Discrete Memoryless Interference Channel: New Outer Bound

Raul Etkin
Hewlett-Packard Laboratories
Palo Alto, CA 94304
Email: raul.etkin@hp.com

Erik Ordentlich
Hewlett-Packard Laboratories
Palo Alto, CA 94304
Email: erik.ordentlich@hp.com

***Abstract*— A new outer bound for the two-user discrete memoryless interference channel is presented. This bound establishes the capacity region of the binary erasure interference channel, whose determination was left open in [6]. The new bound is compared with the best known outer bounds for some additional examples. It is also shown that the new outer bound is tight for a one-sided deterministic interference channel that belongs to the class studied by El Gamal and Costa.**

I. Introduction

The two-user interference channel (IFC) is an information-theoretic model that captures the effects of interference between two users that communicate independent information to their corresponding receivers. The capacity region of the channel gives a fundamental measure of the tradeoff between the rates that can be simultaneously achieved. Unfortunately, except for a few special cases, the capacity region is unknown [1], [2]. There are two IFC models that received particular attention in the literature: the discrete memoryless IFC (DMIFC) and the Gaussian memoryless IFC (GMIFC). Recently, the capacity region of the Gaussian IFC has been characterized with an error smaller than 1 bit/s/Hz per user through the derivation of new outer bounds [3]. However, the DMIFC lacks such a tight characterization for its capacity region. In this work we make progress in the characterization of the DMIFC capacity region by presenting a new outer bound. This bound establishes the capacity region of the binary erasure IFC, whose determination was left open in [5], [6].

The best known achievable region for both the DMIFC and the GMIFC is due to Han and Kobayashi [8]. In order to characterize capacity, it is of interest to obtain outer bounds to the capacity region, such as the ones proposed by Sato [9] and Carleial [10]. These bounds are obtained by allowing cooperation between the transmitters or receivers, and by enhancing the channel so that the actual channel outputs are degraded versions of the enhanced channel outputs. In the enhanced channel, each receiver is able to decode the messages sent by both transmitters. However, in many cases, the receivers in the actual channel cannot completely decode the message sent by the interfering transmitter. In these cases, the bounds of [9] and [10] are loose. In contrast, the bound that we present in this work does not require either receiver to completely decode the interfering signal. In addition to determining the capacity region of the binary erasure IFC, we present some examples that illustrate the performance of the new bound. In particular, the new bound appears to be most useful for channels with only one-sided interference.

Regarding notation we use lowercase letters to denote scalars, uppercase letters to denote random variables, boldface letters to denote vectors, and calligraphic uppercase letters to denote sets. For example x is a scalar, X is a random variable, $\mathbf{x}$ is a vector, $\mathbf{X}_1^n = (X_1, \ldots, X_n)$ is a random vector, and $\mathcal{X}$ is a set. We use $H(\cdot)$, and $I(\cdot;\cdot)$ to denote binary entropy and mutual information. The indicator function is denoted by $1(\cdot)$, and $(x)^+$ is used to denote $\max\{x, 0\}$. The probability mass function (pmf) of the random variable X is written as p_X and, with slight abuse of notation, we use $p(x)$ to denote this function. We use $\mathcal{B}(p)$ to denote the Bernoulli distribution of parameter p. We define $h_i(p_1, \ldots, p_{i-1}) = -\sum_{j=1}^{i} p_j \log_2(p_j)$ for $p_j \geq 0, \sum_{j=1}^{i} p_j = 1$, i.e. the entropy in bits of a random variable with pmf $(p_1, \ldots, p_i)$.

II. Model

We consider a two-user discrete memoryless interference channel. In this model there are two transmitter-receiver pairs, where each transmitter wants to communicate with its corresponding receiver (cf. Figure 1). This channel is defined by the finite input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, the finite output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$, and transition probability matrices $\{p(y_1|x_1, x_2)\}$ and $\{p(y_2|x_1, x_2)\}$. Without loss of generality we assume $\mathcal{X}_1 = \{1, \ldots, k_1\}$ and $\mathcal{X}_2 = \{1, \ldots, k_2\}$.

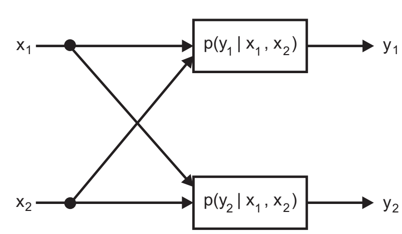

Fig. 1. Two-user discrete memoryless interference channel.

For a given block length n, user i communicates a message $m_i \in \{1, \ldots, 2^{nR_i}\}$ by choosing a codeword from a codebook $\mathcal{C}_{i,n}$, with $|\mathcal{C}_{i,n}| = 2^{nR_i}$. The codewords $\{\mathbf{c}_i(m_i)\}$ of this codebook have symbols chosen from the input alphabet $\mathcal{X}_i$. Receiver i observes the channel outputs $\{y_{i,t} : t = 1, \ldots, n\}$ and uses a decoding function $f_{i,n} : \mathcal{Y}_i^n \to \mathbb{N}$ to get the estimate

$\hat{m}_i$ of the transmitted message m_i. The receiver is in error whenever $\hat{m}_i \neq m_i$. The average probability of error for user i is given by

$$\epsilon_{i,n} = E[P(\hat{m}_i \neq m_i)]$$

where the expectation is taken with respect to the random choice of the transmitted messages m_1 and m_2. Note that due to the interference among users, the probability of error of each user may depend on the codebook of the other user.

A rate pair (R_1, R_2) is achievable if there exists a family of codebook pairs $\{(\mathcal{C}_{1,n}, \mathcal{C}_{2,n})\}_n$ and decoding functions $\{(f_{1,n}(\cdot), f_{2,n}(\cdot)\}_n$, such that the average decoding error probabilities $\epsilon_{1,n}, \epsilon_{2,n}$ go to zero as the block length n goes to infinity.

The capacity region $\mathcal{R}$ of the interference channel is the closure of the set of achievable rate pairs.

III. New Outer Bound

Theorem 1: Let U and Q_i, $i = 1, 2$, be auxiliary random variables with U chosen from the alphabet $\mathcal{U} = \{1, \dots, 7\}$ and Q_i chosen from the alphabets $\mathcal{Q}_i = \{1, \dots, k_i\}$, such that:

$$\begin{aligned} p(x_1, x_2, y_1, y_2, q_1, q_2, u) = p(u)p(q_1|u)p(x_1|q_1, u)p(q_2|u) \\ \cdot p(x_2|q_2, u)p(y_1|x_1, x_2)p(y_2|x_1, x_2) \end{aligned}$$

and for fixed $p_U, p_{Q_1|U}, p_{Q_2|U}, p_{X_1|Q_1,U}, p_{X_2|Q_2,U}$ let

$$\begin{aligned} r_1 =& I(X_1; Y_1|X_2, U), \quad r_2 = I(X_2; Y_2|X_1, U) \\ f_1(\alpha, \beta) =& (1 - \beta)I(Q_2, X_1; Y_1|U) + \beta I(X_1; Y_1|X_2, U) \\ &+ \alpha H(X_2|Q_2, U) \\ f_2(\alpha, \beta) =& (1 - \beta)I(Q_1, X_2; Y_2|U) + \beta I(X_2; Y_2|X_1, U) \\ &+ \alpha H(X_1|Q_1, U). \end{aligned}$$

For arbitrary functions $\beta_1, \beta_2 : \mathbb{R}^+ \to [0, 1]$ define the rate region $\tilde{\mathcal{R}}(\beta_1, \beta_2)$ as:

$$\begin{aligned} \tilde{\mathcal{R}}(\beta_1, \beta_2) = \big\{&(R_1, R_2) : 0 \leq R_1 \leq r_1, 0 \leq R_2 \leq r_2, \\ &R_1 + \alpha R_2 \leq f_1(\alpha, \beta_1(\alpha)), \alpha R_1 + R_2 \leq f_2(\alpha, \beta_2(\alpha)), \\ &\text{for all } \alpha > 0\big\}. \end{aligned}$$

The rate region:

$$\mathcal{R}_{OB}(\beta_1, \beta_2) = \text{Closure} \bigcup_{\substack{p_U, p_{Q_1|U}, p_{Q_2|U}, \\ p_{X_1|Q_1,U}, p_{X_2|Q_2,U}}} \tilde{\mathcal{R}}(\beta_1, \beta_2) \qquad (1)$$

is an outer bound to the capacity region of the discrete memoryless interference channel, i.e. $\mathcal{R} \subseteq \mathcal{R}_{OB}(\beta_1, \beta_2)$.

Proof: First we derive the bounds on the individual rates R_1 and R_2. For any code with block-length n, using Fano's inequality we can write:

$$\begin{aligned} R_1 \leq& \frac{1}{n} I(\mathbf{X}_1^n; \mathbf{Y}_1^n) + \epsilon_n \overset{(a)}{\leq} \frac{1}{n} I(\mathbf{X}_1^n; \mathbf{Y}_1^n, \mathbf{X}_2^n) + \epsilon_n \\ \overset{(b)}{=}& \frac{1}{n} I(\mathbf{X}_1^n; \mathbf{Y}_1^n|\mathbf{X}_2^n) + \epsilon_n \\ =& \frac{1}{n} \left[H(\mathbf{Y}_1^n|\mathbf{X}_2^n) - H(\mathbf{Y}_1^n|\mathbf{X}_1^n, \mathbf{X}_2^n)\right] + \epsilon_n \\ =& \frac{1}{n} \sum_{i=1}^n \big[H(Y_{1,i}|\mathbf{Y}_1^{i-1}, \mathbf{X}_2^n) \\ &- H(Y_{1,i}|\mathbf{Y}_1^{i-1}, \mathbf{X}_1^n, \mathbf{X}_2^n)\big] + \epsilon_n \\ \overset{(c)}{=}& \frac{1}{n} \sum_{i=1}^n \big[H(Y_{1,i}|\mathbf{Y}_1^{i-1}, \mathbf{X}_2^n) - H(Y_{1,i}|X_{1,i}, X_{2,i})\big] + \epsilon_n \\ \overset{(d)}{\leq}& \frac{1}{n} \sum_{i=1}^n \big[H(Y_{1,i}|X_{2,i}) - H(Y_{1,i}|X_{1,i}, X_{2,i})\big] + \epsilon_n \\ =& \frac{1}{n} \sum_{i=1}^n I(X_{1,i}; Y_{1,i}|X_{2,i}) + \epsilon_n \end{aligned}$$

where (a) follows from the non-negativity of mutual information, (b) follows from the independence between $\mathbf{X}_1^n$ and $\mathbf{X}_2^n$, (c) follows from the memoryless property of the channel, (d) follows from the fact that conditioning reduces entropy, and $\epsilon_n \to 0$ as $n \to \infty$. The upper bound r_2 on R_2 can be obtained in a similar way.

We next bound $R_1 + \alpha R_2$. We will use the following inequality:

$$\begin{aligned} H(\mathbf{Y}_1^n|\mathbf{X}_1^n) =& \sum_{i=1}^n H(Y_{1,i}|\mathbf{Y}_1^{i-1}, \mathbf{X}_1^n) \\ \overset{(a)}{\geq}& \sum_{i=1}^n H(Y_{1,i}|\mathbf{X}_2^{i-1}, \mathbf{Y}_1^{i-1}, \mathbf{X}_1^n) \\ \overset{(b)}{=}& \sum_{i=1}^n H(Y_{1,i}|\mathbf{X}_2^{i-1}, X_{1,i}) \end{aligned} \qquad (2)$$

where (a) follows from the fact that extra conditioning reduces conditional entropy, and (b) follows from the equality

$$p(Y_{1,i}|\mathbf{X}_2^{i-1}, \mathbf{Y}_1^{i-1}, \mathbf{X}_1^n) = p(Y_{1,i}|\mathbf{X}_2^{i-1}, X_{1,i})$$

which is a consequence of the memoryless property of the channel.

For any codebook of block-length n, and non-negative β we can write:

$$\begin{aligned} R_1 + \alpha R_2 \leq& \frac{1}{n} \big[I(\mathbf{X}_1^n; \mathbf{Y}_1^n) + \alpha I(\mathbf{X}_2^n; \mathbf{Y}_2^n)\big] + \epsilon_n \\ \overset{(a)}{\leq}& \frac{1}{n} \big[H(\mathbf{X}_1^n) - H(\mathbf{X}_1^n|\mathbf{Y}_1^n) \\ &+ \alpha H(\mathbf{X}_2^n) + \beta I(\mathbf{X}_1^n; \mathbf{X}_2^n|\mathbf{Y}_1^n)\big] + \epsilon_n \\ =& \frac{1}{n} \big\{H(\mathbf{X}_1^n) - H(\mathbf{X}_1^n|\mathbf{Y}_1^n) + \alpha H(\mathbf{X}_2^n) \\ &+ \beta[H(\mathbf{X}_1^n|\mathbf{Y}_1^n) - H(\mathbf{X}_1^n|\mathbf{X}_2^n, \mathbf{Y}_1^n)]\big\} + \epsilon_n \\ =& \frac{1}{n} \big\{(1 - \beta)\big[H(\mathbf{Y}_1^n) - H(\mathbf{Y}_1^n|\mathbf{X}_1^n)\big] + \beta H(\mathbf{X}_1^n) \\ &+ \alpha H(\mathbf{X}_2^n) - \beta H(\mathbf{X}_1^n|\mathbf{X}_2^n, \mathbf{Y}_1^n)\big\} + \epsilon_n \\ =& \frac{1}{n} \big\{(1 - \beta)\big[H(\mathbf{Y}_1^n) - H(\mathbf{Y}_1^n|\mathbf{X}_1^n)\big] \\ &+ \beta I(\mathbf{X}_1^n; \mathbf{X}_2^n, \mathbf{Y}_1^n) + \alpha H(\mathbf{X}_2^n)\big\} + \epsilon_n \\ =& \frac{1}{n} \big\{(1 - \beta)\big[H(\mathbf{Y}_1^n) - H(\mathbf{Y}_1^n|\mathbf{X}_1^n)\big] \\ &+ \beta I(\mathbf{X}_1^n; \mathbf{Y}_1^n|\mathbf{X}_2^n) + \alpha H(\mathbf{X}_2^n)\big\} + \epsilon_n \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n}\big\{(1-\beta)\big[H(\mathbf{Y}_1^n) - H(\mathbf{Y}_1^n|\mathbf{X}_1^n)\big] \\
&\quad + \beta\big[H(\mathbf{Y}_1^n|\mathbf{X}_2^n) - H(\mathbf{Y}_1^n|\mathbf{X}_1^n, \mathbf{X}_2^n)\big] + \alpha H(\mathbf{X}_2^n)\big\} + \epsilon_n \\
&\overset{(b)}{\leq} \frac{1}{n}\sum_{i=1}^{n}\big\{(1-\beta)\big[H(Y_{1,i}) - H(Y_{1,i}|\mathbf{X}_2^{i-1}, X_{1,i})\big] \\
&\quad + \beta\big[H(Y_{1,i}|X_{2,i}) - H(Y_{1,i}|X_{1,i}, X_{2,i})\big] \\
&\quad + \alpha H(X_{2,i}|\mathbf{X}_2^{i-1})\big\} + \epsilon_n \\
&\overset{(c)}{=} \frac{1}{n}\sum_{i=1}^{n}\big\{(1-\beta)H(Y_{1,i}) + \beta I(X_{1,i}; Y_{1,i}|X_{2,i}) \\
&\quad + \big[\alpha H(X_{2,i}|Q_i) - (1-\beta)H(Y_{1,i}|Q_i, X_{1,i})\big]\big\} + \epsilon_n
\end{aligned} \tag{3}
$$

where we added a non-negative mutual information term in (a), and (b) follows from (2), the fact that removing conditioning increases entropy, and the memoryless property of the channel. In step (c) we defined the scalar random variables Q_i with alphabets $\mathcal{Q}_i = \{1, \ldots, |\mathcal{Q}_i|\}$, $|\mathcal{Q}_i| = k_2^{i-1}$, to represent the vectors $\mathbf{X}_2^{i-1}$. Using Carathéodory's theorem [4] and the procedure of [11] one can restrict attention to auxiliary random variables chosen from an alphabet with cardinality $|\mathcal{Q}_i| \leq k_2$. With this cardinality bound, we can let $n \to \infty$ in (3) so that $\epsilon_n \to 0$.

The bound $f_2(\alpha, \beta)$ on $\alpha R_1 + R_2$ can be obtained in a similar way as $f_1(\alpha, \beta)$.

The time sharing random variable U, and the closure and union operations in (1) allow to replace the sums $(1/n)\sum_{i=1}^{n}(\cdot)$ by single letter expressions, and consider every possible distribution on U, Q_1, Q_2, X_1 and X_2 that the codebooks might induce. A standard argument for bounding the cardinality of time sharing random variables can be used here with $U \in \mathcal{U}$, to show that it suffices to take $|\mathcal{U}| \leq 7$. ■

The individual bounds of Theorem 1 are defined in terms of single letter expressions that can be computed with tractable complexity when the alphabet sizes $|\mathcal{X}_1|$ and $|\mathcal{X}_2|$ are small. The individual rate bounds of Theorem 1 can be interpreted as the maximum single user rate that results when a genie provides signal x_i to receiver j ($i \neq j$). The bounds $f_i(\alpha, \beta)$, $i = 1, 2$ require the use of an auxiliary random variable Q_i whose alphabet size can be restricted to $|\mathcal{X}_i|$. The role of Q_i is to capture the effect of the memory in the input codeword $\mathbf{x}_i$. An interesting aspect of the bounds of Theorem 1 is that the alphabet size of these auxiliary random variables can be made quite small.

Another novel feature of the bound $f_1(\alpha, \beta)$ ($f_2(\alpha, \beta)$) is the addition of a non-negative term $\beta I(\mathbf{X}_1^n; \mathbf{X}_2^n|\mathbf{Y}_1^n)$ ($\beta I(\mathbf{X}_1^n; \mathbf{X}_2^n|\mathbf{Y}_2^n)$), where $\beta \in [0, 1]$ can be optimized for each choice of α. At first glance, it is somewhat surprising to see that the addition of a non-negative term can result in a tighter bound. A more traditional approach to deriving upper bounds is to enhance the channel through the aid of a genie that provides side information to one or both receivers. The help of the genie can usually be explained through a non-negative term that is added to the mutual information expressions that determine the upper bound. In our bounds, we can "optimize" the genie for each choice of α.

IV. The Capacity Region of the Binary Erasure Interference Channel

The binary erasure interference channel has input and output alphabets $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y}_1 = \mathcal{Y}_2 = \{0, 1\}$ and is defined by the equations: $y_1 = x_1 \cdot x_2$, $y_2 = x_2$. We see that receiver 2 observes a clean version of x_2, but whenever $x_2 = 0$ the symbol x_1 is "erased" at y_1.

This channel is a deterministic interference channel: the outputs are deterministic functions of the inputs. However, this simple channel does not belong to the class studied in [7], because given x_1, y_1 does not completely determine x_2 (i.e. when $x_1 = 0$, y_1 does not provide any information about x_2).

We specialize the bound of Theorem 1 to this channel. We will only consider the bounds r_2 and $f_1(\alpha, \beta)$ with $\beta = 0$. As will be seen, these bounds are enough to characterize the capacity region. Using $Y_2 = X_2$ we can write:

$$
r_2 = I(X_2; X_2|X_1, U) = H(X_2|X_1, U) \leq H(X_2) \leq 1 \tag{4}
$$

and for any joint distribution of the form given in Theorem 1 we have:

$$
\begin{aligned}
&f_1(\alpha, 0) = I(Q_2, X_1; Y_1|U) + \alpha H(X_2|Q_2, U) \\
&= \sum_{u=1}^{7} p(u)\big[H(Y_1|u) - \sum_{x_1=0}^{1} p(x_1|u)H(Y_1|x_1, Q_2, u) \\
&\quad + \alpha H(X_2|Q_2, u)\big] \\
&\overset{(a)}{=} \sum_{u=1}^{7} p(u)\big[H(Y_1|u) + (\alpha - p(X_1 = 1|u))H(X_2|Q_2, u)\big] \\
&\overset{(b)}{\leq} \max_{u \in \{1,\ldots,7\}} \max_{0 \leq p_1, p_2 \leq 1} \big\{h_2(p_1 \cdot p_2) \\
&\quad + \max_{\substack{0 \leq p_2^a, p_2^b, q \leq 1 \\ p_2^a(1-q) + p_2^b q = p_2}} (\alpha - p_1)\big[(1-q)h_2(p_2^a) + q \cdot h_2(p_2^b)\big]\big\} \\
&\overset{(c)}{=} \max_{0 \leq p_1, p_2 \leq 1} \big\{h_2(p_1 \cdot p_2) + (\alpha - p_1)^+ h_2(p_2)\big\}
\end{aligned} \tag{5}
$$

where (a) is obtained by noting that when $X_1 = 0$, $Y_1 = 0$ and hence the conditional entropy of Y_1 is zero, and when $X_1 = 1$, $Y_1 = X_2$. (b) is obtained by optimizing over the distributions $(X_1|u) \sim \mathcal{B}(p_1)$, $(X_2|u) \sim \mathcal{B}(p_2)$, $(X_2|Q_2 = 0, u) \sim \mathcal{B}(p_2^a)$, $(X_2|Q_2 = 1, u) \sim \mathcal{B}(p_2^b)$, $Q_2 \sim \mathcal{B}(q)$, satisfying the constraint that the marginal distribution of $(X_2|Q_2, u)$ coincides with the distribution of $(X_2|u)$. Finally (c) is obtained by noting that when $a - p_1 < 0$ the solution of the innermost maximization is zero.

When $\alpha > 1/2$ the value of p_1 that optimizes (5) satisfies $p_1^* < \alpha$. Therefore when $\alpha > 1/2$ we obtain:

$$
f_1(\alpha, 0) \leq \max_{0 \leq p_1, p_2 \leq 1} \big\{h_2(p_1 \cdot p_2) + (\alpha - p_1)h_2(p_2)\big\}.
$$

In addition, as $\alpha \to 1/2$ from above, the optimal value $p_2^* \to 1$, while when $\alpha \to \infty$, $p_2^* \to 1/2$. For $1/2 < \alpha < \infty$, $p_2^*(\alpha)$ is a continuous and decreasing function of α.

The bound $R_1 + \alpha R_2 \leq f_1(\alpha, 0)$, true for any α, implies that for fixed $0 < R_2 < 1$ and $\alpha^*(R_2)$ satisfying $h_2(p_2^*(\alpha(R_2))) = R_2$

$$
R_1 \leq f_1(\alpha^*(R_2), 0) - \alpha^*(R_2)R_2
$$

$$\leq \max_{0 \leq p_1 \leq 1} h_2(p_1 \cdot p_2^*(\alpha^*(R_2))) - p_1 R_2. \quad (6)$$

We will show that for any $0 < R_2 < 1$, this upper bound on R_1 is achievable. Fix R_2 and consider a communication scheme in which user 2 communicates with a codebook of block-length n generated from independent and identically distributed (i.i.d.) samples of a $\mathcal{B}(p_2)$ random variable, where $h_2(p_2) = R_2$ with $1/2 < p_2 < 1$. This codebook has 2^{nR_2} codewords. When user 2 communicates using this codebook the vectors $\mathbf{x}_2^n$ have a distribution such that its components are i.i.d. $\mathcal{B}(p_2)$. The resulting communication channel for user 1 is a binary memoryless channel where the input with value 0 is transmitted with no change, but the input with value 1 is converted to a 0 with probability p_2. The capacity of this channel is:

$$C_1(p_2) = \max_{0 \leq p_1 \leq 1} I(X_1;Y_1) = \max_{0 \leq p_1 \leq 1} h(p_1 \cdot p_2) - p_1 h_2(p_2)$$

which can be achieved by user 1 as $n \to \infty$. Therefore, the upper bound (6) can be achieved. It follows that for $0 < R_2 < 1$ we have a characterization of the maximum achievable R_1. Note that $0 \leq R_2 \leq 1$. It follows that the closure of the region

$$\mathcal{R} = \{(R_1, R_2) : 1/2 < p_2 < 1, R_2 = h_2(p_2), \\ R_1 \leq \max_{0 \leq p_1 \leq 1} h(p_1 \cdot p_2) - p_1 h_2(p_2)\} \quad (7)$$

is the capacity region of the binary erasure interference channel (see Figure 2).

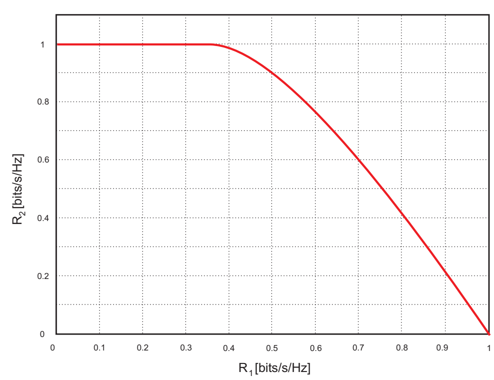

Fig. 2. Capacity region of the binary erasure interference channel.

V. Other Examples

In order to illustrate how the bound of Theorem 1 improves upon the existing outer bounds, but also to understand its limitations we present some additional examples.

1) A one-sided channel with multiplicative interference: Consider a deterministic interference channel with $\mathcal{X}_1 = \{0,1,2\}$, $\mathcal{X}_2 = 1, 2$, $\mathcal{Y}_1 = \{0,1,2,4\}$, $\mathcal{Y}_2 = \{1,2\}$, defined by the equations: $y_1 = x_1 \cdot x_2$, $y_2 = x_2$.

In this channel, receiver 2 observes a clean version of the input signal x_2, while y_1 is affected by interference. In addition, since $y_1 = 0$ whenever $x_1 = 0$ regardless of the value of x_2, this channel does not belong to the class of deterministic interference channels studied in [7].

Before specializing Theorem 1 to this channel, we derive some inner and outer bounds for the capacity region using previous results.

Consider a communication scheme where user 1 communicates using a codebook generated using i.i.d. samples of a random variable with probability mass function (pmf) p_{X_1} with $p_{X_1}(0) = p_0$, $p_{X_1}(1) = p_1$, $p_{X_1}(2) = 1 - p_0 - p_1$ and user 2 communicates using a codebook generated using i.i.d. samples of a random variable with pmf p_{X_2} with $p_{X_2}(1) = 1 - p_2$, $p_{X_2}(2) = p_2$. User 2 can communicate reliably at a rate $R_2 = h_2(p_2)$ as $n \to \infty$. If user 1 treats the interference from user 2 as noise, he can communicate reliably at a rate:

$$R_1 = I(X_1;Y_1) \\ = h_4(p_0, p_1(1-p_2), (1-p_0-p_1)p_2) - (1-p_0)h_2(p_2).$$

As we vary $0 \leq p_0, p_1, p_2 \leq 1$ we get the inner bound corresponding to the scheme that treats interference as noise.

Another inner bound results from requiring that user 1 be able to completely decode the message transmitted by user 2. In this case, the achievable rates are given by the capacity region of the multiple access channel formed by both transmitters and receiver 1. This capacity region is:

$$\mathcal{R}_{MAC_1} = \Big\{(R_1, R_2) : R_1 \leq I(X_1;Y_1|X_2), \\ R_2 \leq I(X_2;Y_1|X_1), R_1 + R_2 \leq I(X_1, X_2;Y_1), \\ \text{for some } (p_0, p_1, p_2) \in [0,1]^3\Big\} \quad (8)$$

where $I(X_1;Y_1|X_2) = h_3(p_0, p_1)$, $I(X_2;Y_1|X_1) = (1 - p_0)h_2(p_2)$, and $I(X_1, X_2;Y_1) = h_4(p_0, p_1(1-p_2), (1-p_0 - p_1)p_2)$.

Furthermore, by time-sharing between treating interference as noise and decoding user 2's message at receiver 1, we can achieve a region that is equal to the convex hull of the union of both achievable regions. We call this region $\mathcal{R}_1$.

With more generality, we could consider a scheme of the type proposed by Han and Kobayashi in [8]. This scheme consists of splitting the information to be sent into two messages: a private message to be decoded only by the intended receiver, and a common message to be decoded by both receivers. In the channel that we are considering, y_2 is independent of x_1 so user 1 can only send private information. Let u_2 and w_2 denote the private and common messages of user 2, and let U_2 and W_2 be the random variables used to generate the codebooks of the Han-Kobayashi (HK) scheme. In addition, let Q be a time sharing random variable used to time-share between multiple schemes. According to Theorem 4.1 of [8] it is enough to consider random variables chosen from alphabets with cardinalities:

$$|\mathcal{U}_2| \leq |\mathcal{X}_2| + 2 = 4, \ |\mathcal{U}_2| \leq \mathcal{X}_2 + 7 = 9, \ |\mathcal{Q}| \leq 11.$$

Furthermore, to compute the HK achievable region we need to consider all possible functions $f : (\mathcal{U}_2 \times \mathcal{W}_2) \to \mathcal{X}_2$ (i.e. 2^{36} functions) that encode the private and common messages into the message transmitted in the channel. It follows that even for the relatively simple channel that we are considering, computing the HK achievable region is a complex task.

We limit our analysis to the following HK scheme. We choose $U_2 \sim \mathcal{B}(u)$, $W_2 \sim \mathcal{B}(w)$, and use the encoding function $x_2 = u_2 \oplus w_2 + 1$, where $\oplus$ denotes modulo 2 addition. After varying $(u, w, p_0, p_1) \in [0,1]^4$ we obtain an achievable region matching $\mathcal{R}_1$. We conjecture that the HK achievable region is equal to $\mathcal{R}_1$.

We can obtain an outer bound to the capacity region by considering a modified interference channel in which a genie provides side information $s_1 = x_2 \cdot 1(x_1 = 0)$ to receiver 1. Since x_2 is a function of (x_1, y_1, s_1), this genie-aided channel belongs to the class of deterministic channels studied in [7], and its capacity region can be explicitly computed. The resulting outer bound is:

$$\begin{aligned}\mathcal{R}_{OB_1} = \big\{&(R_1, R_2) : R_1 \le h_3(p_0, p_1), R_2 \le h_2(p_2),\\ &R_1 + R_2 \le h_4[p_0, p_1(1-p_2), (1-p_0-p_1)p_2]\\ &\quad + p_0 h_2(p_2), \text{ for some } (p_0, p_1, p_2) \in [0,1]^3\big\}.\end{aligned}$$

The same outer bound can be obtained from the bound of [10] Theorem 1, which is more general than the bound of Sato [9].

We specialize Theorem 1 to this channel. We only need to consider the individual constraints r_1, r_2, and $f_1(\alpha,\beta)$, since adding the constraint $f_2(\alpha,\beta)$ does not improve the outer bound. We obtain:

$$\begin{aligned}r_1 &\le \max_{0\le p_0,p_1\le 1} h_3(p_0,p_1) = \log(3)\\ r_2 &\le \max_{0\le p_2\le 1} h_2(p_2) = 1\\ f_1(\alpha,\beta) &\le \max_{0\le p_0,p_1,p_2\le 1} (1-\beta)h_4[p_0, p_1(1-p_2),\\ &\quad (1-p_0-p_1)p_2] + \beta h_3(p_0,p_1)\\ &\quad + (\alpha - (1-\beta)(1-p_0))^+ h_2(p_2)\end{aligned}$$

where the optimization that determines $f_1(\alpha,\beta)$ must be computed numerically for each value of α and β. Furthermore, we can optimize $\beta(\alpha)$ to obtain the smallest outer bound.

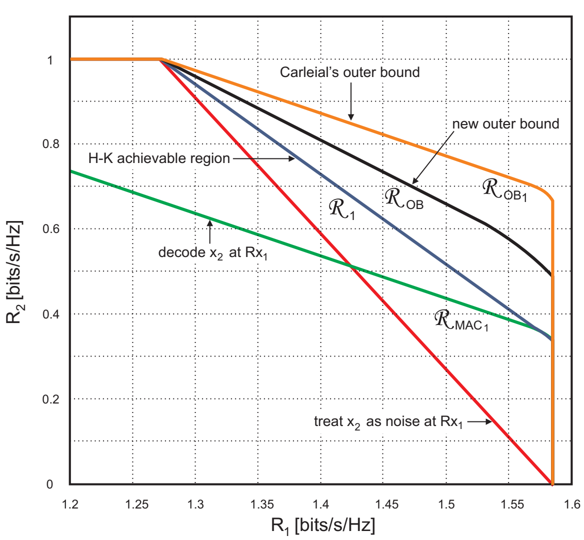

Fig. 3. Achievable region and outer bounds. Note that the axes are drawn in different scales, and the horizontal axis starts at 1.2 [bits/s/Hz].

Figure 3 shows the inner bound $\mathcal{R}_1$ and the outer bounds $\mathcal{R}_{OB_1}$ and $\mathcal{R}_{OB}$. We see that for this channel, the new outer bound $\mathcal{R}_{OB}$ strictly improves on the existing bound $\mathcal{R}_{OB_1}$.

2) The one-sided El Gamal-Costa IFC: Consider a deterministic interference channel defined by:

$$v_1 = g_1(x_1)\ ,\ v_2 = g_2(x_2)\ ,\ y_1 = f_1(x_1, v_2)\ ,\ y_2 = f_2(x_2, v_1)$$

where $g_1(\cdot)$ and $g_2(\cdot)$ are arbitrary functions, while for fixed x_1 (x_2), $f_1(x_1,\cdot)$ ($f_2(x_2,\cdot)$) is injective. The capacity region of this channel has been obtained in [7].

We will show that for this special class of channels, when $y_2 = x_2$ (i.e. $f_2(x_2, v_1) = x_2$) Theorem 1 is tight. We can compute the different bounds of Theorem 1 for this channel for specific values of α and β:

$$\begin{aligned}r_1 &= H(Y_1|X_2) = H(Y_1|V_2,X_2) \le H(Y_1|V_2)\\ r_2 &= H(Y_2|X_1) = H(Y_2)\\ f_1(1,0) &= H(Y_1) + H(X_2|Q_2) - H(Y_1|Q_2,X_1)\\ &= H(Y_1) + H(X_2|Q_2) - H(V_2|Q_2)\\ &= H(Y_1) + H(X_2|V_2,Q_2)\\ &\overset{(a)}{\le} H(Y_1) + H(Y_2|V_2) \qquad (9)\end{aligned}$$

where (a) follows from the assumption $y_2 = x_2$ and removing conditioning. It is easy to check that the right hand sides of (9) coincide with the non-redundant bounds in [7] Theorem 1. It follows that $\mathcal{R}_{OB}(0,0)$ of Theorem 1 is the capacity region of this channel. Note that in this case the bounds of Theorem 1, $f_1(\alpha,\beta)$ for $\alpha \ne 1$, and $f_2(\alpha,\beta)$ are redundant.

REFERENCES

[1] E. C. van der Meulen, "A Survey of Multi-Way Channels in Information Theory: 1961-1976," *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 1-37, Jan. 1977.

[2] E. C. van der Meulen, "Some Reflections on the Interference Channel," *Communications and Cryptography Two Sides of One Tapestry*, pp. 409-421, Kluwer, Boston, MA, 1994, .

[3] R. Etkin, D. Tse, H, Wang, "Gaussian Interference Channel Capacity to Within One Bit: the Symmetric Case," *proceedings of the IEEE Information Theory Workshop*, Chengdu, China, October 22 - 26, 2006.

[4] H. G. Eggleston, "Convexity," *Cambridge University Press*, Cambridge, 1969.

[5] A. B. Carleial, "On the Capacity of Multiple-Terminal Communication Networks," *Technical Report No. 6603-1*, Information Systems Laboratory, Center for Systems Research, Stanford University, Aug. 1975.

[6] P. Vanroose, and E. C. van der Meulen, "Explicit Code Constructions For A Deterministic Interference Channel," *Proceedings of 1991 IEEE International Symposium on Information Theory*, p. 144, June 1991.

[7] 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.

[8] 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.

[9] H. Sato, "Two-User Communication Channels," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 295-304, May 1977.

[10] A. B. Carleial, "Outer Bounds on the Capacity of Interference Channels," *IEEE Trans. Inf. Theory*, vol. 29, no. 4, pp. 602-606, July 1983.

[11] A. Wyner, and J. Ziv, "The Rate-Distortion Function for Source Coding with Side Information at the Decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1-10, Jan. 1976.