The Quantum Multiple-Access Channel With Cribbing Encoders

Uzi Pereg, *Member, IEEE*, Christian Deppe, *Member, IEEE*, and Holger Boche, *Fellow, IEEE*

Abstract—Communication over a quantum multiple-access channel (MAC) with cribbing encoders is considered, whereby Transmitter 2 performs a measurement on a system that is entangled with Transmitter 1. Based on the no-cloning theorem, perfect cribbing is impossible. This leads to the introduction of a MAC model with noisy cribbing. In the causal and non-causal cribbing scenarios, Transmitter 2 performs the measurement before the input of Transmitter 1 is sent through the channel. Hence, Transmitter 2's cribbing may inflict a "state collapse" for Transmitter 1. Achievable regions are derived for each setting. Furthermore, a regularized capacity characterization is established for robust cribbing, *i.e.* when the cribbing system contains all the information of the channel input. Building on the analogy between the noisy cribbing model and the relay channel, a partial decode-forward region is derived for a quantum MAC with non-robust cribbing. For the classical-quantum MAC with cribbing encoders, the capacity region is determined with perfect cribbing of the classical input, and a cutset region is derived for noisy cribbing. In the special case of a classical-quantum MAC with a deterministic cribbing channel, the inner and outer bounds coincide.

Index Terms—Quantum communication, Shannon theory, multiple-access channel, cribbing, relay channel.

I. INTRODUCTION

THE multiple-access channel (MAC) is among the most fundamental and well-understood models in network communication and information theory [1], [2]. The MAC is also referred to as the uplink channel [3], since it is interpreted in cellular communication as the link from the mobiles to the base station [4], and in the satellite-based Internet of Things (IoT), from ground devices to a satellite in space [5]. Furthermore, in a wireless local area network (WLAN), the MAC represents the channel from the terminals to the access point [6, Section 3.2]. In general, the signals of different transmitters may interfere with one another. In particular, in sequential decoding, the receiver first decodes the message of one of the transmitters, while treating the other signals as noise. Then, this estimation can be used in order to reduce the effective noise for the estimation of the next message. Following this interplay, if a cognitive transmitter has access to the signal of another transmitter, this knowledge can be exploited such that the receiver will decode the messages with less noise. Such scenarios naturally arise in wireless systems of cognitive radios [7], [8] and the Internet of Things (IoT) [9]. This motivates the study of the MAC with cribbing encoders, *i.e.* the channel setting where one transmitter has access to the signal of another transmitter [10].

Cooperation in quantum communication networks has been extensively studied in recent years, following both experimental progress and theoretical discoveries [11], [12]. Quantum MACs are considered in the literature in various settings. Winter [13] derived a regularized characterization for the classical capacity region of the quantum MAC (see also [14]). Hsieh *et al.* [15] and Shi *et al.* [16] addressed the model where each transmitter shares entanglement resources with the receiver independently. Furthermore, Boche and Nötzel [17] addressed the cooperation setting of a classical-quantum MAC with conferencing encoders, where the encoders exchange messages between them in a constant rate (see also [18]). Remarkably, Leditzky *et al.* [19] have shown that sharing entanglement between transmitters can strictly increase the achievable rates for a *classical* MAC. The channel construction in [19] is based on a pseudo-telepathy game [20] where quantum strategies guarantee a certain win and outperform classical strategies. The authors of the present paper [21] have recently shown that the dual property does not hold, *i.e.* entanglement between receivers does not increase achievable rates. Related settings of the quantum MAC involve transmission of quantum information [22], [23], error exponents [24], non-additivity effects [25], security [26]–[29], and computation codes [30].

The sixth generation of cellular networks (6G) is expected to achieve gains in terms of latency, resilience, computation power, and trustworthiness in future communication systems, such as the tactile internet [31], which not only transfer data

but also control physical and virtual objects, by using quantum resources [32]. Cooperation between trusted hardware and software components in future communication systems has the potential to isolate untrusted components such that the attack surface of the communication system is substantially reduced [31], [33]. Quantum resources and cooperation offer additional advantages, in terms of performance gains for communication tasks and reduction of attack surface, and are of great potential for 6G networks [34], [35]. It is therefore interesting to investigate cooperation for quantum MACs as a technique to achieve robust efficient protocols for future applications.

The classical MAC with cribbing encoders is a channel model with two transmitters, X_1 and X_2, and a single receiver, Y, where one transmitter has access to the other's transmissions. Willems and van der Meulen [10] introduced this setting and considered different scenarios of full cribbing, *i.e.* with a perfect copy of the other sender's input. Suppose that Transmitter 2 knows X_1. Asnani and Permuter [36] pointed out that for a Gaussian channel, the full cribbing model is degenerate, as it reduces to full cooperation since a noiseless transmission of a continuous signal X_1 allows sending an infinite amount of information from Transmitter 1 to Transmitter 2. This has motivated Asnani and Permuter to consider the MAC with partial cribbing, with a "discretized" version of the other input.

Specifically, Transmitter 2 may have access to $Z = g(X_1)$, instead of X_1, where $g(\cdot)$ is a deterministic function. See [36], [37] for further details. The original work of Willems and van der Meulen [10] included different causality scenarios, where the sender(s) have the ith copy of each other's inputs, either after sending their own transmission at time i (strictly-causal cribbing), before the ith transmission (causal cribbing), or a priori before transmission takes place (non-causal cribbing). Classical cribbing is further studied in [38]–[46].

The MAC with cribbing encoders is closely-related to the relay channel [47], [48]. Even in the classical case, the capacity of the relay channel is an open problem. Savov *et al.* [14], [49] derived a partial decode-forward lower bound for the classical-quantum relay channel, where the relay encodes information in a strictly-causal manner. Recently, Ding *et al.* [50] generalized those results and established the cutset, multihop, and coherent multihop bounds. Communication with the help of environment measurement can be modeled by a quantum channel with a classical relay in the environment [51]. Considering this setting, Smolin *et al.* [52] and Winter [53] determined the environment-assisted quantum capacity and classical capacity, respectively. Savov *et al.* [49] further discussed future research directions of interest, and pointed out that quantum communication scenarios over the relay channel may have applications for the design of quantum repeaters [49, Section V.] (see also [50]). In a recent work by the authors [21], we have considered the quantum broadcast channel with conferencing receivers, and provided an information-theoretic perspective on quantum repeaters through this setting.

In this paper, we consider the quantum MAC $\mathcal{M}_{A_1 A_2 \to B}$ with cribbing encoders. In quantum communication, the description is more delicate. By the no-cloning theorem,

universal copying of quantum states is impossible. Therefore, in the view of quantum mechanics, perfect cribbing is against the laws of nature. As illustrated in Figure 1, if Encoder 1 sends A_1 through the channel, then Encoder 2 is physically prohibited from having a copy of the input state. Hence, we consider the quantum MAC with noisy cribbing, consisting of a concatenation of a cribbing channel and the communication channel (see Figure 1b). Specifically, Encoder 1 sends the input system A_1 through a cribbing channel that has two outputs, A_1' and E. Encoder 2 performs a measurement on the system E, and uses the measurement outcome in order to encode the input state of A_2. Then, the input systems A_1' and A_2 are sent through the communication channel. The model can also be interpreted as if the second transmitter performs a measurement on the environment E of the first transmitter.

The entanglement between the cribbing system E and the communication channel input A_1' has the following implication. If Encoder 2 measures E_i before $A_{1,i}'$ is sent through the channel, as in the causal and non-causal scenarios, then Encoder 2's cribbing measurement may inflict a "state collapse" of Encoder 1's transmission through the communication channel. In other words, in quantum communication, the cribbing operation interferes with Encoder 1's input before it is even transmitted through the communication channel.

We consider the scenarios of strictly-causal, causal, and non-causal cribbing. For a MAC with *robust cribbing*, the cribbing system E includes all the information that is available in A_1'. We derive achievable regions for each setting and establish a regularized capacity characterization for robust cribbing. Building on the analogy between the noisy cribbing model and the relay channel, we further develop a partial decode-forward region for the quantum MAC with strictly-causal non-robust cribbing. For the classical-quantum MAC with cribbing encoders, the capacity region is determined with perfect cribbing of the classical input, and a cutset region is derived for noisy cribbing. In the special case of a classical-quantum MAC with a deterministic cribbing channel, the inner and outer bounds coincide. The setting of noisy cribbing is significantly more challenging, as it is closely-related to the relay channel. The comparison between the relay channel and the MAC with cribbing encoders is further investigated in the present paper. We derive a generalized packing lemma for the MAC. While the lemma does not include cribbing, it is useful in the analysis of the MAC with cribbing encoders.

The paper is organized as follows. In Section II, we give the definitions of the channel model and a brief review of related work. In Section III, we provide the information-theoretic tools for the analysis and state the generalized quantum packing lemma. The main results are given in Section IV and Section V. In the former, we focus on robust cribbing, and in the latter, we introduce partial decode-forward coding for the non-robust case. The proofs are given in the appendix.
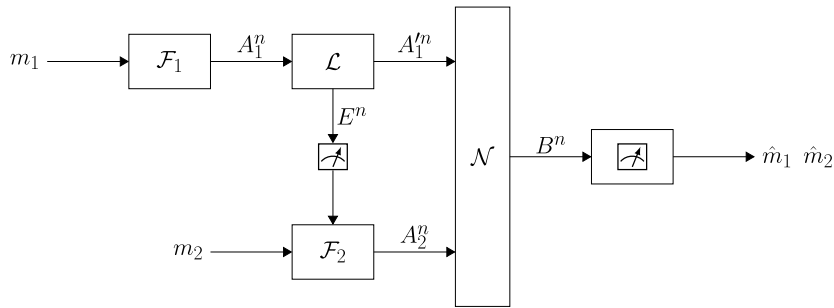
II. DEFINITIONS AND RELATED WORK

A. Notation, States, and Information Measures

We use the following notation conventions. Script letters $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \ldots$ are used for finite sets. Lowercase letters
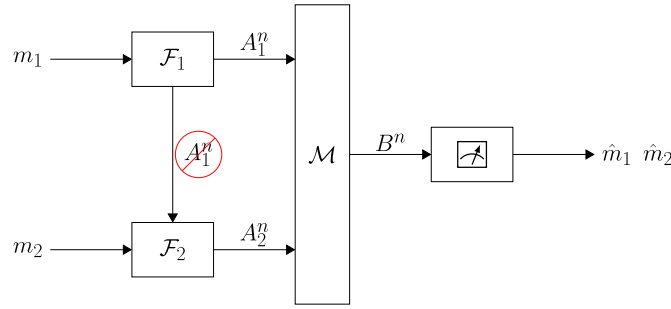
(a) Perfect cribbing violates the laws of quantum mechanics. If Encoder 1 sends A_1 through the channel, then Encoder 2 is physically prohibited from having a copy of the input state.

(b) A quantum MAC $\mathcal{N} \circ \mathcal{L}$ with noisy cribbing. Alice 1 sends the input system A_1^n through n copies of a cribbing channel $\mathcal{L}_{A_1 \to A_1' E}$, producing two outputs, $A_1'^n$ and E^n. Alice 2 performs a measurement on the system E^n, and uses the measurement outcome in order to encode the input state of A_2^n. Then, the input systems $A_1'^n$ and A_2^n are sent through n copies of the communication channel $\mathcal{N}_{A_1' A_2 \to B}$. The model can also be interpreted as if the second transmitter performs a measurement on the environment E^n of the first transmitter.

Fig. 1. The quantum multiple-access channel $\mathcal{M}_{A_1 A_2 \to B}$ with cribbing at Encoder 2.

$x, y, z, \ldots$ represent constants and values of classical random variables, and uppercase letters $X, Y, Z, \ldots$ represent classical random variables. The distribution of a random variable X is specified by a probability mass function (pmf) $p_X(x)$ over a finite set $\mathcal{X}$.

We use $x^j = (x_1, x_2, \ldots, x_j)$ to denote a sequence of letters from $\mathcal{X}$. A random sequence X^n and its distribution $p_{X^n}(x^n)$ are defined accordingly.

The state of a quantum system A is a density operator ρ on the Hilbert space $\mathcal{H}_A$. A density operator is an Hermitian, positive semidefinite operator, with unit trace, *i.e.* $\rho^\dagger = \rho$, $\rho \succeq 0$, and $\mathrm{Tr}(\rho) = 1$. The set of all density operators acting on $\mathcal{H}_A$ is denoted by $\mathscr{D}(\mathcal{H}_A)$. The state is said to be pure if $\rho = |\psi\rangle\langle\psi|$, for some vector $|\psi\rangle \in \mathcal{H}_A$. Define the quantum entropy of the density operator ρ as $H(\rho) \triangleq -\mathrm{Tr}[\rho \log(\rho)]$. Consider the state of a pair of systems A and B on the product space $\mathcal{H}_A \otimes \mathcal{H}_B$. Given a bipartite state σ_{AB}, define the quantum mutual information as

$$I(A; B)_\sigma = H(\sigma_A) + H(\sigma_B) - H(\sigma_{AB}). \tag{1}$$

Furthermore, conditional quantum entropy and mutual information are defined by $H(A|B)_\sigma = H(\sigma_{AB}) - H(\sigma_B)$ and $I(A; B|C)_\sigma = H(A|C)_\sigma + H(B|C)_\sigma - H(AB|C)_\sigma$, respectively. A quantum channel $\mathcal{P}_{A \to B}$ is a completely-positive trace-preserving (cptp) linear map from $\mathscr{D}(\mathcal{H}_A)$ to $\mathscr{D}(\mathcal{H}_B)$.

A measurement of a quantum system is specified in two equivalent manners. When the post-measurement state is irrelevant, we specify a measurement by a positive operator-valued measure (POVM), *i.e.* a set of positive semi-definite operators $\{K_x\}_{x \in \mathcal{X}}$ such that $\sum_x K_x = \mathbb{1}$. According to the Born rule, if the system is in state ρ, then the probability to measure x is $p_X(x) = \mathrm{Tr}(K_x \rho)$. More generally, a measurement is defined by a set of operators $\{W_x\}_{x \in \mathcal{X}}$ such that $\sum_x W_x^\dagger W_x = \mathbb{1}$. If the system is in the state ρ, then the measurement outcome X is distributed by $p_X(x) = \mathrm{Tr}(W_x^\dagger W_x \rho)$ for $x \in \mathcal{X}$. If $X = x$ was measured, then the post-measurement state is $\rho' \equiv (W_x \rho W_x^\dagger)/\mathrm{Tr}(W_x^\dagger W_x \rho)$. Furthermore, the quantum instrument $\mathcal{W}_{A \to \bar{A}X}$ of this measurement is the linear map from the original state, before the measurement, to the joint state of the system after the measurement with the measurement outcome, *i.e.*

$$\mathcal{W}_{A \to \bar{A}X}(\rho) = \sum_{x \in \mathcal{X}} W_x \rho W_x^\dagger \otimes |x\rangle\langle x|. \tag{2}$$

A quantum Markov chain is defined as follows. The quantum systems $A \ominus B \ominus C$ form a Markov chain if there exists a recovery channel $\mathcal{P}_{B \to BC}$ such that $\rho_{ABC} = (\mathrm{id}_A \otimes \mathcal{P}_{B \to BC})(\rho_{AB})$ [54]. In general, this holds if and only if

$$I(A; C|B)_\rho = 0 \tag{3}$$

(see [54, Theorem 5.2]).

B. Quantum Multiple-Access Channel

A quantum multiple-access channel (MAC) maps a quantum state at the senders' system to a quantum state at the receiver. Here, we consider a channel with two transmitters. Formally, a quantum MAC is a cptp map $\mathcal{M}_{A_1 A_2 \to B}$ corresponding to a quantum physical evolution. We assume that the channel is memoryless. That is, if the systems $A_1^n = (A_{1,i})_{i=1}^n$ and $A_2^n = (A_{2,i})_{i=1}^n$ are sent through n channel uses, then the joint input state $\rho_{A_1^n A_2^n}$ undergoes the tensor product mapping $\mathcal{M}_{A_1^n A_2^n \to B^n} \equiv \mathcal{M}_{A_1 A_2 \to B}^{\otimes n}$.

We will consider a MAC with cribbing, where Transmitter 2 can measure a system E that is entangled with Transmitter 1's system. Assume without loss of generality that the quantum MAC can be decomposed as

$$\mathcal{M}_{A_1 A_2 \to B}(\rho_{A_1 A_2}) = \mathrm{Tr}_E \left[(\mathcal{N}_{A_1' A_2 \to B} \circ \mathcal{L}_{A_1 \to A_1' E})(\rho_{A_1 A_2}) \right] \quad (4)$$

for all $\rho_{A_1 A_2}$, and some channel $\mathcal{L}_{A_1 \to A_1' E}$. In general, there always exists such a channel, since we can define $\mathcal{L}_{A_1 \to A_1' E}$ such that the outputs A_1' and E are in a product state and $\rho_{A_1'} = \rho_{A_1}$. However, the interesting case is when the transmitted system A_1' and the cribbing system E are correlated.

The transmitters and the receiver are often called Alice 1, Alice 2, and Bob. In the cribbing setting, Alice 1 first sends the input state of A_1^n through the channel $\mathcal{L}_{A_1^n \to A_1'^n E^n} \equiv \mathcal{L}_{A_1 \to A_1' E}^{\otimes n}$. Alice 2 gains access to the system E^n, and performs a measurement. Then, she encodes the state of her input A_2^n using the measurement outcome, and transmits it to Bob. Henceforth, we will refer to the channel $\mathcal{L}_{A_1 \to A_1' E}$ as the cribbing channel, and to E^n as the cribbing system. We denote the quantum MAC with cribbing by $\mathcal{N}_{A_1' A_2 \to B} \circ \mathcal{L}_{A_1 \to A_1' E}$.

In the sequel, we will also be interested in the following special cases.

Definition 1: Let $\mathcal{N} \circ \mathcal{L}$ be a quantum MAC with cribbing. Given an input state $\theta_{A_1 A_0}$, where A_0 is an arbitrary reference system, let $\rho_{A_1' E A_0} = \mathcal{L}_{A_1 \to A_1' E}(\theta_{A_1 A_0})$ be the output of the cribbing channel. We say that the quantum MAC is with robust cribbing if $A_0 \operatorname{\!-\!} E \operatorname{\!-\!} A_1'$ form a quantum Markov chain.

That is, for every input $\theta_{A_1 A_0}$, there exists a recovery channel $\mathcal{P}_{E \to A_1' E}^\theta$ such that

$$\rho_{A_1' E A_0} = \mathcal{P}_{E \to A_1' E}^\theta(\rho_{E A_0}). \quad (5)$$

We give the following intuition. In the cribbing setting, Alice 2 may have a noisy copy of Alice 1's input. The condition above means that Alice 2's copy is "at least as good" as the one which is transmitted through the channel $\mathcal{N}$. That is, the channel input A_1' does not contain more information than the cribbing system E, which Alice 2 measures. Next, we define the classical-quantum MAC with noiseless cribbing, which is a special case of a quantum MAC with robust cribbing.

Definition 2: The classical-quantum MAC with *noiseless* cribbing is defined by $\mathcal{N} \circ \mathcal{L}$ such that the inputs A_1, A_2 are classical, and the cribbing channel simply copies Alice 1's input, *i.e.* $A_1 \equiv X_1$, $A_2 \equiv X_2$, and $\mathcal{L} = \mathrm{id}_{X_1 \to X_1 X_1}$. Hence, the cribbing system is an exact copy of Alice 1's input, *i.e.*

$$E \equiv A_1' \equiv A_1 \equiv X_1. \quad (6)$$

We denote the classical-quantum MAC with noiseless cribbing by $\mathcal{N}_{X_1 X_2 \to B} \circ \mathrm{id}_{X_1 \to X_1 X_1}$.

If the cribbing observation is noisy, then it may not satisfy the robustness property.

Definition 3: The classical-quantum MAC with a *noisy* cribbing channel $Q_{Z|X_1}$ is defined such that the inputs A_1, A_1', A_2 and the cribbing system E are classical, *i.e.* $A_1 \equiv A_1' \equiv X_1$, $A_2 \equiv X_2$, and $E \equiv Z$, while the cribbing channel is specified by the classical noisy channel $Q_{Z|X_1}$. We denote the classical-quantum MAC with noisy cribbing by $\mathcal{N}_{X_1 X_2 \to B} \circ Q_{Z|X_1}$.

Remark 1: Another simple example of robust cribbing is the case where A_1' does not store information at all, say $|\mathcal{H}_{A_1'}| = 1$. The resulting channel is a basic multi-hop link that concatenates two point-to-point channels. Specifically, Alice 1 sends the input A_1 to Alice 2 via $\mathcal{L}_{A_1 \to E}$, Alice 2 measures E, encodes the input A_2, and send it to Bob over $\mathcal{N}_{A_2 \to B}$.

C. Coding With Cribbing

We consider different scenarios of cribbing. First, we define the setting of a MAC where Alice 2 obtains her measurements of the systems $E_1, E_2, \ldots$ in a causal manner. That is, at time i, Alice 2 can measure the *past and present* systems $E_1, \ldots, E_i$.

Definition 4: A $(2^{nR_1}, 2^{nR_2}, n)$ classical code for the quantum MAC $\mathcal{N}_{A_1' A_2 \to B} \circ \mathcal{L}_{A_1 \to A_1' E}$ with cribbing consists of the following:

- Two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$, assuming 2^{nR_k} is an integer;
- an encoding map $\mathcal{F}_1 : [1 : 2^{nR_1}] \to \mathscr{D}(\mathcal{H}_{A_1}^{\otimes n})$ for Encoder 1;
- a sequence of cribbing POVMs $\mathcal{K}_i = \{K_{E^i}^z, \ z \in \mathcal{Z}\}$, for Encoder 2;
- a sequence of n encoding maps $\mathcal{F}_{2,i} : \mathcal{Z}^i \times [1 : 2^{nR_2}] \to \mathscr{D}(\mathcal{H}_{A_2}^{\otimes i})$, where $i \in [1 : n]$, for Encoder 2; and
- a decoding POVM $\mathcal{D} = \{D_{B^n}^{m_1, m_2}\}$, where the measurement outcome is an index pair (m_1, m_2), with $m_k \in [1 : 2^{nR_k}]$ for $k = 1, 2$.

The sequence of encoding maps needs to be consistent in the sense that the states $\rho_{A_2^i}^{m_2, z^i} \equiv \mathcal{F}_{2,i}(z_1^i, m_2)$ satisfy $\mathrm{Tr}_{A_{2,i+1}^n}(\rho_{A_2^n}^{m_2, z_1^n}) = \rho_{A_2^i}^{m_2, z_1^i}$. We denote the code by $(\mathcal{F}_1, \mathcal{F}_2, \mathcal{K}, \mathcal{D})$.

The communication scheme is depicted in Figure 1b. The sender Alice 1 has the input system A_1^n, the sender Alice 2 has both A_2^n, E^n, and the receiver Bob has B^n. Alice k chooses a message m_k according to a uniform distribution over $[1 : 2^{nR_k}]$, for $k = 1, 2$. To send the message $m_1 \in [1 : 2^{nR_1}]$, Alice 1 encodes the message by $\rho_{A_1^n}^{m_1} \equiv \mathcal{F}_1(m_1)$, and sends her transmission A_1^n through n uses of the cribbing channel. Her transmission induces the following state,

$$\rho_{A_1'^n E^n}^{m_1} = \mathcal{L}^{\otimes n}(\rho_{A_1^n}^{m_1}) \quad (7)$$

where E^n is the cribbing system which will be measured by the second transmitter.

At time i, Alice 2 measures the system E^i using the measurement $\mathcal{K}_i$, and obtains a measurement outcome z_i.

To send the message $m_2 \in [1 : 2^{nR_2}]$, she prepares the input state $\mathcal{F}_{2,i}(m_2, z^i)$ in a causal manner, and sends her transmission. The average joint input state is

$$
\begin{aligned}
& \rho_{A_1'^n A_2^n}^{m_1,m_2} \\
& = \sum_{z^n \in \mathcal{Z}^n} \mathrm{Tr}_{E^n} \left[(\mathrm{id} \otimes K^{z^n}) \rho_{A_1'^n E^n}^{m_1} \right] \otimes \mathcal{F}_{2,n}(m_2, z^n) \quad (8)
\end{aligned}
$$

with $K_{E^n}^{z^n} \equiv K_{E^n}^{z_n} K_{E^{n-1}}^{z_{n-1}} \cdots K_{E_1}^{z_1}$.

Bob receives the channel output systems B^n in the following state,

$$
\rho_{B^n}^{m_1,m_2} = \mathcal{N}^{\otimes n}(\rho_{A_1^n A_2^n}^{m_1,m_2}). \quad (9)
$$

He measures with the decoding POVM $\mathcal{D}$ and obtains an estimate of the message pair $(\hat{m}_1, \hat{m}_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ from the measurement outcome. The average probability of error of the code is

$$
\begin{aligned}
& P_e^{(n)}(\mathcal{F}_1, \mathcal{F}_2, \mathcal{K}, \mathcal{D}) = \\
& \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1=1}^{2^{nR_1}} \sum_{m_2=1}^{2^{nR_2}} \left(1 - \mathrm{Tr}[D^{m_1,m_2} \rho_{B^n}^{m_1,m_2}] \right). \quad (10)
\end{aligned}
$$

A $(2^{nR_1}, 2^{nR_2}, n, \varepsilon)$ classical code satisfies $P_e^{(n)}(\mathcal{F}_1, \mathcal{F}_2, \mathcal{K}, \mathcal{D}) \le \varepsilon$.

A rate pair (R_1, R_2) is called achievable with causal cribbing if for every $\varepsilon > 0$ and sufficiently large n, there exists a $(2^{nR_1}, 2^{nR_2}, n, \varepsilon)$ code. The classical capacity region $\mathcal{C}_{\mathrm{caus}}(\mathcal{N} \circ \mathcal{L})$ of the quantum MAC with causal cribbing is defined as the set of achievable pairs (R_1, R_2), where the subscript 'caus' indicates causal cribbing.

We will also put a considerable focus on strictly-causal and non-causal cribbing. In the strictly-causal setting, Alice 2 can measure the cribbing system only after she has sent her transmission at each time instance. That is, a code with strictly-causal cribbing is defined in a similar manner, where Alice 2 first transmits her input $A_{2,i}$, and only then measures E^i. Therefore, her input state at time i can only depend on the *past* measurements z^{i-1}, and the ith encoding map has the form $\mathcal{F}_{2,i} : \mathcal{Z}^{i-1} \times [1 : 2^{nR_2}] \to \mathscr{D}(\mathcal{H}_{A_2}^{\otimes i})$. We denote the capacity region for this scenario by $\mathcal{C}_{\mathrm{s\text{-}c}}(\mathcal{N} \circ \mathcal{L})$, where the subscript 's-c' stands for strictly-causal cribbing.

With non-causal cribbing, Alice 2 gains access to the entire sequence of systems E^n a priori, *i.e.* before the beginning of her transmission. Thus, she can perform a joint measurement at time $i = 1$, with a single POVM $\mathcal{K} = \{K_{E^n}^\ell\}$, before sending A_2^n. Thereby, she can prepare a joint state $\rho_{A_2^n}^{m_2,\ell}$ of any form. The capacity region with non-causal cribbing is denoted by $\mathcal{C}_{\mathrm{n\text{-}c}}(\mathcal{N} \circ \mathcal{L})$.

Remark 2: The MAC with strictly-causal cribbing is closely related to the relay channel. In particular, if we impose $R_2 = 0$, then Alice 2 can only use the cribbing channel to enhance the transmission of information from Alice 1 to Bob. That is, Alice 2 acts as a relay in this case. In the sequel,

we will introduce methods for quantum cribbing that are inspired by relay coding techniques. It should be noted, however, that as opposed to the relay channel, our cribbing channel ('sender-relay link') is not affected by the transmission of Alice 2 (the 'relay'). Specifically, Savov *et al.* [14], [49]

considered a relay channel $\mathcal{P}_{A_1 A_2 \to EB}$, where the transmitter sends a classical input $A_1 \equiv X_1$, the relay receives E and transmits $A_2 \equiv X_2$, and the destination receiver receives the quantum output B. Hence, the state of E in the relay model is affected by both A_1 and A_2. Here, on the other hand, the cribbing channel acts only on A_1. Nonetheless, it appears that the technical challenges carry over from the relay channel to the present model.

Remark 3: Based on the definitions above, we have the following relation between the capacity regions of the MAC with strictly-causal, causal, and non-causal cribbing: $\mathcal{C}_{\mathrm{s\text{-}c}}(\mathcal{N} \circ \mathcal{L}) \subseteq \mathcal{C}_{\mathrm{caus}}(\mathcal{N} \circ \mathcal{L}) \subseteq \mathcal{C}_{\mathrm{n\text{-}c}}(\mathcal{N} \circ \mathcal{L})$. This follows because, if Alice 2 has causal access to E^i, or to the entire sequence of E^n, then she can choose to measure only E^{i-1} at time i. Hence, every code with strictly-causal cribbing can also be used given causal or non-causal cribbing. Similarly, a causal scheme can be performed given non-causal cribbing.

D. Related Work

We briefly review the capacity result for the quantum MAC without cribbing. A code without cribbing is defined in a similar manner, where Alice 2 is not allowed to measure the cribbing systems E^n. In this case, the model is fully described by $\mathcal{M}_{A_1 A_2 \to B}$, ignoring its components $\mathcal{L}_{A_1 \to A_1' E}$ and $\mathcal{N}_{A_1' A_2 \to B}$ (see (4) and Figure 1b). Denote the capacity region without cribbing by $\mathcal{C}_{\mathrm{none}}(\mathcal{M})$. Define a rate region $\mathcal{R}_{\mathrm{none}}(\mathcal{N} \circ \mathcal{L})$ as follows,

$$
\mathcal{R}_{\mathrm{none}}(\mathcal{M}) = \bigcup_{p_U p_{X_1|U} p_{X_2|U}, \, \theta_{A_1}^{x_1} \otimes \zeta_{A_2}^{x_2}}
$$
$$
\left\{ (R_1, R_2) : \begin{aligned} R_1 &\le I(X_1; B|X_2 U)_\omega \\ R_2 &\le I(X_2; B|X_1 U)_\omega \\ R_1 + R_2 &\le I(X_1 X_2; B|U)_\omega \end{aligned} \right\} \quad (11)
$$

where the union is over the joint distributions of the auxiliary random variables U, X_1, X_2 and the product state collections $\{\theta_{A_1}^{x_1} \otimes \zeta_{A_2}^{x_2}\}$, with

$$
\omega_{UX_1 X_2 B} = \sum_{u,x_1,x_2} p_U(u) p_{X_1|U}(x_1|u) p_{X_2|U}(x_2|u) |u\rangle\langle u|
$$
$$
\otimes |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \otimes \mathcal{M}_{A_1 A_2 \to B}(\theta_{A_1}^{x_1} \otimes \zeta_{A_2}^{x_2}) \quad (12)
$$

and $|\mathcal{U}| = 3$, $|\mathcal{X}_k| \le 3(|\mathcal{H}_{A_k}|^2 + 1)$, for $k = 1, 2$.

Theorem 1 (See [13], [14]): The capacity region of the quantum MAC $\mathcal{M}_{A_1 A_2 \to B}$ without cribbing satisfies

$$
\mathcal{C}_{\mathrm{none}}(\mathcal{M}) = \bigcup_{n=1}^{\infty} \frac{1}{n} \mathcal{R}_{\mathrm{none}}(\mathcal{M}^{\otimes n}). \quad (13)
$$

Remark 4: The union over p_U in the RHS of (11) can be viewed as a convex hull operation, hence the region $\mathcal{R}_{\mathrm{none}}(\mathcal{M})$ is convex. This also has the interpretation of operational time-sharing. Suppose that p_U is a type. By employing a sequence of T codes consecutively, each corresponds to a rate pair $(R_1^{(u)}, R_2^{(u)})$ over a sub-block of length $n \cdot p_U(u)$ for $u \in \{1, \ldots, T\}$, one achieves a rate pair (R_1, R_2) that corresponds to the convex combination of the rates, *i.e.* $R_k = \sum_{u \in \mathcal{U}} p_U(u) R_k^{(u)}$ for $k = 1, 2$. The time-sharing argument implies that the operational capacity region, with or without cribbing, is a convex set in general.

III. INFORMATION THEORETIC TOOLS

To derive our results, we use the quantum version of the method of types properties and techniques. In particular, we will derive a generalized quantum packing lemma for a MAC. While this packing lemma does not include cribbing, it will still be useful in the proof of our main results.

Standard method-of-types concepts are defined as in [55] and [56]. We briefly introduce the notation and basic properties while the detailed definitions can be found in [56, Appendix A]. In particular, given a density operator $\rho = \sum_x p_X(x)|x\rangle\langle x|$ on the Hilbert space $\mathcal{H}_A$, denote the strong δ-typical set that is associated with p_X by $\mathcal{A}^\delta(p_X)$, define the strong δ-typical subspace as the vector space that is spanned by $\{|x^n\rangle : x^n \in \mathcal{A}^\delta(p_X)\}$, and let $\Pi^\delta(\rho)$ be the projector onto this subspace. The following inequalities follow from well-known properties of strong δ-typical sets [57],

$$\text{Tr}(\Pi^\delta(\rho)\rho^{\otimes n}) \geq 1 - \varepsilon \tag{14}$$

$$2^{-n(H(\rho)+c\delta)}\Pi^\delta(\rho) \preceq \Pi^\delta(\rho)\rho^{\otimes n}\Pi^\delta(\rho) \preceq 2^{-n(H(\rho)-c\delta)} \tag{15}$$

$$\text{Tr}(\Pi^\delta(\rho)) \leq 2^{n(H(\rho)+c\delta)} \tag{16}$$

where $c > 0$ is a constant. Furthermore, for $\sigma_B = \sum_x p_X(x)\sigma_B^x$, let $\Pi^\delta(\sigma_B|x^n)$ denote the projector corresponding to the conditional strong δ-typical set given the sequence x^n. Similarly [55],

$$\text{Tr}(\Pi^\delta(\sigma_B|x^n)\sigma_{B^n}^{x^n}) \geq 1 - \varepsilon' \tag{17}$$

$$2^{-n(H(B|X')_\sigma+c'\delta)}\Pi^\delta(\sigma_B|x^n) \preceq \Pi^\delta(\sigma_B|x^n)\sigma_{B^n}^{x^n}\Pi^\delta(\sigma_B|x^n)$$
$$\preceq 2^{-n(H(B|X')_\sigma-c'\delta)}\Pi^\delta(\sigma_B|x^n) \tag{18}$$

$$\text{Tr}(\Pi^\delta(\sigma_B|x^n)) \leq 2^{n(H(B|X')_\sigma+c'\delta)} \tag{19}$$

where $c' > 0$ is a constant, $\sigma_{B^n}^{x^n} = \bigotimes_{i=1}^n \sigma_{B_i}^{x_i}$, and the classical random variable X' is distributed according to the type of x^n. If $x^n \in \mathcal{A}^\delta(p_X)$, then

$$\text{Tr}(\Pi^\delta(\sigma_B)\sigma_{B^n}^{x^n}) \geq 1 - \varepsilon' \tag{20}$$

as well (see [55, Property 15.2.7]).

The definition can be further generalized to the case of $\sigma_B^x = \sum_z p_{Z|X}(z|x)\hat{\sigma}_B^{x,z}$. Given a fixed pair $(x^n, z^n) \in \mathcal{X}^n \times \mathcal{Z}^n$, divide the index set $[1 : n]$ into the subsets $I(a,b|x^n,z^n) = \{i : (x_i, z_i) = (a,b)\}$, for $(a,b) \in \mathcal{X} \times \mathcal{Z}$. The projector onto the conditional strong δ-typical subspace given (x^n, z^n) is

$$\Pi^\delta(\sigma_B|x^n,z^n) \equiv \bigotimes_{a\in\mathcal{X}}\bigotimes_{b\in\mathcal{Z}}\Pi^\delta_{B^{I(a,b|x^n,z^n)}}(\hat{\sigma}_B^{a,b}). \tag{21}$$

Whereas, given a fixed $x^n \in \mathcal{X}^n$, we let $I_1(a|x^n) = \{i : x_i = a\}$, for $a \in \mathcal{X}$, and we define the conditional δ_1-typical projector as

$$\Pi^{\delta_1}(\sigma_B|x^n) \equiv \bigotimes_{a\in\mathcal{X}}\Pi^{\delta_1}_{B^{I_1(a|x^n)}}(\sigma_B^a). \tag{22}$$

We observe that for $\delta < \delta_1$ and sufficiently large n, the conditional typical subspaces are included within each other and we have the following projector inequality,

$$\Pi^\delta(\sigma_B|x^n,z^n) \preceq \Pi^{\delta_1}(\sigma_B|x^n). \tag{23}$$

This property will turn out to be useful in the derivation of the generalized quantum packing lemma below.

A. Generalized Quantum Packing Lemma

Suppose that Alice 1 and Alice 2 wish to send a common message m_0 and two private messages m_1 and m_2, without cribbing. To this end, they construct three codebooks, as follows. The first codebook encodes the common message and consists of 2^{nR_0} codewords $u^n(m_0)$, $m_0 \in [1 : 2^{nR_0}]$. In addition, Alice 1 employs a private codebook with $2^{n(R_0+R_1)}$ codewords, $x_1^n(m_0, m_1)$, $(m_0, m_1) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_1}]$, by which she chooses a quantum state from an ensemble $\{\theta_{x_1^n}\}_{x_1^n \in \mathcal{X}_1^n}$. Similarly, Alice 2 uses $x_2^n(m_0, m_2)$, $(m_0, m_2) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_2}]$, and the quantum states $\{\zeta_{x_2^n}\}_{x_2^n \in \mathcal{X}_2^n}$. The achievability proof is based on random codebook generation, where the codewords are drawn at random according to an input distribution $p_U p_{X_1|U} p_{X_2|U}$. To recover the transmitted message, Bob performs the square-root measurement [58], [59] using a code projector Π and codeword projectors $\Pi_{u^n x_1^n x_2^n}$, $(u^n, x_1^n, x_2^n) \in \mathcal{U}^n \times \mathcal{X}_1^n \times \mathcal{X}_2^n$, which project onto subspaces of the Hilbert space $\mathcal{H}_{B^n}$. The lemma below is a generalized version of the quantum packing lemma by Hsieh, Devetak, and Winter [15].

Lemma 2 (Generalized Quantum Packing Lemma): Let

$$\rho = \sum_{u\in\mathcal{U}} p_U(u) \sum_{x_1\in\mathcal{X}_1}\sum_{x_2\in\mathcal{X}_2} p_{X_1|U}(x_1|u)p_{X_2|U}(x_2|u)\rho_{x_1,x_2} \tag{24}$$

where $\{p_U p_{X_1|U}(x_1|u)p_{X_2|U}(x_2|u), \rho_{x_1 x_2}\}_{(x_1,x_2)\in\mathcal{X}_1\times\mathcal{X}_2}$ is a given ensemble. Furthermore, suppose that there is a code projector Π and codeword projectors Π_{u^n,x_1^n,x_2^n}, Π_{u^n,x_1^n}, Π_{u^n,x_2^n}, Π_{u^n}, for $(u^n, x_1^n, x_2^n) \in \mathcal{A}^\delta(p_U p_{X_1|U} p_{X_2|U})$, that satisfy for every $\alpha > 0$ and sufficiently large n,

$$\text{Tr}(\Pi\rho_{u^n,x_1^n,x_2^n}) \geq 1 - \alpha \tag{a}$$

$$\text{Tr}(\Pi_{u^n,x_1^n,x_2^n}\rho_{u^n,x_1^n,x_2^n}) \geq 1 - \alpha \tag{b}$$

$$\text{Tr}(\Pi_{u^n,x_1^n,x_2^n}) \leq 2^{n\mathsf{e}_0} \tag{c}$$

$$\Pi\rho^{\otimes n}\Pi \preceq 2^{-n(\mathsf{E}_0-\alpha)}\Pi \tag{d}$$

$$\Pi_{u^n,x_2^n}\rho_{u^n,x_2^n}\Pi_{u^n,x_2^n} \preceq 2^{-n(\mathsf{E}_1-\alpha)}\Pi_{u^n,x_2^n} \tag{e}$$

$$\Pi_{u^n,x_1^n}\rho_{u^n,x_1^n}\Pi_{u^n,x_1^n} \preceq 2^{-n(\mathsf{E}_2-\alpha)}\Pi_{u^n,x_1^n} \tag{f}$$

$$\Pi_{u^n}\rho_{u^n}\Pi_{u^n} \preceq 2^{-n(\mathsf{E}_{12}-\alpha)}\Pi_{u^n} \tag{g}$$

$$\Pi_{u^n,x_1^n,x_2^n} \preceq \Pi_{u^n,x_k^n} \preceq \Pi_{u^n} \preceq \Pi \text{ for } k = 1, 2, \tag{h}$$

for some $0 < \mathsf{e}_0 < E_j$ with $\rho_{u^n x_1^n x_2^n} \equiv \bigotimes_{i=1}^n \rho_{u_i x_{1,i} x_{2,i}}$. Then, there exist codewords $u^n(m_0)$, $x_1^n(m_0, m_1)$, and $x_2^n(m_0, m_2)$, for $m_k \in [1 : 2^{nR_k}]$, $k = 0, 1, 2$, and a POVM $\{\Lambda_{m_0,m_1,m_2}\}_{m_k\in[1:2^{nR_k}],k=0,1,2}$ such that

$$\frac{1}{2^{n(R_0+R_1+R_2)}}\sum_{m_0=1}^{2^{nR_0}}\sum_{m_1=1}^{2^{nR_1}}\sum_{m_2=1}^{2^{nR_2}}$$
$$\text{Tr}\left(\Lambda_{m_0,m_1,m_2}\rho_{x_1^n(m_0,m_1),x_2^n(m_0,m_2)}\right)$$
$$\geq 1 - \varepsilon_n(\alpha)$$
$$-4\left(4\cdot 2^{-n(\mathsf{E}_0-\mathsf{e}_0-(R_0+R_1+R_2)-\varepsilon_n'(\alpha))}\right.$$

$$+ 2^{-n(\mathsf{E}_1 - \mathsf{e}_0 - R_1 - \varepsilon_n(\alpha)')} + 2^{-n(\mathsf{E}_2 - \mathsf{e}_0 - R_2 - \varepsilon_n(\alpha)')}$$

$$+ 2^{-n(\mathsf{E}_{12} - \mathsf{e}_0 - (R_1 + R_2) - \varepsilon'_n(\alpha))} \Bigg) \tag{25}$$

where $\varepsilon_n(\alpha)$, $\varepsilon'_n(\alpha)$ tend to zero as $n \to \infty$ and $\alpha \to 0$.

We prove the generalized quantum packing lemma in Appendix A. The lemma can also be derived by following the methods in [50]. While the lemma above does not involve cribbing, it will still be useful in the analysis of our main results. Roughly speaking, for the quantum MAC with cribbing, the cribbing measurement allows Alice 2 to recover a part of Alice 1's message. As this component is known to both Alice 1 and Alice 2, it can be treated as a common message.

IV. MAIN RESULTS

We state our results on the quantum MAC $\mathcal{N}_{A'_1 A_2 \to B} \circ \mathcal{L}_{A_1 \to A'_1 E}$ with cribbing.

A. Strictly-Causal Cribbing

We begin with the MAC with strictly-causal cribbing, where Alice 2 transmits $A_{2,i}$ at time i, and then measures the cribbing system E_i after the transmission. Hence, she only knows the past measurement outcomes z^{i-1} at time i. We derive an achievable region, and give a regularized characterization for the capacity region in the special case of robust cribbing. As a consequence, we determine the capacity region of the classical-quantum MAC with noiseless cribbing. Define

$$\mathcal{R}_{\text{s-c}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L}) = \bigcup_{p_U p_{X_1|U} p_{X_2|U}, \, \theta_{A_1}^{x_1} \otimes \zeta_{A_2}^{x_2}}$$

$$\left\{ \begin{array}{rl} (R_1, R_2): \ R_1 & \leq I(X_1; E|U)_\omega \\ R_2 & \leq I(X_2; B|X_1, U)_\omega \\ R_1 + R_2 & \leq I(X_1 X_2; B)_\omega \end{array} \right\} \tag{26}$$

with

$$\omega_{U X_1 X_2 A'_1 E A_2} =$$

$$\sum_{u, x_1, x_2} p_U(u) p_{X_1|U}(x_1|u) p_{X_2|U}(x_2|u) |u\rangle\langle u|$$

$$\otimes |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \otimes \mathcal{L}_{A_1 \to A'_1 E}(\theta_{A_1}^{x_1}) \otimes \zeta_{A_2}^{x_2}, \tag{27}$$

$$\omega_{U X_1 X_2 B} = \mathcal{N}_{A'_1 A_2 \to B}(\omega_{U X_1 X_2 A'_1 A_2}). \tag{28}$$

The superscript 'DF' stands for decode-forward coding, referring to the coding scheme that achieves this rate region. Our terminology follows the analogy with the relay model (see Remark 2).

Before we state the capacity theorem, we give the following lemma. In principle, one may use the property below in order to compute the region $\mathcal{R}_{\text{s-c}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L})$ for a given channel.

Lemma 3: The union in (26) is exhausted by auxiliary variables U, X_1, X_2 with $|\mathcal{U}| \leq |\mathcal{H}_B|^2 + 2$, $|\mathcal{X}_1| \leq (|\mathcal{H}_{A_1}|^2 + 2)(|\mathcal{H}_B|^4 + 2)$, and $|\mathcal{X}_2| \leq (|\mathcal{H}_{A_2}|^2 + 1)(|\mathcal{H}_B|^2 + 2)$.

The proof of is based on the Fenchel-Eggleston-Carathéodory lemma [60], using similar arguments as in [23]. The details are given in Appendix B.

Theorem 4: Consider a quantum MAC $\mathcal{N}_{A'_1 A_2 \to B} \circ \mathcal{L}_{A_1 \to A'_1 E}$.

1) The rate region $\mathcal{R}_{\text{s-c}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L})$ is achievable for the quantum MAC with strictly-causal cribbing, *i.e.*

$$\mathcal{C}_{\text{s-c}}(\mathcal{N} \circ \mathcal{L}) \supseteq \mathcal{R}_{\text{s-c}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L}). \tag{29}$$

2) Given robust cribbing, the capacity region satisfies

$$\mathcal{C}_{\text{s-c}}(\mathcal{N} \circ \mathcal{L}) = \bigcup_{n=1}^{\infty} \frac{1}{n} \mathcal{R}_{\text{s-c}}^{\text{DF}}(\mathcal{N}^{\otimes n} \circ \mathcal{L}^{\otimes n}). \tag{30}$$

The proof of Theorem 4 is given in Appendix C. To the best of our knowledge, the achievability result above is new even for a classical channel. In the proof, we extend the block Markov coding scheme [10], where Alice 1 encodes two messages in each block in a sequential manner, such that one message is new and the other overlaps with the previous block. Willems and van der Muelen [10] refer to those messages as fresh information and resolution, respectively. Given strictly-causal cribbing, Alice 2 can measure the cribbing systems of the previous block. Thus, Alice 2 decodes the resolution by measuring the previous cribbing block, and encodes the resolution along with her own message. Bob recovers the messages in a reversed order using backward decoding. We refer to this coding scheme as 'decode-forward', since Alice 2 is responsible for decoding the messages of Alice 1, and forwarding them to Bob.

Remark 5: The decode-forward coding scheme relies heavily on the ability of Alice 2 to decode using a cribbing measurement. Our results suggest that this is optimal when the cribbing link is robust. However, as we will discuss in Section V, this is far from optimal when the cribbing link is too noisy. Therefore, we will introduce a new cribbing method which is useful for the non-robust case as well.

As a consequence of Theorem 4, we determine the capacity region of the classical-quantum MAC with a noiseless cribbing channel (see Definition 2).

Corollary 5: The capacity region of the classical-quantum MAC $(\mathcal{N}_{X_1 X_2 \to B} \circ \text{id}_{X_1 \to X_1 X_1})$ with strictly-causal noiseless cribbing is given by

$$\mathcal{C}_{\text{s-c}}(\mathcal{N} \circ \text{id}) =$$

$$\bigcup_{p_U p_{X_1|U} p_{X_2|U}} \left\{ \begin{array}{rl} (R_1, R_2): \ R_1 & \leq H(X_1|U) \\ R_2 & \leq I(X_2; B|X_1, U)_\omega \\ R_1 + R_2 & \leq I(X_1 X_2; B)_\omega \end{array} \right\} \tag{31}$$

with

$$\omega_{U X_1 X_2 B} = \sum_{u, x_1, x_2} p_U(u) p_{X_1|U}(x_1|u) p_{X_2|U}(x_2|u) |u\rangle\langle u|$$

$$\otimes |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \otimes \mathcal{N}_{X_1 X_2 \to B}(x_1, x_2) \tag{32}$$

The proof of Corollary 5 is given in Appendix D. The corollary can be viewed as the classical-quantum counterpart of Willems and van der Muelen's result on the classical MAC with noiseless cribbing [10].

B. Causal and Non-Causal Cribbing

In this section, we address causal and non-causal cribbing. Recall that in the causal setting, Alice 2 measures the cribbing

system E_i at time i, before she transmits. Hence, she knows the past and present measurement outcomes $z^i = (z^{i-1}, z_i)$ at time i. Whereas, in the non-causal case, Alice 2 can perform a joint measurement on E^n a priori, *i.e.* before the beginning of her transmission. Here, as opposed to the model in the previous section, Alice 2's measurement may cause a "state collapse" of Alice 1's input. Hence, it affects the input state for both transmitters. This can be seen in the achievable region below as well. Define

$$\mathcal{R}_{\text{caus}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L}) = \bigcup_{p_U p_{X_1|U}, \, \mathcal{W}_{E \to \bar{E}Z}, \, p_{X_2|Z,U}, \, \theta_{A_1}^{x_1} \otimes \zeta_{A_2}^{x_2}}$$
$$\left\{ \begin{array}{rl} (R_1, R_2) : R_1 & \leq I(X_1; \bar{E}Z|U)_\omega \\ R_2 & \leq I(X_2; B|X_1 U)_\omega \\ R_1 + R_2 & \leq I(X_1 X_2; B)_\omega \end{array} \right\} \quad (33)$$

where the union is over the probability distributions $p_U p_{X_1|U}$ for the auxiliary random variables U and X_1, over the measurement instrument $\mathcal{W}_{E \to \bar{E}Z}(\rho) \equiv \sum_z W_z \rho W_z^\dagger \otimes |z\rangle\langle z|$, the conditional probability distributions $p_{X_2|Z,U}$ for the auxiliary random variable X_2, and the input state collections $\{\theta_{A_1}^{x_1} \otimes \zeta_{A_2}^{x_2}\}$, with

$$\omega_{U X_1 A_1' \bar{E} Z X_2 A_2} = \sum_{u, x_1, z} p_U(u) p_{X_1|U}(x_1|u) |u\rangle\langle u| \otimes |x_1\rangle\langle x_1|$$
$$\otimes W_z \left(\mathcal{L}_{A_1 \to A_1' E}(\theta_{A_1}^{x_1}) \right) W_z^\dagger \otimes |z\rangle\langle z|$$
$$\otimes \left(\sum_{x_2} p_{X_2|Z,U}(x_2|z, u) |x_2\rangle\langle x_2| \otimes \zeta_{A_2}^{x_2} \right), \quad (34)$$
$$\omega_{U X_1 X_2 B} = \mathcal{N}_{A_1' A_2 \to B}(\omega_{U X_1 X_2 A_1' A_2}). \quad (35)$$

Here, $\mathcal{W}_{E \to \bar{E}Z}$ is a quantum instrument of a measurement, where $\bar{E}$ is the post-measurement cribbing system, and Z is the measurement outcome.

Theorem 6: Consider a quantum MAC $\mathcal{N}_{A_1' A_2 \to B} \circ \mathcal{L}_{A_1 \to A_1' E}$.

1) The rate region $\mathcal{R}_{\text{caus}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L})$ is achievable for the quantum MAC with causal cribbing, *i.e.*

$$\mathcal{C}_{\text{caus}}(\mathcal{N} \circ \mathcal{L}) \supseteq \mathcal{R}_{\text{caus}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L}). \quad (36)$$

2) For the classical-quantum MAC $(\mathcal{N}_{X_1 X_2 \to B} \circ \text{id}_{X_1 \to X_1 X_1})$ with noiseless cribbing,

$$\mathcal{C}_{\text{caus}}(\mathcal{N} \circ \text{id}) = \mathcal{C}_{\text{n-c}}(\mathcal{N} \circ \text{id}) =$$
$$\bigcup_{p_{X_1 X_2}} \left\{ \begin{array}{rl} (R_1, R_2) : R_1 & \leq H(X_1) \\ R_2 & \leq I(X_2; B|X_1)_\omega \\ R_1 + R_2 & \leq I(X_1 X_2; B)_\omega \end{array} \right\} \quad (37)$$

with

$$\omega_{X_1 X_2 B} = \sum_{x_1, x_2} p_{X_1 X_2}(x_1, x_2) |x_1, x_2\rangle\langle x_1, x_2|$$
$$\otimes \mathcal{N}_{X_1 X_2 \to B}(x_1, x_2) \quad (38)$$

The proof of Theorem 6 is given in Appendix E. Part 1 seems to be new for classical channels as well, while part 2 is the classical-quantum version of Willems and van der Muelen's result [10].

Remark 6: The classical achievability proof of Willems and van der Meulen [10] is based on the notion of Shannon
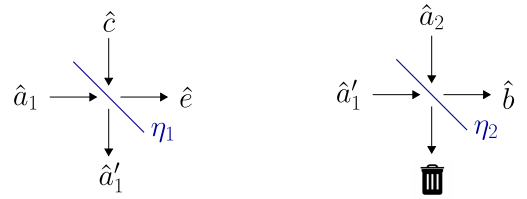

Fig. 2. The beam splitter relation of the single-mode bosonic MAC. The left beam splitter corresponds to the cribbing channel $\mathcal{L}_{A_1 \to A_1' E}$, while the right beam splitter describes the communication channel $\mathcal{N}_{A_1' A_2 \to B}$. Alice 1 encodes the message m_1 by a coherent state, and sends her transmission through n uses of the cribbing channel $\mathcal{L}_{A_1 \to A_1' E}$. The output E^n is the cribbing system which will be measured by the second transmitter in a strictly-causal manner. At time i, Alice 2 encodes a coherent state that depends on her message m_2 and on the previous cribbing measurement outcomes z^{i-1}, she sends the state, and then performs a measurement on E^i to obtain a measurement outcome z_i which she will use in the next time instance. The inputs $A_{1,i}'^n$ and $A_{2,i}$ are sent through the communication channel $\mathcal{N}_{A_1' A_2 \to B}$, and Bob receives the output B_i. Bob performs a measurement on $\hat{B}^n$ to obtain an estimate $(\hat{m}_1, \hat{m}_2)$ of the senders' messages.

strategies, as originally introduced in models of channel uncertainty [61]. For the classical MAC, a strategy is defined as a function $f : \mathcal{X}_1 \to \mathcal{X}_2$ that maps an input symbol of Alice 1 to that of Alice 2. Here, we replace the strategy by a quantum instrument that Alice 2 performs on the cribbing system.

C. Bosonic MAC With Strictly-Causal Cribbing

To demonstrate our results, consider the single-mode bosonic MAC. We extend the finite-dimension result in Theorem 4 to the bosonic channel with infinite-dimension Hilbert spaces based on the discretization limiting argument by Guha *et al.* [62]. A detailed description of (continuous-variable) bosonic systems can be found in [63]. Here, we only define the notation for the quantities that we use. We use hat-notation, *e.g.* $\hat{a}$, $\hat{b}$, $\hat{e}$, to denote operators that act on a quantum state. The single-mode Hilbert space is spanned by the Fock basis $\{|n\rangle\}_{n=0}^\infty$. Each $|n\rangle$ is an eigenstate of the number operator $\hat{n} = \hat{a}^\dagger \hat{a}$, where $\hat{a}$ is the bosonic field annihilation operator. In particular, $|0\rangle$ is the vacuum state of the field. The *creation operator* $\hat{a}^\dagger$ creates an excitation: $\hat{a}^\dagger |n\rangle = \sqrt{n+1}|n+1\rangle$, for $n \geq 0$. Reversely, the *annihilation operator* $\hat{a}$ takes away an excitation: $\hat{a}|n+1\rangle = \sqrt{n+1}|n\rangle$. A coherent state $|\alpha\rangle$, where $\alpha \in \mathbb{C}$, corresponds to an oscillation of the electromagnetic field, and it is the outcome of applying the displacement operator to the vacuum state, *i.e.* $|\alpha\rangle = D(\alpha)|0\rangle$, which resembles the creation operation, with $D(\alpha) \equiv \exp(\alpha \hat{a}^\dagger - \alpha^* \hat{a})$. A thermal state $\tau(N)$ is a Gaussian mixture of coherent states, where

$$\tau(N) \equiv \int_{\mathbb{C}} d^2\alpha \frac{e^{-|\alpha|^2/N}}{\pi N} |\alpha\rangle\langle\alpha|$$
$$= \frac{1}{N+1} \sum_{n=0}^\infty \left(\frac{N}{N+1} \right)^n |n\rangle\langle n| \quad (39)$$

given an average photon number $N \geq 0$.

Consider a bosonic MAC with cribbing encoders, whereby the channel input is a pair of electromagnetic field modes, with annihilation operator $\hat{a}_1$ and $\hat{a}_2$, and the output is a modes with

annihilation operator $\hat{b}$. The annihilation operators correspond to Alice 1, Alice 2, and Bob, respectively. The input-output relation of the bosonic MAC in the Heisenberg picture [64] is given by

$$\hat{e} = \sqrt{\eta_1}\,\hat{a}_1 + \sqrt{1-\eta_1}\,\hat{c} \tag{40}$$

$$\hat{a}'_1 = \sqrt{1-\eta_1}\,\hat{a}_1 - \sqrt{\eta_1}\,\hat{c} \tag{41}$$

and

$$\hat{b} = \sqrt{\eta_2}\,\hat{a}'_1 + \sqrt{1-\eta_2}\,\hat{a}_2 \tag{42}$$

where $\hat{c}$ is associated with the environment noise and the parameter η_k is the transmissivity, $0 \le \eta_k \le 1$, which captures, for instance, the length of the optical fiber and its absorption length [65]. The relations above correspond to the outputs of a beam splitter, as illustrated in Figure 2. It is assumed that the encoder uses a coherent state protocol with an input constraint. That is, the input state is a coherent state $|x_k\rangle$, $x_k \in \mathbb{C}$, for $k = 1, 2$, such that each codeword satisfies $\frac{1}{n}\sum_{i=1}^{n}|x_{k,i}|^2 \le N_{A_k}$.

Based on part 1 of Theorem 4, we derive the following achievable region with strictly-causal cribbing,

$$\mathcal{C}_{\text{s-c}}(\mathcal{N} \circ \mathcal{L}) \supseteq$$
$$\left\{ \begin{array}{l} (R_1, R_2): \\ R_1 \le g(\eta_1 N_{A_1} + (1-\eta_1)N_C) - g((1-\eta_1)N_C) \\ R_2 \le g(\eta_2 \eta_1 N_C + (1-\eta_2)N_{A_2}) - g(\eta_2 \eta_1 N_C) \\ R_1 + R_2 \le g(\eta_2(1-\eta_1)N_{A_1} \\ \quad + \eta_2 \eta_1 N_C + (1-\eta_2)N_{A_2}) - g(\eta_2 \eta_1 N_C) \end{array} \right\}$$
$$\tag{43}$$

where $g(N)$ is the entropy of a thermal state with mean photon number N,

$$g(x) = \begin{cases} (x+1)\log(x+1) - x\log(x) & x > 0 \\ 0 & x = 0. \end{cases} \tag{44}$$

To obtain the region above, set the inputs to be mixed coherent states, $|X_1\rangle$ and $|X_2\rangle$, where X_k is a circularly-symmetric Gaussian random variable with zero mean and variance $N_{A_k}/2$, for $k = 1, 2$, and let $U = 0$.

On the other hand, the capacity region of single-mode Bosonic MAC without cribbing is [66]

$$\mathcal{C}_{\text{none}}(\mathcal{N} \circ \mathcal{L}) =$$
$$\left\{ \begin{array}{l} (R_1, R_2): \\ R_1 \le g(\eta_2(1-\eta_1)N_{A_1} + \eta_2 \eta_1 N_C) - g(\eta_2 \eta_1 N_C) \\ R_2 \le g(\eta_2 \eta_1 N_C + (1-\eta_2)N_{A_2}) - g(\eta_2 \eta_1 N_C) \\ R_1 + R_2 \le g(\eta_2(1-\eta_1)N_{A_1} \\ \quad + \eta_2 \eta_1 N_C + (1-\eta_2)N_{A_2}) - g(\eta_2 \eta_1 N_C) \end{array} \right\}.$$
$$\tag{45}$$

The decode-forward achievable region with strictly-causal cribbing and the capacity region without cribbing are depicted in Figure 3 as the area below the thick blue line and below the dashed red line, respectively. As can be seen in the figure, cribbing can lead to a significant rate gain for Alice 1.
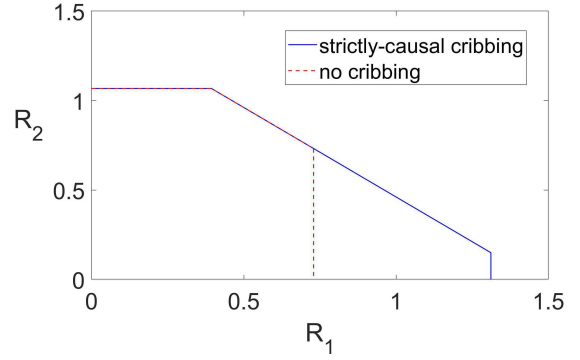

Fig. 3. Achievable regions for the single-mode bosonic MAC. The decode forward achievable region with strictly-causal cribbing is the area below the thick blue line (see (43)). For comparison, the capacity region without cribbing $\mathcal{C}_{\text{none}}(\mathcal{N} \circ \mathcal{L})$ is depicted as the area below the red dashed line (see (45)). The transmission rate of Alice 1 can be significantly higher using cribbing.

V. PARTIAL DECODE-FORWARD

We have pointed out in Remark 5 that if the cribbing system is not robust, then the decode-forward strategy in the previous section is not necessarily optimal. In fact, if the cribbing system E that Alice 2 measures is noisier than the channel input A'_1, then the inner bounds above may be worse than in transmission without cribbing. This is easy to see when E does not contain any information, say $|\mathcal{H}_E| = 1$, in which case the decode-forward region leaves Alice 1 with zero rate,

$$\mathcal{R}_{\text{s-c}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L}) = \mathcal{R}_{\text{caus}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L}) =$$
$$\left\{ (0, R_2): R_2 \le \max_{p_X, \zeta_{A_2}^x} I(X; B) \right\}. \tag{46}$$

In order to treat the case where the cribbing system is noisy, we introduce a quantum cribbing method that is based on relay coding techniques (see Remark 2 on the analogy between the relay channel and the cribbing model). Specifically, we improve the inner bound by incorporating the partial decode-forward technique within our previous cribbing scheme. Consider the quantum MAC with strictly-causal cribbing. Define

$$\mathcal{R}_{\text{s-c}}^{\text{PDF}}(\mathcal{N} \circ \mathcal{L}) = \bigcup_{p_{UV}p_{X_1|U,V}p_{X_2|U,V}\,,\,\theta_{A_1}^{x_1} \otimes \zeta_{A_2}^{x_2}}$$
$$\left\{ \begin{array}{rl} (R_1, R_2): R_1 & \le I(V; E|U)_\omega + I(X_1; B|X_2 UV)_\omega \\ R_2 & \le I(X_2; B|X_1 U)_\omega \\ R_1 + R_2 & \le I(X_1 X_2; B)_\omega \end{array} \right\}$$
$$\tag{47}$$

with

$$\omega_{UVX_1X_2A'_1EA_2} = \sum_{u,v,x_1,x_2} p_{U,V}(u,v)p_{X_1|U,V}(x_1|u,v)$$
$$\cdot\, p_{X_2|U,V}(x_2|u,v)|u,v,x_1,x_2\rangle\langle u,v,x_1,x_2|$$
$$\otimes\, \mathcal{L}_{A_1 \to A'_1 E}(\theta_{A_1}^{x_1}) \otimes \zeta_{A_2}^{x_2}, \tag{48}$$
$$\omega_{UVX_1X_2B} = \mathcal{N}_{A'_1 A_2 \to B}(\omega_{UVX_1X_2A'_1A_2}). \tag{49}$$

The superscript 'PDF' stands for partial decode-forward coding.

Theorem 7: Consider a quantum MAC $\mathcal{N}_{A_1' A_2 \to B} \circ \mathcal{L}_{A_1 \to A_1' E}$. The rate region $\mathcal{R}_{\text{s-c}}^{\text{PDF}}(\mathcal{N} \circ \mathcal{L})$ is achievable for the quantum MAC with strictly-causal cribbing, *i.e.*

$$\mathcal{C}_{\text{s-c}}(\mathcal{N} \circ \mathcal{L}) \supseteq \mathcal{R}_{\text{s-c}}^{\text{PDF}}(\mathcal{N} \circ \mathcal{L}). \tag{50}$$

The proof of Theorem 7 is given in Appendix F. To the best of our knowledge, this is a new result for classical channels as well. As opposed to the analysis for Section IV, the decoder does not fully rely on the cribbing measurement to recover the information from Alice 1. Instead, we use rate-splitting such that part of Alice 1's message is decoded-forward through the cribbing system E, and the remaining part is decoded using A_1'.

For the classical-quantum MAC with *noisy* cribbing, we also prove a cutset upper bound. In this case, the channel inputs are $A_1' \equiv A_1 \equiv X_1$ and $A_2 \equiv X_2$, and the cribbing channel $\mathcal{L}$ is represented by a classical noisy channel $Q_{Z|X_1}$ (see Definition 3). Then, we determine the capacity region in the special case where Z is a deterministic function of X_1, by showing that the partial decode-forward inner bound and the cutset outer bound coincide. Given a classical-quantum MAC $\mathcal{N}_{X_1 X_2 \to B} \circ Q_{Z|X_1}$, define

$$\mathcal{R}_{\text{s-c}}^{\text{CS}}(\mathcal{N} \circ Q) =$$
$$\bigcup_{p_U p_{X_1|U} p_{X_2|U}} \left\{ \begin{array}{rl} (R_1, R_2) : R_1 & \le I(X_1; BZ|X_2 U)_\omega \\ R_2 & \le I(X_2; B|X_1 U)_\omega \\ R_1 + R_2 & \le I(X_1 X_2; B)_\omega \end{array} \right\} \tag{51}$$

with

$$\omega_{U X_1 X_2 Z B} = \sum_{u, x_1, z, x_2} p_U(u) p_{X_1|U}(x_1|u) Q_{Z|X_1}(z|x_1)$$
$$\cdot p_{X_2|U}(x_2|u) |u, x_1, z, x_2\rangle\langle u, x_1, z, x_2|$$
$$\otimes \mathcal{N}_{X_1 X_2 \to B}(x_1, x_2). \tag{52}$$

The superscript 'CS' stands for the cutset outer bound.

Theorem 8: Consider a classical-quantum MAC with a noisy cribbing channel $Q_{Z|X_1}$.

1) The capacity region of the classical-quantum MAC $\mathcal{N}_{X_1 X_2 \to B} \circ Q_{Z|X_1}$ with strictly-causal noisy cribbing is bounded by

$$\mathcal{C}_{\text{s-c}}(\mathcal{N} \circ Q) \subseteq \mathcal{R}_{\text{s-c}}^{\text{CS}}(\mathcal{N} \circ Q). \tag{53}$$

2) If $Q_{Z|X_1}$ is a 0-1 matrix, *i.e.* the cribbing observation Z is a deterministic function of X_1, then

$$\mathcal{C}_{\text{s-c}}(\mathcal{N} \circ Q) = \mathcal{R}_{\text{s-c}}^{\text{PDF}}(\mathcal{N} \circ Q) = \mathcal{R}_{\text{s-c}}^{\text{CS}}(\mathcal{N} \circ Q) =$$
$$\bigcup_{p_U p_{X_1|U} p_{X_2|U}} \left\{ \begin{array}{l} (R_1, R_2) : \\ R_1 \le H(Z|U) + I(X_1; B|X_2 U Z)_\omega \\ R_2 \le I(X_2; B|X_1 U)_\omega \\ R_1 + R_2 \le I(X_1 X_2; B)_\omega \end{array} \right\}. \tag{54}$$

The proof of Theorem 8 is given in Appendix G. Part 1 seems to be new for classical channels as well, while part 2 is the classical-quantum version of the classical result on the classical MAC with partial cribbing [36]. Although, Asnani and Permuter [36] considered a more complex network with cribbing at both encoders.

VI. Summary and Discussion

We consider the quantum MAC $\mathcal{M}_{A_1 A_2 \to B}$ with cribbing encoders. In quantum communication, the description is more delicate. By the no-cloning theorem, universal copying of quantum states is impossible. Therefore, in the view of quantum mechanics, perfect cribbing is against the laws of nature. As illustrated in Figure 1, if Alice 1 sends A_1 through the channel, then Alice 2 is physically prohibited from having a copy of the input state. Hence, we consider the quantum MAC with noisy cribbing, consisting of a concatenation of a cribbing channel $\mathcal{L}_{A_1 \to A_1' E}$ and the communication channel $\mathcal{N}_{A_1' A_2 \to B}$ (see Figure 1b). Specifically, Alice 1 sends her input system A_1 through a cribbing channel that has two outputs, A_1' and E. Alice 2 performs a measurement on the system E, and uses the measurement outcome in order to encode the input state of A_2. Then, the input systems A_1' and A_2 are sent through the communication channel. The model can also be interpreted as if the second transmitter performs a measurement on the environment E of the first transmitter.

We consider the scenarios of strictly-causal, causal, and non-causal cribbing. In strictly-causal cribbing, Alice 2 transmits $A_{2,i}$ at time i, and then measures the cribbing system E_i after the transmission. Hence, she only knows the past measurement outcomes z^{i-1} at time i. In the causal setting, Alice 2 measures the cribbing system E_i, at time i, before she transmits. Hence, she knows the past and present measurement outcomes $z^i = (z^{i-1}, z_i)$ at time i. Whereas, in the non-causal case, Alice 2 can perform a joint measurement on E^n a priori, *i.e.* before the beginning of her transmission. The entanglement between Transmitter 1 and the cribbing system of Transmitter 2 has the following implication. In both the causal and non-causal scenarios, Alice 2 performs the measurement before the input $A_{1,i}'$ is sent through the channel. Hence, the cribbing measurement may inflict a "state collapse" for Alice 1's transmission. In other words, in quantum communication, the cribbing operation interferes with Alice 1's input before it is even transmitted through the communication channel. For a MAC with *robust cribbing*, there exists a recovery channel from E to $E A_1'$. Thereby, the cribbing system E includes all the information that is available in A_1'. We derived achievable regions for each setting and established a regularized capacity characterization for robust cribbing.

The setting of noisy cribbing is significantly more challenging and it is closely-related to the relay channel. We have derived a generalized packing lemma for the MAC. While the lemma does not include cribbing, it is useful in the analysis of the MAC with cribbing encoders. In the setting of the generalized quantum packing lemma in Section III, Alice 1 and Alice 2 send three messages to Bob, a common message m_0 and two private messages m_1 and m_2. It is assumed that there is no cribbing. The generalized quantum packing lemma implies that the capacity region of the quantum MAC $\mathcal{M}_{A_1 A_2 \to B}$ with a common message, without cribbing, is given by the regularization of the following region,

$$\mathcal{R}_{\text{none}}^*(\mathcal{M}) = \bigcup_{p_U p_{X_1|U} p_{X_2|U}, \, \theta_{A_1}^{x_1} \otimes \zeta_{A_2}^{x_2}}$$

$$\left\{\begin{array}{rl} (R_0, R_1, R_2) : R_1 & \leq I(X_1; B|X_2U)_\omega \\ R_2 & \leq I(X_2; B|X_1U)_\omega \\ R_1 + R_2 & \leq I(X_1X_2; B|U)_\omega \\ R_0 + R_1 + R_2 & \leq I(X_1X_2; B)_\omega \end{array}\right\} \quad (55)$$

where the union is as in (11), with $|\mathcal{U}| = |\mathcal{H}_{A_1}|^2 + |\mathcal{H}_{A_2}|^2 + 1$, $|\mathcal{X}_1| \leq |\mathcal{U}|(|\mathcal{H}_{A_k}|^2 + 1)$, for $k = 1, 2$.

This generalizes the classical-quantum result due to Boche and Nötzel [17, Theorem 2].

We further investigated the comparison between the relay channel and the MAC with cribbing encoders in Section V (see Remark 2). Building on the analogy between the noisy cribbing model and the relay channel, we developed a partial decode-forward region for the quantum MAC with strictly-causal non-robust cribbing. For the classical-quantum MAC with cribbing encoders, we have completely determined the capacity region with perfect cribbing of the classical input, and derived a cutset region for noisy cribbing. In the special case of a classical-quantum MAC with a deterministic cribbing channel, the inner and outer bounds coincide.

Cooperation in terms of a quantum MAC offers additional advantages for resilience and trustworthiness, *e.g.* to compensate for jamming attacks by an adversary [67, Section VI.C.]. As previously mentioned, in future communication systems, not only data but also physical and virtual objects will be controlled [31]. This growing demand of communication resources sharpens the need for quantum technology, and it is thus interesting to develop the corresponding theory.

APPENDIX A
PROOF OF LEMMA 2 (GENERALIZED QUANTUM PACKING LEMMA)

Consider the following code construction and decoding measurement.

A. Classical Codebook Construction

(i) Generate 2^{nR_0} independent sequences $u^n(m_0)$, $m_0 \in [1 : 2^{nR_0}]$, at random according to $\prod_{i=1}^n p_U(u_i)$.

(ii) For every m_0, generate 2^{nR_k} conditionally independent sequences $x_k^n(m_0, m_k)$, $m_k \in [1 : 2^{nR_k}]$, according to $\prod_{i=1}^n p_{X_k|U}(x_{k,i}|u_i(m_0))$, where $k = 1, 2$.

We denote the codebooks by $\mathscr{C} = (\mathscr{C}_0, \mathscr{C}_1, \mathscr{C}_2)$, respectively.

B. Decoding POVM

For every u^n, x_1^n, x_2^n, define

$$\Upsilon_{u^n x_1^n x_2^n} = \Pi \Pi_{u^n x_1^n x_2^n} \Pi. \quad (56)$$

Then, define

$$\Lambda_{m_0, m_1, m_2} =$$
$$\left(\sum_{m_0'=1}^{2^{nR_0}} \sum_{m_1'=1}^{2^{nR_1}} \sum_{m_2'=1}^{2^{nR_2}} \Upsilon_{u^n(m_0')x_1^n(m_0',m_1')x_2^n(m_0',m_2')} \right)^{-1/2}$$
$$\cdot \Upsilon_{u^n(m_0)x_1^n(m_0,m_1)x_2^n(m_0,m_2)}$$

$$\cdot \left(\sum_{m_0'=1}^{2^{nR_0}} \sum_{m_1'=1}^{2^{nR_1}} \sum_{m_2'=1}^{2^{nR_2}} \Upsilon_{u^n(m_0')x_1^n(m_0',m_1')x_2^n(m_0',m_2')} \right)^{-1/2} \cdot \quad (57)$$

C. Error Analysis

Denote the message-average error probability by

$$P_e^{(n)}(\mathscr{C}) = \frac{1}{2^{n(R_0+R_1+R_2)}} \sum_{m_0'', m_1'', m_2''} P_e^{(n)}(\mathscr{C}|m_0'', m_1'', m_2'') \quad (58)$$

where the conditional probability of error is defined as

$$P_e^{(n)}(\mathscr{C}|m_0, m_1, m_2) \equiv$$
$$\mathrm{Tr}\left[(\mathbb{1} - \Lambda_{m_0,m_1,m_2}) \rho_{U^n(m_0), X_1^n(m_0,m_1), X_2^n(m_0,m_2)} \right]. \quad (59)$$

We will show that the expected probability of error is bounded by

$$\mathbb{E}_{\mathscr{C}} P_e^{(n)}(\mathscr{C}) \leq 2(\alpha + 3\sqrt{\alpha})$$
$$+ 4\Big(4 \cdot 2^{-n(\mathsf{E}_0 - \mathsf{e}_0 - (R_0+R_1+R_2) - \alpha)}$$
$$+ 2^{-n(\mathsf{E}_1 - \mathsf{e}_0 - R_1 - \alpha)} + 2^{-n(\mathsf{E}_2 - \mathsf{e}_0 - R_2 - \alpha)}$$
$$+ 2^{-n(\mathsf{E}_{12} - \mathsf{e}_0 - (R_1+R_2) - \alpha)} \Big). \quad (60)$$

Fix a message triplet (m_0, m_1, m_2). By the symmetry of the codebook generation, we have

$$\mathbb{E}_{\mathscr{C}} P_e^{(n)}(\mathscr{C})$$
$$= \frac{1}{2^{n(R_0+R_1+R_2)}} \sum_{m_0'', m_1'', m_2''} \mathbb{E}_{\mathscr{C}} P_e^{(n)}(\mathscr{C}|m_0'', m_1'', m_2'')$$
$$= \mathbb{E}_{\mathscr{C}} P_e^{(n)}(\mathscr{C}|m_0, m_1, m_2). \quad (61)$$

Hence, we may assume without loss of generality that the triplet (m_0, m_1, m_2) was sent.

The error analysis begins with the Hayashi-Nagaoka inequality [68, Lemma 2],

$$\mathbb{1} - (S_1 + S_2)^{-1/2} S_1 (S_1 + S_2)^{-1/2} \leq 2(\mathbb{1} - S_1) + 4S_2, \quad (62)$$

for S_1, S_2 such that $0 \preceq S_j \preceq \mathbb{1}$, $j = 1, 2$. Then, by the definition of Λ_{m_0,m_1,m_2} in (57),

$$\mathbb{1} - \Lambda_{m_0,m_1,m_2}$$
$$\preceq 2(\mathbb{1} - \Upsilon_{U^n(m_0)X_1^n(m_0,m_1)X_2^n(m_0,m_2)})$$
$$+ 4 \sum_{\substack{m_0',m_1',m_2' : \\ (m_0',m_1',m_2') \neq (m_0,m_1,m_2)}} \Upsilon_{U^n(m_0')X_1^n(m_0',m_1')X_2^n(m_0',m_2')}. \quad (63)$$

Thus, the expected probability of error satisfies

$$\mathbb{E}_{\mathscr{C}} P_e^{(n)}(\mathscr{C})$$
$$= \mathbb{E}_{\mathscr{C}} P_e^{(n)}(\mathscr{C}|m_0, m_1, m_2)$$
$$\leq 2\mathbb{E}\left\{ \mathrm{Tr}\left[(\mathbb{1} - \Upsilon_{U^n(m_0), X_1^n(m_0,m_1), X_2^n(m_0,m_2)}) \right. \right.$$

$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]\Big\}$$
$$+4 \sum_{\substack{m_0',m_1',m_2': \\ (m_0',m_1',m_2')\neq(m_0,m_1,m_2)}}$$
$$\mathbb{E}\Big\{\mathrm{Tr}\big[\Upsilon_{U^n(m_0'),X_1^n(m_0',m_1'),X_2^n(m_0',m_2')}$$
$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\big]\Big\} \tag{64}$$

where the equality is due to (61), and the inequality follows from (59) and (63).

The first term is now bounded as in the single-user case [15]. We note that the probability that $(U^n(m_0), X_1^n(m_0, m_1), X_2^n(m_0, m_2))$ is not δ-typical, for our fixed message triplet (m_0, m_1, m_2), tends to zero as $n \to \infty$, by the weak law of large numbers. That is,

$$\Pr\Big(U^n(m_0), X_1^n(m_0,m_1), X_2^n(m_0,m_2) \notin$$
$$\mathcal{A}^\delta(P_U P_{X_1|U} P_{X_2})\Big) \leq \zeta_n \tag{65}$$

where ζ_n tends to zero as $n \to \infty$. For every δ-typical realization of $(u^n(m_0), x_1^n(m_0, m_1), x_2^n(m_0, m_2))$, we have

$$\mathrm{Tr}\big[\Upsilon_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}$$
$$\cdot \rho_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}\big]$$
$$= \mathrm{Tr}\big[\Pi\Pi_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}\Pi$$
$$\cdot \rho_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}\big]. \tag{66}$$

Then, by the cyclicity of the trace operation, this equals

$$\mathrm{Tr}\big[\Pi_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}$$
$$\cdot \Pi\rho_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}\Pi\big]$$
$$\geq \mathrm{Tr}\big[\Pi_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}$$
$$\cdot \rho_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}\big]$$
$$- \big\|\Pi_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}$$
$$\cdot \rho_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}\Pi_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}$$
$$- \rho_{u^n(m_0),x_1^n(m_0,m_1),x_2^n(m_0,m_2)}\big\|$$
$$\geq 1 - \alpha - 2\sqrt{\alpha} \tag{67}$$

where the last inequality follows from the gentle operator lemma [69], [55, Lemma 9.4.2] and Assumption (b). It follows that the first term in the RHS of (64) is bounded by $2(\alpha + 2\sqrt{\alpha}) + \zeta_n \leq 2(\alpha + 3\sqrt{\alpha})$ (see (65)), for sufficiently large n.

Next, observe that the sum in the RHS of (64) can be divided as follows,

$$\sum_{\substack{m_0',m_1',m_2': \\ (m_0',m_1',m_2')\neq(m_0,m_1,m_2)}} \mathbb{E}\mathrm{Tr}\Big[\Upsilon_{m_0',m_1',m_2'}$$
$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$= \sum_{m_0'\neq m_0} \mathbb{E}\mathrm{Tr}\Big[\Upsilon_{m_0',m_1,m_2}\rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$+ \sum_{m_1'\neq m_1} \mathbb{E}\mathrm{Tr}\Big[\Upsilon_{m_0,m_1',m_2}\rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$

$$+ \sum_{m_2'\neq m_2} \mathbb{E}\mathrm{Tr}\Big[\Upsilon_{m_0,m_1,m_2'}\rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$+ \sum_{m_1'\neq m_1 \, , \, m_2'\neq m_2}$$
$$\mathbb{E}\mathrm{Tr}\Big[\Upsilon_{m_0,m_1',m_2'}\rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$+ \sum_{m_0'\neq m_0 \, , \, m_1'\neq m_1}$$
$$\mathbb{E}\mathrm{Tr}\Big[\Upsilon_{m_0',m_1',m_2} \cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$+ \sum_{m_0'\neq m_0 \, , \, m_2'\neq m_2}$$
$$\mathbb{E}\mathrm{Tr}\Big[\Upsilon_{m_0',m_1,m_2'} \cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$+ \sum_{m_0'\neq m_0 \, , \, m_1'\neq m_1 \, , \, m_2'\neq m_2}$$
$$\mathbb{E}\mathrm{Tr}\Big[\Upsilon_{m_0',m_1',m_2'} \cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big] \tag{68}$$

The derivation above is analogous to the union of events bound in the classical proof. In particular, the first three sums correspond to the event that exactly one of the messages is decoded erroneously, the next three sums to the event that two of the messages are decoded incorrectly, and the last sum is the probability that all of the decoder's estimates are wrong.

Now, we will bound each of the seven sums in the RHS of (68). Consider the first sum. Then,

$$\mathbb{E}\mathrm{Tr}\Big[\Upsilon_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)}$$
$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$= \mathrm{Tr}\Big[\mathbb{E}\big\{\Pi\Pi_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)}\Pi$$
$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\big\}\Big]. \tag{69}$$

Observe that for every $m_0' \neq m_0$, the random vectors $(U^n(m_0'), X_1^n(m_0', m_1), X_2^n(m_0', m_2))$ and $(U^n(m_0), X_1^n(m_0, m_1), X_2^n(m_0, m_2))$ are statistically independent of each other. Thus, the expression in the square brackets becomes

$$\mathbb{E}\big\{\Pi\Pi_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)}\Pi\big\}$$
$$\cdot \mathbb{E}\big\{\rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\big\}$$
$$= \mathbb{E}\big\{\Pi\Pi_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)}\Pi\big\} \cdot \rho^{\otimes n}. \tag{70}$$

Furthermore, the probability that one of the aforementioned random vectors is not δ-typical tends to zero as $n \to \infty$, by the weak law of large numbers. It follows that

$$\mathbb{E}\mathrm{Tr}\Big[\Upsilon_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)}$$
$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$= \mathrm{Tr}\big[\mathbb{E}\big\{\Pi\Pi_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)}\Pi\big\} \cdot \rho^{\otimes n}\big]$$
$$= \mathbb{E}\mathrm{Tr}\big[\Pi\Pi_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)}\Pi \cdot \rho^{\otimes n}\big]$$
$$= \mathbb{E}\mathrm{Tr}\big[\Pi_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)} \cdot \Pi\rho^{\otimes n}\Pi\big]$$
$$\leq 2^{-n(\mathsf{E}_0-\alpha)}\mathbb{E}\mathrm{Tr}\big[\Pi_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)} \cdot \Pi\big] + \zeta_n$$
$$\leq 2^{-n(\mathsf{E}_0-\alpha)}\mathbb{E}\mathrm{Tr}\big[\Pi_{U^n(m_0'),X_1^n(m_0',m_1),X_2^n(m_0',m_2)}\big] + \zeta_n$$
$$\leq 2^{-n(\mathsf{E}_0-\mathsf{e}_0-\alpha)} + 2\zeta_n \tag{71}$$

where the first and last inequalities hold by Assumptions (c) and (d). Thus, for large n, the first sum in the RHS of (68) is bounded by $2^{-n(\mathsf{E}_0-\mathsf{e}_0-2\alpha)} \cdot 2^{nR_0} = 2^{-n(\mathsf{E}_0-\mathsf{e}_0-R_0-2\alpha)}$. By the same considerations, the last three sums are bounded by $2^{-n(\mathsf{E}_0-\mathsf{e}_0-(R_0+R_1)-2\alpha)}$, $2^{-n(\mathsf{E}_0-\mathsf{e}_0-(R_0+R_2)-2\alpha)}$, and $2^{-n(\mathsf{E}_0-\mathsf{e}_0-(R_0+R_1+R_2)-2\alpha)}$, respectively. Thus, the combination of the first, fifth, sixth, and seventh sums in the RHS of (68) is bounded by $4 \cdot 2^{-n(\mathsf{E}_0-\mathsf{e}_0-(R_0+R_1+R_2)-2\alpha)}$.

It remains to bound the second, third, and fourth sums in the RHS of (68). Moving to the second sum,

$$\mathbb{E}\mathrm{Tr}\Big[\Upsilon_{U^n(m_0),X_1^n(m_0,m_1'),X_2^n(m_0,m_2)}$$
$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$= \mathrm{Tr}\Big[\mathbb{E}\big\{\Pi\Pi\Pi_{U^n(m_0),X_1^n(m_0,m_1'),X_2^n(m_0',m_2)}\Pi$$
$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\big\}\Big]. \qquad (72)$$

Consider a given realization of the codebooks $\mathscr{C}_0 = \{u_0^n(m_0)\}$ and $\mathscr{C}_2 = \{x_2^n(m_0,m_1)\}$. For every $m_1' \neq m_1$, the sequences $X_1^n(m_0,m_1')$ and $X_1^n(m_0,m_1)$ are statistically independent. Thus, when we condition on $\mathscr{C}_0$, $\mathscr{C}_2$, the expression in the square brackets becomes

$$\mathbb{E}_{\mathscr{C}_1}\big\{\Pi\Pi\Pi_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0,m_2)}\Pi\big\}$$
$$\cdot \mathbb{E}_{\mathscr{C}_1}\big\{\rho_{u^n(m_0),X_1^n(m_0,m_1),x_2^n(m_0,m_2)}\big\}$$
$$= \mathbb{E}_{\mathscr{C}_1}\big\{\Pi\Pi\Pi_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0,m_2)}\Pi\big\}$$
$$\cdot \rho_{u^n(m_0),x_2^n(m_0,m_2)}$$
$$= \mathbb{E}_{\mathscr{C}_1}\big\{\Pi_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0,m_2)}\big\} \cdot \rho_{u^n(m_0),x_2^n(m_0,m_2)}$$
$$= \mathbb{E}_{\mathscr{C}_1}\big\{\Pi_{u^n(m_0),x_2^n(m_0,m_2)}\Pi_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0,m_2)}$$
$$\Pi_{u^n(m_0),x_2^n(m_0,m_2)}\big\} \cdot \rho_{u^n(m_0),x_2^n(m_0,m_2)}. \qquad (73)$$

The last two equalities follow from Assumption (h), since $P_0 P_1 P_0 = P_1$ for every pair of projectors P_0 and P_1 such that $P_0 \succeq P_1$. It follows that

$$\mathbb{E}_{\mathscr{C}_1}\mathrm{Tr}\Big[\Upsilon_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0,m_2)}$$
$$\cdot \rho_{u^n(m_0),X_1^n(m_0,m_1),x_2^n(m_0,m_2)}\Big]$$
$$= \mathrm{Tr}\Big[\mathbb{E}_{\mathscr{C}_1}\big\{\Pi_{u^n(m_0),x_2^n(m_0,m_2)}\Pi_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0,m_2)}$$
$$\Pi_{u^n(m_0),x_2^n(m_0,m_2)}\big\} \cdot \rho_{u^n(m_0),x_2^n(m_0,m_2)}\Big]$$
$$= \mathbb{E}_{\mathscr{C}_1}\mathrm{Tr}\Big[\Pi_{u^n(m_0),x_2^n(m_0,m_2)}\Pi_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0,m_2)}$$
$$\Pi_{u^n(m_0),x_2^n(m_0,m_2)} \cdot \rho_{u^n(m_0),x_2^n(m_0,m_2)}\Big]$$
$$= \mathbb{E}_{\mathscr{C}_1}\mathrm{Tr}\Big[\Pi_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0,m_2)}$$
$$\cdot \Pi_{u^n(m_0),x_2^n(m_0,m_2)}\rho_{u^n(m_0),x_2^n(m_0,m_2)}\Pi_{u^n(m_0),x_2^n(m_0,m_2)}\Big]$$
$$\leq 2^{-n(\mathsf{E}_1-\alpha)}\mathbb{E}_{\mathscr{C}_1}\mathrm{Tr}\Big[\Pi_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0',m_2)}$$
$$\cdot \Pi_{u^n(m_0),x_2^n(m_0,m_2)}\Big] + \zeta_n$$
$$\leq 2^{-n(\mathsf{E}_1-\alpha)}\mathbb{E}_{\mathscr{C}_1}\mathrm{Tr}\big[\Pi_{u^n(m_0),X_1^n(m_0,m_1'),x_2^n(m_0,m_2)}\big] + \zeta_n$$
$$\leq 2^{-n(\mathsf{E}_1-\mathsf{e}_0-\alpha)} + 2\zeta_n \qquad (74)$$

by Assumptions (c) and (e). Thus, the second sum in the RHS of (68) is bounded by $2^{-n(\mathsf{E}_1-\mathsf{e}_0-2\alpha)} \cdot 2^{nR_1} = 2^{-n(\mathsf{E}_1-\mathsf{e}_0-R_1-2\alpha)}$. By symmetry, the third sum is bounded by $2^{-n(\mathsf{E}_2-\mathsf{e}_0-R_2-2\alpha)}$.

As for the fourth sum in the RHS of (68),

$$\mathbb{E}\mathrm{Tr}\Big[\Upsilon_{U^n(m_0),X_1^n(m_0,m_1'),X_2^n(m_0,m_2')}$$
$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$= \mathrm{Tr}\Big[\mathbb{E}\big\{\Pi\Pi\Pi_{U^n(m_0),X_1^n(m_0,m_1'),X_2^n(m_0,m_2')}\Pi$$
$$\cdot \rho_{U^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\big\}\Big] \qquad (75)$$

For every $m_1' \neq m_1$ and $m_2' \neq m_2$, the random vectors $(X_1^n(m_0,m_1'), X_2^n(m_0,m_2'))$ and $(X_1^n(m_0,m_1), X_2^n(m_0,m_2))$ are statistically independent. Thus, when we condition on $\mathscr{C}_0 = \{u_0^n(m_0)\}$, the expectation in the square bracket is a product,

$$\mathbb{E}_{\mathscr{C}_1\times\mathscr{C}_2}\big\{\Pi\Pi\Pi_{u^n(m_0),X_1^n(m_0,m_1'),X_2^n(m_0,m_2')}\Pi\big\}$$
$$\cdot \mathbb{E}_{\mathscr{C}_1\times\mathscr{C}_2}\big\{\rho_{u^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\big\}$$
$$= \mathbb{E}_{\mathscr{C}_1\times\mathscr{C}_2}\big\{\Pi\Pi\Pi_{u^n(m_0),X_1^n(m_0,m_1'),X_2^n(m_0,m_2')}\Pi\big\} \cdot \rho_{u^n(m_0)}$$
$$= \mathbb{E}_{\mathscr{C}_1\times\mathscr{C}_2}\big\{\Pi_{u^n(m_0)}\Pi_{u^n(m_0),X_1^n(m_0,m_1'),X_2^n(m_0,m_2')}$$
$$\Pi_{u^n(m_0)}\big\} \cdot \rho_{u^n(m_0)} \qquad (76)$$

based on Assumption (h). It follows that

$$\mathbb{E}_{\mathscr{C}_1\times\mathscr{C}_2}\mathrm{Tr}\Big[\Upsilon_{u^n(m_0),X_1^n(m_0,m_1'),X_2^n(m_0,m_2')}$$
$$\cdot \rho_{u^n(m_0),X_1^n(m_0,m_1),X_2^n(m_0,m_2)}\Big]$$
$$\leq 2^{-n(\mathsf{E}_{12}-\alpha)}\mathbb{E}_{\mathscr{C}_1\times\mathscr{C}_2}\mathrm{Tr}\big[\Pi_{u^n(m_0),X_1^n(m_0,m_1'),X_2^n(m_0,m_2')}$$
$$\cdot \Pi_{u^n(m_0),x_2^n(m_0,m_2)}\big] + \zeta_n$$
$$\leq 2^{-n(\mathsf{E}_{12}-\mathsf{e}_0-2\alpha)} \qquad (77)$$

for sufficiently large n, by Assumptions (c) and (g).

We conclude that the probability of decoding error, averaged over the class of the codebooks, is bounded as in (60). Therefore, there must exist deterministic codebooks with the same error bound, as we were set to prove. $\qquad\square$

APPENDIX B
PROOF OF LEMMA 3 (CARDINALITY BOUNDS)

To bound the alphabet size of the random variables U, X_1, and X_2, we use the Fenchel-Eggleston-Carathéodory lemma [60] and similar arguments as in [23] and [56]. Let

$$d_0 = |\mathcal{H}_B|^2 + 2 \qquad (78)$$
$$d_1 = |\mathcal{U}|(|\mathcal{H}_{A_1}|^2 + 2) \qquad (79)$$
$$d_2 = |\mathcal{U}|(|\mathcal{H}_{A_2}|^2 + 1). \qquad (80)$$

First, we bound the cardinality of $\mathcal{U}$. Consider a given ensemble $\{p_{X_1|U}(x_1|u)p_{X_2|U}(x_2|u), \theta_{A_1}^{x_1} \otimes \zeta_{A_2}^{x_2}\}$. Every density matrix ρ_A has a unique parametric representation $\mathcal{V}(\rho_A)$ of dimension $|\mathcal{H}_A|^2 - 1$. Then, define a map $f_0 : \mathcal{U} \to \mathbb{R}^{d_0}$ by

$$f_0(u) = \Big(\mathcal{V}(\omega_B^u), \; H(B|X_1,X_2,U=u)_\omega,$$
$$I(X_1;E|U=u)_\omega, \; I(X_2;B|X_1U=u)_\omega\Big) \qquad (81)$$

where $\omega_B^u = \sum_{x_1,x_2} p_{X_1|U}(x_1|u) p_{X_2|U}(x_2|u) \ \mathcal{N}(\text{Tr}_E(\mathcal{L}(\theta_{A_1}^{x_1})) \otimes \zeta_{A_2}^{x_2})$. The map f_0 can be extended to a map that acts on probability distributions as follows,

$$F_0 \; : \; p_U(\cdot) \mapsto \sum_{u \in \mathcal{U}} p_U(u) f_0(u) = \Big(\mathcal{V}(\omega_B),$$

$$H(B|X_1 X_2)_\omega, \ I(X_1; E|U)_\omega, \ I(X_2; B|X_1 U)_\omega \Big). \quad (82)$$

According to the Fenchel-Eggleston-Carathéodory lemma [60], any point in the convex closure of a connected compact set within $\mathbb{R}^d$ belongs to the convex hull of d points in the set. Since the map F_0 is linear, it maps the set of distributions on $\mathcal{U}$ to a connected compact set in $\mathbb{R}^{d_0}$. Thus, for every p_U, there exists a probability distribution $p_{\bar{U}}$ on a subset $\bar{\mathcal{U}} \subseteq \mathcal{U}$ of size d_0, such that $F_0(p_{\bar{U}}) = F_0(p_U)$. We deduce that alphabet size can be restricted to $|\mathcal{U}| \leq d_0$, while preserving ω_B, $H(B|X_1 X_2)_\omega$, $I(X_1; E|U)_\rho$, $I(X_2; B|X_1 U)_\rho$, and thus, $H(B)_\omega$ and $I(X_1 X_2; B)_\omega = H(B)_\omega - H(B|X_1 X_2)_\omega$.

Next, we bound the alphabet size for the auxiliary variables X_1 and X_2. For every $u \in \mathcal{U}$, define a map $f_{1,u} : \mathcal{X}_1 \to \mathbb{R}^{|\mathcal{H}_{A_1}|^2 + 2}$ by

$$f_{1,u}(x_1) = \Big(\mathcal{V}(\theta_{A_1}^{x_1}), \ H(E|X_1 = x_1, U = u)_\omega,$$

$$H(B|X_1 = x_1, U = u)_\omega, \ H(B|X_2, X_1 = x_1, U = u)_\omega \Big). \quad (83)$$

Then, the map f_1 is extended to

$$F_{1,u} \; : \; p_{X_1|U}(\cdot|u) \mapsto \sum_{x_1 \in \mathcal{X}_1} p_{X_1|U}(x_1|u) f_{1,u}(x_1) =$$

$$\Big(\mathcal{V}(\theta_{A_1}^u), \ H(E|X_1, U = u), \ H(B|X_1, U = u)_\omega,$$

$$H(B|X_2, X_1, U = u)_\omega \Big) \quad (84)$$

with

$$\theta_{A_1}^u \equiv \sum_{x_1} p_{X_1|U}(x_1|u) \theta_{A_1}^{x_1}. \quad (85)$$

Thus, by Fenchel-Eggleston-Carathéodory lemma [60], for every $p_{X_1|U}(\cdot|u)$, there exists a probability distribution $p_{\bar{X}_1|U}(\cdot|u)$ on a subset $\bar{\mathcal{X}}_1 \subseteq \mathcal{X}_1$ of size $|\mathcal{H}_{A_1}|^2 + 2$, such that $F_{1,u}(p_{\bar{X}_1|U}(\cdot|u)) = F_{1,u}(p_{X_1|U}(\cdot|u)))$. We deduce that alphabet size can be restricted to $|\mathcal{X}_1| \leq d_1$, while preserving $H(E|X_1 U)$, $H(B|X_1 U)_\omega$, $H(B|X_1 X_2)_\omega$, and

$$\omega_{U A_1 X_2 A_2} = \sum_u p_U(u) p_{X_2|U}(x_2|u) |u\rangle\langle u|$$

$$\otimes \theta_{A_1}^u \otimes |x_2\rangle\langle x_2| \otimes \zeta_{A_2}^{x_2} \quad (86)$$

where $\theta_{A_1}^u$ is as in (85). This implies that $\omega_{U A_1' E X_2 A_2} = \mathcal{L}(\omega_{U A_1 X_2 A_2})$, $H(E|U)_\omega$, and $I(X_1; E|U)_\omega = H(E|U)_\omega - H(E|X_1 U)_\omega$ remain the same, and also $\omega_{U X_2 B} \equiv \mathcal{N}_{A_1' A_2 \to B}(\omega_{U X_2 A_1' A_2})$, $H(B)_\omega$, $I(X_2; B|X_1 U)_\omega = H(B|X_1 U)_\omega - H(B|X_1 X_2)_\omega$, and $I(X_1, X_2; B)_\omega = H(B)_\omega - H(B|X_1, X_2)_\omega$. The bound $|\mathcal{X}_2| \leq d_2$ follows from the same argument using the function $f_{2,u} : \mathcal{X}_2 \to \mathbb{R}^{|\mathcal{H}_{A_2}|^2 + 1}$,

$$f_{2,u}(x_2) = \Big(\mathcal{V}(\zeta_{A_2}^{x_2}), \ H(B|X_1, X_2 = x_2, U = u)_\omega,$$

$$H(B|X_1, X_2 = x_2)_\omega \Big) \quad (87)$$

for $u \in \mathcal{U}$. This completes the proof for the cardinality bounds. $\square$

APPENDIX C
PROOF OF THEOREM 4

Consider the quantum MAC $\mathcal{N} \circ \mathcal{L}$ with strictly-causal cribbing.

A. Part 1

We show that for every $\delta_1, \delta_2, \varepsilon_0 > 0$, there exists a $(2^{n(R_1 - \delta_1)}, 2^{n(R_2 - \delta_2)}, n, \varepsilon_0)$ code for $\mathcal{N}_{A_1' A_2 \to B} \circ \mathcal{L}_{A_1 \to A_1' E}$ with strictly-causal cribbing at Encoder 2, provided that $(R_1, R_2) \in \mathcal{R}_{\text{s-c}}^{\text{DF}}(\mathcal{N} \circ \mathcal{L})$. To prove achievability, we extend the classical block Markov coding with backward decoding to the quantum setting, and then apply the quantum packing lemma.

We use T transmission blocks, where each block consists of n channel uses. In particular, with strictly-causal cribbing, Encoder 2 has access to the cribbing measurements from the previous blocks. Each transmitter sends $T - 1$ messages. Let us fix $m_k(0) = m_k(T) \equiv 1$ for $k = 1, 2$. Alice 1 sends $(m_1(j))_{j=1}^{T-1}$, and Alice 2 $(m_2(j))_{j=1}^{T-1}$. Hence, the coding rate for User k is $\left(\frac{T-1}{T} \right) R_k$, which tends to R_k as the number of blocks T grows to infinity.

Let $\{ p_U(u) p_{X_1|U}(x_1|u) p_{X_2|U}(x_2|u), \theta^{x_1} \otimes \zeta^{x_2} \}$ be a given ensemble over $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$.

Define

$$\omega_{A_1' E}^{x_1} = \mathcal{L}_{A_1 \to A_1' E}(\theta^{x_1}), \quad (88)$$

$$\omega_B^{x_1, x_2} = \mathcal{N}_{A_1' A_2 \to B}(\omega_{A_1'}^{x_1} \otimes \zeta^{x_2}). \quad (89)$$

In addition, let

$$\theta^u = \sum_{x_1 \in \mathcal{X}_1} p_{X_1|U}(x_1|u) \theta^{x_1} \quad (90)$$

$$\zeta^u = \sum_{x_2 \in \mathcal{X}_2} p_{X_2|U}(x_2|u) \zeta^{x_2} \quad (91)$$

$$\omega_{A_1' E}^u = \mathcal{L}_{A_1 \to A_1' E}(\theta^u) \quad (92)$$

$$\omega_B^u = \mathcal{N}_{A_1' A_2 \to B}(\omega_{A_1'}^u \otimes \zeta^u) \quad (93)$$

for $u \in \mathcal{U}$. Hence, $\omega_{A_1' E}$ and ω_B are the corresponding average states.

The code construction, encoding with cribbing, and decoding procedures are described below.

B. Classical Codebook Construction

(i) Generate 2^{nR_1} independent sequences $u^n(m_0)$, $m_0 \in [1 : 2^{nR_1}]$, at random according to $\prod_{i=1}^n p_U(u_i)$.

(ii) For every m_0, generate 2^{nR_1} conditionally independent sequences $x_1^n(m_0, m_1)$, $m_1 \in [1 : 2^{nR_1}]$, according to $\prod_{i=1}^n p_{X_1|U}(x_{1,i}|u_i(m_0))$.

(iii) For every m_0, generate 2^{nR_2} conditionally independent sequences $x_2^n(m_0, m_2)$, $m_2 \in [1 : 2^{nR_2}]$, according to $\prod_{i=1}^n p_{X_2|U}(x_{2,i}|u_i(m_0))$.

C. Encoding and Decoding

1) Encoder 1: To send the messages $(m_1(j))_{j=1}^{T-1}$, Alice 1 performs the following. In block j, set

$$m_0 = m_1(j-1), \ m_1 = m_1(j). \tag{94}$$

Then, prepare the state

$$\rho_{A_1^n(j)} = \theta^{x_1^n(m_0, m_1)} \tag{95}$$

and send $A_1^n(j)$, for $j \in [1:T]$. As the jth transmission goes through the cribbing channel $\mathcal{L}_{A_1 \to A_1' E}$, we have

$$\rho_{A_1'^n(j)E^n(j)} = \omega_{A_1'^n E^n}^{x_1^n(m_0, m_1)}. \tag{96}$$

The second transmitter can access the cribbing system $E^n(j)$, which are entangled with Alice 1's transmission. Cribbing is performed by a sequence of measurements that recover the messages of Alice 1. In each block, the choice of the cribbing measurement depends on the outcome in the previous block. With strictly-causal encoding, Alice 2 must perform the cribbing measurement at the end of the block, after she has already sent $A_2^n(j)$.

2) Encoder 2: To send the messages $(m_2(j))_{j=1}^{T-1}$, Alice 2 performs the following. Fix $\widetilde{m}_1(0) \equiv 1$. In block j, given the previous cribbing estimate $\widetilde{m}_1(j-1)$, do as follows, for $j = 1, 2, \ldots, T$.

(i) Set

$$m_0 = \widetilde{m}_1(j-1), \ m_2 = m_2(j). \tag{97}$$

(ii) Prepare the state

$$\rho_{A_2^n(j)} = \zeta^{x_2^n(m_0, m_2)} \tag{98}$$

and send $A_2^n(j)$, for $j \in [1:T]$.

(iii) Measure the next cribbing estimate $\widetilde{m}_1(j)$ by applying a POVM $\mathcal{K}_{E^n(j)|m_0} = \{K_{m_1|m_0}\}$ that will be chosen later.

3) Backward Decoding: The decoder recovers the messages using sequential measurements as well. Yet, the order is backwards, *i.e.* the measurement of the jth message of Alice 1 is chosen based on the estimate of $m_1(j+1)$. Fix $\hat{m}_1(0) = \hat{m}_1(T) \equiv 1$.

In block j, for $j = T, T-1, \ldots, 2$, the decoder uses the previous estimate of $\hat{m}_1(j)$, and measures $\hat{m}_1(j-1)$ and $\hat{m}_2(j)$ using a POVM $\mathcal{D}_{B^n(j)|\hat{m}_1} = \{D_{m_0, m_2|\hat{m}_1}\}$ with $\hat{m}_1 = \hat{m}_1(j)$. Finally, in block 1, the decoder measures $\hat{m}_2(1)$ with a POVM $\mathcal{D}_{B^n(1)|m_0, m_1} = \{D'_{m_2|\hat{m}_0, \hat{m}_1}\}$, with $\hat{m}_0 = 1$ and $\hat{m}_1 = \hat{m}_1(1)$. The decoding POVMs will also be specified later.

D. Analysis of Probability of Error

Let $\delta > 0$. We use the notation $\varepsilon_i(\delta)$, $i = 1, 2, \ldots$, for terms that tend to zero as $\delta \to 0$. By symmetry, we may assume without loss of generality that the transmitters send the messages $M_1(j) = M_2(j) = 1$. Consider the following events,

$$\mathcal{E}_0(j) = \{(U^n(M_1(j-1)), X_1^n(M_1(j-1), 1),$$
$$X_2^n(M_1(j-1), 1)) \notin \mathcal{A}^{\delta_1}(p_{U, X_1, X_2})\} \tag{99}$$

$$\mathcal{E}_1(j) = \{\widetilde{M}_1(j) \neq 1\} \tag{100}$$

$$\mathcal{E}_2(j) = \{\widehat{M}_1(j-1) \neq 1\} \tag{101}$$

$$\mathcal{E}_3(j) = \{\widehat{M}_2(j) \neq 1\} \tag{102}$$

for $j \in [1:T]$, with $\delta_1 \equiv \delta/(2|\mathcal{X}_2||\mathcal{U}|)$. By the union of events bound, the probability of error is bounded by

$$P_{e|m_1(j)=m_2(j)=1}^{(Tn)}(\mathcal{F}_1, \mathcal{F}_2, \mathcal{K}, \mathcal{D})$$
$$\leq \sum_{j=1}^{T} \Pr(\mathcal{E}_0(j)) + \sum_{j=1}^{T} \Pr(\mathcal{E}_1(j) \mid \mathcal{E}_0^c(j) \cap \mathcal{E}_1^c(j-1))$$
$$+ \sum_{j=1}^{T} \Pr(\mathcal{E}_2(j) \cup \mathcal{E}_3(j) \mid \mathcal{E}_0^c(j) \cap \mathcal{E}_1^c(j) \cap \mathcal{E}_2^c(j+1)) \tag{103}$$

where the conditioning on $M_1(j) = M_2(j) = 1$ is omitted for convenience of notation. By the weak law of large numbers, the probability terms $\Pr(\mathcal{E}_0(j))$ tend to zero as $n \to \infty$.

To bound the second sum, which is associated with the cribbing measurements, we use the quantum packing lemma. Alice 2's measurement is effectively a decoder for the marginal cribbing channel $\mathcal{L}_{A_1 \to E}^{(1)}$, which is defined by $\mathcal{L}_{A_1 \to E}^{(1)}(\rho_{A_1}) \equiv \mathrm{Tr}_{A_1'}(\mathcal{L}_{A_1 \to E A_1'}(\rho_{A_1}))$. Given $\mathcal{E}_0^c(j)$, we have that $(U^n(M_1(j-1)), X_1^n(M_1(j-1), 1)) \in \mathcal{A}^\delta(p_{U, X_1})$. Now, observe that

$$\mathrm{Tr}\left[\Pi^\delta(\omega_E|u^n, x_1^n)\omega_{E^n}^{u^n, x_1^n}\right] \geq 1 - \varepsilon_1(\delta) \tag{104}$$

$$\Pi^\delta(\omega_E|u^n)\omega_{E^n}^{u^n}\Pi^\delta(\omega_E|u^n) \preceq 2^{-n(H(E|U)_\omega - \varepsilon_1(\delta))}\Pi^\delta(\omega_E) \tag{105}$$

$$\mathrm{Tr}\left[\Pi^\delta(\omega_E|u^n, x_1^n)\right] \leq 2^{n(H(E|U, X_1)_\omega + \varepsilon_1(\delta))} \tag{106}$$

$$\mathrm{Tr}\left[\Pi^\delta(\omega_E|u^n)\omega_{E^n}^{u^n, x_1^n}\right] \geq 1 - \varepsilon_1(\delta) \tag{107}$$

for all $(u^n, x_1^n) \in \mathcal{A}^\delta(p_{U, X_1})$, by (17)-(20), respectively. Since the codebooks are statistically independent of each other, we have by the single-user quantum packing lemma [15], [56, Lemma 12], that there exists a POVM $K_{m_1|u^n}$ such that $\Pr(\mathcal{E}_1(j) \mid \mathcal{E}_0^c(j) \cap \mathcal{E}_1^c(j-1)) \leq 2^{-n(I(X_1;E|U)_\rho - R_1 - \varepsilon_2(\delta))}$, which tends to zero as $n \to \infty$, provided that

$$R_1 < I(X_1; E|U)_\omega - \varepsilon_2(\delta). \tag{108}$$

We move to the last sum in the RHS of (103). Here, we use our generalized packing lemma. Suppose that $\mathcal{E}_1^c(j) \cap \mathcal{E}_2^c(j+1)$ occurred, namely Encoder 2 measured the correct $M_1(j-1)$ and the decoder measured the correct $M_1(j)$. Furthermore,

$$\mathrm{Tr}(\Pi^\delta(\omega_B)\omega_{B^n}^{u^n, x_1^n, x_2^n}) \geq 1 - \varepsilon_3(\delta) \tag{a'}$$

$$\mathrm{Tr}(\Pi^\delta(\omega_B|u^n, x_1^n, x_2^n)\omega_{B^n}^{u^n, x_1^n, x_2^n}) \geq 1 - \varepsilon_3(\delta) \tag{b'}$$

$$\mathrm{Tr}(\Pi^\delta(\omega_B|u^n, x_1^n, x_2^n)) \leq 2^{nH(B|X_1 X_2)_\omega} \tag{c'}$$

$$\Pi^\delta(\omega_B)\omega_B^{\otimes n}\Pi^\delta(\omega_B) \preceq 2^{-n(H(B)_\omega - \alpha)}\Pi^\delta(\omega_B) \tag{d'}$$

and

$$\Pi^\delta(\omega_B|u^n, x_2^n)\omega_{B^n}^{u^n, x_2^n}\Pi^\delta(\omega_B|u^n, x_2^n)$$
$$\preceq 2^{-n(H(B|X_2 U)_\omega - \alpha)}\Pi^\delta(\omega_B|u^n, x_2^n) \tag{e'}$$

$$\Pi^\delta(\omega_B|u^n, x_1^n)\omega_{B^n}^{u^n, x_1^n}\Pi^\delta(\omega_B|u^n, x_1^n)$$
$$\preceq 2^{-n(H(B|X_1U)_\omega - \alpha)}\Pi^\delta(\omega_B|u^n, x_1^n) \quad \text{(f')}$$

$$\Pi^\delta(\omega_B|u^n)\omega_{B^n}^{u^n}\Pi^\delta(\omega_B|u^n)$$
$$\preceq 2^{-n(H(B|U)_\omega - \alpha)}\Pi^\delta(\omega_B|u^n) \quad \text{(g')}$$

$$\Pi^\delta(\omega_B|u^n, x_1^n, x_2^n)$$
$$\preceq \Pi^\delta(\omega_B|u^n, x_k^n) \preceq \Pi^\delta(\omega_B|u^n) \preceq \Pi^\delta(\omega_B)$$
$$\text{for } k = 1, 2, \quad \text{(h')}$$

for all $(x_1^n, x_1^n) \in \mathcal{A}^{\delta/2}(p_{X_1, X_2})$, by (17)-(20). Therefore, by Lemma 2, there exists a POVM $D_{m_0, m_2|m_1}$ such that

$$\Pr\left(\mathcal{E}_2(j) \cup \mathcal{E}_3(j) \mid \mathcal{E}_0^c(j) \cap \mathcal{E}_1^c(j) \cap \mathcal{E}_2^c(j+1)\right)$$
$$\leq 4 \cdot 2^{-n(I(X_1X_2;B)_\rho - (R_0+R_2) - \varepsilon_4(\delta))}$$
$$+ 2^{-n(I(X_2;B|X_1U)_\omega - R_2 - \varepsilon_4(\delta))}$$
$$+ 2^{-n(I(X_1X_2;B|U)_\omega - R_2 - \varepsilon_4(\delta))} \quad \text{(109)}$$

where we set $R_0 \leftarrow R_1$ and $R_1 \leftarrow 0$, since we are decoding $(m_0, m_2) \equiv (M_1(j-1), M_2(j))$ while conditioning on $m_1 \equiv M_1(j)$. The last bound tends to zero as $n \to \infty$, provided that

$$R_2 < I(X_2;B|X_1U)_\omega - \varepsilon_4(\delta) \quad \text{(110)}$$

and

$$R_1 + R_2 < I(X_1X_2;B)_\omega - \varepsilon_4(\delta). \quad \text{(111)}$$

This completes the achievability proof.

E. Part 2

Consider the quantum MAC with strictly-causal *robust* cribbing. To show that rate pairs in $\frac{1}{\kappa}\mathcal{R}_{\text{s-c}}^{\text{DF}}((\mathcal{N} \circ \mathcal{L})^{\otimes \kappa})$ are achievable, one may employ the coding scheme from part 1 for the product MAC $(\mathcal{N} \circ \mathcal{L})^{\otimes \kappa}$, where κ is arbitrarily large. Now, we show the converse part using standard considerations along with the quantum Markov chain property for robust cribbing (see Definition 1).

Suppose that Alice 1 chooses m_1 uniformly at random, and prepares an input state $\rho_{A_1^n}^{m_1}$. Upon sending the systems A_1^n through the cribbing channel, we have $\rho_{A_1'^n E^n}^{m_1} = \mathcal{L}_{A_1^n \to A_1'^n E^n}(\rho_{A_1^n}^{m_1})$. Before preparing the state of her system $A_{2,i}$, Alice 2 can measure the cribbing systems E^{i-1}, and obtain an outcome z_{i-1}. Hence, Alice 2 prepares the input state $\rho_{A_2^i}^{m_2, z^{i-1}}$. Then, $A_{1,i}'$ and $A_{2,i}$ are sent through the MAC $\mathcal{N}_{A_1' A_2 \to B}$. Bob receives the output systems B^n and performs a measurement in order to obtain an estimate $(\hat{m}_1, \hat{m}_2)$ of the message pair.

Consider a sequence of codes $(\mathcal{F}_{1n}, \mathcal{F}_{2n}, \mathcal{K}_n, \mathcal{D}_n)$ such that the average probability of error tends to zero, hence the error probabilities $\Pr\left(\widehat{M}_1 \neq M_1\right)$, $\Pr\left((\widehat{M}_1, \widehat{M}_2) \neq (M_1, M_2)\right)$, $\Pr\left(\widehat{M}_2 \neq M_2|M_1\right)$ are bounded by some α_n which tends to zero as $n \to \infty$. By Fano's inequality [70], it follows that

$$H(M_1|\widehat{M}_1) \leq n\varepsilon_n \quad \text{(112)}$$
$$H(M_1, M_2|\widehat{M}_1, \widehat{M}_2) \leq n\varepsilon_n' \quad \text{(113)}$$
$$H(M_2|\widehat{M}_2, M_1) \leq n\varepsilon_n'' \quad \text{(114)}$$

where $\varepsilon_n, \varepsilon_n', \varepsilon_n''$ tend to zero as $n \to \infty$. Hence,

$$nR_1 = H(M_1) = I(M_1; \widehat{M}_1)_\rho + H(M_1|\widehat{M}_1)$$
$$\leq I(M_1; \widehat{M}_1)_\rho + n\varepsilon_n$$
$$\leq I(M_1; B^n)_\rho + n\varepsilon_n$$
$$\leq I(M_1; A_1'^n E^n)_\rho + n\varepsilon_n \quad \text{(115)}$$

where the second inequality follows from the Holevo bound (see Ref. [57, Theo. 12.1]), and the last inequality follows from the data processing inequality. Now, given robust cribbing, the systems $M_1 \diamond E^n \diamond A_1'^n$ form a quantum Markov chain, by Definition 1. Hence, $I(M_1; A_1'^n|E^n)_\rho = 0$ and $I(M_1; A_1'^n E^n)_\rho = I(M_1; E^n)_\rho$. Thus,

$$nR_1 \leq I(M_1; E^n)_\rho + n\varepsilon_n. \quad \text{(116)}$$

By the same considerations, we also have

$$n(R_1 + R_2) = I(M_1M_2; \widehat{M}_1, \widehat{M}_2)_\rho + H(M_1M_2|\widehat{M}_1, \widehat{M}_2)$$
$$\leq I(M_1M_2; B^n)_\rho + n\varepsilon_n' \quad \text{(117)}$$

and

$$nR_2 = I(M_2; \widehat{M}_2|M_1)_\rho + H(M_2|\widehat{M}_2, M_1)$$
$$\leq I(M_2; B^n|M_1)_\rho + n\varepsilon_n''. \quad \text{(118)}$$

The proof for the regularized region follows from (116)-(118) by defining $X_1^n = f_1(M_1)$, $X_2^n = f_2(M_2)$, and $U^n = \emptyset$, where f_k is an arbitrary one-to-one map from $[1:2^{nR_k}]$ to $\mathcal{X}_k^n$, for $k = 1, 2$. □

APPENDIX D
PROOF OF COROLLARY 5

Consider the classical-quantum MAC $\mathcal{N}_{X_1X_2 \to B} \circ \text{id}_{X_1 \to X_1X_1}$ with strictly-causal noiseless cribbing. Since the cribbing system stores a perfect copy of Alice 1's classical input in this case, *i.e.* $E \equiv X_1$, we have $I(X_1; E|U)_\omega = H(X_1|U)$. Thereby, achievability immediately follows from part 1 of Theorem 4. Note that the classical-quantum setting with noiseless cribbing is a special case of robust cribbing. Hence, we can use the derivation of part 2 of Theorem 4 as well.

As for the converse proof, let $X_1^n \equiv \mathcal{F}_1(M_1)$ and $X_{2,i} \equiv \mathcal{F}_{2,i}(M_2, X_1^{i-1})$ denote the channel inputs. Since the encoding map $\mathcal{F}_1$ is classical, there exists a random element S_1 that controls the encoding function. That is, $X_1^n \equiv f_1(M_1, S_1)$, where f_1 is a deterministic function, and S_1 is statistically independent of the message. Define

$$U_i \equiv (X_1^{i-1}, S_1). \quad \text{(119)}$$

Then, the transmission rate of Alice 1 satisfies

$$nR_1 = H(M_1)$$
$$= H(M_1|S_1)$$
$$= H(X_1^n|S_1)$$
$$= \sum_{i=1}^n H(X_{1,i}|U_i) \quad \text{(120)}$$

where the second equality holds since S_1 is independent of the message, the third since X_1^n is a deterministic function of (M_1, S_1), and the last equality follows from the entropy chain rule (see (119)).

Next, we bound the transmission rate of Alice 2 as follows,

$$
\begin{aligned}
nR_2 &= H(M_2) \\
&= H(M_2|M_1, S_1) \\
&= I(M_2; \widehat{M}_2|M_1, S_1) + H(M_2|\widehat{M}_2, M_1, S_1) \\
&\leq I(M_2; \widehat{M}_2|M_1, S_1) + H(M_2|\widehat{M}_2, M_1) \\
&\leq I(M_2; B^n|M_1, S_1)_\rho + n\varepsilon_n''
\end{aligned}
\tag{121}
$$

where the first inequality holds since conditioning cannot increase entropy, and the second follows from the data processing inequality and Fano's inequality (see (114)). Using the chain rule, the last bound can be expressed as

$$
\begin{aligned}
nR_2 &\leq \sum_{i=1}^n I(M_2; B_i|M_1, S_1, B^{i-1})_\rho + n\varepsilon_n'' \\
&= \sum_{i=1}^n I(M_2; B_i|M_1, S_1, X_{1,i}, X_1^{i-1}, B^{i-1})_\rho + n\varepsilon_n''
\end{aligned}
\tag{122}
$$

as X_1^n is a deterministic function of (M_1, S_1). Observe that the mutual information summand is then bounded by

$$
\begin{aligned}
&I(M_2; B_i|M_1, S_1, X_{1,i}, X_1^{i-1}, B^{i-1})_\rho \\
&\leq I(M_2, X_{2,i}; B_i|M_1, S_1, X_{1,i}, X_1^{i-1}, B^{i-1})_\rho \\
&= H(B_i|M_1, S_1, X_{1,i}, X_1^{i-1}, B^{i-1})_\rho \\
&\quad - H(B_i|X_{1,i}, X_{2,i}, M_1, M_2, S_1, X_1^{i-1}, B^{i-1})_\rho \\
&\leq H(B_i|S_1, X_{1,i}, X_1^{i-1})_\rho \\
&\quad - H(B_i|X_{1,i}, X_{2,i}, M_1, M_2, S_1, X_1^{i-1}, B^{i-1})_\rho.
\end{aligned}
\tag{123}
$$

Consider the second term, and observe that given $X_{1,i} = x_{1,i}$, $X_{2,i} = x_{2,i}$, the output system B_i is in the state $\mathcal{N}(x_{1,i}, x_{2,i})$ and it has no correlation with M_1, M_2, S_1, X_1^{i-1}, and B^{i-1}. That is, given $X_{1,i}$ and $X_{2,i}$, the system B_i is in a product state with the joint system of $(M_1, M_2, S_1, X_1^{i-1}, B^{i-1})$. Thus, the second term equals $H(B_i|X_{1,i}, X_{2,i})_\rho$. Similarly, $H(B_i|S_1, X_{1,i}, X_{2,i}, X_1^{i-1})_\rho = H(B_i|X_{1,i}, X_{2,i})_\rho$ as well, which implies

$$
\begin{aligned}
&H(B_i|X_{1,i}, X_{2,i}, M_1, M_2, S_1, X_1^{i-1}, B^{i-1})_\rho \\
&= H(B_i|X_{1,i}, X_{2,i}, S_1, X_1^{i-1})_\rho.
\end{aligned}
\tag{124}
$$

Therefore, by (122)-(124), along with the definition of U_i in (119),

$$
\begin{aligned}
nR_2 &\leq \sum_{i=1}^n [H(B_i|X_{1,i}, U_i)_\rho - H(B_i|X_{1,i}, X_{2,i}, U_i)_\rho] + n\varepsilon_n'' \\
&= \sum_{i=1}^n I(X_{2,i}; B_i|X_{1,i}, U_i)_\rho + n\varepsilon_n''.
\end{aligned}
\tag{125}
$$

The bound on the sum-rate is straightforward. Indeed, by (117),

$$
n(R_1 + R_2) \leq \sum_{i=1}^n I(M_1 M_2; B_i|B^{i-1})_\rho + n\varepsilon_n'
$$

$$
\begin{aligned}
&\leq \sum_{i=1}^n I(M_1 M_2 B^{i-1}; B_i)_\rho + n\varepsilon_n' \\
&\leq \sum_{i=1}^n I(M_1 M_2 B^{i-1} X_{1,i} X_{2,i}; B_i)_\rho + n\varepsilon_n' \\
&= \sum_{i=1}^n I(X_{1,i} X_{2,i}; B_i)_\rho + n\varepsilon_n'
\end{aligned}
\tag{126}
$$

since $I(M_1 M_2 B^{i-1}; B_i|X_{1,i} X_{2,i})_\rho = 0$.

To complete the proof, consider a time index K that is drawn uniformly at random, from $[1 : n]$, independently of M_1, M_2, and S_1. Then, by (120), (125), and (126),

$$
R_1 = H(X_{1,K}|U_K, K)
\tag{127}
$$

$$
R_2 - \varepsilon_n'' \leq I(X_{2,K}; B_K|X_{1,K}, U_K, K)_\rho
\tag{128}
$$

$$
R_1 + R_2 - \varepsilon_n' \leq I(X_{1,K} X_{2,K}; B_K|K)_\rho.
\tag{129}
$$

Furthermore, define a joint state $\omega_{U X_1 X_2 B}$ by identifying $U \equiv (K, U_K)$, $X_1 \equiv X_{1,K}$, $X_2 \equiv X_{2,K}$, and $B \equiv B_K$. That is,

$$
\begin{aligned}
\omega_{U X_1 X_2 B} &\equiv \sum_{i=1}^n \sum_{u_i, x_1, x_2} \frac{1}{n} p_{U_i}(u_i) p_{X_{1,i}|U_i}(x_1|u_i) \\
&\quad \cdot p_{X_{2,i}|U_i}(x_2|u_i)|i, u_i\rangle\langle i, u_i| \otimes |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \\
&\quad \otimes \mathcal{N}(x_1, x_2).
\end{aligned}
\tag{130}
$$

Thus, the individual rates are bounded by

$$
R_1 = H(X_1|U)
\tag{131}
$$

$$
R_2 - \varepsilon_n'' \leq I(X_2; B|X_1, U)_\omega.
\tag{132}
$$

As for the sum-rate bound in (129),

$$
\begin{aligned}
R_1 + R_2 - \varepsilon_n' &\leq I(X_1 X_2; B|K)_\omega \\
&\leq I(X_1 X_2 K; B)_\omega \\
&= I(X_1 X_2; B)_\omega + I(K; B|X_1 X_2)_\omega \\
&= I(X_1 X_2; B)_\omega
\end{aligned}
\tag{133}
$$

as $I(K; B|X_1 X_2)_\omega = 0$ since the channel has a memoryless product form. This completes the proof of Corollary 5. $\square$

APPENDIX E
PROOF OF THEOREM 6

A. Part 1

Consider the quantum MAC $\mathcal{N} \circ \mathcal{L}$ with causal cribbing. Since the achievability proof is similar to the derivation for the strictly-causal setting in Appendix E, we only give the outline. As before, we use T transmission blocks to send $T - 1$ messages for each user, $(m_1(j))_{j=1}^{T-1}$ and $(m_2(j))_{j=1}^{T-1}$. Let $\{p_U p_{X_1|U}, \theta^{x_1}\}$ be a given ensemble over $\mathcal{H}_{A_1}$. Furthermore, consider a measurement instrument $\mathcal{W}_{E \to \bar{E} Z}$, and let $\{p_{X_2|Z,U}(\cdot|z, u), \zeta^{x_2}\}$, for $z \in \mathcal{Z}$ and $u \in \mathcal{U}$, be a collection of ensembles over $\mathcal{H}_{A_2}$.

Define

$$
\sigma_{A_1' E}^{x_1} = \mathcal{L}_{A_1 \to A_1' E}(\theta^{x_1}),
\tag{134}
$$

$$
\omega_{A_1' \bar{E} Z}^{x_1} = \mathcal{W}_{E \to \bar{E} Z}(\omega_{A_1' E}^{x_1})
\tag{135}
$$

$$
\omega_B^{x_1, x_2} = \mathcal{N}_{A_1' A_2 \to B}(\omega_{A_1'}^{x_1} \otimes \zeta^{x_2}).
\tag{136}
$$

In addition, for every $u \in \mathcal{U}$ and measurement outcome $z \in \mathcal{Z}$, let

$$\theta^u = \sum_{x_1 \in \mathcal{X}_1} p_{X_1|U}(x_1|u)\theta^{x_1} \tag{137}$$

$$\zeta^{u,z} = \sum_{x_2 \in \mathcal{X}_2} p_{X_2|U,Z}(x_2|u,z)\zeta^{x_2} \tag{138}$$

and consider the corresponding post-measurement states,

$$\omega_{A_1'\bar{E}}^{u,z} = W_z \mathcal{L}_{A_1 \to A_1'E}(\theta^u) W_z^\dagger / \mathrm{Tr}\left(W_z^\dagger W_z \mathcal{L}_{A_1 \to A_1'E}(\theta^u)\right) \tag{139}$$

$$\omega_B^{u,z} = \mathcal{N}_{A_1'A_2 \to B}(\omega_{A_1'}^{u,z} \otimes \zeta^{u,z}). \tag{140}$$

Hence, $\omega_{A_1'\bar{E}Z}$, and ω_B are the corresponding average states. The code is described below.

B. Classical Codebook Construction

(i) Generate 2^{nR_1} independent sequences $u^n(m_0)$, $m_0 \in [1:2^{nR_1}]$, at random according to $\prod_{i=1}^n p_U(u_i)$.

(ii) For every m_0, generate 2^{nR_1} conditionally independent sequences $x_1^n(m_0, m_1)$, $m_1 \in [1:2^{nR_1}]$, according to $\prod_{i=1}^n p_{X_1|U}(x_{1,i}|u_i(m_0))$.

(iii) For every m_0 and measurement sequence $z^n \in \mathcal{Z}^n$, generate 2^{nR_2} conditionally independent sequences $x_2^n(m_0, m_2, z^n)$, $m_2 \in [1:2^{nR_2}]$, according to $\prod_{i=1}^n p_{X_2|Z,U}(x_{2,i}|z_i, u_i(m_0))$.

C. Encoding and Decoding

1) Encoder 1: To send the messages $(m_1(j))_{j=1}^{T-1}$, Alice 1 performs the following. In block j, set

$$m_0 = m_1(j-1), \ m_1 = m_1(j). \tag{141}$$

Then, prepare the state

$$\rho_{A_1^n(j)} = \theta^{x_1^n(m_0, m_1)} \tag{142}$$

and send $A_1^n(j)$, for $j \in [1:T]$. As the jth transmission goes through the cribbing channel $\mathcal{L}_{A_1 \to A_1'E}$, we have

$$\rho_{A_1'^n(j)E^n(j)} = \sigma_{A_1'^n E^n}^{x_1^n(m_0, m_1)}. \tag{143}$$

The second transmitter can access the cribbing system $E^n(j)$, which are entangled with Alice 1's transmission. Cribbing is performed by a sequence of measurements that recover the messages of Alice 1. In each block, the choice of the cribbing measurement depends on the outcome in the previous block. With causal encoding, Alice 2 can prepare her input at time i based on the measurement outcome of E^i.

2) Encoder 2: To send the messages $(m_2(j))_{j=1}^{T-1}$, Alice 2 performs the following. Fix $\widetilde{m}_1(0) \equiv 1$. In block j, given the previous cribbing estimate $\widetilde{m}_1(j-1)$, do as follows, for $j = 1, 2, \ldots, T$.

(i) Apply the measurement instrument $\mathcal{W}_{E^n \to \bar{E}^n Z^n} \equiv \mathcal{W}_{E \to \bar{E}Z}^{\otimes n}$ to the cribbing system $E^n(j)$. As a result, Alice 2 obtains a measurement outcome $z^n(j)$. Hence, the average post-measurement state is

$$\rho_{A_1'^n(j)\bar{E}^n(j)Z^n(j)} = \omega_{A_1'^n \bar{E}^n Z^n}^{x_1^n(m_0, m_1)}. \tag{144}$$

(ii) Set

$$m_0 = \widetilde{m}_1(j-1), \ m_2 = m_2(j), \ z^n = z^n(j). \tag{145}$$

(iii) Prepare the state

$$\rho_{A_2^n(j)} = \zeta^{x_2^n(m_0, m_2, z^n)} \tag{146}$$

and send $A_2^n(j)$, for $j \in [1:T]$.

(iv) Measure the next cribbing estimate $\widetilde{m}_1(j)$ by applying a POVM $\mathcal{K}_{\bar{E}^n(j)Z^n(j)|m_0} = \{K_{m_1|m_0}\}$, which will be chosen later, on the joint system $\bar{E}^n(j), Z^n(j)$ following step (i).

D. Backward Decoding

The decoder recovers the messages in a backward order, *i.e.* the measurement of the jth message of Alice 1 is chosen based on the estimate of $m_1(j+1)$. Fix $\hat{m}_1(0) = \hat{m}_1(T) \equiv 1$.

In block j, for $j = T, T-1, \ldots, 2$, the decoder uses the previous estimate of $\hat{m}_1(j)$, and measures $\hat{m}_1(j-1)$ and $\hat{m}_2(j)$ using a POVM $\mathcal{D}_{B^n(j)|\hat{m}_1} = \{D_{m_0, m_2|\hat{m}_1}\}$ with $\hat{m}_1 = \hat{m}_1(j)$. Finally, in block 1, the decoder measures $\hat{m}_2(1)$ with a POVM $\mathcal{D}_{B^n(1)|m_0, m_1} = \{D'_{m_2|\hat{m}_0, \hat{m}_1}\}$, with $\hat{m}_0 = 1$ and $\hat{m}_1 = \hat{m}_1(1)$. The decoding POVMs will also be specified later.

E. Analysis of Probability of Error

Let $\delta > 0$. We use the notation $\varepsilon_i(\delta)$, $i = 1, 2, \ldots$, for terms that tend to zero as $\delta \to 0$. By symmetry, we may assume without loss of generality that the transmitters send the messages $M_1(j) = M_2(j) = 1$. Consider the following events,

$$\mathcal{E}_0(j) = \{(U^n(M_1(j-1)), X_1^n(M_1(j-1), 1),$$
$$X_2^n(M_1(j-1), 1, Z^n(j))) \notin \mathcal{A}^{\delta_1}(p_{U, X_1, X_2})\} \tag{147}$$

$$\mathcal{E}_1(j) = \{\widetilde{M}_1(j) \neq 1\} \tag{148}$$

$$\mathcal{E}_2(j) = \{\widehat{M}_1(j-1) \neq 1\} \tag{149}$$

$$\mathcal{E}_3(j) = \{\widehat{M}_2(j) \neq 1\} \tag{150}$$

for $j \in [1:T]$, with $\delta_1 \equiv \delta/(2|\mathcal{X}_2||\mathcal{U}|)$. By the union of events bound, the probability of error is bounded by

$$P_{e|m_1(j)=m_2(j)=1}^{(Tn)}(\mathcal{F}_1, \mathcal{F}_2, \mathcal{K}, \mathcal{D})$$
$$\leq \sum_{j=1}^T \Pr\left(\mathcal{E}_0(j)\right) + \sum_{j=1}^T \Pr\left(\mathcal{E}_1(j) \mid \mathcal{E}_0^c(j) \cap \mathcal{E}_1^c(j-1)\right)$$
$$+ \sum_{j=1}^T \Pr\left(\mathcal{E}_2(j) \cup \mathcal{E}_3(j) \mid \mathcal{E}_0^c(j) \cap \mathcal{E}_1^c(j) \cap \mathcal{E}_2^c(j+1)\right) \tag{151}$$

where the conditioning on $M_1(j) = M_2(j) = 1$ is omitted for convenience of notation. By the weak law of large numbers, the probability terms $\Pr\left(\mathcal{E}_0(j)\right)$ tend to zero as $n \to \infty$.

To bound the second sum, which is associated with the cribbing measurements, we use the quantum packing lemma. Alice 2's measurement is effectively a decoder for the marginal cribbing channel $\mathcal{K}_{E \to \bar{E}Z} \circ \mathcal{L}_{A_1 \to E}^{(1)}$. Given $\mathcal{E}_0^c(j)$, we have

that $(U^n(M_1(j-1)), X_1^n(M_1(j-1),1)) \in \mathcal{A}^\delta(p_{U,X_1})$. Now, observe that

$$\mathrm{Tr}\left[\Pi^\delta(\omega_{\bar{E}Z}|u^n,x_1^n)\omega_{\bar{E}^n Z^n}^{u^n,x_1^n}\right] \geq 1 - \varepsilon_1(\delta) \tag{152}$$

$$\Pi^\delta(\omega_{\bar{E}Z}|u^n)\omega_{\bar{E}^n Z^n}^{u^n}\Pi^\delta(\omega_{\bar{E}Z}|u^n) \preceq$$
$$2^{-n(H(\bar{E}Z|U)_\omega - \varepsilon_1(\delta))}\Pi^\delta(\omega_{\bar{E}Z}) \tag{153}$$

$$\mathrm{Tr}\left[\Pi^\delta(\omega_{\bar{E}Z}|u^n,x_1^n)\right] \leq 2^{n(H(\bar{E}Z|U,X_1)_\omega + \varepsilon_1(\delta))} \tag{154}$$

$$\mathrm{Tr}\left[\Pi^\delta(\omega_{\bar{E}Z}|u^n)\omega_{\bar{E}^n Z^n}^{u^n,x_1^n}\right] \geq 1 - \varepsilon_1(\delta) \tag{155}$$

for all $(u^n, x_1^n) \in \mathcal{A}^\delta(p_{U,X_1})$, by (17)-(20), respectively. Since the codebooks are statistically independent of each other, we have by the single-user quantum packing lemma [15] [56, Lemma 12], that there exists a POVM $K_{m_1|u^n}$ such that $\mathrm{Pr}\left(\mathscr{E}_1(j) \mid \mathscr{E}_0^c(j) \cap \mathscr{E}_1^c(j-1)\right) \leq 2^{-n(I(X_1;\bar{E}Z|U)_\rho - R_1 - \varepsilon_2(\delta))}$, which tends to zero as $n \to \infty$, provided that

$$R_1 < I(X_1; \bar{E}Z|U)_\omega - \varepsilon_2(\delta). \tag{156}$$

The last sum in the RHS of (151) tends to zero as in Appendix E, based on our generalized packing lemma (see Lemma 2), for

provided that $R_2 < I(X_2; B|X_1U)_\omega$ and $R_1 + R_2 < I(X_1X_2; B)_\omega$. This completes the proof outline for part 1.

F. Part 2

Consider the classical-quantum MAC $\mathcal{N}_{X_1X_2 \to B} \circ \mathrm{id}_{X_1 \to X_1X_1}$ with either causal or non-causal noiseless cribbing. Observe that it suffices to show the direct part for causal cribbing, and the converse part for the non-causal setting, since $\mathcal{C}_{\mathrm{caus}}(\mathcal{N} \circ \mathcal{L}) \subseteq \mathcal{C}_{\mathrm{n\text{-}c}}(\mathcal{N} \circ \mathcal{L})$ (see Remark 3). As the cribbing system stores a perfect copy of Alice 1's classical input in this case, *i.e.* $E \equiv \bar{E} \equiv Z \equiv X_1$, we have $I(X_1; \bar{E}Z|U)_\omega = H(X_1|U)$. Thereby, achievability immediately follows from part 1 of the theorem.

Now, we show the converse part for non-causal cribbing. Suppose that Alice k chooses a message M_k uniformly at random, for $k = 1, 2$. Alice 1 transmits $X_1^n = \mathcal{F}_1(M_1)$. Thereby, Alice 2 measures X_1^n from the cribbing system and transmits $X_2^n = \mathcal{F}_2(M_2, X_1^n)$. Then, X_1^n and X_2^n are sent through n copies of the classical-quantum MAC $\mathcal{N}_{X_1X_2 \to B}$. Bob receives the output systems B^n and performs a measurement in order to obtain an estimate $(\widehat{M_1}, \widehat{M_2})$ of the message pair. Consider a sequence of codes such that the average probability of error tends to zero. Since the encoding map $\mathcal{F}_1$ is classical, there exists a random element S_1 such that $X_1^n \equiv f_1(M_1, S_1)$, where f_1 is a deterministic function. Then, Alice 1's transmission rate satisfies

$$nR_1 = H(X_1^n|S_1)$$
$$\leq \sum_{i=1}^n H(X_{1,i}). \tag{157}$$

To bound the second transmission rate and the rate sum, we apply Fano's inequality as in Appendix E, hence

$$nR_2 \leq I(M_2; B^n|M_1, S_1)_\rho + n\varepsilon_n'' \tag{158}$$

(see (121)). Using the chain rule, the last bound can be expressed as

$$nR_2 \leq \sum_{i=1}^n I(M_2; B_i|M_1, S_1, B^{i-1})_\rho + n\varepsilon_n''$$
$$= \sum_{i=1}^n I(M_2; B_i|M_1, S_1, X_{1,i}, B^{i-1})_\rho + n\varepsilon_n'' \tag{159}$$

since $X_{1,i}$ is a deterministic function of (M_1, S_1). Observe that

$$I(M_2; B_i|M_1, S_1, X_{1,i}, B^{i-1})_\rho$$
$$\leq I(M_2, X_{2,i}; B_i|M_1, S_1, X_{1,i}, B^{i-1})_\rho$$
$$= H(B_i|M_1, S_1, X_{1,i}, B^{i-1})_\rho$$
$$\quad - H(B_i|X_{1,i}, X_{2,i}, M_1, M_2, S_1, B^{i-1})_\rho$$
$$\leq H(B_i|X_{1,i})_\rho - H(B_i|X_{1,i}, X_{2,i}, M_1, M_2, S_1, B^{i-1})_\rho. \tag{160}$$

Consider the second term, and observe that given $X_{1,i}$ and $X_{2,i}$, the system B_i is in a product state with the joint system of (M_1, M_2, S_1, B^{i-1}). Thus, the second term equals $H(B_i|X_{1,i}, X_{2,i})_\rho$. Therefore, by (159)-(160),

$$nR_2 \leq \sum_{i=1}^n I(X_{2,i}; B_i|X_{1,i})_\rho + \varepsilon_n''. \tag{161}$$

Similarly, the rate sum is bounded by

$$n(R_1 + R_2) \leq \sum_{i=1}^n I(X_{1,i}X_{2,i}; B_i)_\rho + n\varepsilon_n' \tag{162}$$

(see (126)).

To obtain the single-letter converse, consider a time index K that is drawn uniformly at random, from $[1:n]$, independently of M_1, M_2, and S_1. By (157), (161), and (162),

$$R_1 \leq H(X_{1,K}|K) \tag{163}$$
$$R_2 - \varepsilon_n'' \leq I(X_{2,K}; B_K|X_{1,K}, K)_\rho \tag{164}$$
$$R_1 + R_2 - \varepsilon_n' \leq I(X_{1,K}X_{2,K}; B_K|K)_\rho. \tag{165}$$

Furthermore, define a joint state $\omega_{X_1X_2B}$ by identifying $X_1 \equiv X_{1,K}$, $X_2 \equiv X_{2,K}$, and $B \equiv B_K$. That is, $\omega_{KX_1X_2B} \equiv \sum_{i=1}^n \sum_{x_1,x_2} \frac{1}{n} p_{X_{1,i}}(x_1) p_{X_{2,i}}(x_2) |i\rangle\langle i| \otimes |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \otimes \mathcal{N}(x_1, x_2)$. Thus, the individual rates are bounded by

$$R_1 = H(X_1|K)$$
$$\leq H(X_1) \tag{166}$$

and

$$R_2 - \varepsilon_n'' \leq I(X_2; B|X_1, K)_\omega$$
$$\leq I(X_2, K; B|X_1)_\omega$$
$$= I(X_2; B|X_1)_\omega, \tag{167}$$

and the rate sum by

$$R_1 + R_2 - \varepsilon_n' \leq I(X_1X_2; B|K)_\omega$$
$$\leq I(X_1X_2K; B)_\omega$$
$$= I(X_1X_2; B)_\omega \tag{168}$$

as $I(K; B|X_1 X_2)_\omega = 0$ since the channel has a memoryless product form. This completes the proof of Theorem 6. $\square$

APPENDIX F
PROOF OF THEOREM 7

Consider the quantum MAC $\mathcal{N} \circ \mathcal{L}$ with strictly-causal cribbing. Here, we prove the partial-decode forward bound on the capacity region. We show that for every $\delta_1, \delta_2, \varepsilon_0 > 0$, there exists a $(2^{n(R_1-\delta_1)}, 2^{n(R_2-\delta_2)}, n, \varepsilon_0)$ code for $\mathcal{N}_{A_1' A_2 \to B} \circ \mathcal{L}_{A_1 \to A_1' E}$ with strictly-causal cribbing at Encoder 2, provided that $(R_1, R_2) \in \mathcal{R}_{\text{s-c}}^{\text{PDF}}(\mathcal{N} \circ \mathcal{L})$. To prove achievability, we extend the classical block Markov coding with backward decoding to the quantum setting, and then apply the quantum packing lemma.

As before, we use T transmission blocks, where each block consists of n channel uses. Given strictly-causal cribbing, Encoder 2 has access to the cribbing measurements from the previous blocks. Alice 1 sends $T - 1$ pairs of messages, and Alice 2 sends $T - 1$ messages. Specifically, Alice 1 sends the pairs $(\ell_1(j), m_1'(j))_{j=1}^{T-1}$ at rates $(R_0, R_1 - R_0)$, with $R_0 < R_1$, while Alice 2 sends $(m_2(j))_{j=1}^{T-1}$ at rate R_2. Let us fix $\ell_1(0) = m_1'(0) = m_2(0) \equiv 1$ and $\ell_1(T) = m_1'(T) = m_2(T) \equiv 1$. Hence, the coding rate for User k is $\left(\frac{T-1}{T}\right) R_k$, which tends to R_k as the number of blocks T grows to infinity.

Let $\{p_U(u) p_{X_0|U} p_{X_1|X_0,U}(x_1|x_0, u) p_{X_2|X_0,U}(x_2|x_0, u), \theta^{x_1} \otimes \zeta^{x_2}\}$ be a given ensemble over $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$.

Define

$$\omega_{A_1' E}^{x_1} = \mathcal{L}_{A_1 \to A_1' E}(\theta^{x_1}), \qquad (169)$$

$$\omega_B^{x_1, x_2} = \mathcal{N}_{A_1' A_2 \to B}(\omega_{A_1'}^{x_1} \otimes \zeta^{x_2}). \qquad (170)$$

In addition, let

$$\theta^{x_0, u} = \sum_{x_1 \in \mathcal{X}_1} p_{X_1|X_0, U}(x_1|x_0, u) \theta^{x_1} \qquad (171)$$

$$\zeta^{x_0, u} = \sum_{x_2 \in \mathcal{X}_2} p_{X_2|X_0, U}(x_2|x_0, u) \zeta^{x_2} \qquad (172)$$

$$\omega_{A_1' E}^{x_0, u} = \mathcal{L}_{A_1 \to A_1' E}(\theta^{x_0, u}) \qquad (173)$$

$$\omega_B^{x_0, u} = \mathcal{N}_{A_1' A_2 \to B}(\omega_{A_1'}^{x_0, u} \otimes \zeta^{x_0, u}) \qquad (174)$$

for $u \in \mathcal{U}$ and $x_0 \in \mathcal{X}_0$. Hence, $\omega_{A_1' E}$ and ω_B are the corresponding average states.

The partial decode-forward coding scheme is described below and depicted in Figure 4.

A. Classical Codebook Construction

(i) Generate 2^{nR_0} independent sequences $u^n(\ell_0)$, $\ell_0 \in [1 : 2^{nR_0}]$, at random according to $\prod_{i=1}^n p_U(u_i)$.

(ii) For every ℓ_0, generate 2^{nR_0} conditionally independent sequences $x_0^n(\ell_0, \ell_1)$, $\ell_1 \in [1 : 2^{nR_0}]$, according to $\prod_{i=1}^n p_{X_0|U}(x_{0,i}|u_i(\ell_0))$.

(iii) For every (ℓ_0, ℓ_1), generate $2^{n(R_1-R_0)}$ conditionally independent sequences $x_1^n(\ell_0, \ell_1, m_1')$, $m_1' \in [1 : 2^{n(R_1-R_0)}]$, according to $\prod_{i=1}^n p_{X_1|U}(x_{1,i}|x_{0,i}(\ell_0, \ell_1), u_i(\ell_0))$.

(iv) For every ℓ_0, generate 2^{nR_2} conditionally independent sequences $x_2^n(\ell_0, m_2)$, $m_2 \in [1 : 2^{nR_2}]$, according to $\prod_{i=1}^n p_{X_2|U}(x_{2,i}|u_i(m_0))$.

B. Encoding and Decoding

1) *Encoder 1:* To send the messages $(\ell_1(j), m_1'(j))_{j=1}^{T-1}$, Alice 1 performs the following. In block j, set

$$\ell_0 = \ell_1(j-1), \ \ell_1 = \ell_1(j), \ m_1' = m_1'(j). \qquad (175)$$

Then, prepare the state

$$\rho_{A_1^n(j)} = \theta^{x_1^n(\ell_0, \ell_1, m_1')} \qquad (176)$$

and send $A_1^n(j)$, for $j \in [1 : T]$. See the third row in Figure 4. As the jth transmission goes through the cribbing channel $\mathcal{L}_{A_1 \to A_1' E}$, we have

$$\rho_{A_1'^n(j) E^n(j)} = \omega_{A_1'^n E^n}^{x_1^n(\ell_0, \ell_1, m_1')}. \qquad (177)$$

The second transmitter can access the cribbing system $E^n(j)$, which are entangled with Alice 1's transmission. Cribbing is performed by a sequence of measurements that recover the messages of Alice 1. In each block, the choice of the cribbing measurement depends on the outcome in the previous block. With strictly-causal encoding, Alice 2 must perform the cribbing measurement at the end of the block, after she has already sent $A_2^n(j)$.

2) *Encoder 2:* To send the messages $(m_2(j))_{j=1}^{T-1}$, Alice 2 performs the following. Fix $\widetilde{\ell}_1(0) \equiv 1$. In block j, given the previous cribbing estimate $\widetilde{\ell}_1(j-1)$, do as follows, for $j = 1, 2, \ldots, T$.

(i) Set

$$\ell_0 = \widetilde{\ell}_1(j-1), \ m_2 = m_2(j). \qquad (178)$$

(ii) Prepare the state

$$\rho_{A_2^n(j)} = \zeta^{x_2^n(\ell_0, m_2)} \qquad (179)$$

and send $A_2^n(j)$, for $j \in [1 : T]$. See the fourth and fifth rows in Figure 4.

(iii) Measure the next cribbing estimate $\widetilde{\ell}_1(j)$ by applying a POVM $\mathcal{K}_{E^n(j)|\ell_0} = \{K_{\ell_1|u^n}\}$, with $u^n \equiv u^n(\ell_0)$, that will be chosen later.

Here, the decoder performs two measurements on each output block. As in [71] and [56], we use the gentle measurement lemma to verify that the decoder's measurement does not cause a "state collapse" at the output.

3) *Sequential Decoding:* To estimate Alice 1 and Alice 2's messages, Bob performs the following.

(i) First, the decoder recovers the messages $(\ell_1(j))$ and $(m_2(j))$, using backward decoding. That is, the measurement of the jth message is chosen based on the estimate of $\ell_1(j + 1)$. See the bottom part of Figure 4. Fix $\ell_1(0) = \ell_1(T) \equiv 1$.
In block j, for $j = T, T-1, \ldots, 1$, the decoder uses the previous estimate of $\hat{\ell}_1(j)$, and measures $\hat{\ell}_1(j - 1)$ and $\hat{m}_2(j)$ using a POVM $\mathcal{D}_{B^n(j)|\ell_1} = \{D_{\ell_0, m_2|\ell_1}\}$ with $\ell_1 = \hat{\ell}_1(j)$.

(ii) Next, the decoder recovers the messages $(m_1'(j))$, going in the forward direction. For $j = 1, 2, \ldots, T$, the decoder uses the estimate of $\hat{\ell}_1(j - 1)$, $\hat{\ell}_1(j)$, and $\hat{m}_2(j)$ from the previous step, and measures $\hat{m}_1'(j)$ using a POVM $\mathcal{G}_{B^n(j)|\ell_0, \ell_1, m_2} = \{G_{m_1'|x_0^n, x_2^n}\}$, based on the knowledge

Block	1	2	$\cdots$	$T-1$	T
U^n	$u^n(1)$	$u^n(\ell_1(1))$	$\cdots$	$u^n(\ell_1(T-2))$	$u^n(\ell_1(T-1))$
X_0^n	$x_0^n(1,\ell_1(1))$	$x_0^n(\ell_1(1),\ell_1(2))$	$\cdots$	$x_0^n(\ell_1(T-2),\ell_1(T-1))$	$x_0^n(\ell_1(T-1),1)$
A_1	$x_1^n(1,\ell_1(1),m_1'(1))$	$x_1^n(\ell_1(1),\ell_1(2),m_1'(2))$	$\cdots$	$x_1^n(\ell_1(T-2),\ell_1(T-1),m_1'(T-1))$	$x_1^n(\ell_1(T-1),1,1)$
E	$\widetilde{\ell}_1(1) \rightarrow$	$\widetilde{\ell}_1(2) \rightarrow$	$\cdots$	$\widetilde{\ell}_1(T-1) \rightarrow$	$\emptyset$
A_2	$x_2^n(1,m_2(1))$	$x_2^n(\widetilde{\ell}_1(1),m_2(2))$	$\cdots$	$x_2^n(\widetilde{\ell}_1(T-2),m_2(T-1))$	$x_2^n(\widetilde{\ell}_1(T-1),1)$
B	$\emptyset$ $\hat{m}_2(1)$ $\hat{m}_1'(1)$	$\leftarrow \hat{\ell}_1(1)$ $\hat{m}_2(2)$ $\hat{m}_1'(2)$	$\cdots$ $\cdots$ $\cdots$	$\leftarrow \hat{\ell}_1(T-2)$ $\hat{m}_2(T-1)$ $\hat{m}_1'(T-1)$	$\leftarrow \hat{\ell}_1(T-1)$ $\emptyset$ $\emptyset$

Fig. 4. Partial decode-forward cribbing scheme. The block index $j \in [1:T]$ is indicated at the top. In the following rows, we have the corresponding elements: (1), (2) auxiliary sequences; (3) codewords of Alice 1; (4) cribbing estimates by Alice 2; (5) codewords of Alice 2; (6) estimated messages at the decoder. The arrows in the fourth row indicate that the Alice 2 measures and encodes forward with respect to the block index, while the arrows in the sixth row indicate that Bob decodes backwards.

of $\ell_0 = \hat{\ell}_1(j-1)$, $\ell_1 = \hat{\ell}_1(j)$, $m_2 = \hat{m}_2(j)$, $x_0^n = x_0^n(\ell_0,\ell_1)$, and $x_2^n = x_2^n(\ell_0,m_2)$.

C. Analysis of Probability of Error

Let $\delta > 0$. We use the notation $\varepsilon_i(\delta)$, $i = 1, 2, \ldots$, for terms that tend to zero as $\delta \to 0$. By symmetry, we may assume without loss of generality that the transmitters send the messages $L_1(j) = M_1'(j) = M_2(j) = 1$. Consider the following events,

$$\mathscr{E}_0(j) = \{(U^n(L_1(j-1)), X_0^n(L_1(j-1),1),$$
$$X_1^n(L_1(j-1),1,1), X_2^n(L_1(j-1),1))$$
$$\notin \mathcal{A}^{\delta_1}(p_{U,X_0,X_1,X_2})\} \tag{180}$$

$$\mathscr{E}_1(j) = \{\widetilde{L}_1(j) \neq 1\} \tag{181}$$
$$\mathscr{E}_2(j) = \{\hat{L}_1(j-1) \neq 1\} \tag{182}$$
$$\mathscr{E}_3(j) = \{\widehat{M}_2(j)) \neq 1\} \tag{183}$$
$$\mathscr{E}_4(j) = \{\widehat{M}_1'(j) \neq 1\} \tag{184}$$

for $j \in [1:T]$, with $\delta_1 \equiv \delta/(2|\mathcal{X}_2||\mathcal{U}|)$. By the union of events bound, the probability of error is bounded by

$$P_{e|m_1(j)=m_2(j)=1}^{(Tn)}(\mathcal{F}_1, \mathcal{F}_2, \mathcal{K}, \mathcal{D})$$
$$\leq \sum_{j=1}^{T} \Pr(\mathscr{E}_0(j)) + \sum_{j=1}^{T} \Pr(\mathscr{E}_1(j) \mid \mathscr{E}_0^c(j) \cap \mathscr{E}_1^c(j-1))$$
$$+ \sum_{j=1}^{T} \Pr(\mathscr{E}_2(j) \cup \mathscr{E}_3(j) \mid \mathscr{E}_0^c(j) \cap \mathscr{E}_1^c(j) \cap \mathscr{E}_2^c(j+1))$$
$$+ \sum_{j=1}^{T} \Pr(\mathscr{E}_4(j) \mid \mathscr{E}_0^c(j) \cap \mathscr{E}_1^c(j) \cap \mathscr{E}_2^c(j+1)$$
$$\cap \mathscr{E}_2^c(j) \cap \mathscr{E}_3^c(j)) \tag{185}$$

where the conditioning on $L_1(j) = M_1'(j) = M_2(j) = 1$ is omitted for convenience of notation. By the weak law of large numbers, the first sum tends to zero as $n \to \infty$. To bound the second and third sums, we use the arguments in Appendix F,

replacing m_0, m_1 by ℓ_0, ℓ_1, respectively, and thus, replacing R_1 and X_1 by R_0 and X_0, respectively. Thus, there exists a cribbing measurement $K_{\ell_1|u^n}$ such that the second sum tends to zero if

$$R_0 < I(X_0; E|U)_\omega - \varepsilon_2(\delta), \tag{186}$$

by the single-user packing lemma (see (108)). Furthermore, there exists a POVM $D_{\ell_0,m_2|\ell_1}$ such that the third sum tends to zero as $n \to \infty$, provided that

$$R_2 < I(X_2; B|X_0 U)_\omega - \varepsilon_4(\delta) \tag{187}$$

and

$$R_0 + R_2 < I(X_0 X_2; B)_\omega - \varepsilon_4(\delta) \tag{188}$$

by Lemma 2 and the arguments in Appendix F that lead to (110)-(111).

It remains to show that the fourth sum in the RHS of (185) tends to zero as well. As in [71] and [56], we observe that due to the packing lemma inequality (25), the gentle measurement lemma [69], [72] implies that the post-measurement state $\tilde{\rho}_{B^n}$ is close to the original state ρ_{B^n} in the sense that

$$\frac{1}{2}\|\tilde{\rho}_{B^n} - \rho_{B^n}\|_1 \leq \left(4 \cdot 2^{-n(I(X_0 X_2; B)_\rho - (R_0 + R_2) - \varepsilon_4(\delta))} \right.$$
$$+ 2^{-n(I(X_2; B|X_0 U)_\omega - R_2 - \varepsilon_4(\delta))}$$
$$\left. + 2^{-n(I(X_0 X_2; B|U)_\omega - R_2 - \varepsilon_4(\delta))}\right)^{1/2}$$
$$\leq \varepsilon_5(\delta) \tag{189}$$

for sufficiently large n and rates as in (187)-(188). Therefore, the distribution of measurement outcomes when $\tilde{\rho}_{B^n}$ is measured is roughly the same as if the measurements $\mathcal{D}_{B^n(j)|L_1(j-1)}$ were never performed. To be precise, the difference between the probability of a measurement outcome $\hat{m}_1'$ when $\tilde{\rho}_{B^n}$ is measured and the probability when ρ_{B^n} is measured is bounded by $\varepsilon_5(\delta)$ in absolute value (see [55, Lem. 9.11]). Furthermore,

$$\text{Tr}\left[\Pi^\delta(\omega_B|x_1^n, x_0^n, x_2^n)\omega_{B_1^n}^{x_1^n,x_0^n,x_2^n}\right] \geq 1 - \varepsilon_6(\delta) \tag{190}$$

$$\Pi^\delta(\omega_B|x_0^n, x_2^n)\omega_{B^n}^{x_1^n, x_0^n, x_2^n}\Pi^\delta(\omega_B|x_0^n, x_2^n) \preceq$$
$$2^{-n(H(B|X_0 X_2)_\omega - \varepsilon_6(\delta))}\Pi^\delta(\omega_B|x_0^n, x_2^n) \tag{191}$$

$$\text{Tr}\left[\Pi^\delta(\omega_B|x_1^n, x_0^n, x_2^n)\right] \leq 2^{n(H(B|X_1 X_0 X_2)_\omega + \varepsilon_6(\delta))} \tag{192}$$

$$\text{Tr}\left[\Pi^\delta(\omega_B|x_0^n, x_2^n)\omega_{B^n}^{x_1^n, x_0^n, x_2^n}\right] \geq 1 - \varepsilon_6(\delta) \tag{193}$$

for all $(x_0^n, x_1^n, x_2^n) \in \mathcal{A}^{\delta/2}(p_{X_0, X_1, X_2})$, by (17)-(20), respectively. Thus, by the single-user quantum packing lemma [15] [56, Lemma 12], there exists a POVM $G_{m_1'|x_0^n, x_2^n}$ such that the error term in the fourth sum tends to zero as $n \to \infty$, provided that

$$R_1 - R_0 < I(X_1; B|X_0 X_2)_\omega - \varepsilon_2(\delta). \tag{194}$$

We have thus shown that a rate pair (R_1, R_2) is achievable if

$$R_1 > R_0$$
$$R_0 < I(X_0; E|U)_\omega$$
$$R_1 - R_0 < I(X_1; B|X_0 X_2)_\omega$$
$$R_0 + R_2 < I(X_0 X_2; B)_\omega$$
$$R_2 < I(X_2; B|X_0 U)_\omega \tag{195}$$

(see (186), (187), (188) and (194)). By eliminating R_0, we obtain the following region

$$R_1 < I(X_0; E|U)_\omega + I(X_1; B|X_0 X_2)_\omega$$
$$R_1 + R_2 < I(X_0 X_1 X_2; B)_\omega = I(X_1 X_2; B)_\omega$$
$$R_2 < I(X_2; B|X_0 U)_\omega. \tag{196}$$

Then, set $X_0 = (U, V)$. This completes the achievability proof for the partial decode-forward inner bound. □

APPENDIX G
PROOF OF THEOREM 8

A. Part 1

Consider the classical-quantum MAC $\mathcal{N}_{X_1 X_2 \to B} \circ Q_{Z|X_1}$ with strictly-causal noisy cribbing. Suppose that Alice 1 chooses m_1 uniformly at random, and prepares an input state $\rho_{A_1^n}^{m_1}$. Consider a sequence of codes $(\mathcal{F}_{1n}, \mathcal{F}_{2n}, \mathcal{K}_n, \mathcal{D}_n)$ such that the average probability of error tends to zero, hence the error probabilities $\Pr\left(\widehat{M}_1 \neq M_1|M_2\right)$, $\Pr\left(\widehat{M}_2 \neq M_2|M_1\right)$, and $\Pr\left((\widehat{M}_1, \widehat{M}_2) \neq (M_1, M_2)\right)$, are bounded by some α_n which tends to zero as $n \to \infty$. By Fano's inequality [70], it follows that

$$H(M_1|\widehat{M}_1, M_2) \leq n\varepsilon_n \tag{197}$$

$$H(M_2|\widehat{M}_2, M_1) \leq n\varepsilon_n' \tag{198}$$

$$H(M_1, M_2|\widehat{M}_1, \widehat{M}_2) \leq n\varepsilon_n'' \tag{199}$$

where $\varepsilon_n, \varepsilon_n', \varepsilon_n''$ tend to zero as $n \to \infty$. Hence,

$$
\begin{aligned}
nR_1 &= H(M_1|M_2, S_1, S_2) \\
&= I(M_1; \widehat{M}_1|M_2, S_2)_\rho + H(M_1|\widehat{M}_1, M_2, S_1, S_2) \\
&\leq H(M_1|M_2, S_1, S_2) \\
&= I(M_1; \widehat{M}_1|M_2, S_2)_\rho + H(M_1|\widehat{M}_1, M_2) \\
&\leq I(M_1; \widehat{M}_1|M_2, S_1, S_2)_\rho + n\varepsilon_n \\
&\leq I(M_1; B^n|M_2, S_1, S_2)_\rho + n\varepsilon_n \tag{200}
\end{aligned}
$$

where the last inequality follows from the Holevo bound (see Ref. [57, Theo. 12.1]). Now,

$$
\begin{aligned}
&I(M_1; B^n|M_2, S_1, S_2)_\rho \\
&\leq I(M_1; B^n Z^n|M_2 S_1 S_2)_\rho \\
&= \sum_{i=1}^n I(M_1; B_i Z_i|B^{i-1} Z^{i-1} M_2 S_1 S_2)_\rho \\
&= \sum_{i=1}^n I(M_1 X_{1,i}; B_i Z_i|B^{i-1} Z^{i-1} M_2 S_1 S_2 X_{2,i})_\rho \tag{201}
\end{aligned}
$$

where the last equality holds since $X_{1,i}$ and $X_{2,i}$ are deterministic functions of (M_1, S_1) and (M_2, Z^{i-1}, S_2), respectively. Observe that

$$
\begin{aligned}
&I(M_1 X_{1,i}; B_i Z_i|B^{i-1} Z^{i-1} M_2 S_1 S_2 X_{2,i})_\rho \\
&= H(B_i Z_i|B^{i-1} Z^{i-1} M_2 S_1 S_2 X_{2,i})_\rho \\
&\quad - H(B_i Z_i|B^{i-1} Z^{i-1} M_1 M_2 S_1 S_2 X_{1,i} X_{2,i})_\rho \\
&\leq H(B_i Z_i|X_{2,i})_\rho \\
&\quad - H(B_i Z_i|B^{i-1} Z^{i-1} M_1 M_2 S_1 S_2 X_{1,i} X_{2,i})_\rho \tag{202}
\end{aligned}
$$

since conditioning cannot increase entropy. The second term equals $H(B_i Z_i|X_{1,i} X_{2,i})_\rho$ because, given $(X_{1,i}, X_{2,i}) = (x_1, x_2)$, the joint cribbing and output system $B_i Z_i$ has no correlation with $B^{i-1} Z^{i-1} M_1 M_2 S_1 S_2$. Thus, by (200)-(202),

$$R_1 \leq \frac{1}{n}\sum_{i=1}^n I(X_{1,i}; B_i Z_i|X_{2,i})_\rho + \varepsilon_n. \tag{203}$$

Similarly,

$$R_2 \leq \frac{1}{n}\sum_{i=1}^n I(X_{2,i}; B_i|X_{1,i})_\rho + \varepsilon_n' \tag{204}$$

and

$$R_1 + R_2 = \frac{1}{n}\sum_{i=1}^n I(X_{1,i} X_{2,i}; B_i)_\rho + \varepsilon_n' \tag{205}$$

Defining a time index U that is drawn from $[1 : n]$ independently uniformly at random, it follows that

$$R_1 - \varepsilon_n \leq I(X_1; BZ|X_2 U)_\omega \tag{206}$$
$$R_2 - \varepsilon_n' \leq I(X_2; B|X_1 U)_\omega \tag{207}$$

and

$$
\begin{aligned}
R_1 + R_2 - \varepsilon_n'' &\leq I(X_1 X_2; B|U)_\omega \\
&\leq I(X_1 X_2 U; B)_\omega \\
&\leq I(X_1 X_2; B)_\omega. \tag{208}
\end{aligned}
$$

where $\omega_{U X_1 Z X_2 B}$ is defined as

$$
\begin{aligned}
\omega_{U X_1 Z X_2 B} &\equiv \frac{1}{n}\sum_{i=1}^n \Big(|i\rangle\langle i| \otimes \sum_{x_1, x_2} p_{X_{1,i}}(x_1) Q(z|x_1) \\
&\quad \cdot p_{X_{2,i}}(x_2)|x_1, z, x_2\rangle\langle x_1, z, x_2| \otimes \mathcal{N}(x_1, x_2)\Big). \tag{209}
\end{aligned}
$$

B. Part 2

Suppose that $Z = g(X_1)$, where $g : \mathcal{X}_1 \to \mathcal{Z}$ is a deterministic function. We determine the capacity region using the inner bound in Theorem 7 and the outer bound in part 1. Observe that it suffices to consider the first rate, since the first inequality is the only difference between the inner and outer bounds (*cf.* (47) and (51)). Consider the partial-decode forward inner bound in Theorem 7. Since $Z = g(X_1)$, we can set $V = Z$ in the RHS of (47). Hence, the first rate R_1 is bounded by

$$I(V; Z|U) = H(Z|U). \tag{210}$$

The converse part follows from part 1, as

$$
\begin{aligned}
I(X_1; BZ|X_2U)_\omega &= I(X_1; Z|X_2U) + I(X_1; B|X_2UZ)_\omega \\
&\leq H(Z|X_2U) + I(X_1; B|X_2UZ)_\omega \\
&\leq H(Z|U) + I(X_1; B|X_2UZ)_\omega. \tag{211}
\end{aligned}
$$

This completes the proof of Theorem 8. $\qquad\square$

References

[1] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probab.*, vol. 2, no. 5, pp. 805–814, Oct. 1974.

[2] P. Li and J. Xu, "Fundamental rate limits of UAV-enabled multiple access channel with trajectory optimization," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 458–474, Jan. 2020.

[3] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple Access Techniques for 5G Wireless Networks and Beyond*, vol. 159. Cham, Switzerland: Springer, 2019.

[4] M. Shirvanimoghaddam *et al.*, "Massive non-orthogonal multiple access for cellular IoT: Potentials and limitations," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 55–61, Sep. 2017.

[5] R. De Gaudenzi, O. Del Rio Herrero, G. Gallinaro, S. Cioni, and P.-D. Arapoglou, "Random access schemes for satellite networks, from VSAT to M2M: A survey," *Int. J. Satell. Commun. Netw.*, vol. 36, no. 1, pp. 66–107, 2018.

[6] K. Terplan and P. A. Morreale, *The Telecommunications Handbook*. Danvers, MA, USA: CRC Press, 2018.

[7] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, Apr. 2009.

[8] X. Chen, H.-H. Chen, and W. Meng, "Cooperative communications for cognitive radio networks—From theory to applications," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1180–1192, 3rd Quart., 2014.

[9] B. Lyu, Z. Yang, H. Guo, F. Tian, and G. Gui, "Relay cooperation enhanced backscatter communication for Internet-of-Things," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2860–2871, Apr. 2019.

[10] F. M. J. Willems and E. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 3, pp. 313–327, May 1985.

[11] P. Loock *et al.*, "Extending quantum links: Modules for fiber- and memory-based quantum repeaters," *Adv. Quantum Technol.*, vol. 3, no. 11, Nov. 2020, Art. no. 1900141.

[12] R. Bassoli *et al.*, *Quantum Communication Networks* (Foundations in Signal Processing, Communications and Networking), vol. 23. Cham, Switzerland: Springer, 2021.

[13] A. Winter, "The capacity of the quantum multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 3059–3065, Nov. 2001.

[14] I. Savov, "Network information theory for classical-quantum channels," Ph.D. dissertation, School Comput. Sci., McGill Univ., Montreal, QC, Canada, 2012.

[15] M.-H. Hsieh, I. Devetak, and A. Winter, "Entanglement-assisted capacity of quantum multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3078–3090, Jul. 2008.

[16] H. Shi, M.-H. Hsieh, S. Guha, Z. Zhang, and Q. Zhuang, "Entanglement-assisted capacity regions and protocol designs for quantum multiple-access channels," *npj Quantum Inf.*, vol. 7, no. 1, pp. 1–9, Dec. 2021.

[17] H. Boche and J. Nötzel, "The classical-quantum multiple access channel with conferencing encoders and with common messages," *Quantum Inf. Process.*, vol. 13, no. 12, pp. 2595–2617, Dec. 2014.

[18] S. Diadamo and H. Boche, "The simultaneous identification capacity of the classical-quantum multiple access channel with stochastic encoders for transmission," 2019, *arXiv:1903.03395*.

[19] F. Leditzky, M. A. Alhejji, J. Levin, and G. Smith, "Playing games with multiple access channels," *Nature Commun.*, vol. 11, no. 1, pp. 1–5, Dec. 2020.

[20] G. Brassard, A. Broadbent, and A. Tapp, "Quantum pseudo-telepathy," *Found. Phys.*, vol. 35, no. 11, pp. 1877–1907, 2005.

[21] U. Pereg, C. Deppe, and H. Boche, "Quantum broadcast channels with cooperating decoders: An information-theoretic perspective on quantum repeaters," *J. Math. Phys.*, vol. 62, no. 6, Jun. 2021, Art. no. 062204.

[22] J. Yard, "Simultaneous classical-quantum capacities of quantum multiple access channels," Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, Mar. 2005.

[23] J. Yard, P. Hayden, and I. Devetak, "Capacity theorems for quantum multiple-access channels: Classical-quantum and quantum-quantum capacity regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3091–3113, Jul. 2008.

[24] M. Hayashi and N. Cai, "Exponent for classical-quantum multiple access channel," 2017, *arXiv:1701.02939*.

[25] L. Czekaj and P. Horodecki, "Nonadditivity effects in classical capacities of quantum multiple-access channels," 2008, *arXiv:0807.3977*.

[26] H. Aghaee and B. Akhbari, "Classical-quantum multiple access wiretap channel," in *Proc. 16th Int. ISC Iranian Soc. Cryptol. Conf. Inf. Secur. Cryptol. (ISCISC)*, Aug. 2019, pp. 99–103.

[27] H. Boche, G. Janßen, and S. Saeedinaeeni, "Universal superposition codes: Capacity regions of compound quantum broadcast channel with confidential messages," *J. Math. Phys.*, vol. 61, no. 4, Apr. 2020, Art. no. 042204.

[28] T. Das, K. Horodecki, and R. Pisarczyk, "Secure communication over generalised quantum multiple access channels," 2021, *arXiv:2106.13310*.

[29] S. Chakraborty, A. Nema, and P. Sen, "One-shot inner bounds for sending private classical information over a quantum MAC," 2021, *arXiv:2105.06100*.

[30] M. Hayashi and A. Vazquez-Castro, "Computation-aided classical-quantum multiple access to boost network communication speeds," 2021, *arXiv:2105.14505*.

[31] G. P. Fettweis and H. Boche, "6G: The personal tactile internet- and open questions for information theory," *IEEE BITS Inf. Theory Mag.*, early access, Oct. 12, 2021, doi: 10.1109/MBITS.2021.3118662.

[32] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?" *Nature Electron.*, vol. 3, no. 1, pp. 20–29, Jan. 2020.

[33] G. P. Fettweis and H. Boche, "On 6G and trustworthiness," *Commun. ACM*, 2022.

[34] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 118–125, Aug. 2020.

[35] F. Fitzek and H. Boche, "6G-life: Digital transformation and sovereignty of future communication networks," *IEEE Network*, vol. 35, no. 6, pp. 3–4, Nov/Dec. 2021.

[36] H. Asnani and H. H. Permuter, "Multiple-access channel with partial and controlled cribbing encoders," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2252–2266, Apr. 2013.

[37] T. Kopetz, H. H. Permuter, and S. Shamai, "Multiple access channels with combined cooperation and partial cribbing," *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 825–848, Feb. 2016.

[38] N. Helal, M. Bloch, and A. Nosratinia, "Cooperative resolvability and secrecy in the cribbing multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 66, no. 9, pp. 5429–5447, Sep. 2020.

[39] N. Helal and A. Nosratinia, "Multiple access wiretap channel with cribbing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 739–743.

[40] W. Huleihel and Y. Steinberg, "Channels with cooperation links that may be absent," *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 5886–5906, Sep. 2017.

[41] W. Huleihel and Y. Steinberg, "Multiple access channel with unreliable cribbing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 1491–1495.

[42] Y. Steinberg, "Channels with cooperation links that may be absent," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 1947–1951.

[43] M. Zamanighomi, M. J. Emadi, F. S. Chaharsooghi, and M. R. Aref, "Multiple access channel with correlated channel states and cooperating encoders," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Oct. 2011, pp. 628–632.

[44] E. Amir and Y. Steinberg, "The multiple access channel with correlated sources and cribbing encoders," in *Proc. Int. Zürich Seminar Inf. Commun. (IZS)*, 2012, pp. 8–10.

[45] A. Bracher and A. Lapidoth, "Feedback, cribbing, and causal state information on the multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7627–7654, Dec. 2014.

[46] S. I. Bross and A. Lapidoth, "The state-dependent multiple-access channel with states available at a cribbing encoder," in *Proc. IEEE 26th Conv. Electric. Electron. Engin. Isr.*, Nov. 2010, pp. 665–669.

[47] J. Shimonovich, A. Somekh-Baruch, and S. Shamai, "Cognitive cooperative communications on the multiple access channel," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Sep. 2013, pp. 1–5.

[48] R. Kolte, A. Ozgur, and H. Permuter, "State-dependent multiple-access channels with partially cribbing encoders," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 21–25.

[49] I. Savov, M. M. Wilde, and M. Vu, "Partial decode-forward for quantum relay channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, USA, Jul. 2012, pp. 731–735.

[50] D. Ding, H. Gharibyan, P. Hayden, and M. Walter, "A quantum multi-party packing lemma and the relay channel," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3500–3519, Jun. 2020.

[51] P. Hayden and C. King, "Correcting quantum channels by measuring the environment," *Quantum Inf. Comput.*, vol. 5, no. 2, pp. 156–160, 2005.

[52] J. A. Smolin, F. Verstraete, and A. Winter, "Entanglement of assistance and multipartite state distillation," *Phys. Rev. A, Gen. Phys.*, vol. 72, Nov. 2005, Art. no. 052317.

[53] A. Winter, "On environment-assisted capacities of quantum channels," *Markov Process Rel. Fields*, vol. 13, nos. 1–2, pp. 297–314, 2007.

[54] D. Sutter, "Approximate quantum Markov chains," in *Approximate Quantum Markov Chains*. Cham, Switzerland: Springer, 2018, pp. 75–100.

[55] M. M. Wilde, *Quantum Information Theory*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2017.

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

[57] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[58] A. S. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 269–273, Jan. 1998.

[59] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A, Gen. Phys.*, vol. 56, p. 131, Jul. 1997.

[60] H. G. Eggleston, "Convexity," *J. London Math. Soc.*, vol. 1, no. 1, pp. 183–186, 1966.

[61] G. Keshet, Y. Steinberg, and N. Merhav, "Channel coding in the presence of side information," *Found. Trends Commun. Inf. Theory*, vol. 4, no. 6, pp. 445–586, Jan. 2007.

[62] S. Guha, J. H. Shapiro, and B. I. Erkmen, "Classical capacity of bosonic broadcast communication and a minimum output entropy conjecture," *Phys. Rev. A, Gen. Phys.*, vol. 76, no. 3, Sep. 2007, Art. no. 032303.

[63] C. Weedbrook *et al.*, "Gaussian quantum information," *Rev. Modern Phys.*, vol. 84, no. 2, pp. 621–669, May 2012.

[64] A. S. Holevo and R. F. Werner, "Evaluating capacities of bosonic Gaussian channels," *Phys. Rev. A, Gen. Phys.*, vol. 63, Feb. 2001, Art. no. 032312.

[65] J. Eisert and M. M. Wolf, "Gaussian quantum channels," in *Quantum Information with Continuous Variables of Atoms and Light*. Singapore: World Scientific, 2007, pp. 23–42.

[66] B. J. Yen and J. H. Shapiro, "Multiple-access bosonic communications," *Phys. Rev. A, Gen. Phys.*, vol. 72, no. 6, Dec. 2005, Art. no. 062312.

[67] U. Pereg and Y. Steinberg, "The arbitrarily varying channel with colored Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 3781–3817, Jun. 2021.

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

[69] A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2481–2485, Nov. 1999.

[70] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

[71] U. Pereg, "Communication over quantum channels with parameter estimation," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2020, pp. 1–6.

[72] T. Ogawa and H. Nagaoka, "Making good codes for classical-quantum channel coding via quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2261–2266, Jun. 2007.

Uzi Pereg (Member, IEEE) received the B.Sc. degree *(summa cum laude)* in electrical engineering from the Azrieli College of Engineering, Jerusalem, Israel, in 2011, and the M.Sc. and Ph.D. degrees from the Technion—Israel Institute of Technology, Haifa, Israel, in 2015 and 2019, respectively. He is currently a Post-Doctoral Researcher with the Institute for Communications Engineering, Technical University of Munich, and the Munich Center for Quantum Science and Technology (MCQST), Munich, Germany. In 2020, he joined the Theory Group of the German Federal Government (BMBF) project for the design and analysis of quantum communication and repeater systems. His research interests are in the areas of quantum communications, information theory, and coding theory. He was a recipient of the 2018 Pearl Award for outstanding research work in the field of communications, the 2018 KLA-Tencor Award for Excellent Conference Paper, the 2018–2019 Viterbi Fellowship, the 2020–2021 Israel CHE Fellowship for Quantum Science and Technology, and the 2022 Seed Funding Grant for Exceptional Junior Researchers of the Munich Center for Quantum Science and Technology.

Christian Deppe (Member, IEEE) received the Dipl.Math. and Dr.Math. degrees in mathematics from the Universität Bielefeld, Bielefeld, Germany, in 1996 and 1998, respectively. He was a Research and Teaching Assistant with the Fakultät für Mathematik, Universität Bielefeld, from 1998 to 2010. From 2011 to 2013, he was the Project Leader of the project Sicherheit und Robustheit des Quanten-Repeaters of the Federal Ministry of Education and a Research with the Fakultät für Mathematik, Universität Bielefeld. In 2014, he was supported by the DFG Project at the Institute of Theoretical Information Technology, Technische Universität München. In 2015, he had a temporary professorship with the Fakultät für Mathematik und Informatik, Friedrich-Schiller Universität Jena. He is currently the Project Leader of the project Abhörsichere Kommunikation über Quanten-Repeater of the Federal Ministry of Education and a Research at Fakultät für Mathematik, Universität Bielefeld. Since 2018, he has been with the Department of Communications Engineering, Technische Universität München.

Holger Boche (Fellow, IEEE) received the Dipl.-Ing. degree in electrical engineering, the Graduate degree in mathematics, and the Dr.-Ing. degree in electrical engineering from the Technische Universität Dresden, Dresden, Germany, in 1990, 1992, and 1994, respectively, the master's degree from the Friedrich-Schiller Universität Jena, Jena, Germany, in 1997, and the Dr.Rer.Nat. degree in pure mathematics from the Technische Universität Berlin, Berlin, Germany, in 1998.

In 1997, he joined the Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute (HHI), Berlin. From 2002 to 2010, he was a Full Professor of mobile communication networks with the Institute for Communications Systems, Technische Universität Berlin. In 2003, he became the Director of the Fraunhofer German-Sino Laboratory for Mobile Communications, Berlin. In 2004, he became the Director of the Fraunhofer Institute for Telecommunications, HHI. He was a Visiting Professor with ETH Zürich, Zürich, Switzerland, from 2004 to 2006 (Winter), and with KTH Royal Institute of Technology, Stockholm, Sweden, in 2005 (Summer). He has been a member and an Honorary Fellow of the TUM Institute for Advanced Study, Munich, Germany, since 2014. Since 2018, he has been a Founding Director of the Center for Quantum Engineering, Technische Universität München. Since 2021, he has been jointly leading the BMBF Research Hub 6G-life with Frank Fitzek. He is currently a Full Professor with the Institute of Theoretical Information Technology, Technische Universität München, Munich, Germany, which he joined in October 2010. Among his publications is the recent book, *Information Theoretic Security and Privacy of Information Systems* (Cambridge University Press, 2017).

Prof. Boche is a member of the IEEE Signal Processing Society SPCOM and SPTM Technical Committees. He was an Elected Member of the German Academy of Sciences (Leopoldina) in 2008 and to the Berlin Brandenburg Academy of Sciences and Humanities in 2009. He was a recipient of the Research Award "Technische Kommunikation" from the Alcatel SEL Foundation in October 2003, the "Innovation Award" from the Vodafone Foundation in June 2006, and the Gottfried Wilhelm Leibniz Prize from the Deutsche Forschungsgemeinschaft (German Research Foundation) in 2008. He was a co-recipient of the 2006 IEEE Signal Processing Society Best Paper Award and a recipient of the 2007 IEEE Signal Processing Society Best Paper Award. He was the General Chair of the Symposium on Information Theoretic Approaches to Security and Privacy at IEEE GlobalSIP 2016.