A Limiting Formula for the Interference Channel Capacity Region

Mohammad J. Rezaeian and Alex Grant
Institute for Telecommunications Research
University of South Australia
Mawson Lakes Blvd
Mawson Lakes SA 5095 Australia
e-mail: mreza@spri.levels.unisa.edu.au

Abstract — A limiting expression for the capacity region of the interference channel is found using a sequence of achievable regions starting from the best previously known single letter achievable region. This formula allows approximation of the capacity region by fewer terms of the sequence, as compared to previous limiting formulas.

I. INTRODUCTION

For the general model of the interference channel the capacity region is a limiting expression [1]. The widest class of interference channels with a single letter capacity region is the strong interference channel [2]. The best general single-letter achievable region was given by Han and Kobayashi [3]. In this paper we define a sequence of achievable regions starting from the Han and Kobayashi region that monotonically approaches the capacity region. Sufficiently large terms of this sequence can in principle give precise approximations to the capacity region. This work was motivated by Kraemer [4, Theorem 5.1], although we do not use the concept of directed information. The basic idea behind the new regions is to apply Han and Kobayashi to extended interference channels, in which l-sequences of inputs and outputs for the original channel are regarded as symbols of the extended channel.

In the next section the basic definitions for the interference channel model and its capacity region are given. In Section III the capacity region is derived by a limiting formula.

Random variables $x \in X$ with support $X$ shall be shown upper case, with realizations lower case, $x$. Boldface $X \in X^{m \times n}$ (with realizations $x$) is a $m \times n$ random matrix. $\mathbb{R}_+$ are the non-negative reals.

II. DISCRETE MEMORYLESS INTERFERENCE CHANNEL

Definition 1. A discrete memoryless interference channel $([X_1, X_2, Y_1, Y_2, \omega(y_1, y_2|x_1, x_2)]_{IC})$ consists of two input alphabets $X_1, X_2$, two output alphabets $Y_1, Y_2$ and a channel transition probability matrix $\omega(y_1, y_2|x_1, x_2)$.

Given $\omega(y_1, y_2|x_1, x_2)$ with marginals $\omega_1(y_1|x_1, x_2), \omega_2(y_2|x_1, x_2)$ and any integer $n > 0$, define the transmission probability matrix $\omega^n(y_1, y_2|x_1, x_2) = \prod_{i=1}^{n} \omega(y_{1i}, y_{2i}|x_{1i}, x_{2i})$ with marginals $\omega_1^n(y_1|x_1, x_2), \omega_2^n(y_2|x_1, x_2)$.

Definition 2. A $(n, M_1, M_2, \lambda)$ code for $[X_1, X_2, Y_1, Y_2, \omega(y_1, y_2|x_1, x_2)]_{IC}$ consists of codebooks $C_1 = \{x_1(i) | i = 1, \ldots, M_1\}$ and $C_2 = \{x_2(j) | j = 1, \ldots, M_2\}$ with corresponding disjoint decoding regions $A_i \subset Y_1^n$ and $B_j \subset Y_2^n$ for $i = 1, 2, \ldots, M_1; j = 1, 2, \ldots, M_2$, such that the average probabilities of error $\epsilon_1$ and $\epsilon_2$ for the first and second receiver obey

$$\epsilon_1 = \frac{1}{M_1 M_2} \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} \omega_1^n(A_i^n | x_1(i), x_2(j)) \leq \lambda$$

$$\epsilon_2 = \frac{1}{M_1 M_2} \sum_{j=1}^{M_2} \sum_{i=1}^{M_1} \omega_2^n(B_j^n | x_1(i), x_2(j)) \leq \lambda.$$  

Definition 3. $(R_1, R_2) \in \mathbb{R}_+^2$ is an approachable rate for the interference channel if for any $\eta > 0, 0 < \lambda < 1$ and for any sufficiently large $n$, there exists a $(n, M_1, M_2, \lambda)$ code such that $\log M_1/n \geq R_1 - \eta$ and $\log M_2/n \geq R_2 - \eta$.

The set of all approachable rates for an interference channel is called the capacity region, denoted by $C_{IC}$. A region in which all points are approachable is an achievable region. Using Definition 2, we can easily verify the following lemma.

Lemma 1. Let $l_1 \leq L_1, j \leq L_2, k \leq N_1', m \leq N_2'$ be index integers. If there exist $n$-sequences $x_1(i, k), x_2(j, m)$ and regions $A_{ikm} \subset Y_1^n, B_{jkm} \subset Y_2^n$ such that

$$\frac{1}{L_1' N_1' L_2' N_2'} \sum_{k=1}^{N_1'} \sum_{m=1}^{N_2'} \sum_{i=1}^{L_1'} \sum_{j=1}^{L_2'} \omega_1^n(A_{ikm} | x_1(i, k), x_2(j, m)) \leq \lambda$$

$$\frac{1}{L_1' N_1' L_2' N_2'} \sum_{k=1}^{N_1'} \sum_{m=1}^{N_2'} \sum_{i=1}^{L_1'} \sum_{j=1}^{L_2'} \omega_2^n(B_{jkm} | x_1(i, k), x_2(j, m)) \leq \lambda$$

then there exists a $(n, L_1' N_1' L_2' N_2', \lambda)$ code for $[X_1, X_2, Y_1, Y_2, \omega(y_1, y_2|x_1, x_2)]_{IC}$.

Proof. Substituting $A_{ik}$ with $\cup_m A_{ikm}$ and $B_{jm}$ with $\cup_k B_{jkm}$ in equation (2), we obtain (1) with the index mapping $ik \rightarrow i, jm \rightarrow j$ and $M_1 = L_1' N_1', M_2 = L_2' N_2'$. 

□
Lemma 2. Consider $Y \in \mathcal{Y}$, $U \in \mathcal{U}$, $V \in \mathcal{V}$, $W \in \mathcal{W}$ and $Q \in \mathcal{Q}$ with $P(Y, U, V, W, Q) = P(Q)P(U|Q)P(V|Q)P(W|Q)P(Y|U, V, W, Q)$. Define random matrices $U \in \mathcal{U}^{M_1 \times n}$, $V \in \mathcal{V}^{M_2 \times n}$, $W \in \mathcal{W}^{M_3 \times n}$ and vector $Q \in \mathcal{Q}^n$ with

$$P(U, V, W, Q) = P(Q)P(U|Q)P(V|Q)P(W|Q),$$

where $P(Q) = \prod_i P(Q_i)$, $P(U|Q) = \prod_{i,j} P(U_{ij}|Q_j)$ and similar for $P(W|Q)$ and $P(V|Q)$. We call $(U, V, W)$ a random codebook (codewords $u(i), i = 1, \ldots, M_1$ etc. are the rows). For fixed codebook $(u, v, w, q)$, let $\varepsilon(u, v, w, q)$ be the minimum probability of error over selection of decoding regions for a time-varying memoryless three-user multiple-access channel with transition probability $P(Y|U, V, W, q_t)$ at time $t$.

$\varepsilon \equiv \mathbb{E}[\varepsilon]$ can be arbitrary small by sufficiently large $n$ if $\log M_i/n = R_i - \eta$, $i = 1, 2, 3$ for some $\eta > 0$, and some $(R_1, R_2, R_3)$ satisfying

$$R_1 \leq I(Y; U | V, W, Q),$$
$$R_1 \leq I(Y; V | U, W, Q),$$
$$R_1 \leq I(Y; W | U, V, Q),$$
$$R_1 + R_2 \leq I(Y; U, V | W, Q),$$
$$R_2 + R_3 \leq I(Y; V, W | U, Q),$$
$$R_1 + R_2 + R_3 \leq I(Y; U, V, W | Q).$$

Proof. Follow the typical sequence proof technique of the main theorem in [3].

III. INTERFERENCE CHANNEL CAPACITY REGION

Here we define a sequence of regions $C_i \subseteq C_{i+1}$, where $C_1$ is the Han and Kobayashi region [3]. We will show in Theorem 1 that $C_1$ is achievable for any $l$ and that in the limit $l \to \infty$ it is the capacity region.

Definition 4. For $[\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, \omega(y_1, y_2 | x_1, x_2)]_{\text{SC}}$ define the region $C_1$ as follows.

1. Let $\tilde{Y}_1 \in \mathcal{Y}_1, \tilde{Y}_2 \in \mathcal{Y}_2$. It will be useful to interchangeably think of these objects as random $l$-vectors, or random variables with cardinality $|\mathcal{Y}_i|^l$. Thus $\tilde{Y}_1 = (\tilde{Y}_{11}, \ldots, \tilde{Y}_{1n})$ shall be an $n$-sequence of these vectors. $Y_1 = (\tilde{Y}_{11}, \tilde{Y}_{12}, \ldots, \tilde{Y}_{1l})$. These notations will also be used for other random variables, $\tilde{X}_1$ etc.

2. Define auxiliary random variables $W_1 \in \mathcal{W}_1, W_2 \in \mathcal{W}_2, U_1 \in U_1, U_2 \in U_2, Q \in \mathcal{Q}$ with cardinalities $|W_1| = (|X_1| + l), |U_1| = (|X_1| + 2)^l$, $|U_2| = (|X_1| + 2)^l$, $i = 1, 2$ and $|Q| = 11$.

3. Let $\mathcal{P}$ be the set of all distributions $P(\tilde{Y}_1, \tilde{Y}_2, W_1, U_1, U_2, Q)$ for which there exists functions $f_1 : U_1 \times W_1 \times Q \to \mathcal{X}_1^l$ and $f_2 : U_2 \times W_2 \times Q \to \mathcal{X}_2^l$ such that

$$P(\tilde{Y}_1, \tilde{Y}_2, W_1, U_1, U_2, Q) = P(Q)P(U_1 | Q)P(U_2 | Q)P(W_1 | Q)P(W_2 | Q)P(U_1 | Q)P(U_2 | Q)$$

$$\omega(\tilde{Y}_1, \tilde{Y}_2, f_1(U_1, W_1, Q), f_2(U_2, W_2, Q))$$

4. For any $P \in \mathcal{P}$ define $\mathcal{R}(P) \subseteq \mathbb{R}_+^l$ as the set of pairs $(R_1, R_2)$ such that $(R_1, R_2) = (S_1 + T_1, S_2 + T_2)$ for some $(S_1, T_1, S_2, T_2) \in \mathbb{R}_+^4$ satisfying

$$S_1 + T_1 \leq I(\tilde{Y}_1; U_1 | W_1, Q)/l,$$
$$T_1 \leq I(\tilde{Y}_1; W_1 | U_1, Q)/l,$$
$$S_1 \leq I(\tilde{Y}_1; W_1 | U_1, Q)/l,$$
$$S_1 + T_1 + T_2 \leq I(\tilde{Y}_1; W_1, W_2 | U_2, Q)/l,$$
$$S_2 \leq I(\tilde{Y}_2; U_2 | W_2, Q)/l,$$
$$T_2 \leq I(\tilde{Y}_2; W_2 | U_2, Q)/l,$$
$$T_2 \leq I(\tilde{Y}_2; W_2, U_2 | Q)/l,$$
$$T_1 + T_2 \leq I(\tilde{Y}_2; W_2, W_1 | U_1, Q)/l,$$
$$S_2 + T_2 \leq I(\tilde{Y}_2; W_1, U_2 | Q)/l,$$
$$S_2 + T_1 \leq I(\tilde{Y}_2; U_2, W_2 | Q)/l.$$

5. $C_1 = \text{closure}(\bigcup_{P \in \mathcal{P}} \mathcal{R}(P))$

Theorem 1. $\lim_{l \to \infty} C_1 = C_{IC}$

Proof (direct): It suffices to show that for each $P \in \mathcal{P}$ and any $(S_1, T_1, S_2, T_2)$ satisfying (3), any $\lambda > 0$ and all sufficiently large $n$, there exists a $(n, L_1^1 N_1^1, L_1^2 N_1^2, \lambda)$ code for the channel $[\mathcal{X}_1^l, \mathcal{X}_2^l, \mathcal{Y}_1^l, \mathcal{Y}_2^l, \zeta(\tilde{y}_1, \tilde{y}_2 | \tilde{x}_1, \tilde{x}_2)]_{\text{SC}}$, where $\zeta(\tilde{y}_1, \tilde{y}_2 | \tilde{x}_1, \tilde{x}_2) = \omega'(\tilde{y}_1, \tilde{y}_2 | \tilde{x}_1, \tilde{x}_2)$ (recalling that $\tilde{y}_1$, etc. are in fact l-vectors), such that for some $\eta > 0$

$$\log L_1/n = S_1 - \eta,$$
$$\log N_1/n = T_1 - \eta.$$
Let the functions $f_1^k : \mathcal{Y}_1^k \times \mathcal{W}_1^k \times \mathcal{Q}_2 \to \mathcal{X}_1^k$ and $f_2^k : \mathcal{Y}_2^k \times \mathcal{W}_2^k \times \mathcal{Q}_2 \to \mathcal{X}_2^k$ be defined by coordinate-wise application of $f_1$ and $f_2$. For given $q$, $w_1$, $w_2$, $u_1$ and $u_2$ let

$$
\epsilon_1 = \min_{A_{ikm} \subset \mathcal{Y}_1^k \times \mathcal{W}_1^k \times \mathcal{Q}_2} \frac{1}{N_1 N_2 N_q} \sum_{i=1}^{L_1} \sum_{k=1}^{L_2} \sum_{m=1}^{N_q} f_1^k (u_1(i), w_1(k), q, f_2^k (u_2(j), w_2(m), q), q)
$$

$$
\epsilon_2 = \min_{B_{jkm} \subset \mathcal{Y}_2^k \times \mathcal{W}_2^k \times \mathcal{Q}_2} \frac{1}{N_1 N_2 N_q} \sum_{j=1}^{L_1} \sum_{k=1}^{L_2} \sum_{m=1}^{N_q} f_1^k (u_1(i), w_1(k), q, f_2^k (u_2(j), w_2(m), q), q)
$$

where $\epsilon_1 (A_{ikm})$ and $\epsilon_2 (B_{jkm})$ are the marginal transmission probability matrices.

Let $\epsilon_1 = E[\epsilon_1]$ and $\epsilon_2 = E[\epsilon_2]$ and $\epsilon_1 (q) = E[\epsilon_1 | q]$ and $\epsilon_2 (q) = E[\epsilon_2 | q]$, according to (5). We can show that

$$
\sum_{u_2} P(u_2 | q) \frac{1}{L_q} \sum_{i=1}^{L_1} \omega_1^k (A_{ikm}^c | u_1(i), w_1(k), q, f_2^k (u_2(j), w_2(m), q)) = \omega_1^k (A_{ikm} | u_1(i), w_1(k), w_2(m), q)
$$

where

$$
P(a | q) \omega_1^k (y_1 | u_1(i), w_1(k), q, f_2^k (u_2(j), w_2(m), q)) = \prod_{i=1}^{L_1} P(a_i | q_i) \text{ (note } a_i \in U_2).
$$

Therefore

$$
\epsilon_1 (q) = \sum_{u_1} \sum_{u_2} P(w_1 | q) P(w_2 | q) \frac{1}{L_q} \sum_{i=1}^{L_1} \omega_1^k (A_{ikm}^c | u_1(i), w_1(k), w_2(m), q)
$$

Moreover,

$$
\omega_1^k (y_1 | u_1(i), w_1(k), w_2(m), q) = \prod_{t=1}^{n} \omega_1 (y_{1t} | u_{1it}, w_{1kt}, w_{2mt}, q_t),
$$

where

$$
\omega_1 (y_{1t} | u_{1t}, w_{1t}, w_{2t}, q) = \sum_{u_2} P(u_2 | q) \omega_1^k (y_{1t} | u_1(i), w_1(k), q, f_2 (u_2, w_2, q)).
$$

Therefore $\epsilon_1 (q)$ is the average probability of error for a memoryless time-varying three-user multiple-access channel having transition probability $\omega_1 (y_{1t} | u_{1t}, w_{1t}, w_{2t}, q_t)$ at time $t$. Applying Lemma 2, $\epsilon_1$ can be bounded arbitrarily if $\log L_1 / n = S_1' - \eta$, $\log L_2 / n = T_1' - \eta$ and $\log N_q / n = T_2 - \eta$ for some $\eta > 0$ and some $(S_1', T_1', T_2')$ satisfying:

$$
S_1' \leq I(Y_1; U_1 | W_1, W_2, Q)
$$

$$
T_1' \leq I(Y_1; W_1 | U_1, W_2, Q)
$$

$$
T_1' \leq I(Y_1; U_1 | W_1, W_2, Q)
$$

$$
S_1' + T_1' \leq I(Y_1; U_1, W_1 | W_2, Q)
$$

$$
S_1' + T_1' \leq I(Y_1; U_1, W_1 | W_2, Q)
$$

(10)

for the distribution

$$
P_1 (Y_1, W_1, W_2, U_1, Q) = P(Q) P(W_1 | Q) P(W_2 | Q) P(U_1 | Q)
$$

$$
\times \sum_{u_2} P(u_2 | Q) \omega_1 (Y_1 | f_1 (u_1, W_1, Q), f_2 (u_2, W_2, Q)).
$$

(11)

Equivalently, $\epsilon_1$ can be bounded arbitrarily if (4) is satisfied for some $\eta > 0$ and some $(S_1, T_1, T_2)$ satisfying the first seven inequalities in (3) for the distribution (11). Following an entirely similar procedure, $\epsilon_2$ can be bounded arbitrary if (4) is satisfied for some $\eta > 0$ and some $(S_2, T_1, T_2)$ satisfying the last seven inequalities in (3), for the distribution

$$
P_2 (Y_2, W_1, W_2, U_2, Q) = P(Q) P(W_1 | Q) P(W_2 | Q) P(U_2 | Q)
$$

$$
\times \sum_{u_1} P(u_1 | Q) \omega_2 (Y_2 | f_1 (u_1, W_1, Q), f_2 (u_2, W_2, Q)).
$$

(12)

But $P_1$ and $P_2$ are both marginals of the same given distribution $P \in \mathcal{P}$. So for $P \in \mathcal{P}$ and any $(S_1, T_1, S_2, T_2)$ satisfying (3), $\epsilon_1$, $\epsilon_2$ can be bounded arbitrarily for sufficiently large $n$, if for some $\eta > 0$, (4) is satisfied.

$\epsilon_1 \leq \lambda / 2$ and $\epsilon_2 \leq \lambda / 2$ implies existence of $q$, $w_1$, $w_2$, $u_1$, $u_2$ such that $\epsilon_1 \leq \lambda$ and $\epsilon_2 \leq \lambda$ and therefore existance of sequences $X_1(i, k) = f_1^k (u_1(i), w_1(k), q)$, and $X_2(j, m) = f_2^j (u_2(j), w_2(m), q)$ for $1 \leq i \leq L_1, 1 \leq j \leq L_2, 1 \leq k \leq N_1, 1 \leq m \leq N_2$ satisfying (2) (by $L_1' = L_1, \cdots$) and as a result of Lemma 1, the existence of an $(n, L_1 N_1, L_2 N_2, \lambda)$ code for $[X_1', X_2', Y_1', Y_2', \gamma_1, \gamma_2 | x_1, x_2]_{10}c$.

Proof (converse): Consider $P_c \subset \mathcal{P}$, defined by all distributions $(Y_1, Y_2, W_1, W_2, U_1, U_2, Q)$ that can be factored as $P(Q)P(W_1)P(W_2)P(U_1)P(U_2)\omega_1(Y_1, Y_2 | f_1 (U_1, W_1, Q), f_2 (U_2, W_2, Q))$ where $P(W_1), P(W_2), P(Q)$ have all their probability mass in one point (say $w_1', w_2', q_\lambda$). For such distributions $\mathcal{R}(P)$ is defined by $R_c \leq I(Y_1; U_2)/l$ for $i = 1, 2$.

Introduce random vectors $\tilde{X}_1$, and $\tilde{X}_2$ joint with the other random variables by $\tilde{x}_1 = f_1 (u_1, w_1', q_\lambda)$ and $\tilde{x}_2 = f_2 (u_2, w_2', q_\lambda)$. We have the Markov chains
\( \hat{Y}_1 \rightarrow U_1 \rightarrow \hat{X}_1 \) and \( \hat{Y}_2 \rightarrow U_2 \rightarrow \hat{X}_2 \), and hence
\[ I(\hat{Y}_i; \hat{X}_i) \leq I(\hat{Y}_i; U_i) \] for \( i = 1, 2 \) where the left hand side mutual informations are on \( P_1(\hat{X}_1)P_2(\hat{X}_2)\omega^j(\hat{Y}_1, \hat{Y}_2 | \hat{X}_1, \hat{X}_2) \) and \( P_1(\hat{X}_1)P_2(\hat{X}_2) \) are distributions induced by \( P(U_1)P(U_2) \) and the functions \( \tilde{x}_1 = f_1(u_1, w^*_1, q^*) \) and \( \tilde{x}_2 = f_2(u_2, w^*_2, q^*) \).

Because \( |X^*_1| < |U_1| \) and \( |X^*_2| < |U_2| \), there exist functions \( f_1(u_1, w^*_1, q^*) \) and \( f_2(u_2, w^*_2, q^*) \) such that the set of distributions \( P(\hat{X}_1)P(\hat{X}_2) \) ranges over all possible distributions \( P(\hat{X}_1)P(\hat{X}_2) \), for all possible distributions \( P(U_1)P(U_2) \).

Letting \( \mathcal{R}^i(P(\tilde{x}_1), P(\tilde{x}_2)) \) be the region \( R_i \leq I(\hat{Y}_i; \hat{X}_i)/l, \ i = 1, 2 \) for \( P(\hat{X}_1)P(\hat{X}_2)\omega^j(\hat{Y}_1, \hat{Y}_2 | \hat{X}_1, \hat{X}_2) \), we have:

\[
\lim_{l \to \infty} \bigcup_{P(\tilde{x}_1), P(\tilde{x}_2)} \mathcal{R}^i(P(\tilde{x}_1), P(\tilde{x}_2)) \subseteq \lim_{l \to \infty} \bigcup_{P \in \mathcal{P}_c} \mathcal{R}(P) \subseteq \lim_{l \to \infty} \bigcup_{P \in \mathcal{P}} \mathcal{R}(P). \quad (13)
\]

But the left hand side is just another limiting expression for the capacity region \([1]\). \hfill \Box

Acknowledgments

This work is financially supported by The Sir Ross and Sir Keith Smith Center for Aviation Operation Research.

References


