An Achievable Rate Region for Interference Channels with Conferencing

Yi Cao and Biao Chen
Department of EECS, Syracuse University, Syracuse, NY 13244
Email: ycao01@syr.edu, bichen@ecs.syr.edu

I. INTRODUCTION

The capacity region of an interference channel (IC), where the information sources at the two transmitters are statistically independent, has been a long standing problem. Carleial was the first to use the superposition code idea [1] to obtain an inner bound for IC. This inner bound was later improved by Han and Kobayashi [2] who gave an achievable rate region that is the largest reported to this date. Recently, a simplified description of the Han-Kobayashi (HK) rate region for the general IC is derived by Chong-Motani-Garg in [3].

A related and less well investigated problem is when the information sources at the two transmitters are correlated, i.e., interference channel with common information (ICCI) [4], [5]. In [4], an achievable rate region, an outer bound, and a limiting expression for the capacity region were obtained. Later, the capacity region of this channel under strong interference was found in [5]. Recently, improved achievable regions for general ICCI [6], [7] and three new outer bounds for the capacity region of Gaussian ICCI [8] were proposed. However, all those results are based on the assumption that the common message is available noncausally.

In this work, we investigate the problem of user cooperation in interference channels for the causal case. Here, each user not only transmits his own message to the intended receiver, but also serves as a relay to help transmit part of the other user’s message. We apply the superposition block Markov encoding, which was used previously for the relay channel [9] and for user cooperation in multiple access channels [10]. Our proposed achievable rate region is a generalized form of the HK region for IC [2], the capacity region of degraded relay channels [9], and the capacity region of the Gaussian vector broadcast channel (GVBC) [11].

This paper is organized as follows. In section II, we present the channel model and review some existing results. In section III, we propose an achievable region for general IC with transmitter conferencing. In section IV, numerical examples are used to compare the proposed region with the HK region and the capacity region of GVBC. We conclude in section V.

II. PRELIMINARIES AND EXISTING RESULTS

A. Definitions

A memoryless discrete IC with conferencing (ICC) is denoted by \((X_1, X_2, p, Y_1, Y_2, \tilde{Y}_1, \tilde{Y}_2)\), where \(X_1, X_2\) are two finite alphabet sets for the channel input, \(Y_1, Y_2\) are two finite alphabet sets for the channel output, \(\tilde{Y}_1, \tilde{Y}_2\) are two finite alphabet sets for the received signals at the transmitters (which also serve as relays), and \(p\) is the channel transition probability \(p(y_1, y_2, \tilde{y}_1, \tilde{y}_2 | x_1, x_2)\). Here we assume the channel is memoryless and encoders 1 and 2 are allowed to depend noncausally on their own messages and the past values of \(\tilde{y}_2\) and \(\tilde{y}_1\). Let \(\mathcal{M}_1 = \{1, 2, \cdots, M_1\}\) and \(\mathcal{M}_2 = \{1, 2, \cdots, M_2\}\) be the message sets of sender 1 and sender 2, respectively. Thus, for \(w_1 \in \mathcal{M}_1\) and \(w_2 \in \mathcal{M}_2\), the joint probability mass function of \(\mathcal{M}_1 \times \mathcal{M}_2 \times X_1^n \times X_2^n \times Y_1^n \times Y_2^n \times \tilde{Y}_1^n \times \tilde{Y}_2^n\) is given by

\[
p(w_1, w_2, x_1, x_2, y_1, y_2, \tilde{y}_1, \tilde{y}_2) = p(w_1)p(w_2)\prod_{i=1}^n p(x_1 | w_1, \tilde{y}_{i-1})\prod_{i=1}^n p(x_2 | w_2, \tilde{y}_{i-1})p(y_1 | w_1, y_i, \tilde{y}_{i-1})p(y_2 | w_2, y_i, \tilde{y}_{i-1}, x_{1i}, x_{2i})
\]

Suppose \(w_1 \in \mathcal{M}_1\) and \(w_2 \in \mathcal{M}_2\) are sent by transmitters 1 and 2 respectively, \(g_1\) and \(g_2\) are the decoding functions at receivers 1 and 2; the average probabilities of decoding error of this channel are defined as

\[
P_{e,1}^{(n)} = \frac{1}{M_1 M_2} \sum_{w_1, w_2} Pr(g_1(Y_1) \neq w_1 | w_1, w_2) \quad (2)
\]

\[
P_{e,2}^{(n)} = \frac{1}{M_1 M_2} \sum_{w_1, w_2} Pr(g_2(Y_2) \neq w_2 | w_1, w_2) \quad (3)
\]
The capacity region of ICC is the closure of all rate pairs \((R_1, R_2)\) such that \(P_{e_1}^{(n)} \to 0, P_{e_2}^{(n)} \to 0\) as codeword length \(n \to \infty\), where \(R_1 = \frac{1}{n} \log M_1\) and \(R_2 = \frac{1}{n} \log M_2\).

### B. Existing Results

1) Chong-Motani-Garg recently derived a simplified description of the HK region for IC [3], as summarized below.

**Proposition 1:** Let \(P_r^*\) be the set of probability distributions \(P_r^*\) that factor as

\[
P_r^*(q, u_1, u_2, x_1, x_2) = p(q)p(x_1, u_1|q)p(x_2|u_2|q).
\]

(4)

For a fixed \(P_r^* \in P_r^\ast\), let \(\mathcal{R}_{HK}^c(P_r^*)\) be the set of \((R_1, R_2)\) satisfying (9)-(15) in Theorem 2 of [3]. Then \(\bigcup_{P_r^* \in P_r^\ast} \mathcal{R}_{HK}^c(P_r^*)\) is equivalent to the HK region.

2) The capacity of the degraded relay channel is given in proposition 2 [9].

**Proposition 2:** A relay channel consists of an input \(x_1\), a relay output \(y_1\), a channel output \(y\), and a relay sender \(x_2\) (whose transmission is allowed to depend on the past symbols of \(y_1\)). If \(y\) is a degraded form of \(y_1\) [9], then

\[
C = \max_{P(x_1, x_2)} \min \{I(X_1, X_2; Y), I(X_1; Y_1|X_2)\}.
\]

(5)

3) The capacity region of GVBC is computed using a covariance matrix constraint on the inputs \(X = (X_1, X_2)^T\) of the form \(E[XX^T] \leq S\). In order to mimic the individual power constraints \(P_1, P_2\) on the two users for the vector case, the input covariance matrix \(S\) is of the form \(S = \begin{pmatrix} P_1 & c \\ c & P_2 \end{pmatrix}\), for some \(-\sqrt{P_1P_2} \leq c \leq \sqrt{P_1P_2}\). Then, the capacity region of GVBC is given below [11].

**Proposition 3:** For each such \(S\) and all positive semi-definite matrices \(B\) and \(D\), where \(B + D \leq S\), both rate pairs

\[
R_1 \leq \frac{1}{2} \log \left( \frac{|H_1B + H_1D + Q_s|}{|Q_s|} \right), R_2 \leq \frac{1}{2} \log \left( \frac{|H_2(B + D)H_2^T + Q_s|}{|H_2B + H_2D + Q_s|} \right)
\]

and

\[
R_1 \leq \frac{1}{2} \log \left( \frac{|H_1(B + D)H_1^T + Q_s|}{|H_1B + H_1D + Q_s|} \right), R_2 \leq \frac{1}{2} \log \left( \frac{|H_2D + H_2^T + Q_s|}{|Q_s|} \right)
\]

(6)

(7)

are achievable, where \(H_1 = (1, a_{21})\) and \(H_2 = (a_{12}, 1)\). The convex hull of the union of these pairs over all possible \(S, B\) and \(D\) matrices is the capacity region of GVBC.

### III. Main Results

We first give a brief outline of our encoding-decoding strategy. We split each user’s message into two parts: \(M\) and \(W\), where \(M\) is to be sent directly to the intended receiver, and \(W\) is the cooperative message to be sent to the receiver via the cooperation of the other user (relay). Our cooperation strategy is based on superposition block Markov encoding with the assumption that \(W\) can be perfectly decoded by the relay. The purpose of introducing \(M\) is to achieve a reasonable rate region (no less than IC without conferencing) even when the conferencing channel is poor. For the message \(M\), we apply simultaneous superposition coding [2] and further split it into two parts: private message \(V\) and common message \(U\).

For the cooperation in transmitting \(W\), we jointly consider \(B\) blocks, each of \(n\) symbols. Each user transmits a sequence of \(B - 1\) messages \(w_1, \cdot \cdot \cdot, w_{B-1}\) in \(B\) blocks, with no new message in the last block. Note that as \(B \to \infty\), \((B - 1)/B\) is arbitrarily close to 1, hence the penalty on rate is negligible. Now we take user 1 as an example to show the whole process. Suppose there are \(2nR_{10}\) codewords of \(W_1\) for user 1 to transmit. We establish a random partition by randomly throwing them into \(2nR_{10}\) cells. This partition is made known to both transmitters and receivers. Suppose user 1 sends message \(w_{1,b-1}\) at block \(b-1\). At the end of block \(b-1\), following our assumption, user 2 can perfectly decode \(w_{1,b-1}\) and calculate the cell index \(s_{1b}\) to which \(w_{1,b-1}\) belongs. At block \(b\), both user 1 and 2 spend some power transmitting \(s_{1b}\). This provides the basis for cooperatively resolving the remaining \(Y_1\) uncertainty about \(w_{1,b-1}\). After decoding \(s_{1b}\) at the end of block \(b\), receiver \(Y_2\) intersects its ambiguity set \(D(y_1(b-1))\) i.e., the set of all codewords \(w\) that are jointly typical with \(y_{1,b-1}\) [9] with cell \(s_{1b}\) and gets the unique correct codeword \(w_{1,b-1}\) with a probability close to 1.

Since both users 1 and 2 can perfectly decode each other’s messages \(W_1, W_2\) and then calculate the corresponding cell indices \(S_1, S_2\), we can employ dirty paper coding (DPC) to transmit \(V_1, U_1\) and \(S_1\) treating \(S_2\) as a known interference at transmitter 1. Similarly at transmitter 2, we can transmit \(V_2, U_2\) and \(S_2\) treating \(S_1\) as a known interference. Thus, introducing auxiliary random variables \(M_1, N_1, G_1, H_1\) and \(M_2, N_2, G_2, H_2\) for DPC, we summarize the achievable region for ICC in the theorem below.

**Theorem 1:** Let \(Z_1 = (Y_1, Y_2, Y_1, X_1, X_2, M_1, N_1, G_1, H_1, V_2, U_2, W_2, S_2, Q)\) and let \(P_r^*\) be the set of distribution on \(Z_1\) that can be factored into the form

\[
p(q)p(u_2|q)p(w_2|q)p(s_2|q)p(v_2|w_2s_2q)\]

\[
\times p(n_1|s_2q)p(g_1|s_2q)p(h_1|s_2q)p(m_1|n_1h_1s_2q)\]

\[
\times p(x_1|m_1g_1s_2q)p(x_2|v_2w_2h_1q)p(y_1|y_2|y_1|y_2|x_1|z_2)
\]

Let \(S(Z_1)\) be the set of \((R_1, R_2)\) such that \(R_1 = R_{11} + R_{12} + R_{13}\) and \(R_2 = R_{22} + R_{21} + R_{23}\) satisfying:

\[
R_{11} \leq L_{11} - I(M_1; S_2|N_1H_1) (9)
\]

\[
R_{12} \leq L_{12} - I(N_1; S_2) (10)
\]

\[
R_{13} \leq L_{13} - I(G_1; S_2) (11)
\]

\[
R_{10} \leq L_{10} - I(H_1; S_2) (12)
\]

\[
L_{11} \leq I(Y_1N_1H_1U_2; M_1) (13)
\]

\[
L_{11} + L_{12} \leq I(Y_1H_1U_2; M_1N_1H_1) (14)
\]

\[
L_{11} + L_{10} \leq I(Y_1N_1U_2; M_1H_1) (15)
\]

\[
L_{11} + L_{21} \leq I(Y_1N_1H_1; M_1U_2) (16)
\]

\[
L_{11} + L_{12} + L_{10} \leq I(Y_1U_2; M_1N_1H_1) (17)
\]

\[
L_{11} + L_{21} + L_{23} \leq I(Y_1N_1H_1; M_1H_1U_2) (18)
\]

\[
L_{11} + L_{12} + L_{21} + L_{23} \leq I(Y_1; M_1N_1H_1U_2) (20)
\]

\[
L_{13} \leq R_{10} + I(Y_1M_1N_1H_1U_2; G_1) (21)
\]

\[
L_{13} \leq I(Y_1H_1S_2; G_1) (22)
\]
Let $\mathcal{R}_1^* = \bigcup_{Z_1 \in \mathcal{P}^*} S(Z_1)$. Swap index 1 and 2 in all of the above statements and inequalities and we get $\mathcal{R}_2^* = \bigcup_{Z_2 \in \mathcal{P}^*} S(Z_2)$. Then the achievable region $\mathcal{R}^* = \text{convhull}(\mathcal{R}_1^* \cup \mathcal{R}_2^*)$ and the cardinality $||Q|| \leq 33$. 

**Proof:** We only need to prove the achievability of $\mathcal{R}^*_2$. 

**Codebook Generation:** Let $q = (q^{(1)}, \ldots, q^{(n)})$ be a random sequence of $Q^n$ distributed according to $\prod_{i=1}^n p(q^{(i)})$. Generate $2^{nR_2}$ i.i.d (independently and identically distributed) codewords $w_2(j_2)$ for common messages, $2^{nR_2}$ i.d codewords $s_2(l_2)$ for cell indices, and $2^{nR_2}$ i.d codewords $w_2(k_2)$ for cooperative messages according to $\prod_{i=1}^n p(w_2^{(i)} q^{(i)})$, respectively. For each pair of $(w_2(j_2), s_2(l_2))$, generate $2^{nR_2}$ i.i.d codewords $w_2(j_2, j_2, l_2)$ for private messages according to $\prod_{i=1}^n p(w_2^{(i)} q^{(i)} | s_2^{(i)} q^{(i)})$. Generate $2^{nL_1}$ i.i.d codewords $m_1(n_1)$ for common messages according to $\prod_{i=1}^n p(n_1^{(i)} q^{(i)})$ and randomly place them into $2^{nR_{10}}$ bins; generate $2^{nL_1}$ i.i.d codewords $h_1(l_1)$ for cell indices according to $\prod_{i=1}^n p(h_1^{(i)} q^{(i)})$ and randomly place them into $2^{nR_{10}}$ bins; generate $2^{nL_1}$ i.i.d codewords $g_1(\psi_1)$ for cooperative messages according to $\prod_{i=1}^n p(g_1^{(i)} q^{(i)})$ and randomly place them into $2^{nR_{13}}$ bins. For each pair of $(m_1(n_1), h_1(l_1))$, generate $2^{nL_1}$ i.i.d codewords $m_1(n_1, l_1, l_1)$ for private messages according to $\prod_{i=1}^n p(m_1^{(i)} n_1^{(i)} h_1^{(i)} q^{(i)})$ and randomly place them into $2^{nR_{11}}$ bins.

To apply superposition block Markov encoding, we also need two random partitions. Randomly place the above generated $2^{nR_2}$ codewords $w_2(k_2)$ into $2^{nR_2}$ cells, and those $2^{nL_1}$ codewords $g_1(\psi_1)$ into $2^{nR_{10}}$ cells.

**Encoding:** In block b, user 2 wants to send new indices $i_{2b}, j_{2b}$ and $k_{2b}$. For cooperatively resolving the remaining $Y_2$ uncertainty about $w_2(k_{2b}, b-1)$ in the previous block $b-1$, it also sends the cell index of $w_2(k_{2b}-1, b)$, denoted by $l_{2b}$. At the same time, user 1 wants to send new indices $i_{1b}, j_{1b}, k_{1b}$ and the cell index of $g_1(\psi_1, b-1)$, denoted by $l_{1b}$. Since user 1 can also perfectly calculate $l_{2b}$ at the end of block $b-1$, it looks into bins $j_{1b}, k_{1b}$ and $l_{1b}$ for codewords $m_1(n_1)_{1b}$ and $g_1(\psi_1)_{1b}$ and $h_1(l_1)_{1b}$ that are jointly typical with $s_2(l_{2b})$, respectively. For the previously found $(n_1(\eta_{1b}), h_1(\omega_{1b}))$, encoder 1 looks into bin $i_{1b}$ for codeword $m_1(\xi_{1b}, \eta_{1b}, \omega_{1b})$ such that $(q_{s_2(l_{2b})}, n_1(\eta_{1b}), h_1(\omega_{1b}), m_1(\xi_{1b}, \eta_{1b}, \omega_{1b}))$ are jointly typical. For the above bin searching, if there is more than one such codeword, pick the one with the smallest index; if there is no such codeword, declare an error. Then, user 1 sends $x_1$ generated according to $\prod_{i=1}^n p(x_1^{(i)} | m_1(\xi_{1b}, \eta_{1b}, \omega_{1b})^{(i)} g_1(\psi_1)^{(i)} s_2^{(i)}(l_{2b})^{(i)})$ and user 2 sends $x_2$ generated according to $\prod_{i=1}^n p(x_2^{(i)} | v_2^{(i)}(i_{2b}, j_{2b}, k_{2b})^{(i)} w_2^{(i)}(k_{2b})^{(i)} h_1(\omega_{1b})^{(i)} q^{(i)})$.

**Decoding:** User 2, as a relay to user 1, wants to correctly recover the new index $k_{1b}$ sent in block $b$. Since it already knows $h_1(\omega_{1b})$ and $s_2(l_{2b})$ during encoding, it looks for all the sequences $g_1(\psi_1)$, such that

$$\{q_{s_2(l_{2b})}, h_1(\omega_{1b}), g_1(\psi_1), y_{1b}\} \in A_{i}^{(n)} (Q S H_2 G_1 Y_{1b})$$

If those sequences satisfying (35) have the same bin index $k_{1b}$, we declare $k_{1b} = k_{1b}$. Otherwise, we declare an error. On the other hand, user 1 determines the unique $w_2(k_{2b})$, such that

$$\{q_{s_2(l_{2b})}, h_1(\omega_{1b}), w_2(k_{2b}), y_{1b}\} \in A_{i}^{(n)} (Q S H_2 W Y_{2b})$$

At the receiver side, we assume $Y_1$ knows $i_{1b}, j_{1b}, l_{1b}, j_{2b}, l_{2b}$, and it can construct $m_1(\xi_{1b}, \eta_{1b}, \omega_{1b}, \xi_{2b}, \eta_{2b}, \omega_{2b})$, $(n_1(\eta_{1b}), h_1(\omega_{1b}), u_2(j_{2b}, l_{2b})$, which are jointly typical with $y_{1b}, l_{1b}$. Now it wants to first decode bin indices $i_{1b}, j_{1b}, l_{1b}$ and the common message index $j_{2b}$. It looks for $m_1(\xi_{1b}, \eta_{1b}, \omega_{1b}), m_1(\eta_{1b}), h_1(\omega_{1b}), u_2(j_{2b})$, such that

$$\{q, m_1(\xi_{1b}, \eta_{1b}, \omega_{1b}), m_1(\eta_{1b}), h_1(\omega_{1b}), u_2(j_{2b}), y_{1b}\} \in A_{i}^{(n)} (Q M_1 N_1 H_2 U_2 Y_{1b})$$

If those sequences satisfying (35) have the same bin indices and message index respectively, we declare $i_{1b} = i_{1b}, j_{1b} = j_{1b}, k_{1b} = l_{1b}$, and $j_{2b} = j_{2b}$. Otherwise, declare an error. Assuming cell index $l_{1b}$ is successfully decoded at $Y_1$, then we declare $k_{1b} = k_{1b}, j_{1b}$ if those sequences $g_1(\psi_1) \in C(l_{1b}) \cap D(y_{1b} | b) = \{k_{1b} \}$ is the same bin index $k_{1b}$. Here $C(l_{1b})$ denotes the set of $g_1(\psi_1)$ in cell $l_{1b}$, and $D(y_{1b} | b)$ is the ambiguity set, i.e., sequences of $g_1(\psi_1)$ such that

$$\{q, m_1(\xi_{1b}, \eta_{1b}, \omega_{1b}, i_{1b}, \omega_{1b}, \xi_2, \eta_{2b}, \omega_{2b}), m_1(\xi_{1b}, \eta_{1b}, \omega_{1b}) \} \in A_{i}^{(n)} (Q M_1 N_1 H_2 U_2 G_1 Y_{1b})$$

For $Y_2$, the decoding process is the same and we skip the details.

**Analysis of error probability:** We first consider $P_r^{(n)}$ and we still use the story in block $b$. Let $P_0$ denote the probability that there is no $m$ in bin $i_{1b}$, such that $(q, s_2(l_{2b}), n_1(\eta_{1b}), h_1(\omega_{1b}), m_1(\xi_{1b}, \eta_{1b}, \omega_{1b}))$ are jointly typical. Then,

$$P_0 \leq (1 - 2^{-n(L_1 + S_2 N_1 H_1 Q + 3r - L_1 + 1)} (1 / n))^{2^{n(L_1 + S_1 D_1)} / n^{S_1 D_1}}$$

So, (9) guarantees $P_0 \to 0$ as $n \to \infty$. Similarly, bounds (10)-(12) guarantee that encoder 1 can find codewords $n_1(\eta_{1b})$,...
$g_1(\psi_1b)$ and $h_1(\omega_1b)$, which are jointly typical with $s_2(l_2b)$, respectively.

Now we calculate the error probability for user 2 (as a relay) to decode $k_{1b}$. Denote the sent codeword $g_1(\psi_1b)$ as $g_1(k_{1b}, k^*)$ since it is picked from bin $k_{1b}$. $k^*$ denotes the index of $g_1(\psi_1b)$ in bin $k_{1b}$. Let $E_1(k_1, k)$ denote the event (33) and let $P_1$ denote the probability for user 2 to make a decoding error. Then

\[
P_1 = Pr\{E_1(k_{1b}, k^*) \text{ or } \bigcup_{k_{1b}\neq k_1} E_1(k_1, k)\} \leq Pr\{E_1(k_{1b}, k^*)\} + \sum_{k_{1b}\neq k_1} Pr\{E_1(k_1, k)\}
\]

(39)

(40)

(41)

For $k_1 \neq k_{1b}$, we know

\[
Pr\{E_1(k_1, k)\} = \sum_{(q_{n-2}, h_1, g_1(1)) \in A(n)} p(q) p(g_1|q)p(s_2h_1y_{1b}|q) \leq \frac{|A(n)|^2}{n} \cdot \frac{2^{-(H(Q) - \epsilon)} + 2^{-(H(G_1) - \epsilon)} + 2^{-(H(S_2H_1y_1) - \epsilon)}}{2^{-(H(Q) + H(G_1) + H(S_2H_1y_1) - H(QG_1S_2H_1y_1) - 4\epsilon)}} \leq 2^{-n(I(S_2H_1y_1G_1|Q) - 4\epsilon)}
\]

Therefore, $P_1 \leq \epsilon + 2^{-n(I(S_2H_1y_1G_1|Q) - L_{13} - 4\epsilon)}$. $\epsilon$ can be arbitrarily small by letting $n \to \infty$. Thus, bound (22) assures $P_1 \to 0$ as $n \to \infty$.

For the decoding of $i_{1b}, j_{1b}, l_{1b}$ and $j_{2b}$ by $Y_1$, it is a direct application of the simultaneous superposition coding [2]. However, regarding our codebook generation scheme, particularly the construction of $m_1$, we will get a somewhat simpler description, similar to that of Chong-Motani-Garg [3]. This leads to the bounds (13)-(20) and we skip the details here.

Let $E_2(k_1, k)$ denote the event that $g_1(k_1, k) \in C(l_{1b}) \cap D(y_1(b-1))$ and $E_2(k_1, k)\in C(l_{2b})$. We also define an indicator function $I(k_1, k)$. If $g_1(k_1, k)$ satisfies (36), $I(k_1, k) = 1$; otherwise, $I(k_1, k) = 0$. The number of sequences in $D(y_1(b-1))$ with bin index $k_1 \neq k_{1b-1}$ is

\[
||D(y_1(b-1))|| = \sum_{k_1 \neq k_{1b-1, k}} E(I(k_1, k)) = \sum_{k_1 \neq k_{1b-1, k}} Pr\{E_2(k_1, k)\} \leq \sum_{k_1 \neq k_{1b-1, k}} 2^{-n(I(Y_1M_{1b}H_1U_2G_1|Q) - 4\epsilon)} \leq 2^{-n(I(Y_1M_{1b}H_1U_2G_1|Q) - L_{13} - 4\epsilon)}
\]

(42)

(43)

(44)

(45)

(46)

Now, let $P_2$ denote the probability of error for $Y_1$ to decode $k_{1, b-1}$. Denote the actually sent codeword $g_1(\psi_{1, b-1})$ in block $b-1$ as $g_1(k_{1, b-1}, k^*)$. Then,

\[
P_2 = Pr\{E_2^c(k_{1, b-1}, k^*) \text{ or } \bigcup_{k_{1b} \neq k_{1, b-1}} E_2(k_1, k)\} \leq \epsilon + \sum_{k_1 \neq k_{1, b-1}} Pr\{E_2(k_1, k)\}
\]

(47)

(48)

(49)

(50)

Bound (21) guarantees $P_2 \to 0$ as $n \to \infty$. Thus, with bounds (9)-(22), it is guaranteed that $Y_1$ will correctly decode $i_{1b}, j_{1b}, l_{1b}, j_{2b}$ and $k_{1b-1}$ at the end of block $b$ with a probability arbitrarily close to 1. Then, the information state of receiver $Y_1$ propagates forward, yielding the total decoding error probability $P_{e, 1}^{(n)} \to 0$ as $n \to \infty$. The analysis of $P_{e, 2}^{(n)}$ is similar to $P_{e, 1}^{(n)}$, which leads to bounds (23)-(32). Thus, $R_1^*$ is achievable. Proof of $R_2^*$ is identical, hence $R^*$ is achievable for ICC via time sharing.

The cardinality bound on $Q$, i.e., $|Q| \leq 33$ is obtained by applying the Carathéodory Theorem to a set of inequalities that bound the rate pair $(R_1, R_2)$, obtained via Fourier-Motzkin elimination to Eqs. (9)-(32). Q.E.D.

**Remarks:** From the encoding-decoding strategy of the above theorem, the achievable region $R^*$ is actually a generalization of the HK region for IC, the capacity region of degraded relay channels, and the capacity region of GVBC. It reduces to those extreme cases under the conditions elaborated below.

1) When the conferencing channel between the two users is very poor, the bound in (22) and (32) can be very small. In this case, allocating power for cooperation will actually reduce the rate otherwise achievable via direct transmission. As a result, the encoders will not allocate any power to transmit $W_1$ and $W_2$, so $W_1 = W_2 = S_1 = S_2 = 0$ and $G_1 = G_2 = H_1 = H_2 = 0$. Then, both $R_1^*$ and $R_2^*$ reduce to the region in Proposition 1, which equals to the HK region.

2) When the conferencing channel between the two users is good enough, it is not necessary to transmit messages directly to the receiver, because cooperative transmission with the other user will always yield a better rate. In this case, the encoders will let $V_1 = V_2 = U_1 = U_2 = 0$ and $M_1 = M_2 = N_1 = N_2 = 0$. Now, if user 2 refrains from transmitting its own message and only serves as a relay to user 1 (i.e., $W_2 = S_2 = G_2 = H_2 = 0$), both $R_1^*$ and $R_2^*$ reduce to the capacity region of the degraded relay channel in Proposition 2. Similarly, if user 1 serves only as a relay to user 2, it also reduces to the capacity region of the degraded relay channel.

3) When the conferencing channel between the two users is ideal (i.e., the conferencing channel capacity is infinite), the bounds in (22) and (32) are no longer needed. So, $W_1 = W_2 = G_1 = G_2 \approx 0$. Combining the result in case 2 that $V_1 = V_2 = U_1 = U_2 = 0$ and $M_1 = M_2 = N_1 = N_2 = 0$, we can easily check that $R_1^*$ reduces to

\[
R_1^* \leq I(Y_1; H_1|Q) - I(H_1; S_2|Q)
\]

(51)

\[
R_2^* \leq I(Y_2; S_2|Q)
\]

(52)
and $R_2^*$ reduces to
\[
R_1 \leq I(Y_1; S_1|Q) \tag{53}
\]
\[
R_2 \leq I(Y_2; H_2|Q) - I(H_2; S_1|Q) \tag{54}
\]
For the Gaussian case, the rate region (51)-(52) becomes (6) and (53)-(54) becomes (7). So, in this case, $R^*$ reduces to the capacity region of GVBC.

4) During the review process of this paper, we became aware of [12], which essentially tackles the same problem using a different approach. In [12], user cooperation results in a common information (in the sense of [4]) at the encoders and the decoder uses backward decoding (similar to that of [10]) instead of the random partitioning (i.e., binning) we use in our approach. Except for some extreme cases, it appears no subset relation can be established. The obtained achievable region in [12] is simpler as it does not involve a large number of auxiliary variables; however, the scheme in [12] is strictly suboptimal for certain extreme cases (e.g., degraded relay channels, IC with degraded message sets with weak interference, and MIMO BC) whereas our achievable region can be easily shown to be optimal in each of these cases.

IV. NUMERICAL EXAMPLES

The standard form of a Gaussian interference channel is:
\[
Y_1 = X_1 + a_{21}X_2 + Z_1
\]
\[
Y_2 = a_{12}X_1 + X_2 + Z_2
\]
where $Z_1$ and $Z_2$ are arbitrarily correlated zero mean, unit variance Gaussian random variables. Suppose the power constraints of $X_1$ and $X_2$ are $P_1$ and $P_2$, respectively. For the conferencing channel with perfect echo cancellation, we have
\[
\tilde{Y}_1 = K_1X_1 + Z_1, \quad \tilde{Y}_2 = K_2X_2 + Z_2
\]
where $Z_1$ and $Z_2$ are both zero mean, unit variance Gaussian variables. By reciprocity, we assume the channel coefficient $K_1 = K_2$. Since the computation of $R^*$ is formidable, here we constrain all the inputs to be Gaussian distributed and set $Q = \phi$ in order to compare our region with $G^*$ in (5.9) of [2] and the capacity region of GVBC. We denote this modified region as $R$. Consider, that for certain $\alpha_t, \beta_t, \gamma_t, \theta_t, \mu_t \in [0, 1]$, with $\alpha_t + \beta_t + \gamma_t + \theta_t + \mu_t = 1$, where $t = 1, 2$, the following hold:
\[
U_t \sim N(0, \beta_t P_t), W_t \sim N(0, \gamma_t P_t), S_t \sim N(0, 1) \tag{58}
\]
\[
V_t = V_t' + U_t + \sqrt{\theta_t P_t S_t}, \quad \text{where} \quad V_t' \sim N(0, \alpha_t P_t) \tag{59}
\]
\[
X_1 = V_1 + W_1 + \sqrt{\mu_t} S_1, \quad X_2 = V_2 + W_2 + \sqrt{\mu_2} P_2 S_1 \tag{60}
\]
After applying Fourier-Motzkin Elimination on those bounds (9)-(32), we find that for each set of $(\alpha_1, \beta_1, \gamma_1, \theta_1, \mu_1)$ and $(\alpha_2, \beta_2, \gamma_2, \theta_2, \mu_2)$, both $S(Z_1)$ and $S(Z_2)$ are delimited by straight lines of slope $0, -\frac{1}{2}, -1, -2, \infty$ as in the original HK region. Exhausting all the parameters between $[0, 1]$, and taking the convex hull of all those $S(Z_1)$ and $S(Z_2)$, we get the achievable region $R$ for ICC in Fig.2.

Fig. 2. Comparison of $R$ with HK region and Gaussian vector broadcast channel capacity. $P_1 = 6, P_2 = 1.5, a_{12} = a_{21} = 0.74$

Remarks:

1) When there is no conferencing between the two users, the achievable region reduces to the HK region. When the quality of the conferencing channel improves, it increases our achievable region for ICC within the limit of the capacity of GVBC.

2) When the channel coefficient is $K_1 = K_2 = 4$, the region $R$ is already very close to the upper bound; when $K_1 = K_2 = 1$, which is equal to the channel coefficient of the transmitter to the receiver, cooperation achieves a slightly better rate region than independent transmission.

3) For the channel coefficient $K_1 = K_2 = 4$, the corresponding relay channels (i.e., one of the users only serves as a relay) are degraded, thus the intercepts of the bound at both axes are the capacities of respective relay channels.

REFERENCES