\Theta}_B(1)\right)^2 \right\} \tag{65}
$$

by (60).

It remains to bound the last term for $\alpha \in (0, \frac{1}{2})$:

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{C}}\left\{ P\left(\Theta_B(1), \hat{\Theta}_B(1)\right)^2 \right\} \\
&\overset{(a)}{=} \mathbb{E}_{\mathbf{C}}\left\{ P\left(\mathrm{Tr}_{TK_1K_0L}\left[\mathcal{N}_{SA\to B}\left(\tilde{W}^{(1)}(\psi^{u(1)}_{SS_0K_1K_0})\tilde{W}^{\dagger(1)}\right)\right], \frac{1}{2^{R_S}} \sum_{\ell\in\mathsf{L}} \rho_B^{v(1,\ell),u(1)} \right)^2 \right\} \\
&= \mathbb{E}_{\mathbf{C}}\left\{ P\left(\mathrm{Tr}_{TK_1K_0L}\left[\mathcal{N}_{SA\to B}\left(\tilde{W}^{(1)}(\psi^{u(1)}_{SS_0K_1K_0})\tilde{W}^{\dagger(1)}\right)\right], \frac{1}{2^{R_S}} \sum_{\ell\in\mathsf{L}} \mathrm{Tr}_{TK_1K_0}\left[\mathcal{N}_{SA\to B}\left(\rho^{v(1,\ell),u(1)}_{SATK_1K_0}\right)\right] \right)^2 \right\} \\
&\overset{(b)}{\leq} \mathbb{E}_{\mathbf{C}}\left\{ P\left(\phi^{(1)}_{SATK_1K_0L}, \tilde{W}^{(1)}(\psi^{u(1)}_{SS_0K_1K_0})\tilde{W}^{\dagger(1)} \right)^2 \right\} \\
&\overset{(c)}{=} \mathbb{E}_{\mathbf{C}}\left\{ P\left(\frac{1}{2^{R_S}} \sum_{\ell\in\mathsf{L}} \rho_S^{v(1,\ell),u(1)}, \sigma_S^{u(1)} \right)^2 \right\} \\
&\overset{(d)}{=} \mathbb{E}_{\mathbf{C}}\left\{ P\left(\tau_{S|\mathcal{C}_1}, \sigma_S^{u(1)} \right)^2 \right\} \\
&= \mathbb{E}_{\mathbf{C}}\left\{ 1 - F^2\left(\tau_{S|\mathcal{C}_1}, \sigma_S^{u(1)} \right) \right] \\
&\overset{(e)}{\leq} \mathbb{E}_{\mathbf{C}}\left\{ 1 - 2^{-\tilde{D}_{1+\alpha}\left(\tau_{S|\mathcal{C}_1}\|\sigma_S^{u(1)}\right)} \right\}
\end{aligned}
$$

$$\overset{(f)}{\leq} \ln 2\mathbb{E}_{\mathbf{C}}\left\{\tilde{D}_{1+\alpha}\left(\tau_{S|\mathcal{C}_1}\|\sigma_S^{u(1)}\right)\right\}$$
$$\overset{(g)}{\leq} \frac{1}{\alpha}\frac{\nu_2^{\alpha}}{2^{\alpha R_S}}2^{\alpha\tilde{D}_{1+\alpha}(\rho_{VUS}\|\rho_{V-U-S})}, \tag{66}$$

where

(a) follows by substituting the definition of $\Theta_B(m)$ and $\hat{\Theta}_B(m)$ in (58) and (59), respectively,
(b) from the fact that monotonicity of the purified distance, with respect to the quantum channel $\mathrm{Tr}_{TK_1K_0}\mathcal{N}_{SA\to B}(\cdot)$, and
(c) from (22).

Furthermore,

(d) holds as we introduce the notation $\tau_{S|\mathcal{C}_1}$ to represent for the average state: $\tau_{S|\mathcal{C}_1} \equiv \frac{1}{2^{R_S}}\sum_{\ell\in\mathcal{L}}\rho_S^{v(1,\ell),u(1)}$, and since $\sum_v p_{V|U}(v|u)\,[\mathrm{Tr}_A\,\rho_{SA}^{v,u}] = \mathrm{Tr}_{S_0}\,\sigma_{SS_0}^{u}$, and
(e) as $\tilde{D}_\alpha(\cdot\|\cdot)$ is monotonically increasing in α, thus $F^2(\rho,\sigma) = 2^{-\tilde{D}_{1/2}(\rho\|\sigma)} \geq 2^{-\tilde{D}_{1+\alpha}(\rho\|\sigma)}$ for every pair of quantum states $\rho,\sigma\in\mathscr{D}(\mathcal{H})$ and $\alpha > -\frac{1}{2}$ (see [40, Corollary 4.3]).
(f) follows from the inequality $1-2^{-x}\leq x\ln 2$, and
(g) by applying Lemma 3.

Proposition 2 then follows from (65)-(66). □

Acknowledgments

The authors thank Gerhard Kramer (TUM) for useful discussions. MK and UP were supported by ISF, Grants n. 939/23 and 2691/23, DIP n. 2032991, Ollendorff-Minerva Center n. 86160946, and HD Quantum Center n. 2033613. UP was also supported by Chaya Chair n. 8776026, and VATAT Program for Quantum Science and Technology n. 86636903.

References

[1] S. S. Pradhan, A. Padakandla, and F. Shirani, "An algebraic and probabilistic framework for network information theory," *Found. Trends® Commun. Inf. Theory*, vol. 18, no. 2, pp. 173–379, 2020.
[2] H. ZivariFard, M. R. Bloch, and A. Nosratinia, "Keyless covert communication via channel state information," *IEEE Trans. Inf. Theory*, vol. 68, no. 8, pp. 5440–5474, 2022.
[3] W. Kim, Y. Ahn, J. Kim, and B. Shim, "Towards deep learning-aided wireless channel estimation and channel state information feedback for 6G," *J. Commun. Netw.*, vol. 25, no. 1, pp. 61–75, 2023.
[4] P. Belzig, "Fully quantum arbitrarily varying channel coding for entanglement-assisted communication," in *2024 IEEE Int. Symp. Inf. Theory (ISIT)*, 2024, pp. 1011–1016.
[5] H. Boche, R. F. Schaefer, H. V. Poor, and F. H. P. Fitzek, "On the need of neuromorphic twins to detect denial-of-service attacks on communication networks," *IEEE/ACM Trans. Netw.*, vol. 32, no. 4, pp. 2875–2887, 2024.
[6] C.-X. Wang, Y. Yang, J. Huang, X. Gao, T. J. Cui, and L. Hanzo, "Electromagnetic information theory: Fundamentals and applications for 6g wireless communication systems," *IEEE Wireless Commun.*, vol. 31, no. 5, pp. 279–286, 2024.
[7] F. Luo, S. Khan, B. Jiang, and K. Wu, "Vision transformers for human activity recognition using wifi channel state information," *IEEE Internet of Things J.*, vol. 11, no. 17, pp. 28 111–28 122, 2024.
[8] Y. Chen and H. Boche, "Robust joint message and state transmission under arbitrarily varying jamming," *arXiv preprint arXiv:2501.10896*, 2025.
[9] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, vol. 2, no. 4, Oct. 1958.
[10] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Control Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.
[11] M. H. M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, pp. 439–441, May. 1983.
[12] A. Bennatan, D. Burshtein, G. Caire, and S. Shamai, "Superposition coding for side-information channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1872–1889, Apr. 2006.
[13] V. Ramachandran, S. R. B. Pillai, and V. M. Prabhakaran, "Joint state estimation and communication over a state-dependent Gaussian multiple access channel," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6743–6752, Jul. 2019.
[14] U. Pereg and Y. Steinberg, "The arbitrarily varying channel under constraints with side information at the encoder," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 861–887, 2019.
[15] M. Ahmadipour, M. Kobayashi, M. Wigger, and G. Caire, "An information-theoretic approach to joint sensing and communication," *IEEE Trans. Inf. Theory*, vol. 70, no. 2, pp. 1124–1146, 2024.
[16] G. Keshet, Y. Steinberg, N. Merhav *et al.*, "Channel coding in the presence of side information," *Found. Trends® Commun. Inf. Theory*, vol. 4, no. 6, pp. 445–586, 2008.
[17] T. Weissman, "Capacity of channels with action-dependent states," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5396–5411, Oct. 2010.
[18] Y. Kim, "Writing on dirty memory." Ph.D. dissertation, Carnegie Mellon University, USA, 2016.

[19] C. Heegard and A. A. E. Gamal, “On the capacity of computer memory with defects,” *IEEE Trans. Inf. Theory*, vol. 29, pp. 731–739, Sep. 1983.

[20] K. Kittichokechai, T. J. Oechtering, and M. Skoglund, “Multi-stage coding for channels with a rewrite option and reversible input,” in *Proc. IEEE Int. Symp. Inf. Theory*, 2012, pp. 3063–3067.

[21] H. Asnani, H. Permuter, and T. Weissman, “Probing capacity,” *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7317–7332, Nov. 2011.

[22] K. Kittichokechai, T. J. Oechtering, and M. Skoglund, “Coding with action-dependent side information and additional reconstruction requirements,” *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6355–6367, Nov. 2015.

[23] Y. Steinberg and T. Weissman, “The degraded broadcast channel with action-dependent states,” in *Proc. IEEE Int. Symp. Inf. Theory*, 2012, pp. 596–600.

[24] B. Ahmadi and O. Simeone, “On channels with action-dependent states,” in *Proc. IEEE Inf. Theory Workshop*, 2012, pp. 167–171.

[25] L. Dikstein, H. H. Permuter, and S. S. Shamai, “MAC with action-dependent state information at one encoder,” *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 173–188, Jan. 2015.

[26] B. Dai, A. J. H. Vinck, Y. Luo, and X. Tang, “Wiretap channel with action-dependent channel state information,” *Entropy*, vol. 15, no. 2, pp. 445–473, 2013.

[27] B. Dai, C. Li, Y. Liang, Z. Ma, and S. S. Shamai, “Impact of action-dependent state and channel feedback on Gaussian wiretap channels,” *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3435–3455, Jun. 2020.

[28] T. Welling, O. Günlü, and A. Yener, “Transmitter actions for secure integrated sensing and communication,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2024, pp. 2580–2585.

[29] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti, “Experimental demonstration of long-distance continuous-variable quantum key distribution,” *Nature Photonics*, vol. 7, no. 5, p. 378, 2013.

[30] A. Orieux and E. Diamanti, “Recent advances on integrated quantum communications,” *J. Opt.*, vol. 18, no. 8, p. 083002, Aug. 2016.

[31] L. Petit *et al.*, “Universal quantum logic in hot silicon qubits,” *Nature*, vol. 580, no. 7803, pp. 355–359, Apr. 2020.

[32] C. Wang and A. Rahman, “Quantum-enabled 6G wireless networks: Opportunities and challenges,” *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 58–69, 2022.

[33] R. Bassoli, F. H. Fitzek, and H. Boche, “Quantum communication networks for 6G,” *Photon. Quantum Technol. Sci. Appl.*, vol. 2, pp. 715–736, 2023.

[34] F. Dupuis, “The capacity of quantum channels with side information at the transmitter,” in *Proc. IEEE Int. Symp. Inf. Theory*, 2009, pp. 948–952.

[35] U. Pereg, “Entanglement-assisted capacity of quantum channels with side information,” 2019, arXiv:1909.09992.

[36] A. Anshu, M. Hayashi, and N. A. Warsi, “Secure communication over fully quantum Gel’fand-Pinsker wiretap channel,” *IEEE Trans. Inf. Theory*, vol. 66, no. 9, pp. 5548–5566, Jun. 2020.

[37] H. ZivariFard, R. A. Chou, and X. Wang, “Covert communication with positive rate over state-dependent quantum channels,” in *Proc. IEEE Inf. Theory Workshop (ITW)*, 2024, pp. 711–716.

[38] U. Pereg, “Communication over quantum channels with parameter estimation,” *IEEE Trans. Inf. Theory*, vol. 68, no. 1, pp. 359–383, Jan. 2022.

[39] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, “On quantum Rényi entropies: A new generalization and some properties,” *J. Math. Phys.*, vol. 54, no. 12, Dec. 2013.

[40] M. Tomamichel, *Quantum Information Processing with Finite Resources: Mathematical Foundations*. Springer-Verlag, 2015.

[41] M. Hayashi, *Quantum Information Theory: Mathematical Foundation*, 2nd ed. Springer, 2017.

[42] M. Hayashi and H. Nagaoka, “General formulas for capacity of classical-quantum channels,” *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1753–1768, Jun. 2003.