

The Two-User Compound Interference Channel

Adnan Raja, *Member, IEEE*, Vinod M. Prabhakaran, *Member, IEEE*, and Pramod Viswanath, *Member, IEEE*

Abstract—We introduce the two-user finite state compound interference channel. The main contributions involve both novel inner and outer bounds. For the Gaussian case, we characterize its capacity region to within one bit. The inner bound is multilevel superposition coding but the decoding of the levels is opportunistic, depending on the channel state. The genie aided outer bound is motivated by the typical error events of the achievable scheme.

Index Terms—Compound channel, interference channel, multi-level superposition coding, slow fading channel, universal scheme.

I. INTRODUCTION

THE focus of this paper is the communication scenario depicted in Fig. 1. Two transmitter-receiver pairs communicate reliably in the face of interference. The discrete time complex baseband model is:

$$y_1[m] = h_{11}x_1[m] + h_{21}x_2[m] + z_1[m] \quad (1)$$

$$y_2[m] = h_{12}x_1[m] + h_{22}x_2[m] + z_2[m]. \quad (2)$$

Here m is the time index, y_k is the signal at receiver k while x_k is the signal sent out by the transmitter k (with $k = 1, 2$). The noise sequences $\{z_1[m], z_2[m]\}_m$ are memoryless complex Gaussian with zero mean and unit variance. The transmitters are subject to average power constraints:

$$\sum_{m=1}^N |x_k[m]|^2 \leq NP_k, \quad k = 1, 2, \quad \forall N \geq 1. \quad (3)$$

The complex parameters $\{h_{k\ell}, \ell = 1, 2, k = 1, 2\}$ model the channel coefficients between the pairs of transmitters and receivers. They do not vary with time but the transmitters and receivers have different information about them.

- Receiver k is exactly aware of the two channel coefficients h_{1k}, h_{2k} ; this models *coherent* communication.
- Transmitters are only *coarsely* aware of the channel coefficients: the transmitters know that the channel coefficients

Manuscript received May 06, 2008; revised May 12, 2009. Current version published October 21, 2009. This work was supported in part by the National Science Foundation under Grant NSF CCR 02-37549, by Office of Naval Research under Grant Navy N00014-089-1-0088, and by a Vodafone graduate fellowship. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Toronto, ON, Canada, July 2008.

The authors are with the Department of Electrical and Computer Engineering, and the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: araja2@illinois.edu; vinodmp@illinois.edu; pramodv@illinois.edu).

Communicated by L. Zheng, Associate Editor for Communications.

Color version of Figure 4 in this paper is available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2009.2030475

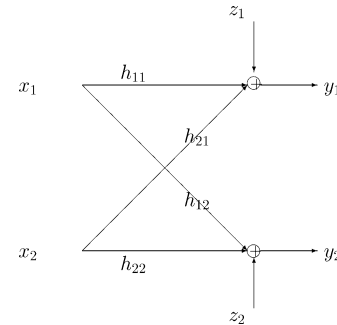


Fig. 1. The two-user Gaussian interference channel.

belong to a finite set. Specifically, both the transmitters know that

$$(h_{1k}, h_{2k}) \in \mathcal{A}_k, \quad k = 1, 2. \quad (4)$$

This models potential partial feedback to the transmitters regarding the channel coefficients.

A more general compound channel model allows for all four channel parameters to jointly take on different choices:

$$(h_{11}, h_{12}, h_{21}, h_{22}) \in \mathcal{A}. \quad (5)$$

However, since the receivers do not cooperate in the interference channel, it turns out that the setting in (5) is no more general than the one in (4). This is explored in Section VIII.

The key problem of interest is the characterization of the capacity region: the set of rate pairs at which arbitrarily reliable communication between the two transmitter-receiver pairs. The “compound” aspect of the channel is in insisting that the receivers be able to decode the messages of interest with arbitrarily high probability, no matter which of the finite states the channel coefficients take on. Our main result is a characterization of the capacity region up to one bit.

A special instance of the problem studied here is the classical two user Gaussian interference channel: in a recent work, Etkin, Tse, and Wang [4] showed that a *single* superposition coding scheme (a specific choice among the broad class of schemes first identified by Han and Kobayashi [1]) achieves performance within one bit of the capacity region. The transmission involved splitting the data into two parts—one *public* and the other *private*—and linearly superposing them. The idea is that the public data stream is decoded by both the receivers while the private data stream only by the receiver of interest. The key identity of the proposed superposition scheme is the following: the power allocated to the private stream is such that it appears at exactly the same level as the background noise at the unintended receiver (the idea is that since the private data stream is being

treated as noise at the unintended receiver, there is no extra incentive to reduce its level even further than that of the additive noise). A novel outer bound developed in [4] showed that this simple superposition scheme is within one bit of the capacity region.

Implementation of the specific superposition scheme proposed above requires each transmitter to be aware of the interference level it is causing to the unintended receiver. In the context of the compound channel being studied here, the transmitter is not aware of the interference level; this poses an obstacle to adopting the idea of appropriately choosing the power level of the private data stream. One possibility could be to set the power level of the private data stream based on the strongest interfering link level (among the set of possible choices)—this would ensure that it is only received *below* noise level when the interfering link level takes on the other possible choices. However, this approach might be too pessimistic and its closeness to optimality is unclear.

We circumvent this problem by proposing the following novel twist to the general superposition coding scheme. Our main idea is best described when the interference links (h_{12} and h_{21} in Fig. 1) take on only two possible values and the direct links are fixed (i.e., the sets \mathcal{A}_1 and \mathcal{A}_2 have cardinality of two, cf. (4)). We now superpose *three* data streams at each transmitter. Two of them, public and private, are as earlier: all receivers in all channel states decode the public message while only the receiver of interest decodes the private message (no matter the channel state, again). The novelty is in the third data stream that we will call *semi-public*: this data stream is decoded by the unintended receiver only when the interference link is the stronger of the two choices (and treated as noise otherwise). As such, this data stream is neither fully private nor public (the unintended receiver either treats it as noise or decodes it based on the channel state) and the nomenclature is chosen to highlight this feature.

The power split rule is the following: the power of the private stream is set such that at the higher of the interference link levels, it is received at the unintended receiver at the same level as the additive noise. The power of the semi-public data stream is set such that it is received at the unintended receiver at the same level as the additive noise only when the interference link level is at the lower of the two possible choices. The rationale is that the semi-public data stream is not decoded only when the interference link level is at the lower of the two possible choices, and thus it can transmit higher power than if its power is restricted by the higher of the interference link levels. This approach scales naturally when the interference link levels can take on more than two possible choices (the number of splits of the data stream is one more than the cardinality of the set of possible choices).

We derive novel outer bounds to show that our simple achievable scheme is within one bit of the capacity region. Our outer bounds are genie aided and are based on the clues provided by the *typical error events* in the achievable scheme. This approach sheds operational insight into the nature of the outer bounds even in the noncompound version (thus eliminating the “guesswork” involved in the derivation, cf. Section IV of [5]).

The paper is organized as follows: we start with a simple two-state compound interference channel. In this setting, both

the direct and interference link levels can take on only one of two possible values (so the sets \mathcal{A}_1 and \mathcal{A}_2 have cardinality two). Using a somewhat abstract setting (described in Section II) that features the Gaussian problem of interest as a special case, we present our main results (both inner and outer bounds) for this two-state compound interference channel. Our definition of the abstract setting is motivated by that chosen in [5] and could be viewed as a natural compound version of the interference channel studied by Telatar and Tse [5]. This is done in Section III. We discuss the insights garnered from these results in the context of the simpler noncompound interference channel in Section VI. Next, we are ready to set up the model and describe the solution the more general finite state interference channel; we do this first in the abstract setting (Section VII) followed by specializing to the Gaussian scenario of interest (Section VIII).

II. MODEL

Consider a two-user, two-state compound memoryless interference channel depicted as shown in Fig. 2. There are two transmitters which want to reliably communicate independent messages to two corresponding receivers. Each receiver can be, independently, in one of the two possible states denoted by α and β , thereby leading to four possible channel realizations. We assume for our model that the interfering link corresponding to each state exhibits degradedness, i.e., the input to the channel from the first transmitter at any discrete time $X_1 \in \mathcal{X}_1$ passes through a degraded discrete memoryless broadcast channel: the two outputs of the degraded broadcast channel are $S_{1\alpha} \in \mathcal{S}_1$ and (the degraded version) $S_{1\beta} \in \mathcal{S}_1$. Similarly, at any time, the input to the channel from transmitter 2 $X_2 \in \mathcal{X}_2$ produces $S_{2\alpha} \in \mathcal{S}_2$ and a degraded version $S_{2\beta} \in \mathcal{S}_2$ of it. The channel to any one of the two receivers is decided by the *state* of that receiver: here there are only two states α and β . Once the state is decided, it is fixed for the entire duration of communication. When the first receiver is in state α , the output at any time is

$$Y_{1\alpha} = f_{1\alpha}(X_1, S_{2\alpha}) \in \mathcal{Y}_1. \tag{6}$$

Similarly, when the first receiver is in state β , the output at any time is

$$Y_{1\beta} = f_{1\beta}(X_1, S_{2\beta}) \in \mathcal{Y}_1. \tag{7}$$

Here $f_{1\alpha}$ and $f_{1\beta}$ are deterministic functions such that for every $x_1 \in \mathcal{X}_1, s_2 \in \mathcal{S}_2$, and $\eta = \alpha, \beta$, the following function is invertible:

$$f_{1\eta}(x_1, \cdot) : \mathcal{S}_2 \rightarrow \mathcal{Y}_1.$$

Likewise, the outputs of user-2 under the two possible states the channel to it can take are defined using similar deterministic functions $f_{2\alpha}$ and $f_{2\beta}$.

We allow each receiver to be in potentially different states, and they are both aware of the state they are in. A pair of communication rates (R_1, R_2) is said to be *achievable* if for every $\epsilon > 0$, there are block length n encoders

$$\text{enc}_k : \{1, \dots, M_k\} \rightarrow \mathcal{X}_k^n, M_k \geq 2^{n(R_k - \epsilon)} \tag{8}$$

$k = 1, 2,$

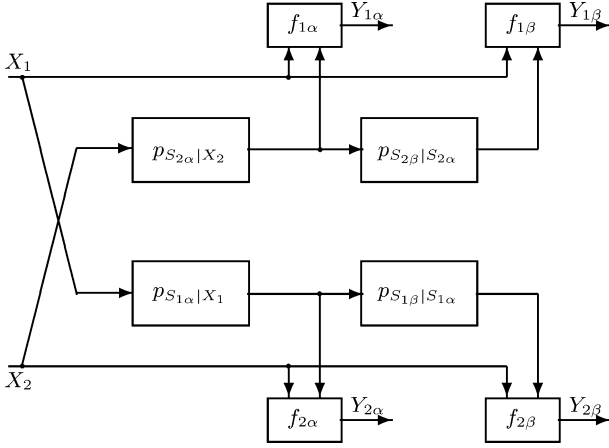


Fig. 2. A two-state compound channel model.

and decoders

$$\text{dec}_{k\eta} : \mathcal{Y}_k^n \rightarrow \{1, \dots, M_k\}, \quad k = 1, 2, \quad \eta = \alpha, \beta \quad (9)$$

such that

$$\begin{aligned} & \frac{1}{M_1 M_2} \sum_{m_1, m_2} \Pr(\text{dec}_{k\eta}(Y_{k\eta}^n) = m_k \\ & k = 1, 2, \eta = \alpha, \beta \mid X_k^n = \text{enc}_k(m_k), k = 1, 2) \\ & \geq 1 - \epsilon. \end{aligned} \quad (10)$$

We are interested in the capacity region \mathcal{C} , which is the set of all achievable (R_1, R_2) pairs. We can make a few observations as follows.

- The channel described here can be thought of as a natural generalization of that studied in [5].
- An important special case occurs when the channels $p_{S_{k\alpha}|X_k}$ and $p_{S_{k\beta}|S_{k\alpha}}$ are *deterministic* for both $k = 1, 2$. This channel is a compound version of the deterministic channel considered by El Gamal and Costa [2] with the interference in state β being a deterministic function of the interference in state α .
- The compound Gaussian interference channel, with the cardinality of both the sets \mathcal{A}_1 and \mathcal{A}_2 restricted to 2 (in the notation introduced in Section 1), is a special instance of the model in Fig. 2. We start with a compound Gaussian interference channel with

$$\begin{aligned} (h_{11}, h_{21}) & \in \{(h_{11\alpha}, h_{21\alpha}), (h_{11\beta}, h_{21\beta})\} \\ (h_{22}, h_{12}) & \in \{(h_{22\alpha}, h_{12\alpha}), (h_{22\beta}, h_{12\beta})\}. \end{aligned}$$

Further, without loss of generality, we can assume that

$$|h_{21\alpha}| \geq |h_{21\beta}| \quad (11)$$

$$|h_{12\alpha}| \geq |h_{12\beta}|. \quad (12)$$

With the following assignment, we see that the model in Fig. 2 can capture the Gaussian model in Fig. 1:

$$S_{1\alpha} = h_{12\alpha} X_1 + Z_2 \quad (13)$$

$$S_{1\beta} = \frac{h_{12\beta}}{h_{12\alpha}} S_{1\alpha} + \left(1 - \left|\frac{h_{12\beta}}{h_{12\alpha}}\right|^2\right)^{1/2} Z_2' \quad (14)$$

$$S_{2\alpha} = h_{21\alpha} X_2 + Z_1 \quad (15)$$

$$S_{2\beta} = \frac{h_{21\beta}}{h_{21\alpha}} S_{2\alpha} + \left(1 - \left|\frac{h_{21\beta}}{h_{21\alpha}}\right|^2\right)^{1/2} Z_1' \quad (16)$$

$$Y_{1\alpha} = f_{1\alpha}(X_1, S_{2\alpha}) = h_{11\alpha} X_1 + S_{2\alpha} \quad (17)$$

$$Y_{1\beta} = f_{1\beta}(X_1, S_{2\beta}) = h_{11\beta} X_1 + S_{2\beta} \quad (18)$$

$$Y_{2\alpha} = f_{2\alpha}(X_2, S_{1\alpha}) = h_{22\alpha} X_2 + S_{1\alpha} \quad (19)$$

$$Y_{2\beta} = f_{2\beta}(X_2, S_{1\beta}) = h_{22\beta} X_2 + S_{1\beta}. \quad (20)$$

Here $Z_1, Z_1', Z_2,$ and Z_2' are independent complex Gaussian random variables with unit variance.

III. MAIN RESULT

Our main results on the two-state compound interference channel are the following:

- we first show the performance of an achievable scheme and hence characterize an inner-bound;
- next, we give an outer-bound to the capacity region and quantify the gap between the outer-bound and the achievable scheme;
- specializing to the compound deterministic interference channel, we completely characterize the capacity region;
- specializing to the compound Gaussian interference channel, we characterize the capacity region up to a gap of 1 bit (at all operating SNR values and all channel parameter values).

A. Inner-Bound: Achievable Scheme

The achievable scheme is characterized by \mathcal{P} , the set of random variables

$$(Q, X_1, U_{1\alpha}, U_{1\beta}, X_2, U_{2\alpha}, U_{2\beta}) \quad (21)$$

such that the following Markov chain is satisfied:

$$U_{1\beta} - U_{1\alpha} - X_1 - Q - X_2 - U_{2\alpha} - U_{2\beta}.$$

Alternatively, the joint probability distribution function factors as

$$\begin{aligned} & p(q, x_1, u_{1\alpha}, u_{1\beta}, x_2, u_{2\alpha}, u_{2\beta}) \\ & = p(q)p(x_1|q)p(u_{1\alpha}|x_1)p(u_{1\beta}|u_{1\alpha}) \\ & \quad \times p(x_2|q)p(u_{2\alpha}|x_2)p(u_{2\beta}|u_{2\alpha}). \end{aligned} \quad (22)$$

Our achievable scheme is a multilevel superposition coding one and can be viewed as a generalization of the two-level superposition coding scheme of Chong *et al.* [3]. The random coding method can be intuitively described as follows, using the ‘‘cloud-center’’ analogy from Cover and Thomas (see [9, Sec. 14.6.3]); a formal statement and its proof follow later. The random variables $U_{1\beta}$ and $U_{2\beta}$ are used to generate the *outermost* code books (with rate $R_{1\beta}$ and $R_{2\beta}$, respectively) for the two users. These messages encoded via these code books are decoded by both receivers and, as such, can be interpreted as *public* information. Next, the random variables $U_{1\alpha}$ and $U_{2\alpha}$

are used to generate the next level of code books (with rate $R_{1\alpha}$ and $R_{2\alpha}$, respectively). The messages encoded via these code books are decoded by the receiver with stronger interference (i.e., $R_{X_{k\alpha}}$) but treated as noise by the receiver with weaker interference (i.e., $R_{X_{k\beta}}$); as such, these messages can be viewed as *semi-public* information. Finally, the messages encoded via the inner most code books (rates R_{1p} and R_{2p}) are only decoded by the receiver of interest; thus this constitutes *private* information.

Given $P \in \mathcal{P}$, we define the six-dimensional region as shown in (23)–(55) at the bottom of the page. We define the two-dimensional region

$$\begin{aligned} \mathcal{R}_{\text{in}}(P) \triangleq & \left\{ (R_1, R_2) : \right. \\ & R_1 = R_{1p} + R_{1\alpha} + R_{1\beta}, \\ & R_2 = R_{2p} + R_{2\alpha} + R_{2\beta} \\ & \left. (R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta}) \in \mathcal{R}_{\text{in}}^{(6)}(P) \right\}. \end{aligned} \quad (56)$$

In other words $\mathcal{R}_{\text{in}}(P)$ is the projection of the six-dimensional polytope $\mathcal{R}_{\text{in}}^{(6)}(P)$. One approach to take the projection, is to do the Fourier–Motzkin elimination, as done for the basic superposition coding scheme in the context of the regular (non-compound) interference channel [3]. Doing this explicitly is rather cumbersome as the inequalities here are much more in number than the inequalities that were handled by Chong *et al.* in [3].

Theorem 1: The capacity region \mathcal{C} satisfies

$$\mathcal{C} \supseteq \bigcup_{P \in \mathcal{P}} \mathcal{R}_{\text{in}}(P). \quad (57)$$

Proof: A formal description of the achievable scheme and the proof of this theorem are available in Section IV-A. \square

Particularizing, we restrict ourselves to a subset of \mathcal{P} defined as follows. Given random variables (Q, X_1, X_2) such that X_1 and X_2 are conditionally independent when conditioned on Q , we define random variables $U_{1\alpha}$ and $U_{1\beta}$ which take values in \mathcal{S}_1 , and $U_{2\alpha}$ and $U_{2\beta}$ which take values in \mathcal{S}_2 . They are jointly

$$\begin{aligned} \mathcal{R}_{\text{in}}^{(6)}(P) \triangleq & \{ (R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta}) : \text{satisfying (24)–(55)} \} \quad (23) \\ & R_{1p} \leq I(Y_{1\beta}; X_1 | U_{1\alpha}, U_{2\beta}, Q) = \gamma_{11} \quad (24) \\ & R_{2\beta} + R_{1p} \leq I(Y_{1\beta}; X_1, U_{2\beta} | U_{1\alpha}, Q) = \gamma_{12} \quad (25) \\ & R_{1\alpha} + R_{1p} \leq I(Y_{1\beta}; X_1 | U_{1\beta}, U_{2\beta}, Q) = \gamma_{13} \quad (26) \\ & R_{2\beta} + R_{1\alpha} + R_{1p} \leq I(Y_{1\beta}; X_1, U_{2\beta} | U_{1\beta}, Q) = \gamma_{14} \quad (27) \\ & R_{1\beta} + R_{1\alpha} + R_{1p} \leq I(Y_{1\beta}; X_1 | U_{2\beta}, Q) = \gamma_{15} \quad (28) \\ & R_{2\beta} + R_{1\beta} + R_{1\alpha} + R_{1p} \leq I(Y_{1\beta}; X_1, U_{2\beta} | Q) = \gamma_{16} \quad (29) \\ & R_{1p} \leq I(Y_{1\alpha}; X_1 | U_{1\alpha}, U_{2\alpha}, Q) = \delta_{11} \quad (30) \\ & R_{2\alpha} + R_{1p} \leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{1\alpha}, U_{2\beta}, Q) = \delta_{12} \quad (31) \\ & R_{2\beta} + R_{2\alpha} + R_{1p} \leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{1\alpha}, Q) = \delta_{13} \quad (32) \\ & R_{1\alpha} + R_{1p} \leq I(Y_{1\alpha}; X_1 | U_{1\beta}, U_{2\alpha}, Q) = \delta_{14} \quad (33) \\ & R_{2\alpha} + R_{1\alpha} + R_{1p} \leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{1\beta}, U_{2\beta}, Q) = \delta_{15} \quad (34) \\ & R_{2\beta} + R_{2\alpha} + R_{1\alpha} + R_{1p} \leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{1\beta}, Q) = \delta_{16} \quad (35) \\ & R_{1\beta} + R_{1\alpha} + R_{1p} \leq I(Y_{1\alpha}; X_1 | U_{2\alpha}, Q) = \delta_{17} \quad (36) \\ & R_{2\alpha} + R_{1\beta} + R_{1\alpha} + R_{1p} \leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{2\beta}, Q) = \delta_{18} \quad (37) \\ & R_{2\beta} + R_{2\alpha} + R_{1\beta} + R_{1\alpha} + R_{1p} \leq I(Y_{1\alpha}; X_1, U_{2\alpha} | Q) = \delta_{19} \quad (38) \\ & R_{2p} \leq I(Y_{2\beta}; X_2 | U_{2\alpha}, U_{1\beta}, Q) = \gamma_{21} \quad (39) \\ & R_{1\beta} + R_{2p} \leq I(Y_{2\beta}; X_2, U_{1\beta} | U_{2\alpha}, Q) = \gamma_{22} \quad (40) \\ & R_{2\alpha} + R_{2p} \leq I(Y_{2\beta}; X_2 | U_{2\beta}, U_{1\beta}, Q) = \gamma_{23} \quad (41) \\ & R_{1\beta} + R_{2\alpha} + R_{2p} \leq I(Y_{2\beta}; X_2, U_{1\beta} | U_{2\beta}, Q) = \gamma_{24} \quad (42) \\ & R_{2\beta} + R_{2\alpha} + R_{2p} \leq I(Y_{2\beta}; X_2 | U_{1\beta}, Q) = \gamma_{25} \quad (43) \\ & R_{1\beta} + R_{2\beta} + R_{2\alpha} + R_{2p} \leq I(Y_{2\beta}; X_2, U_{1\beta} | Q) = \gamma_{26} \quad (44) \\ & R_{2p} \leq I(Y_{2\alpha}; X_2 | U_{2\alpha}, U_{1\alpha}, Q) = \delta_{21} \quad (45) \\ & R_{1\alpha} + R_{2p} \leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{2\alpha}, U_{1\beta}, Q) = \delta_{22} \quad (46) \\ & R_{1\beta} + R_{1\alpha} + R_{2p} \leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{2\alpha}, Q) = \delta_{23} \quad (47) \\ & R_{2\alpha} + R_{2p} \leq I(Y_{2\alpha}; X_2 | U_{2\beta}, U_{1\alpha}, Q) = \delta_{24} \quad (48) \\ & R_{1\alpha} + R_{2\alpha} + R_{2p} \leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{2\beta}, U_{1\beta}, Q) = \delta_{25} \quad (49) \\ & R_{1\beta} + R_{1\alpha} + R_{2\alpha} + R_{2p} \leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{2\beta}, Q) = \delta_{26} \quad (50) \\ & R_{2\beta} + R_{2\alpha} + R_{2p} \leq I(Y_{2\alpha}; X_2 | U_{1\alpha}, Q) = \delta_{27} \quad (51) \\ & R_{1\alpha} + R_{2\beta} + R_{2\alpha} + R_{2p} \leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{1\beta}, Q) = \delta_{28} \quad (52) \\ & R_{1\beta} + R_{1\alpha} + R_{2\beta} + R_{2\alpha} + R_{2p} \leq I(Y_{2\alpha}; X_2, U_{1\alpha} | Q) = \delta_{29} \quad (53) \\ & R_{1p} + R_{1\alpha} + R_{1\beta} \geq 0 \quad (54) \\ & R_{2p} + R_{2\alpha} + R_{2\beta} \geq 0. \quad (55) \end{aligned}$$

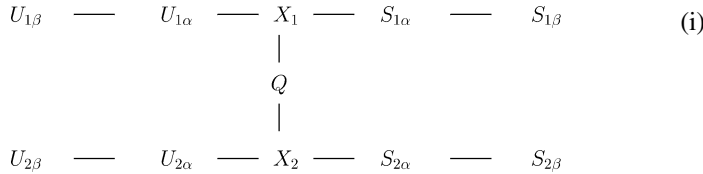


Fig. 3. The Markov chain satisfied by the random variables involved.

distributed with (Q, X_1, X_2) according to the conditional distribution

$$\begin{aligned}
 & p(u_{1\alpha}, u_{1\beta}, u_{2\alpha}, u_{2\beta} | q, x_1, x_2) \\
 &= p_{S_{1\alpha} | X_1}(u_{1\alpha} | x_1) p_{S_{1\beta} | S_{1\alpha}}(u_{1\beta} | u_{1\alpha}) \\
 & \quad \times p_{S_{2\alpha} | X_2}(u_{2\alpha} | x_2) p_{S_{2\beta} | S_{2\alpha}}(u_{2\beta} | u_{2\alpha}). \quad (58)
 \end{aligned}$$

The Markov chain property satisfied by the random variables is pictorially depicted in Fig. 3. Further $U_{k\alpha}$ and $S_{k\alpha}$ have the same marginal distribution and similarly $U_{k\beta}$ and $S_{k\beta}$ also have the same marginal distribution. Our choice is motivated by the choice in the paper by Telatar and Tse [5]. Every member of this family is uniquely determined by joint random variables (Q, X_1, X_2) such that $X_1 - Q - X_2$ is a Markov chain. We will henceforth denote the corresponding regions $R_{\text{in}}^{(6)}(P)$ by $R_{\text{in}}^{(6)}(Q, X_1, X_2)$ and $R_{\text{out}}(P)$ by $R_{\text{out}}(Q, X_1, X_2)$. We now have the natural result as follows.

Corollary 2:

$$\mathcal{C} \supseteq \bigcup_{Q, X_1, X_2} \mathcal{R}_{\text{in}}(Q, X_1, X_2) \quad (59)$$

where the union is over all (Q, X_1, X_2) such that $X_1 - Q - X_2$ is a Markov chain.

Proof: Follows directly from Theorem 1. \square

Remark 1: We observe that the Fourier–Motzkin elimination procedure to implement the projection operation in obtaining $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ would yield only a finite set of inequalities. This procedure is explained in much detail in [6]. The procedure eliminates one variable at each step, but in the process giving rise to potentially a larger set of inequalities. If there are N inequalities to begin with, then after one elimination step the number of inequalities can be upper bounded by $(N/2)^2$. Since, in our case we have to eliminate 4 variables and $N = 30$ to begin with, we can upper bound the number of inequalities remaining by $30^{16}/2^{30}$. Note that this upper bound can be very weak, as at every stage a number of these inequalities are redundant and can be discarded right away. Further, the right hand sides of these remaining inequalities would be linear functions of $p(q)$ and for a fixed $Q = q_0$ the right hand sides form a closed set of finite dimensions. Thus, by Carathéodory’s theorem, we can conclude that the cardinality of Q can taken to be finite without loss of generality in the union in (59) and can be upper bounded by $30^{16}/2^{30}$.

B. Outer-Bound

Theorem 3: For every (Q, X_1, X_2) such that $X_1 - Q - X_2$ is a Markov chain, the region $\mathcal{R}_{\text{out}}(Q, X_1, X_2) \subseteq \mathbb{R}_+^2$ described in Section 5.1 by (143) is such that:

$$\mathcal{C} \subseteq \bigcup_{Q, X_1, X_2} \mathcal{R}_{\text{out}}(Q, X_1, X_2), \quad (60)$$

where the union is over all (Q, X_1, X_2) such that $X_1 - Q - X_2$ is a Markov chain.

(ii) If $(R_1, R_2) \in \mathcal{R}_{\text{out}}(Q, X_1, X_2)$, then $(R_1 - \Delta_1, R_2 - \Delta_2) \in \mathcal{R}_{\text{in}}(Q, X_1, X_2)$, where

$$\begin{aligned}
 & \Delta_1(Q, X_1, X_2) \\
 &= \max(I(X_2; S_{2\alpha} | U_{2\alpha}), I(X_2; S_{2\beta} | U_{2\beta})) \quad (61)
 \end{aligned}$$

$$\begin{aligned}
 & \Delta_2(Q, X_1, X_2) \\
 &= \max(I(X_1; S_{1\alpha} | U_{1\alpha}), I(X_1; S_{1\beta} | U_{1\beta})) \quad (62)
 \end{aligned}$$

in which the random variables are jointly distributed according to (58) and the channel conditional distributions.

Proof: The definition of $\mathcal{R}_{\text{out}}(Q, X_1, X_2)$ in Section V-A is motivated by the *external representation* of $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ that we obtain in Section IV-B. Part(i) is proved in Section V-A. Part(ii) is proved in Section V-B. \square

C. Special Cases

Our model captures two important special cases:

- the compound deterministic interference channel;
- the compound Gaussian interference channel,

as discussed in Section II. Thus our results apply to these cases (readily for the deterministic channel, and with an appropriate approximation result to the continuous alphabet Gaussian channel). Moreover, the structure afforded by these special cases allows us to derive further insight into the nature of the general results derived earlier.

1) *Compound Deterministic Interference Channel:* In this instance, the capacity region is exactly described.

Corollary 4: For the deterministic compound interference channel, the inner bound in Theorem 1 is the capacity region.

Proof: The proof is elementary. When the channel is deterministic, we see that the gap claimed by Theorem 3

$$\Delta_1(Q, X_1, X_2) = \Delta_2(Q, X_1, X_2) = 0. \quad (63)$$

This completes the proof. \square

2) *Two-State Compound Gaussian Interference Channel:* For the Gaussian version, we can characterize the capacity to within one-bit.

Corollary 5: For the two-state compound Gaussian interference channel, the achievable region of Theorem 1 is within at most one bit of the capacity region.

Proof: For the Gaussian channel, each of the mutual information terms in the expressions for $\Delta_1(Q, X_1, X_2)$ and $\Delta_2(Q, X_1, X_2)$ can be upper bounded by 1 bit. To see this, note that $S_{1\alpha} = h_{1\alpha}X_1 + Z_2$ by (13). $U_{1\alpha}$ by our choice has the same marginal distribution as $S_{1\alpha}$ and further given X_1 is independent of $S_{1\alpha}$. Therefore $U_{1\alpha} = h_{1\alpha}X_1 + Z_2'$, where Z_2 and Z_2' are independent and identically distributed memoryless complex Gaussian random variables. Hence

$$I(X_1; S_{1\alpha} | U_{1\alpha}) = h(S_{1\alpha} | U_{1\alpha}) - h(Z_2)$$

$$\leq h(S_{1\alpha} - U_{1\alpha}) - h(Z_2) = 1.$$

Similarly

$$I(X_1; S_{1\beta} | U_{1\beta}) \leq 1 \quad (64)$$

$$I(X_2; S_{2\alpha} | U_{2\alpha}) \leq 1 \quad (65)$$

$$I(X_2; S_{2\beta} | U_{2\beta}) \leq 1. \quad (66)$$

□

Additionally, we can use Gaussian code books to get to within one bit of the capacity.

Corollary 6: For the two-state compound Gaussian interference channel

$$\mathcal{C} \subseteq \mathcal{R}_{\text{out}}(Q^*, X_1^*, X_2^*) \quad (67)$$

where $Q^* = 1$, $X_1^* \sim \mathcal{CN}(0, P_1)$, $X_2^* \sim \mathcal{CN}(0, P_2)$.

This implies that $\mathcal{R}_{\text{in}}(Q^*, X_1^*, X_2^*)$ is within one-bit of the capacity region \mathcal{C} of the two-state Gaussian compound interference channel.

Proof: See Section V-C. □

IV. AN ACHIEVABLE SCHEME

We will present a natural, and novel, achievable scheme first. We will evaluate the set of reliable communication rates using this strategy and hence characterize an inner bound to the capacity region; this will complete the proof of Theorem 1. Next, we will see some important geometric properties of the achievable rate region.

A. Proof Of Theorem 1

Our coding scheme is a natural generalization of the scheme of Chong *et al.* [3]. Since there are two possible states for both receivers, each encoder now sends two sets of common information, with the receivers opportunistically decoding the common information (depending on the state). We choose the random variables corresponding to the two sets of common information in a degraded manner, following the same ordering of degradation of the interferences under the two states (cf.

$$X_k - S_{k\alpha} - S_{k\beta}, \quad k = 1, 2). \quad (68)$$

Fix a $P \in \mathcal{P}$.

Codebook Generation.

Generate a codeword Q^n of length n , generating each element independent and identically distributed (i.i.d.) according to $\prod_{i=1}^n p(q_i)$. For the codeword Q^n , generate $2^{nR_{1\beta}}$ independent codewords

$$U_{1\beta}^n(j_1), \quad j_1 \in \{1, 2, \dots, 2^{nR_{1\beta}}\} \quad (69)$$

generating each element i.i.d. according to $\prod_{i=1}^n p(u_{1\beta i} | q_i)$. For each of the codewords $U_{1\beta}^n(j_1)$, generate $2^{nR_{1\alpha}}$ independent codewords

$$U_{1\alpha}^n(j_1, k_1), \quad k_1 \in \{1, 2, \dots, 2^{nR_{1\alpha}}\} \quad (70)$$

generating each element i.i.d. according to $\prod_{i=1}^n p(u_{1\alpha i} | u_{1\beta i}, q_i)$. For each of the codewords $U_{1\alpha}^n(j_1, k_1)$, generate $2^{nR_{1p}}$ independent codewords

$$X_1^n(j_1, k_1, l_1), \quad l_1 \in \{1, 2, \dots, 2^{nR_{1p}}\} \quad (71)$$

generating each element i.i.d. according to $\prod_{i=1}^n p(x_1 | u_{1\alpha i}, u_{1\beta i}, q_i)$. Similarly generate code books

$$U_{2\beta}^n(j_2), \quad j_2 \in \{1, 2, \dots, 2^{nR_{2\beta}}\} \quad (72)$$

$$U_{2\alpha}^n(j_2, k_2), \quad k_2 \in \{1, 2, \dots, 2^{nR_{2\alpha}}\} \quad (73)$$

$$X_2^n(j_2, k_2, l_2), \quad l_2 \in \{1, 2, \dots, 2^{nR_{2p}}\}. \quad (74)$$

The codebook generation is pictorially represented in Fig. 4.

Encoding.

Transmitter Tx_1 sends $X_1^n(j_1, k_1, l_1)$ to communicate the message indexed by (j_1, k_1, l_1) . Transmitter Tx_2 sends $X_2^n(j_2, k_2, l_2)$ to communicate the message indexed by (j_2, k_2, l_2) .

Decoding.

The receivers do joint typical set decoding. Let $A_\epsilon^{(n)}(\Omega)$ denote the set of jointly typical sequences ω^n where Ω is the probability space containing the entire collection of random variables.

Receiver $\text{Rx}_{1\beta}$ determines a *unique* $(\hat{j}_1, \hat{k}_1, \hat{l}_1)$ and any \hat{j}_2 such that

$$\begin{aligned} & (Q^n, U_{1\beta}^n(\hat{j}_1), U_{1\alpha}^n(\hat{j}_1, \hat{k}_1), X_1^n(\hat{j}_1, \hat{k}_1, \hat{l}_1), U_{2\beta}^n(\hat{j}_2), Y_{1\beta}^n) \\ & \in A_\epsilon^{(n)}(Q, U_{1\beta}, U_{1\alpha}, X_1, U_{2\beta}, Y_{1\beta}). \end{aligned}$$

It declares an error if it fails to find such a choice.

Receiver $\text{Rx}_{1\alpha}$ determines a *unique* $(\hat{j}_1, \hat{k}_1, \hat{l}_1)$ and any (\hat{j}_2, \hat{k}_2) such that

$$\begin{aligned} & (Q^n, U_{1\beta}^n(\hat{j}_1), U_{1\alpha}^n(\hat{j}_1, \hat{k}_1), X_1^n(\hat{j}_1, \hat{k}_1, \hat{l}_1), U_{2\beta}^n(\hat{j}_2) \\ & U_{2\alpha}^n(\hat{j}_2, \hat{k}_2), Y_{1\alpha}^n) \in A_\epsilon^{(n)}(Q, U_{1\beta}, U_{1\alpha}, X_1, U_{2\beta}, U_{2\alpha}, Y_{1\alpha}). \end{aligned}$$

It declares an error if it fails to find such a choice.

Similar decoding is done by receivers $\text{Rx}_{2\beta}$ and $\text{Rx}_{2\alpha}$.

From the analysis of the probability of error, we show in Appendix A that the rate vector

$$(R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta})$$

is achievable if it satisfies conditions (75)–(110) shown at the top of the following page.

$$\begin{aligned}
R_{1p} &\leq I(Y_{1\beta}; X_1 | U_{1\alpha}, U_{2\beta}, Q), & \text{if } R_{1p} > 0 & \quad (75) \\
R_{2\beta} + R_{1p} &\leq I(Y_{1\beta}; X_1, U_{2\beta} | U_{1\alpha}, Q), & \text{if } R_{1p} > 0 & \quad (76) \\
R_{1\alpha} + R_{1p} &\leq I(Y_{1\beta}; X_1 | U_{1\beta}, U_{2\beta}, Q), & \text{if } R_{1\alpha} + R_{1p} > 0 & \quad (77) \\
R_{2\beta} + R_{1\alpha} + R_{1p} &\leq I(Y_{1\beta}; X_1, U_{2\beta} | U_{1\beta}, Q), & \text{if } R_{1\alpha} + R_{1p} > 0 & \quad (78) \\
R_{1\beta} + R_{1\alpha} + R_{1p} &\leq I(Y_{1\beta}; X_1 | U_{2\beta}, Q), & \text{if } R_{1\beta} + R_{1\alpha} + R_{1p} > 0 & \quad (79) \\
R_{2\beta} + R_{1\beta} + R_{1\alpha} + R_{1p} &\leq I(Y_{1\beta}; X_1, U_{2\beta} | Q), & \text{if } R_{1\beta} + R_{1\alpha} + R_{1p} > 0 & \quad (80) \\
R_{1p} &\leq I(Y_{1\alpha}; X_1 | U_{1\alpha}, U_{2\alpha}, Q), & \text{if } R_{1p} > 0 & \quad (81) \\
R_{2\alpha} + R_{1p} &\leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{1\alpha}, U_{2\beta}, Q), & \text{if } R_{1p} > 0 & \quad (82) \\
R_{2\beta} + R_{2\alpha} + R_{1p} &\leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{1\alpha}, Q), & \text{if } R_{1p} > 0 & \quad (83) \\
R_{1\alpha} + R_{1p} &\leq I(Y_{1\alpha}; X_1 | U_{1\beta}, U_{2\alpha}, Q), & \text{if } R_{1\alpha} + R_{1p} > 0 & \quad (84) \\
R_{2\alpha} + R_{1\alpha} + R_{1p} &\leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{1\beta}, U_{2\beta}, Q), & \text{if } R_{1\alpha} + R_{1p} > 0 & \quad (85) \\
R_{2\beta} + R_{2\alpha} + R_{1\alpha} + R_{1p} &\leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{1\beta}, Q), & \text{if } R_{1\alpha} + R_{1p} > 0 & \quad (86) \\
R_{1\beta} + R_{1\alpha} + R_{1p} &\leq I(Y_{1\alpha}; X_1 | U_{2\alpha}, Q), & \text{if } R_{1\beta} + R_{1\alpha} + R_{1p} > 0 & \quad (87) \\
R_{2\alpha} + R_{1\beta} + R_{1\alpha} + R_{1p} &\leq I(Y_{1\alpha}; X_1, U_{2\alpha} | U_{2\beta}, Q), & \text{if } R_{1\beta} + R_{1\alpha} + R_{1p} > 0 & \quad (88) \\
R_{2\beta} + R_{2\alpha} + R_{1\beta} + R_{1\alpha} + R_{1p} &\leq I(Y_{1\alpha}; X_1, U_{2\alpha} | Q), & \text{if } R_{1\beta} + R_{1\alpha} + R_{1p} > 0 & \quad (89) \\
R_{2p} &\leq I(Y_{2\beta}; X_2 | U_{2\alpha}, U_{1\beta}, Q), & \text{if } R_{2p} > 0 & \quad (90) \\
R_{1\beta} + R_{2p} &\leq I(Y_{2\beta}; X_2, U_{1\beta} | U_{2\alpha}, Q), & \text{if } R_{2p} > 0 & \quad (91) \\
R_{2\alpha} + R_{2p} &\leq I(Y_{2\beta}; X_2 | U_{2\beta}, U_{1\beta}, Q), & \text{if } R_{2\alpha} + R_{2p} > 0 & \quad (92) \\
R_{1\beta} + R_{2\alpha} + R_{2p} &\leq I(Y_{2\beta}; X_2, U_{1\beta} | U_{2\beta}, Q), & \text{if } R_{2\alpha} + R_{2p} > 0 & \quad (93) \\
R_{2\beta} + R_{2\alpha} + R_{2p} &\leq I(Y_{2\beta}; X_2 | U_{1\beta}, Q), & \text{if } R_{2\beta} + R_{2\alpha} + R_{2p} > 0 & \quad (94) \\
R_{1\beta} + R_{2\beta} + R_{2\alpha} + R_{2p} &\leq I(Y_{2\beta}; X_2, U_{1\beta} | Q), & \text{if } R_{2\beta} + R_{2\alpha} + R_{2p} > 0 & \quad (95) \\
R_{2p} &\leq I(Y_{2\alpha}; X_2 | U_{2\alpha}, U_{1\alpha}, Q), & \text{if } R_{2p} > 0 & \quad (96) \\
R_{1\alpha} + R_{2p} &\leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{2\alpha}, U_{1\beta}, Q), & \text{if } R_{2p} > 0 & \quad (97) \\
R_{1\beta} + R_{1\alpha} + R_{2p} &\leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{2\alpha}, Q), & \text{if } R_{2p} > 0 & \quad (98) \\
R_{2\alpha} + R_{2p} &\leq I(Y_{2\alpha}; X_2 | U_{2\beta}, U_{1\alpha}, Q), & \text{if } R_{2\alpha} + R_{2p} > 0 & \quad (99) \\
R_{1\alpha} + R_{2\alpha} + R_{2p} &\leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{2\beta}, U_{1\beta}, Q), & \text{if } R_{2\alpha} + R_{2p} > 0 & \quad (100) \\
R_{1\beta} + R_{1\alpha} + R_{2\alpha} + R_{2p} &\leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{2\beta}, Q), & \text{if } R_{2\alpha} + R_{2p} > 0 & \quad (101) \\
R_{2\beta} + R_{2\alpha} + R_{2p} &\leq I(Y_{2\alpha}; X_2 | U_{1\alpha}, Q), & \text{if } R_{2\beta} + R_{2\alpha} + R_{2p} > 0 & \quad (102) \\
R_{1\alpha} + R_{2\beta} + R_{2\alpha} + R_{2p} &\leq I(Y_{2\alpha}; X_2, U_{1\alpha} | U_{1\beta}, Q), & \text{if } R_{2\beta} + R_{2\alpha} + R_{2p} > 0 & \quad (103) \\
R_{1\beta} + R_{1\alpha} + R_{2\beta} + R_{2\alpha} + R_{2p} &\leq I(Y_{2\alpha}; X_2, U_{1\alpha} | Q), & \text{if } R_{2\beta} + R_{2\alpha} + R_{2p} > 0 & \quad (104) \\
R_{1p} &\geq 0 & & \quad (105) \\
R_{1\alpha} &\geq 0 & & \quad (106) \\
R_{1\beta} &\geq 0 & & \quad (107) \\
R_{2p} &\geq 0 & & \quad (108) \\
R_{2\alpha} &\geq 0 & & \quad (109) \\
R_{2\beta} &\geq 0 & & \quad (110)
\end{aligned}$$

Note that (75)–(80) are the decodability conditions at $R_{X_{1\beta}}$; (81)–(89) are the decodability conditions at $R_{X_{1\alpha}}$; (90)–(95) are the decodability conditions at $R_{X_{2\beta}}$; (96)–(104) are the decodability conditions at $R_{X_{2\alpha}}$ and (105)–(110) are stating the fact that the rates are nonnegative real numbers.

Define

$$\tilde{\mathcal{R}}_{\text{in}}^{(6)}(P) \triangleq \{(R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta}) : \text{satisfies (75)–(110)}\} \quad (111)$$

and its projection onto the two dimension space (R_1, R_2) by $\tilde{\mathcal{R}}_{\text{in}}(P)$.

Lemma 7:

$$\mathcal{R}_{\text{in}}(P) \subseteq \tilde{\mathcal{R}}_{\text{in}}(P).$$

Proof: See Appendix B. \square

Thus, we have shown that the capacity region \mathcal{C} satisfies

$$\mathcal{C} \supseteq \bigcup_{P \in \mathcal{P}} \mathcal{R}_{\text{in}}(P). \quad (112)$$

In particular, restricting to a subfamily of \mathcal{P} , where given random variables (Q, X_1, X_2) such that $X_1 - Q - X_2$ is a Markov chain and $(U_{1\alpha}, U_{1\beta}, U_{2\alpha}, U_{2\beta})$ are defined by (58), we get

$$\mathcal{C} \supseteq \bigcup_{(Q, X_1, X_2)} \mathcal{R}_{\text{in}}(Q, X_1, X_2). \quad (113)$$

This completes the proof of Theorem 1. \square

B. Dual Representation Of $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$

We have noted in Remark 1 that a finite set of inequalities (half-planes) are sufficient to describe $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$. It was also pointed out that it is tedious to characterize $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ explicitly. Nevertheless, we would like to derive some useful insights into their properties from the dual representation of

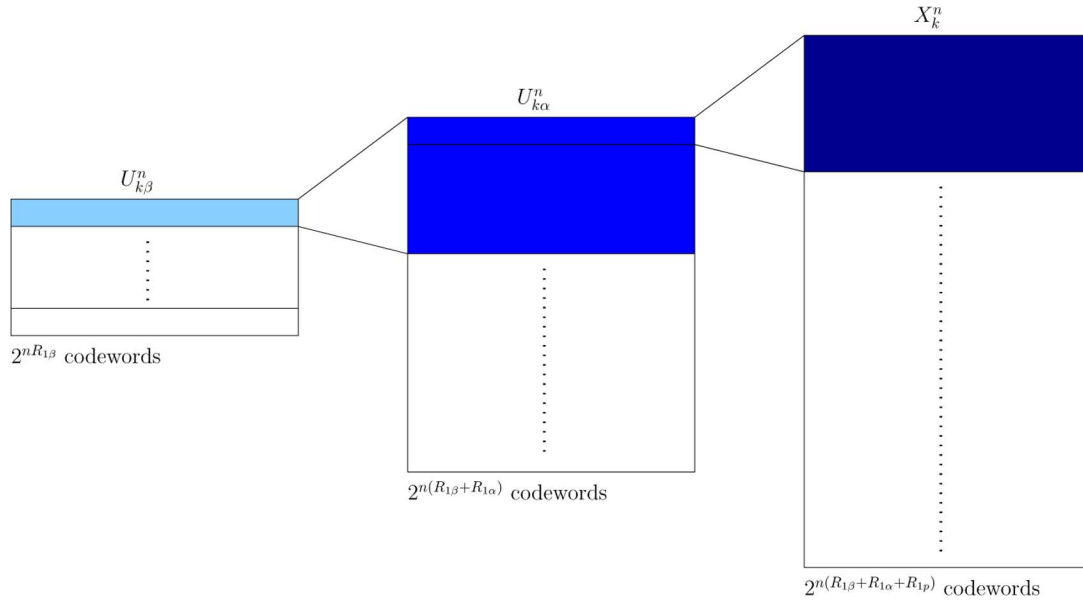


Fig. 4. The multilevel superposition coding scheme.

$\mathcal{R}_{\text{in}}(Q, X_1, X_2)$. These will prove useful in deriving the outer bound.

We begin by noting that $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ is a closed and bounded convex region. (In fact, we know that it is a polyhedron.) The *external representation theorem* of classical Convex set theory (see Theorem 18.8, [8]) states that “an n -dimensional closed convex set in \mathbb{R}^n is the intersection of the closed half-spaces tangent to it”. Thus

$$\begin{aligned} \mathcal{R}_{\text{in}}(Q, X_1, X_2) &= \{(R_1, R_2) : aR_1 + bR_2 \\ &\leq c^*(a, b | (Q, X_1, X_2)), \quad \forall (a, b) \in \mathbb{R}^2\}. \end{aligned} \quad (114)$$

Here, $c^*(a, b | (Q, X_1, X_2))$ is the support function (Section 13, [8]) of $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ and is defined as the solution of the following linear program:

$$\begin{aligned} \text{Max} \quad & aR_1 + bR_2, \\ \text{s.t.} \quad & (R_1, R_2) \in \mathcal{R}_{\text{in}}(Q, X_1, X_2). \end{aligned} \quad (115)$$

Since $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ is the projection of the six-dimensional region $\mathcal{R}_{\text{in}}^{(6)}(Q, X_1, X_2)$, the linear program (115) is equivalent to the linear program (116) shown at the bottom of the page.

The *dual* of the linear program in (116) sheds important geometric information. Let us denote the dual-variables associated with the inequalities (24)–(29) by $\nu_{11}, \dots, \nu_{16}$, with (30)–(38)

by $\mu_{11}, \dots, \mu_{19}$, with (39)–(44) by $\nu_{21}, \dots, \nu_{26}$, with (45)–(53) by $\mu_{21}, \dots, \mu_{29}$ and with (54)–(55) by ω_1 and ω_2 .

Define $\Lambda_{(a,b)} \subset \mathbb{R}^{32}$ by (117)–(124) shown at the bottom of the following page.

For any $\lambda \in \Lambda_{(a,b)}$ define

$$c_{(\lambda,a,b)}^{(\text{in})}(Q, X_1, X_2) \triangleq \sum_{j=1}^2 \left(\sum_{i=1}^6 \nu_{ji} \gamma_{ji} + \sum_{i=1}^9 \mu_{ji} \delta_{ji} \right). \quad (125)$$

The dual linear program is then given by

<p>The dual linear program :</p> $\begin{aligned} \min \quad & c_{(\lambda,a,b)}^{(\text{in})}(Q, X_1, X_2) \\ \text{s.t.} \quad & \lambda \in \Lambda_{(a,b)}. \end{aligned} \quad (126)$
--

By the strong duality theorem

$$\begin{aligned} c^*(a, b | (Q, X_1, X_2)) &= \min \left\{ c_{(\lambda,a,b)}^{(\text{in})}(Q, X_1, X_2) \mid \lambda \in \Lambda_{(a,b)} \right\}. \end{aligned} \quad (127)$$

Therefore

$$\begin{aligned} &\{(R_1, R_2) : aR_1 + bR_2 \leq c^*(a, b | (Q, X_1, X_2))\} \\ &= \bigcap_{\lambda \in \Lambda_{(a,b)}} \{(R_1, R_2) : aR_1 + bR_2 \\ &\leq c_{(\lambda,a,b)}^{(\text{in})}(Q, X_1, X_2)\}. \end{aligned} \quad (128)$$

<p>The primal linear program:</p> $\begin{aligned} \text{Max} \quad & aR_{1p} + aR_{1\alpha} + aR_{1\beta} + bR_{2p} + bR_{2\alpha} + bR_{2\beta} \\ \text{s.t.} \quad & (R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta}) \in \mathcal{R}_{\text{in}}^{(6)}(Q, X_1, X_2). \end{aligned} \quad (116)$
--

Using (114) and (128), $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ can now be described as

$$\begin{aligned} \mathcal{R}_{\text{in}}(Q, X_1, X_2) &= \{(R_1, R_2) : \\ &aR_1 + bR_2 \leq c_{(\lambda, a, b)}^{(\text{in})}(Q, X_1, X_2) \\ &\forall \lambda \in \Lambda_{(a, b)}, \quad \forall (a, b) \in \mathbb{R}^2\}. \end{aligned} \quad (129)$$

This dual representation is essentially characterizing $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ as an intersection of an infinite number of hyper-planes. These hyper-planes are linear inequalities which are linear combinations of (24)–(55). The dual variables are characterizing the way in which the linear combination is taken.

We make two additional important observations.

- The dual variables ω_1 and ω_2 correspond to the inequalities (54) and (55), respectively. These inequalities correspond to just the trivial hyper-planes, viz. $R_1 \geq 0$ and $R_2 \geq 0$, and does not contribute to any other hyper-planes of $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$.
- From Remark 1, we know that only a finite number of inequalities are sufficient to characterize $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$.

These observations lead to the following Lemma.

Lemma 8:

$$\mathcal{R}_{\text{in}}(Q, X_1, X_2) = \{(R_1, R_2) : R_1 \geq 0, R_2 \geq 0 \\ aR_1 + bR_2 \leq c_{(\lambda, a, b)}^{(\text{in})}(Q, X_1, X_2), \forall (\lambda, a, b) \in \Gamma\} \quad (130)$$

where

$$\Gamma = \{(\lambda, a, b) \mid (a, b) \in \mathbb{R}_+^2, \lambda \in \Lambda'_{(a, b)}\} \quad (131)$$

and

$$\Lambda'_{(a, b)} \triangleq \{\lambda \in \Lambda_{(a, b)} : \omega_1 = 0, \omega_2 = 0\}. \quad (132)$$

Further Γ has finite cardinality.

Proof: Every inequality used to define $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ in (129) is described by parameters (a, b) and $\lambda \in \Lambda_{(a, b)}$. Note that this set of inequalities includes the following two inequalities:

$$-R_1 \leq 0 \quad (133)$$

$$-R_2 \leq 0. \quad (134)$$

Consider any inequality, other than the two special ones above, described by (a, b) and $\lambda \in \Lambda_{(a, b)}$, such that $\lambda \notin \Lambda'_{(a, b)}$:

$$aR_1 + bR_2 \leq c_{(\lambda, a, b)}^{(\text{in})}. \quad (135)$$

Define

$$(\tilde{a}, \tilde{b}) \triangleq (a + \omega_1, b + \omega_2). \quad (136)$$

Consider $\tilde{\lambda} \in \Lambda'_{(\tilde{a}, \tilde{b})}$, obtained by replacing ω_1 and ω_2 in λ by 0. Now we have

$$c_{(\lambda, a, b)}^{(\text{in})} = c_{(\tilde{\lambda}, \tilde{a}, \tilde{b})}^{(\text{in})}. \quad (137)$$

Therefore

$$\begin{aligned} (a + \omega_1)R_1 + (b + \omega_2)R_2 \\ = \tilde{a}R_1 + \tilde{b}R_2 \leq c_{(\tilde{\lambda}, \tilde{a}, \tilde{b})}^{(\text{in})} = c_{(\lambda, a, b)}^{(\text{in})}. \end{aligned} \quad (138)$$

The above inequality, along with $R_1 \geq 0$ and $R_2 \geq 0$, implies (135). Therefore we have that (135) is redundant.

Thus we have proved that inequalities that are characterized by a $\lambda \notin \Lambda'_{(a, b)}$ are redundant and can be removed. It also follows from (118)–(124) that $\Lambda'_{(a, b)}$ is an empty set if either a or b is less than 0. Thus we only need to consider inequalities characterized by (a, b) , where $a \geq 0$ and $b \geq 0$. The finiteness of the cardinality of Γ follows from Remark 1. This completes the proof. \square

We end this section by stating the following proposition, that will be used in proving the outer bound.

$$\Lambda_{(a, b)} \triangleq \{(\{\nu_{1i}\}_1^6, \{\mu_{1i}\}_1^9, \{\nu_{2i}\}_1^6, \{\mu_{2i}\}_1^9, \{\omega_i\}_1^2) : \text{satisfying (118)–(124)}\} \quad (117)$$

$$\sum_{i=1}^6 \nu_{1i} + \sum_{i=1}^9 \mu_{1i} - \omega_1 = a \quad (118)$$

$$\sum_{i=3}^6 \nu_{1i} + \sum_{i=4}^9 \mu_{1i} + (\mu_{22} + \mu_{25} + \mu_{28}) + (\mu_{23} + \mu_{26} + \mu_{29}) - \omega_1 = a \quad (119)$$

$$\sum_{i=5}^6 \nu_{1i} + \sum_{i=7}^9 \mu_{1i} + (\mu_{23} + \mu_{26} + \mu_{29}) + (\nu_{22} + \nu_{24} + \nu_{26}) - \omega_1 = a \quad (120)$$

$$\sum_{i=1}^6 \nu_{2i} + \sum_{i=1}^9 \mu_{2i} - \omega_2 = b \quad (121)$$

$$\sum_{i=3}^6 \nu_{2i} + \sum_{i=4}^9 \mu_{2i} + (\mu_{12} + \mu_{15} + \mu_{18}) + (\mu_{13} + \mu_{16} + \mu_{19}) - \omega_2 = b \quad (122)$$

$$\sum_{i=5}^6 \nu_{2i} + \sum_{i=7}^9 \mu_{2i} + (\mu_{13} + \mu_{16} + \mu_{19}) + (\nu_{12} + \nu_{14} + \nu_{16}) - \omega_2 = b \quad (123)$$

$$\mu_{ij}, \nu_{ij}, \omega_i \geq 0. \quad (124)$$

Proposition 1: For any $\lambda \in \Lambda'_{(a,b)}$

$$\begin{aligned} & (\mu_{11} + \mu_{12} + \mu_{13}) + (\nu_{11} + \nu_{12}) \\ &= (\mu_{22} + \mu_{23} + \mu_{25} + \mu_{26} + \mu_{28} + \mu_{29}) \end{aligned} \quad (139)$$

$$\begin{aligned} & (\mu_{14} + \mu_{15} + \mu_{16}) + (\mu_{22} + \mu_{25} + \mu_{28}) + (\nu_{13} + \nu_{14}) \\ &= (\nu_{22} + \nu_{24} + \nu_{26}) \end{aligned} \quad (140)$$

$$\begin{aligned} & (\mu_{21} + \mu_{22} + \mu_{23}) + (\nu_{21} + \nu_{22}) \\ &= (\mu_{12} + \mu_{13} + \mu_{15} + \mu_{16} + \mu_{18} + \mu_{19}) \end{aligned} \quad (141)$$

$$\begin{aligned} & (\mu_{24} + \mu_{25} + \mu_{26}) + (\mu_{12} + \mu_{15} + \mu_{18}) + (\nu_{23} + \nu_{24}) \\ &= (\nu_{12} + \nu_{14} + \nu_{16}). \end{aligned} \quad (142)$$

Proof: The result follows directly from (118)–(123) and the definition of $\Lambda'_{(a,b)}$. \square

V. OUTER BOUND

Our goal in this section is to show that, if (R_1, R_2) is achievable then there exist random variables (Q, X_1, X_2) , where $X_1 - Q - X_2$ is a Markov chain, and a region

$$\begin{aligned} & \mathcal{R}_{\text{out}}(Q, X_1, X_2) \\ & \triangleq \left\{ (R_1, R_2) : R_1 \geq 0, R_2 \geq 0 \right. \\ & \quad \left. aR_1 + bR_2 \leq c_{(\lambda,a,b)}^{(\text{out})}(Q, X_1, X_2), \quad \forall (\lambda, a, b) \in \Gamma \right\} \end{aligned} \quad (143)$$

such that

$$(R_1, R_2) \in \mathcal{R}_{\text{out}}(Q, X_1, X_2). \quad (144)$$

The term $c_{(\lambda,a,b)}^{(\text{out})}(Q, X_1, X_2)$ is defined in Section V-A. Note that our definition of $\mathcal{R}_{\text{out}}(Q, X_1, X_2)$ is inspired by the characterization of $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ that we have obtained through Lemma 8.

Further, quantifying the difference between $c_{(\lambda,a,b)}^{(\text{in})}(Q, X_1, X_2)$ and $c_{(\lambda,a,b)}^{(\text{out})}(Q, X_1, X_2)$ will give us the gap between the inner and the outer bounds.

A. Proof of Theorem 3(i)

Suppose there is a sequence of encoders at rates (R_1, R_2) , sequenced by the block length n , and decoders with probability of error going to 0 as $n \rightarrow \infty$. Fix the block length n and consider the corresponding code book. Let $X_1^n, X_2^n, S_1^n, S_2^n, Y_1^n, Y_2^n$ be the random variables induced by the channel and encoders for uniformly distributed messages, independent across the two users. We define random variables $U_{1\alpha}^n$ which is obtained by passing X_1^n through an independent copy of the channel $p_{S_{1\alpha}|X_1}$, and $U_{1\beta}^n$ by passing the $U_{1\alpha}^n$ so obtained through an independent copy of the channel $p_{S_{1\beta}|S_{1\alpha}}$. Similarly, we also define $U_{2\alpha}^n$ and $U_{2\beta}^n$ from X_2^n and independent copies of $p_{S_{2\alpha}|X_2}$ and $p_{S_{2\beta}|S_{2\alpha}}$. Mathematically, we have the following Markov chain:

$$\begin{aligned} & U_{k\beta i} - U_{k\alpha i} - X_{ki} - S_{k\alpha i} - S_{k\beta i} \\ & \quad \forall 1 \leq i \leq n, k = 1, 2. \end{aligned} \quad (145)$$

And further the marginal distribution of $U_{k\beta i}$'s and $U_{k\alpha i}$'s is the same as the marginal distribution of $S_{k\beta i}$'s and $S_{k\alpha i}$'s respectively.

Since the probability of error goes to 0 as $n \rightarrow \infty$, by Fano's inequality there exists a sequence $\epsilon_n \rightarrow 0$ such that, for every $(\lambda, a, b) \in \Gamma$

$$\begin{aligned} & n(aR_1 + bR_2 - (a+b)\epsilon_n) \\ & \stackrel{(1)}{\leq} \left(\sum_{i=1}^6 \nu_{1i} \right) I(X_1^n; Y_{1\beta}^n) \\ & \quad + \left(\sum_{i=1}^9 \mu_{1i} \right) I(X_1^n; Y_{1\alpha}^n) \\ & \quad + \left(\sum_{i=1}^6 \nu_{2i} \right) I(X_2^n; Y_{2\beta}^n) \\ & \quad + \left(\sum_{i=1}^9 \mu_{2i} \right) I(X_2^n; Y_{2\alpha}^n) \\ & \stackrel{(2)}{\leq} \sum_{i=1}^6 \nu_{1i} I(X_1^n; Y_{1\beta}^n, V_{1\beta}^n) \\ & \quad + \sum_{i=1}^9 \mu_{1i} I(X_1^n; Y_{1\alpha}^n, V_{1\alpha}^n) \\ & \quad + \sum_{i=1}^6 \nu_{2i} I(X_2^n; Y_{2\beta}^n, V_{2\beta}^n) \\ & \quad + \sum_{i=1}^9 \mu_{2i} I(X_2^n; Y_{2\alpha}^n, V_{2\alpha}^n). \end{aligned} \quad (146)$$

Note that for $\lambda \in \Lambda'_{(a,b)}$, we have $\omega_1 = \omega_2 = 0$. Therefore

$$a = \sum_{i=1}^6 \nu_{1i} + \sum_{i=1}^9 \mu_{1i}, \quad (147)$$

$$b = \sum_{i=1}^6 \nu_{2i} + \sum_{i=1}^9 \mu_{2i}. \quad (148)$$

We use this in step (1) along with Fano's inequality for decodability under different states of the receiver. In step (2), we consider *genies* which provide different side-information V 's to the decoders. Consider, for instance, the term $\nu_{11} I(X_1^n; Y_{1\beta}^n, V_{1\beta}^n)$. We will choose the side-information $V_{1\beta}^n$ in such a way that we can form a correspondence between this term and the term contributed to the inner bound by the right hand side of the constraint (24). In particular, we choose the genie provided side-information $V_{1\beta}^n$ to match the error-event corresponding to (24). More specifically, we note that the corresponding error-event is when receiver-1 in state β correctly decodes the other user's common information $U_{2\beta}$, and its own common information $(U_{1\beta}, U_{1\alpha})$, but makes an error in decoding its private message. Hence, the genie provides the side-information $(U_{1\alpha}^n, U_{1\beta}^n, U_{2\beta}^n)$ which can be shrunk to $V_{1\beta}^n = (U_{1\alpha}^n, U_{2\beta}^n)$ because of the Markov relationship between $X_1^n, U_{1\alpha}^n$, and $U_{1\beta}^n$. Now, we expand the term $I(X_1^n; Y_{1\beta}^n, V_{1\beta}^n)$ to get (149). We can repeat these two steps for every term in (146); the (expanded) upper bounds on all the terms are given in (149)–(178) shown on the next two pages.

$$I(X_1^n; Y_{1\beta}^n, V_{11\beta}^n) \triangleq I(X_1^n; Y_{1\beta}^n, U_{1\alpha}^n, U_{2\beta}^n) = H(Y_{1\beta}^n | U_{1\alpha}^n, U_{2\beta}^n) - H(S_{2\beta}^n | U_{2\beta}^n) + H(U_{1\alpha}^n) - H(U_{1\alpha}^n | X_1^n) \quad (149)$$

$$I(X_1^n; Y_{1\beta}^n, V_{12\beta}^n) \triangleq I(X_1^n; Y_{1\beta}^n, U_{1\alpha}^n) = H(Y_{1\beta}^n | U_{1\alpha}^n) - H(S_{2\beta}^n) + H(U_{1\alpha}^n) - H(U_{1\alpha}^n | X_1^n) \quad (150)$$

$$I(X_1^n; Y_{1\beta}^n, V_{13\beta}^n) \triangleq I(X_1^n; Y_{1\beta}^n, U_{1\beta}^n, U_{2\beta}^n) = H(Y_{1\beta}^n | U_{1\beta}^n, U_{2\beta}^n) - H(S_{2\beta}^n | U_{2\beta}^n) + H(U_{1\beta}^n) - H(U_{1\beta}^n | X_1^n) \quad (151)$$

$$I(X_1^n; Y_{1\beta}^n, V_{14\beta}^n) \triangleq I(X_1^n; Y_{1\beta}^n, U_{1\beta}^n) = H(Y_{1\beta}^n | U_{1\beta}^n) - H(S_{2\beta}^n) + H(U_{1\beta}^n) - H(U_{1\beta}^n | X_1^n) \quad (152)$$

$$I(X_1^n; Y_{1\beta}^n, V_{15\beta}^n) \triangleq I(X_1^n; Y_{1\beta}^n, U_{2\beta}^n) = H(Y_{1\beta}^n | U_{2\beta}^n) - H(S_{2\beta}^n | U_{2\beta}^n) \quad (153)$$

$$I(X_1^n; Y_{1\beta}^n, V_{16\beta}^n) \triangleq I(X_1^n; Y_{1\beta}^n) = H(Y_{1\beta}^n) - H(S_{2\beta}^n) \quad (154)$$

$$I(X_1^n; Y_{1\alpha}^n, V_{11\alpha}^n) \triangleq I(X_1^n; Y_{1\alpha}^n, U_{1\alpha}^n, U_{2\alpha}^n) = H(Y_{1\alpha}^n | U_{1\alpha}^n, U_{2\alpha}^n) - H(S_{2\alpha}^n | U_{2\alpha}^n) + H(U_{1\alpha}^n) - H(U_{1\alpha}^n | X_1^n) \quad (155)$$

$$I(X_1^n; Y_{1\alpha}^n, V_{12\alpha}^n) \triangleq I(X_1^n; Y_{1\alpha}^n, U_{1\alpha}^n, U_{2\beta}^n) = H(Y_{1\alpha}^n | U_{1\alpha}^n, U_{2\beta}^n) - H(S_{2\alpha}^n | U_{2\beta}^n) + H(U_{1\alpha}^n) - H(U_{1\alpha}^n | X_1^n) \quad (156)$$

$$I(X_1^n; Y_{1\alpha}^n, V_{13\alpha}^n) \triangleq I(X_1^n; Y_{1\alpha}^n, U_{1\alpha}^n) = H(Y_{1\alpha}^n | U_{1\alpha}^n) - H(S_{2\alpha}^n) + H(U_{1\alpha}^n) - H(U_{1\alpha}^n | X_1^n) \quad (157)$$

$$I(X_1^n; Y_{1\alpha}^n, V_{14\alpha}^n) \triangleq I(X_1^n; Y_{1\alpha}^n, U_{1\beta}^n, U_{2\alpha}^n) = H(Y_{1\alpha}^n | U_{1\beta}^n, U_{2\alpha}^n) - H(S_{2\alpha}^n | U_{2\alpha}^n) + H(U_{1\beta}^n) - H(U_{1\beta}^n | X_1^n) \quad (158)$$

$$I(X_1^n; Y_{1\alpha}^n, V_{15\alpha}^n) \triangleq I(X_1^n; Y_{1\alpha}^n, U_{1\beta}^n, U_{2\beta}^n) = H(Y_{1\alpha}^n | U_{1\beta}^n, U_{2\beta}^n) - H(S_{2\alpha}^n | U_{2\beta}^n) + H(U_{1\beta}^n) - H(U_{1\beta}^n | X_1^n) \quad (159)$$

$$I(X_1^n; Y_{1\alpha}^n, V_{16\alpha}^n) \triangleq I(X_1^n; Y_{1\alpha}^n, U_{1\beta}^n) = H(Y_{1\alpha}^n | U_{1\beta}^n) - H(S_{2\alpha}^n) + H(U_{1\beta}^n) - H(U_{1\beta}^n | X_1^n) \quad (160)$$

$$I(X_1^n; Y_{1\alpha}^n, V_{17\alpha}^n) \triangleq I(X_1^n; Y_{1\alpha}^n, U_{2\alpha}^n) = H(Y_{1\alpha}^n | U_{2\alpha}^n) - H(S_{2\alpha}^n | U_{2\alpha}^n) \quad (161)$$

$$I(X_1^n; Y_{1\alpha}^n, V_{18\alpha}^n) \triangleq I(X_1^n; Y_{1\alpha}^n, U_{2\beta}^n) = H(Y_{1\alpha}^n | U_{2\beta}^n) - H(S_{2\alpha}^n | U_{2\beta}^n) \quad (162)$$

$$I(X_1^n; Y_{1\alpha}^n, V_{19\alpha}^n) \triangleq I(X_1^n; Y_{1\alpha}^n) = H(Y_{1\alpha}^n) - H(S_{2\alpha}^n) \quad (163)$$

$$I(X_2^n; Y_{2\beta}^n, V_{21\beta}^n) \triangleq I(X_2^n; Y_{2\beta}^n, U_{2\alpha}^n, U_{1\beta}^n) = H(Y_{2\beta}^n | U_{2\alpha}^n, U_{1\beta}^n) - H(S_{1\beta}^n | U_{1\beta}^n) + H(U_{2\alpha}^n) - H(U_{2\alpha}^n | X_2^n) \quad (164)$$

$$I(X_2^n; Y_{2\beta}^n, V_{22\beta}^n) \triangleq I(X_2^n; Y_{2\beta}^n, U_{2\alpha}^n) = H(Y_{2\beta}^n | U_{2\alpha}^n) - H(S_{1\beta}^n) + H(U_{2\alpha}^n) - H(U_{2\alpha}^n | X_2^n) \quad (165)$$

$$I(X_2^n; Y_{2\beta}^n, V_{23\beta}^n) \triangleq I(X_2^n; Y_{2\beta}^n, U_{2\beta}^n, U_{1\beta}^n) = H(Y_{2\beta}^n | U_{2\beta}^n, U_{1\beta}^n) - H(S_{1\beta}^n | U_{1\beta}^n) + H(U_{2\beta}^n) - H(U_{2\beta}^n | X_2^n) \quad (166)$$

$$I(X_2^n; Y_{2\beta}^n, V_{24\beta}^n) \triangleq I(X_2^n; Y_{2\beta}^n, U_{2\beta}^n) = H(Y_{2\beta}^n | U_{2\beta}^n) - H(S_{1\beta}^n) + H(U_{2\beta}^n) - H(U_{2\beta}^n | X_2^n) \quad (167)$$

$$I(X_2^n; Y_{2\beta}^n, V_{25\beta}^n) \triangleq I(X_2^n; Y_{2\beta}^n, U_{1\beta}^n) = H(Y_{2\beta}^n | U_{1\beta}^n) - H(S_{1\beta}^n | U_{1\beta}^n) \quad (168)$$

$$I(X_2^n; Y_{2\beta}^n, V_{26\beta}^n) \triangleq I(X_2^n; Y_{2\beta}^n) = H(Y_{2\beta}^n) - H(S_{1\beta}^n) \quad (169)$$

$$I(X_2^n; Y_{2\alpha}^n, V_{21\alpha}^n) \triangleq I(X_2^n; Y_{2\alpha}^n, U_{2\alpha}^n, U_{1\alpha}^n) = H(Y_{2\alpha}^n | U_{2\alpha}^n, U_{1\alpha}^n) - H(S_{1\alpha}^n | U_{1\alpha}^n) + H(U_{2\alpha}^n) - H(U_{2\alpha}^n | X_2^n) \quad (170)$$

$$I(X_2^n; Y_{2\alpha}^n, V_{22\alpha}^n) \triangleq I(X_2^n; Y_{2\alpha}^n, U_{2\alpha}^n, U_{1\beta}^n) = H(Y_{2\alpha}^n | U_{2\alpha}^n, U_{1\beta}^n) - H(S_{1\alpha}^n | U_{1\beta}^n) + H(U_{2\alpha}^n) - H(U_{2\alpha}^n | X_2^n) \quad (171)$$

$$I(X_2^n; Y_{2\alpha}^n, V_{23\alpha}^n) \triangleq I(X_2^n; Y_{2\alpha}^n, U_{2\alpha}^n) = H(Y_{2\alpha}^n | U_{2\alpha}^n) - H(S_{1\alpha}^n) + H(U_{2\alpha}^n) - H(U_{2\alpha}^n | X_2^n) \quad (172)$$

$$\begin{aligned}
 I(X_2^n; Y_{2\alpha}^n, V_{24\alpha}^n) &\triangleq I(X_2^n; Y_{2\alpha}^n, U_{2\beta}^n, U_{1\alpha}^n) = H(Y_{2\alpha}^n | U_{2\beta}^n, U_{1\alpha}^n) - H(S_{1\alpha}^n | U_{1\alpha}^n) \\
 &\quad + H(U_{2\beta}^n) - H(U_{2\beta}^n | X_2^n) \quad (173)
 \end{aligned}$$

$$\begin{aligned}
 I(X_2^n; Y_{2\alpha}^n, V_{25\alpha}^n) &\triangleq I(X_2^n; Y_{2\alpha}^n, U_{2\beta}^n, U_{1\beta}^n) = H(Y_{2\alpha}^n | U_{2\beta}^n, U_{1\beta}^n) - H(S_{1\alpha}^n | U_{1\beta}^n) \\
 &\quad + H(U_{2\beta}^n) - H(U_{2\beta}^n | X_2^n) \quad (174)
 \end{aligned}$$

$$\begin{aligned}
 I(X_2^n; Y_{2\alpha}^n, V_{26\alpha}^n) &\triangleq I(X_2^n; Y_{2\alpha}^n, U_{2\beta}^n) = H(Y_{2\alpha}^n | U_{2\beta}^n) - H(S_{1\alpha}^n) \\
 &\quad + H(U_{2\beta}^n) - H(U_{2\beta}^n | X_2^n) \quad (175)
 \end{aligned}$$

$$I(X_2^n; Y_{2\alpha}^n, V_{27\alpha}^n) \triangleq I(X_2^n; Y_{2\alpha}^n, U_{1\alpha}^n) = H(Y_{2\alpha}^n | U_{1\alpha}^n) - H(S_{1\alpha}^n | U_{1\alpha}^n) \quad (176)$$

$$I(X_2^n; Y_{2\alpha}^n, V_{28\alpha}^n) \triangleq I(X_2^n; Y_{2\alpha}^n, U_{1\beta}^n) = H(Y_{2\alpha}^n | U_{1\beta}^n) - H(S_{1\alpha}^n | U_{1\beta}^n) \quad (177)$$

$$I(X_2^n; Y_{2\alpha}^n, V_{29\alpha}^n) \triangleq I(X_2^n; Y_{2\alpha}^n) = H(Y_{2\alpha}^n) - H(S_{1\alpha}^n). \quad (178)$$

Continuing with our outer bound derivation, from (146)

$$n(aR_1 + bR_2 - (a + b)\epsilon_n)$$

$$\begin{aligned}
 &\stackrel{(a)}{\leq} \sum_{i=1}^6 \nu_{1i} H(Y_{1\beta}^n | V_{1i\beta}^n) + \sum_{i=1}^9 \mu_{1i} H(Y_{1\alpha}^n | V_{1i\alpha}^n) \\
 &\quad + H(U_{1\alpha}^n) \{(\nu_{11} + \nu_{12}) + (\mu_{11} + \mu_{12} + \mu_{13})\} \\
 &\quad + H(U_{1\beta}^n) \{(\nu_{13} + \nu_{14}) + (\mu_{14} + \mu_{15} + \mu_{16})\} \\
 &\quad - H(S_{1\alpha}^n) \{(\mu_{23} + \mu_{26} + \mu_{29})\} \\
 &\quad - H(S_{1\beta}^n) \{(\nu_{22} + \nu_{24} + \nu_{26})\} \\
 &\quad - H(U_{1\alpha}^n | X_1^n) \{(\nu_{11} + \nu_{12}) + (\mu_{11} + \mu_{12} + \mu_{13})\} \\
 &\quad - H(U_{1\beta}^n | X_1^n) \{(\nu_{13} + \nu_{14}) + (\mu_{14} + \mu_{15} + \mu_{16})\} \\
 &\quad - H(S_{1\alpha}^n | U_{1\alpha}^n) \{(\mu_{21} + \mu_{24} + \mu_{27})\} \\
 &\quad - H(S_{1\beta}^n | U_{1\beta}^n) \{(\nu_{21} + \nu_{23} + \nu_{25})\} \\
 &\quad - H(S_{1\alpha}^n | U_{1\beta}^n) \{(\mu_{22} + \mu_{25} + \mu_{28})\} \\
 &\quad + \sum_{i=1}^6 \nu_{2i} H(Y_{2\beta}^n | V_{2i\beta}^n) + \sum_{i=1}^9 \mu_{2i} H(Y_{2\alpha}^n | V_{2i\alpha}^n) \\
 &\quad + H(U_{2\alpha}^n) \{(\nu_{21} + \nu_{22}) + (\mu_{21} + \mu_{22} + \mu_{23})\} \\
 &\quad + H(U_{2\beta}^n) \{(\nu_{23} + \nu_{24}) + (\mu_{24} + \mu_{25} + \mu_{26})\} \\
 &\quad - H(S_{2\alpha}^n) \{(\mu_{13} + \mu_{16} + \mu_{19})\} \\
 &\quad - H(S_{2\beta}^n) \{(\nu_{12} + \nu_{14} + \nu_{16})\} \\
 &\quad - H(U_{2\alpha}^n | X_2^n) \{(\nu_{21} + \nu_{22}) + (\mu_{21} + \mu_{22} + \mu_{23})\} \\
 &\quad - H(U_{2\beta}^n | X_2^n) \{(\nu_{23} + \nu_{24}) + (\mu_{24} + \mu_{25} + \mu_{26})\} \\
 &\quad - H(S_{2\alpha}^n | U_{2\alpha}^n) \{(\mu_{11} + \mu_{14} + \mu_{17})\} \\
 &\quad - H(S_{2\beta}^n | U_{2\beta}^n) \{(\nu_{11} + \nu_{13} + \nu_{15})\} \\
 &\quad - H(S_{2\alpha}^n | U_{2\beta}^n) \{(\mu_{12} + \mu_{15} + \mu_{18})\} \quad (179)
 \end{aligned}$$

$$\begin{aligned}
 &\stackrel{(b)}{\leq} \sum_{i=1}^6 \nu_{1i} H(Y_{1\beta}^n | V_{1i\beta}^n) + \sum_{i=1}^9 \mu_{1i} H(Y_{1\alpha}^n | V_{1i\alpha}^n) \\
 &\quad + H(U_{1\alpha}^n) \{(\mu_{22} + \mu_{25} + \mu_{28})\} \\
 &\quad - H(U_{1\beta}^n) \{(\mu_{22} + \mu_{25} + \mu_{28})\} \\
 &\quad - H(S_{1\alpha}^n | U_{1\beta}^n) \{(\mu_{22} + \mu_{25} + \mu_{28})\} \\
 &\quad - H(U_{1\alpha}^n | X_1^n) \{(\mu_{22} + \mu_{25} + \mu_{28})\} \\
 &\quad + (\mu_{23} + \mu_{26} + \mu_{29}) \\
 &\quad - H(U_{1\beta}^n | X_1^n) \{(\nu_{22} + \nu_{24} + \nu_{26})
 \end{aligned}$$

$$\begin{aligned}
 &\quad - (\mu_{22} + \mu_{25} + \mu_{28})\} \\
 &\quad - H(S_{1\alpha}^n | U_{1\alpha}^n) \{(\mu_{21} + \mu_{24} + \mu_{27})\} \\
 &\quad - H(S_{1\beta}^n | U_{1\beta}^n) \{(\nu_{21} + \nu_{23} + \nu_{25})\} \\
 &\quad + \sum_{i=1}^6 \nu_{2i} H(Y_{2\beta}^n | V_{2i\beta}^n) + \sum_{i=1}^9 \mu_{2i} H(Y_{2\alpha}^n | V_{2i\alpha}^n) \\
 &\quad + H(U_{2\alpha}^n) \{(\mu_{12} + \mu_{15} + \mu_{18})\} \\
 &\quad - H(U_{2\beta}^n) \{(\mu_{12} + \mu_{15} + \mu_{18})\} \\
 &\quad - H(S_{2\alpha}^n | U_{2\beta}^n) \{(\mu_{12} + \mu_{15} + \mu_{18})\} \\
 &\quad - H(U_{2\alpha}^n | X_2^n) \{(\mu_{12} + \mu_{15} + \mu_{18}) + (\mu_{13} + \mu_{16} + \mu_{19})\} \\
 &\quad - H(U_{2\beta}^n | X_2^n) \{(\nu_{12} + \nu_{14} + \nu_{16}) - (\mu_{12} + \mu_{15} + \mu_{18})\} \\
 &\quad - H(S_{2\alpha}^n | U_{2\alpha}^n) \{(\mu_{11} + \mu_{14} + \mu_{17})\} \\
 &\quad - H(S_{2\beta}^n | U_{2\beta}^n) \{(\nu_{11} + \nu_{13} + \nu_{15})\} \quad (180)
 \end{aligned}$$

$$\begin{aligned}
 &\stackrel{(c)}{\leq} \sum_{i=1}^6 \nu_{1i} H(Y_{1\beta}^n | V_{1i\beta}^n) + \sum_{i=1}^9 \mu_{1i} H(Y_{1\alpha}^n | V_{1i\alpha}^n) \\
 &\quad - H(U_{1\alpha}^n | X_1^n) \{(\mu_{22} + \mu_{25} + \mu_{28}) + (\mu_{23} + \mu_{26} + \mu_{29})\} \\
 &\quad - H(U_{1\beta}^n | X_1^n) \{(\nu_{22} + \nu_{24} + \nu_{26})\} \\
 &\quad - H(S_{1\alpha}^n | X_1^n) \{(\mu_{21} + \mu_{24} + \mu_{27})\} \\
 &\quad - H(S_{1\beta}^n | X_1^n) \{(\nu_{21} + \nu_{23} + \nu_{25})\} \\
 &\quad + \sum_{i=1}^6 \nu_{2i} H(Y_{2\beta}^n | V_{2i\beta}^n) + \sum_{i=1}^9 \mu_{2i} H(Y_{2\alpha}^n | V_{2i\alpha}^n) \\
 &\quad - H(U_{2\alpha}^n | X_2^n) \{(\mu_{12} + \mu_{15} + \mu_{18}) + (\mu_{13} + \mu_{16} + \mu_{19})\} \\
 &\quad - H(U_{2\beta}^n | X_2^n) \{(\nu_{12} + \nu_{14} + \nu_{16})\} \\
 &\quad - H(S_{2\alpha}^n | X_2^n) \{(\mu_{11} + \mu_{14} + \mu_{17})\} \\
 &\quad - H(S_{2\beta}^n | X_2^n) \{(\nu_{11} + \nu_{13} + \nu_{15})\} \quad (181)
 \end{aligned}$$

$$\begin{aligned}
 &\stackrel{(d)}{=} \sum_{i=1}^6 \nu_{1i} H(Y_{1\beta}^n | V_{1i\beta}^n) + \sum_{i=1}^9 \mu_{1i} H(Y_{1\alpha}^n | V_{1i\alpha}^n) \\
 &\quad - H(S_{1\alpha}^n | X_1^n) \left(\sum_{i=1}^9 \mu_{2i} \right) \\
 &\quad - H(S_{1\beta}^n | X_1^n) \left(\sum_{i=1}^6 \nu_{2i} \right) \\
 &\quad + \sum_{i=1}^6 \nu_{2i} H(Y_{2\beta}^n | V_{2i\beta}^n) + \sum_{i=1}^9 \mu_{2i} H(Y_{2\alpha}^n | V_{2i\alpha}^n)
 \end{aligned}$$

$$\begin{aligned}
& -H(S_{2\alpha}^n | X_2^n) \left(\sum_{i=1}^9 \mu_{1i} \right) \\
& -H(S_{2\beta}^n | X_2^n) \left(\sum_{i=1}^6 \nu_{1i} \right). \quad (182)
\end{aligned}$$

Here, we have the following.

- To get inequality (a), we used (149)–(178) in (146) and collected the terms together.
- For equality (b), we used Proposition 1 along with the fact that $U_{k\beta}^n$ and $U_{k\alpha}^n$ has the same marginal distribution as $S_{k\beta}^n$ and $S_{k\alpha}^n$, respectively. Therefore

$$H(U_{1\alpha}^n) = H(S_{1\alpha}^n) \quad (183)$$

$$H(U_{1\beta}^n) = H(S_{1\beta}^n) \quad (184)$$

$$H(U_{2\alpha}^n) = H(S_{2\alpha}^n) \quad (185)$$

$$H(U_{2\beta}^n) = H(S_{2\beta}^n). \quad (186)$$

- Inequality (c) follows from the fact that conditioning reduces entropy. In particular,

$$\begin{aligned}
& H(U_{1\alpha}^n) - H(U_{1\beta}^n) - H(S_{1\alpha}^n | U_{1\beta}^n) \\
& = -H(U_{1\beta}^n | S_{1\alpha}^n) \leq -H(U_{1\beta}^n | X_1^n) \quad (187)
\end{aligned}$$

$$\begin{aligned}
& H(U_{2\alpha}^n) - H(U_{2\beta}^n) - H(S_{2\alpha}^n | U_{2\beta}^n) \\
& = -H(U_{2\beta}^n | S_{2\alpha}^n) \leq -H(U_{2\beta}^n | X_2^n). \quad (188)
\end{aligned}$$

- For equality (d), we used

$$\begin{aligned}
& H(U_{1\alpha}^n | X_1^n) = H(S_{1\alpha}^n | X_1^n) \\
& H(U_{1\beta}^n | X_1^n) = H(S_{1\beta}^n | X_1^n) \\
& H(U_{2\alpha}^n | X_2^n) = H(S_{2\alpha}^n | X_2^n) \\
& H(U_{2\beta}^n | X_2^n) = H(S_{2\beta}^n | X_2^n).
\end{aligned}$$

Now, we “single-letterize” using the chain rule along with the fact that the channel is memoryless and conditioning reduces entropy

$$\begin{aligned}
& aR_1 + bR_2 - (a+b)\epsilon_n \\
& \leq \frac{1}{n} \sum_{q=1}^n \left\{ \sum_{i=1}^6 \nu_{1i} H(Y_{1\beta}(q) | V_{1i\beta}(q)) \right. \\
& \quad + \sum_{i=1}^9 \mu_{1i} H(Y_{1\alpha}(q) | V_{1i\alpha}(q)) \\
& \quad - H(S_{1\alpha}(q) | X_1(q)) \left(\sum_{i=1}^9 \mu_{2i} \right) \\
& \quad \left. - H(S_{1\beta}(q) | X_1(q)) \left(\sum_{i=1}^6 \nu_{2i} \right) \right. \\
& \quad + \sum_{i=1}^6 \nu_{2i} H(Y_{2\beta}(q) | V_{2i\beta}(q)) \\
& \quad \left. + \sum_{i=1}^9 \mu_{2i} H(Y_{2\alpha}(q) | V_{2i\alpha}(q)) \right\}
\end{aligned}$$

$$\begin{aligned}
& -H(S_{2\alpha}(q) | X_2(q)) \left(\sum_{i=1}^9 \mu_{1i} \right) \\
& -H(S_{2\beta}(q) | X_2(q)) \left(\sum_{i=1}^6 \nu_{1i} \right) \}. \quad (189)
\end{aligned}$$

$$\begin{aligned}
& aR_1 + bR_2 - (a+b)\epsilon_n \\
& \leq \sum_{i=1}^6 \nu_{1i} H(Y_{1\beta} | V_{1i\beta}, Q) \\
& \quad + \sum_{i=1}^9 \mu_{1i} H(Y_{1\alpha} | V_{1i\alpha}, Q) \\
& \quad - \left(\sum_{i=1}^9 \mu_{2i} \right) H(S_{1\alpha} | X_1, Q) \\
& \quad - \left(\sum_{i=1}^6 \nu_{2i} \right) H(S_{1\beta} | X_1, Q) \\
& \quad + \sum_{i=1}^6 \nu_{2i} H(Y_{2\beta} | V_{2i\beta}, Q) \\
& \quad + \sum_{i=1}^9 \mu_{2i} H(Y_{2\alpha} | V_{2i\alpha}, Q) \\
& \quad - \left(\sum_{i=1}^9 \mu_{1i} \right) H(S_{2\alpha} | X_2, Q) \\
& \quad - \left(\sum_{i=1}^6 \nu_{1i} \right) H(S_{2\beta} | X_2, Q) \quad (190)
\end{aligned}$$

where we set $(Q, U_{1\beta}, U_{1\alpha}, X_1, S_{1\beta}, S_{1\alpha}, U_{2\beta}, U_{2\alpha}, X_2, S_{2\beta}, S_{2\alpha})$ to be joint random variables such that Q is uniformly distributed over $\{1, 2, \dots, n\}$ and,

$$\begin{aligned}
& \Pr(U_{1\beta}, U_{1\alpha}, X_1, S_{1\alpha}, S_{1\beta}, U_{2\beta}, U_{2\alpha}, X_2, S_{2\alpha}, S_{2\beta} | Q = q) \\
& = \Pr(U_{1\beta}(q), U_{1\alpha}(q), X_1(q), S_{1\alpha}(q), S_{1\beta}(q) \\
& \quad U_{2\beta}(q), U_{2\alpha}(q), X_2(q), S_{2\alpha}(q), S_{2\beta}(q)) \quad (191)
\end{aligned}$$

for $1 \leq q \leq n$. Since the messages are independent for the two users, so are $X_1(q)$ and $X_2(q)$. Therefore, (Q, X_1, X_2) satisfies the Markov chain $X_1 - Q - X_2$. Further because of our choice of $(U_{1\beta}(q), U_{1\alpha}(q), U_{1\beta}(q), U_{1\alpha}(q))$, the random variables satisfy the condition (58). Hence the random variables $(Q, X_1, U_{1\alpha}, U_{1\beta}, X_2, U_{2\alpha}, U_{2\beta})$ belong to the sub-family of \mathcal{P} that we described earlier, whose elements are defined by (Q, X_1, X_2) .

The above step is done for all $(\lambda, a, b) \in \Gamma$. Since the cardinality of Γ is finite, we can use Carathéodory's theorem along the lines of Remark 1 to bound the cardinality of Q and make it independent of n . Taking $n \rightarrow \infty$, we get $\forall (\lambda, a, b) \in \Gamma$

$$aR_1 + bR_2 \leq c_{(\lambda, a, b)}^{(\text{out})}(Q, X_1, X_2) \quad (192)$$

where

$$c_{(\lambda, a, b)}^{(\text{out})}(Q, X_1, X_2) \triangleq \sum_{i=1}^6 \nu_{1i} H(Y_{1\beta} | V_{1i\beta}, Q)$$

$$\begin{aligned}
 & + \sum_{i=1}^9 \mu_{1i} H(Y_{1\alpha} | V_{1i\alpha}, Q) \\
 & - \left(\sum_{i=1}^9 \mu_{2i} \right) H(S_{1\alpha} | X_1, Q) \\
 & - \left(\sum_{i=1}^6 \nu_{2i} \right) H(S_{1\beta} | X_1, Q) \\
 & + \sum_{i=1}^6 \nu_{2i} H(Y_{2\beta} | V_{2i\beta}, Q) \\
 & + \sum_{i=1}^9 \mu_{2i} H(Y_{2\alpha} | V_{2i\alpha}, Q) \\
 & - \left(\sum_{i=1}^9 \mu_{1i} \right) H(S_{2\alpha} | X_2, Q) \\
 & - \left(\sum_{i=1}^6 \nu_{1i} \right) H(S_{2\beta} | X_2, Q).
 \end{aligned} \tag{193}$$

We are now ready to formally define $\mathcal{R}_{\text{out}}(Q, X_1, X_2)$

$$\mathcal{R}_{\text{out}}(Q, X_1, X_2) \triangleq \{(R_1, R_2) : R_1 \geq 0, R_2 \geq 0, aR_1 + bR_2 \leq c_{(\lambda, a, b)}^{(\text{out})}(Q, X_1, X_2), \forall (\lambda, a, b) \in \Gamma\}.$$

We have proved that if (R_1, R_2) is achievable, then

$$(R_1, R_2) \in \bigcup_{Q, X_1, X_2} \mathcal{R}_{\text{out}}(Q, X_1, X_2). \tag{194}$$

This completes the proof. \square

B. Proof of Theorem 3 (ii)

For a given (Q, X_1, X_2) such that $X_1 - Q - X_2$ is a Markov chain, we need to quantify the gap between $\mathcal{R}_{\text{out}}(Q, X_1, X_2)$ and $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$, which are defined by (130) and (143), respectively. In order to do this, we quantify the gap between $c_{(\lambda, a, b)}^{(\text{out})}(Q, X_1, X_2)$ and $c_{(\lambda, a, b)}^{(\text{in})}(Q, X_1, X_2)$.

$$\begin{aligned}
 & c_{(\lambda, a, b)}^{(\text{out})}(Q, X_1, X_2) - c_{(\lambda, a, b)}^{(\text{in})}(Q, X_1, X_2) \\
 & = (H(S_{1\alpha} | U_{1\alpha}, Q) - H(S_{1\alpha} | X_1, Q)) \left(\sum_{i=1}^9 \mu_{2i} \right) \\
 & + (H(S_{1\beta} | U_{1\beta}, Q) - H(S_{1\beta} | X_1, Q)) \left(\sum_{i=1}^6 \nu_{2i} \right) \\
 & + (H(S_{2\alpha} | U_{2\alpha}, Q) - H(S_{2\alpha} | X_2, Q)) \left(\sum_{i=1}^9 \mu_{1i} \right) \\
 & + (H(S_{2\beta} | U_{2\beta}, Q) - H(S_{2\beta} | X_2, Q)) \left(\sum_{i=1}^6 \nu_{1i} \right) \\
 & = I(S_{1\alpha}; X_1 | U_{1\alpha}, Q) \left(\sum_{i=1}^9 \mu_{2i} \right) \\
 & + I(S_{1\beta}; X_1 | U_{1\beta}, Q) \left(\sum_{i=1}^6 \nu_{2i} \right)
 \end{aligned}$$

$$\begin{aligned}
 & + I(S_{2\alpha}; X_2 | U_{2\alpha}, Q) \left(\sum_{i=1}^9 \mu_{1i} \right) \\
 & + I(S_{2\beta}; X_2 | U_{2\beta}, Q) \left(\sum_{i=1}^6 \nu_{1i} \right) \\
 & \leq b \max(I(S_{1\alpha}; X_1 | U_{1\alpha}, Q), I(S_{1\beta}; X_1 | U_{1\beta}, Q)) \\
 & + a \max(I(S_{2\alpha}; X_2 | U_{2\alpha}, Q), I(S_{2\beta}; X_2 | U_{2\beta}, Q)) \\
 & \leq a\Delta_1(Q, X_1, X_2) + b\Delta_2(Q, X_1, X_2).
 \end{aligned} \tag{195}$$

Here $\Delta_1(Q, X_1, X_2)$ and $\Delta_2(Q, X_1, X_2)$ are defined as follows:

$$\begin{aligned}
 \Delta_1(Q, X_1, X_2) & \triangleq \max(I(S_{2\alpha}; X_2 | U_{2\alpha}), I(S_{2\beta}; X_2 | U_{2\beta})), \\
 \Delta_2(Q, X_1, X_2) & \triangleq \max(I(S_{1\alpha}; X_1 | U_{1\alpha}), I(S_{1\beta}; X_1 | U_{1\beta})).
 \end{aligned}$$

This completes the proof of Theorem 3. \square

C. Proof Of Corollary 6

Consider the two-state compound Gaussian interference channel. For this special case, we have the following result that identifies the Gaussian code books to be sufficient.

Lemma 9:

$$\mathcal{R}_{\text{out}}(Q, X_1, X_2) \subseteq \mathcal{R}_{\text{out}}(Q^*, X_1^*, X_2^*) \tag{196}$$

where $Q^* = 1$, $X_1^* \sim \mathcal{CN}(0, P_1)$, $X_2^* \sim \mathcal{CN}(0, P_2)$.

We note for easy reference that $\mathcal{R}_{\text{out}}(Q, X_1, X_2)$ is defined in (143).

Proof: It suffices to show that

$$c_{(\lambda, a, b)}^{(\text{out})}(Q, X_1, X_2) \leq c_{(\lambda, a, b)}^{(\text{out})}(Q^*, X_1^*, X_2^*),$$

where $c_{(\lambda, a, b)}^{(\text{out})}(Q, X_1, X_2)$ is as defined in (193).

$$\begin{aligned}
 c_{(\lambda, a, b)}^{(\text{out})}(Q, X_1, X_2) & = \sum_{i=1}^6 \nu_{1i} h(Y_{1\beta} | V_{1i\beta}, Q) \\
 & + \sum_{i=1}^9 \mu_{1i} h(Y_{1\alpha} | V_{1i\alpha}, Q) \\
 & - \left(\sum_{i=1}^9 \mu_{2i} \right) h(S_{1\alpha} | X_1, Q) \\
 & - \left(\sum_{i=1}^6 \nu_{2i} \right) h(S_{1\beta} | X_1, Q) \\
 & + \sum_{i=1}^6 \nu_{2i} h(Y_{2\beta} | V_{2i\beta}, Q) \\
 & + \sum_{i=1}^9 \mu_{2i} h(Y_{2\alpha} | V_{2i\alpha}, Q) \\
 & - \left(\sum_{i=1}^9 \mu_{1i} \right) h(S_{2\alpha} | X_2, Q)
 \end{aligned}$$

$$- \left(\sum_{i=1}^6 \nu_{1i} \right) h(S_{2\beta} | X_2, Q)$$

The terms $h(S_{1\alpha} | X_1, Q)$, $h(S_{1\beta} | X_1, Q)$, $h(S_{2\alpha} | X_2, Q)$ and $h(S_{2\beta} | X_2, Q)$ are the differential entropies of complex Gaussian noise with known variance and are readily handled. Let us now turn to the term $h(Y_{1\beta} | V_{11\beta}, Q)$:

$$h(Y_{1\beta} | V_{11\beta}, Q) = h(Y_{1\beta} | U_{1\alpha} U_{2\beta}, Q) \quad (197)$$

$$= \sum_q p(q) h(Y_{1\beta} | U_{1\alpha} U_{2\beta}, Q = q) \quad (198)$$

$$\stackrel{(a)}{\leq} \sum_q p(q) h(Y_{1\beta q}^* | U_{1\alpha q}^* U_{2\beta q}^*) \quad (199)$$

$$= \sum_q p(q) h(h_{11\beta} X_{1q}^* + h_{21\beta} X_{2q}^* + Z_1 | h_{12\alpha} X_{1q}^* + Z_2, h_{12\beta} X_{2q}^* + Z_1') \quad (200)$$

$$= \sum_q p(q) \log \left(\frac{|h_{11\beta}|^2 P_{1q}}{|h_{12\alpha}|^2 P_{1q} + 1} + \frac{|h_{21\beta}|^2 P_{2q}}{|h_{12\beta}|^2 P_{2q} + 1} + 1 \right) \quad (201)$$

$$\stackrel{(b)}{\leq} \log \left(\frac{|h_{11\beta}|^2 P_1}{|h_{12\alpha}|^2 P_1 + 1} + \frac{|h_{21\beta}|^2 P_2}{|h_{12\beta}|^2 P_2 + 1} + 1 \right) \quad (202)$$

$$= h(Y_{1\beta}^* | U_{1\alpha}^* U_{2\beta}^*, Q^*). \quad (203)$$

Here,

- in step (a), we denoted

$$\mathbb{E}[|X_k|^2 | Q = q] = P_{kq}, \quad k = 1, 2 \quad (204)$$

and used the fact that conditional differential entropy is maximized with the Gaussian distribution under a covariance constraint (Lemma 1 [7]). The * denotes jointly Gaussian random variables with the same covariance constraints;

- in step (b), we used Jensen's inequality.

A similar argument follows for the other terms. To conclude, we have shown that

$$c_{(\lambda, a, b)}^{(\text{out})}(Q, X_1, X_2) \leq c_{(\lambda, a, b)}^{(\text{out})}(Q^*, X_1^*, X_2^*). \quad (205)$$

This completes the proof. \square

Finally, we can readily see the proof of Corollary 6. This is because

$$\bigcup_{(Q, X_1, X_2)} \mathcal{R}_{\text{out}}(Q, X_1, X_2) = \mathcal{R}_{\text{out}}(Q^*, X_1^*, X_2^*) \quad (206)$$

as a direct consequence of Lemma 9.

VI. DISCUSSION: INSIGHTS ON THE NONCOMPOUND INTERFERENCE CHANNEL

In this section we consider the noncompound interference channel model introduced in [5]; this is a specific instance of our model and is obtained by setting $\alpha = \beta$. Our results, when specialized to this instance provide an alternative derivation of the results of Chong *et al.* [3] and Tse and Telatar [5]. Below we briefly sketch our results with an aim to compare and contrast the different proofs. The goal is not only to give better insight into existing results, but also to give an idea on how our new proof technique scales more naturally to the 2-state compound interference channel (and in general to the n -state compound interference channel that we will describe in the next section). We first describe the achievable scheme and the inner bound. Following that, we will describe the outer-bound, focusing on contrasts between the different approaches.

A. Achievable Scheme

The special case of the noncompound version is obtained by setting

$$S_{k\beta} = S_{k\alpha} = S_k \quad (207)$$

and, correspondingly,

$$U_{k\beta} = U_{k\alpha} = U_k \quad (208)$$

for $k = 1, 2$. We also set

$$R_{k\alpha} = 0. \quad (209)$$

We rename $R_{k\beta}$ as T_k and R_{kp} as S_k to be consistent with the notation of Chong *et al.* [3].

The superposition achievable scheme can now be described by joint random variables

$$P = (Q, U_1, X_1, U_2, X_2) \quad (210)$$

with the joint distribution factoring as

$$p(q)p(x_1 | q)p(x_2 | q)p(u_1 | x_1 q)p(u_2 | x_2 q). \quad (211)$$

From Section IV-A, it follows that any rate vector (S_1, T_1, S_2, T_2) that satisfies

$$S_1 \leq I(Y_1; X_1 | U_1, U_2, Q), \quad \text{if } S_1 > 0 \quad (212)$$

$$T_2 + S_1 \leq I(Y_1; X_1 U_2 | U_1, Q), \quad \text{if } S_1 > 0 \quad (213)$$

$$T_1 + S_1 \leq I(Y_1; X_1 | U_2, Q), \quad \text{if } T_1 + S_1 > 0 \quad (214)$$

$$T_2 + T_1 + S_1 \leq I(Y_1; X_1 U_2 | Q), \quad \text{if } T_1 + S_1 > 0 \quad (215)$$

$$S_2 \leq I(Y_2; X_2 | U_2, U_1, Q), \quad \text{if } S_2 > 0 \quad (216)$$

$$T_1 + S_2 \leq I(Y_2; X_2, U_1 | U_2, Q), \quad \text{if } S_2 > 0 \quad (217)$$

$$T_2 + S_2 \leq I(Y_2; X_2 | U_1, Q), \quad \text{if } T_2 + S_2 > 0 \quad (218)$$

$$T_1 + T_2 + S_2 \leq I(Y_2; X_2, U_1 | Q), \quad \text{if } T_2 + S_2 > 0 \quad (219)$$

$$S_1, T_1, S_2, T_2 \geq 0 \quad (220)$$

is achievable. Define the four-dimensional region

$$\tilde{\mathcal{R}}_{\text{in}}^{(4)}(P) \triangleq \{(S_1, T_1, S_2, T_2) : \text{satisfies (212)–(220)}\} \quad (221)$$

and its projection on the two-dimensional space

$$\tilde{\mathcal{R}}_{\text{in}}(P) \triangleq \{(R_1, R_2) : R_k = S_k + T_k, k = 1, 2\}. \quad (222)$$

On the other hand, define

$$\mathcal{R}_{\text{in}}^{(4)}(P) \triangleq \{(S_1, T_1, S_2, T_2) : \text{satisfies (224)–(233)}\}. \quad (223)$$

$$S_1 \leq I(Y_1; X_1 | U_1, U_2, Q) \quad (224)$$

$$T_2 + S_1 \leq I(Y_1; X_1 U_2 | U_1, Q) \quad (225)$$

$$T_1 + S_1 \leq I(Y_1; X_1 | U_2, Q) \quad (226)$$

$$T_2 + T_1 + S_1 \leq I(Y_1; X_1 U_2 | Q) \quad (227)$$

$$S_2 \leq I(Y_2; X_2 | U_2, U_1, Q) \quad (228)$$

$$T_1 + S_2 \leq I(Y_2; X_2, U_1 | U_2, Q) \quad (229)$$

$$T_2 + S_2 \leq I(Y_2; X_2 | U_1, Q) \quad (230)$$

$$T_1 + T_2 + S_2 \leq I(Y_2; X_2, U_1 | Q) \quad (231)$$

$$S_1 + T_1 \geq 0 \quad (232)$$

$$S_2 + T_2 \geq 0. \quad (233)$$

Let its projection on the two-dimensional space

$$\mathcal{R}_{\text{in}}(P) \triangleq \{(R_1, R_2) : R_k = S_k + T_k, k = 1, 2\}. \quad (234)$$

In [3, Th. 2], the authors explicitly evaluated the constraints that define this set and described it as the “compact version” of the Han–Kobayashi region [1] (which results from a somewhat different coding strategy, as compared to the superposition coding one). However, we know from Lemma 7 that

$$\mathcal{R}_{\text{in}}(P) \subseteq \tilde{\mathcal{R}}_{\text{in}}(P). \quad (235)$$

We thus conclude the alternate proof of [3](Theorem 2). Our approach differs from the approach of Chong *et al.* [3] in two ways.

- It is instructive to observe the similarities and differences between the 4-dimensional achievable region $\tilde{\mathcal{R}}_{\text{in}}^{(4)}(P)$ to the one in [3] (Lemma 3). First, the inequalities involved are the same. However, several of these constraints are inactive when the boundary conditions on the data rates bite. We can immediately conclude that our achievable region ($\tilde{\mathcal{R}}_{\text{in}}^{(4)}(P)$) is in general a superset of the region in [3]. This is somewhat surprising since the encoding method in both cases is superposition coding. The differences result due to our careful consideration of the error events in the decoding process.
- Chong *et al.* [3] described the two-dimensional region explicitly by carrying out the somewhat tedious algorithmic procedure of Fourier–Motzkin elimination. Further, they showed that a potentially bigger region (the compact description region) is achievable by time-sharing between two other schemes defined by $(Q, \emptyset, X_1, U_2, X_2)$ and $(Q, U_1, X_1, \emptyset, X_2)$. In our approach, we entirely avoid describing the two-dimensional region explicitly. Further, we showed that there is no need to time-share between any other schemes, to achieve $\mathcal{R}_{\text{in}}(P)$.

B. Outer Bound

For a given $P = (Q, U_1, X_1, U_2, X_2)$, the inner-bound region in Chong *et al.* [3] is described by seven linear inequalities involving R_1 and R_2 . In [5], Telatar and Tse picked a specific choice of (U_1, U_2) given by

$$p(u_1, u_2 | q, x_1, x_2) = p_{S_1 | X_1}(u_1 | x_1) p_{S_2 | X_2}(u_2 | x_2). \quad (236)$$

In deriving the outer bound, Telatar and Tse [5] gave extra information to the receivers (the so-called “genie-aided” approach) to handle the seven inequalities. The rationale to what side information the genie should provide to handle the different linear inequalities was somewhat speculative (cf. Section IV [5]).

Our approach avoids an explicit representation of the inner-bound. This higher level description allowed us (cf. Section IV-B) to show that any inequality involved in the projected region can be obtained by linear combination of the inequalities (224)–(231). Further, each inequality in (224)–(231) arises from a *typical error event* consideration. We now have the operational insight into what side information to give when. We demonstrate this process in the instance of (224). This inequality must be satisfied to ensure that the Receiver 1 decodes its own private message, on the condition that it can decode both the public messages correctly. This suggests that corresponding to this inequality, we may give the side information (U_1^n, U_2^n) . A similar argument handles each of the other inequalities (224)–(231).

VII. N-STATE COMPOUND INTERFERENCE CHANNEL

In this section we consider the natural extension of the two-state compound interference channel to an N -state compound interference channel. Our earlier results (both inner and outer bounds) also generalize naturally to the more general N -state model.

A. Model

The N -state compound interference channel is depicted in Fig. 5. Each receiver can be in one of the N possible states denoted by $\alpha_1, \alpha_2, \dots, \alpha_N$.

B. Results

We can characterize the inner bound and outer bounds to the capacity region in a way similar to the two-state compound channel.

Inner Bound.

Our coding scheme is $N + 1$ -level superposition coding. This is much along the lines of the three-level superposition coding employed for the two-state compound interference channel. The coding scheme is characterized by jointly distributed random variables

$$(Q, X_1, U_{1\alpha_1}, \dots, U_{1\alpha_N}, X_2, U_{2\alpha_1}, \dots, U_{2\alpha_N}) \quad (237)$$

which satisfy the Markov chain

$$U_{1\alpha_N} - \dots - U_{1\alpha_1} - X_1 - Q - X_2 - U_{2\alpha_1} - \dots - U_{2\alpha_N}. \quad (238)$$

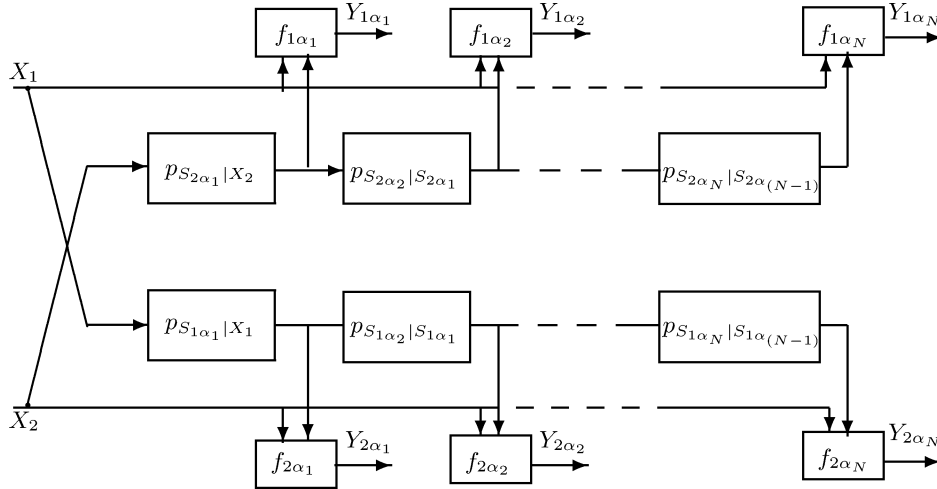


Fig. 5. The N -state compound interference channel model.

As earlier, we restrict ourselves to a subfamily of the jointly distributed random variables uniquely determined by (Q, X_1, X_2) in the following way:

Given (Q, X_1, X_2) , we pick random variables

$$\{U_{k\alpha_n}, \quad n = 1, \dots, N, \quad k = 1, 2\} \quad (239)$$

such that they have the same joint distribution as

$$\{S_{k\alpha_n}, \quad n = 1, \dots, N, \quad k = 1, 2\} \quad (240)$$

but are independent of them.

Using these random variables, we generate the $(N + 1)$ -level superposition random code books for each user with rates $(R_{1\alpha_N}, \dots, R_{1\alpha_1}, R_{1p})$ and $(R_{2\alpha_N}, \dots, R_{2\alpha_1}, R_{2p})$ respectively.

The decoding at each receiver is jointly typical set decoding. It is similar to the decoding described for the two-state. Each receiver tries to decode fully all of its own messages, but only partially decodes the other (interfering) user. This strategy can be seen as an opportunistic strategy where the extent of the interference that the receiver decodes depends upon the level of interference it sees.

The remainder description of the achievable rate region follows the same development pattern as for the two-state compound channel. It would be impractical (in terms of the length of the descriptions) to explicitly detail this description. As such, we briefly itemize the main points in the achievable region description below.

- We first have an achievable rate region $\tilde{\mathcal{R}}_{\text{in}}^{2(N+1)}(Q, X_1, X_2)$ in $2(N + 1)$ dimensions along the same lines as (111) (we have avoided the explicit description of the linear inequalities describing the region due to the tedium and length involved in doing so). As earlier, let $\tilde{\mathcal{R}}_{\text{in}}(Q, X_1, X_2)$ be the projection onto the two-dimensional space (R_1, R_2) where

$$\begin{aligned} R_1 &= R_{1\alpha_N} + \dots + R_{1\alpha_1} + R_{1p}, \\ R_2 &= R_{2\alpha_N} + \dots + R_{2\alpha_1} + R_{2p}. \end{aligned} \quad (241)$$

We have that $\tilde{\mathcal{R}}_{\text{in}}(Q, X_1, X_2)$ is achievable.

- We next define $\mathcal{R}_{\text{in}}^{2(N+1)}(Q, X_1, X_2)$ as a generalization of (23) and define its projection onto the two dimensional space $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$. Lemma 7 can be appropriately generalized to show that

$$\mathcal{R}_{\text{in}}(Q, X_1, X_2) \subseteq \tilde{\mathcal{R}}_{\text{in}}(Q, X_1, X_2) \quad (242)$$

thus proving that $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ is also achievable.

- We next characterize an external representation of $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$, using an appropriate generalization of (130) to the N -state model). In other words, we represent it as an intersection of hyperplanes, where the inequality used to define the hyperplane can be obtained as a linear combination of the inequalities used to define $\mathcal{R}_{\text{in}}^{2(N+1)}(Q, X_1, X_2)$.

Outer Bound.

An outer-bound $\mathcal{R}_{\text{out}}(Q, X_1, X_2)$ can be derived with an external representation that is similar to the corresponding one for the inner bound $\mathcal{R}_{\text{in}}(Q, X_1, X_2)$ (this step is a natural generalization of (143)). In deriving the outer bound, we use appropriate genie-aided techniques (that involve providing suitable side information to the receiver). Again, what side information is shared is decided based on the typical error events which lead to the corresponding inequality in the inner bound.

Gap.

Finally, we characterize the gap between the outer and inner bounds to the capacity region for the n -state compound channel, in much the same way as we did for the two-state compound channel. This is stated formally below.

Theorem 10: For the N -state compound interference channel of Fig. 5, if (R_1, R_2) is in the outer bound to the capacity region, then $(R_1 - \Delta_1, R_2 - \Delta_2)$ is achievable, where

$$\begin{aligned} \Delta_1(Q, X_1, X_2) &= \max_{1 \leq n \leq N} I(X_2; S_{2\alpha_n} | U_{2\alpha_n}) \end{aligned} \quad (243)$$

$$\begin{aligned} \Delta_2(Q, X_1, X_2) &= \max_{1 \leq n \leq N} I(X_1; S_{1\alpha_n} | U_{1\alpha_n}). \end{aligned} \quad (244)$$

Specializing to the deterministic version, we can see that this gap is zero and hence the capacity region is characterized exactly. Specializing to the Gaussian version, we can see that this gap is no more than one bit. This completes the extension to the N -state compound channel scenario.

C. Discussion

A few comments on the structure and properties of the achievable scheme are in order here.

- Note that the structure of the achievable scheme (or the power split in the Gaussian scheme), which is characterized by the joint random variables

$$(Q, X_1, U_{1\alpha_1}, \dots, U_{1\alpha_N}, X_2, U_{2\alpha_1}, \dots, U_{2\alpha_N}) \quad (245)$$

depends *only on the interference states and not on the deterministic functions* $f_{k\alpha_n}$. The functions $f_{k\alpha_n}$ however may still help in determining the actual achievable rate region.

We highlight this point by considering the case when each of the degraded interference channels in our model are identity, i.e.

$$S_{k\alpha_2} = S_{k\alpha_1} = \dots = S_{k\alpha_N}, \quad k = 1, 2. \quad (246)$$

For this model the “compoundness” of the channel is only due to the functions $f_{k\alpha_n}$. Indeed, only two levels of superposition coding suffice, much as in the noncompound version of the problem.

- Let us assume

$$S_{1\alpha_2} = S_{1\alpha_1}. \quad (247)$$

Then our achievable scheme sets

$$U_{1\alpha_2} = U_{1\alpha_1}. \quad (248)$$

This implies that the level of the code book corresponding to $U_{1\alpha_2}$ is “degenerate” and that we might as well set

$$R_{1\alpha_2} = 0. \quad (249)$$

Suppose, however that

$$f_{1\alpha_2} \neq f_{1\alpha_1} \quad (250)$$

and hence the two receiver states $Y_{1\alpha_1}$ and $Y_{1\alpha_2}$ are not the same. While the receiver in either state adopts the same decoding technique (with respect to the level of interference it decodes), the higher dimensional constraints on the rate vector, as imposed by the decoding condition for each state, are different. Nevertheless, we see that for the Gaussian case one of these states is always worse than the other and thus would be the critical bottleneck in determining the achievable rates; this is done next.

VIII. THE COMPOUND GAUSSIAN INTERFERENCE CHANNEL

A. Model

The single-antenna Gaussian interference channel is parametrized by the complex channel parameters $(h_{11}, h_{21}, h_{12}, h_{22})$. The compound Gaussian interference channel lets the channel parameters take values from a set \mathcal{A} —finite or infinite

$$\mathcal{A} = \{(h_{11}, h_{21}, h_{12}, h_{22})\} \subseteq \mathbb{C}^4. \quad (251)$$

Without loss of generality, we can assume that the cross-link gains take real values. To see this, consider the signal at receiver 1

$$Y_1 = h_{11}X_1 + h_{12}X_2 + Z_1. \quad (252)$$

Since our model assumes R-CSI, the receiver 1 can rotate the received signal by $h_{12}^*/|h_{12}|$.

Define

$$\mathcal{A}_k \triangleq \{(h_{1k}, h_{2k}) \mid (h_{11}, h_{21}, h_{12}, h_{22}) \in \mathcal{A}\} \\ k = 1, 2. \quad (253)$$

Observe that the channels from the two transmitters to the receiver k are defined solely by the parameters (h_{1k}, h_{2k}) . Therefore, the set \mathcal{A}_k is the set of states that the receiver k can take. Now define \mathcal{A}' as

$$\mathcal{A}' \triangleq \mathcal{A}_1 \times \mathcal{A}_2 \\ = \{(h_{11}, h_{21}, h_{12}, h_{22}) \mid (h_{11}, h_{21}) \\ \in \mathcal{A}_1, (h_{12}, h_{22}) \in \mathcal{A}_2\}. \quad (254)$$

In other words \mathcal{A}' allows for *all combinations* of the possible states for both the receivers. Let $\mathcal{C}(\mathcal{A})$ denote the capacity region of the compound channel defined by the set \mathcal{A} . We have the following proposition.

Proposition 2:

$$\mathcal{C}(\mathcal{A}) = \mathcal{C}(\mathcal{A}').$$

Proof: Note that $\mathcal{A} \subseteq \mathcal{A}'$. Thus it is clear that any scheme that works for the compound channel \mathcal{A}' also works for the compound channel \mathcal{A} . However, since the two receivers do not cooperate, only the *marginal* channels to each receiver decide the decodability of any communication scheme. We now conclude that a scheme that works for the compound channel \mathcal{A} also works for the compound channel \mathcal{A}' . This completes the proof. \square

In the light of this observation, without loss of generality, we need only to consider compound channels whose state set \mathcal{A} decomposes as $\mathcal{A}_1 \times \mathcal{A}_2$.

B. Finite State Compound Channel

Let us first assume that the cardinality of \mathcal{A} (or equivalently \mathcal{A}_1 and \mathcal{A}_2) is finite. In Section II we saw that the case where the cardinality of \mathcal{A}_1 and \mathcal{A}_2 is restricted to two is captured by

the two-state compound interference channel of Fig. 2. Similarly, the general case where $|\mathcal{A}_1|$ and $|\mathcal{A}_2|$ are finite (with cardinality no more than N) is captured by the N -state compound interference channel of Fig. 5. We see this formally below. The key point is the infinitely divisible nature of Gaussian statistics. This aspect was used to show that the scalar Gaussian broadcast channel is always stochastically degraded (see [9, Sec. 14.1.3]). In a similar vein, the compound scalar Gaussian interference channel can always be supposed to have degraded interference states.

We begin by noting that if

$$|\mathcal{A}_1| \neq |\mathcal{A}_2| \quad (255)$$

then we can add redundant duplicate copies to one of the sets, so that

$$|\mathcal{A}_1| = |\mathcal{A}_2|. \quad (256)$$

Therefore, without loss of generality, we suppose this is true

$$|\mathcal{A}_1| = |\mathcal{A}_2| = N. \quad (257)$$

Then, *order* the finite sets \mathcal{A}_1 and \mathcal{A}_2 such that

$$\begin{aligned} h_{21\alpha_1} &\geq h_{21\alpha_2} \geq \dots \geq h_{21\alpha_N}, \\ h_{12\alpha_1} &\geq h_{12\alpha_2} \geq \dots \geq h_{12\alpha_N}. \end{aligned}$$

Next, we do the following substitution to reduce the finite state Gaussian interference channel to the model of Fig. 5.

$$S_{1\alpha_1} = h_{12\alpha_1} X_1 + Z_{2\alpha_1} \quad (258)$$

$$S_{1\alpha_n} = \frac{h_{12\alpha_n}}{h_{12\alpha_{(n-1)}}} S_{1\alpha_{(n-1)}} + \left(1 - \left(\frac{h_{12\alpha_n}}{h_{12\alpha_{(n-1)}}}\right)^2\right)^{1/2} Z_{2\alpha_n} \quad n = 2, \dots, N \quad (259)$$

$$S_{2\alpha_1} = h_{21\alpha_1} X_2 + Z_{1\alpha_1} \quad (260)$$

$$S_{2\alpha_n} = \frac{h_{21\alpha_n}}{h_{21\alpha_{(n-1)}}} S_{2\alpha_{(n-1)}} + \left(1 - \left(\frac{h_{21\alpha_n}}{h_{21\alpha_{(n-1)}}}\right)^2\right)^{1/2} Z_{1\alpha_n} \quad n = 2, \dots, N. \quad (261)$$

$$Y_{1\alpha_n} = f_{1\alpha_n}(X_1, S_{2\alpha_n}) = h_{11\alpha_n} X_1 + S_{2\alpha_n}, \quad 1 \leq n \leq N \quad (262)$$

$$Y_{2\alpha_n} = f_{2\alpha_n}(X_2, S_{1\alpha_n}) = h_{22\alpha_n} X_2 + S_{1\alpha_n}, \quad 1 \leq n \leq N. \quad (263)$$

Here $Z_{k\alpha_n}$'s are independent complex Gaussian random variables with unit variance. Note that the function $f_{k\alpha_n}$ captures the direct-link gains h_{11} and h_{22} . The channels $p(S_{k\alpha_n} | S_{k\alpha_{(n-1)}})$ capture the cross-link gains h_{12} and h_{21} as well as the additive noise.

Theorem 11: For the finite state compound Gaussian interference channel, multilevel superposition coding with Gaussian code books and opportunistic decoding depending on the interference state is within 1 bit of the capacity region.

Proof: We have shown earlier in this section that any finite state Gaussian interference channel is captured as a special case of the model in Fig. 5. Specializing the result of Theorem 10 to the Gaussian case, we have that the multilevel superposition coding is within 1 bit of the capacity. Further, it suffices to only consider Gaussian code books in the superposition code (along the same lines as Corollary 6). \square

C. Infinite State Compound Channel

We, next, consider the case of the compound interference channel with an infinite state set \mathcal{A} . The idea is to approximate the set by a quantized finite state \mathcal{A}^N . The finer the quantization is, the better the approximation will be. We first make a few key observations.

- 1) The number of levels needed in the superposition coding scheme only depends on the number of distinct values the cross link takes, i.e., if h_{12} takes N_1 distinct values and h_{21} takes N_2 distinct values then at receiver 1 we need a superposition coding scheme with $N_1 + 1$ levels and at receiver-2 we need a superposition coding scheme with $N_2 + 1$ levels.
- 2) The capacity of the compound interference channel depends only on the magnitude of the direct link gains and not the phase and further for fixed values of the cross-link gains, it is monotonically increasing with it. As a result, suppose

$$(h_{11}, h_{21}), (h'_{11}, h_{21}) \in \mathcal{A}_1, \quad |h_{11}| \leq |h'_{11}|. \quad (264)$$

These correspond to two states of the receiver 1, which differ only in the direct link gain, but have the same cross link gain. As observed in the previous section, for either of the two states, the receiver adopts the *same* decoding method. Further, since we have restricted ourselves to Gaussian code books, we see that the performance is restricted only by the state that has the weaker of the two direct links. Therefore, at any receiver, for a fixed cross link value the direct link which is the weakest is the bottleneck.

We can discard the state (h'_{11}, h_{21}) to reduce the set \mathcal{A}_1 . The compound channel with the reduced state-sets and the original compound channel have the same capacity and a scheme that works for the reduced state-set also works for the original compound channel.

We now succinctly describe the quantization procedure but will leave out the finer details of the proof. For the state set \mathcal{A} , we define the quantized state set \mathcal{A}^N obtained by taking an N -level quantization of the cross link gains and the maximum quantization interval of length δ . As $N \rightarrow \infty$, $\delta \rightarrow 0$. We consider a scheme for this N -state compound channel \mathcal{A}^N . This is an N -level superposition coding scheme with Gaussian code-books. We know that this scheme achieves within 1 bit of the capacity of the compound finite state compound interference channel \mathcal{A}^N .

We look at its performance when the channel is actually an infinite state compound channel \mathcal{A} . The cross links do not take just the quantized state value but instead take values in an interval

of length at most δ . However, the decoding (i.e., the number of interfering streams the receiver will decode) is done assuming that the receiver took value from the quantized state. The right hand side of the constraint inequalities on the sub-rates (which are similar to the right hand side of (30)–(53)) must now decrease to incorporate this difference. Assuming that these terms are continuous functions of the cross link gains with bounded derivatives they would decrease by a maximum amount of δ' , where $\delta' \rightarrow 0$ as $N \rightarrow \infty$. It can be shown that if (R_1, R_2) can be achieved by this scheme for the finite-state compound interference channel \mathcal{A}^N , then $(R_1 - \delta', R_2 - \delta')$ can be achieved by the this scheme for the infinite-state compound interference channel \mathcal{A} .

Note that the capacity of the finite-state compound interference channel \mathcal{A}^N is itself an upper bound to the capacity of the compound interference channel \mathcal{A} . Hence by taking a quantization large enough, we can achieve the capacity of the compound interference channel \mathcal{A} to within 1 bit of its capacity.

APPENDIX A ANALYSIS OF PROBABILITY OF ERROR

In the following we consider the decodability conditions at receiver $R_{X_{1\beta}}$ only. A very similar analysis applies to the other receiver-state pairs.

Due to the symmetry of the random code book generation, the probability of error averaged over the ensemble of random random code books, does not depend on which codeword was sent. Hence, without loss of generality, we can assume that the messages indexed by

$$(j_1, k_1, l_1) = (1, 1, 1), \quad (j_2, k_2, l_2) = (1, 1, 1) \quad (265)$$

were sent by the two transmitters respectively. Let us define the following event:

$$E_{jklm} = \left\{ \begin{aligned} &(Q^n, U_{1\beta}^n(j), U_{1\alpha}^n(j, k), \\ &X_1^n(j, k, l), U_{2\beta}^n(m), Y_{1\beta}^n) \\ &\in A_e^{(n)}(Q, U_{1\beta}, U_{1\alpha}, X_1, U_{2\beta}, Y_{1\beta}) \end{aligned} \right\}.$$

Letting $P_e^{(n)}$ denote the probability of decoding error at $R_{X_{1\beta}}$ we have

$$\begin{aligned} P_e^{(n)} &= P\left((\cup_m E_{111m})^c \cup \cup_{(j,k,l) \neq (1,1,1)} E_{jklm} \right) \quad (266) \\ &\leq \underbrace{P((\cup_m E_{111m})^c)}_{(a)} \\ &\quad + \underbrace{\sum_{l \neq 1} P(E_{11l1})}_{(b)} + \underbrace{\sum_{l \neq 1, m \neq 1} P(E_{11lm})}_{(c)} \\ &\quad + \underbrace{\sum_{k \neq 1, l} P(E_{1kl1})}_{(d)} + \underbrace{\sum_{k \neq 1, l, m \neq 1} P(E_{1klm})}_{(e)} \end{aligned}$$

$$+ \underbrace{\sum_{j \neq 1, k, l} P(E_{jkl1})}_{(f)} + \underbrace{\sum_{j \neq 1, k, l, m \neq 1} P(E_{jklm})}_{(g)}. \quad (267)$$

The final inequality used the union bound. Let us consider each term in (267) and study the conditions needed to make it go to 0 asymptotically (in n).

- It is straightforward to see that (a) goes to 0 as $n \rightarrow \infty$.
- Now consider (b). We begin by noting that $l \in \{1, \dots, 2^{nR_{1\beta}}\}$. Therefore if $R_{1\beta} = 0$ then (b) = 0. Else

$$(b) \leq 2^{nR_{1\beta}} 2^{-n(I(Y_{1\beta}; X_1 | U_{1\alpha}, U_{2\beta}, Q) - 5\epsilon)}. \quad (268)$$

Therefore for (b) to go to 0 as $n \rightarrow \infty$, we must have

$$R_{1\beta} \leq I(Y_{1\beta}; X_1 | U_{1\alpha}, U_{2\beta}, Q), \quad \text{if } R_{1\beta} > 0. \quad (269)$$

- Similarly, (c) is 0 if $R_{1\beta} = 0$ or $R_{2\beta} = 0$. Else, it must be that

$$R_{2\beta} + R_{1\beta} \leq I(Y_{1\beta}; X_1, U_{2\beta} | U_{1\alpha}, Q). \quad (270)$$

It is important to note that if $R_{2\beta} = 0$, but $R_{1\beta} > 0$ then, (270) is redundant because of (269). Therefore for (c) to go to 0 as $n \rightarrow \infty$ (assuming that (b) goes to 0 too), we must have,

$$R_{2\beta} + R_{1\beta} \leq I(Y_{1\beta}; X_1, U_{2\beta} | U_{1\alpha}, Q), \quad \text{if } R_{1\beta} > 0. \quad (271)$$

Similarly for (d), (e), (f) and (g), we must have

$$R_{1\alpha} + R_{1\beta} \leq I(Y_{1\beta}; X_1 | U_{1\beta}, U_{2\beta}, Q) \quad \text{if } R_{1\alpha} + R_{1\beta} > 0 \quad (272)$$

$$R_{2\beta} + R_{1\alpha} + R_{1\beta} \leq I(Y_{1\beta}; X_1, U_{2\beta} | U_{1\beta}, Q) \quad \text{if } R_{1\alpha} + R_{1\beta} > 0 \quad (273)$$

$$R_{1\beta} + R_{1\alpha} + R_{1\beta} \leq I(Y_{1\beta}; X_1 | U_{2\beta}, Q) \quad \text{if } R_{1\beta} + R_{1\alpha} + R_{1\beta} > 0 \quad (274)$$

$$R_{2\beta} + R_{1\beta} + R_{1\alpha} + R_{1\beta} \leq I(Y_{1\beta}; X_1, U_{2\beta} | Q), \quad \text{if } R_{1\beta} + R_{1\alpha} + R_{1\beta} > 0 \quad (275)$$

respectively.

APPENDIX B PROOF OF LEMMA 7

Consider any $(R_1, R_2) \in \mathcal{R}_{\text{in}}(P)$. Then there exists an

$$(R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta}) \in \mathcal{R}_{\text{in}}^{(6)}(P) \quad (276)$$

such that,

$$R_1 = R_{1p} + R_{1\alpha} + R_{1\beta} \quad \text{and} \quad R_2 = R_{2p} + R_{2\alpha} + R_{2\beta}. \quad (277)$$

We will find a

$$(\tilde{R}_{1p}, \tilde{R}_{1\alpha}, \tilde{R}_{1\beta}, \tilde{R}_{2p}, \tilde{R}_{2\alpha}, \tilde{R}_{2\beta}) \in \tilde{\mathcal{R}}_{\text{in}}^{(6)}(P) \quad (278)$$

such that,

$$\begin{aligned}\tilde{R}_{1p} + \tilde{R}_{1\alpha} + \tilde{R}_{1\beta} &= R_{1p} + R_{1\alpha} + R_{1\beta} \\ \tilde{R}_{2p} + \tilde{R}_{2\alpha} + \tilde{R}_{2\beta} &= R_{2p} + R_{2\alpha} + R_{2\beta}\end{aligned}$$

by the following algorithmic procedure.

Step 1a) For $k = 1, 2$, if $R_{k\beta} < 0$ then

$$R_{k\alpha} \leftarrow R_{k\alpha} + R_{k\beta}; \quad R_{k\beta} \leftarrow 0. \quad (279)$$

Step 1b) For $k = 1, 2$, if $R_{k\alpha} < 0$ then

$$R_{kp} \leftarrow R_{kp} + R_{k\alpha}; \quad R_{k\alpha} \leftarrow 0. \quad (280)$$

Step 2a) For $k = 1, 2$, if $R_{kp} < 0$ then

$$R_{k\alpha} \leftarrow R_{k\alpha} + R_{kp}; \quad R_{kp} \leftarrow 0. \quad (281)$$

Step 2b) For $k = 1, 2$, if $R_{k\alpha} < 0$ then

$$R_{k\beta} \leftarrow R_{k\beta} + R_{k\alpha}; \quad R_{k\alpha} \leftarrow 0. \quad (282)$$

First up, we note that at each step we are ensuring that $R_{1p} + R_{1\alpha} + R_{1\beta}$ and $R_{2p} + R_{2\alpha} + R_{2\beta}$ stay invariant. Next, note that if

$$(R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta}) \in \mathcal{R}_{\text{in}}^{(6)}(P) \quad (283)$$

and $(R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta})$ are all nonnegative, then it follows by definition of $\mathcal{R}_{\text{in}}^{(6)}(P)$ and $\tilde{\mathcal{R}}_{\text{in}}^{(6)}(P)$ that

$$(R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta}) \in \tilde{\mathcal{R}}_{\text{in}}^{(6)}(P) \quad (284)$$

and hence

$$(R_1, R_2) \in \tilde{\mathcal{R}}_{\text{in}}(P). \quad (285)$$

Claim 1: At the end of Step 1b, the new $(R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta})$ still remains in $\mathcal{R}_{\text{in}}^{(6)}(P)$ and satisfies

$$R_{1\alpha}, R_{1\beta}, R_{2\alpha}, R_{2\beta} \geq 0. \quad (286)$$

Proof: Consider Step 1a). Note that in this step we are potentially increasing $R_{k\beta}$, but the rest of the components either remain the same or decrease. Also note that in this step, we are keeping $R_{k\beta} + R_{k\alpha}$ invariant. Therefore, we only need to ensure that the inequalities among (24)–(53) that have $R_{k\beta}$, but not $R_{k\alpha}$ are not violated. This can be verified to be true, because of the polymatroidal nature of each block of the inequalities in (24)–(53). The argument is similar for Step 1b). \square

Claim 2: At the end of Step 2b), the new

$$(R_{1p}, R_{1\alpha}, R_{1\beta}, R_{2p}, R_{2\alpha}, R_{2\beta})$$

is in $\tilde{\mathcal{R}}_{\text{in}}^{(6)}(P)$.

Proof: Note that in Step 2a), the only component that potentially increases is R_{kp} , and so we might be violating one of

the following constraints: (24), (25), (30)–(32), (39), (40), and (45)–(47). However, by setting $R_{kp} = 0$, these violated constraints no longer matter for $\tilde{\mathcal{R}}_{\text{in}}^{(6)}(P)$. The argument is similar for Step 2b). Note that at the end of Step 2b), we have ensured that all the components are nonnegative. \square

REFERENCES

- [1] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.
- [2] A. A. El Gamal and M. H. M. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 343–346, Mar. 1982.
- [3] 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, Jul. 2008.
- [4] R. H. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.
- [5] E. Telatar and D. Tse, "Bounds on the capacity region of a class of interference channels," in *Proc. IEEE Int. Symp. Inf. Theory 2007*, Nice, France, Jun. 2007.
- [6] K. Kobayashi and T. S. Han, "A further consideration on the HK and the CMG regions for the interference channel," in *Proc. Inf. Theory Applicat. Workshop 2007*, San Diego, CA, Jan. 2007.
- [7] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inform. Theory*, vol. 33, no. 5, pp. 711–716, Sep. 1987.
- [8] R. Tyrrell Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1970.
- [9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, ser. Telecommunications. New York: Wiley, 1991.

Adnan Raja (M'06) received the B.Tech. degree from the Department of Electrical Engineering at the Indian Institute of Technology, Bombay, India, in 2006 and the M.S. degree from the Department of Electrical and Computer at the University of Illinois at Urbana-Champaign, in 2008, where he is currently pursuing the Ph.D. degree.

His research interests are in information theory and wireless communication.

Vinod Prabhakaran (M'07) received the Ph.D. degree from the Electrical Engineering and Computer Science (EECS) Department, University of California, Berkeley, in 2007.

He is currently a Postdoctoral Researcher at Coordinated Science Laboratory, University of Illinois, Urbana-Champaign. His research interests are in information theory, wireless communication, distributed signal processing and communication, and secure communication.

Dr. Prabhakaran has received the Tong Leong Lim Pre-Doctoral Prize and the Demetri Angelakos Memorial Achievement Award from the EECS Department, University of California, Berkeley.

Pramod Viswanath (M'00) received the Ph.D. degree from the Electrical Engineering and Computer Science (EECS) Department, the University of California at Berkeley in 2000.

He is an Associate Professor in the Electrical and Computer Engineering (ECE) Department at the University of Illinois at Urbana-Champaign, Urbana, IL. He was a Member of Technical Staff at Flarion Technologies until August 2001 before joining the ECE Department at the University of Illinois, Urbana-Champaign.

Dr. Viswanath is a recipient of the Eliahu Jury Award from the EECS Department of University of California (UC), Berkeley (2000), the Bernard Friedman Award from the Mathematics department of UC Berkeley (2000), and the NSF CAREER Award (2003). He was an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY from 2006 to 2008.