Feedback Strategies for White Gaussian Interference Networks

Gerhard Kramer, *Member, IEEE*

***Abstract*—A white Gaussian interference network is a channel with T transmitters and R receivers where the received symbols are linear combinations of the transmitted symbols and white Gaussian noise. This paper considers the case where K messages are transmitted through the network in a point-to-point manner, i.e., each message is encoded by exactly one transmitter and is destined for exactly one receiver. It is further assumed that feedback is available so that each transmitter sees the outputs of the receivers to which it is sending messages. Communication strategies based on the discrete Fourier transform (DFT) are developed that perform well for such networks. For multiple-access channels ($K = T$, $R = 1$) with equal transmitter powers the strategies achieve the feedback sum-rate capacity if the powers are beyond some threshold. For the same channels with fixed transmitter powers and large K, the achievable sum-rate is approximately $(\log \log K)/2$ larger than the sum-rate capacity without feedback. For broadcast channels ($T = 1$, $K = R$) with strong symmetries, the strategies achieve a monotonically increasing sum-rate with K. For interference channels ($K = T = R$) with strong interference, the strategies significantly enlarge the no-feedback capacity region by "correlation routing."**

***Index Terms*—Broadcast channels, capacity, feedback communication, Gaussian noise, interference, multiple-access communication.**

I. Introduction

AARON WYNER devoted three of his first eight papers in the IEEE TRANSACTIONS ON INFORMATION THEORY to study the Gaussian channel with noiseless feedback [1]–[3]. The first two of these papers considered Schalkwijk and Kailath's "ingenious scheme for transmitting information through this channel" [1, p. 129], [4]. This scheme dramatically reduces the complexity and delay required to achieve a desired error probability as compared to the best codes without feedback. In fact, the error probability decreases *doubly* exponentially in the number of transmitted symbols rather than the usual single exponential decrease. Wyner's contribution in [1] was to show that, unfortunately, a *peak* rather than *average* power constraint makes the error decrease single exponential again. Thus, because feedback does not increase capacity [5], [6], it might seem that feedback has little to offer for single-transmitter, single-receiver, memoryless channels.

The situation changes for channels with many transmitters and receivers: now feedback sometimes increases capacity [7, Sec. 9], [8]. Thus, one can expect that feedback is useful even with a peak power constraint. However, evaluating the capacity of multiterminal channels seems to be a difficult problem and there are few known conclusive results. Most existing results are either loose bounds or are based on models for which there is a full (high-rate and noiseless) feedback link from the receivers to the transmitters, e.g., [9]–[12].

We will follow tradition and restrict our attention to channels with full feedback. There are two practical reasons for doing this. First, code design is simpler with full feedback than without. Hopefully, the structure of good codes with full feedback will give hints on how to code with realistic feedback. Second, the capacity of such channels provides an outer bound on the capacity with realistic feedback. If full feedback does not increase the capacity by much, one might choose not to incorporate feedback into the code design. On the other hand, a substantial increase in capacity suggests that one should consider codes that utilize feedback.

A. Channel Model

We consider a class of channels called *interference networks* [13]–[15]. An interference network with T transmitters and R receivers is defined by the conditional probability distribution $P(y_1, \ldots, y_R|x_1, \ldots, x_T)$ where the random variables $X_1, \ldots, X_T$ and $Y_1, \ldots, Y_R$ take on values in the respective alphabets $\mathcal{X}_1, \ldots, \mathcal{X}_T$ and $\mathcal{Y}_1, \ldots, \mathcal{Y}_R$ (see [15]). We consider the case for which these alphabets are the set of real (or complex) numbers and

$$Y_r = \left(\sum_{t=1}^{T} \gamma_{rt} X_t\right) + Z_r, \qquad r = 1, \ldots, R \tag{1}$$

where the γ_{rt} are real (or complex) numbers and the $Z_1, \ldots, Z_R$ are independent, zero-mean, Gaussian random variables with unit variance (in the complex case $Z_r = Z_{Rr} + jZ_{Ir}$, where Z_{Rr} and Z_{Ir} are independent, zero-mean, Gaussian random variables with variance $1/2$ and $j = \sqrt{-1}$). Equivalently

$$\underline{Y} = \Gamma \underline{X} + \underline{Z} \tag{2}$$

where Γ is the $R \times T$ matrix of the γ_{rt} and $\underline{X}$, $\underline{Y}$, $\underline{Z}$ are column vectors of the respective X_t, Y_r, Z_r. For example, Fig. 1 depicts an interference network with

$$\Gamma = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1/2 & 1 \end{bmatrix}. \tag{3}$$

Manuscript received February 19, 2001; revised November 15, 2001. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Sorrento, Italy, June 25–30, 2000.

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

Communicated by S. Shamai, Guest Editor.

Publisher Item Identifier S 0018-9448(02)04024-5.

0018-9448/02$17.00 © 2002 IEEE

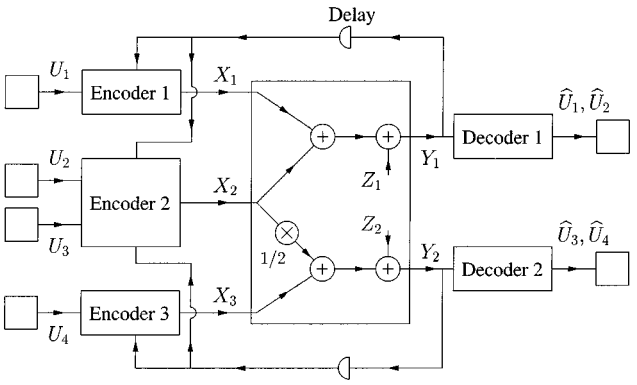


Fig. 1. An interference network with $K = 4, T = 3$, and $R = 2$.

The channel is used N times and there are T power constraints on the transmissions:

$$\sum_{n=1}^{N} \mathrm{E}[|X_{tn}|^2]/N \le P_t, \qquad t = 1, \ldots, T \tag{4}$$

where $\mathrm{E}[\cdot]$ denotes expectation and X_{tn} is the input of transmitter t at time n. We similarly write Y_{rn} and Z_{rn} for the respective output and noise of receiver r at time n.

Our aim is to design good transmission schemes for the following scenario. Suppose K messages $U_1, \ldots, U_K$ are transmitted through the network. The message transmission is assumed to be *point-to-point*, i.e., each message U_k is encoded by exactly one transmitter and is destined for exactly one receiver. There is full feedback in the sense that transmitter t sees, with a small delay, the network outputs of those receivers to which it is transmitting messages.

More formally, let U_k be a sequence of B_k coin-tossing bits so that its rate is B_k/N bits per use, or $R_k := \log(2)B_k/N$ nats per use where $\log(\cdot)$ is the natural logarithm. We collect the indexes of the messages sent by transmitter t and received by receiver r in the respective *encoding index set* $\mathcal{E}_t$ and *decoding index set* $\mathcal{D}_r$. For example, in Fig. 1 we have $\mathcal{E}_1 = \{1\}$, $\mathcal{E}_2 = \{2, 3\}$, $\mathcal{E}_3 = \{4\}$, and $\mathcal{D}_1 = \{1, 2\}$, $\mathcal{D}_2 = \{3, 4\}$. The point-to-point restriction means that the $\mathcal{E}_t$ are disjoint for all t and the $\mathcal{D}_r$ are disjoint for all r. The full-feedback restriction implies that terminal t can make X_{tn} a function of the messages $\{U_k\colon k \in \mathcal{E}_t\}$ and the feedback $\{\underline{Y}_r^{n-1}\colon\ r \text{ satisfies } \mathcal{E}_t \cap \mathcal{D}_r \neq \emptyset\}$, where $\underline{Y}_r^{n-1}$ is the column vector of $Y_{r1}, Y_{r2}, \ldots, Y_{r(n-1)}$. A feedback code is often called a *strategy* (see [7, Sec. 15]). Our feedback strategies should achieve high rates while transmitting the messages *reliably*, i.e., with arbitrarily small positive error probability. The closure of the set of rate-tuples $(R_1, R_2, \ldots, R_K)$ at which one can transmit reliably is called the *capacity region*.

For the purpose of describing our strategies we require that each transmitter send at most one message to any receiver, i.e., $\mathcal{E}_t \cap \mathcal{D}_r$ has at most one element for any t, r. This is not a crucial restriction since one can consider the messages in $\mathcal{E}_t \cap \mathcal{D}_r$ to be a single message with more bits. Next, observe that any transmitter that sends no message receives no feedback, and any receiver that has no message destined for it sends no feedback. Such terminals can be ignored so we assume that each $\mathcal{E}_t$ and $\mathcal{D}_r$ is nonempty. Note, however, that transmitters and receivers with empty $\mathcal{E}_t$ and $\mathcal{D}_r$ are useful in other networks, e.g., the relay in a relay channel [10].

B. Feedback Strategies

Our approach to coding is to generalize the strategies of Schalkwijk, Kailath, Butman, and Ozarow [4], [16], [11] (see also the early work of Elias and Green [17], [18]). A common theme in these papers is to encode by mapping the message U_k onto a point θ_k on the real line and to correct the receiver's linear minimum mean squared error (LMMSE) estimate $\hat{\theta}_{kn}$ of this point at each channel use. More precisely, let $\epsilon_{kn} = \hat{\theta}_{kn} - \theta_k$ be the estimation error after n channel uses. The transmitters send power-normalized versions X_{kn} of the ϵ_{kn}. A technique closely related to this idea had been used for discrete memoryless channels as early as [19].

An extension of this approach was presented in [16] where Gaussian channels with *colored* noise are considered. For such channels, it turns out that one improves performance by judiciously multiplying the ϵ_{1n} by ± 1. The aim of this *phase modulation* is to align the phase of X_{1n} with the phase of the noise symbols, thereby increasing the mutual information between the transmit and receive symbol sequences. We call this modulation technique *modulated estimate correction* or simply MEC.

The MEC technique was adapted to the two-transmitter multiple-access channel (MAC) with feedback in [11] and the two-receiver broadcast channel with feedback in [20]. The idea for the MAC is to have transmitter 1 send ϵ_{1n} *without* modulation, while transmitter 2 sends ϵ_{2n} *with* modulation by ± 1. The purpose of the modulation is to align the phases of X_{1n} and X_{2n} so as to increase the mutual information between the channel input and output symbol sequences. It was found that, in the capacity-achieving steady state, the modulating sequence of transmitter 2 is $+1, -1, +1, -1, \ldots$ [11].

More recently, the MEC technique was extended to the T-transmitter MAC with $T \ge 3$ in [21], [22]. However, these strategies do not achieve capacity. We propose other strategies that achieve capacity in some cases and are useful for the much larger class of channels considered here. Our two key ideas can be summarized as follows. First, we consider a *complex* channel model rather than a real one. Second, we give message k a distinct *modulation frequency* $(k-1)/K$, i.e., the modulating sequence of message k is

$$1, e^{(2\pi j(k-1)/K)}, e^{2\cdot(2\pi j(k-1)/K)}, e^{3\cdot(2\pi j(k-1)/K)}, \ldots .$$

Furthermore, we simplify the notation and analyses of [4], [16], [11], [20] to be able to deal with many messages, transmitters, and receivers. One of the main results of the simplifications is a matrix recursion that is useful for all the interference networks described above.

This paper is organized as follows. Section II introduces notation and presents capacity outer bounds. In Section III, we describe the MEC strategies in detail and derive a covariance matrix recursion. In Section IV, we introduce *Fourier* MEC strategies that are based on the discrete Fourier transform (DFT). Section V develops further results for the MAC. Section VI presents results for other interference networks: the broadcast channel, the interference channel, and the network depicted in Fig. 1. In

Section VII, we derive bounds on the decoding error probability. Section VIII concludes the paper by discussing open problems and extensions of the ideas presented here.

II. Preliminaries

A. Notation

Let $\underline{A}^T$ and $\underline{A}^H$ denote the respective transpose and complex-conjugate transpose of the column vector $\underline{A}$. The covariance matrix of $\underline{A}$ is $Q_{\underline{A}} := \mathrm{E}[(\underline{A}-\mu_{\underline{A}})(\underline{A}-\mu_{\underline{A}})^H]$ where $\mu_{\underline{A}} = \mathrm{E}[\underline{A}]$. We define

$$V_{\underline{A}} := |Q_{\underline{A}}|$$

where $|M|$ denotes the determinant of the matrix M. $V_{\underline{A}}$ can be considered to be a kind of variance. The entropy of a real, length-N, Gaussian random vector $\underline{A}$ with nonzero $V_{\underline{A}}$ is (see, e.g., [23, pp. 230, 234])

$$H(\underline{A}) = \tfrac{1}{2}\log\left((2\pi e)^N V_{\underline{A}}\right).$$

We will use the same notation for entropy and mutual information as in [24, p. 18, 20].

The LMMSE estimate of a *zero-mean* complex random vector $\underline{A} = [A_1 A_2 \cdots A_J]^T$ given the *zero-mean* complex random vector $\underline{B} = [B_1 B_2 \cdots B_L]^T$ has the form $\hat{\underline{A}} = M \cdot \underline{B}$ for some matrix M. The orthogonality principle [25, p. 386] ensures that $\mathrm{E}[(\hat{\underline{A}}-\underline{A})\underline{B}^H] = 0$ so that $MQ_{\underline{B}} = \mathrm{E}[\underline{A}\,\underline{B}^H]$. Thus, if $Q_{\underline{B}}$ is invertible we have [25, p. 382]

$$\hat{\underline{A}} = \left(\mathrm{E}[\underline{A}\,\underline{B}^H]\cdot Q_{\underline{B}}^{-1}\right)\cdot \underline{B}. \tag{5}$$

We will consider only invertible $Q_{\underline{B}}$. We define the conditional "variance" of $\underline{A}$ given $\underline{B}$ as

$$V_{\underline{A}|\underline{B}} := V_{\underline{A}-\hat{\underline{A}}}.$$

The covariance matrix of $[A_1 \cdots A_J\ B_1 \cdots B_L]^T$ is denoted by $Q_{\underline{A}\,\underline{B}}$, and we write $V_{\underline{A}\,\underline{B}}$ for $|Q_{\underline{A}\,\underline{B}}|$. One can verify that, for nonzero $|Q_{\underline{B}}|$

$$V_{\underline{A}|\underline{B}} = \frac{|Q_{\underline{A}\,\underline{B}}|}{|Q_{\underline{B}}|} = \frac{V_{\underline{A}\,\underline{B}}}{V_{\underline{B}}}, \tag{6}$$

where the first equality follows by the Schur complement formula [26, p. 22]

$$\begin{aligned}|Q_{\underline{A}\,\underline{B}}| &= |Q_{\underline{A}}|\cdot|Q_{\underline{B}} - \mathrm{E}[\underline{B}\,\underline{A}^H]Q_{\underline{A}}^{-1}\,\mathrm{E}[\underline{A}\,\underline{B}^H]| \\ &= |Q_{\underline{B}}|\cdot|Q_{\underline{A}} - \mathrm{E}[\underline{A}\,\underline{B}^H]\,Q_{\underline{B}}^{-1}\,\mathrm{E}[\underline{B}\,\underline{A}^H]|.\end{aligned} \tag{7}$$

We sometimes deal with random vectors whose entries A_k have indexes in a set $\mathcal{S}$. We denote this vector by $\underline{A}^{(\mathcal{S})}$, e.g., $\underline{A}^{(\{1,3\})} = [A_1 \quad A_3]^T$. Similarly, for a matrix A let $A^{(\mathcal{S})}$ be the submatrix of A containing the entries of the rows and columns in $\mathcal{S}$.

Next, the *Hadamard product* of two matrices A and B, denoted $A \odot B$, is the matrix whose entries are the entry-by-entry products of the corresponding entries of A and B [26, p. 455]. Two simple properties involving the Hadamard product are

$$|A \odot (\underline{b}\,\underline{c}^H)| = |A| \cdot \prod_k b_k\, c_k^* \tag{8}$$

$$\left[A \odot \left(\underline{b}\,\underline{c}^H\right)\right]\cdot D = \left[A \odot \left(\underline{b}\,\underline{1}^T\right)\right]\cdot\left[D \odot \left(\underline{c}^*\,\underline{1}^T\right)\right] \tag{9}$$

for any matrices A, D and vectors $\underline{b}$, $\underline{c}$, where $*$ denotes complex conjugation. We will also make use of a Hadamard *quotient*, i.e., the matrix whose entries are the entry-by-entry quotients of A and B. We denote this matrix by

$$\odot\frac{A}{B}.$$

We denote by $\sqrt{A}$, $\exp(\underline{A})$, and $\log(\underline{A})$ the matrices and vectors whose entries are the *entry-by-entry* square roots, exponentials, and logarithms of the entries of A and $\underline{A}$, respectively. We denote by $\mathrm{diag}\,(A)$ the column vector of the diagonal entries of A. The notation $\underline{0}$ and $\underline{1}$ is used for the respective all-zeros and all-ones vectors.

Finally, consider the length-K DFT with which one can associate the $K \times K$ matrix F having entries

$$f_{mn} = \exp(2\pi j(m-1)(n-1)/K).$$

We let $\underline{f}_n$ denote column $(n \bmod K + 1)$ of F and define $F_n := \underline{f}_n\,\underline{f}_n^H$.

B. Real and Complex Networks

The interference network is a *memoryless* channel. The memoryless nature of the channel makes the capacity region independent of the amount of delay in the feedback links as long as all delays are finite. One can thus, without loss of optimality, group the transmitter and feedback symbols into length-L blocks and use the feedback only after each such block. For example, suppose that Γ has *real* entries. If the network has *real* symbol alphabets then one can choose $L = 2$ to simulate any strategy for the same network with *complex* symbol alphabets. Similarly, if the network has *complex* symbol alphabets then one can run any pair of *real* strategies in parallel on the real and imaginary parts of the X_{tn}. We thus have the result that the capacity region of a real interference network and its complex counterpart coincide, as long as one normalizes the rates by the number of real dimensions. From here on, we therefore consider only complex networks.

C. MAC Outer Bounds on the Capacity Region

Suppose that the R receivers can cooperate, so that one has a multiple-access channel (MAC) with output $\underline{Y} = [Y_1 \cdots Y_R]^T$. The capacity region of the original network is contained in that of this MAC. In turn, an outer bound on the capacity region of the MAC is the intersection of the rate-tuples satisfying (see [9, eq. (9)]

$$\sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{E}_t} R_k \le I\left(\underline{X}^{(\mathcal{T})};\,\underline{Y}\,\middle|\,\underline{X}^{(\mathcal{T}^C)}\right) \tag{10}$$

where $\mathcal{T}$ is any subset of $\{1, 2, \ldots, T\}$, $\mathcal{T}^C$ is the complement of $\mathcal{T}$ in $\{1, 2, \ldots, T\}$, and $\underline{X} = [X_1 \cdots X_T]^T$ is the vector of network inputs. The only restriction on the distribution $P_{\underline{X}}$ of $\underline{X}$ is that it satisfy the power constraints, i.e., $\mathrm{E}[|X_t|^2] \le P_t$ for all t. However, a conditional maximum-entropy theorem (see

[9, Lemma 1]) lets one restrict attention to zero-mean Gaussian $P_{\underline{X}}$. The bound (10) is then

$$\sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{E}_t} R_k \le \tfrac{1}{2}\log\left(V_{\underline{Y}|\underline{X}^{(\mathcal{T}^C)}}\right) \tag{11}$$

where $Q_{\underline{X}}$ satisfies $\mathrm{E}[|X_t|^2]\le P_t$ for all t. We use the natural logarithm throughout the paper, so that the rate units are nats per use per dimension.

Consider first the broadcast channel ($T=1$). The only determinant we need in (11) is $V_{\underline{Y}}=|I+P_1\cdot\Gamma\Gamma^H|$, so that

$$\sum_{k=1}^{K} R_k \le \frac{1}{2}\log\left[1+P_1\sum_{r=1}^{R}|\gamma_{r1}|^2\right]. \tag{12}$$

We derive a better bound than (12) in Appendix E.

The computation of (11) becomes more involved as T increases because of the maximization over $Q_{\underline{X}}$. However, the best $Q_{\underline{X}}$ sometimes has a simple form when the network has many symmetries. In particular, consider a MAC ($R=1$) having $\gamma_{1t}=1$ and $P_t=P$ for all t. Suppose we want to bound the sum rate $\sum_{k=1}^{K}R_k$, and suppose that $Q_{\underline{X}}$ maximizes the sum rate. One can swap any two of the entries in $\underline{X}$ and get the same sum rate because of the network symmetries. Furthermore, one can complex-conjugate any X_t and still get the same sum rate. The "average" of these $Q_{\underline{X}}$ cannot give a worse sum rate by the concavity of $\log(|Q_{\underline{X}}^{(\mathcal{A})}|/|Q_{\underline{X}}^{(\mathcal{B})}|)$ in $Q_{\underline{X}}^{(\mathcal{A})}$ for $\mathcal{B}\subseteq\mathcal{A}$ (see [23, p. 506]). Thus, we may restrict attention to $Q_{\underline{X}}$ of the form

$$Q_{\underline{X}} = P\cdot\begin{bmatrix}1&\rho&\rho&\cdots&\rho\\ \rho&1&\rho&\cdots&\rho\\ \rho&\rho&1&\cdots&\rho\\ \vdots&\vdots&\vdots&\ddots&\vdots\\ \rho&\rho&\rho&\cdots&1\end{bmatrix} \tag{13}$$

where ρ is a nonnegative *real* number. This approach is the same as that used in [9, Sec. IV]. Using (13) in (11) we arrive at [9, eq. (45)]

$$\sum_{k=1}^{K} R_k \le \frac{K}{2|\mathcal{T}|}\log\left[1+\frac{1+(K-1)\rho}{1+(K-|\mathcal{T}|-1)\rho}(1-\rho)|\mathcal{T}|P\right] \tag{14}$$

for $|\mathcal{T}|=1,2\ldots,K$ and some ρ, $0\le\rho\le 1$. For example, consider $|\mathcal{T}|=K$ for which $\rho=1$ gives the largest sum rate. Using (14), we have

$$\begin{aligned}\sum_{k=1}^{K} R_k &\le \frac{1}{2}\log\left(1+K^2P\right)\\ &\le \frac{1}{2}\log\left(K+K^2P\right)\\ &= \frac{1}{2}\log\left(1+KP\right)+\frac{1}{2}\log K.\end{aligned} \tag{15}$$

Observe that the first term in (15) is the sum-rate capacity without feedback so that feedback increases the capacity by at most $(\log K)/2$. This bound is a little better than the $[\log(K+1)]/2$ bound in [27, Theorem 4.1]. However, the latter bound is also valid for nonwhite Gaussian noise. Another simple bound is the factor-of-two bound in [9].

The reasoning used to arrive at (13) can be used to show that this $Q_{\underline{X}}$ maximizes the sum rate of other networks with strong symmetries, e.g., any network where $P_t=P$ for all t and where Γ is the all-ones matrix. The resulting bounds have the same form as (14) with P replaced by RP. To see this, consider

$$\begin{aligned}V_{\underline{Y}|X_1\ldots X_k} &= |Q_{\underline{Y}|X_1\ldots X_k}|\\ &= \left|I+\underline{1}\,\underline{1}^T\cdot\mathrm{E}[Y_1^2|X_1\ldots X_k]\right|\\ &= 1+R\cdot\mathrm{E}[Y_1^2|X_1\ldots X_k].\end{aligned}$$

Now observe that for the MAC we have

$$V_{Y_1|X_1\ldots X_k} = 1+\mathrm{E}[Y_1^2|X_1\ldots X_k].$$

Thus, we obtain (14) with the second term in the square brackets multiplied by the number of receivers R.

Another channel for which (13) maximizes the sum rate is the interference channel ($T=R$) where $P_t=P$ for all t and where Γ has diagonal entries 1 and off-diagonal entries γ. For example, if $K=2$ and γ is real we have

$$\begin{aligned}R_1+R_2 &\le \log\left(1+P(1+\gamma^2)(1-\rho^2)\right),\\ R_1+R_2 &\le x\frac{1}{2}\log\left[1+2P(1+\gamma^2+2\gamma\rho)\right.\\ &\qquad\left.+P^2(1-\gamma^2)^2(1-\rho^2)\right].\end{aligned} \tag{16}$$

These bounds will be used in Section VI-B.

III. LMMSE Estimation and MEC Coding

We describe two transmission schemes: *full* LMMSE estimation and *memoryless* LMMSE estimation. The former scheme has receiver r using the entire vector $\underline{Y}_r^n$ available to it at time n to estimate θ_k, $k\in\mathcal{D}_r$. This approach seems difficult to analyze and simulate. The second scheme has receiver r estimating $\epsilon_{k(n-1)}$, $k\in\mathcal{D}_r$, by using Y_{rn} only. This approach is easier to analyze and we use it to obtain the results of Sections IV–VI.

A. Full LMMSE Estimation

The messages U_k are mapped onto points θ_k in the complex plane. For now we simply assume that the θ_k are random variables with zero mean and finite variance. We shall return to the effect of the distribution of θ_k in Section VII.

After the nth channel use, receiver r outputs the LMMSE estimates

$$\hat{\theta}_{kn} = \mathrm{E}[\theta_k\,(\underline{Y}_r^n)^H]\cdot Q_{\underline{Y}_r^n}^{-1}\cdot\underline{Y}_r^n \tag{17}$$

where $k\in\mathcal{D}_r$. The estimation error is $\epsilon_{kn}=\hat{\theta}_{kn}-\theta_k$ and we set $\hat{\theta}_{k0}=0$. By the orthogonality principle [25, p. 386] we have

$$\mathrm{E}[\epsilon_{kn}^*\,\underline{Y}_r^n]=\underline{0}. \tag{18}$$

At time n transmitter t corrects the receivers' estimates by sending a weighted sum of the errors $\epsilon_{k(n-1)}$, i.e.,

$$X_{tn}=\sum_{k\in\mathcal{E}_t} W_{kn} \tag{19}$$

where

$$W_{kn} = \left(\sqrt{P_{Wkn}/\sigma^2_{k(n-1)}} \cdot m^*_{kn}\right) \cdot \epsilon_{k(n-1)} \tag{20}$$

and where P_{Wkn} is a nonnegative real number, $\sigma^2_{k(n-1)} = \mathrm{E}[|\epsilon_{k(n-1)}|^2]$, and m_{kn} is a *modulation coefficient*. Both the P_{Wkn} and the m_{kn} are chosen before transmission and are known to all receivers. Of course, the choice of the P_{Wkn} and m_{kn} must take into account the power constraints, i.e.,

$$\sum_{n=1}^{N} \mathrm{E}[|X_{tn}|^2] = \sum_{n=1}^{N} \mathrm{E}\left[\left|\sum\nolimits_{k\in\mathcal{E}_t} W_{kn}\right|^2\right] \le P_t. \tag{21}$$

After the N network uses, the receivers choose their $\hat{U}_k$ as those U_k whose points θ_k are closest to the respective $\hat{\theta}_{kN}$.

Consider the error variance $\sigma^2_{kn} = \mathrm{E}[|\epsilon_{k(n-1)}|^2]$. In Appendix A we show that

$$\sigma^2_{kn} = \sigma^2_{k(n-1)} \cdot \frac{V_{Y_{rn}|\theta_k \underline{Y}_r^{n-1}}}{V_{Y_{rn}|\underline{Y}_r^{n-1}}}. \tag{22}$$

Note that the "uncertainty" about the message point decreases exponentially with n if the fraction on the right-hand of (22) approaches a fixed point value less than one. The "instantaneous rate" of decrease is

$$\begin{aligned} R_{kn} &:= \frac{1}{2}\log\left(\sigma^2_{k(n-1)}/\sigma^2_{kn}\right) \\ &= \frac{1}{2}\log\left(V_{Y_{rn}|\underline{Y}_r^{n-1}} \Big/ V_{Y_{rn}|\theta_k \underline{Y}_r^{n-1}}\right) \end{aligned} \tag{23}$$

where the $1/2$ in front of the logarithm normalizes the rate by the number of real dimensions. We will show in Section VII that one can send at a data rate R_k close to the average $\sum_{n=1}^{N} R_{kn}/N$ for large N. Consider next the K^2 correlation coefficients

$$\rho_{(kk')n} = \frac{\mathrm{E}[\epsilon_{kn}\epsilon^*_{k'n}]}{\sqrt{\sigma^2_{kn}\sigma^2_{k'n}}} \tag{24}$$

$1 \le k \le K, 1 \le k' \le K$. Let r and r' be the receivers decoding the respective U_k and $U_{k'}$ (this convention is used throughout the paper). The recursion for $\rho_{(kk')n}$ is (see Appendix A)

$$\begin{aligned} \rho_{(kk')n} = \exp(R_{kn}+R_{k'n}) \cdot \Big[&\rho_{(kk')(n-1)} \\ &- \underline{c}^T_{(kr')n} Q^{-1}_{\underline{Y}^n_{r'}} \underline{c}^*_{(k'r')n} - \underline{c}^T_{(kr)n} Q^{-1}_{\underline{Y}^n_r} \underline{c}^*_{(k'r)n} \\ &+ \underline{c}^T_{(kr)n} Q^{-1}_{\underline{Y}^n_r} \mathrm{E}[(\underline{Y}^n_r)(\underline{Y}^n_{r'})^H] Q^{-1}_{\underline{Y}^n_{r'}} \underline{c}^*_{(k'r')n}\Big] \end{aligned} \tag{25}$$

where

$$\underline{c}_{(k\tilde{r})n} = \mathrm{E}\left[\epsilon_{k(n-1)}(\underline{Y}^n_{\tilde{r}})^*\right] \Big/ \sqrt{\sigma^2_{k(n-1)}}. \tag{26}$$

B. Memoryless LMMSE Estimation

The performance analysis for the LMMSE estimator (17) is complicated because successive channel outputs are *not* independent in general. An approach more amenable to analysis is to have receiver r estimate $\epsilon_{k(n-1)}$ by using Y_{rn} only (see [20]). The estimate of θ_k is then simply the recursive

$$\hat{\theta}_{kn} = \hat{\theta}_{k(n-1)} - \hat{\epsilon}_{k(n-1)}.$$

The transmitters continue to use MEC and send a weighted sum of the modulated ϵ_{kn}.

Receiver r's memoryless LMMSE estimates are

$$\hat{\epsilon}_{k(n-1)} = \frac{\mathrm{E}[\epsilon_{k(n-1)}Y^*_{rn}]}{V_{Y_{rn}}} \cdot Y_{rn}. \tag{27}$$

The recursion (22) simplifies to (see Appendix A)

$$\sigma^2_{kn} = \sigma^2_{k(n-1)} \cdot \frac{V_{Y_{rn}|W_{kn}}}{V_{Y_{rn}}}. \tag{28}$$

The new "instantaneous rate" is thus

$$\begin{aligned} R_{kn} &= \frac{1}{2}\log\left(\sigma^2_{k(n-1)}/\sigma^2_{kn}\right) \\ &= \frac{1}{2}\log\left(V_{Y_{rn}}/V_{Y_{rn}|W_{kn}}\right). \end{aligned} \tag{29}$$

The recursion (25) simplifies to (see Appendix A)

$$\begin{aligned} \rho_{(kk')n} = \exp(R_{kn}+R_{k'n}) \Bigg[&\rho_{(kk')(n-1)} - \frac{c_{(kr')n}c^*_{(k'r')n}}{V_{Y_{r'n}}} \\ &- \frac{c_{(kr)n}c^*_{(k'r)n}}{V_{Y_{rn}}} + \frac{c_{(kr)n}c^*_{(k'r')n}}{V_{Y_{rn}}V_{Y_{r'n}}}\mathrm{E}[Y_{rn}Y^*_{r'n}]\Bigg] \end{aligned} \tag{30}$$

where

$$c_{(k\tilde{r})n} = \mathrm{E}\left[\epsilon_{k(n-1)}Y^*_{\tilde{r}n}\right] \Big/ \sqrt{\sigma^2_{k(n-1)}}. \tag{31}$$

For further analysis and computational convenience, the recursions (30) are collected in one *matrix* recursion

$$\begin{aligned} Q_n = \exp(\underline{R}_n)\exp(\underline{R}_n)^T \odot \Bigg[&Q_{n-1} - \left(\odot \frac{C_n \odot D_n}{\underline{1}\cdot \underline{V}^T_{\underline{Y}_n}}\right) \\ &- \left(\odot \frac{C_n \odot D_n}{\underline{1}\cdot \underline{V}^T_{\underline{Y}_n}}\right)^H + \left(\odot \frac{D_n \odot D_n^H \odot Q_{\underline{Y}_n}}{\underline{V}_{\underline{Y}_n}\underline{V}^T_{\underline{Y}_n}}\right)\Bigg] \end{aligned} \tag{32}$$

where Q_n is the $K\times K$ matrix of the $\rho_{(kk')n}$, $\underline{R}_n$ is the length-K vector of the R_{kn}, C_n is the $K \times K$ matrix whose k, k' entry is $c_{(kr')n}$, D_n is the $K \times K$ matrix $\underline{1}\cdot \mathrm{diag}\,(C_n)^H$, $\underline{V}_{\underline{Y}_n}$ is the length-K vector of the $V_{Y_{rn}}$ (with $V_{Y_{rn}}$ appearing in position k), $\underline{Y}_n$ is the length-K vector whose kth entry is Y_{rn}, and $Q_{\underline{Y}_n} = \mathrm{E}[\underline{Y}_n\underline{Y}_n^H]$. The steps required to compute this recursion are listed in Appendix B, and they require a few more identities.

First, let M_T to be the $K \times T$ matrix whose k, t entry is 1 if message U_k is sent by transmitter t, and is 0 otherwise. Similarly, let M_R to be the $K \times R$ matrix whose k, r entry is 1 if message U_k is decoded by receiver r, and is 0 otherwise. We define

$$\tilde{\Gamma} := M_R \cdot \Gamma \cdot M_T^T \tag{33}$$

i.e., $\tilde{\Gamma}$ is the $K \times K$ matrix whose k, k' entry is $\tilde{\gamma}_{kk'} = \gamma_{rt'}$, where (as usual) r is the receiver decoding U_k and t' is the transmitter sending $U_{k'}$. For example, the network of Fig. 1 has

$$\begin{aligned} \tilde{\Gamma} &= \begin{bmatrix}1&0\\1&0\\0&1\\0&1\end{bmatrix} \cdot \begin{bmatrix}1&1&0\\0&1/2&1\end{bmatrix} \cdot \begin{bmatrix}1&0&0\\0&1&0\\0&1&0\\0&0&1\end{bmatrix}^T \\ &= \begin{bmatrix}1&1&1&0\\1&1&1&0\\0&1/2&1/2&1\\0&1/2&1/2&1\end{bmatrix}. \end{aligned} \tag{34}$$

Furthermore, we define

$$G_n := \tilde{\Gamma}^H \odot \left[\left(\sqrt{\underline{P}_{Wn}} \odot \underline{m}_n\right)\cdot \underline{1}^T\right] \tag{35}$$

where $\sqrt{\underline{P}_{Wn}}$ and $\underline{m}_n$ are the length-K vectors of the respective $\sqrt{P_{Wkn}}$ and m_{kn}. By inserting (1), (19), and (20) into (31) we have

$$c_{(k\tilde{r})n} = \sum_{\ell=1}^{K} \rho_{(k\ell)(n-1)} \cdot \left[\tilde{\gamma}^*_{\tilde{r}\ell} \sqrt{P_{W\ell n}}\, m_{\ell n}\right]$$
$$\Rightarrow \quad C_n = Q_{n-1} \cdot G_n. \tag{36}$$

Similarly, one finds that (see Appendix A)

$$Q_{\underline{Y}_n} = G_n^H Q_{n-1} G_n + M_R M_R^T \tag{37}$$
$$\underline{V}_{\underline{Y}_n} = \mathrm{diag}\left(Q_{\underline{Y}_n}\right) \tag{38}$$
$$\underline{V}_{\underline{Y}_n|\underline{W}_n} = \underline{V}_{\underline{Y}_n} - |\mathrm{diag}\,(C_n)|^2 \tag{39}$$
$$\underline{R}_n = \frac{1}{2}\log\left(\odot \frac{\underline{V}_{\underline{Y}_n}}{\underline{V}_{\underline{Y}_n|\underline{W}_n}}\right) \tag{40}$$

where $\underline{V}_{\underline{Y}_n|\underline{W}_n}$ is the length-K vector of the $V_{Y_{rn}|W_{kn}}$. Finally, the power constraints are, for $t = 1, \ldots, T$

$$\left(\sqrt{\underline{P}_{Wn}^{(\mathcal{E}_t)}} \odot \underline{m}_n^{(\mathcal{E}_t)}\right)^H \cdot Q_{n-1}^{(\mathcal{E}_t)} \cdot \left(\sqrt{\underline{P}_{Wn}^{(\mathcal{E}_t)}} \odot \underline{m}_n^{(\mathcal{E}_t)}\right) \le P_t. \tag{41}$$

The T bounds (41) can be written in matrix form as

$$\mathrm{diag}\left(\left[\left(\sqrt{\underline{P}_{Wn}} \odot \underline{m}_n\right) \cdot M_T\right]^H \cdot Q_{n-1} \cdot \left[\left(\sqrt{\underline{P}_{Wn}} \odot \underline{m}_n\right) \cdot M_T\right]\right) \le \underline{P} \tag{42}$$

where $\underline{P}$ is the length-T vector of the P_t, and $\underline{a} \le \underline{b}$ means $a_t \le b_t$ for all t.

Terminal t can meet its power constraint by distributing the power among its messages in a fixed manner and normalizing X_{tn} before transmission. In other words, terminal t chooses a power distribution $\underline{P}_{W0}^{(\mathcal{E}_t)}$ and computes

$$\tilde{P}_{tn} = \left(\sqrt{\underline{P}_{W0}^{(\mathcal{E}_t)}} \odot \underline{m}_n^{(\mathcal{E}_t)}\right)^H \cdot Q_{n-1}^{(\mathcal{E}_t)} \cdot \left(\sqrt{\underline{P}_{W0}^{(\mathcal{E}_t)}} \odot \underline{m}_n^{(\mathcal{E}_t)}\right) \tag{43}$$

before each channel use. The $\underline{P}_{Wn}^{(\mathcal{E}_t)}$ are then set to

$$\underline{P}_{Wn}^{(\mathcal{E}_t)} = \underline{P}_{W0}^{(\mathcal{E}_t)} \cdot (P_t/\tilde{P}_{tn}) \tag{44}$$

which ensures that $\mathrm{E}[|X_{tn}|^2] = P_t$.

C. A Second Matrix Recursion

In addition to the $\rho_{(kk')n}$ we will also consider the correlation coefficients

$$\tilde{\rho}_{(kk')n} = \frac{\mathrm{E}[W_{kn} W^*_{k'n}]}{\sqrt{P_{Wkn} P_{Wk'n}}}. \tag{45}$$

Substituting (20) into (45), we have

$$\tilde{\rho}_{(kk')n} = \rho_{(kk')(n-1)} \cdot m^*_{kn} m_{k'n}.$$

We again collect the $\tilde{\rho}_{(kk')n}$ into a matrix $\tilde{Q}_n$, so that

$$\tilde{Q}_n = Q_{n-1} \odot M_n^*$$

where $M_n = \underline{m}_n \underline{m}_n^H$. Furthermore, define

$$\tilde{G}_n := \tilde{\Gamma}^H \odot \left(\sqrt{\underline{P}_{Wn}} \cdot \underline{1}^T\right)$$
$$\tilde{C}_n := \tilde{Q}_n \cdot \tilde{G}_n$$
$$\tilde{D}_n := \underline{1} \cdot \mathrm{diag}\,(\tilde{C}_n)^H.$$

In fact, $\tilde{G}_n$, $\tilde{C}_n$, and $\tilde{D}_n$ are the respective $G_n \odot (\underline{m}_n^{-1} \cdot \underline{1}^T)$, $C_n \odot (\underline{m}_n^* \cdot \underline{1}^T)$, and $D_n \odot (\underline{1} \cdot \underline{m}_n^T)$. We also have

$$Q_{\underline{Y}_n} = \tilde{G}_n^H \tilde{Q}_n \tilde{G}_n + M_R M_R^T$$
$$\underline{V}_{\underline{Y}_n} = \mathrm{diag}\left(Q_{\underline{Y}_n}\right)$$
$$\underline{V}_{\underline{Y}_n|\underline{W}_n} = \underline{V}_{\underline{Y}_n} - |\mathrm{diag}\,(\tilde{C}_n)|^2$$
$$\tilde{P}_{tn} = \sqrt{\left(\underline{P}_{W0}^{(\mathcal{E}_t)}\right)^T} \cdot \tilde{Q}_n^{(\mathcal{E}_t)} \cdot \sqrt{\underline{P}_{W0}^{(\mathcal{E}_t)}}.$$

Applying (9) several times, we find that (32) becomes

$$\tilde{Q}_{n+1} = \odot \frac{M^*_{n+1}}{M^*_n} \odot \exp(\underline{R}_n) \exp(\underline{R}_n)^T \odot \left[\tilde{Q}_n - \left(\odot \frac{\tilde{C}_n \odot \tilde{D}_n}{\underline{1} \cdot \underline{V}^T_{\underline{Y}_n}}\right) - \left(\odot \frac{\tilde{C}_n \odot \tilde{D}_n}{\underline{1} \cdot \underline{V}^T_{\underline{Y}_n}}\right)^H + \left(\odot \frac{\tilde{D}_n \odot \tilde{D}_n^H \odot Q_{\underline{Y}_n}}{\underline{V}_{\underline{Y}_n} \underline{V}^T_{\underline{Y}_n}}\right)\right]. \tag{46}$$

The advantage of (46) over (32) is that the modulation coefficients have been moved into a single matrix: the first matrix on the right-hand side of (46).

IV. Fourier MEC

The only design issue remaining is the choice of modulation coefficients. We choose

$$m_{kn} = \exp\left(j\,\frac{2\pi(k-1)}{K}\,(n-1)\right) \tag{47}$$

i.e., each message in the network is given a distinct *modulation frequency*. Another way to look at this is that the modulation coefficients are chosen periodically from the rows of the length-K DFT matrix. *We will consider only the modulation coefficients (47) from here on.* However, one can show that other choices for the m_{kn} will yield the same performance as (47)—for instance, if K factors as $K_a K_b$ then one can replace the length-K DFT by a two-dimensional DFT. Taking this idea further, if $K = 2^m$, one can modulate using the length-2^m Hadamard transform. We call all such transmission techniques *Fourier MEC*.

A. Existence of a Fixed Point

Consider the recursion (46) with (47) so that $\underline{m}_n = \underline{f}_n$ and $M_n = F_n$ (see the end of Section II for the definition of $\underline{f}_n$ and F_n). We then have

$$\odot \frac{M^*_{n+1}}{M^*_n} = F_{-1}. \tag{48}$$

Inserting (48) into (46), we have

$$\tilde{Q}_{n+1} = g(\tilde{Q}_n)$$

where $g(\cdot)$ is defined by (46). Note that by using Fourier MEC the mapping $g(\cdot)$ does not depend on n. Furthermore, one can show that $g(\cdot)$ is *continuous* with respect to the Frobenius matrix norm $\|Q\| := \sqrt{\sum_{k,k'} |q_{kk'}|^2}$. These facts lead to the following fixed-point result.

Theorem 1: The mapping (46) has at least one fixed point, i.e., there exists a correlation coefficient matrix Q such that $g(Q) = Q$.

Proof: Let $\mathcal{Q}$ be the set of $K \times K$ correlation coefficient matrices. Equivalently, $\mathcal{Q}$ is the set of positive semidefinite Hermitian matrices whose diagonal entries are 1 [28, p. 338]. The set $\mathcal{Q}$ is clearly bounded with respect to the Frobenius matrix norm. This set is also closed because the limit of any sequence $\{Q_n\}$ with $Q_n \in \mathcal{Q}$ will be positive semidefinite, Hermitian, and have unit diagonals. Furthermore, $\mathcal{Q}$ is convex because the set of positive semidefinite Hermitian matrices is convex, and because unit diagonals are retained by convex combinations. Since $g(\cdot)$ is a continuous map of $\mathcal{Q}$ into itself, the Brouwer fixed point theorem [29, p. 780] gives the desired result. □

B. Fixed Point Properties

Suppose $\tilde{Q}$ is a fixed point of (46), and let $\underline{R}$, $\tilde{C}$, $\tilde{D}$, etc., be the corresponding fixed-point vectors and matrices. Suppose further that we restrict our attention to receiver r. The submatrix $\tilde{Q}^{(\mathcal{D}_r)}$ of $\tilde{Q}$ is, using (46) and $\mathrm{E}[Y_r Y_{r'}] = V_{Y_r}$ for $r, r' \in \mathcal{D}_r$

$$\tilde{Q}^{(\mathcal{D}_r)} = F_{-1}^{(\mathcal{D}_r)} \odot \left[\odot \frac{V_{Y_r} \tilde{Q}^{(\mathcal{D}_r)} - \underline{\tilde{c}}^{(\mathcal{D}_r)} \left(\underline{\tilde{c}}^{(\mathcal{D}_r)}\right)^H}{\sqrt{\underline{V}_{Y|\underline{W}}^{(\mathcal{D}_r)} \left(\underline{V}_{Y|\underline{W}}^{(\mathcal{D}_r)}\right)^T}} \right] \tag{49}$$

where $\underline{\tilde{c}}^{(\mathcal{D}_r)}$ is any column of $\tilde{C}^{(\mathcal{D}_r)}$. Applying (7) and (8), we have

$$\begin{aligned} \left|\tilde{Q}^{(\mathcal{D}_r)}\right| &= \frac{V_{Y_r}^{|\mathcal{D}_r|}}{\prod_{k \in \mathcal{D}_r} V_{Y_r|W_k}} \cdot \left| \tilde{Q}^{(\mathcal{D}_r)} - \frac{\underline{\tilde{c}}^{(\mathcal{D}_r)} \left(\underline{\tilde{c}}^{(\mathcal{D}_r)}\right)^H}{V_{Y_r}} \right| \\ &= \frac{V_{Y_r}^{|\mathcal{D}_r|}}{\prod_{k \in \mathcal{D}_r} V_{Y_r|W_k}} \cdot \frac{V_{\underline{W}^{(\mathcal{D}_r)}|Y_r}}{\prod_{k \in \mathcal{D}_r} P_{Wk}} \\ &= \frac{V_{Y_r}^{|\mathcal{D}_r|-1}}{\prod_{k \in \mathcal{D}_r} V_{Y_r|W_k}} \cdot \left|\tilde{Q}^{(\mathcal{D}_r)}\right| \cdot V_{Y_r|\underline{W}^{(\mathcal{D}_r)}} \end{aligned} \tag{50}$$

where $|\mathcal{D}_r|$ is the cardinality of $\mathcal{D}_r$. Thus, for any receiver r we have

$$\left(\frac{V_{Y_r}}{V_{Y_r|\underline{W}^{(\mathcal{D}_r)}}}\right)^{|\mathcal{D}_r|-1} = \prod_{k \in \mathcal{D}_r} \frac{V_{Y_r|W_k}}{V_{Y_r|\underline{W}^{(\mathcal{D}_r)}}}. \tag{51}$$

This equation is the same as

$$I(\underline{W}^{(\mathcal{D}_r)}; Y_r) = \frac{\sum_{k \in \mathcal{D}_r} I(\underline{W}^{(\mathcal{D}_r - \{k\})}; Y_r | W_k)}{|\mathcal{D}_r| - 1} \tag{52}$$

if all the random variables are Gaussian, where $\mathcal{A} - \mathcal{B}$ denotes the set difference $\{a: a \in \mathcal{A}, a \notin \mathcal{B}\}$. Equation (52) has an interesting interpretation for the MAC; this is discussed in Sections V-B and V-C.

C. A Symmetric Case: $\gamma_{rt} = 1$

Suppose we have an interference network with $\gamma_{rt} = 1$ for all r and t. The matrix $\tilde{\Gamma}$ is then the $K \times K$ all-ones matrix and (46) simplifies to

$$\tilde{Q}_{n+1} = F_{-1} \odot \left[\odot \frac{V_{Y_{1n}} \tilde{Q}_n - (\underline{\tilde{c}}_n \underline{\tilde{c}}_n^H) \odot A_n}{\sqrt{\underline{V}_{Y_{1n}|\underline{W}_n} \underline{V}_{Y_{1n}|\underline{W}_n}^T}} \right] \tag{53}$$

where $\underline{\tilde{c}}_n = \tilde{Q}_n \cdot \sqrt{\underline{P}_{Wn}}$, $\underline{V}_{Y_{1n}|\underline{W}_n} = V_{Y_{1n}} \underline{1} - |\underline{\tilde{c}}_n|^2$ and

$$A_n = \underline{1}\,\underline{1}^T + \left(\underline{1}\,\underline{1}^T - M_R M_R^T\right) / V_{Y_{1n}}. \tag{54}$$

We will encounter this simplification for the MAC, and certain broadcast and interference channels.

V. Multiple-Access Channels (MACs)

A MAC is an interference network with $K = T$ and $R = 1$. For convenience, we replace P_t by $P_t |\gamma_{1t}|^2$ and set $\Gamma = [1 \ \cdots \ 1]$. We further simplify notation by replacing Y_1 by Y, W_{kn} by X_{kn}, and so on. The MAC has the symmetry discussed in Section IV-C with $M_R = \underline{1}$. The recursion (53) is thus

$$\tilde{Q}_{n+1} = F_{-1} \odot \left[\odot \frac{V_{Y_n} \tilde{Q}_n - \underline{\tilde{c}}_n \underline{\tilde{c}}_n^H}{\sqrt{\underline{V}_{Y_n|\underline{X}_n} \underline{V}_{Y_n|\underline{X}_n}^T}} \right] \tag{55}$$

where $\underline{\tilde{c}}_n = \tilde{Q}_n \cdot \sqrt{\underline{P}}$. In fact, for the MAC, the full LMMSE estimator and the memoryless LMMSE estimator coincide. This is easily seen by computing

$$\mathrm{E}[Y_n Y_{n-m}^*] = \sum_{k=1}^{K} \sqrt{\frac{P_k}{\sigma_{k(n-1)}^2}}\, m_{kn}^* \, \mathrm{E}\left[\epsilon_{n-1} Y_{n-m}^*\right] = 0$$

for $m > 0$, where the second equality follows by (18). Thus, the Y_n are uncorrelated and the full LMMSE estimator is simply the memoryless one.

A. One Transmitter

If $K = 1$, we have the additive white Gaussian noise (AWGN) channel. The rate (29) is

$$R_{1n} = \tfrac{1}{2} \log(1 + P_1) \tag{56}$$

which is the capacity of the channel without or with feedback [5], [6]. We show in Section VII that (56) implies that MEC coding combined with LMMSE estimation achieves the capacity of the AWGN channel with feedback [4], [16].

B. Two Transmitters

If $K = 2$, we are considering the same problem as Ozarow in [11]. The recursion for the $(1, 2)$ entry of $\tilde{Q}_n$ is

$$\tilde{\rho}_{(12)(n+1)} = -\frac{V_{Y_n} \tilde{\rho}_{(12)n} - \tilde{c}_{1n} \tilde{c}_{2n}^*}{\sqrt{V_{Y_n|X_{1n}} V_{Y_n|X_{2n}}}} \tag{57}$$

which is basically the same as in [11, eq. (11)]. Furthermore, any fixed point must satisfy (51) which is

$$V_Y = V_{Y|X_1} \cdot V_{Y|X_2}.$$

This equation is the same as in [11, eq. (12)] with $\sigma^2 = 1$. Ozarow showed that the MEC strategies attain this fixed point, and that this fixed point gives the maximum possible sum rate. The reason this happens can be seen from (52) which is

$$I(X_1 X_2; Y) = I(X_1; Y | X_2) + I(X_2; Y | X_1). \tag{58}$$

Equation (58) gives the precise condition that yields the best sum-rate upper bound using (10). Thus, Fourier MEC achieves one point on the capacity region boundary. Ozarow further showed that a modification of MEC coding achieves all points inside the entire capacity region.

C. Three and More Transmitters

The case $K \geq 3$ was considered in [21], [22] where improvements over the no-feedback rates are found for $K = 3$. However, the strategies presented there do not achieve capacity and perform worse than the no-feedback capacity for $K > 3$. We show that Fourier MEC performs better in the sense that capacity can be achieved. For simplicity, we consider only the equal-power case, i.e., $P_t = P$ for all t. We claim that $\tilde{Q}_n$ is then circulant for all n; equivalently, $\tilde{Q}_n$ has as eigenvectors the columns of F [30, p. 202]. The claim is certainly true for $n = 1$, since $\tilde{Q}_1$ is the $K \times K$ identity matrix. We proceed by induction and assume that $\tilde{Q}_n$ has the eigenvectors $\underline{f}_{k-1}$ with the corresponding eigenvalues λ_{kn}, $k = 1, \ldots, K$. We then have

$$\underline{\tilde{c}}_n = (\sqrt{P}\lambda_{1n})\underline{1} \tag{59}$$

$$V_{Y_n} = PK\lambda_{1n} + 1 \tag{60}$$

$$\underline{V}_{Y_n|\underline{X}_n} = [P\lambda_{1n}(K - \lambda_{1n}) + 1]\underline{1}. \tag{61}$$

Inserting these identities into (55), we have

$$\tilde{Q}_{n+1} = \frac{F_{-1} \odot \left[(PK\lambda_{1n} + 1)\tilde{Q}_n - P\lambda_{1n}^2 \underline{1}\,\underline{1}^T\right]}{P\lambda_{1n}(K - \lambda_{1n}) + 1} \tag{62}$$

where the Hadamard quotient has become a scalar division. Multiplying by $F^{-1} = F^H/K$ on the left and by F on the right, we find that $F^{-1}\tilde{Q}_{n+1}F$ is a *diagonal* matrix with entries

$$\lambda_{k(n+1)} = \begin{cases} \frac{(PK\lambda_{1n}+1)\lambda_{(k+1)n}}{P\lambda_{1n}(K-\lambda_{1n})+1}, & k = 1, \ldots, K-1 \\ \frac{\lambda_{1n}}{P\lambda_{1n}(K-\lambda_{1n})+1}, & k = K. \end{cases} \tag{63}$$

This result implies that $\tilde{Q}_{n+1}$ also has the columns of F as eigenvectors and that $\tilde{Q}_{n+1}$ is circulant, as claimed. Equation (63) is thus the recursion for the eigenvalues of $\tilde{Q}_n$.

Computer experiments suggest that (63) converges to a fixed point. However, convergence seems difficult to prove (see [11]). Instead of proving convergence, consider the outer bound (14) for $t = K$ and $t = K - 1$

$$\sum_{k=1}^{K} R_k \leq \tfrac{1}{2} \log(1 + PK\lambda) \tag{64}$$

$$\sum_{k=1}^{K} R_k \leq \frac{K}{2(K-1)} \log[1 + P\lambda(K - \lambda)] \tag{65}$$

where $\lambda = 1 + (K-1)\rho$. We prove the following lemma.

Lemma 1: The eigenvalue recursion (63) has a unique fixed point. Furthermore, at this fixed point equality holds in both (64) and (65), where λ is the fixed-point value of λ_{1n}.

Proof: Theorem 1 assures us that a fixed point exists, so let λ_k be a fixed-point value of λ_{kn} for all k. Solving (63) in terms of λ_1 we have, for all k

$$\lambda_k = \frac{\lambda_1}{P\lambda_1(K - \lambda_1) + 1} \left[\frac{PK\lambda_1 + 1}{P\lambda_1(K - \lambda_1) + 1}\right]^{K-k}. \tag{66}$$

The term in square brackets is at least 1, so that λ_1 must be the largest eigenvalue. Taking derivatives with respect to λ_1, we find that λ_k increases with λ_1 when $\lambda_1 \geq 0$. Thus, $\sum_{k=1}^{K} \lambda_k$ increases with λ_1, and there exists a unique positive λ_1 for which $\sum_{k=1}^{K} \lambda_k = K$. This λ_1 gives the desired fixed point.

Next, note that the right-hand sides of (64) and (65) are equal if

$$(PK\lambda + 1)^{K-1} = [P\lambda(K - \lambda) + 1]^K. \tag{67}$$

Setting $k = 1$ in (66), we find that $\lambda = \lambda_1$ is a solution to this equation. The corresponding sum rate is

$$\begin{aligned} \sum_{k=1}^{K} R_k &= \sum_{k=1}^{K} \frac{1}{2} \log\left(V_Y / V_{Y|X_k}\right) \\ &= \frac{1}{2} \log\left(\left[\frac{PK\lambda_1 + 1}{P\lambda_1(K - \lambda_1) + 1}\right]^K\right) \\ &= \frac{1}{2} \log(PK\lambda_1 + 1) \end{aligned} \tag{68}$$

where the first equality follows by (29), the second by (60) and (61), and the third by (67). □

An important next question is whether one can force the recursion (63) into the fixed point of Lemma 1. Appendix C describes a way of doing this: one judiciously chooses the transmit powers for the first $K - 1$ channel uses. This leads to the following capacity result.

Theorem 2: Fourier MEC achieves the feedback sum-rate capacity of the K-transmitter MAC with equal transmitter powers P if P satisfies

$$\left(PK^2/2 + 1\right)^{K-1} \leq \left(PK^2/4 + 1\right)^K. \tag{69}$$

For these powers, the feedback sum-rate capacity is given by (68) where λ_1 is the solution of (67) for $1 \leq \lambda \leq K$. A stronger condition than (69) is $P \geq 2^{K+1}/K^2$.

Proof: Consider the bounds (64) and (65). Observe that $\log(PK\lambda+1)$ increases with λ, $[K/(K-1)]\log[P\lambda(K-\lambda)+1]$ is convex-$\cap$ in λ, and that these two functions intersect for some λ between 1 and K. The latter function is maximized by $\lambda = K/2$. Thus, the fixed point of Lemma 1 achieves the sum-rate capacity if $\lambda_1 \geq K/2$. Equivalently, we require that at $\lambda = K/2$ the bound (64) be at least as tight as (65), which gives (69). The capacity expression (68) follows by the above discussion. Finally, rewriting (69), we have

$$P \geq \frac{1}{K^2/4 + 1/P} \left[\frac{2 + 4/(PK^2)}{1 + 4/(PK^2)}\right]^{K-1}.$$

The term in front of the square brackets is less than $4/K^2$ and the term inside the square brackets is less than 2. Thus, $P \geq 2^{K+1}/K^2$ suffices to ensure capacity is achieved. □

The powers satisfying (69) are shown in Fig. 2 for up to 11 transmitters. We see that these powers are well-approximated by $P = 2^{K+1}/K^2$ for large K. For example, for $K = 8$ and $K = 9$, the smallest powers satisfying (69) are the respective $P = 7.715$ and $P = 12.393$, while the respective $2^{K+1}/K^2$ are 8.000 and 12.642.

Fig. 3 shows the achievable rates of (68) when $P = 10$, as well as the outer bound (14) and the sum-rate capacity without feedback. The capacity with feedback is achieved for $K \leq 8$. For $K > 8$ we do not know whether the achievable rate curve or the outer bound is too loose, or whether both are too loose. However, the bounds are not too far apart and one always improves

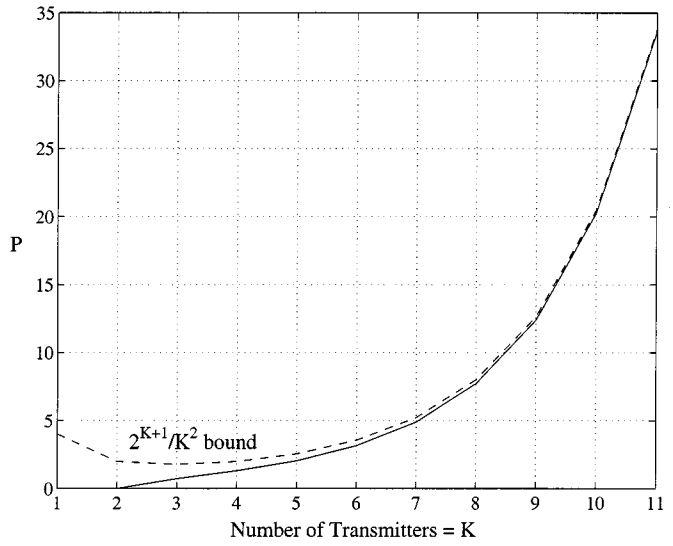


Fig. 2. Power beyond which Fourier MEC is known to achieve the sum-rate capacity for a MAC with equal transmitter powers.

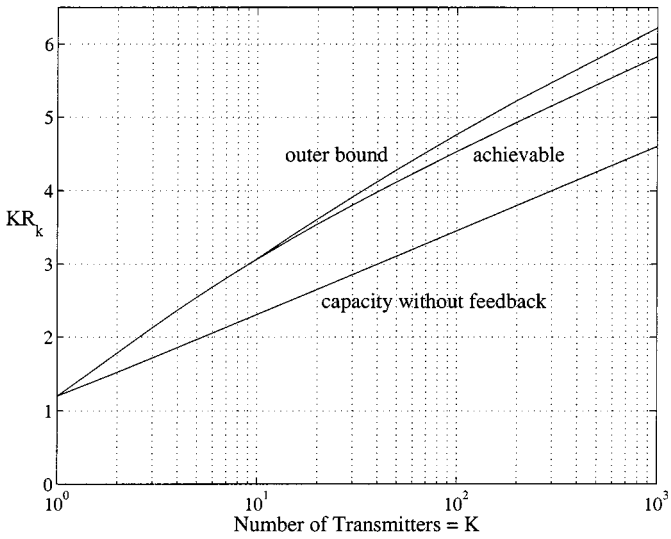


Fig. 3. Sum rate achievable for MACs with $P = 10$. The rate units are nats per use per dimension.

over the no-feedback case. In fact, we prove the following result.

Theorem 3: Consider the MAC with feedback and with equal transmitter powers. The difference between the sum-rate achieved by Fourier MEC and the sum-rate capacity without feedback grows as $(\log \log K)/2$ with K when K is large.

Proof: We bound the fixed-point λ of Lemma 1. Equation (67) can be rewritten as

$$1 = (PK\lambda + 1) \exp\left\{K \log\left[\frac{P\lambda(K-\lambda)+1}{PK\lambda+1}\right]\right\}. \quad (70)$$

Applying $1 - 1/x \le \log(x) \le x - 1$, we have

$$(PK\lambda+1)\exp\left[\frac{-PK\lambda^2}{P\lambda(K-\lambda)+1}\right] \le 1 \le (PK\lambda+1)\exp\left[\frac{-PK\lambda^2}{PK\lambda+1}\right].$$

We use $P\lambda(K-\lambda)+1 > P\lambda(K-\lambda)$ and $1 \le \lambda \le K$ to obtain

$$PK \exp\left[\frac{-K\lambda}{K-\lambda}\right] < 1 \le (PK^2+1)\exp\left[\frac{-PK\lambda}{PK+1}\right].$$

These bounds can be rewritten as

$$\frac{K\log(PK)}{K+\log(PK)} < \lambda \le \left(\frac{PK+1}{PK}\right)\log(PK^2+1). \quad (71)$$

Thus, for large K (and $P > 0$) we have $\log \lambda \approx \log \log K$. Inserting this result into (68) we have

$$\sum_{k=1}^{K} R_k \approx \frac{1}{2}\log K + \frac{1}{2}\log\log K. \quad (72)$$

The difference between (72) and the no-feedback capacity $[\log(1+PK)]/2 \approx (\log K)/2$ grows as $(\log \log K)/2$. □

This theorem means that the feedback sum-rate capacity is *at least* $(\log \log K)/2$ larger than the no-feedback sum-rate capacity for large K. On the other hand, (15) implies that the sum-rate capacity increases by *at most* $(\log K)/2$.

Finally, we comment on why the achievable sum-rate of Fourier MEC and the upper bound (14) are not the same for large K (or low SNR). Consider (52) for $K = 3$ which is

$$I(X_1X_2X_3;Y) = \tfrac{1}{2}[I(X_1X_2;Y|X_3) + I(X_1X_3;Y|X_2) + I(X_2X_3;Y|X_1)]. \quad (73)$$

Fourier MEC achieves the rate point satisfying (73). Furthermore, for high SNR, (73) is the precise condition that gives the best sum-rate upper bound using (14). However, for low SNR, the appropriate condition using (14) becomes

$$\frac{1}{2}[I(X_1X_2;Y|X_3) + I(X_1X_3;Y|X_2) + I(X_2X_3;Y|X_1)] = I(X_1;Y|X_2X_3) + I(X_2;Y|X_1X_3) + I(X_3;Y|X_1X_2).$$

The rates satisfying (73) might therefore be suboptimal. However, we emphasize again that we do not know whether the achievable rates of Fourier MEC or the outer bound is too loose, or whether both are too loose. It is certainly plausible that Fourier MEC achieves capacity for all SNRs.

VI. Other Interference Networks

A. Broadcast Channels

A broadcast channel is an interference network with $K = R$ and $T = 1$ so that $\Gamma = [\gamma_{11}\ \gamma_{21}\ \dots\ \gamma_{R1}]^T$. The two-receiver case was considered in [20], where the performance of the memoryless LMMSE estimator was analyzed. Other results are reviewed in [31, Sec. 4].

We simplify notation by replacing γ_{r1} by γ_r, P_1 by P, and so on. We will consider only the case $\gamma_r = 1$ for simplicity. The channel then has the symmetry discussed in Section IV-C with $M_R = I$. The correlation coefficient recursion is given by (53) with $V_{Y_{1n}} = P + 1$ and

$$A_n = \begin{bmatrix} 1 & a & \cdots & a \\ a & 1 & \cdots & a \\ \vdots & & \ddots & \vdots \\ a & a & \cdots & 1 \end{bmatrix} \quad (74)$$

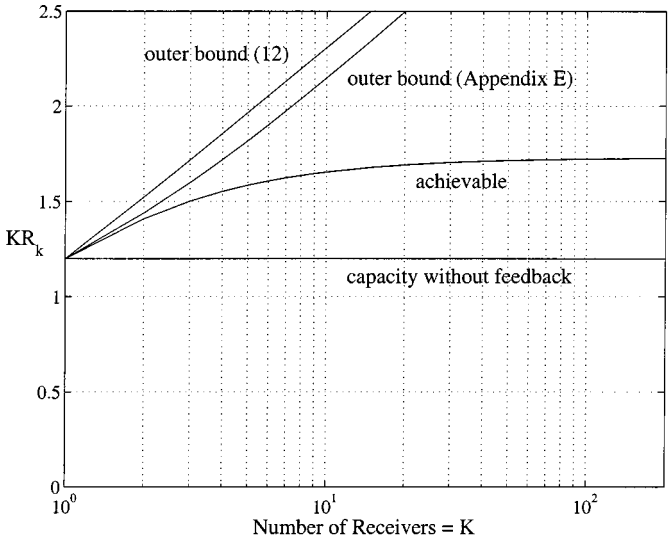


Fig. 4. Sum rate achievable for broadcast channels with $\gamma_r = 1$ and $P = 10$. The rate units are nats per use per dimension.

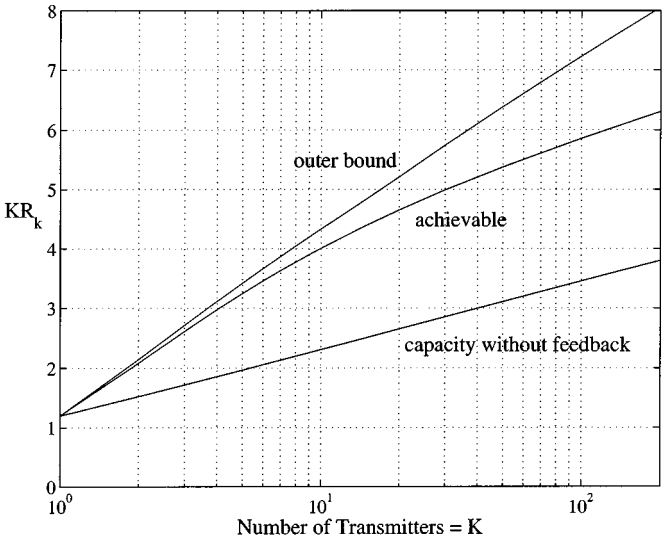


Fig. 5. Sum rate achievable for interference channels with $\gamma_{rt} = 1$ and $P = 10$. The rate units are nats per use per dimension.

where $a = (P+2)/(P+1)$. One can check that for the broadcast channel the full LMMSE estimator of Section III is *not* the same as the memoryless one.[1] One might, therefore, expect the memoryless estimator to be suboptimal, and in fact [20] shows that there are better estimators.

We continue by distributing the transmit power equally among the K messages, i.e., $\underline{P}_{W0} = \underline{1}$. One again finds that $\tilde{Q}_n$ is circulant for all n (see [32]). The corresponding eigenvalue recursion is

$$\lambda_{k(n+1)} = \begin{cases} \frac{(P+1)\lambda_{(k+1)n} - P(1-a)\lambda_{1n}/K}{(P+1) - P\lambda_{1n}/K}, & k < K \\ \frac{(P+1)\lambda_{1n} - P[1+(K-1)a]\lambda_{1n}/K}{(P+1) - P\lambda_{1n}/K}, & k = K \end{cases} \tag{75}$$

where we have used $V_{Y_{1n}|W_{1n}} = (P+1) - P\lambda_{1n}/K$.

Computer experiments suggest that (75) converges to a fixed point and Fig. 4 shows the resulting rates when $P = 10$. In Appendix D, we prove that the sum rate of any fixed point increases with K and approaches 1.7306 nats/use/dimension. The sum rate without feedback is $\log(11)/2 = 1.1989$ nats/use/dimension, so that feedback boosts the sum rate by up to 44%. Fig. 4 also depicts the outer bound (12) and a slightly better outer bound derived in Appendix E. The latter bound is a generalization of a bound in [20]. However, neither bound seems useful for large K.

B. Interference Channels

An interference channel is an interference network with $K = T = R$ for which transmitter t sends U_t to receiver t (see [7, Sec. 14], [13]–[15]). We will again consider only $P_t = P$ for all t.

The first case that we will investigate is where Γ is the all-ones matrix. The channel then has the symmetry discussed in Section IV-C with $M_R = I$. The correlation coefficient recursion is given by (53) with A_n having the same form as (74) except that a is replaced by $a_n = 1 + 1/V_{Y_{1n}}$. One can again show that $\tilde{Q}_n$ is circulant for all n and that the eigenvalue recursion is

$$\lambda_{k(n+1)} = \begin{cases} \frac{(PK\lambda_{1n}+1)\lambda_{(k+1)n} - P(1-a_n)\lambda_{1n}^2}{P\lambda_{1n}(K-\lambda_{1n})+1}, & k < K \\ \frac{(PK\lambda_{1n}+1)\lambda_{1n} - P[1+(K-1)a_n]\lambda_{1n}^2}{P\lambda_{1n}(K-\lambda_{1n})+1}, & k = K \end{cases} \tag{76}$$

where we have used $V_{Y_{kn}} = PK\lambda_{1n} + 1$ and

$$V_{Y_{kn}|W_{kn}} = P\lambda_{1n}(K - \lambda_{1n}) + 1, \qquad \text{for all } k.$$

It is not surprising that computer experiments suggest that (76) converges to a fixed point. Properties of such a fixed point are derived in Appendix F, where it is shown that for large K the sum rate is approximately $(\log K)/2 + \log\log K$. This rate is about $\log\log K$ larger than the sum-rate capacity without feedback, which is $\log(1+PK)/2 \approx (\log K)/2$.

Fig. 5 depicts the sum rates for $P = 10$, as well as the sum rate capacity without feedback and the outer bound computed using (14). Observe that feedback increases capacity more than for the MAC. For example, for $K = 10$ the sum rate increases by about 74% (for the MAC the corresponding increase is about 33%). The outer bound shows that not much improvement is possible over Fourier MEC for $P = 10$ and $K \leq 10$.

Consider next the case $K = 2$ and

$$\Gamma = \begin{bmatrix} 1 & \gamma \\ \gamma & 1 \end{bmatrix}$$

where γ is a nonnegative *real* number. The recursion for $\tilde{\rho}_n := \tilde{\rho}_{(12)n}$ is (see (46))

$$\tilde{\rho}_{n+1} = -\frac{V_{Y_{1n}}}{V_{Y_{1n}|X_{1n}}} \left[\tilde{\rho}_n - 2\frac{P(1+\gamma\tilde{\rho}_n)^2}{V_{Y_{1n}}} + \frac{P^2(1+\gamma\tilde{\rho}_n)^2(2\gamma + \tilde{\rho}_n + \gamma\tilde{\rho}_n)}{V_{Y_{1n}}^2} \right] \tag{77}$$

where

$$V_{Y_{1n}} = P(1 + \gamma^2 + 2\gamma\tilde{\rho}_n) + 1$$

and

$$V_{Y_{1n}|X_{1n}} = P\gamma^2(1 - \tilde{\rho}_n^2) + 1.$$

[1]The difference between the full LMMSE estimator and the memoryless one is not mentioned in [20]. However, Ozarow told the author that he had been aware of the difference.

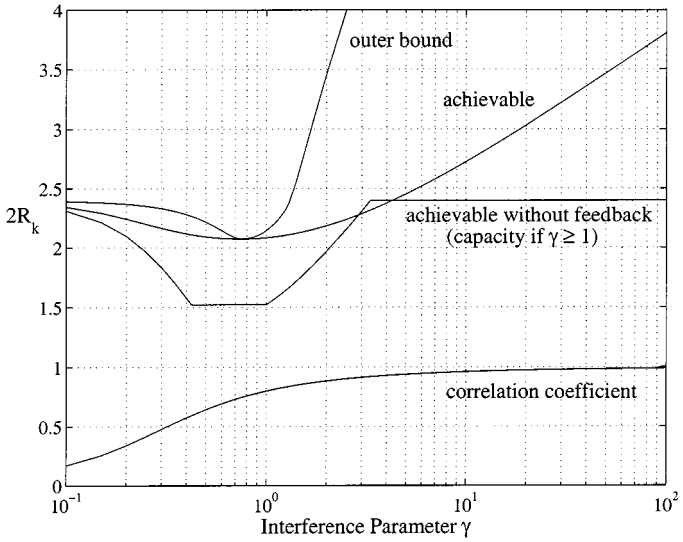


Fig. 6. Sum rate achievable for interference channels with $K = 2$ and $P = 10$. The rate units are nats per use per dimension.

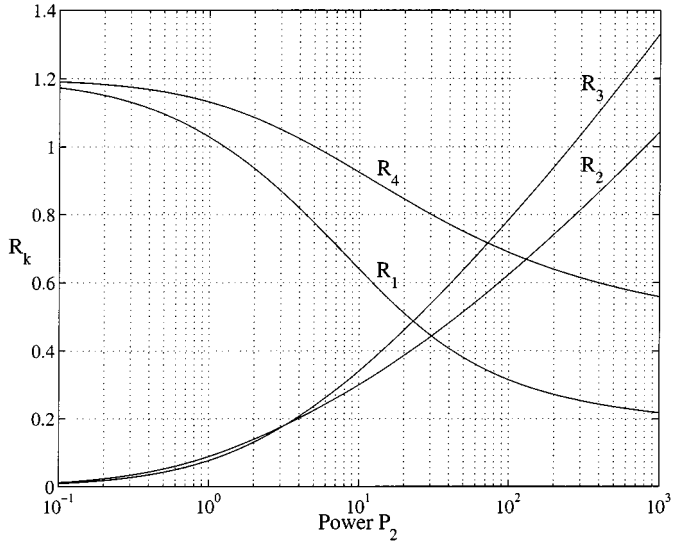


Fig. 7. Rates achievable for the network of Fig. 1 with $P_1 = P_3 = 10$ and $\underline{P}_{W0}^{(\{2\})} = [1\ 3]^T$. The rate units are nats per use per dimension.

Again, computer experiments suggest that (77) converges to a fixed point. Fig. 6 depicts the resulting rates as a function of the interference parameter γ. Observe that the sum rate decreases from $\log(11) = 2.3979$ when $\gamma = 0$ to some minimum, and then keeps increasing as γ increases. This behavior is similar to the no-feedback curve except when γ is large. We remark that the sum rate capacity without feedback is known only if $\gamma = 0$ or $\gamma \geq 1$. For $0 < \gamma < 1$ we have plotted the best known achievable rates (see [15] and [33, Fig. 4]; other interesting results can be found in [34]–[36]). Note that our strategies are not optimal since they are not as good as the best no-feedback codes for some γ. However, one might be able to improve the rates by using the full LMMSE estimator rather than the memoryless one.

Fig. 6 also shows the correlation coefficients $\tilde{\rho}_n$. Note that $\tilde{\rho}_n$ approaches 1 as γ increases. This is not surprising when one considers that a good strategy for large γ is to let each transmitter route its message via the other transmitter. For example, transmitter 1 could send its message to receiver 2, which then sends this message to transmitter 2 via the noiseless feedback link, and finally transmitter 2 sends the message to receiver 1. Fourier MEC mimics this procedure by building up correlation between the transmitted symbols. This approach could be termed *correlation routing*, and it is most effective when there is very strong interference. For instance, for $\gamma = 1000$ the sum rate 4.9510 nats/use/dimension is achievable which more than doubles the capacity without feedback.

Finally, Fig. 6 shows the outer bound computed from (16). The bound does not seem useful for large γ. However, the bound seems to touch the achievable rate curve when both curves are at their minimum. If so, this would mean that Fourier MEC achieves the sum rate capacity for some γ satisfying $0 < \gamma < 1$ when $P = 10$ (again, the capacity without feedback is currently not known for any such γ).

C. A Four-Message, Three-Transmitter, Two-Receiver Network

The last network we will consider is that of Fig. 1. We demonstrate that the recursion (32) "works" for nonequal powers and rates by considering the case with $P_1 = P_3 = 10$, $\underline{P}_{W0}^{(\{2\})} = [1\ 3]^T$, and variable P_2. Computer experiments suggest that (76) converges to a fixed point, and Fig. 7 shows how the resulting R_k change as a function of P_2. Observe that $R_1 \leq R_4$ because the interference caused by X_2 is larger at Y_1 than it is at Y_3. For small P_2, we have $R_2 \approx R_3$ because, although we have allocated U_3 three times more power than U_2, the gain $\gamma_{22} = 1/2$ reduces this power by a factor of four. For large P_2, we have $R_2 < R_3$; this is probably because U_2 gets less power than U_3, and because X_1 and X_2 interfere with each other more than X_2 and X_3.

VII. Error Probability

Let $\overline{R_k}$ be the average rate $\sum_{n=1}^{N} R_{kn}/N$. We have, using (29)

$$\sigma_{kN}^2 = \sigma_{k0}^2 \cdot \exp(-2N\overline{R_k}). \tag{78}$$

An error occurs if the message point θ_k is mistaken for another message point θ_k', where $\theta_k' \neq \theta_k$. We place the points on a square grid whose corners are at $(\pm 1 \pm j)/2$, and choose the distance between neighboring points as $\Delta = (2^{B_k} - 1)^{-1}$ so that there are 2^{2B_k} points in total. The data rate is thus B_k/N bits per use per dimension, i.e., $R_k = \log(2)B_k/N$ nats per use per dimension.

After the Nth channel use we decode by choosing the message point closest to $\hat{\theta}_{kN}$. The error probability P_e is then bounded by

$$P_e \leq \Pr\left(|\epsilon_{kN}| \geq \Delta/2\right) \tag{79}$$

where we have an inequality because the region of points closest to θ_k is larger than the circle of radius $\Delta/2$ about θ_k. Moreover, we know that the variance of ϵ_{kN} is σ_{kN}^2, so applying Chebyshev's inequality [24, p. 127] we have

$$\begin{aligned} P_e &\leq 4\sigma_{kN}^2/\Delta^2 \\ &\leq 4\sigma_{k0}^2 \cdot \exp\left(-2N[\overline{R_k} - \log(2)B_k/N]\right) \\ &= 4\sigma_{k0}^2 \cdot \exp\left(-2N(\overline{R_k} - R_k)\right). \end{aligned} \tag{80}$$

Observe that if $R_k < \overline{R_k}$ then the error probability decreases exponentially with N. Thus, we can transmit messages at rates up to $\overline{R_k}$ with arbitrarily small positive error probability. The error exponent is simply $2(\overline{R_k} - R_k)$.

A. Guassian-Distributed Message points

The bound (80) is valid for any message distribution with finite variance. Now suppose that θ_k is *Gaussian* and that an error is said to occur if $\hat{\theta}_{kN}$ is not in the circle of radius $\Delta/2$ about θ_k. We can then write (79) as (see [37, p. 55])

$$\begin{aligned} P_e &\le \exp\left(-\frac{\Delta^2}{8\sigma_{kN}^2}\right), \\ &\le \exp\left[-\frac{1}{8\sigma_{k0}^2}\exp\left(2N(\overline{R_k}-R_k)\right)\right]. \end{aligned} \tag{81}$$

Thus, for $R_k < \overline{R_k}$ the error probability decreases *doubly* exponentially with N. This type of error decrease was discovered in [4]. However, in practice one will have *peak*-power constraints, in which case Wyner showed that only a single-exponential decrease of P_e with N is possible [1].

VIII. Concluding Remarks

Fourier MEC was shown to perform well for several interference networks with strong symmetries. In particular, the strategies achieve the sum-rate capacity for MACs as long as the transmitter powers are all the same and beyond some threshold. However, up to now we were unsuccessful in analyzing the same networks without these symmetries. We list some of the problems we feel deserve further study.

- The convergence of the recursion (32) and/or (46) should be characterized. Interestingly, these recursions converged for every interference network we considered.
- The sum-rate capacity of the MAC with symmetric and *low* powers is still open. The solution might require better outer bounds than (14).
- The MAC with symmetric powers and *asymmetric rates* should be considered in more detail. For example, it would be interesting if the natural extension of the approach used in [11, p. 626] achieved all points in the capacity region of the MAC for powers beyond some threshold.
- Outer bounds and achievability results for the MAC, broadcast channel, and interference channel with *asymmetric powers* should be derived.
- Noisy feedback should be addressed. Perhaps strategies related to Fourier MEC can improve transmission rates for networks with imperfect feedback.

The following approaches may prove useful in solving some of these problems.

- We expect the full LMMSE estimator to give better rates than the memoryless one for the broadcast and interference channels. Alternatively, one may wish to consider LMMSE estimators with memory M, i.e., the receivers produce their $\hat{\theta}_{kn}$ by estimating $\hat{\epsilon}_{k(n-1)}$ using $Y_{rn}, Y_{r(n-1)}, \ldots Y_{r(n-M)}$.
- Section II-B explained how, in going from real to complex networks, one is basically grouping symbols into blocks of length 2. Extending this idea, one could group complex symbols into blocks of any length and use *unitary* matrices to "rotate" these blocks. Observe that any unitary matrix can be written in the form e^{jA}, where A is a Hermitian matrix [28, p. 486]—compare this with (47). Unitary matrices as "modulation coefficients" may also prove useful for interference networks with *vector* inputs and outputs.

Finally, we remark that many of the results presented here can be applied to other channels. For example, MEC applies unchanged to channels with average *symbol*, rather than *block*, power constraints. Furthermore, much of the analysis remains the same for channels with non-Gaussian noise as long as the channel has a real (or complex) input alphabet. For example, suppose that the noise Z_{rn} has Laplacian statistics. Nothing prevents one from using MEC, and all of the results in Sections III, IV, and VII apply since they do not require Z_{rn} to be Gaussian.

Appendix A
LMMSE Recursions

The variances for the *full* LMMSE estimator are, using (6)

$$\begin{aligned} \sigma_{kn}^2 &= V_{\theta_k|\underline{Y}_r^n} = \frac{V_{\theta_k \underline{Y}_r^n}}{V_{\underline{Y}_r^n}} \\ &= \frac{V_{\underline{Y}_r^{n-1}} \cdot V_{\theta_k|\underline{Y}_r^{n-1}} \cdot V_{Y_{rn}|\theta_k \underline{Y}_r^{n-1}}}{V_{\underline{Y}_r^{n-1}} \cdot V_{Y_{rn}|\underline{Y}_r^{n-1}}} \\ &= \sigma_{k(n-1)}^2 \cdot \frac{V_{Y_{rn}|\theta_k \underline{Y}_r^{n-1}}}{V_{Y_{rn}|\underline{Y}_r^{n-1}}}. \end{aligned}$$

This proves (22). Next, let

$$\hat{\epsilon}_{k(n-1)} = \mathrm{E}[\epsilon_{k(n-1)}(\underline{Y}_r^n)^H] \cdot Q_{\underline{Y}_r^n}^{-1} \cdot \underline{Y}_r^n$$

be the LMMSE estimate of $\epsilon_{k(n-1)}$ given $\underline{Y}_r^n$. It is not difficult to show that $\hat{\theta}_{kn} = \hat{\theta}_{k(n-1)} - \hat{\epsilon}_{k(n-1)}$ so that

$$\epsilon_{kn} = \hat{\theta}_{kn} - \theta_k = \epsilon_{k(n-1)} - \hat{\epsilon}_{k(n-1)}.$$

For the correlation coefficients, we thus have

$$\begin{aligned} \mathrm{E}[\epsilon_{kn}\epsilon_{k'n}^*] &= \mathrm{E}\left[[\epsilon_{k(n-1)} - \hat{\epsilon}_{k(n-1)}][\epsilon_{k'(n-1)} - \hat{\epsilon}_{k'(n-1)}]^*\right] \\ &= \sqrt{\sigma_{k(n-1)}^2 \sigma_{k'(n-1)}^2} \cdot \Big[\rho_{(kk')(n-1)} \\ &\quad - \underline{c}_{(kr')n}^T Q_{\underline{Y}_{r'}^n}^{-1} \underline{c}_{(k'r')n}^* - \underline{c}_{(kr)n}^T Q_{\underline{Y}_r^n}^{-1} \underline{c}_{(k'r)n}^* \\ &\quad + \underline{c}_{(kr)n}^T Q_{\underline{Y}_r^n}^{-1} \mathrm{E}[(\underline{Y}_r^n)(\underline{Y}_{r'}^n)^H] Q_{\underline{Y}_{r'}^n}^{-1} \underline{c}_{(k'r')n}^*\Big] \end{aligned}$$

where

$$\underline{c}_{(k\tilde{r})n} = \mathrm{E}[\epsilon_{k(n-1)}(\underline{Y}_{\tilde{r}}^n)^*] \Big/ \sqrt{\sigma_{k(n-1)}^2}.$$

From this one has (25).

For the *memoryless* LMMSE estimator, we have

$$\begin{aligned} \sigma_{kn}^2 &= V_{\epsilon_{k(n-1)}|Y_{rn}} = \frac{V_{\epsilon_{k(n-1)} Y_{rn}}}{V_{Y_{rn}}} \\ &= V_{\epsilon_{k(n-1)}} \cdot \frac{V_{Y_{rn}|\epsilon_{k(n-1)}}}{V_{Y_{rn}}} \\ &= \sigma_{k(n-1)}^2 \cdot \frac{V_{Y_{rn}|W_{kn}}}{V_{Y_{rn}}} \end{aligned}$$

where we have used $V_{\underline{A}|c\cdot\underline{B}} = V_{\underline{A}|\underline{B}}$ for any constant c. This proves (28). In (25), we can replace $Q_{\underline{Y}_r^n}$ by $V_{Y_{rn}}$, and $\underline{c}_{(k\tilde{r})n}$ by

$$c_{(k\tilde{r})n} = \mathrm{E}[\epsilon_{k(n-1)} Y_{\tilde{r}n}^*] \Big/ \sqrt{\sigma_{k(n-1)}^2}.$$

This gives

$$\rho_{(kk')n} = \exp(R_{kn} + R_{k'n}) \cdot \left[\rho_{(kk')(n-1)} - \frac{c_{(kr')n} c_{(k'r')n}^*}{V_{Y_{r'n}}} - \frac{c_{(kr)n} c_{(k'r)n}^*}{V_{Y_{rn}}} + \frac{c_{(kr)n} c_{(k'r')n}^*}{V_{Y_{rn}} V_{Y_{r'n}}} \mathrm{E}[Y_{rn} Y_{r'n}^*] \right]$$

which is the same as (30). Furthermore

$$\begin{aligned} \mathrm{E}[Y_{rn} Y_{r'n}^*] &= \delta_{rr'} + \sum_{k,\,k'} \gamma_{rk}\, \mathrm{E}[W_{kn} W_{k'n}^*] \gamma_{r'k'}^* \\ &= \delta_{rr'} + \sum_{k,\,k'} \left(\gamma_{rk}^* \sqrt{P_{Wkn}} m_{kn} \right)^* \rho_{kk'(n-1)} \\ &\quad \cdot \left(\gamma_{r'k'}^* \sqrt{P_{Wk'n}} m_{k'n} \right) \end{aligned}$$

where $\delta_{rr'} = 1$ if $r = r'$ and is 0 otherwise. This implies (37). Finally, we have

$$\begin{aligned} V_{Y_{rn}|W_{kn}} &= V_{Y_{rn}} - \frac{|\mathrm{E}[W_{kn} Y_{rn}^*]|^2}{P_{Wkn} |m_{kn}|^2} \\ &= V_{Y_{rn}} - |\mathrm{E}[\epsilon_{kn} Y_{rn}^*]|^2 / \sigma_{k(n-1)}^2 \\ &= V_{Y_{rn}} - |c_{(kr)n}|^2. \end{aligned}$$

From this, one obtains (39).

Appendix B
Computing a Matrix Recursion

The steps required to compute (32) are listed below. Note that (32) does not necessarily give a correlation coefficient matrix if Q_{n-1} is not a correlation coefficient matrix. For example, consider the AWGN channel ($T = R = K = 1$) for which $Q_n = 1$ for all n. Suppose that for some reason, e.g., round-off error, we had $Q_{n-1} = 1 - \delta_{n-1}$. Then (55) is

$$\begin{aligned} Q_n &= \frac{(P+1)Q_{n-1} - P Q_{n-1}^2}{(P+1) - P Q_{n-1}^2} \\ &= (1 - \delta_{n-1}) \frac{1 + P\delta_{n-1}}{1 + P\delta_{n-1}(2 - \delta_{n-1})} \end{aligned}$$

and if $0 < \delta_{n-1} < 1$ we have $1 > Q_{n-1} > Q_n$. In fact, we find that $\delta_n \to 1$ and $Q_n \to 0$ as n increases. This problem can be avoided by ensuring that Q_n is a correlation coefficient matrix for all n. We did this by

- setting the diagonals of Q_n to be 1 and
- replacing Q_n by $(Q_n + Q_n^H)/2$

after each recursion update. The second step ensures that Q_n is Hermitian.

Initialize:

- set $n = 1$ and $Q_0 = I$
- compute $\tilde{\Gamma} = M_R \cdot \Gamma \cdot M_T^T$
- choose P_t and a power distribution $\underline{P}_{W0}^{(\mathcal{E}_t)}$, $t = 1, \ldots, T$.

Step n:

- Choose the modulation coefficients $\underline{m}_n$
- Ensure the power constraints are met for $t = 1, \ldots, T$

$$\tilde{P}_{tn} = \left(\sqrt{\underline{P}_{W0}^{(\mathcal{E}_t)}} \odot \underline{m}_n^{(\mathcal{E}_t)} \right)^H \cdot Q_{n-1}^{(\mathcal{E}_t)} \cdot \left(\sqrt{\underline{P}_{W0}^{(\mathcal{E}_t)}} \odot \underline{m}_n^{(\mathcal{E}_t)} \right)$$

$$\underline{P}_{Wn}^{(\mathcal{E}_t)} = \underline{P}_{W0}^{(\mathcal{E}_t)} \cdot (P_t / \tilde{P}_{tn}).$$

- Update the matrices:

$$\begin{aligned} G_n &= \tilde{\Gamma}^H \odot \left[\left(\sqrt{\underline{P}_{Wn}} \odot \underline{m}_n \right) \cdot \underline{1}^T \right] \\ C_n &= Q_{n-1} \cdot G_n \\ D_n &= \underline{1} \cdot \mathrm{diag}\,(C_n)^H \\ Q_{\underline{Y}_n} &= G_n^H\, Q_{n-1}\, G_n + M_R\, M_R^T \\ \underline{V}_{\underline{Y}_n} &= \mathrm{diag}\,(Q_{\underline{Y}_n}) \\ \underline{V}_{\underline{Y}_n|\underline{W}_n} &= \underline{V}_{\underline{Y}_n} - |\mathrm{diag}\,(C_n)|^2 \\ \underline{R}_n &= \frac{1}{2} \log \left(\odot \frac{\underline{V}_{\underline{Y}_n}}{\underline{V}_{\underline{Y}_n|\underline{W}_n}} \right). \end{aligned}$$

- Update the correlation coefficient matrix $Q_n = f(Q_{n-1}, \underline{m}_n)$, where $f(\cdot)$ is given by (32).
- Increment n by one.

Appendix C
Powers for MAC Eigenvalues

Suppose we want the recursion (63) to yield some desired values $\lambda_{1K}, \lambda_{2K}, \ldots, \lambda_{KK}$ at time $n = K$. We replace P by P_n at time n and apply (63) $K - 1$ times to get

$$\lambda_{kK} = \frac{1}{D} \prod_{\substack{n=1 \\ n \neq k-1}}^{K-1} [P_n K \lambda_{1n} + 1]$$

for all k, where

$$D = \prod_{n=1}^{K-1} [P_n \lambda_{1n} (K - \lambda_{1n}) + 1].$$

We thus have, for $n = 1, 2, \ldots, K-1$

$$\lambda_{1K} / \lambda_{(n+1)K} = P_n K \lambda_{1n} + 1.$$

Solving for P_n, we have

$$P_n = \frac{1}{K \lambda_{1n}} \left(\lambda_{1K} / \lambda_{(n+1)K} - 1 \right) \tag{82}$$

where λ_{1n} can be computed recursively via $\lambda_{11} = 1$ and

$$\lambda_{1n} = \prod_{m=1}^{n-1} \frac{P_m K \lambda_{1m} + 1}{P_m \lambda_{1m} (K - \lambda_{1m}) + 1}. \tag{83}$$

For instance, suppose that $K = 3$ and that we want to be in the $P = 10$ fixed-point $\lambda_{13} = 2.3055$, $\lambda_{23} = 0.5590$, $\lambda_{33} = 0.1355$ after two channel uses. Repeated application of (82) and (83) gives

$$\begin{aligned} n = 1: \quad & P_1 = 1.0414, \\ & \lambda_{12} = 1.3378, \quad \lambda_{22} = 1.3378, \quad \lambda_{32} = 0.3244; \\ n = 2: \quad & P_2 = 3.9903, \\ & \lambda_{13} = 2.3055, \quad \lambda_{23} = 0.5590, \quad \lambda_{33} = 0.1355. \end{aligned}$$

APPENDIX D
BROADCAST CHANNEL FIXED POINT PROPERTIES

Using (75), any fixed point satisfies

$$\lambda_k = \frac{(P+1)\lambda_{k+1} + P\lambda_1/(K(P+1))}{(P+1) - P\lambda_1/K}$$

for $1 \le k \le K-1$. Rewriting this equation, we obtain

$$\lambda_{k+1} = \lambda_k \left(1 - \frac{P\lambda_1}{K(P+1)}\right) - \frac{P\lambda_1}{K(P+1)^2}.$$

Solving in terms of λ_1 for all k, we have

$$\lambda_k = \left(\lambda_1 + \frac{1}{P+1}\right)\left(1 - \frac{P\lambda_1}{K(P+1)}\right)^{k-1} - \frac{1}{P+1}. \quad (84)$$

Note that the λ_k decrease with k so that λ_1 is the largest eigenvalue. Equation (75) also specifies that

$$\lambda_K = \frac{(P+1)\lambda_1 - P\left[1 + (K-1)\frac{P+2}{P+1}\right]\lambda_1/K}{(P+1) - P\lambda_1/K}. \quad (85)$$

We set $k = K$ in (84) and equate the result with (85) to obtain

$$\lambda_1 = (P+1)^2\left(\lambda_1 + \frac{1}{P+1}\right)\left[1 - \frac{P\lambda_1}{(P+1)K}\right]^K - (P+1).$$

Observe that the term in square brackets increases with K, and is bounded from above by $\exp[-P\lambda_1/(P+1)]$. Thus, λ_1 increases with K which means that the sum rate increases with K. Furthermore, λ_1 approaches from below that $\tilde{\lambda}_1$ satisfying

$$\tilde{\lambda}_1 = (P+1)^2\left(\tilde{\lambda}_1 + \frac{1}{P+1}\right)\exp\left(\frac{-P\tilde{\lambda}_1}{P+1}\right) - (P+1). \quad (86)$$

For example, for $P = 10$ we have $\tilde{\lambda}_1 = 3.8073$. Thus, for $K > (P+1)/(P\tilde{\lambda}_1)$ the sum rate is bounded by

$$\sum_{k=1}^{K} R_k = K \cdot \frac{1}{2}\log\left(V_{Y_1}/V_{Y_1|W_1}\right) < \frac{K}{2}\log\left(\frac{P+1}{(P+1) - P\tilde{\lambda}_1/K}\right). \quad (87)$$

But because the right-hand side of (87) decreases with K and the sum rate increases with K, a better outer bound for any finite K is to determine the sum rate that will be approached as K increases:

$$\sum_{k=1}^{K} R_k < \lim_{\tilde{K}\to\infty}\frac{\tilde{K}}{2}\log\left(\frac{P+1}{(P+1) - P\tilde{\lambda}_1/\tilde{K}}\right) = \frac{P\tilde{\lambda}_1}{2(P+1)}. \quad (88)$$

For example, for $P = 10$ the sum rate will approach 1.7306 from below as K increases.

APPENDIX E
A BROADCAST CHANNEL CAPACITY OUTER BOUND

An outer bound on the $K = 2$ capacity region was given in [20], where the channel was made *physically* degraded by giving receiver 1 both Y_1 and Y_2. The extra information of receiver 1 obviously enlarges the capacity region. Furthermore, the capacity region of a physically degraded broadcast channel is not increased by feedback (the proofs of [38] and [39] can be extended to vector reception) and is the set of rate pairs (R_1, R_2) satisfying

$$0 \le R_1 \le \frac{1}{2}\log\left(1 + P\alpha(|\gamma_1|^2 + |\gamma_2|^2)\right)$$
$$0 \le R_2 \le \frac{1}{2}\log\left(1 + \frac{P(1-\alpha)|\gamma_2|^2}{1 + P\alpha|\gamma_1|^2}\right)$$

where α takes on any value between 0 and 1.

One can generalize this outer bound for $K > 2$. The idea is to give receiver k the outputs $Y_k, Y_{k+1}, \ldots, Y_K$ to get a K-receiver physically degraded broadcast channel. The capacity region of this channel with and without feedback is the set of rate-tuples $(R_1, \ldots, R_K)$ satisfying

$$0 \le R_k \le \frac{1}{2}\log\left(1 + \frac{P\alpha_k\sum_{\ell=k}^{K}|\gamma_\ell|^2}{1 + P\sum_{\ell=1}^{k-1}\alpha_\ell|\gamma_\ell|^2}\right) \quad (89)$$

for $k = 1, \ldots, K$, $\alpha_k \ge 0$, and $\sum_{k=1}^{K}\alpha_k = 1$ (the proof of this result follows from generalizations of the results of [39]–[41]).

We would like to find the best equal-rate point in (89). We do this by making (89) the same for all k

$$P\alpha_1\sum_{\ell=1}^{K}|\gamma_\ell|^2 = \frac{P\alpha_k\sum_{\ell=k}^{K}|\gamma_\ell|^2}{1 + P\sum_{\ell=1}^{k-1}\alpha_\ell|\gamma_\ell|^2}$$

for $2 \le k \le K$. From this equation we have

$$\alpha_k = \alpha_1\left(\sum_{\ell=1}^{K}|\gamma_\ell|^2\right)\left[1 + P\sum_{\ell=1}^{k-1}\alpha_\ell|\gamma_\ell|^2\right]\Big/\left(\sum_{\ell=k}^{K}|\gamma_\ell|^2\right).$$

Note that the α_k increase with k, and that they are all zero if $\alpha_1 = 0$. Thus, there is a unique α_1, $0 < \alpha_1 \le 1/K$, so that $\sum_{k=1}^{K}\alpha_k = 1$. This α_1 can be found by binary search and gives an outer bound on the equal rates.

APPENDIX F
INTERFERENCE CHANNEL FIXED POINT PROPERTIES

Consider the interference channel where Γ is the all-ones matrix. We use the same approach as in Appendix D to find that any fixed point of (76) has

$$\lambda_k = \left(\lambda_1 + \frac{1}{PK\lambda_1 + 1}\right)\left(1 - \frac{P\lambda_1^2}{PK\lambda_1 + 1}\right)^{k-1} - \frac{1}{PK\lambda_1 + 1}. \quad (90)$$

for all k. Again, observe that the λ_k decrease with k so that λ_1 is the largest eigenvalue. Equation (76) also specifies that

$$\lambda_K = \frac{(PK\lambda_1 + 1)\lambda_1 - P\left[1 + (K-1)\frac{PK\lambda_1 + 2}{PK\lambda_1 + 1}\right]\lambda_1^2}{P\lambda_1(K - \lambda_1) + 1}. \quad (91)$$

We set $k = K$ in (90) and equate the result with (91) to obtain

$$\lambda_1 = \frac{(PK\lambda_1 + 1)^2}{PK + 1}\cdot\left(\lambda_1 + \frac{1}{PK\lambda_1 + 1}\right)\cdot\left[1 - \frac{P\lambda_1^2}{PK\lambda_1 + 1}\right]^K - \frac{1}{PK + 1}. \quad (92)$$

Rewriting (92), we have

$$\frac{PK\lambda_1^2+\lambda_1^2+\lambda_1}{PK\lambda_1^2+\lambda_1+1} = \lambda_1(PK\lambda_1+1) \cdot \exp\left[K\log\left(\frac{P\lambda_1(K-\lambda_1)+1}{PK\lambda_1+1}\right)\right]. \tag{93}$$

Note the similarity of this equation to (70). The left-hand side of (93) is at least 1 and less than $1+1/(PK)$ because $1 \le \lambda_1 \le K$. Thus, we have

$$1 \le \lambda_1(PK\lambda_1+1)\exp\left(\frac{-PK\lambda_1}{PK+1}\right)$$
$$1+\frac{1}{PK} > \lambda_1(PK\lambda_1+1)\exp\left(\frac{-K\lambda_1}{K-\lambda_1}\right)$$

where we have used $1-1/x \le \log(x) \le x-1$. Rewriting these bounds, we obtain

$$\lambda_1 \le \left(\frac{PK+1}{PK}\right)\log[\lambda_1(PK\lambda_1+1)] \tag{94}$$

$$\lambda_1 > \frac{K\log\left[\frac{PK\lambda_1(PK\lambda_1+1)}{PK+1}\right]}{K+\log\left[\frac{PK\lambda_1(PK\lambda_1+1)}{PK+1}\right]}. \tag{95}$$

Since $1 \le \lambda_1 \le K$, we can further bound

$$\frac{K\log(PK)}{K+\log(PK)} < \lambda_1 \le \left(\frac{PK+1}{PK}\right)\log\left[K(PK^2+1)\right] \tag{96}$$

which is almost the same as (71). Thus, for large K we have $\log\lambda_1 \approx \log\log K$. This result was good enough for the MAC (see (71) and (72)), but we here require bounds on λ_1 directly. To this end, we use (96) to further bound the right-hand sides of (94) and (95). The result is that for large K, we have $\lambda_1 \approx \log K + 2\log\log K$. A similar analysis for the MAC shows that $\lambda_1 \approx \log K + \log\log K$. Thus, the fixed point λ_1 increases somewhat faster for the interference channel than for the MAC. The fixed-point sum rate is

$$\sum_{k=1}^{K} R_k = K\log\left(V_{Y_1}/V_{Y_1|X_1}\right)/2 = -\frac{1}{2}\log\left[\left(1-\frac{P\lambda_1^2}{PK\lambda_1+1}\right)^K\right].$$

For large K this sum rate is about $\lambda_1/2$, which is approximately $(\log K)/2 + \log\log K$.

ACKNOWLEDGMENT

The author wishes to thank J. Mazo and the two anonymous referees for their valuable comments.

REFERENCES

[1] A. D. Wyner, "On the Schalkwijk–Kailath coding scheme with a peak energy constraint," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 129–134, Jan. 1968.

[2] L. A. Shepp, J. K. Wolf, A. D. Wyner, and J. Ziv, "Binary communication over the Gaussian channel using feedback with a peak energy constraint," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 476–478, July 1969.

[3] T. T. Kadota, A. D. Wyner, and J. Ziv, "Bound on the average transmission time for sequential estimation systems containing an additive Gaussian noise channel," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 514–515, Sept. 1971.

[4] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive white noise channels with feedback—Part I: No bandwidth constraint," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 172–182, Apr. 1966.

[5] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inform. Theory*, vol. IT-2, pp. 221–238, Sept. 1956. Reprinted in *Claude Elwood Shannon: Collected Papers*, N. J. A. Sloane and A. D. Wyner, Eds. Piscataway, NJ: IEEE Press, 1993.

[6] R. L. Dobrushin, "Information transmission in a channel with feedback," *Theory Probab. Appl.*, vol. 34, pp. 367–383, Dec. 1958.

[7] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability*, J. Neyman, Ed. Berkeley, CA: Univ. Calif. Press, 1961, vol. 1, pp. 611–644. Reprinted in *Claude Elwood Shannon: Collected Papers*, N. J. A. Sloane and A. D. Wyner, Eds. Piscataway, NJ: IEEE Press, 1993.

[8] N. T. Gaarder and J. K. Wolf, "The capacity region of a multiple-access discrete memoryless channel can increase with feedback," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 100–102, Jan. 1975.

[9] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 711–716, Sept. 1987.

[10] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 572–584, Sept. 1979.

[11] L. H. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 623–629, July 1984.

[12] F. J. Willems, "On multiple access channels with feedback," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 842–845, Nov. 1984.

[13] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probab.*, vol. 2, pp. 805–814, Oct. 1974.

[14] M. L. Ulrey, "The capacity region of a channel with s senders and r receivers," *Inform. Contr.*, vol. 29, pp. 185–203, 1975.

[15] A. B. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 60–70, Jan. 1978.

[16] S. Butman, "A general formulation of linear feedback communication systems with solutions," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 392–400, May 1969.

[17] P. Elias, "Channel capacity without coding," in *Lectures on Communication System Theory*. New York: McGraw-Hill, 1961, pp. 363–368.

[18] P. E. Green, Jr., "Feedback communication systems," in *Lectures on Communication System Theory*. New York: McGraw-Hill, 1961, pp. 345–363.

[19] M. Horstein, "Sequential transmission using noiseless feedback," *IEEE Trans. Inform. Theory*, vol. IT-9, pp. 136–143, July 1963.

[20] L. H. Ozarow and S. K. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 667–671, July 1984.

[21] M. S. Iacobucci and M. G. Di Benedetto, "A feedback code for the multiple access channel (MAC): A case study," in *Proc. 1997 IEEE Global Communications Conf., Communications Theory Mini Conf.*, Phoenix, AZ, Nov. 3–8, 1997, pp. 128–132.

[22] ——, "Design and performance of a code for the multiple access channel with feedback," in *Proc. 1998 IEEE Int. Symp. Information Theory*, Cambridge, MA, Aug. 16–21, 1998, p. 125.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[24] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[25] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[26] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[27] S. Pombra and T. M. Cover, "Non white Gaussian multiple access channels with feedback," *IEEE Trans. Inform. Theory*, vol. 40, pp. 885–892, May 1994.

[28] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[29] I. N. Bronshtein and K. A. Semendyayev, *Handbook of Mathematics*, 3rd ed. Berlin, Germany: Springer-Verlag, 1997.

[30] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[31] T. M. Cover, "Comments on broadcast channels," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2524–2530, Oct. 1998.

[32] G. Kramer, "Modulated estimate correction for the white Gaussian broadcast channel with feedback," in *Proc. 34th Annu. Conf. Informatopn Science and Systems*, Princeton, NJ, Mar. 15–17, 2000, pp. TA4-17–TA4-20.
[33] A. B. Carleial, "Outer bounds on the capacity of interference channels," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 602–606, July 1983.
[34] 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.
[35] M. H. M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 607–615, Sept. 1985.
[36] E. C. van der Meulen, *Some Reflections on the Interference Channel*. Boston, MA: Kluwer, 1994, pp. 409–421.
[37] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965.
[38] A. El Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 379–381, May 1978.
[39] ——, "The capacity of the physically degraded Gaussian broadcast channel with feedback," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 508–511, July 1981.
[40] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 197–207, Mar. 1973.
[41] ——, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 279–280, Mar. 1974.