Classical Communication Over a Quantum Interference Channel

Omar Fawzi, Patrick Hayden, *Member, IEEE*, Ivan Savov, Pranab Sen, and Mark M. Wilde, *Member, IEEE*

***Abstract*—Calculating the capacity of interference channels is a notorious open problem in classical information theory. Such channels have two senders and two receivers, and each sender would like to communicate with a partner receiver. The capacity of such channels is known exactly in the settings of "very strong" and "strong" interference, while the Han–Kobayashi coding strategy gives the best known achievable rate region in the general case. Here, we introduce and study the quantum interference channel, a natural generalization of the interference channel to the setting of quantum information theory. We restrict ourselves for the most part to channels with two classical inputs and two quantum outputs in order to simplify the presentation of our results (though generalizations of our results to channels with quantum inputs are straightforward). We are able to determine the exact classical capacity of this channel in the settings of "very strong" and "strong" interference, by exploiting Winter's successive decoding strategy and a novel two-sender quantum simultaneous decoder, respectively. We provide a proof that a Han–Kobayashi strategy is achievable with Holevo information rates, up to a conjecture regarding the existence of a three-sender quantum simultaneous decoder. This conjecture holds for a special class of quantum multiple-access channels with average output states that commute, and we discuss some other variations of the conjecture that hold. Finally, we detail a connection between the quantum interference channel and prior work on the capacity of bipartite unitary gates.**

***Index Terms*—Classical communication, quantum interference channel, quantum Shannon theory, quantum simultaneous decoding, quantum successive decoding, unitary gate capacity.**

I. Introduction

CLASSICAL information theory came as a surprise to the communication engineers of the 1940s and 1950s [43], [58]. It was astonishing that two-terminal noisy communication channels generally have a nonzero capacity at which two parties can communicate error-free in the asymptotic limit of many channel uses, and furthermore, that the computation of this capacity is a straightforward convex optimization problem [53]—many consider the achievements of Shannon to be among the great scientific accomplishments of the last century. Soon after this accomplishment, Shannon laid the foundations for multiuser information theory, and he claimed that a three-terminal communication channel with two senders and one receiver also has a simple, elegant solution [54], [58]. Some time later, Liao [38] and Ahlswede [2] provided a formal proof of the capacity of this multiple access channel without any knowledge of Shannon's unpublished solution. The beauty of information theory in these two settings is that it offers elementary solutions to problems that, at the outset, seem to be extraordinarily difficult to solve.

The situation for more general communication scenarios in multiuser information theory is not as simple and elegant as it is for single-sender, single-receiver channels and multiple access channels [17]. For example, the capacity of the interference channel is one of the notorious open problems in classical information theory [37]. The interference channel refers to the setting in which a noisy communication channel connects two senders to two receivers, and each sender's goal is to communicate with a partner receiver. Each sender's transmission can interfere with the other's, and this is one reason (among many) that the problem is difficult to solve in the general case. This channel arises naturally in the context of data transmission over interfering wireless links or digital subscriber lines [37]. Shannon himself introduced the problem and attempted to solve it [54], but it is the later work of others that would provide ongoing improvements to the inner and outer bounds for the capacity of the interference channel [8], [20], [36], [46]–[49].

Carleial offered the first surprising result for the interference channel [8], by demonstrating that each sender can achieve the same rates of communication as if there is no interference at all if the interference from the other sender's transmission is "very strong." Carleial's solution is to have each receiver decode the other sender's message first and follow by decoding the partner sender's message, rather than each receiver simply treating the other sender's transmission as noise. Thus, Carleial's strategy demonstrates that we can achieve improved communication rates by taking advantage of interference rather than treating it as an obstacle. Sato then gave a full characterization of the capacity of the Gaussian interference channel in the setting of "strong" interference [49], by appealing to an earlier result of Ahlswede regarding the capacity of a compound multiple access channel [2]. Han and Kobayashi independently found Sato's result, and they built on these insights and applied them to the most general setting (not necessarily "strong" or "very strong" interference) by allowing for each decoder to partially decode

Manuscript received September 20, 2011; accepted February 09, 2012. Date of publication February 22, 2012; date of current version May 15, 2012. This paper was presented in part at the 49th Annual Allerton Conference, September 2011, and the 2012 Quantum Information Processing Conference, December 2011.

O. Fawzi, P. Hayden, I. Savov, and M. M. Wilde are with the School of Computer Science, McGill University, Montréal, QC H3A 2A7, Canada (e-mail: omar.fawzi@mail.mcgill.ca; patrick@cs.mcgill.ca; ivan.savov@gmail.com; mwilde@gmail.com).

P. Sen is with the School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai 400005, India (e-mail: pgdsen@tifr.res.in).

Communicated by A. Holevo, Associate Editor for Quantum Information Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2012.2188620

the other sender's message and use this information to better decode the message intended for them [20]. The resulting achievable rate region is known as the Han–Kobayashi rate region, and it is currently the best known inner bound on the capacity of the interference channel.[1]

The model of the interference channel as stated in the aforementioned works is an important practical model for data transmission in a noisy two-input, two-output network, but it ignores a fundamental aspect of the physical systems employed to transmit these data. At bottom, these physical systems operate according to the laws of quantum mechanics [41], and ultimately, at some level, these laws govern how noise can affect these systems. Now, for many systems (macroscopic ones in particular), these laws are not necessary to describe the dynamics of encoding, transmission, and decoding, and one could argue in this case that there is not any benefit of recasting information theory as a *quantum* information theory because it would only add a layer of complexity to the theory. However, there are examples of natural physical systems, such as fiber-optic cables or free space channels, for which quantum information theory offers a boost in capacity if the coding scheme makes clever use of quantum mechanics [18]. Thus, it is important to determine the information capacities of quantum channels, given that the physical carriers of information are quantum and quantum effects often give a boost in capacity. In [18], it is shown that a receiver making use of a collective measurement operating on all of the channel outputs has an improvement in performance over a receiver decoding with single-channel-output measurements. Additionally, there are existential arguments for examples of channels in which entanglement at the encoder can improve performance, leading to superadditive effects that simply cannot occur in classical information theory [25].

The quantum-mechanical approach to information theory has shed a new light on the very nature of information, and researchers have made much progress on this front in the past few decades [41]. Perhaps the most fundamental problem in quantum information theory is the task of transmitting bits over a quantum channel. Holevo and Schumacher–Westmoreland (HSW) offered independent proofs that the Holevo information, one generalization of Shannon's mutual information, is an achievable rate for classical data transmission over a quantum channel [30], [50]. Many researchers thought for some time that the Holevo information of a quantum channel would be equal to its classical capacity, but recent work has demonstrated that the answer to the most fundamental question of the classical capacity of a quantum channel remains wide open in the general case [25], [29].

Soon after the HSW result, quantum information theorists began exploring other avenues, one of which is multiuser quantum information theory. Winter proved that the capacity region of a quantum multiple access channel is a natural generalization of the classical solution, in which we can replace Shannon information rates with Holevo information rates [62]. It was not obvious at the outset that this solution would be possible—after all, any retrieval of data from a quantum system inevitably disturbs the state of the system, suggesting that successive decoding strategies employed in the classical case might not work for quantum systems [11]. But Winter overcame this obstacle by realizing that a so-called "gentle" or "tender" measurement, a measurement with an outcome that succeeds with high probability, effectively causes no disturbance to the state in the asymptotic limit of many channel uses. Later, Yard *et al.* considered various capacities of a quantum broadcast channel [63], and they found results that are natural generalizations of results from classical multiuser information theory [5], [17]. In parallel with these developments, researchers have considered many generalizations of the above settings, depending on the form of the transmitted information [12], [28], [31], [32], [39], [55], whether assisting resources are available [4], [15], [33], [56] or whether the sender and receiver would like to tradeoff different resources against each other [13], [14], [34].

II. Summary of Results

In this paper, we introduce the quantum interference channel, a natural generalization of the interference channel to the quantum domain. We at first restrict our discussion to a particular *ccqq* quantum interference channel, which has two classical inputs and two quantum outputs. This restriction simplifies the presentation, and a straightforward extension of our results leads to results for a general quantum interference channel with quantum inputs and quantum outputs. We summarize our main results as follows.

1) Our first contribution is an exact characterization of the capacity region of a *ccqq* quantum interference channel with "very strong" interference—the result here is a straightforward generalization of Carleial's result from [8].
2) Our second contribution is a different exact characterization of the capacity of a *ccqq* channel that exhibits "strong interference." This result employs a novel quantum simultaneous decoder for quantum multiple access channels with two classical inputs and one quantum output.
3) Our next contribution is a quantization of the Han–Kobayashi achievable rate region, up to a conjecture regarding the existence of a quantum simultaneous decoder for quantum multiple access channels with three classical inputs and one quantum output. We prove that a three-sender quantum simultaneous decoder exists in the special case where the induced channel to each receiver has average output states that commute, but we have not been able to prove the existence of such a decoder in the general case (neither is it clear how to leverage the proof of the two-sender simultaneous decoder). We prove that a certain rate region described in terms of min-entropies [44], [45] is achievable for the general noncommuting case, and our suspicion is that a proof for the most general case should exist and will bear similarities to these proofs. The existence of such a simultaneous decoder immediately implies that the senders and receivers can achieve the rates on the Han–Kobayashi inner bound. This conjecture is also closely related to the "multiparty typicality" conjecture formulated in [16].

[1] Chong *et al.* subsequently proposed another achievable rate region originally thought to improve the Han–Kobayashi rate region [9], but later work demonstrated that the Chong–Motani–Garg achievable rate region is equivalent to the Han–Kobayashi region [10], [35].

4) We also describe an achievable rate region for the quantum interference channel based on a successive decoding and rate-splitting strategy [52].
5) We supply an outer bound on the capacity of the quantum interference channel, similar to Sato's outer bound from [47].
6) Finally, we discuss the connection between prior work on the capacity of unitary gates [3], [22]–[24] and the capacity of the quantum interference channel. The quantum interference channel that we consider in this last contribution is an isometry, in which the two inputs and two outputs are quantum and the channel acts as a noiseless evolution from the senders to the receivers.

We structure this paper as follows. We first introduce the notation used in the rest of this paper. We then detail the general information processing task that two senders and two receivers are trying to accomplish using the quantum interference channel. Section V discusses the connection between the multiple access channel and the interference channel, and we prove the existence of a quantum simultaneous decoder for the multiple access channel with two classical inputs and one quantum output. This section also states a conjecture regarding the existence of a quantum simultaneous decoder with three classical inputs and one quantum output, and we prove that it exists for a special case. We also discuss an achievable rate region in terms of min-entropies, and we remark briefly on many avenues that we pursued in an attempt to prove this conjecture. Section VI presents our results regarding the quantum interference channel. We first determine the capacity of the quantum interference channel if the channel has "very strong" interference and follow with the capacity when the channel exhibits "strong" interference. We next show how to achieve the Han–Kobayashi inner bound, by exploiting the conjecture regarding the existence of a three-sender quantum simultaneous decoder. We then present a set of achievable rates obtained using successive decoding and rate-splitting. This section ends with an outer bound on the capacity of the quantum interference channel. Section VII presents our final contribution regarding the connection to unitary gate capacities, and the conclusion summarizes our findings and states open lines of pursuit for the quantum interference channel.

III. Notation

We denote quantum systems as A, B, and C and their corresponding Hilbert spaces as $\mathcal{H}^A$, $\mathcal{H}^B$, and $\mathcal{H}^C$ with respective dimensions d_A, d_B, and d_C. We denote pure states of the system A with a *ket* $|\phi\rangle^A$ and the corresponding density operator as $\phi^A = |\phi\rangle\langle\phi|^A$. All kets that are quantum states have unit norm, and all density operators are positive semidefinite with unit trace. We model our lack of access to a quantum system with the partial trace operation. That is, given a two-qubit state ρ^{AB} shared between Alice and Bob, we can describe Alice's state with the reduced density operator:

$$\rho^A = \mathrm{Tr}_B\left\{\rho^{AB}\right\}$$

where Tr_B denotes a partial trace over Bob's system. Let

$$H(A)_\rho \equiv -\mathrm{Tr}\left\{\rho^A \log \rho^A\right\}$$

be the von Neumann entropy of the state ρ^A. For a state σ^{ABC}, we define the quantum conditional entropy

$$H(A|B)_\sigma \equiv H(AB)_\sigma - H(B)_\sigma$$

the quantum mutual information

$$I(A;B)_\sigma \equiv H(A)_\sigma + H(B)_\sigma - H(AB)_\sigma$$

and the conditional quantum mutual information

$$I(A;B|C)_\sigma \equiv H(A|C)_\sigma + H(B|C)_\sigma - H(AB|C)_\sigma.$$

Quantum operations are completely positive trace-preserving maps $\mathcal{N}^{A'\to B}$, which accept input states in A' and output states in B. In order to describe the "distance" between two quantum states, we use the notion of *trace distance*. The trace distance between states σ and ρ is

$$\|\sigma - \rho\|_1 = \mathrm{Tr}\,|\sigma - \rho|$$

where $|X| = \sqrt{X^\dagger X}$. Two states that are similar have trace distance close to zero, whereas states that are perfectly distinguishable have trace distance equal to two. Throughout this paper, logarithms and exponents are taken base two unless otherwise specified. The Appendix reviews several important properties of typical sequences and typical subspaces.

IV. Information Processing Task

We first discuss the information processing task that two senders and two receivers are trying to accomplish with the quantum interference channel. We assume that they have access to many independent uses of a particular type of channel with two classical inputs and two quantum outputs. A *ccqq* quantum interference channel is the following map:

$$x, y \to \rho^{B_1 B_2}_{x,y} \tag{1}$$

where the inputs x and y produce a density operator $\rho^{B_1 B_2}_{x,y}$ that exists on quantum systems B_1 and B_2. Receiver 1 has access to system B_1, and Receiver 2 has access to system B_2. An $(n, R_1 - \delta, R_2 - \delta, \epsilon)$ quantum interference channel code consists of three steps: encoding, transmission, and decoding.

Encoding: Sender 1 chooses a message l from a message set $\mathcal{L} = \{1, 2, \ldots, |\mathcal{L}|\}$, where $|\mathcal{L}| = 2^{n(R_1-\delta)}$, and Sender 2 similarly chooses a message m from a message set $\mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}$, where $|\mathcal{M}| = 2^{n(R_2-\delta)}$, where δ is some arbitrarily small positive number. Senders 1 and 2 then encode their messages as codewords of the following form:

$$\begin{aligned} x^n(l) &\equiv x_1(l)\, x_2(l)\, \cdots\, x_n(l) \\ y^n(m) &\equiv y_1(m)\, y_2(m)\, \cdots\, y_n(m). \end{aligned}$$

Transmission: They both input each letter of their codewords to a single use of the channel in (1), leading to an n-fold tensor product state of the following form at the output:

$$\rho^{B_1^n B_2^n}_{x^n(l),y^n(m)} \equiv \rho^{B_{1,1}B_{2,1}}_{x_1(l),y_1(m)} \otimes \rho^{B_{1,2}B_{2,2}}_{x_2(l),y_2(m)} \otimes \cdots \otimes \rho^{B_{1,n}B_{2,n}}_{x_n(l),y_n(m)}.$$

Receiver 1 has access to systems $B_{1,i}$ for all $i \in \{1, \ldots, n\}$, and Receiver 2 has access to systems $B_{2,i}$.

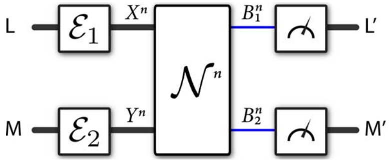

Fig. 1. Information processing task for the quantum interference channel. Let $\mathcal{N}$ represent the quantum interference channel with two classical inputs X and Y and two quantum outputs B_1 and B_2. Sender 1 selects a message l to transmit (modeled by a random variable L), and Sender 2 selects a message m to transmit (modeled by M). Each sender encodes their message as a codeword and transmits the codeword over many independent uses of a quantum interference channel. The receivers each receive the quantum outputs of the channel and perform a measurement to determine the message that their partner sender transmitted.

Decoding: Receiver 1 performs a measurement on his systems in order to determine the message of Sender 1, and Receiver 2 similarly performs a measurement to obtain Sender 2's message. More specifically, Receiver 1 performs a positive operator-valued measure (POVM) $\{\Lambda_l\}_{l\in\{1,\ldots,|\mathcal{L}|\}}$ where Λ_l is a positive operator for all l and $\sum_l \Lambda_l = I$, and Receiver 2 performs a POVM $\{\Gamma_m\}_{m\in\{1,\ldots,|\mathcal{M}|\}}$ with similar conditions holding for the operators in this set. Fig. 1 depicts all of these steps.

The probability of the receivers correctly decoding a particular message pair (l, m) is as follows:

$$\Pr\{L' = l,\ M' = m \mid L = l,\ M = m\} = \\ \mathrm{Tr}\left\{(\Lambda_l \otimes \Gamma_m)\, \rho_{x^n(l),y^n(m)}^{B_1^n B_2^n}\right\}$$

and so the probability of incorrectly decoding that message pair is

$$\begin{aligned} p_e(l,m) &\equiv \Pr\{(L',M') \neq (l,m) \mid L = l,\ M = m\} \\ &= \mathrm{Tr}\left\{(I - \Lambda_l \otimes \Gamma_m)\, \rho_{x^n(l),y^n(m)}^{B_1^n B_2^n}\right\} \end{aligned}$$

where L and M indicate random variables corresponding to the senders' choice of messages and the primed random variables correspond to the classical outputs of the receivers' measurements. The quantum interference channel code is ϵ-good if the average probability of error $\overline{p}_e$ is bounded from above by ϵ:

$$\begin{aligned} \overline{p}_e &\equiv \frac{1}{|\mathcal{L}||\mathcal{M}|} \sum_{l,m} p_e(l,m) \\ &= \frac{1}{|\mathcal{L}||\mathcal{M}|} \sum_{l,m} \mathrm{Tr}\left\{(I - \Lambda_l \otimes \Gamma_m)\, \rho_{x^n(l),y^n(m)}^{B_1^n B_2^n}\right\} \leq \epsilon. \end{aligned}$$

A rate pair (R_1, R_2) is *achievable* if there exists an $(n, R_1 - \delta, R_2 - \delta, \epsilon)$ quantum interference channel code for all $\delta, \epsilon > 0$ and sufficiently large n. The *capacity region* of the quantum interference channel is the closure of the set of all achievable rates.

V. Classical Communication Over the Quantum Multiple Access Channel

There is a strong connection between the multiple access channel and the interference channel. In fact, inner bounds for the capacity of an interference channel can be obtained by requiring the two receivers to decode both messages. Such a strategy naturally defines two multiple access channels that share the same senders [20], [17].[2] It is thus important to understand two different coding approaches for obtaining the capacity of the multiple access channel.

A. Successive Decoding

A first approach to achieve the capacity of the multiple access channel is to exploit a successive decoding strategy [11], [17], where the receiver first decodes the message of one sender while treating the other sender's transmission as noise. The receiver then decodes the message of the other sender by exploiting the decoded information as side information. This strategy achieves one "corner point" of the capacity region, and a symmetric strategy, where the receiver decodes in the opposite order, achieves the other corner point. They can achieve any rate pair between these two corner points with a time-sharing strategy, in which they exploit successive decoding in one order for a fraction of the channel uses and they exploit successive decoding in the opposite order for the remaining fraction of the channel uses. They can achieve the other boundary points and the interior of the capacity region by resource wasting.

Winter exploited this approach for the quantum multiple access channel [62], essentially by using a random coding argument and by showing that a measurement to determine the first sender's message causes a negligible disturbance of the channel output state. Hsieh *et al.* followed up on this result by showing how to perform entanglement-assisted classical communication over a quantum multiple access channel [33].

Theorem 1 (Successive Decoding [62]):
Let $x, y \rightarrow \rho_{x,y}$ be a ccq channel from two senders to a single receiver. Let $p_X(x)$ and $p_Y(y)$ be respective input distributions that each sender uses to create random codebooks of the form $\{X^n(l)\}_{l\in[1,\ldots,L]}$ and $\{Y^n(m)\}_{m\in[1,\ldots,M]}$. Suppose that the rates $R_1 = \frac{1}{n}\log_2(L) + \delta$ and $R_2 = \frac{1}{n}\log_2(M) + \delta$ (where $\delta > 0$) satisfy

$$\begin{aligned} R_1 &\leq I(X;B)_\rho \\ R_2 &\leq I(Y;B|X)_\rho \end{aligned}$$

where the Holevo information quantities are with respect to a classical-quantum state of the form

$$\rho^{XYB} \equiv \sum_{x,y} p_X(x) p_Y(y) \left|x\right\rangle \left\langle x\right|^X \otimes \left|y\right\rangle \left\langle y\right|^Y \otimes \rho_{x,y}^B. \quad (2)$$

Then, there exist two POVMs $\{\Lambda_l\}$ and $\left\{\Gamma_m^{(l)}\right\}$ acting in successive order such that the expectation of the average probability of correct detection is arbitrarily close to one:

$$\mathbb{E}\left\{\frac{1}{LM}\sum_{l,m}\mathrm{Tr}\left\{\sqrt{\Gamma_m^{(l)}}\sqrt{\Lambda_l}\rho_{X^n(l)Y^n(m)}\sqrt{\Lambda_l}\sqrt{\Gamma_m^{(l)}}\right\}\right\} \\ \geq 1 - \epsilon$$

where the expectation is with respect to X^n and Y^n.

[2] The setting in which both receivers decode both messages of the two senders is the same as the setting for the compound multiple access channel [2].

B. Quantum Simultaneous Decoding

Another approach to achieve the capacity of the multiple access channel is for the receiver to use a simultaneous decoder (sometimes referred to as a jointly typical decoder in the independent identically distributed setting), which decodes the messages of all senders at the same time rather than in succession [11], [17]. On one hand, simultaneous decoding is more complex than successive decoding because it considers all tuples of messages, but on the other hand, it is more powerful than a successive decoding strategy because it can decode at any rates provided that the rates are in the capacity region (also there is no need for time-sharing).

With such a strategy and for two senders, there are four different types of errors that can occur—one of these we can bound with a standard typicality argument and the other three correspond to the bounds on the capacity region of the channel. This strategy is our approach given below, and we can prove that a quantum simultaneous decoder exists for multiple access channels with two classical inputs and one quantum output. Though, for a three-sender quantum multiple access channel, we are only able to prove that a quantum simultaneous decoder exists in the special case where the averaged output states commute. Thus, we leave the general case stated as a conjecture.

1) Two-Sender Quantum Simultaneous Decoding: This section contains the proof of the two-sender quantum simultaneous decoder. We should mention that Sen arrived at this result with a different technique [51].

Theorem 2 (Two-Sender Quantum Simultaneous Decoding): Let $x, y \rightarrow \rho_{x,y}$ be a ccq channel from two senders to a single receiver. Let $p_X(x)$ and $p_Y(y)$ be respective input distributions that each sender uses to create random codebooks of the form $\{X^n(l)\}_{l\in[1,\ldots,L]}$ and $\{Y^n(m)\}_{m\in[1,\ldots,M]}$. Suppose that the rates $R_1 = \frac{1}{n}\log_2(L)+\delta$ and $R_2 = \frac{1}{n}\log_2(M)+\delta$ (where $\delta > 0$) satisfy the following inequalities:

$$R_1 \leq I(X;B|Y)_\rho \tag{3}$$

$$R_2 \leq I(Y;B|X)_\rho \tag{4}$$

$$R_1+R_2 \leq I(XY;B)_\rho \tag{5}$$

where the entropies are with respect to a state of the form in (2). Then, there exists a simultaneous decoding POVM $\{\Lambda_{l,m}\}$ such that the expectation of the average probability of error is bounded from above by ϵ for all $\epsilon > 0$ and sufficiently large n.

Proof: Suppose that the channel is a ccq channel of the form $x, y \rightarrow \rho_{x,y}$ and that the two senders have independent distributions $p_X(x)$ and $p_Y(y)$. These distributions induce the following averaged output states:

$$\rho_x \equiv \sum_y p_Y(y)\rho_{x,y} \tag{6}$$

$$\rho_y \equiv \sum_x p_X(x)\rho_{x,y} \tag{7}$$

$$\rho \equiv \sum_{x,y} p_X(x)p_Y(y)\rho_{x,y}. \tag{8}$$

Codeword Selection: Senders 1 and 2 choose codewords $\{X^n(l)\}_{l\in\{1,\ldots,L\}}$ and $\{Y^n(m)\}_{m\in\{1,\ldots,M\}}$ independently and randomly according to the product distributions $p_{X^n}(x^n)$ and $p_{Y^n}(y^n)$.

POVM Construction: Let $\Pi^n_{\rho,\delta}$ be the typical projector for the tensor power state $\rho^{\otimes n}$ defined by (8). Let $\Pi^n_{\rho_{y^n},\delta}$ be the conditionally typical projector for the tensor product state ρ_{y^n} defined by (7) for n uses of the channel. Let $\Pi^n_{\rho_{x^n},\delta}$ be the conditionally typical projector for the tensor product state ρ_{x^n} defined by (6) for n uses of the channel. Let $\Pi^n_{\rho_{x^n,y^n},\delta}$ be the conditionally typical projector for the tensor product state ρ_{x^n,y^n} defined as the output of the n channels when codewords x^n and y^n are input. (We are using the "weak" definitions of these projectors as defined in the Appendix.) In what follows, we make the following abbreviations:

$$\begin{aligned}\Pi &\equiv \Pi^n_{\rho,\delta}\\ \Pi_{y^n} &\equiv \Pi^n_{\rho_{y^n},\delta}\\ \Pi_{x^n} &\equiv \Pi^n_{\rho_{x^n},\delta}\\ \Pi_{x^n,y^n} &\equiv \Pi^n_{\rho_{x^n,y^n},\delta}.\end{aligned}$$

The detection POVM $\{\Lambda_{l,m}\}$ has the following form:

$$\begin{aligned}\Lambda_{l,m} &\equiv \left(\sum_{l',m'}\Pi'_{l',m'}\right)^{-\frac{1}{2}}\Pi'_{l,m}\left(\sum_{l',m'}\Pi'_{l',m'}\right)^{-\frac{1}{2}}\\ \Pi'_{l,m} &\equiv \Pi\,\Pi_{X^n(l)}\,\Pi_{X^n(l),Y^n(m)}\,\Pi_{X^n(l)}\,\Pi.\end{aligned} \tag{9}$$

Observe that the operator $\Pi'_{l,m}$ is a positive operator and thus $\{\Lambda_{l,m}\}$ is a valid POVM.

Error Analysis: The average error probability of the code has the following form:

$$\overline{p}_e \equiv \frac{1}{LM}\sum_{l,m}\mathrm{Tr}\left\{(I-\Lambda_{l,m})\,\rho_{X^n(l),Y^n(m)}\right\}. \tag{10}$$

We instead analyze the expectation of the average error probability, where the expectation is with respect to the random choice of code:

$$\begin{aligned}&\mathbb{E}_{X^n,Y^n}\{\overline{p}_e\}\\ &\equiv \mathbb{E}_{X^n,Y^n}\left\{\frac{1}{LM}\sum_{l,m}\mathrm{Tr}\left\{(I-\Lambda_{l,m})\,\rho_{X^n(l),Y^n(m)}\right\}\right\}\\ &= \frac{1}{LM}\sum_{l,m}\mathbb{E}_{X^n,Y^n}\left\{\mathrm{Tr}\left\{(I-\Lambda_{l,m})\,\rho_{X^n(l),Y^n(m)}\right\}\right\}.\end{aligned}$$

Due to the symmetry of the code construction (the fact that the expectation $\mathbb{E}_{X^n,Y^n}\left\{\mathrm{Tr}\left\{(I-\Lambda_{l,m})\,\rho_{X^n(l),Y^n(m)}\right\}\right\}$ is independent of the particular message pair (l,m)), it suffices to analyze the expectation of the average error probability for the first message pair (1, 1):

$$\mathbb{E}_{X^n,Y^n}\{\overline{p}_e\} = \mathbb{E}_{X^n,Y^n}\left\{\mathrm{Tr}\left\{(I-\Lambda_{1,1})\,\rho_{X^n(1),Y^n(1)}\right\}\right\}.$$

We now begin our error analysis. In what follows, we abbreviate X^n as X and Y^n as Y in order to save space. We first bound the above error probability as

$$\begin{aligned}&\mathbb{E}_{XY}\{\overline{p}_e\}\\ &\leq \mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{(I-\Lambda_{1,1})\,\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}\end{aligned}$$

$$+\,\mathbb{E}_{XY}\left\{\left\|\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}-\rho_{X(1),Y(1)}\right\|_1\right\} \quad (11)$$
$$\leq \mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{(I-\Lambda_{1,1})\ \Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\} + 2\sqrt{\epsilon} \quad (12)$$

where the first inequality follows from the inequality

$$\mathrm{Tr}\left\{\Lambda\rho\right\} \leq \mathrm{Tr}\left\{\Lambda\sigma\right\} + \|\rho-\sigma\|_1 \quad (13)$$

which holds for all ρ, σ, and Λ such that $0 \leq \rho, \sigma, \Lambda \leq I$. The second inequality follows from the properties of weak conditionally typical subspaces and the Gentle Operator Lemma for ensembles, by taking n to be sufficiently large (a discussion of these properties is in the Appendix). The idea behind this first bound on the error probability is that we require the projector $\Pi_{Y(1)}$ in order to remove some of large eigenvalues of an averaged version of $\rho_{X(1),Y(1)}$, and this point in the proof seems to be the most opportune time to insert it.

The Hayashi–Nagaoka operator inequality applies to a positive operator T and an operator S where $0 \leq S \leq I$ [27], [26]:

$$I-(S+T)^{-\frac{1}{2}}\,S\,(S+T)^{-\frac{1}{2}} \leq 2\,(I-S)+4T.$$

Choosing

$$S = \Pi'_{1,1}$$
$$T = \sum_{(l,m)\neq(1,1)} \Pi'_{l,m}$$

we can apply the above operator inequality to bound the first term in (12) as

$$\mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{(I-\Lambda_{1,1})\ \Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}$$
$$\leq 2\,\mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\left(I-\Pi'_{1,1}\right)\ \Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}$$
$$+\,4\sum_{(l,m)\neq(1,1)} \mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi'_{l,m}\,\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}. \quad (14)$$

We first consider bounding the term in the second line above. Consider that

$$\mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi'_{1,1}\,\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}$$
$$= \mathbb{E}_{XY}\{\mathrm{Tr}\{\Pi\,\Pi_{X(1)}\,\Pi_{X(1),Y(1)}\,\Pi_{X(1)}\,\Pi\ \Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\}\}$$
$$\geq \mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi_{X(1),Y(1)}\,\rho_{X(1),Y(1)}\right\}\right\}$$
$$-\,\mathbb{E}_{XY}\left\{\left\|\Pi\,\rho_{X(1),Y(1)}\,\Pi-\rho_{X(1),Y(1)}\right\|_1\right\}$$
$$-\,\mathbb{E}_{XY}\left\{\left\|\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}-\rho_{X(1),Y(1)}\right\|_1\right\}$$
$$-\,\mathbb{E}_{XY}\left\{\left\|\Pi_{X(1)}\,\rho_{X(1),Y(1)}\,\Pi_{X(1)}-\rho_{X(1),Y(1)}\right\|_1\right\}$$
$$\geq 1-\epsilon-6\sqrt{\epsilon}. \quad (15)$$

The aforementioned inequalities follow by employing the Gentle Operator Lemma for ensembles, (13), and the following inequalities that follow from the discussion in the Appendix:

$$\mathbb{E}_{XY}\left\{\mathrm{Tr}\{\Pi_{X(1)}\,\rho_{X(1),Y(1)}\}\right\} \geq 1-\epsilon \quad (16)$$
$$\mathbb{E}_{XY}\left\{\mathrm{Tr}\{\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\}\right\} \geq 1-\epsilon \quad (17)$$
$$\mathbb{E}_{XY}\left\{\mathrm{Tr}\{\Pi\,\rho_{X(1),Y(1)}\}\right\} \geq 1-\epsilon. \quad (18)$$
$$\mathbb{E}_{XY}\left\{\mathrm{Tr}\{\Pi_{X(1),Y(1)}\,\rho_{X(1),Y(1)}\}\right\} \geq 1-\epsilon. \quad (19)$$

This bound then implies that

$$\mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\left(I-\Pi'_{1,1}\right)\ \Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\} \leq \epsilon+6\sqrt{\epsilon}. \quad (20)$$

The bound in (14) reduces to the following one after applying (20):

$$\mathbb{E}_{XY}\left\{\overline{p}_e\right\} \leq 2\left(\epsilon+6\sqrt{\epsilon}\right)$$
$$+4\sum_{(l,m)\neq(1,1)} \mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi'_{l,m}\,\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}.$$

We can expand the doubly indexed sum in the above expression as

$$\sum_{(l,m)\neq(1,1)} \mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi'_{l,m}\,\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\} =$$
$$\sum_{l\neq1}\mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi'_{l,1}\,\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}$$
$$+\sum_{m\neq1}\mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi'_{1,m}\,\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}$$
$$+\sum_{l\neq1,\ m\neq1} \mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi'_{l,m}\,\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}. \quad (21)$$

We begin by bounding the term in the second line above. Consider the following chain of inequalities:

$$\sum_{l\neq1}\mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi'_{l,1}\Pi_{Y(1)}\,\rho_{X(1),Y(1)}\,\Pi_{Y(1)}\right\}\right\}$$
$$=\sum_{l\neq1}\mathbb{E}_{Y}\left\{\mathrm{Tr}\left\{\mathbb{E}_{X}\left\{\Pi'_{l,1}\right\}\ \Pi_{Y(1)}\mathbb{E}_{X}\left\{\rho_{X(1),Y(1)}\right\}\Pi_{Y(1)}\right\}\right\}$$
$$=\sum_{l\neq1}\mathbb{E}_{Y}\left\{\mathrm{Tr}\left\{\mathbb{E}_{X}\left\{\Pi'_{l,1}\right\}\ \Pi_{Y(1)}\,\rho_{Y(1)}\,\Pi_{Y(1)}\right\}\right\}$$
$$\leq 2^{-n[H(B|Y)-\delta]}\sum_{l\neq1}\mathbb{E}_{Y}\left\{\mathrm{Tr}\left\{\mathbb{E}_{X}\left\{\Pi'_{l,1}\right\}\ \Pi_{Y(1)}\right\}\right\}$$
$$= 2^{-n[H(B|Y)-\delta]}\sum_{l\neq1}\mathbb{E}_{XY}\left\{\mathrm{Tr}\left\{\Pi'_{l,1}\,\Pi_{Y(1)}\right\}\right\}. \quad (22)$$

The first equality follows because $X(l)$ and $X(1)$ are independent—the senders choose the code randomly in such a way that this is true. The second equality follows because $\mathbb{E}_X\left\{\rho_{X(1),Y(1)}\right\} = \rho_{Y(1)}$. The first inequality follows by applying the following operator inequality for weak conditionally typical subspaces:

$$\Pi_{y^n}\ \rho_{y^n}\ \Pi_{y^n} \leq 2^{-n[H(B|Y)-\delta]}\ \Pi_{y^n}.$$

The last equality is from factoring out the expectation. We now focus on the expression inside the expectation

$$\mathrm{Tr}\left\{\Pi'_{l,1}\,\Pi_{Y(1)}\right\}$$
$$= \mathrm{Tr}\left\{\Pi\,\Pi_{X(l)}\,\Pi_{X(l),Y(1)}\,\Pi_{X(l)}\,\Pi\,\Pi_{Y(1)}\right\}$$

$$\begin{aligned} &= \text{Tr}\left\{\Pi_{X(l),Y(1)}\, \Pi_{X(l)}\, \Pi\, \Pi_{Y(1)}\, \Pi\, \Pi_{X(l)}\right\} \\ &\le \text{Tr}\left\{\Pi_{X(l),Y(1)}\right\} \\ &\le 2^{n[H(B|XY)+\delta]}. \end{aligned}$$

The first equality is from substitution. The second equality is from cyclicity of trace. The first inequality is from

$$\Pi_{x^n}\, \Pi\, \Pi_{y^n}\, \Pi\, \Pi_{x^n} \le \Pi_{x^n}\, \Pi\, \Pi_{x^n} \le \Pi_{x^n} \le I.$$

The final inequality follows from the bound on the rank of the weak conditionally typical projector (see the Appendix).

Substituting back into (22), we have

$$\begin{aligned} &\sum_{l\neq 1} \mathbb{E}_{XY}\left\{\text{Tr}\left\{\Pi'_{l,1}\Pi_{Y(1)}\, \rho_{X(1),Y(1)}\, \Pi_{Y(1)}\right\}\right\} \\ &\quad\le 2^{-n[H(B|Y)-\delta]} \sum_{l\neq 1} 2^{n[H(B|XY)+\delta]} \\ &\quad\le 2^{-n[H(B|Y)-\delta]}\, 2^{n[H(B|XY)+\delta]}\, L \\ &\quad= 2^{-n[I(X;B|Y)-2\delta]}\, L. \end{aligned}$$

We employ a different argument to bound the term in the third line of (21). Consider the following chain of inequalities:

$$\begin{aligned} &\sum_{m\neq 1} \mathbb{E}_{XY}\left\{\text{Tr}\left\{\Pi'_{1,m}\, \Pi_{Y(1)}\, \rho_{X(1),Y(1)}\, \Pi_{Y(1)}\right\}\right\} \\ &\quad= \sum_{m\neq 1} \mathbb{E}_{X}\{\text{Tr}\{\mathbb{E}_{Y}\{\Pi'_{1,m}\}\mathbb{E}_{Y}\{\Pi_{Y(1)}\, \rho_{X(1),Y(1)}\, \Pi_{Y(1)}\}\}\}. \end{aligned} \tag{23}$$

This equality follows from the fact that $Y(m)$ and $Y(1)$ are independent. We now focus on bounding the operator $\mathbb{E}_{Y}\left\{\Pi'_{1,m}\right\}$ inside the trace:

$$\begin{aligned} &\mathbb{E}_{Y}\left\{\Pi'_{1,m}\right\} \\ &\quad= \mathbb{E}_{Y}\left\{\Pi\, \Pi_{X(1)}\, \Pi_{X(1),Y(m)}\, \Pi_{X(1)}\, \Pi\right\} \\ &\quad\le 2^{n[H(B|XY)+\delta]}\, \mathbb{E}_{Y}\left\{\Pi\, \Pi_{X(1)}\, \rho_{X(1),Y(m)}\, \Pi_{X(1)}\, \Pi\right\} \\ &\quad= 2^{n[H(B|XY)+\delta]}\, \Pi\, \Pi_{X(1)}\, \mathbb{E}_{Y}\left\{\rho_{X(1),Y(m)}\right\}\, \Pi_{X(1)}\, \Pi \\ &\quad= 2^{n[H(B|XY)+\delta]}\, \Pi\, \Pi_{X(1)}\, \rho_{X(1)}\, \Pi_{X(1)}\, \Pi \\ &\quad\le 2^{n[H(B|XY)+\delta]}\, 2^{-n[H(B|X)-\delta]}\, \Pi\, \Pi_{X(1)}\, \Pi \\ &\quad= 2^{-n[I(Y;B|X)-2\delta]}\, \Pi\, \Pi_{X(1)}\, \Pi \\ &\quad= 2^{-n[I(Y;B|X)-2\delta]}\, I. \end{aligned} \tag{24}$$

The first equality follows by substitution. The first inequality follows from the following operator inequality:

$$\begin{aligned} \Pi_{x^n,y^n} &\le 2^{n[H(B|XY)+\delta]}\, \Pi_{x^n,y^n}\, \rho_{x^n,y^n}\, \Pi_{x^n,y^n} \\ &= 2^{n[H(B|XY)+\delta]}\, \Pi_{x^n,y^n}\, \sqrt{\rho_{x^n,y^n}}\sqrt{\rho_{x^n,y^n}}\, \Pi_{x^n,y^n} \\ &= 2^{n[H(B|XY)+\delta]}\, \sqrt{\rho_{x^n,y^n}}\Pi_{x^n,y^n}\sqrt{\rho_{x^n,y^n}} \\ &\le 2^{n[H(B|XY)+\delta]}\, \rho_{x^n,y^n}. \end{aligned}$$

The second equality follows because Π and $\Pi_{X(1)}$ are constants with respect to the expectation over Y. The third equality follows because $\mathbb{E}_{Y}\left\{\rho_{X(1),Y(m)}\right\} = \rho_{X(1)}$, and the second inequality follows from the operator inequality

$$\Pi_{x^n}\, \rho_{x^n}\, \Pi_{x^n} \le 2^{-n[H(B|X)-\delta]}\Pi_{x^n}.$$

The final inequality follows from

$$\Pi\, \Pi_{x^n}\, \Pi \le \Pi \le I.$$

Substituting the operator inequality in (24) into (23), we have

$$\begin{aligned} &\sum_{m\neq 1} \mathbb{E}_{XY}\left\{\text{Tr}\left\{\Pi'_{1,m}\, \Pi_{Y(1)}\, \rho_{X(1),Y(1)}\, \Pi_{Y(1)}\right\}\right\} \\ &\quad\le 2^{-n[I(Y;B|X)-2\delta]} \sum_{m\neq 1} \mathbb{E}_{XY}\left\{\text{Tr}\left\{\Pi_{Y(1)}\rho_{X(1),Y(1)}\Pi_{Y(1)}\right\}\right\} \\ &\quad\le 2^{-n[I(Y;B|X)-2\delta]} \sum_{m\neq 1} \mathbb{E}_{XY}\left\{\text{Tr}\left\{\rho_{X(1),Y(1)}\right\}\right\} \\ &\quad\le 2^{-n[I(Y;B|X)-2\delta]}\, M. \end{aligned}$$

The second inequality follows because $\Pi_{y^n} \le I$.

Finally, we obtain a bound on the term in the last line of (21) with a slightly different argument:

$$\begin{aligned} &\sum_{l\neq 1, m\neq 1} \mathbb{E}_{XY}\left\{\text{Tr}\left\{\Pi'_{l,m}\, \Pi_{Y(1)}\, \rho_{X(1),Y(1)}\, \Pi_{Y(1)}\right\}\right\} \\ &= \sum_{\substack{l\neq 1,\\ m\neq 1}} \mathbb{E}_{Y}\left\{\text{Tr}\left\{\mathbb{E}_{X}\left\{\Pi'_{l,m}\right\}\Pi_{Y(1)}\mathbb{E}_{X}\left\{\rho_{X(1),Y(1)}\right\}\Pi_{Y(1)}\right\}\right\} \\ &= \sum_{l\neq 1,\ m\neq 1} \mathbb{E}_{Y}\left\{\text{Tr}\left\{\mathbb{E}_{X}\left\{\Pi'_{l,m}\right\}\, \Pi_{Y(1)}\, \rho_{Y(1)}\, \Pi_{Y(1)}\right\}\right\} \\ &\le \sum_{l\neq 1,\ m\neq 1} \mathbb{E}_{Y}\left\{\text{Tr}\left\{\mathbb{E}_{X}\left\{\Pi'_{l,m}\right\}\, \rho_{Y(1)}\right\}\right\} \\ &= \sum_{\substack{l\neq 1,\\ m\neq 1}} \mathbb{E}_{XY}\left\{\text{Tr}\left\{\Pi\, \Pi_{X(l)}\, \Pi_{X(l),Y(m)}\, \Pi_{X(l)}\, \Pi\, \rho_{Y(1)}\right\}\right\} \\ &= \sum_{\substack{l\neq 1,\\ m\neq 1}} \mathbb{E}_{X}\{\text{Tr}\{\Pi\Pi_{X(l)}\mathbb{E}_{Y}\{\Pi_{X(l)Y(m)}\}\Pi_{X(l)}\Pi\mathbb{E}_{Y}\{\rho_{Y(1)}\}\}\} \\ &= \sum_{\substack{l\neq 1,\\ m\neq 1}} \mathbb{E}_{X}\left\{\text{Tr}\left\{\Pi_{X(l)}\, \mathbb{E}_{Y}\left\{\Pi_{X(l),Y(m)}\right\}\, \Pi_{X(l)}\, \Pi\, \rho^{\otimes n}\, \Pi\right\}\right\}. \end{aligned}$$

The first equality follows from the independence of $X(l)$ and $X(1)$. The second equality follows because $\mathbb{E}_{X}\left\{\rho_{X(1),Y(1)}\right\} = \rho_{Y(1)}$. The first inequality follows from the fact that ρ_{y^n} and Π_{y^n} commute and thus $\Pi_{y^n}\, \rho_{y^n}\, \Pi_{y^n} = \sqrt{\rho_{y^n}}\, \Pi_{y^n}\, \sqrt{\rho_{y^n}} \le \rho_{y^n}$. The third equality follows from factoring out the expectation and substitution of the definition of $\Pi'_{l,m}$. The fourth equality follows from the independence of $Y(m)$ and $Y(1)$. The last equality follows because $\mathbb{E}_{Y}\left\{\rho_{Y(1)}\right\} = \rho^{\otimes n}$ and from cyclicity of trace. Continuing, we have

$$\begin{aligned} &\le 2^{-n[H(B)-\delta]}\times \\ &\quad \sum_{l\neq 1,\ m\neq 1} \text{Tr}\left\{\mathbb{E}_{X}\left\{\Pi_{X(l)}\, \mathbb{E}_{Y}\left\{\Pi_{X(l),Y(m)}\right\}\, \Pi_{X(l)}\right\}\, \Pi\right\} \\ &= 2^{-n[H(B)-\delta]}\times \\ &\quad \sum_{l\neq 1,\ m\neq 1} \mathbb{E}_{XY}\left\{\text{Tr}\left\{\Pi_{X(l),Y(m)}\, \Pi_{X(l)}\, \Pi\, \Pi_{X(l)}\right\}\right\} \\ &\le 2^{-n[H(B)-\delta]} \sum_{l\neq 1,\ m\neq 1} \mathbb{E}_{XY}\left\{\text{Tr}\left\{\Pi_{X(l),Y(m)}\right\}\right\} \\ &\le 2^{-n[H(B)-\delta]}\, 2^{n[H(B|XY)+\delta]}\, LM \\ &= 2^{-n[I(XY;B)-2\delta]}\, LM. \end{aligned} \tag{25}$$

The first inequality is from the following operator inequality:

$$\Pi\,\rho^{\otimes n}\,\Pi \leq 2^{-n[H(B)-\delta]}\Pi.$$

The second equality is from cyclicity of trace and factoring out the expectations. The second inequality is from the operator inequality

$$\Pi_{x^n}\,\Pi\,\Pi_{x^n} \leq \Pi_{x^n} \leq I.$$

The final inequality is from the bound on the rank of the weak conditionally typical projector.

Combining everything together, we get the following bound on the expectation of the average error probability:

$$\mathbb{E}_{XY}\left\{\overline{p}_e\right\} \leq 2\left(\epsilon + 7\sqrt{\epsilon}\right) + 4\,L\,2^{-n[I(X;B|Y)-2\delta]} + 4\,M\,2^{-n[I(Y;B|X)-2\delta]} + 4\,LM\,2^{-n[I(XY;B)-2\delta]}.$$

Thus, we can choose the message sizes to be as follows:

$$L = 2^{n[R_1-3\delta]}$$
$$M = 2^{n[R_2-3\delta]}$$

so that the expectation of the average error probability vanishes in the asymptotic limit whenever the rates R_1 and R_2 obey the following inequalities:

$$R_1 - \delta < I(X;B|Y)$$
$$R_2 - \delta < I(Y;B|X)$$
$$R_1 + R_2 - 4\delta < I(XY;B).$$

■

A casual glance at the above proof might lead one to believe it is just a straightforward extension of the "usual" proofs of the HSW theorem [12] [30], [33], [50], [60], but it differs from these and extends them non trivially in several regards. First, we choose the square-root POVM in (9) in a particular way—specifically, the layering of projectors is such that the projector of size $\approx 2^{nH(B|XY)}$ is surrounded by the projector of size $\approx 2^{H(B|X)}$, which itself is surrounded by the projector of size $\approx 2^{nH(B)}$. If one were to place the projector of size $\approx 2^{nH(B|Y)}$ somewhere in the square-root POVM, this leads to difficulties with noncommutative projectors (discussed in earlier versions of this paper on the arXiv). So, our second observation is to instead "smooth" the state by the projector of size $\approx 2^{nH(B|Y)}$ before applying the Hayashi–Nagaoka operator inequality. The above combination seems to be just the right trick for applying independence of the codewords after invoking the Hayashi–Nagaoka operator inequality. The final way in which our proof differs from earlier ones is that we analyze each of the four errors in a different way (these four types of errors occur after the application of the Hayashi–Nagaoka operator inequality). This asymmetry does not occur in the error analysis of the classical multiple access channel (see [17, pp. 4–15]), but for the moment, it seems to be necessary in the quantum case due to the general noncommutativity of typical projectors. Many of these observations are present in Sen's proof of the aforementioned theorem [51], but his proof introduces several new techniques (interestingly, he does not exploit the familiar square-root POVM or the Hayashi–Nagaoka operator inequality).

We obtain the following simple corollary of Theorem 2 by a technique called "coded time-sharing" [20], [17]. The main idea is to pick a sequence q^n according to a product distribution $p_{Q^n}(q^n)$ and then pick the codeword sequences x^n and y^n according to $p_{X^n|Q^n}(x^n|q^n)$ and $p_{Y^n|Q^n}(y^n|q^n)$, respectively (so that x^n and y^n are conditionally independent when given q^n). In the proof, all typical projectors are conditional on q^n, and we take the expectation over the time-sharing variable Q as well when bounding the expectation of the average error probability. Thus, we omit the proof of the below corollary.

Corollary 3:

Suppose that the rates R_1 and R_2 satisfy the following inequalities:

$$R_1 \leq I(X;B|YQ)_\rho \tag{26}$$
$$R_2 \leq I(Y;B|XQ)_\rho \tag{27}$$
$$R_1 + R_2 \leq I(XY;B|Q)_\rho \tag{28}$$

where the entropies are with respect to a state of the following form:

$$\rho^{QXYB} \equiv \sum_{x,y,q} p_Q(q)\,p_{X|Q}(x|q)\,p_{Y|Q}(y|q)\,|q\rangle\langle q|^Q \otimes |x\rangle\langle x|^X \otimes |y\rangle\langle y|^Y \otimes \rho^B_{x,y}.$$

Then, if the codebooks for Senders 1 and 2 are chosen as described previously, there exists a corresponding simultaneous decoding POVM $\{\Lambda_{l,m}\}$ such that the expectation of the average probability of error is bounded above by ϵ for all $\epsilon > 0$ and sufficiently large n.

2) Conjecture for Three-Sender Quantum Simultaneous Decoding: We now state our conjecture regarding the existence of a quantum simultaneous decoder for a quantum multiple access channel with three classical inputs. We state the conjecture for a three-sender quantum multiple access channel because this form is the one required for the proof of the Han–Kobayashi achievable rate region [20].

Conjecture 4 (Existence of a Three-Sender Quantum Simultaneous Decoder): Let $x,\ y,\ z \rightarrow \rho_{x,y,z}$ be a cccq quantum multiple access channel, where Sender 1 has access to the x input, Sender 2 has access to the y input, and Sender 3 has access to the z input. Let p_X, p_Y, and p_Z be distributions on the inputs. Define the following random code: let $\{X^n(k)\}_{k\in\{1,\ldots,K\}}$ be independent random variables distributed according to the product distribution p_{X^n} and similarly and independently let $\{Y^n(l)\}_{l\in\{1,\ldots,L\}}$ and $\{Z^n(m)\}_{m\in\{1,\ldots,M\}}$ be independent random variables distributed according to product distributions p_{Y^n} and p_{Z^n}. The rates of communication are $R_1 = \frac{1}{n}\log_2(K) + \delta$, $R_2 = \frac{1}{n}\log_2(L) + \delta$, and $R_3 = \frac{1}{n}\log_2(M) + \delta$, respectively, where $\delta > 0$. Suppose that these rates obey the following inequalities:

$$R_1 \leq I(X;B|YZ)_\rho$$
$$R_2 \leq I(Y;B|XZ)_\rho$$
$$R_3 \leq I(Z;B|XY)_\rho$$

$$\begin{aligned} R_1 + R_2 &\leq I\left(XY;B|Z\right)_\rho \\ R_1 + R_3 &\leq I\left(XZ;B|Y\right)_\rho \\ R_2 + R_3 &\leq I\left(YZ;B|X\right)_\rho \\ R_1 + R_2 + R_3 &\leq I\left(XYZ;B\right)_\rho \end{aligned}$$

where the Holevo information quantities are with respect to the following classical-quantum state:

$$\rho^{XYZB} \equiv \sum_{x,y,z} p_X(x) p_Y(y) p_Z(z) \left|x\right\rangle\left\langle x\right|^X \otimes \left|y\right\rangle\left\langle y\right|^Y \otimes \left|z\right\rangle\left\langle z\right|^Z \otimes \rho^B_{x,y,z}. \quad (29)$$

Then, there exists a decoding POVM $\{\Lambda_{l,m,k}\}_{l,m,k}$ such that the expectation of the average probability of error is bounded above by ϵ for all $\epsilon > 0$ and sufficiently large n:

$$\mathbb{E}\left\{ \frac{1}{KLM} \sum_{k,l,m} \mathrm{Tr}\left\{ \left(I - \Lambda_{k,l,m}\right) \rho_{X^n(k),Y^n(l),Z^n(m)} \right\} \right\} \leq \epsilon$$

where the expectation is with respect to X^n, Y^n, and Z^n.

The importance of this conjecture stems not only from the fact that a proof of it would be helpful in achieving a "quantized" version of the Han–Kobayashi achievable rate region, but also because such a proof might more broadly be helpful for "quantizing" other results in network classical information theory. Indeed, many coding theorems in network classical information theory exploit a simultaneous decoding approach (sometimes known as jointly typical decoding) [17]. Also, Dutil and Hayden have recently put forward a related conjecture known as the "multiparty typicality" conjecture [16], and it is likely that a proof of Conjecture 4 could aid in producing a proof of the multiparty typicality conjecture or vice versa.

3) Special Cases of the Conjecture: We now offer two theorems that are variations of the above conjecture that do hold for three-sender multiple access channels. The first is a special case in which we assume that certain averaged output states commute, and the second is one in which certain bounds contain min-entropies. It seems likely that an eventual proof of Conjecture 4, should one be found, will involve steps similar to those presented below, albeit with some crucial additional ideas.

4) Commuting Case: We prove a special case of Conjecture 4 in which we assume that certain averaged output states commute. First, let us define the following states:

$$\begin{aligned} \rho_{x,z} &\equiv \sum_y p_y(y)\ \rho_{x,y,z} \\ \rho_{y,z} &\equiv \sum_x p_x(x)\ \rho_{x,y,z} \\ \rho_{x,y} &\equiv \sum_z p_z(z)\ \rho_{x,y,z} \\ \rho_x &\equiv \sum_z p_z(z)\ \rho_{x,z} \\ \rho_y &\equiv \sum_x p_x(x)\ \rho_{x,y} \\ \rho_z &\equiv \sum_y p_y(y)\ \rho_{y,z} \\ \rho &\equiv \sum_{x,y,z} p_x(x)\ p_y(y)\ p_z(z)\ \rho_{x,y,z}. \end{aligned}$$

Theorem 5 (Averaged State Commuting Case): Consider the same setup as in Conjecture 4, with the additional assumption that certain averaged states commute: $[\rho_{x,z}, \rho_{y,z}] = [\rho_{x,y}, \rho_{y,z}] = [\rho_{x,y}, \rho_{x,z}] = 0$ for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$, and $z \in \mathcal{Z}$. Then, there exists a quantum simultaneous decoder in the sense described in Conjecture 4.

Proof: The proof exploits some ideas from Theorem 2. Thus, we merely describe the key points of the proof.

We randomly and independently choose codewords for the three senders according to the respective product distributions $p_{X^n}(x^n)$, $p_{Y^n}(y^n)$, and $p_{Z^n}(z^n)$. We define the detection POVM to be of the following form:

$$\Lambda_{k,l,m} \equiv \left(\sum_{k',l',m'} \Pi'_{k',l',m'} \right)^{-1/2} \Pi'_{k,l,m} \left(\sum_{k',l',m'} \Pi'_{k',l',m'} \right)^{-1/2} \quad (30)$$

where

$$\begin{aligned} \Pi'_{k,l,m} &\equiv M^\dagger_{x^n(k),y^n(l),z^n(m)} M_{x^n(k),y^n(l),z^n(m)}, \\ M_{x^n,y^n,z^n} &\equiv \Pi_{x^n,y^n,z^n}\ \Pi_{x^n,y^n}\ \Pi_{x^n,z^n}\ \Pi_{y^n,z^n} \times \\ &\quad \Pi_{x^n}\ \Pi_{y^n}\ \Pi_{z^n}\ \Pi \end{aligned}$$

and each of the above projectors are conditionally typical projectors defined with a similar shorthand from the proof of Theorem 2. Observe that all of the conditionally typical projectors Π_{x^n,y^n}, Π_{x^n,z^n}, Π_{y^n,z^n}, Π_{x^n}, Π_{y^n}, Π_{z^n}, and Π are mutually commuting from the assumption of the theorem. We analyze the expectation of the average error probability, and due to the symmetry of the code construction, it suffices to analyze this error probability for the first message triple (1, 1, 1):

$$\mathbb{E}_{X^n,Y^n,Z^n}\left\{ \mathrm{Tr}\left\{ \left(I - \Lambda_{1,1,1}\right)\ \rho_{X^n(1),Y^n(1),Z^n(1)} \right\} \right\}.$$

Our first move is to "unravel" the operator $I - \Lambda_{1,1,1}$ by means of the Hayashi–Nagaoka operator inequality, so that

$$I - \Lambda_{1,1,1} \leq 2\left(I - \Pi'_{1,1,1}\right) + \ 4 \sum_{(k,l,m) \neq (1,1,1)} \Pi'_{k,l,m}.$$

The first error with the operator $I - \Pi'_{1,1,1}$ under the trace can be bounded from above by some $f(\epsilon)$ where $\lim_{\epsilon \to 0} f(\epsilon) = 0$, by employing the trace inequality in (13) and the Gentle Operator Lemma for ensembles. We can expand the triply indexed sum for the second error into seven different types of errors. We delineate the different errors in the following table:

k	l	m
$*$	1	1
1	$*$	1
1	1	$*$
$*$	$*$	1
$*$	1	$*$
1	$*$	$*$
$*$	$*$	$*$

(31)

where $*$ denotes some message other than the first one (implying an incorrect decoding). Each of these we can bound by averaging over the state $\rho_{X^n(1),Y^n(1),Z^n(1)}$ and commuting the appropriate projector to be closest to the state. For example, consider the first error term. We have that $X^n(k)$ and $X^n(1)$ are independent. Bring the expectation over X^n inside of the trace and average over the state $\rho_{X^n(1),Y^n(1),Z^n(1)}$ to get $\rho_{Y^n(1),Z^n(1)}$. Commute $\Pi_{Y^n(1),Z^n(1)}$ to be closest to the state on both sides and exploit the operator inequality $\Pi_{y^n,z^n}\ \rho_{y^n,z^n}\ \Pi_{y^n,z^n} \leq 2^{-n[H(B|YZ)-\delta]}\ \Pi_{y^n,z^n}$. After a few steps, we end up with the bound $2^{-n[I(X;B|YZ)-2\delta]}\ K$. The other six bounds proceed in a similar fashion, demonstrating that Conjecture 4 holds true for this special case. ■

5) Min-Entropy Case: A simple modification of the proof of Theorem 2 allows us to achieve rates expressible in terms of min-entropies [45], [44], for arbitrary quantum channels. The min-entropy $H_{\min}(B)_\rho$ of a quantum state ρ^B is equal to the negative logarithm of its maximal eigenvalue:

$$H_{\min}(B)_\rho \equiv -\log\left(\inf_{\lambda\in\mathbb{R}}\{\lambda : \rho \leq \lambda I\}\right)$$

and the conditional min-entropy of a classical-quantum state $\rho^{XB} \equiv \sum_x p_X(x)\,|x\rangle\,\langle x|^X \otimes \rho_x^B$ with classical system X and quantum system B is as follows [44]:

$$H_{\min}(B|X)_\rho \equiv \inf_{x\in\mathcal{X}} H_{\min}(B)_{\rho_x}.$$

This definition of conditional min-entropy, where the conditioning system is classical, implies the following operator inequality:

$$\forall\, x \quad \rho_x^B \leq 2^{-H_{\min}(B|X)_\rho} I^B. \tag{32}$$

The following theorem gives an achievable rate region for a three-sender quantum simultaneous decoder. The entropy differences in (33)–(34) and (36)–(37) of the following theorem may not necessarily be positive for all states because the conditional quantum min-entropy can be less than the conditional von Neumann entropy. Nevertheless, there are some states for which these rates are positive, and Example 7 gives a channel for which the min-entropy rates are equivalent to the von Neumann entropy rates.

Theorem 6 (Min-Entropy Case):

Consider the same setup as in Conjecture 4. There exists a quantum simultaneous decoder in the sense described in Conjecture 4 that achieves the following rate region:

$$R_1 \leq H_{\min}(B|ZY) - H(B|XYZ) \tag{33}$$
$$R_2 \leq H_{\min}(B|XZ) - H(B|XYZ) \tag{34}$$
$$R_3 \leq I(Z;B|XY) \tag{35}$$
$$R_1 + R_2 \leq H_{\min}(B|Z) - H(B|XYZ) \tag{36}$$
$$R_2 + R_3 \leq H_{\min}(B|X) - H(B|XYZ) \tag{37}$$
$$R_1 + R_3 \leq I(XZ;B|Y) \tag{38}$$
$$R_1 + R_2 + R_3 \leq I(XYZ;B). \tag{39}$$

Other variations of the above achievable rate region are possible by permuting the variables X, Y, and Z in the above expressions.

Proof: The main idea for this proof is to exploit a decoding POVM of the form in (30), with $\Pi'_{k,l,m}$ chosen to be as follows:

$$\Pi'_{k,l,m} = \Pi\ \Pi_{y^n(l)}\ \Pi_{x^n(k),y^n(l)}\ \Pi_{x^n(k),y^n(l),z^n(m)} \times \\ \Pi_{x^n(k),y^n(l)}\ \Pi_{y^n(l)}\ \Pi. \tag{40}$$

We can bound the expectation of the average error probability again by exploiting the Hayashi–Nagaoka operator inequality. After doing so, the first error with the operator $I - \Pi'_{1,1,1}$ under the trace can be bounded from above by some $f(\epsilon)$ where $\lim_{\epsilon\to 0} f(\epsilon) = 0$, by employing the trace inequality in (13) and the Gentle Operator Lemma for ensembles. The second error again breaks into the seven errors of the form in (31). We discuss below how to handle each of these errors:

1) $X^n(k)$ and $X^n(1)$ are independent. Bring the expectation over X^n inside of the trace and average over the state $\rho_{X^n}(1)$, $Y^n(1)$, $Z^n(1)$ to get $\rho_{Y^n(1),Z^n(1)}$. The state $\rho_{Y^n(1),Z^n(1)}$ is bounded from above by $2^{-nH_{\min}(B|YZ)}$ and proceed to upper bound this error by $2^{-n[H_{\min}(B|ZY)-H(B|XYZ)]}\ K$.
2) $Y^n(l)$ and $Y^n(1)$ are independent. Bring the expectation over Y^n inside of the trace and average over the state $\rho_{X^n(1),Y^n(1),Z^n(1)}$ to get $\rho_{X^n(1),Z^n(1)}$. The state $\rho_{X^n(1),Z^n(1)}$ is bounded from above by $2^{-nH_{\min}(B|XZ)}$ and proceed to upper bound this error by $2^{-n[H_{\min}(B|XZ)-H(B|XYZ)]}\ L$.
3) $Z^n(m)$ and $Z^n(1)$ are independent. Exploit the operator inequality $\Pi_{x^n,y^n,z^n} \leq 2^{n[H(B|XYZ)+\delta]}\ \rho_{x^n,y^n,z^n}$, bring the expectation over Z^n inside of the trace and average over the state $\rho_{X^n(1),Y^n(1),Z^n(M)}$ to get $\rho_{X^n(1),Y^n(1)}$. Exploit the operator inequality $\Pi_{x^n,y^n}\ \ \rho_{x^n,y^n}\ \ \Pi_{x^n,y^n} \leq 2^{-n[H(B|XY)-\delta]}\ \Pi_{x^n,y^n}$. We can then upper bound this error by $2^{-n[I(Z;B|XY)-2\delta]}\ M$.
4) $X^n(k)$ and $X^n(1)$ are independent, and so are $Y^n(l)$ and $Y^n(1)$. Bring the expectations over X^n and Y^n inside of the trace and average over the state $\rho_{X^n(1),Y^n(1),Z^n(1)}$ to get $\rho_{Z^n(1)}$. The state $\rho_{Z^n(1)}$ is bounded from above by $2^{-nH_{\min}(B|Z)}$ and proceed to upper bound this error by $2^{-n[H_{\min}(B|Z)-H(B|XYZ)]}\ KL$.
5) $Y^n(l)$ and $Y^n(1)$ are independent, and so are $Z^n(m)$ and $Z^n(1)$. Bring the expectations over Y^n and Z^n inside of the trace and average over the state $\rho_{X^n(1),Y^n(1),Z^n(1)}$ to get $\rho_{X^n(1)}$. The state $\rho_{X^n(1)}$ is bounded from above by $2^{-nH_{\min}(B|X)}$ and proceed to upper bound this error by $2^{-n[H_{\min}(B|X)-H(B|XYZ)]}\ LM$.
6) $X^n(k)$ and $X^n(1)$ are independent, and so are $Z^n(m)$ and $Z^n(1)$. Exploit the operator inequality $\Pi_{x^n,y^n,z^n} \leq 2^{n[H(B|XYZ)+\delta]}\ \rho_{x^n,y^n,z^n}$, bring the expectation over Z^n inside of the trace and average over the state $\rho_{X^n(1),Y^n(1),Z^n(M)}$ to get $\rho_{X^n(1),Y^n(1)}$. Exploit the operator inequality $\Pi_{x^n,y^n}\ \rho_{x^n,y^n}\ \Pi_{x^n,y^n} \leq \rho_{x^n,y^n}$. Bring the expectation over X^n inside of the trace and average over the state $\rho_{X^n(1),Y^n(1)}$ to get $\rho_{Y^n(1)}$. Exploit the operator inequality $\Pi_{y^n}\ \ \rho_{y^n}\ \ \Pi_{y^n} \leq 2^{-n[H(B|Y)-\delta]}\ \Pi_{y^n}$. We can then upper bound this error by $2^{-n[I(XZ;B|Y)-2\delta]}\ KM$.

7) All variables are independent. Bring the expectations over X^n, Y^n, and Z^n inside of the trace and average over the state $\rho_{X^n(1),Y^n(1),Z^n(1)}$ to get $\rho^{\otimes n}$. Exploit the operator inequality $\Pi \; \rho^{\otimes n} \; \Pi \leq 2^{-n[H(B)-\delta]} \; \Pi$ and proceed to upper bound this error by $2^{-n[I(XYZ;B)-2\delta]} \; KLM$. ■

Example 7: We now provide an example of a cccq quantum multiple access channel for which a quantum simultaneous decoder can achieve its capacity region. We show that the min-entropy rates in (33)–(39) of Theorem 6 are equal to the von Neumann entropy rates from Conjecture 4. By Winter's results in [62] for a cccq multiple access channel, this implies that the min-entropy rate region is equivalent to the capacity region for this particular channel. Consider a channel that takes three bits x, y, and z as input and outputs one of the four "BB84" states:

$$000 \rightarrow |0\rangle, \quad 001 \rightarrow |+\rangle, \quad 010 \rightarrow |1\rangle, \quad 011 \rightarrow |-\rangle,$$
$$100 \rightarrow |1\rangle, \quad 101 \rightarrow |-\rangle, \quad 110 \rightarrow |0\rangle, \quad 111 \rightarrow |+\rangle.$$

A classical-quantum state on which we evaluate information quantities is

$$\rho^{XYZB} \equiv \sum_{x,y,z=0}^{1} p_X(x) p_Y(y) p_Z(z) \, |x\rangle \langle x|^X \otimes |y\rangle \langle y|^Y \otimes |z\rangle \langle z|^Z \otimes \psi^B_{x,y,z}$$

where $\psi^B_{x,y,z}$ is one of $|0\rangle$, $|1\rangle$, $|+\rangle$, or $|-\rangle$ depending on the choice of the bits x, y, and z. The conditional entropy $H(B|XYZ)_\rho$ vanishes for this state because the state is pure when conditioned on the classical registers X, Y, and Z. So it is only necessary to compare $H_{\min}(B|ZY)$ with $H(B|ZY)$, $H_{\min}(B|XZ)$ with $H(B|XZ)$, $H_{\min}(B|Z)$ with $H(B|Z)$, and $H_{\min}(B|X)$ with $H(B|X)$. We choose $p_X(x)$, $p_Y(y)$, and $p_Z(z)$ to be the uniform distribution. This gives the following reduced state on Z, Y, and B:

$$\begin{aligned}\frac{1}{4}|00\rangle\langle 00|^{ZY} &\otimes \frac{1}{2}\left(|0\rangle\langle 0|^B + |1\rangle\langle 1|^B\right) \\ &+ \frac{1}{4}|01\rangle\langle 01|^{ZY} \otimes \frac{1}{2}\left(|+\rangle\langle +|^B + |-\rangle\langle -|^B\right) \\ &+ \frac{1}{4}|10\rangle\langle 10|^{ZY} \otimes \frac{1}{2}\left(|0\rangle\langle 0|^B + |1\rangle\langle 1|^B\right) \\ &+ \frac{1}{4}|11\rangle\langle 11|^{ZY} \otimes \frac{1}{2}\left(|+\rangle\langle +|^B + |-\rangle\langle -|^B\right)\end{aligned}$$

for which it is straightforward to show that certain entropies take their maximal value of one bit: $H_{\min}(B|ZY) = H(B|ZY) = 1$ and $H_{\min}(B|Z) = H(B|Z) = 1$. We also have the following reduced state on X, Z, and B:

$$\begin{aligned}\frac{1}{4}|00\rangle\langle 00|^{XZ} &\otimes \frac{1}{2}\left(|0\rangle\langle 0|^B + |1\rangle\langle 1|^B\right) \\ &+ \frac{1}{4}|01\rangle\langle 01|^{XZ} \otimes \frac{1}{2}\left(|+\rangle\langle +|^B + |-\rangle\langle -|^B\right) \\ &+ \frac{1}{4}|10\rangle\langle 10|^{XZ} \otimes \frac{1}{2}\left(|0\rangle\langle 0|^B + |1\rangle\langle 1|^B\right) \\ &+ \frac{1}{4}|11\rangle\langle 11|^{XZ} \otimes \frac{1}{2}\left(|+\rangle\langle +|^B + |-\rangle\langle -|^B\right)\end{aligned}$$

for which the other entropies take their maximal value of one bit: $H_{\min}(B|XZ) = H(B|XZ) = 1$ and $H_{\min}(B|X) = H(B|X) = 1$. Furthermore, we can show that the conditional entropy $H(B|XY)_\rho$ takes it maximum value of $H_2(\cos^2(\pi/8))$ when $p_X(x)$ and $p_Y(y)$ are uniform (where $H_2(p) \equiv -p\log_2 p - (1-p)\log_2(1-p)$). Thus, the region achievable with min-entropies in (33)–(39) of Theorem 6 is equivalent to the capacity region for this channel:

$$\begin{aligned} R_1 &\leq 1 \\ R_2 &\leq 1 \\ R_3 &\leq H_2\left(\cos^2(\pi/8)\right) \\ R_1 + R_2 &\leq 1 \\ R_2 + R_3 &\leq 1 \\ R_1 + R_3 &\leq 1 \\ R_1 + R_2 + R_3 &\leq 1. \end{aligned}$$

6) Other Attempts at Proving Conjecture 4: We have attempted to prove Conjecture 4 in many different ways, and this section briefly summarizes these attempts. We again mention that our quantum simultaneous decoding conjecture seems related to the multiparty typicality conjecture from [16].

We have attempted to prove Conjecture 4 by exploiting the asymmetric hypothesis testing techniques from [40] and [59]. The problem with these approaches in the multiple access setting is that the POVM selected in the operational definitions of the quantum relative entropy is optimal for one type of error in (21), but it is not necessarily optimal for the other two types of errors. The hypothesis testing approaches from [6] and [7] also do not appear to be of much help for our goals here because they involve an infimum over the choice of the second state in the quantum relative entropy.

Another attempt is to improve the achievable rate region of Theorem 6, by replacing min-entropies with *smooth* min-entropies [44]. In fact, the smooth min-entropy is known to approach the von Neumann entropy in the case of a large number of independent and identically distributed random variables [44], [57]. To prove the conjecture, it would be sufficient to find a state $\tilde{\rho}^{X^nY^nZ^nB^n}$ that is close to $\rho^{X^nY^nZ^nB^n}$—which corresponds to n independent copies of the state ρ^{XYZB} in (29)—that simultaneously satisfies $H_{\min}(B|ZY)_{\tilde{\rho}} \geq H^{\epsilon}_{\min}(B|ZY)_\rho$, $H_{\min}(B|XZ)_{\tilde{\rho}} \geq H^{\epsilon}_{\min}(B|XZ)_\rho$, $H_{\min}(B|Z)_{\tilde{\rho}} \geq H^{\epsilon}_{\min}(B|Z)_\rho$, $H_{\min}(B|X)_{\tilde{\rho}} \geq H^{\epsilon}_{\min}(B|X)_\rho$. Here, $H^{\epsilon}_{\min}(B|X)_\rho$ refers to the ϵ-smooth min-entropy, which is the maximum of $H_{\min}(B|X)_{\rho'}$ over all states ρ' on XB that are ϵ-close to ρ; see [44] for a precise definition. In the proof of Theorem 6, we would replace the output of the channel ρ by $\tilde{\rho}$ before applying the Hayashi–Nagaoka operator inequality and the min-entropy terms would approach the von Neumann entropy terms we are looking for.

VI. Quantum Interference Channel

This section contains some of the main results of this paper, the inner and outer bounds on the capacity of a *ccqq* quantum interference channel of the following form:

$$x_1, x_2 \rightarrow \rho^{B_1B_2}_{x_1,x_2} \tag{41}$$

where Sender 1 has access to the classical x_1 input, Sender 2 has access to the classical x_2 input, Receiver 1 has access to the B_1 quantum system, and Receiver 2 has access to the B_2 quantum system. The first inner bound that we prove is similar to the result of Carleial for "very strong" interference. We then prove a quantum simultaneous decoding inner bound and give the capacity of the channel whenever it exhibits "strong" interference. The main inner bound is the Han–Kobayashi achievable rate region with Shannon information quantities replaced by Holevo information quantities, and this inner bound relies on Conjecture 4 for its proof. The outer bound in Section VI-B is similar to an outer bound in the classical case due to Sato [46].

A. Inner Bounds

As mentioned earlier, the interference channel naturally induces two multiple access channels with the same senders. Thus, one possible coding strategy for the interference channel is to build a codebook for each multiple access channel that is decodable for *both* receivers. In fact, most—if not all—known coding strategies for the interference channel are based on this idea. It is important to say here that we have to use the *same* codebook for both multiple access channels. For this reason, using the *existence* of good codes achieving all tuples in the capacity region is not sufficient.

1) Very Strong Interference: A setting for which we can determine the capacity of a ccqq interference channel is the setting of "very strong" interference (see [17, pp. 6–11]). The conditions for "very strong" interference are that the following information inequalities should hold for all distributions $p_{X_1}(x_1)$ and $p_{X_2}(x_2)$:

$$I(X_1;B_1|X_2)_\rho \leq I(X_1;B_2)_\rho \tag{42}$$

$$I(X_2;B_2|X_1)_\rho \leq I(X_2;B_1)_\rho \tag{43}$$

where $\rho^{X_1X_2B_1B_2}$ is a state of the following form:

$$\rho^{X_1X_2B_1B_2} \equiv \sum_{x_1,x_2} p_{X_1}(x_1)\, p_{X_2}(x_2)\, |x_1\rangle\langle x_1|^{X_1} \otimes |x_2\rangle\langle x_2|^{X_2} \otimes \rho_{x_1,x_2}^{B_1B_2}. \tag{44}$$

The information inequalities in (42) and (43) imply that the interference is so strong that it is possible for each receiver to decode the other sender's message before decoding the message intended for him. These conditions are a generalization of Carleial's conditions for a classical Gaussian interference channel [8].

Theorem 8 (Very Strong Interference):

Let a ccqq quantum interference channel as in (41) be given, and suppose that it has "very strong" interference as in (42) and (43). Then, the channel's capacity region is the union of all rates R_1 and R_2 satisfying the following inequalities:

$$R_1 \leq I(X_1;B_1|X_2Q)_\rho$$
$$R_2 \leq I(X_2;B_2|X_1Q)_\rho$$

where the union is over input distributions

$$p_Q(q)\, p_{X_1|Q}(x_1|q)\; p_{X_2|Q}(x_2|q).$$

Proof: Our proof technique is to apply Winter's successive decoder from Lemma 1, so that each receiver first decodes the message of the other sender, followed by decoding the message of the partner sender. More specifically, Senders 1 and 2 randomly choose a codebook of size $L \approx 2^{nI(X_1;B_1|X_2Q)}$ and $M \approx 2^{nI(X_2;B_2|X_1Q)}$, respectively. The choice of random code is such that Receiver 1 can first decode the message m because the message m is distinguishable whenever the message set size M is less than $2^{nI(X_2;B_1|Q)}$ and the very strong interference condition in (42) guarantees that this holds. Receiver 1 then uses X_2 as side information to decode message l from Sender 1. Receiver 2 performs similar steps by exploiting the very strong interference condition in (43). The random choice of code guarantees that the expectation of the average error probability is arbitrarily small, and this furthermore guarantees the existence of a particular code with arbitrarily small average error probability. The converse of this theorem follows by the same reasoning as Carleial [8], [17]—the outer bound follows by considering that the conditional mutual information rates in the statement of the theorem are what they could achieve if Senders 1 and 2 maximize their rates individually. ■

Example 9: We now consider an example of a ccqq quantum interference channel with two classical inputs and two quantum outputs:

$$00 \to |00\rangle^{B_1B_2} \tag{45}$$

$$01 \to \cos(\theta)\,|01\rangle^{B_1B_2} + \sin(\theta)\,|10\rangle^{B_1B_2} \tag{46}$$

$$10 \to -\sin(\theta)\,|01\rangle^{B_1B_2} + \cos(\theta)\,|10\rangle^{B_1B_2} \tag{47}$$

$$11 \to |11\rangle^{B_1B_2}. \tag{48}$$

The first classical input is for Sender 1, and the second classical input is for Sender 2. This transformation results if the two senders input one of the four classical states $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ to a "θ-SWAP" unitary transformation that takes this computational basis to the output basis in (45)–(48).

We would like to determine an interval for the parameter θ for which the channel exhibits "very strong" interference. In order to do so, we need to consider classical-quantum states of the following form:

$$\rho^{X_1X_2B_1B_2} \equiv \sum_{x_1,x_2=0}^{1} p_{X_1}(x_1)\, p_{X_2}(x_2)\, |x_1\rangle\langle x_1|^{X_1} \otimes |x_2\rangle\langle x_2|^{X_2} \otimes \psi_{x_1,x_2}^{B_1B_2} \tag{49}$$

where $\psi_{x_1,x_2}^{B_1B_2}$ is one of the pure output states in (45)–(48). We should then check whether the conditions in (42) and (43) hold for all distributions $p_{X_1}(x_1)$ and $p_{X_2}(x_2)$. We can equivalently express these conditions in terms of von Neumann entropies as follows:

$$H(B_1|X_2)_\rho - H(B_1|X_1X_2)_\rho \leq H(B_2)_\rho - H(B_2|X_1)_\rho$$
$$H(B_2|X_1)_\rho - H(B_2|X_1X_2)_\rho \leq H(B_1)_\rho - H(B_1|X_2)_\rho$$

and thus, it suffices to calculate six entropies for states of the form in (49). After some straightforward calculations, we find the results in (50)–(54), as shown at the bottom of the next page,

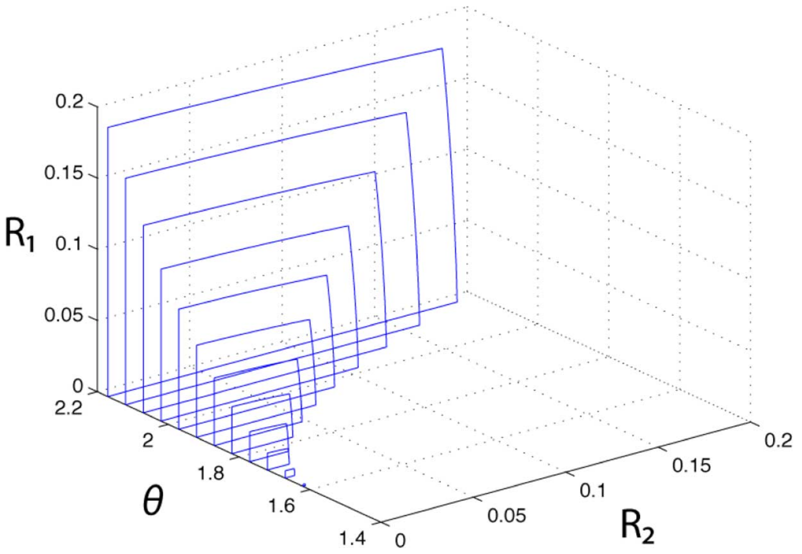

Fig. 2. Capacity region of the "θ-SWAP" interference channel for various values of θ such that the channel exhibits "very strong" interference. The capacity region is largest when θ gets closer to 2.18, and it vanishes when $\theta = \pi/2$ because the channel becomes a full SWAP (at this point, Receiver i gets no information from Sender i, where $i \in \{1, 2\}$).

where $H_2(p)$ is the binary entropy function. We numerically checked for particular values of θ whether the conditions (42) and (43) hold for all distributions $p_{X_1}(x_1)$ and $p_{X_2}(x_2)$, and we found that they hold when $\theta \in [0.96, 2.18] \cup [4.10, 5.32]$ (the latter interval in the union is approximately a shift of the first interval by π). The interval $[0.96, 2.18]$ contains $\theta = \pi/2$, the value of θ for which the capacity should vanish because the transformation is equivalent to a full SWAP (the channel at this point has "too strong" interference). We compute the capacity region given in Theorem 8 for several values of θ in the interval $\theta \in [\pi/2, 2.18]$ (it is redundant to evaluate for other intervals because the capacity region is symmetric about $\pi/2$ and it is also equivalent for the two π-shifted intervals $[0.96, 2.18]$ and $[4.1, 5.32]$). Fig. 2 plots these capacity regions for several values of θ in the interval $[\pi/2, 2.18]$.

2) Quantum Simultaneous Decoding Inner Bound: The two-sender quantum simultaneous decoder from Theorem 2 and Corollary 3 allows us to establish a nontrivial inner bound on the capacity of the quantum interference channel. The strategy is simply to consider the induced multiple access channels to each receiver and choose the rates low enough such that each receiver can decode the messages from *both* senders [2], [17]. This gives us the following theorem.

Theorem 10 (Simultaneous Decoding Inner Bound): Let a ccqq quantum interference channel as in (41) be given. Then, an achievable rate region is the union of all rates R_1 and R_2 satisfying the following inequalities:

$$\begin{aligned} R_1 &\leq \min\left\{I(X_1; B_1|X_2Q)_\rho, I(X_1; B_2|X_2Q)_\rho\right\} \\ R_2 &\leq \min\left\{I(X_2; B_2|X_1Q)_\rho, I(X_2; B_1|X_1Q)_\rho\right\} \\ R_1 + R_2 &\leq \min\left\{I(X_1X_2; B_1|Q)_\rho, I(X_1X_2; B_2|Q)_\rho\right\} \end{aligned}$$

where the union is over input distributions $p_Q(q)\, p_{X_1|Q}(x_1|q)\, p_{X_2|Q}(x_2|q)$.

Proof: The proof exploits the two-sender quantum simultaneous decoder from Corollary 3. We first generate a time-sharing sequence q^n according to the product distribution $p_{Q^n}(q^n)$. Let Sender 1 generate a codebook $\{X_1^n(m_1)\}_{m_1}$ independently and randomly according to the distribution $p_{X_1|Q}(x_1|q)$, and let Sender 2 generate a codebook $\{X_2^n(m_2)\}_{m_2}$ with the distribution $p_{X_2|Q}(x_2|q)$. The induced ccq multiple access channel to Receiver 1 is $x_1, x_2 \rightarrow \rho^{B_1}_{x_1,x_2}$, and the induced channel to Receiver 2 is $x_1, x_2 \rightarrow \rho^{B_2}_{x_1,x_2}$. Corollary 3 states that there exists a simultaneous decoding

$$\begin{aligned} H(B_1|X_1X_2)_\rho &= H(B_2|X_1X_2)_\rho = \left(p_{X_1}(0)p_{X_2}(1) + p_{X_1}(1)p_{X_2}(0)\right) H_2\left(\cos^2(\theta)\right) && (50) \\ H(B_1)_\rho &= H_2\left(p_{X_1}(0) + \left(p_{X_1}(1)p_{X_2}(0) - p_{X_1}(0)p_{X_2}(1)\right)\sin^2(\theta)\right) && (51) \\ H(B_2)_\rho &= H_2\left(p_{X_2}(0) + \left(p_{X_1}(0)p_{X_2}(1) - p_{X_1}(1)p_{X_2}(0)\right)\sin^2(\theta)\right) && (52) \\ H(B_2|X_1)_\rho &= p_{X_1}(0)H_2\left(p_{X_2}(1)\cos^2(\theta)\right) + p_{X_1}(1)H_2\left(p_{X_2}(0)\cos^2(\theta)\right) && (53) \\ H(B_1|X_2)_\rho &= p_{X_2}(0)H_2\left(p_{X_1}(1)\cos^2(\theta)\right) + p_{X_2}(1)H_2\left(p_{X_1}(0)\cos^2(\theta)\right) && (54) \end{aligned}$$

POVM $\{\Lambda_{m_1,m_2}\}$ for Receiver 1 (corresponding to the random choice of code) such that (55), as shown at the bottom of the page, holds as long as

$$\begin{aligned} R_1 &\leq I(X_1; B_1 | X_2 Q)_\rho \\ R_2 &\leq I(X_2; B_1 | X_1 Q)_\rho \\ R_1 + R_2 &\leq I(X_1 X_2; B_1 | Q)_\rho . \end{aligned}$$

Similarly, we can invoke Corollary 3 to show that there is a simultaneous decoding POVM $\{\Gamma_{m_1,m_2}\}$ for Receiver 2 such that (56), as shown at the bottom of the page, holds as long as

$$\begin{aligned} R_1 &\leq I(X_1; B_2 | X_2 Q)_\rho \\ R_2 &\leq I(X_2; B_2 | X_1 Q)_\rho \\ R_1 + R_2 &\leq I(X_1 X_2; B_2 | Q)_\rho . \end{aligned}$$

Thus, if we choose the rates as given in the statement of the theorem, then all six of the aforementioned inequalities are satisfied, implying that the inequality in (57), as shown at the bottom of the page, holds. Invoking the following operator inequality:

$$I - \Lambda^{B_1^n}_{m_1,m_2} \otimes \Gamma^{B_2^n}_{m_1,m_2} \leq I - \Lambda^{B_1^n}_{m_1,m_2} + I - \Gamma^{B_2^n}_{m_1,m_2}$$

and derandomizing the expectation implies the existence of a code upon which all parties can agree. The agreed upon code has vanishing error probability in the asymptotic limit. ■

3) Strong Interference: The simultaneous decoding inner bound from the previous section allows us to determine the capacity of a ccqq interference channel in the setting of "strong" interference (see [17, pp. 6–12]). The conditions for "strong" interference are that the following information inequalities should hold for all distributions $p_{X_1}(x_1)$ and $p_{X_2}(x_2)$:

$$I(X_1; B_1 | X_2)_\rho \leq I(X_1; B_2 | X_2)_\rho \tag{58}$$

$$I(X_2; B_2 | X_1)_\rho \leq I(X_2; B_1 | X_1)_\rho \tag{59}$$

where $\rho^{X_1 X_2 B_1 B_2}$ is a state of the form in (44).

Theorem 11 (Strong Interference):

Let a ccqq quantum interference channel as in (41) be given which satisfies the condition of "strong interference" as in (58) and (59). Then, the capacity region of such a channel is the union of all rates R_1 and R_2 satisfying the following inequalities:

$$\begin{aligned} R_1 &\leq I(X_1; B_1 | X_2 Q)_\rho \\ R_2 &\leq I(X_2; B_2 | X_1 Q)_\rho \\ R_1 + R_2 &\leq \min \left\{ I(X_1 X_2; B_1 | Q)_\rho , I(X_1 X_2; B_2 | Q)_\rho \right\} \end{aligned}$$

where the union is over input distributions

$$p_Q(q)\, p_{X_1|Q}(x_1|q)\; p_{X_2|Q}(x_2|q) .$$

Proof: The proof exploits the quantum simultaneous decoding inner bound from Theorem 10 and the strong interference conditions in (58) and (59). The matching outer bound follows from similar reasoning as in [17, pp. 6–13], though using quantum information inequalities rather than classical ones. ■

4) Han–Kobayashi Achievable Rate Region: The following result provides an achievable rate region for the reliable transmission of classical data over a *ccqq* quantum interference channel (assuming Conjecture 4 regarding the existence of a quantum simultaneous decoder). We should mention that this result was subsequently proved by Sen [51] without relying on Conjecture 4. The statement of the theorem generates codes constructed from a single copy of a ccqq quantum interference channel. We can obtain the regularization of the region by blocking the channel k times and constructing codes from the blocked channel (for any finite k).

Theorem 12 (Achievable Rate Region for the Quantum Interference Channel):

Assume Conjecture 4 holds. Let $\mathcal{S}_\theta$ be the set of tuples of nonnegative reals (S_1, S_2, T_1, T_2) such that

$$S_1 \leq I(U_1; B_1 | W_1 W_2)_\theta \tag{60}$$

$$T_1 \leq I(W_1; B_1 | U_1 W_2)_\theta \tag{61}$$

$$\mathbb{E}_{X_1^n, X_2^n, Q^n} \left\{ \frac{1}{M_1 M_2} \sum_{m_1,m_2} \operatorname{Tr} \left\{ (I - \Lambda_{m_1,m_2})\, \rho^{B_1^n}_{X_1^n(m_1), X_2^n(m_2)} \right\} \right\} \leq \frac{\epsilon}{2} \tag{55}$$

$$\mathbb{E}_{X_1^n, X_2^n, Q^n} \left\{ \frac{1}{M_1 M_2} \sum_{m_1,m_2} \operatorname{Tr} \left\{ (I - \Gamma_{m_1,m_2})\, \rho^{B_2^n}_{X_1^n(m_1), X_2^n(m_2)} \right\} \right\} \leq \frac{\epsilon}{2} \tag{56}$$

$$\mathbb{E}_{X_1^n, X_2^n, Q^n} \left\{ \frac{1}{M_1 M_2} \sum_{m_1,m_2} \operatorname{Tr} \left\{ \left[\left(I - \Lambda^{B_1^n}_{m_1,m_2} \right) + \left(I - \Gamma^{B_2^n}_{m_1,m_2} \right) \right] \rho^{B_1^n B_2^n}_{X_1^n(m_1), X_2^n(m_2)} \right\} \right\} \leq \epsilon \tag{57}$$

$$T_2 \leq I(W_2; B_1|U_1 W_1)_\theta \tag{62}$$
$$S_1 + T_1 \leq I(U_1 W_1; B_1|W_2)_\theta \tag{63}$$
$$S_1 + T_2 \leq I(U_1 W_2; B_1|W_1)_\theta \tag{64}$$
$$T_1 + T_2 \leq I(W_1 W_2; B_1|U_1)_\theta \tag{65}$$
$$S_1 + T_1 + T_2 \leq I(U_1 W_1 W_2; B_1)_\theta \tag{66}$$
$$S_2 \leq I(U_2; B_2|W_1 W_2)_\theta \tag{67}$$
$$T_1 \leq I(W_1; B_2|U_2 W_2)_\theta \tag{68}$$
$$T_2 \leq I(W_2; B_2|U_2 W_1)_\theta \tag{69}$$
$$S_2 + T_1 \leq I(U_2 W_1; B_2|W_2)_\theta \tag{70}$$
$$S_2 + T_2 \leq I(U_2 W_2; B_2|W_1)_\theta \tag{71}$$
$$T_1 + T_2 \leq I(W_1 W_2; B_2|U_2)_\theta \tag{72}$$
$$S_2 + T_1 + T_2 \leq I(U_2 W_1 W_2; B_2)_\theta \tag{73}$$

where θ is a state of the following form:

$$\theta^{U_1 U_2 W_1 W_2 B_1 B_2} \equiv \sum_{u_1, u_2, w_1, w_2} p_{U_1}(u_1)\, p_{U_2}(u_2)\, p_{W_1}(w_1)\, p_{W_2}(w_2) \; |u_1\rangle\langle u_1|^{U_1} \otimes |u_2\rangle\langle u_2|^{U_2} \otimes |w_1\rangle\langle w_1|^{W_1} \otimes |w_2\rangle\langle w_2|^{W_2} \otimes \rho^{B_1 B_2}_{f_1(u_1, w_1), f_2(u_2, w_2)} \tag{74}$$

and $f_1 : \mathcal{U}_1 \times \mathcal{W}_1 \rightarrow \mathcal{X}_1$ and $f_2 : \mathcal{U}_2 \times \mathcal{W}_2 \rightarrow \mathcal{X}_2$ are arbitrary functions. A rate region is achievable if for all $\epsilon > 0$ and sufficiently large n, there exists a code with vanishing average error probability as given in (75), shown at the bottom of the page, where $\rho_{f_1^n(u_1^n(i), w_1^n(k)), f_2^n(u_2^n(j), w_2^n(m))}$ represents the encoded state, i is a "personal" message of Sender 1, k is a "common" message of Sender 1, j is a "personal" message of Sender 2, m is a "common" message of Sender 2, $\{\Lambda_{i,k,m}\}$ is the POVM of Receiver 1, and $\{\Gamma_{j,k,m}\}$ is the POVM of Receiver 2. An achievable rate region for the quantum interference channel $x_1, x_2 \rightarrow \rho_{x_1, x_2}$ is the set of all rates $(S_1 + T_1, S_2 + T_2)$, where $(S_1, S_2, T_1, T_2) \in \mathcal{S}_\theta$ and θ is a state of the form in (74).

Proof: We merely need to set up how the senders select a code randomly and the rest of the proof follows by reasoning similar to that of Han and Kobayashi [20], although we require an application of Conjecture 4. Fig. 3 depicts the Han–Kobayashi coding strategy. Sender 1 generates 2^{nS_1} "personal" codewords $\{u_1^n(i)\}_{i \in [1, \ldots, L_1]}$ according to the distribution $p_{U_1^n}(u_1^n)$ and 2^{nT_1} "common" codewords $\{w_1^n(k)\}_{k \in [1, \ldots, M_1]}$ according to the distribution $p_{W_1^n}(w_1^n)$. Sender 2 generates 2^{nS_2} "personal" codewords $\{u_2^n(j)\}_{j \in [1, \ldots, L_2]}$ according to the distribution $p_{U_2^n}(u_2^n)$ and 2^{nT_2} "common" codewords $\{w_2^n(m)\}_{m \in [1, \ldots, M_2]}$ according to the distribution $p_{W_2^n}(w_2^n)$. Receiver 1 "sees" a three-input multiple access channel after tracing over Receiver 2's system, and the relevant state for randomly selecting a code is many copies of $\mathrm{Tr}_{B_2}\left\{\theta^{U_1 U_2 W_1 W_2 B_1 B_2}\right\}$. Receiver 2 "sees" a three-input multiple access channel after tracing over Receiver 1's system, and the relevant state for randomly selecting a code is many copies of $\mathrm{Tr}_{B_1}\left\{\theta^{U_1 U_2 W_1 W_2 B_1 B_2}\right\}$. Observe that these states are of the form needed to apply Conjecture 4. A direct application of Conjecture 4 to the state $\mathrm{Tr}_{B_2}\left\{\theta^{U_1 U_2 W_1 W_2 B_1 B_2}\right\}$ shows that there exists a POVM that can distinguish the common messages of both senders and the personal message of Sender 1 provided that (60)–(66) hold. Similarly, a direct application of Conjecture 4 to the state $\mathrm{Tr}_{B_1}\left\{\theta^{U_1 U_2 W_1 W_2 B_1 B_2}\right\}$ shows that there exists a POVM that can distinguish the common messages of both senders and the personal message of Sender 2 provided that (67)–(73) hold. We obtain the bounds in (76) and (77), as shown at the bottom of the page, on the expectation of the average error probability for each code, provided that the rates satisfy the inequalities in (60)–(73). We then sum the two expectations of the average error probabilities together. Since the expectation is bounded above by some arbitrarily small, positive number ϵ, there exists a particular code such that the bound in (78), shown at the bottom of the next page, holds. We finally apply the bound

$$I - \Lambda_{i,k,m} \otimes \Gamma_{j,k,m} \leq (I - \Lambda_{i,k,m}) + (I - \Gamma_{j,k,m})$$

that holds for any two commuting positive operators each less than or equal to the identity, to get the bound in (75) on the average error probability. This demonstrates that any rate pair

$$\frac{1}{L_1 L_2 M_1 M_2} \sum_{i,j,k,m} \mathrm{Tr}\left\{(I - \Lambda_{i,k,m} \otimes \Gamma_{j,k,m})\, \rho_{f_1^n(u_1^n(i), w_1^n(k)), f_2^n(u_2^n(j), w_2^n(m))}\right\} \leq \epsilon \tag{75}$$

$$\mathbb{E}\left\{\frac{1}{L_1 M_1 M_2} \sum_{i,k,m} \mathrm{Tr}\left\{(I - \Lambda_{i,k,m})\, \rho_{f_1^n(u_1^n(i), w_1^n(k)), f_2^n(u_2^n(j), w_2^n(m))}\right\}\right\} \leq \frac{\epsilon}{2} \tag{76}$$

$$\mathbb{E}\left\{\frac{1}{L_2 M_1 M_2} \sum_{j,k,m} \mathrm{Tr}\left\{(I - \Gamma_{j,k,m})\, \rho_{f_1^n(u_1^n(i), w_1^n(k)), f_2^n(u_2^n(j), w_2^n(m))}\right\}\right\} \leq \frac{\epsilon}{2} \tag{77}$$

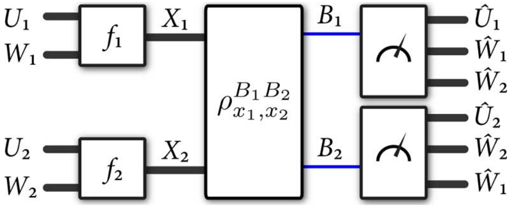

Fig. 3. Han–Kobayashi coding strategy. Sender 1 selects codewords according to a "personal" random variable U_1 and a "common" random variable W_1. She then acts on U_1 and W_1 with some deterministic function f_1 that outputs a variable X_1 which serves as a classical input to the interference channel. Sender 2 uses a similar encoding. Receiver 1 performs a measurement to decode both variables of Sender 1 and the common random variable W_2 of Sender 2. Receiver 2 acts similarly. The advantage of this coding strategy is that it makes use of interference in the channel by having each receiver partially decode what the other sender is transmitting. Theorem 12 gives the rates that are achievable assuming that Conjecture 4 holds.

$(S_1 + T_1, S_2 + T_2)$ is achievable for the quantum interference channel (up to Conjecture 4). ■

Extending the strategies of the previous section and this section to the case of a quantum interference channel with quantum inputs and quantum outputs is straightforward. The senders have the choice to prepare density operators, conditional on classical inputs, as input to this general quantum interference channel, and this extra preprocessing for preparation effectively induces a *ccqq* quantum interference channel for which they are coding. Thus, the achievable rate regions include an extra degree of freedom in the choice of density operators at the inputs. Also, Theorems 8 and 11 are no longer optimal in the case of "very strong" or "strong" interference because entanglement at the individual encoders could increase capacity for certain interference channels [25].

5) Rates Achievable by Successive Decoding: In Section V-A on the multiple access channel, we saw that a successive decoding strategy can be used to achieve certain rate tuples. Then, by time-sharing between the different codes achieving these rates, it is possible to construct good codes for the full capacity region of the multiple access channel. To obtain an inner bound for the interference channel, one could try to use these codes for the two induced multiple access channels. However, this strategy is not well adapted in this setting because the codebooks obtained for the two multiple access channels are not necessarily the same for fixed rates R_1 and R_2. In addition, decoding a codebook constructed by time-sharing between two codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ assumes that both $\mathcal{C}_1$ and $\mathcal{C}_2$ are decodable, and these codes do in general depend on the properties of the channel for which one is coding. For this reason, a time-sharing strategy that works for one of the induced multiple access channels might not work for the other one.

It is, however, possible to use successive decoding strategies for an interference channel in the following way. We start by considering a strategy where both receivers are asked to decode both messages, i.e., we are dealing with the compound multiple access channel. Such a strategy defines an achievable rate region known as the "successive decoding inner bound" for the interference channel (cf., [17, pp. 6--7]). Suppose that Receiver 1 starts by decoding the message of Sender 2 and then the message of Sender 1, and Receiver 2 does the same. We can describe the decode orderings of the receivers by the two permutations $\pi_1 = (2,1)$ and $\pi_2 = (2,1)$. In this case, we know that the random code defined by picking 2^{nR_1} and 2^{nR_2} codewords independently according to the product distributions $p^n_{X_1^n}$ and $p^n_{X_2^n}$ is decodable on average for Receiver 1 provided $R_1 < I(X_1;B_1|X_2)$ and $R_2 < I(X_2;B_1)$. Moreover, it is decodable on average for Receiver 2 provided $R_1 < I(X_2;B_2|X_1)$ and $R_2 < I(X_2;B_2)$. Thus, the rate pairs $R_1 < \min\{I(X_1;B_1|X_2), I(X_1;B_2|X_2\})$ and $R_2 < \min\{I(X_2;B_1), I(X_2;B_2)\}$ are all achievable for the interference channel. Recall that Receiver 2 is actually not interested in the message sent by Sender 1. The only reason to decode the message of Sender 1 is to be able to decode the message of Sender 2 at a higher rate. It is thus useless to require Receiver 2 to decode the message of Sender 1 after decoding the message of Sender 2.

The above ordering shows that the rate pairs R_1, R_2, where $R_1 < I(X_1;B_1|X_2)$ and $R_2 < \min\{I(X_2;B_1),\ I(X_2;B_2)\}$ are all achievable for the interference channel. Naturally, we can do the same for all decode orderings π_1, π_2 and we can achieve rates arbitrarily close to the following points:

$$P_1 = (I(X_1;B_1|X_2), \min\{I(X_2;B_1), I(X_2;B_2)\}) \quad (79)$$

$$\begin{aligned} P_2 = (&\min\{I(X_1;B_1|X_2), I(X_1;B_2)\} \\ &\min\{I(X_2;B_1), I(X_2;B_2|X_1)\}) \end{aligned} \quad (80)$$

$$P_3 = (\min\{I(X_1;B_1), I(X_1;B_2)\}, I(X_2;B_2|X_1)) \quad (81)$$

$$P_4 = (I(X_1;B_1), I(X_2;B_2)). \quad (82)$$

Of course, one can use time-sharing between these different codes for the interference channel to obtain other achievable rates. These rates are illustrated in the RHS of Fig. 4.

Improving Rates Using Rate-Splitting: As can be seen in Fig. 4, the region defined by the convex hull of the points (79)–(82) is, in general, smaller than the simultaneous decoding inner bound. A natural question is whether it is possible to obtain the simultaneous decoding inner bound, or even more generally, the full Han–Kobayashi rate region using a more sophisticated successive decoding argument. There exists an attempt to answer this question for the classical interference channel [52]. This attempt exploits rate-splitting [19] and a careful analysis of the geometrical structure of the 4-D region (corresponding to the two natural multiple access channels defined by the interference channel) that projects down to the 2-D Chong–Motani–Garg region [9]. The Chong–Motani–Garg region is known to be equivalent to the Han–Kobayashi region

$$\frac{1}{L_1 L_2 M_1 M_2} \sum_{i,j,k,m} \mathrm{Tr}\left\{\left[(I-\Lambda_{i,k,m}) + (I-\Gamma_{j,k,m})\right] \rho_{f_1^n\left(u_1^n(i), w_1^n(k)\right), f_2^n\left(u_2^n(j), w_2^n(m)\right)}\right\} \le \epsilon \quad (78)$$

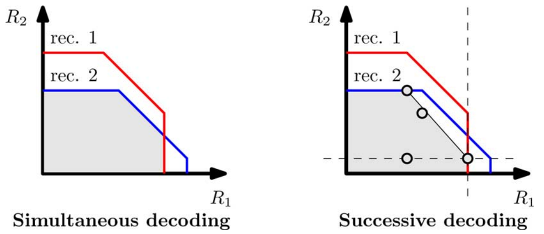

Fig. 4. These plots show achievable rates regions for the interference channel for simultaneous decoding and successive decoding strategies with fixed input distributions. Using a simultaneous decoding strategy, it is possible to achieve the intersection of the two regions of the corresponding multiple access channels. Using a successive decoding strategy, we obtain four achievable rate points that correspond to the possible decoding orders for the two multiple access channels. The solid red and blue lines outline the different multiple access channel achievable rate regions, and the shaded gray areas outline the achievable rate regions for the two different decoding strategies.

when considering all possible input distributions [10], [35]. The argument of [52] rests on an assumption that the change of the code distribution dictated by applying the rate-splitting technique at the convenience of some receiver does not affect the other receiver's decoding ability. Unfortunately, this assumption does not hold in general. We explain this issue in greater detail in the following paragraphs.

Consider an input distribution $p_X(x)$ on some alphabet $\mathcal{X}$. Let $\mathcal{C}_0$ be the codebook obtained by picking 2^{nR} independent codewords of length n distributed according to $p_{X^n}(x^n)$. A split of $p_X(x)$ consists of a function $f : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ and distributions $p_U(u)$ and $p_V(v)$ such that $f(U, V) \sim p_X(x)$, where $U \sim p_U(u)$ and $V \sim p_V(v)$ are independent [19]. The rate-splitting technique in general refers to following coding strategy. Generate a code $\mathcal{C}_U$ from the distribution $p_{U^n}(u^n)$ consisting of 2^{nR_U} independent codewords and a code $\mathcal{C}_V$ from the distribution $p_{V^n}(v^n)$ consisting of 2^{nR_V} independent codewords, where $R_U + R_V = R$. The codebook $\mathcal{C}_{\text{split}}$ is defined as $\{f^n(u^n, v^n) : (u^n, v^n) \in \mathcal{C}_U \times \mathcal{C}_V\}$. Note that $\mathcal{C}_{\text{split}}$ contains $2^{n(R_U+R_V)} = 2^{nR}$ codewords. Furthermore, the codewords of $\mathcal{C}_{\text{split}}$ are all distributed according to $p_{X^n}(x^n)$. The difference between this codebook and $\mathcal{C}_0$ is that the codewords in $\mathcal{C}_{\text{split}}$ are *not* pairwise independent because two codewords in $\mathcal{C}_{\text{split}}$ could arise from the same u^n and $v_1^n \neq v_2^n$ where $u^n \in \mathcal{C}_U$ and v_1^n, $v_2^n \in \mathcal{C}_V$.

Now we describe how to choose the rates R_U and R_V. Suppose that $R = I(X; Y)$ where Y is the output of a channel on input X. Then, a natural choice for R_U and R_V is $R_U = I(U; Y)$ and $R_V = I(V; Y|U)$ because $I(X; Y) = I(U; Y) + I(V; Y|U)$. Observe that the values of R_U and R_V depend on the channel. Consider now a code for an interference channel, where X is to be decoded by both receivers. Such an additional requirement arises for example for the common messages in the Han–Kobayashi inner bound strategy. Let $R = I(X; Y_1)$ and $R \leq I(X; Y_2)$. Using the codebook $\mathcal{C}_0$, both receivers are able to decode X. However, when coding for a multiple access channel with output Y_1, we might want to split $p_X(x)$ into $p_U(u)$ and $p_V(v)$ and use the codebook $\mathcal{C}_{\text{split}}$ for X with rates $R_U = I(U; Y_1)$ and $R_V = I(V; Y_1|U)$ instead of using $\mathcal{C}_0$ [19]. We perform this split because we want to get a noncorner point of the rate region for the multiple access channel with output Y_1 only using successive decoding. In this case, Receiver 1 can decode with small error probability. We should, however, keep in mind that we are coding for an interference channel and we also want Receiver 2 to decode X. The problem is that it is possible that $R_U = I(U; Y_1) > I(U; Y_2)$, in which case Receiver 2 cannot decode U and thus cannot decode X. In this case, the code obtained by splitting according to the first receiver's prescription is not a good code for the second receiver and hence not a good code for the interference channel.

One can however use rate-splitting to obtain potentially better rates than the four points (79)–(82) that can be achieved using a simple successive decoding strategy. In fact, splitting the two inputs of the interference channel as in the Han–Kobayashi strategy into a "personal" and a "common" part and requiring each receiver to decode both common parts induces two three-user multiple access channels. One can naturally use all 6×6 pairs of decoding orders to obtain an achievable rate pair for the interference channel. Fig. 5 shows some rates that can be achieved using such a strategy for a classical Gaussian interference channel.

Of course, it is possible to split the inputs even further, leading to two six-user multiple access channels. An interesting open question is to determine whether such a strategy can achieve the full Han–Kobayashi region—such a result would be important for the quantum interference channel because it would immediately lead to a way to achieve the analogous Han–Kobayashi region without employing Conjecture 4.

B. Outer Bound

We also give a simple outer bound for the capacity of the quantum interference channel. This result follows naturally from a classical result of Sato's [46], where he observes that any code for the quantum interference channel also gives codes for three quantum multiple access channel subproblems, one for Receiver 1, another for Receiver 2, and a third for the two receivers considered together. Thus, if we have an outer bound on the underlying quantum multiple access channel capacities [62], then we can trivially get an outer bound on the quantum interference channel capacity. We omit the following theorem's proof because of its similarity to Sato's proof.

Theorem 13:
Consider the Sato region defined as follows:

$$\mathcal{R}_{\text{Sato}}(\mathcal{N}) \overset{\triangle}{=} \bigcup_{p_Q(q)p_1(x_1|q)p_2(x_2|q)} \{(R_1, R_2)\} \qquad (83)$$

where R_1 and R_2 are rates satisfying the following inequalities:

$$R_1 \leq I(X_1; B_1|X_2 Q)_\theta \qquad (84)$$

$$R_2 \leq I(X_2; B_2|X_1 Q)_\theta \qquad (85)$$

$$R_1 + R_2 \leq I(X_1 X_2; B_1 B_2|Q)_\theta. \qquad (86)$$

The aforementioned entropic quantities are with respect to the following state:

$$\theta^{Q X_1 X_2 B_1 B_2} \equiv \sum_{q, x_1, x_2} p_Q(q) p_1(x_1|q) p_2(x_2|q) \; |q\rangle\langle q|^Q \otimes |x_1\rangle\langle x_1|^{X_1} \otimes |x_2\rangle\langle x_2|^{X_2} \otimes \rho_{x_1 x_2}^{B_1 B_2}. \qquad (87)$$

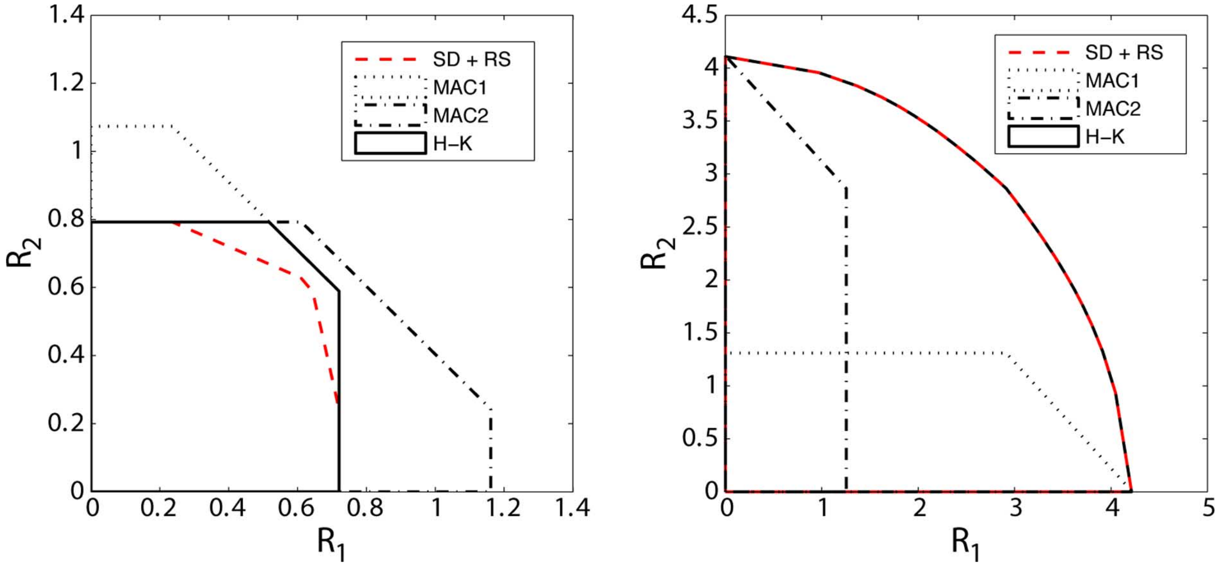

Fig. 5. These two figures plot rate pairs that the senders and receivers in a classical Gaussian interference channel can achieve with successive decoding and rate-splitting $(\mathrm{SD} + \mathrm{RS})$. The figures compare these rates with those achievable by the Han–Kobayashi (HK) coding strategy, while also plotting the regions corresponding to the two induced multiple access channels to each receiver (MAC1 and MAC2). The LHS figure demonstrates that, for a particular choice of signal-to-noise ratio (SNR) and interference to noise (INR) parameters $(\mathrm{SNR1} = 1.7, \mathrm{SNR2} = 2, \mathrm{INR1} = 3.4, \mathrm{INR2} = 4)$, successive decoding with rate-splitting does not perform as well as the Han–Kobayashi strategy. The RHS figure demonstrates that, for a different choice of parameters $(\mathrm{SNR1} = 343, \mathrm{SNR2} = 296, \mathrm{INR1} = 5, \mathrm{INR2} = 5)$, the two strategies perform equally well.

Then, the region $\mathcal{R}_{\mathrm{Sato}}$ forms an outer bound on the capacity region of the quantum interference channel.

VII. Connection to Unitary Gate Capacities

Considerable effort has been devoted to the problem of establishing the information theoretic capacities of an interaction $U : C \otimes D \rightarrow C \otimes D$ between two quantum systems [3], [22]–[24]. One imagines that Charlie controls the system represented by the C Hilbert space while Donna controls D, and that they would like to exploit U to communicate or establish correlations. (More generally, the interaction might be modeled by a Hamiltonian, but that situation can be reduced to the unitary case.) Since U has two inputs and two outputs, this is a special case of a quantum interference channel, and so Theorem 12 will yield achievable rates for classical communication over U and, as we shall see, significantly more.

When U is thought of as an interference channel (say, with quantum inputs A_1 and A_2 and quantum outputs B_1 and B_2 as discussed at the end of Section VI-A4), Charlie plays the roles of both Sender 1 and Receiver 2, while Donna plays the roles of both Sender 2 and Receiver 1 (Fig. 6 depicts this communication scenario). Theorem 12 then gives achievable rates for simultaneous Charlie-to-Donna and Donna-to-Charlie classical communication over U. Indeed, it appears to provide the first nontrivial protocol accomplishing this task for general bidirectional channels. (Earlier protocols assumed free shared entanglement between Charlie and Donna [3].) To apply the theorem, it suffices to identify $A_1 = B_2 = C$ and $A_2 = B_1 = D$ in the interference channel $\mathcal{N}^{A_1A_2 \rightarrow B_1B_2}(\rho) = U\rho U^{\dagger}$. The communication rates achievable for the θ-SWAP channel of Example 9, for instance, apply equally well to this setting.

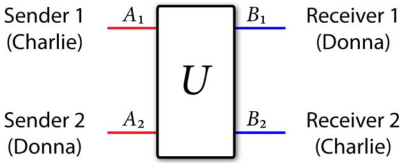

Fig. 6. Connection between a quantum interference channel and a bidirectional unitary gate. The quantum interference channel has quantum inputs A_1 and A_2 and quantum outputs B_1 and B_2. We can identify Sender 1 and Receiver 2 as Charlie and Sender 2 and Receiver 1 as Donna to make a connection with the bidirectional unitary gate setting.

The fact that Charlie and Donna are each both sender and receiver gives them some flexibility not available for general interference channels. Most importantly, in this "bidirectional" setting, they are permitted to use U sequentially, reinserting their outputs into the channel in each successive round [3]. Codes for standard interference channels, on the other hand, cannot take advantage of this flexibility, and so finding the optimal tradeoff between forward and backward communication will likely require codes specifically tailored to the bidirectional setting.

As an interference channel, U is also special because the *only* noise is due to interference: the channel itself is noiseless. Because U does not leak information to an environment, communication can be made coherent at essentially no cost. This allowed Harrow and Leung to establish the following remarkable result, which we state informally using resource inequalities [13]. Recall that $[c \rightarrow c]$ represents a classical bit of communication from Charlie to Donna, $[q \rightarrow q]$ one qubit of communication from Charlie to Donna, and $[q \rightarrow qq]$ one cobit from Charlie to Donna, that is, the isometry $\sum_x \alpha_x |x\rangle^C \rightarrow \sum_x \alpha_x |x\rangle^C |x\rangle^D$. $[c \leftarrow c]$ [21], $[q \leftarrow q]$ and $[qq \leftarrow q]$ represent the same resources but with Donna the sender and Charlie the receiver. Finally, $[qq]$

represents a single shared ebit. For a rigorous definition of resource inequalities, see [13] and [24].

Theorem 14 [22]:

For any bipartite unitary (or isometry) U and R_1, $R_2 \geq 0$, each of the following resource inequalities is equivalent:

$$\langle U \rangle \geq R_1[c \rightarrow c] + R_2[c \leftarrow c] + E[qq] \tag{88}$$

$$\langle U \rangle \geq R_1[q \rightarrow qq] + R_2[qq \leftarrow q] + E[qq] \tag{89}$$

$$\langle U \rangle \geq \frac{R_1}{2}[q \rightarrow q] + \frac{R_2}{2}[q \leftarrow q] + \left(E - \frac{R_1 + R_2}{2} \right)[qq]. \tag{90}$$

Note that the inequalities need only hold in the limit of a large number of uses of U and might require the catalytic use of resources. Still, they imply that for bidirectional channels, the codes we have designed for sending classical data can also be used to send cobits, ebits, and even qubits. In particular, any rates of classical communication that are achievable can automatically be upgraded to cobit communication rates. While our codes should be effective for cobit communication, they have not been designed to generate entanglement. While they can do so at the rate $R_1 + R_2$ by virtue of the fact that a cobit can be used to generate an ebit, that process might be inefficient. In fact, Harrow and Leung have even exhibited a particular channel with C and D each consisting of k qubits for which $R_1 + R_2$ can never exceed $O(\log k)$ but for which E can be larger than $k-1$ [23]. For that channel, our codes would produce an amount of entanglement exponentially smaller than optimal. Rectifying that problem would require modifying the interference channel codes we developed in this paper to also establish shared randomness between the two receivers; such shared randomness would automatically become entanglement in the bidirectional unitary setting.

VIII. Outlook

Calculating the capacity of the interference channel in the classical setting has been an open problem for many years now, and calculating the capacity of the quantum interference channel will be at least as difficult to solve. We have proved that a quantum simultaneous decoder exists for a multiple access channel with two senders, and we have given some evidence that it should exist for channels with three senders. This conjecture holds at least in the case of a quantum multiple access channel in which certain averages of the channel outputs commute. If this conjecture holds in the general case, it immediately implies that the Han–Kobayashi rate region, expressed in terms of Holevo information quantities, is an achievable rate region for the quantum interference channel. Note that even though the general conjecture is still open, the Han–Kobayashi rate region was recently shown to be achievable [51].

Even though Theorem 12 is now known to hold [51], it would still be very interesting to prove Conjecture 4. A proof of this conjecture would probably have important consequences for multiuser quantum information theory since it would allow for many classical information theory results based on simultaneous decoding to be adapted to the quantum setting. It could also likely prove an entanglement-assisted version of a quantum simultaneous decoder by exploiting the coding techniques from [33], and this would in turn lead to another interesting generalization of the Han–Kobayashi rate region where we assume that senders share entanglement with their partner receivers. Reference [63] made progress in this direction by proving the existence of a quantum simultaneous decoder for an entanglement-assisted quantum multiple access channel with two senders, though the three-sender case is still open.

Also, just as there are many different capacities for a single-sender single-receiver quantum channel, we would expect that there are many interesting capacities that we could study for a quantum interference channel. In fact, we initially attempted to use some of the well-known decoupling techniques for the case of quantum information transmission over the quantum interference channel [28], [1], but we were not able to achieve nontrivial rates.

Another important question to consider for the *quantum* interference channel is as follows: Is there anything that quantum mechanics can offer to improve upon the Han–Kobayashi achievable rate region? Quantum effects might play some unexpected role for the quantum interference channel and allow us to achieve a rate region that is superior to the well-known Han–Kobayashi rate region.

Finally, it could be that quantum simultaneous decoding is not necessary in order to achieve the Han–Kobayashi region. In fact, our first attempt at the proof of Theorem 12 was to quantize the successive decoding method from [52], by exploiting the coding techniques from [14] and [62] tailored for classical communication. But we found an issue with the technique in [52] even for the classical interference channel because rate-splitting at the convenience of one receiver affects the other receiver's decoding abilities. Thus, it remains open to determine if a successive decoding strategy can achieve the Han–Kobayashi rate region.

Appendix

Consider a density operator ρ with the following spectral decomposition:

$$\rho = \sum_x p_X(x) \left| x \right\rangle \left\langle x \right| .$$

The weakly typical subspace is defined as the span of all vectors such that the sample entropy $\overline{H}\left(x^n\right)$ of their classical label is close to the true entropy $H(X)$ of the distribution $p_X(x)$ [41], [60]:

$$T_\delta^{X^n} \equiv \text{span}\left\{ \left| x^n \right\rangle : \left| \overline{H}\left(x^n\right) - H(X) \right| \leq \delta \right\}$$

where

$$\overline{H}\left(x^n\right) \equiv -\frac{1}{n}\log\left(p_{X^n}\left(x^n\right)\right)$$

$$H(X) \equiv -\sum_x p_X(x) \log p_X(x).$$

The projector $\Pi_{\rho,\delta}^n$ onto the typical subspace of ρ is defined as

$$\Pi_{\rho,\delta}^n \equiv \sum_{x^n \in T_\delta^{X^n}} \left| x^n \right\rangle \left\langle x^n \right|$$

where we have "overloaded" the symbol $T_\delta^{X^n}$ to refer also to the set of δ-typical sequences:

$$T_\delta^{X^n} \equiv \left\{x^n : \left|\overline{H}\left(x^n\right) - H(X)\right| \leq \delta\right\}.$$

The three important properties of the typical projector are as follows:

$$\begin{aligned}\operatorname{Tr}\left\{\Pi_{\rho,\delta}^n \rho^{\otimes n}\right\} &\geq 1-\epsilon \\ \operatorname{Tr}\left\{\Pi_{\rho,\delta}^n\right\} &\leq 2^{n[H(X)+\delta]} \\ 2^{-n[H(X)+\delta]}\Pi_{\rho,\delta}^n \leq \Pi_{\rho,\delta}^n \rho^{\otimes n} \Pi_{\rho,\delta}^n &\leq 2^{-n[H(X)-\delta]}\Pi_{\rho,\delta}^n\end{aligned}$$

where the first property holds for arbitrary ϵ, $\delta > 0$ and sufficiently large n.

Consider an ensemble $\{p_X(x), \rho_x\}_{x\in\mathcal{X}}$ of states. Suppose that each state ρ_x has the following spectral decomposition:

$$\rho_x = \sum_y p_{Y|X}\left(y|x\right)\left|y_x\right\rangle\left\langle y_x\right|.$$

Consider a density operator ρ_{x^n} which is conditional on a classical sequence $x^n \equiv x_1 \cdots x_n$:

$$\rho_{x^n} \equiv \rho_{x_1} \otimes \cdots \otimes \rho_{x_n}.$$

We define the weak conditionally typical subspace as the span of vectors (conditional on the sequence x^n) such that the sample conditional entropy $\overline{H}\left(y^n|x^n\right)$ of their classical labels is close to the true conditional entropy $H\left(Y|X\right)$ of the distribution $p_{Y|X}\left(y|x\right)p_X(x)$ [41], [61]:

$$T_\delta^{Y^n|x^n} \equiv \operatorname{span}\left\{\left|y_{x^n}^n\right\rangle : \left|\overline{H}\left(y^n|x^n\right) - H\left(Y|X\right)\right| \leq \delta\right\}$$

where

$$\begin{aligned}\overline{H}\left(y^n|x^n\right) &\equiv -\frac{1}{n}\log\left(p_{Y^n|X^n}\left(y^n|x^n\right)\right) \\ H\left(Y|X\right) &\equiv -\sum_x p_X(x)\sum_y p_{Y|X}\left(y|x\right)\log p_{Y|X}\left(y|x\right).\end{aligned}$$

The projector $\Pi_{\rho_{x^n},\delta}$ onto the weak conditionally typical subspace of ρ_{x^n} is as follows:

$$\Pi_{\rho_{x^n},\delta} \equiv \sum_{y^n \in T_\delta^{Y^n|x^n}} \left|y_{x^n}^n\right\rangle\left\langle y_{x^n}^n\right|$$

where we have again overloaded the symbol $T_\delta^{Y^n|x^n}$ to refer to the set of weak conditionally typical sequences:

$$T_\delta^{Y^n|x^n} \equiv \left\{y^n : \left|\overline{H}\left(y^n|x^n\right) - H\left(Y|X\right)\right| \leq \delta\right\}.$$

The three important properties of the weak conditionally typical projector are as follows:

$$\begin{aligned}\mathbb{E}_{X^n}\left\{\operatorname{Tr}\left\{\Pi_{\rho_{X^n},\delta}\rho_{X^n}\right\}\right\} &\geq 1-\epsilon \\ \operatorname{Tr}\left\{\Pi_{\rho_{x^n},\delta}\right\} &\leq 2^{n[H(Y|X)+\delta]} \\ 2^{-n[H(Y|X)+\delta]}\,\Pi_{\rho_{x^n},\delta} &\leq \Pi_{\rho_{x^n},\delta}\,\rho_{x^n}\,\Pi_{\rho_{x^n},\delta} \\ &\leq 2^{-n[H(Y|X)-\delta]}\,\Pi_{\rho_{x^n},\delta}\end{aligned}$$

where the first property holds for arbitrary ϵ, $\delta > 0$ and sufficiently large n, and the expectation is with respect to the distribution $p_{X^n}\left(x^n\right)$.

Lemma 15 (Gentle Operator Lemma for Ensembles [42], [61], [60]):

Given an ensemble $\{p_X(x), \rho_x\}$ with expected density operator $\rho \equiv \sum_x p_X(x)\rho_x$, suppose that an operator Λ such that $I \geq \Lambda \geq 0$ succeeds with high probability on the state ρ:

$$\operatorname{Tr}\left\{\Lambda\rho\right\} \geq 1-\epsilon.$$

Then, the subnormalized state $\sqrt{\Lambda}\rho_x\sqrt{\Lambda}$ is close in expected trace distance to the original state ρ_x:

$$\mathbb{E}_X\left\{\left\|\sqrt{\Lambda}\rho_X\sqrt{\Lambda} - \rho_X\right\|_1\right\} \leq 2\sqrt{\epsilon}.$$

Acknowledgment

We acknowledge discussions with Charlie Bennett, Kamil Brádler, Nilanjana Datta, Frédéric Dupuis, Saikat Guha, Aram Harrow, Min-Hsiu Hsieh, Debbie Leung, Will Matthews, Marco Piani, Eren Şaşoglu, Graeme Smith, John Smolin, Jon Tyson, Mai Vu, Andreas Winter, and Shen Chen Xu. P. Hayden acknowledges support from the Canada Research Chairs program, the Perimeter Institute, CIFAR, FQRNT's INTRIQ, MITACS, NSERC, ONR through grant N000140811249, and QuantumWorks. M. M. Wilde acknowledges support from the MDEIE (Québec) PSR-SIIRI international collaboration grant. I. Savov acknowledges support from FQRNT and NSERC.

References

[1] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, "The mother of all protocols: Restructuring quantum information's family tree," *Proc. Royal Soc. A*, vol. 465, no. 2108, pp. 2537–2563, 2009.

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

[3] C. H. Bennett, A. W. Harrow, D. W. Leung, and J. A. Smolin, "On the capacities of bipartite Hamiltonians and unitary gates," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1895–1911, Aug. 2003.

[4] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Trans. Inf. Theory*, vol. 48, no. 10, pp. 2637–2655, Oct. 2002.

[5] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 2, pp. 197–207, Mar. 1973.

[6] I. Bjelaković, J.-D. Deuschel, T. Krüger, R. Seiler, R. Siegmund-Schultze, and A. Szkoła, *A Quantum Version of Sanov's Theorem*, 2004 [Online]. Available: arXiv:quant-ph/0412157

[7] F. G. S. L. Brandao and M. B. Plenio, "A generalization of quantum Stein's lemma," *Commun. Math. Phys.*, vol. 295, pp. 791–791, 2010.

[8] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 5, pp. 569–569, Sep. 1975.

[9] H.-F. Chong, M. Motani, and H. K. Garg, "A comparison of two achievable rate regions for the interference channel," in *Proc. USCD-ITA Workshop*, San Diego, CA, Feb. 2006.

[10] H.-F. Chong, M. Motani, H. K. Garg, and H. E. Gamal, "On the Han-Kobayashi region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3188–3195, 2008.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley-Interscience, 1991.

[12] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 44–55, Jan. 2005.

[13] I. Devetak, A. W. Harrow, and A. Winter, "A resource framework for quantum Shannon theory," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4587–4618, Oct. 2008.

[14] I. Devetak and P. W. Shor, "The capacity of a quantum channel for simultaneous transmission of classical and quantum information," *Commun. Math. Phys.*, vol. 256, pp. 287–303, 2005.

[15] F. Dupuis, P. Hayden, and K. Li, "A father protocol for quantum broadcast channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2946–2956, Jun. 2010 [Online]. Available: arXiv:quant-ph/0612155

[16] N. Dutil, "Multiparty quantum protocols for assisted entanglement distillation," Ph.D. dissertation, McGill Univ., Montreal, QC, Canada, 2011 [Online]. Available: arXiv:1105.4657

[17] A. E. Gamal and Y.-H. Kim, *Lecture Notes on Network Information Theory*, 2010 [Online]. Available: arXiv:1001.3404

[18] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, "Classical capacity of the lossy bosonic channel: The exact solution," *Phys. Rev. Lett.*, vol. 92, no. 2, pp. 027902–027902, Jan. 2004.

[19] A. J. Grant, B. Rimoldi, R. L. Urbanke, and P. A. Whiting, "Rate-splitting multiple access for discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 873–890, Mar. 2001.

[20] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.

[21] A. Harrow, "Coherent communication of classical messages," *Phys. Rev. Lett.*, vol. 92, no. 9, pp. 097902–097902, 2004.

[22] A. W. Harrow and D. W. Leung, "Bidirectional coherent classical communication," *Quantum Inf. Comput.*, vol. 5, no. 4–5, pp. 380–395, 2005.

[23] A. W. Harrow and D. W. Leung, "An exponential separation between the entanglement and communication capacities of a bipartite unitary interaction," in *Proc. IEEE Inf. Theory Workshop*, 2008, pp. 381–385.

[24] A. W. Harrow and P. W. Shor, "Time reversal and exchange symmetries of unitary gate capacities," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 462–475, Jan. 2010.

[25] M. B. Hastings, "Superadditivity of communication capacity using entangled inputs," *Nature Phys.*, vol. 5, pp. 255–257, 2009.

[26] M. Hayashi, *Quantum Information: An Introduction*. New York: Springer, 2006.

[27] 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.

[28] P. Hayden, M. Horodecki, A. Winter, and J. Yard, "A decoupling approach to the quantum capacity," *Open Syst. Inf. Dyn.*, vol. 15, pp. 7–19, Mar. 2008.

[29] P. Hayden and A. Winter, "Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$," *Commun. Math. Phys.*, vol. 284, no. 1, pp. 263–280, Nov. 2008.

[30] 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.

[31] M. Horodecki, J. Oppenheim, and A. Winter, "Partial quantum information," *Nature*, vol. 436, pp. 673–676, 2005.

[32] M. Horodecki, J. Oppenheim, and A. Winter, "Quantum state merging and negative information," *Commun. Math. Phys.*, vol. 269, pp. 107–136, 2007.

[33] 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.

[34] M.-H. Hsieh and M. M. Wilde, "Entanglement-assisted communication of classical and quantum information," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4682–4704, Sep. 2010 [Online]. Available: arXiv:0811.4227

[35] K. Kobayashi and T. S. Han, "A further consideration on the HK and the CMG regions for the interference channel," in *Proc. Inf. Theory Appl. Workshop*, 2007.

[36] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 581–586, Mar. 2004.

[37] G. Kramer, "Review of rate regions for interference channels," in *Proc. Int. Zurich Seminar Commun.*, 2006, pp. 162–165.

[38] H. H.-J. Liao, "Multiple access channels," Ph.D. dissertation, Univ. Hawaii, Honolulu, HI, 1972.

[39] S. Lloyd, "Capacity of the noisy quantum channel," *Phys. Rev. A*, vol. 55, no. 3, pp. 1613–1622, Mar. 1997.

[40] M. Mosonyi and N. Datta, "Generalized relative entropies and the capacity of classical-quantum channels," *J. Math. Phys.*, vol. 50, no. 7, pp. 072104–072104, 2009.

[41] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[42] 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.

[43] J. R. Pierce, "The early days of information theory," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 1, pp. 3–8, Jan. 1973.

[44] R. Renner, "Security of quantum key distribution," Ph.D. dissertation, Swiss Federal Inst. Technol., Zurich, Switzerland, 2005.

[45] A. Rényi, "On measures of information and entropy," in *Proc. 4th Berkeley Symp. Math., Statist., Probability*, 1960, pp. 547–561.

[46] H. Sato, "Two-user communication channels," *IEEE Trans. Inf. Theory*, vol. 23, no. 3, pp. 295–304, May 1977.

[47] H. Sato, "An outer bound to the capacity region of broadcast channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 374–377, May 1978.

[48] H. Sato, "On degraded Gaussian two-user channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 5, pp. 637–640, Sep. 1978.

[49] H. Sato, "The capacity of the Gaussian interference channel under strong interference (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 6, pp. 786–788, Nov. 1981.

[50] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol. 56, no. 1, pp. 131–138, Jul. 1997.

[51] P. Sen, *Achieving the Han-Kobayashi inner bound for the quantum interference channel by sequential decoding*, 2011 [Online]. Available: arXiv:1109.0802

[52] E. Şaşoğlu, "Successive cancellation for cyclic interference channels," in *Proc. IEEE Inf. Theory Workshop*, Porto, Portugal, May 2008, pp. 36–40.

[53] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[54] C. E. Shannon, J. Neyman, Ed., "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Statist. Probability*, , Berkeley, CA, Jun.–Jul. 20–30, 1961, vol. 1, pp. 611–644.

[55] P. W. Shor, "The quantum channel capacity and coherent information," presented at the Proc. MSRI Workshop Quantum Comput., 2002.

[56] G. Smith, J. Smolin, and A. Winter, "The quantum capacity with symmetric side channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4208–4217, Sep. 2008.

[57] M. Tomamichel, R. Colbeck, and R. Renner, "A fully quantum asymptotic equipartition property," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5840–5847, Dec. 2009.

[58] S. Verdu, "Fifty years of Shannon theory," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2057–2078, Oct. 1998.

[59] L. Wang and R. Renner, *One-shot classical-quantum capacity and hypothesis testing*, Jul. 2010 [Online]. Available: arXiv:1007.5456

[60] M. M. Wilde, *From classical to quantum Shannon theory*, Jun. 2011 [Online]. Available: arXiv:1106.1445

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

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

[63] S. C. Xu and M. M. Wilde, *Sequential, successive, and simultaneous decoders for entanglement-assisted classical communication*, Jul. 2011 [Online]. Available: arXiv:1107.1347

[64] J. Yard, P. Hayden, and I. Devetak, *Quantum broadcast channels*, Mar. 2006 [Online]. Available: arXiv:quant-ph/0603098

Omar Fawzi received a Licence (B.S.) in Computer Science and Mathematics from the École Normale Supérieure de Lyon in 2006 and a Master (M.S.) in Computer Science from Université Paris 7 in 2008. He is currently a Ph.D. student in the School of Computer Science, McGill University. His research interests revolve around (quantum) information theory.

Patrick Hayden (M'04) received his doctorate from the University of Oxford in 2001 and was subsequently a postdoctoral fellow at the California Institute of Technology (Caltech) until 2004.

He is an associate professor in McGill University's School of Computer Science and a Distinguished Research Chair of the Perimeter Institute for Theoretical Physics. He is a member of the publications committee of the IEEE Information Theory Society. His research focuses on quantum information theory and its applications to other areas of physics and computer science.

Ivan Savov is a graduate student in the School of Computer Science at McGill University in Montreal, Canada. Previously, he received a B.Eng. degree in electrical engineering in 2005 and a M.Sc. degree in Physics in 2008, both from McGill University. His research interests include network information theory, quantum information theory, error correcting codes, and machine learning.

Pranab Sen received the Ph.D. degree in computer science from the School of Technology and Computer Science at the Tata Institute of Fundamental Research, Mumbai, India, in 2001. Currently, he is a Reader at the School of Technology and Computer Science at the Tata Institute of Fundamental Research, Mumbai, India and a Visiting Professor with the School of Computer Science, McGill University, Montreal, QC, Canada. His current research interests are in quantum algorithms, classical and quantum communication complexity, and quantum Shannon theory.

Mark M. Wilde (M'99) was born in Metairie, Louisiana, USA. He received the B.S. degree in computer engineering from Texas A&M University, College Station, Texas, in 2002, the M.S. degree in electrical engineering from Tulane University, New Orleans, Louisiana, in 2004, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, California, in 2008.

He was a Visiting Researcher at NEC Laboratories America, Princeton, New Jersey, and at the Centre for Quantum Technologies, Singapore, during the latter half of 2008. He then was employed as a Quantum Information Scientist at Science Applications International Corporation during 2009. Currently, he is a Postdoctoral Fellow at the School of Computer Science, McGill University, Montreal, QC, Canada. He has published over 50 articles and preprints in the area of quantum information processing. His current research interests are in quantum Shannon theory and quantum error correction.

Dr. Wilde is a member of the American Physical Society, a writer for The Quantum Times, and has been a reviewer for the IEEE TRANSACTIONS ON INFORMATION THEORY and the IEEE International Symposium on Information Theory.