[31] G. W. Rogers, D. J. Marchette, and C. E. Priebe, "A procedure for model complexity selection in semiparametric mixture model density estimation," Naval Surface Warfare Center, Dahlgren Division, Dahlgren, VA, Tech. Rep., 2002.
[32] A. Barron, L. Birgé, and P. Massart, "Risk bounds for model selection via penalization," *Probab. Theory Related Fields*, vol. 113, pp. 301–413, 1999.
[33] G. Castellan, "Sélection d'histogrammes à l'aide d'un critère de type Akaike," *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 330, pp. 729–732, 2000.
[34] P. Massart, "Some applications of concentration inequalities to statistics," *Ann. Fac. Sci. Toulouse Math.*, vol. 9, pp. 245–303, 2000.
[35] J. Rissanen, "On Bayesian analysis of mixtures with an unknown number of components," *Automatica*, vol. 14, pp. 465–471, 1978.
[36] ——, "Stochastic complexity (with discussion)," *J. Roy. Statist. Soc. Ser. B Statist. Methodol.*, vol. 49, pp. 223–239, 1987.
[37] J. Rissanen, T. Speed, and B. Yu, "Density estimation by stochastic complexity," *IEEE Trans. Inform. Theory*, vol. 38, pp. 315–323, Mar. 1992.
[38] L. Devroye, *A Course in Density Estimation*. Boston, MA: Birkhäuser, 1987.

Outer Bounds on the Capacity of Gaussian Interference Channels

Gerhard Kramer, *Member, IEEE*

***Abstract*—Two outer bounds on the capacity region of the two-user Gaussian interference channel (IFC) are derived. The idea of the first bound is to let a genie give each receiver just enough information to decode both messages. This bound unifies and improves the best known outer bounds of Sato and Carleial. Furthermore, the bound extends to discrete memoryless IFCs and is shown to be equivalent to another bound of Carleial. The second bound follows directly from existing results of Costa and Sato and possesses certain optimality properties for weak interference.**

***Index Terms*—Capacity, Gaussian noise, interference.**

I. Introduction

A basic open problem in Shannon theory is the capacity of the interference channel (IFC). A two-user IFC has two transmitters and two receivers, and transmitter t sends its message to receiver t for $t = 1, 2$. The transmissions interfere with each other, and this seems to make the problem hard to solve. The state-of-the-art is today basically the same as at the time of the survey article [1]. The purpose of this correspondence is to unify and improve certain capacity outer bounds.

A memoryless IFC is formally defined by the alphabets $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}_1$, $\mathcal{Y}_2$, and the conditional probability distribution $P(y_1, y_2|x_1, x_2)$ where $x_t \in \mathcal{X}_t$ and $y_t \in \mathcal{Y}_t$, $t = 1, 2$. The channel is used N times, and is *time invariant* and *memoryless* in the sense that

$$P(y_{1n}, y_{2n}|x_1^n, x_2^n, y_1^{n-1}, y_2^{n-1}) = P_{Y_1Y_2|X_1X_2}(y_{1n}, y_{2n}|x_{1n}, x_{2n}) \tag{1}$$

where x_t^n denotes the sequence $x_{t1}, x_{t2}, \ldots, x_{tn}$ and $n = 1, 2, \ldots, N$. As in (1), we use the convention of dropping subscripts on the probability distributions when the arguments are the lower case versions of the random variables. Transmitter t sends a B_t message U_t to receiver t so that the rates are $R_t = B_t/N$ bits per use. The *capacity region* $\mathcal{C}_{\mathrm{IFC}}$ is the closure of the set of rate pairs (R_1, R_2) for which the receivers can decode their messages with arbitrarily small positive error probability.

We will consider primarily the *Gaussian* IFC that has real input and output alphabets, and whose outputs are linear combinations of the transmitted symbols and white Gaussian noise. Furthermore, there are power constraints on the transmissions

$$\sum_{n=1}^{N} \mathrm{E}[X_{tn}^2]/N \le P_t, \qquad t = 1, 2. \tag{2}$$

One can reduce a Gaussian IFC to so-called *standard form* [2], i.e., the received symbols are

$$Y_1 = X_1 + c_{21}X_2 + Z_1 \tag{3}$$
$$Y_2 = c_{12}X_1 + X_2 + Z_2 \tag{4}$$

where c_{21} and c_{12} are real numbers, and Z_1 and Z_2 are independent, zero-mean, unit-variance, Gaussian random variables (see [2, Sec.V-B]).

The capacity of the standard-form Gaussian IFC is currently known only if $c_{21} = c_{12} = 0$ or if $c_{21}^2 \ge 1$ and $c_{12}^2 \ge 1$. The former case is trivial. For the latter case, called *strong interference*, $\mathcal{C}_{\mathrm{IFC}}$ is the set of (R_1, R_2) satisfying

$$0 \le R_1 \le \log(P_1 + 1)/2 \tag{5}$$
$$0 \le R_2 \le \log(P_2 + 1)/2 \tag{6}$$
$$0 \le R_1 + R_2 \le \log(P_1 + c_{21}^2 P_2 + 1)/2 \tag{7}$$
$$0 \le R_1 + R_2 \le \log(P_2 + c_{12}^2 P_1 + 1)/2. \tag{8}$$

One can prove this result by the following argument [3]–[5]: for any point (R_1, R_2) inside the capacity region both receivers will, in fact, be able to decode both messages U_1 and U_2. For example, if receiver 1 can decode U_1 it can determine X_1^N and compute

$$\widetilde{Y}_1^N = c_{12}X_1^N + X_2^N + Z_1^N/c_{21}.$$

But $\widetilde{Y}_1^N$ is *less noisy* than Y_2^N because $c_{21}^2 \ge 1$. Thus, with high probability, receiver 1 can decode U_2 if receiver 2 can decode U_2. Repeating this argument for receiver 2 we find that the problem is equivalent to the *compound* IFC, also called the compound multiple-access channel (MAC), that is discussed in [6, Sec.3, p. 14]–[8, p. 288]. The capacity region of this channel is given by (5)–(8). For historical reasons, and because of the remarkable nature of the capacity region, one sometimes distinguishes between *strong* and *very strong* interference. The interference is said to be very strong if the bounds (7) and (8) can be dropped, i.e., if $c_{21}^2 \ge P_1 + 1$ and $c_{12}^2 \ge P_2 + 1$. We refer to [1], [4], [5], [9] for a more detailed discussion of such cases.

This correspondence is organized as follows. In Section II, we review the existing capacity outer bounds for Gaussian IFCs with $0 < c_{21}^2 < 1$ or $0 < c_{12}^2 < 1$. In Section III, we derive an improved bound. Section IV develops another bound that is even better. In Section V, we extend the first bound to discrete memoryless IFCs and recover a bound proposed by Carleial in [10, Theorem 1]. Finally, Section VI suggests possible extensions of our ideas.

II. Existing Capacity Outer Bounds

Two outer bounds on $\mathcal{C}_{\mathrm{IFC}}$ were described by Sato [11] and Carleial [10]. Sato's bound applies to both discrete memoryless and Gaussian IFCs.

Manuscript received July 19, 2001; revised January 28, 2003. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Washington, DC, June 2001.

The author is with Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974 USA (e-mail: gkr@bell-labs.com).

Communicated by I. E. Telatar, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2004.825249

0018-9448/04$20.00 © 2004 IEEE

Proposition 1 (Sato): $\mathcal{C}_{\mathrm{IFC}}$ is contained within the convex hull of the set of rate pairs (R_1, R_2) satisfying

$$R_1 \le I(X_1; \widetilde{Y}_1 | X_2) \tag{9}$$
$$R_2 \le I(X_2; \widetilde{Y}_2 | X_1) \tag{10}$$
$$R_1 + R_2 \le I(X_1 X_2; \widetilde{Y}_1 \widetilde{Y}_2) \tag{11}$$

where X_1 and X_2 are independent, and $\widetilde{Y}_1$ and $\widetilde{Y}_2$ are any random variables satisfying

$$P_{\widetilde{Y}_1|X_1X_2}(y_1|x_1, x_2) = P(y_1|x_1, x_2)$$

and

$$P_{\widetilde{Y}_2|X_1X_2}(y_2|x_1, x_2) = P(y_2|x_1, x_2)$$

for all x_1, x_2, y_1, y_2.

Proof: $\mathcal{C}_{\mathrm{IFC}}$ depends only on $P(y_1|x_1, x_2)$ and $P(y_2|x_1, x_2)$ so one can replace Y_1 and Y_2 by $\widetilde{Y}_1$ and $\widetilde{Y}_2$ [11]. The bounds (9) and (10) follow by giving receivers 1 and 2 the inputs X_2 and X_1, respectively. The bound (11) follows by letting the receivers cooperate. The convex hull operation is needed by arguments similar to [8, p. 277]. □

For the Gaussian channel, $\widetilde{Y}_1$ and $\widetilde{Y}_2$ are the respective Y_1 and Y_2 with Z_1 and Z_2 replaced by *arbitrarily correlated*, zero-mean, unit-variance, Gaussian random variables $\widetilde{Z}_1$ and $\widetilde{Z}_2$. Furthermore, the maximum-entropy theorem [12, p. 234] lets one restrict attention to zero-mean Gaussian X_1 and X_2 with variances P_1 and P_2, respectively.

Proposition 2 (Carleial): If $0 < c_{21}^2 \le 1$ for the Gaussian IFC then $\mathcal{C}_{\mathrm{IFC}}$ is contained within the set of (R_1, R_2) satisfying

$$R_1 \le \log(P_1 + 1)/2 \tag{12}$$
$$R_2 \le \log(P_2 + 1)/2 \tag{13}$$
$$R_1 + R_2 \le \log(P_1/c_{21}^2 + P_2 + 1)/2. \tag{14}$$

If $0 < c_{12}^2 \le 1$, then one can add the sum-rate bound obtained by swapping the indexes 1 and 2 in (14).

Proof: The bounds (12) and (13) are valid even if a genie gives receivers 1 and 2 the respective X_2 and X_1. The bound (14) follows by replacing Y_1 in (3) by the improved $\widetilde{Y}_1 = X_1 + c_{21}X_2 + c_{21}Z_1$. Converting this channel to standard form, receiver 1 sees $\widetilde{Y}_1/c_{21} = \widetilde{X}_1 + X_2 + Z_1$, where $\widetilde{X}_1 = X_1/c_{21}$ so that transmitter 1's power is P_1/c_{21}^2. This new channel to receiver 1 has strong interference and applying (7) we obtain (14). Finally, if $0 < c_{12}^2 \le 1$, we use the same type of argument to swap the indexes 1 and 2 in (14). □

It turns out that neither of these two bounds implies the other—in some cases Sato's bound is tighter than Carleial's and in other cases the opposite is true.

A third useful bound applies to the special class of *degraded* IFCs [11, Sec.V], [13], [14]. A degraded IFC has the outputs

$$Y_1 = X_1 + X_2 + Z_1 \tag{15}$$
$$Y_2 = X_1 + X_2 + Z_2 \tag{16}$$

where Z_1 and Z_2 are independent, zero-mean, Gaussian random variables with respective variances N_1 and N_2. Observe that this channel is *not* in standard form. The following bound was presented by Sato in [13] and is tight when $N_1 \ge N_2$ and $R_2 = \log(1 + P_2/N_2)/2$.

Proposition 3 (Sato): For a degraded Gaussian IFC with $N_1 \ge N_2$, the capacity region $\mathcal{C}_{\mathrm{IFC}}$ is contained within the set of (R_1, R_2) satisfying

$$0 \le R_1 \le \log[1 + P_1'/(P_2' + N_1)]/2 \tag{17}$$
$$0 \le R_2 \le \log(1 + P_2'/N_2)/2 \tag{18}$$

where $P_1' + P_2' = P_1 + P_2$.

Proof: Suppose we enlarge $\mathcal{C}_{\mathrm{IFC}}$ by allowing the transmitters to cooperate, in the sense that they can control $X_1 + X_2$ under the sum power constraint $P_1 + P_2$. We then have a degraded broadcast channel [15], [16] whose capacity region is given by (17) and (18). □

Finally, we describe a result of Costa [9]. First, by setting $c_{12} = 0$ one enlarges $\mathcal{C}_{\mathrm{IFC}}$. (To see this, observe that if $c_{12} = 0$ then receiver 2 can choose one of transmitter 1's codewords at random, multiply this codeword by $c_{12}' \ne 0$, and add the result to its received signal to simulate the $c_{12}' \ne 0$ case.) Next, by a series of manipulations, one can show that the resulting "Z-interference" channel has the same capacity region as a degraded IFC. The following proposition summarizes the result.

Proposition 4 (Costa): The capacity region of a standard-form Gaussian IFC with $c_{12} = 0$ and $c_{21} > 0$ is the same as the capacity region of a degraded Gaussian IFC with P_1 changed to P_1/c_{21}^2, P_2 left unchanged, $N_1 = 1/c_{21}^2$, and $N_2 = 1$.

Proof: See [9, Fig. 6]. □

III. An Improved Outer Bound

Our first bound improves on Propositions 1 and 2.

Theorem 1: $\mathcal{C}_{\mathrm{IFC}}$ is contained within the set of (R_1, R_2) satisfying (12), (13), and

$$R_1 + R_2 \le \frac{1}{2}\log\left[\left(P_1 + c_{21}^2 P_2 + 1\right)\left(\frac{P_2 + 1}{\min(c_{21}^2, 1)P_2 + 1}\right)\right]. \tag{19}$$

One can further add the sum-rate bound obtained by swapping the indexes 1 and 2 in (19).

We give two different proofs of this theorem. The first proof is based on a genie-aided argument, and the second is based on linear minimum mean-squared error (LMMSE) estimation.

A. A Genie-Aided Proof

Suppose that (R_1, R_2) and N are such that receivers 1 and 2 can correctly decode U_1 and U_2 with high probability from Y_1^N and Y_2^N, respectively. Suppose further that a genie gives receiver 1 the following side information:

$$\widetilde{Y}_1 = d_1X_1 + d_2X_2 + d_3Z_1 + d_4\widetilde{Z}_1 \tag{20}$$

where d_1, d_2, d_3, d_4 are some real numbers and $\widetilde{Z}_1$ is a zero-mean, unit-variance, Gaussian random variable independent of all other random variables. By hypothesis, receiver 1 can determine X_1^N. Receiver 1 can further attempt to decode U_2 by forming the following approximation to Y_2:

$$\begin{aligned}\widehat{Y}_2 &= \alpha Y_1 + \beta\widetilde{Y}_1 + (c_{12} - \alpha - \beta d_1)X_1 \\ &= c_{12}X_1 + (\alpha c_{21} + \beta d_2)X_2 + (\alpha + \beta d_3)Z_1 + (\beta d_4)\widetilde{Z}_1\end{aligned} \tag{21}$$

where α and β are some real numbers chosen so that $\alpha c_{21} + \beta d_2 = 1$. Assuming that c_{21}^2 is positive, one can simply choose $\alpha = (1 - \beta d_2)/c_{21}$ so that

$$\begin{aligned}&\widehat{Y}_2 = c_{12}X_1 + X_2 + \widehat{Z}_2, \text{ where} \\ &\widehat{Z}_2 = \left(c_{21}^{-1} + \beta\left(d_3 - d_2c_{21}^{-1}\right)\right)Z_1 + (\beta d_4)\widetilde{Z}_1.\end{aligned} \tag{22}$$

We now invoke the same type of argument used to determine $\mathcal{C}_{\mathrm{IFC}}$ with strong interference. Suppose that $\widehat{Y}_2$ is at most as noisy as Y_2, i.e., the noise $\widehat{Z}_2$ has variance 1 or less. In this case, receiver 1 can correctly decode U_2 with high probability (recall that receiver 2 can correctly decode U_2 with high probability by hypothesis) and one has a MAC from X_1 and X_2 to Y_1. We can therefore apply a standard outer bound on the sum rate for a Gaussian MAC [11, Sec. III].

Lemma 1: If $c_{21}^2 > 0$ then $\mathcal{C}_{\mathrm{IFC}}$ is contained within the set of (R_1, R_2) satisfying

$$R_1 + R_2 \leq \min_{\substack{d_1, d_2 \\ d_3, d_4}} I(X_1 X_2; Y_1 \widetilde{Y}_1) \tag{23}$$

where X_1 and X_2 are independent, zero-mean, Gaussian random variables with $\mathrm{E}[X_t^2] \leq P_t$ for $t = 1, 2$, $\widetilde{Y}_1$ is given by (20) and, for some real β

$$\left(c_{21}^{-1} + \beta\left(d_3 - d_2 c_{21}^{-1}\right)\right)^2 + (\beta d_4)^2 \leq 1. \tag{24}$$

Before optimizing (23) we again consider Sato's and Carleial's bounds.

Lemma 2: If $c_{21}^2 > 0$ then the sum-rate bound of Proposition 1 is a special case of (23) in Lemma 1. If $0 < c_{21}^2 \leq 1$ then the sum-rate bound of Proposition 2 is also a special case of (23) in Lemma 1.

Proof: Consider the case $d_1 = c_{12}$ and $d_2 = 1$ so that

$$\widetilde{Y}_1 = c_{12} X_1 + X_2 + d_3 Z_1 + d_4 \widetilde{Z}_1.$$

Choosing $\beta = 1$ simplifies (24) to $d_3^2 + d_4^2 \leq 1$. This means that $d_3 Z_1 + d_4 \widetilde{Z}_1$ can have unit variance and, moreover, can be made arbitrarily correlated with Z_1. We thus have the sum-rate bound of Proposition 1 as a special case of (23).

Consider next the case $d_1 = d_2 = 0$ and $d_4 = 1$ so that $\widetilde{Y}_1 = d_3 Z_1 + \widetilde{Z}_1$. The bound (23) becomes

$$R_1 + R_2 \leq \log\left[\left(P_1 + c_{21}^2 P_2 + 1\right)\left(d_3^2 + 1\right) - d_3^2\right]/2 \tag{25}$$

and (24) simplifies to

$$(c_{21}^{-1} + \beta d_3)^2 + \beta^2 \leq 1.$$

The tightest bound results when $\beta = \pm\sqrt{1 - c_{21}^2}$, $d_3 = \mp\sqrt{c_{21}^{-2} - 1}$, and

$$R_1 + R_2 \leq \log\left(P_1/c_{21}^2 + P_2 + 1\right)/2. \tag{26}$$

We thus have the sum-rate bound of Proposition 2 as a special case of (23). □

We proceed to optimize (23) with respect to d_1, d_2, d_3, d_4, and β. First, it is clear that (23) is minimized when (24) holds with equality. The resulting quadratic in β is

$$\beta^2[(d_2 - c_{21} d_3)^2 + c_{21}^2 d_4^2] - 2\beta(d_2 - c_{21} d_3) + (1 - c_{21}^2) = 0. \tag{27}$$

Suppose that $d_4 \neq 0$ so that the β^2 coefficient is positive. In this case, β has a real solution if the quadratic's discriminant is nonnegative. Equivalently, we require that

$$(d_2 - c_{21} d_3)^2 \geq d_4^2\left(1 - c_{21}^2\right). \tag{28}$$

Next, if $d_4 \neq 0$, the information $I(X_1 X_2; Y_1 \widetilde{Y}_1)$ is

$$\frac{1}{2}\log\left[\left(P_1 + c_{21}^2 P_2 + 1\right)\left(d_1^2 P_1 + d_2^2 P_2 + d_3^2 + d_4^2\right) - (d_1 P_1 + c_{21} d_2 P_2 + d_3)^2\right] - \frac{1}{2}\log\left(d_4^2\right). \tag{29}$$

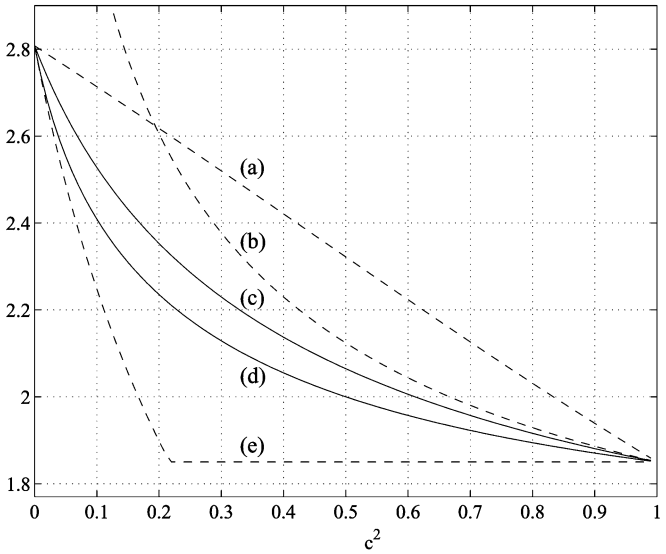


Fig. 1. Bounds on $\boldsymbol{R}_1 + \boldsymbol{R}_2$ for $\boldsymbol{P}_1 = \boldsymbol{P}_2 = \mathbf{6}$ and $\mathbf{0} \leq \boldsymbol{c}^2 \leq \mathbf{1}$. *(a):* Sato's upper bound. *(b):* Carleial's upper bound. *(c):* Upper bound of Theorem 1. *(d):* Upper bound of Theorem 2. *(e):* Lower bound. The rate units are bits per channel use. Note the $\boldsymbol{y}$-axis range.

Minimizing (29) with respect to d_1, we have

$$d_1 = \frac{c_{21} d_2 P_2 + d_3}{c_{21}^2 P_2 + 1}. \tag{30}$$

Inserting (30) into (29) we have

$$\frac{1}{2}\log\left[\left(P_1 + c_{21}^2 P_2 + 1\right)\left(1 + \frac{P_2(d_2 - c_{21} d_3)^2}{d_4^2(c_{21}^2 P_2 + 1)}\right)\right]. \tag{31}$$

Note that we can set $d_1 = 0$ by choosing $d_3 = -c_{21} d_2 P_2$, and that this does not restrict the optimization of (31). Thus, we could have refrained from using X_1 in (20). This is not surprising because the bounding technique assumes that receiver 1 can determine X_1 anyway. However, it is interesting that $d_1 = 0$ is *not* always best, i.e., including X_1 in $\widetilde{Y}_1$ can, in general, reduce the information of receiver 1.

If $c_{21}^2 \geq 1$, then (28) is true for any choice of d_2, d_3, d_4. The information (31) is thus minimized by choosing $d_2 = d_3 = 0$. If $0 < c_{21}^2 < 1$, then (31) is minimized by choosing d_2, d_3, and d_4 so that equality holds in (28). For example, one can set $d_4 = 1$, $d_3 = -c_{21} d_2 P_2$ (which makes $d_1 = 0$), and

$$d_2 = \frac{\sqrt{1 - c_{21}^2}}{c_{21}^2 P_2 + 1}. \tag{32}$$

This proves Theorem 1.

Note that Theorem 1 is true for *any* c_{21}. Furthermore, all rate-pairs satisfying (12), (13) and (19) are approachable at receiver 1 when $c_{21}^2 \geq 1$ [4], [5]. Fig. 1 shows what happens when $0 \leq c_{21}^2 \leq 1$ and $P_1 = P_2 = 6$ (note the y-axis range); the dashed curves are taken from [10, Fig. 4]. Note that Theorem 1 improves noticeably on the other two outer bounds.

Fig. 2 shows the bounds for $P_1 = P_2 = 6$ and $c^2 := c_{21}^2 = c_{12}^2 = 0.3$. The rate units are bits per channel use. The inner bound was formed by drawing lines between the three points marked with crosses—these points are taken from [2] (see also [9]).

B. A Proof Based on LMMSE Estimation

Suppose we make $\widetilde{Y}_1$ the error of the LMMSE estimate of Y_2 given X_1 and Y_1. We have

$$\widetilde{Y}_1 = Y_2 - \mathrm{E}\left[\begin{bmatrix} X_1 & Y_1 \end{bmatrix} \quad Y_2\right] \cdot Q^{-1}_{\begin{bmatrix} X_1 & Y_1 \end{bmatrix}} \cdot \begin{bmatrix} X_1 \\ Y_1 \end{bmatrix} \tag{33}$$

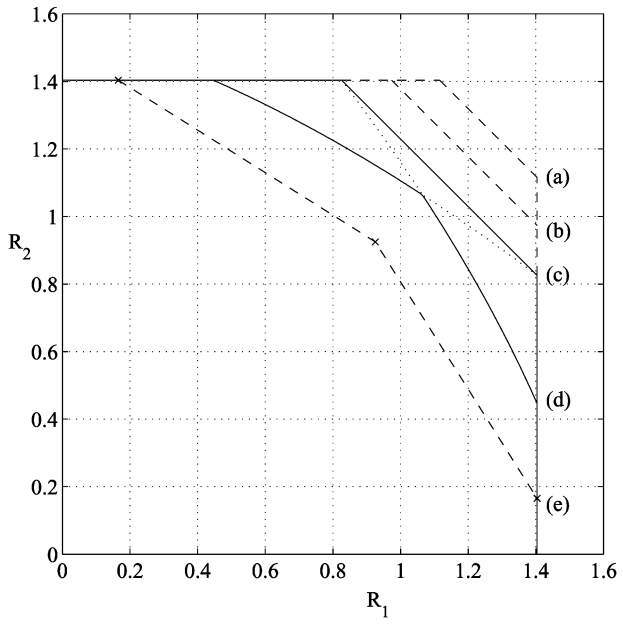


Fig. 2. Bounds on $\mathcal{C}_{\mathrm{IFC}}$ for $P_1 = P_2 = 6$ and $c^2 = 0.3$: *(a):* Sato's outer bound. *(b):* Carleial's outer bound. *(c):* Outer bound of Theorem 1. *(d):* Outer bound of Theorem 2. *(e):* Inner bound. The rate units are bits per channel use.

where

$$Q_{[X_1 \quad Y_1]} = \mathrm{E}\left[[X_1 \quad Y_1]^T [X_1 \quad Y_1]\right]$$

and T denotes transposition (see [17, Ch. 12]). Receiver 1 can obviously compute Y_2 from Y_1, X_1, and $\widetilde{Y}_1$. Furthermore, the orthogonality principle ensures that $\mathrm{E}[\widetilde{Y}_1 \quad [X_1 \quad Y_1]] = [0 \quad 0]$, and the maximum entropy theorem [12, p. 234] ensures that Gaussian X_1 and X_2 maximize the mutual information in (23). We thus choose X_1 and X_2 to be Gaussian random variables and compute

$$\begin{aligned} I(X_1X_2; Y_1\widetilde{Y}_1) &= I(X_1X_2; Y_1) + I(X_2; \widetilde{Y}_1|X_1Y_1) \\ &= I(X_1X_2; Y_1) + I(X_2; Y_2|X_1Y_1) \\ &= \frac{1}{2}\log\left(V_{Y_1} \cdot V_{Y_2|X_1Y_1}/V_{Z_2|Z_1}\right) \end{aligned} \tag{34}$$

where V_A is the variance of A and $V_{A|B}$ is the variance of A given B. Since $V_{Y_1} = P_1 + c_{21}^2 P_2 + 1$ it remains to minimize $V_{Y_2|X_1Y_1}/V_{Z_2|Z_1}$. Let $E[Z_1Z_2] = \rho$, $|\rho| \leq 1$, so that

$$\begin{aligned} \frac{V_{Y_2|X_1Y_1}}{V_{Z_2|Z_1}} &= \frac{V_{(X_2+Z_2)|(c_{21}X_2+Z_1)}}{V_{Z_2|Z_1}} \\ &= \frac{P_2\left(c_{21}^2 - 2\rho c_{21} + 1\right) + (1-\rho^2)}{(1-\rho^2)\left(c_{21}^2 P_2 + 1\right)}. \end{aligned} \tag{35}$$

Minimizing with respect to ρ yields $\rho = c_{21}$ for $c_{21}^2 \leq 1$. Thus, we have

$$\frac{V_{Y_2|X_1Y_1}}{V_{Z_2|Z_1}} = \frac{P_2 + 1}{c_{21}^2 P_2 + 1}. \tag{36}$$

Inserting (36) into (34) gives Theorem 1 for $c_{21}^2 \leq 1$.

IV. A Bound Based on Degraded IFCs

Our second bound combines Propositions 3 and 4. This combination was actually described in [9, p. 612] but was applied only to investigate certain extreme points of $\mathcal{C}_{\mathrm{IFC}}$. We show that this bound is more useful and is in a sense optimal for weak interference.

Theorem 2: If $0 < c_{21}^2 \leq 1$ then $\mathcal{C}_{\mathrm{IFC}}$ is contained within the set of (R_1, R_2) satisfying

$$R_1 \leq \log[1 + P_1'/(P_2' + 1/c_{21}^2)]/2 \tag{37}$$

$$R_2 \leq \log(1 + P_2')/2 \tag{38}$$

where $P_1' + P_2' = P_1/c_{21}^2 + P_2$. If $0 < c_{12}^2 \leq 1$, then one can add the bound obtained by swapping the indexes 1 and 2 in (37) and (38).

Proof: Set $c_{12} = 0$ to enlarge the capacity region. Proposition 4 assures us that we can consider the degraded IFC with P_1 replaced by P_1/c_{21}^2 and $N_1 = 1/c_{21}^2$. Finally, apply Proposition 3 to this degraded IFC. □

Consider, for instance, the point with $R_1 = R_2$. The best choice for P_2' in Theorem 2 is then

$$P_2^* = \frac{-(1 + c_{21}^2) + \sqrt{(1 + c_{21}^2)^2 + 4c_{21}^2(P_1 + P_2 c_{21}^2)}}{2c_{21}^2} \tag{39}$$

which gives $R_1 = R_2 \leq \log(1 + P_2^*)/2$. This point gives the best sum-rate bound using Theorem 2 if $P_1 = P_2$. For example, Fig. 1 shows the resulting sum-rate bound when $P_1 = P_2 = 6$ and $c_{12}^2 = c_{21}^2 = c^2$. Note that Theorem 2 is better than Theorem 1 for all c^2. In fact, we have the following result that was kindly provided to us by Dr. Igal Sason of the Technion.

Theorem 3: The bound of Theorem 2 is better than the bound of Theorem 1 when $0 < c_{12}^2 \leq 1$ or $0 < c_{21}^2 \leq 1$.

Proof: Suppose that $0 < c_{21}^2 \leq 1$. We can choose $P_1' = P_1/c_{21}^2$ and $P_2' = P_2$ so that (37) and (38) give (19). The same approach can be used if $0 < c_{12}^2 \leq 1$. □

Next, observe from Fig. 1 that Theorem 2 seems to give the correct slope at $c^2 = 0$. To check this, we use Theorem 2 with both c_{21}^2 and c_{12}^2. For the c_{12}^2 bound, we replace P_1' and P_2' with P_1'' and P_2''. The best sum rate upper bound is determined by solving the following set of equations for P_1', P_2', P_1'', and P_2'':

$$\frac{1}{2}\log\left(1 + \frac{P_1'}{P_2' + c_{21}^{-2}}\right) = \frac{1}{2}\log\left(1 + P_1''\right) \tag{40}$$

$$\frac{1}{2}\log\left(1 + P_2'\right) = \frac{1}{2}\log\left(1 + \frac{P_2''}{P_1'' + c_{12}^{-2}}\right) \tag{41}$$

$$P_1' + P_2' = P_1/c_{21}^2 + P_2 \tag{42}$$

$$P_1'' + P_2'' = P_1 + P_2/c_{12}^2. \tag{43}$$

The result is

$$\begin{aligned} R_1 + R_2 \leq \frac{1}{2}\log\Bigg[1 + P_1 + P_2 + \frac{B}{2} \\ - \frac{1}{2}\sqrt{B^2 - 4P_1P_2\left(\frac{1}{c_{21}^2} - 1\right)\left(\frac{1}{c_{12}^2} - 1\right)}\Bigg] \end{aligned} \tag{44}$$

where

$$B = \left(\frac{1}{c_{21}^2 c_{12}^2} - 1\right) + P_1\left(\frac{1}{c_{21}^2} - 1\right) + P_2\left(\frac{1}{c_{12}^2} - 1\right). \tag{45}$$

Finally, we set $c_{12}^2 = \gamma\, c_{21}^2$ for some $\gamma > 0$ and take the derivative of the right-hand side of (44) with respect to c_{21}^2. The resulting slope at $c_{21}^2 = 0$ is

$$\frac{-P_1P_2}{2} \cdot \left(\frac{1}{1 + P_1} + \frac{\gamma}{1 + P_2}\right) \cdot \log(e). \tag{46}$$

Consider now the lower bound in Fig. 1 for small c_{21}^2. This bound is achieved by considering the interference as noise [9, p. 608], and the resulting sum rate is

$$R_1 + R_2 = \log\left[1 + P_1/\left(1 + c_{21}^2 P_2\right)\right]/2 + \log\left[1 + P_2/\left(1 + \gamma\, c_{21}^2 P_1\right)\right]/2. \quad (47)$$

We consider (47) as a function of c_{21}^2 whose slope at $c_{21}^2 = 0$ is the same as (46). We thus have the following theorem.

Theorem 4: For the Gaussian IFC with $c_{12}^2 = \gamma\, c_{21}^2$, $\gamma > 0$, the slope of the $R_1 + R_2$ capacity curve as a function of c_{21}^2 at $c_{21}^2 = 0$ is given by (46).

For example, the choice $P_1 = P_2 = 6$ and $\gamma = 1$ gives the slope -7.420 if the rate units are bits per channel use. Theorem 4 implies that Gaussian distributed inputs are basically optimal for very weak interference. This seems related to the findings of [18].

Finally, Fig. 2 shows the bound of Theorem 2 for $P_1 = P_2 = 6$ and $c^2 = 0.3$. Observe that Theorem 2 is better than Theorem 1 for all rate pairs. However, Theorem 2 does *not* include the extreme points of [9]. This suggests that some generalization of the techniques of [9] will improve on Theorem 2. For example, one could try to determine the slope of the capacity region boundary at the extreme points of [9].

V. Extension to Discrete Memoryless Interference Channels

Theorem 1 can be extended to discrete memoryless IFCs (DM-IFCs) that have discrete and finite alphabets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2$. We show that this approach gives the same bound Carleial derived in [10, Theorem 1]. Thus, Theorem 1 could have been found from the extension of [10, Theorem 1] to Gaussian IFCs.

We begin by generalizing Lemma 1 to DM-IFCs. The idea is to find a $\widetilde{Y}_1$ so that X_1, Y_1, and $\widetilde{Y}_1$ give a $\widehat{Y}_2$ that has the same statistics as Y_2. One way of making this precise is that receiver 1 inputs X_1, Y_1, and $\widetilde{Y}_1$ to a discrete memoryless channel (DMC) $P(\widehat{y}_2 \,|\, x_1, y_1, \widetilde{y}_1)$ and observes $\widehat{Y}_2$. We will use the notation $\widehat{Y}_2 \sim Y_2$ to mean that $\widehat{Y}_2$ is statistically equivalent to Y_2 given X_1 and X_2, i.e., for each y_2 in the range of Y_2 there is a $\widehat{y}_2$ in the range of $\widehat{Y}_2$ for which $P(\widehat{y}_2|x_1, x_2) = P(y_2|x_1, x_2)$ for all x_1 and x_2.

Lemma 3: $\mathcal{C}_{\mathrm{IFC}}$ is contained within the set of (R_1, R_2) satisfying

$$R_1 + R_2 \le I(X_1 X_2; Y_1 \widetilde{Y}_1) \quad (48)$$

where X_1 and X_2 are independent, $\widetilde{Y}_1$ is the output of a DMC $P(\widetilde{y}_1 \,|\, x_1, x_2, y_2)$, and there is a DMC $P(\widehat{y}_2 \,|\, x_1, y_1, \widetilde{y}_1)$ for which $\widehat{Y}_2 \sim Y_2$.

Proof: Suppose that receivers 1 and 2 can decode their messages. Receiver 1 can then form X_1^N. Suppose next that a genie gives $\widetilde{Y}_1$ to receiver 1, which then passes X_1, Y_1, and $\widetilde{Y}_1$ through $P(\widehat{y}_2 \,|\, x_1, y_1, \widetilde{y}_1)$ to obtain $\widehat{Y}_2 \sim Y_2$. Hence, receiver 1 can decode transmitter 2's message and we have a MAC to receiver 1. A standard outer bound on the sum rate for a discrete memoryless MAC [12, p. 395] yields (48). □

Lemma 3 can be combined with other standard bounds to give the following genie-aided theorem. The statistical model of this theorem is depicted in Fig. 3.

Theorem 5: $\mathcal{C}_{\mathrm{IFC}}$ is contained within the set $\mathcal{R}_G(\widetilde{Y}_1, \widetilde{Y}_2)$ of rate-pairs (R_1, R_2) satisfying

$$R_1 \le I(X_1; Y_1 | X_2 V) \quad (49)$$
$$R_2 \le I(X_2; Y_2 | X_1 V) \quad (50)$$
$$R_1 + R_2 \le I(X_1 X_2; Y_1 \widetilde{Y}_1 | V) \quad (51)$$

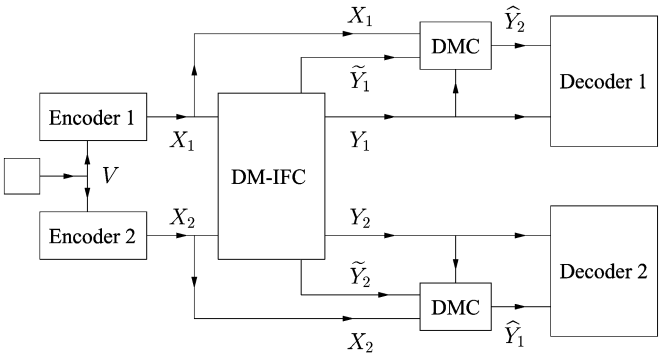


Fig. 3. The statistical model for Theorem 5.

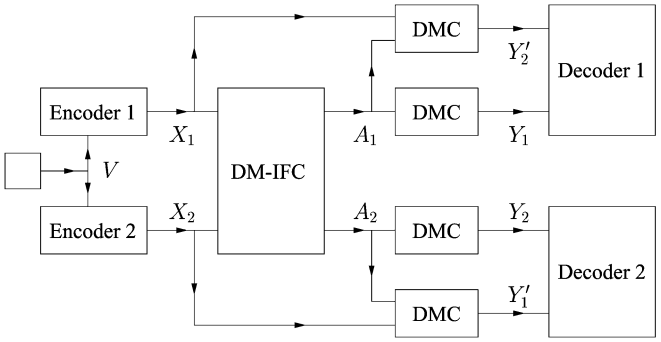


Fig. 4. The statistical model for Proposition 5.

$$R_1 + R_2 \le I(X_1 X_2; Y_2 \widetilde{Y}_2 | V) \quad (52)$$

where $\widetilde{Y}_1$ and $\widetilde{Y}_2$ are generated by some distribution $P(\widetilde{y}_1, \widetilde{y}_2 \,|\, x_1, x_2, y_1, y_2)$, there are DMCs $P(\widehat{y}_2 \,|\, x_1, y_1, \widetilde{y}_1)$ and $P(\widehat{y}_1 \,|\, x_2, y_2, \widetilde{y}_2)$ for which $\widehat{Y}_2 \sim Y_2$ and $\widehat{Y}_1 \sim Y_1$, $P(x_1, x_2 \,|\, v)$ factors as $P(x_1 \,|\, v)\, P(x_2 \,|\, v)$, and V takes on at most four values.

Proof: The DMCs assure that both receivers can decode both messages. Thus, we have a compound IFC and can apply the capacity results from [8, p. 289] with the individual rate bounds replaced by (49) and (50). The cardinality bound on V follows from [8, p. 310]. □

We next state a result from [10, Theorem 1] in a slightly different form. The difference from [10] is that we use a time-sharing random variable V rather than a convex hull operation (see [8, p. 289]). The statistical model is depicted in Fig. 4 and is basically the same as [10, Fig. 2].

Proposition 5 (Carleial): $\mathcal{C}_{\mathrm{IFC}}$ is contained within the set $\mathcal{R}_C(A_1, A_2)$ of rate-pairs (R_1, R_2) satisfying

$$R_1 \le I(X_1; Y_1 | X_2 V) \quad (53)$$
$$R_2 \le I(X_2; Y_2 | X_1 V) \quad (54)$$
$$R_1 + R_2 \le I(X_1 X_2; A_1 | V) \quad (55)$$
$$R_1 + R_2 \le I(X_1 X_2; A_2 | V) \quad (56)$$

where A_1 and A_2 are the outputs of a DM-IFC $P(a_1, a_2 \,|\, x_1, x_2)$, there are DMCs $P(y_1 \,|\, a_1)$ and $P(y_2 \,|\, a_2)$ producing the respective Y_1 and Y_2, there are DMCs $P(y_1' \,|\, x_2, a_2)$ and $P(y_2' \,|\, x_1, a_1)$ for which $Y_1' \sim Y_1$ and $Y_2' \sim Y_2$, $P(x_1, x_2 \,|\, v)$ factors as $P(x_1 \,|\, v)\, P(x_2 \,|\, v)$, and V takes on at most four values.

The motivation for introducing the four DMCs in Proposition 5 is to give both receivers enough information to decode both messages, i.e., the DMCs make the channel in Fig. 4 a compound IFC. This might lead

one to suspect that Proposition 5 and Theorem 5 are closely related. In fact, they turn out to be equivalent.

Theorem 6:

$$\mathcal{R}_G(\widetilde{Y}_1, \widetilde{Y}_2) \supseteq \mathcal{R}_C\left([Y_1, \widetilde{Y}_1], [Y_2, \widetilde{Y}_2]\right) \tag{57}$$

$$\mathcal{R}_C(A_1, A_2) \supseteq \mathcal{R}_G(A_1, A_2) \tag{58}$$

where $\widetilde{Y}_1$ and $\widetilde{Y}_2$ satisfy the conditions of Theorem 5, and A_1 and A_2 satisfy the conditions of Proposition 5.

Proof: Suppose we are given $\widetilde{Y}_1$ and $\widetilde{Y}_2$ satisfying the conditions of Theorem 5. We choose $P(a_1, a_2 \mid x_1, x_2)$ so that $A_1 = [Y_1, \widetilde{Y}_1]$ and $A_2 = [Y_2, \widetilde{Y}_2]$. We further choose the DMCs of Proposition 5 so that $Y_2' \sim \widehat{Y}_2$ and $Y_1' \sim \widehat{Y}_1$ (see Fig. 4). Observe that the bounds (51) and (52) are the same as the bounds (55) and (56). This proves (57).

Next, suppose we are given A_1 and A_2 satisfying the requirements of Proposition 5. We can choose $P(\widetilde{y}_1, \widetilde{y}_2 \mid x_1, x_2, y_1, y_2)$ so that $[\widetilde{Y}_1, \widetilde{Y}_2] = [A_1, A_2]$. We can further choose the DMCs of Theorem 5 so that $\widehat{Y}_2 \sim Y_2'$ and $\widehat{Y}_1 \sim Y_1'$ (see Fig. 3). We thus have

$$I(X_1X_2; Y_r\widetilde{Y}_r) = I(X_1X_2; \widetilde{Y}_r) = I(X_1X_2; A_r)$$

because of the Markov relation $X_1X_2 - A_r - Y_r$ for $r = 1, 2$. This proves (58). □

Theorem 6 implies that Proposition 5 and Theorem 5 give the same bounds. Thus, the Gaussian generalization of Proposition 5 will yield Theorem 1. For example, consider $\widetilde{Y}_1$ in (20) with $d_1 = 0$, d_2 as in (32), $d_3 = -c_{21}d_2P_2$, and $d_4 = 1$. Now let $V = 0$, $A_1 = [Y_1, \widetilde{Y}_1]$, and choose the four single-output channels in Fig. 4 as specified in the proof of Theorem 6 . The bound (55) is then the same as the bound (19) in Theorem 1 for $0 \leq c_{21}^2 \leq 1$.

VI. Extensions

There are several possible extensions of the ideas presented here. First, there might be genies other than (20) which give tighter outer bounds than those found here. Second, recall that Theorem 2 is *not* the best known outer bound for all rate pairs. We expect that some generalization of the techniques of [9] will give a better bound. Third, one could extend the bounding technique to interference networks [2] having more than two transmitters and receivers by giving several side-information random variables to each receiver. The optimization of the parameters will be more difficult than for the two-receiver case, though. Finally, one could generalize the bounds to channels where the transmitters and receivers have more than one antenna.

Acknowledgment

The author would like to thank Dr. Igal Sason for his suggestions.

References

[1] E. C. van der Meulen, *Some Reflections on the Interference Channel*. Boston, MA: Kluwer, 1994, pp. 409–421.
[2] A. B. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 60–70, Jan. 1978.
[3] H. Sato, "On the capacity region of a discrete two-user channel for strong interference," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 377–379, May 1978.
[4] ——, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 786–788, Nov. 1981.
[5] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 49–60, Jan. 1981.
[6] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probab.*, vol. 2, pp. 805–814, Oct. 1974.
[7] E. C. van der Meulen, "A survey of multi-way channels in information theory: 1961–1976," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 1–37, Jan. 1977.
[8] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Channels*. Budapest, Hungary: Akadémiai Kiadó, 1981.
[9] M. H. M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 607–615, Sept. 1985.
[10] A. B. Carleial, "Outer bounds on the capacity of interference channels," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 602–606, July 1983.
[11] H. Sato, "Two-user communication channels," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 295–304, May 1977.
[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[13] H. Sato, "On degraded Gaussian two-user channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 637–640, Sept. 1978.
[14] R. Benzel, "The capacity region of a class of discrete additive degraded interference channels," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 228–231, Mar. 1979.
[15] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 2–14, Jan. 1972.
[16] P. P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 279–280, Mar. 1974.
[17] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[18] M. S. Pinsker, V. V. Prelov, and S. Verdú, "Sensitivity of channel capacity," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1877–1888, Nov. 1995.