On Cognitive Interference Networks

Amos Lapidoth
ETH Zurich
CH-8092 Zurich, Switzerland
Email: lapidoth@isi.ee.ethz.ch

Shlomo Shamai (Shitz)
Technion, Haifa
Haifa 32000, Israel
Email: sschlomo@ee.technion.ac.il

Michèle A. Wigger
ETH Zurich
CH-8092 Zurich, Switzerland
Email: wigger@isi.ee.ethz.ch

***Abstract*— We study the high-power asymptotic behavior of the sum-rate capacity of multi-user interference networks with an equal number of transmitters and receivers. We assume that each transmitter is cognizant of the message it wishes to convey to its corresponding receiver and also of the messages that a subset of the other transmitters wish to send. The receivers are assumed not to be able to cooperate in any way so that they must base their decision on the signal they receive only. We focus on the network's pre-log, which is defined as the limiting ratio of the sum-rate capacity to half the logarithm of the transmitted power.**

We present both upper and lower bounds on the network's pre-log. The lower bounds are based on a linear partial-cancellation scheme which entails linearly transforming Gaussian codebooks so as to eliminate the interference in a subset of the receivers.

***Inter alias*, the bounds give a complete characterization of the networks and side-information settings that result in a full pre-log, i.e., in a pre-log that is equal to the number of transmitters (and receivers) as well as a complete characterization of networks whose pre-log is equal to the full pre-log minus one. They also fully characterize networks where the full pre-log can only be achieved if each transmitter knows the messages of all users, i.e., when the side-information is "full".**

I. Introduction

In this paper we study communication scenarios that arise in wireless networks when multiple spatially-separated transmitters communicate to multiple spatially-separated receivers.

Consider a situation where K non-cooperating transmitters, labeled $\{1,\ldots,K\}$, wish to communicate with K non-cooperating receivers, labeled $\{1,\ldots,K\}$, where Receiver j wants to learn Message M_j for each $1 \leq j \leq K$. Here $\{M_j\}_{j=1}^K$ are independent with M_j being uniformly distributed over the set $\{1,\ldots,\lfloor e^{nR_j}\rfloor\}$, where n denotes the block-length of transmission and R_j is the rate of transmission to Receiver j.

We assume that each transmitter is cognizant of a subset of the messages $\{M_1,\ldots,M_K\}$ and denote the set of indices of the messages known to Transmitter k by $\mathcal{S}_k$, $k \in \mathcal{K} = \{1,\ldots,K\}$. Also, we assume that the labeling of the transmitters is such that Transmitter k knows Message M_k and hence $\{k\} \subseteq \mathcal{S}_k \subseteq \mathcal{K}$. Transmitter k computes its sequence of inputs at times 1 to n, $\mathbf{X}_k^n \triangleq (X_k(1),\ldots,X_k(n))^{\mathsf{T}}$ as a function of the set of Messages $\{M_j\}_{j\in\mathcal{S}_k}$.

A setting where every transmitter knows all the involved messages—i.e., where $\mathcal{S}_k = \mathcal{K}$ for all $k \in \mathcal{K}$—will be called the *full side-information* setting, and a setting where every Transmitter k is cognizant only of the Message M_k—i.e., where $\mathcal{S}_k = \{k\}$ for all $k \in \mathcal{K}$—will be called the *no side-information* setting. The *full side-information* setting is also called "fully cognitive network", and it corresponds to a broadcast channel with multiple receivers. The *no side-information* setting is also called "non-cognitive network" and is a generalization of the two-user interference channel to more than two transmitters and more than two receivers. A network with neither *full side-information* nor *no side-information* is called a *partial side-information* network. We will refer to any of the above settings as interference networks.

The interference networks are described by a fixed channel matrix $\mathsf{H} \in \mathbb{R}^{K\times K}$, where $\mathbb{R}$ denotes the set of real numbers, as follows. Denote the output signals observed at Receivers 1 through K at the discrete time-t by $Y_1(t)$ through $Y_K(t)$. The output vector at time-t $\mathbf{Y}(t) \triangleq (Y_1(t),\ldots,Y_K(t))^{\mathsf{T}}$ is given by

$$\mathbf{Y}(t) = \mathsf{H}\mathbf{x}(t) + \mathbf{Z}(t), \quad 1 \leq t \leq n, \tag{1}$$

where $\mathbf{x}(t) \triangleq (x_1(t),\ldots,x_K(t))^{\mathsf{T}}$ is the time-t input vector consisting of the inputs at Transmitters 1 through K, and where $\{\mathbf{Z}(t)\}$ is a sequence of independent and identically distributed (IID) Gaussian random vectors of zero-mean and covariance matrix I_K. (Here I_K denotes the identity matrix of dimension K.) Throughout the paper the channel matrix H is assumed to be of full rank.

For each transmitter we impose the same average block power constraint on the sequence of channel inputs, i.e.,

$$\frac{1}{n}\mathsf{E}\left[\sum_{t=1}^{n} X_k^2(t)\right] \leq P, \quad k \in \mathcal{K}. \tag{2}$$

We say that a rate-tuple $(R_1,\ldots,R_K)$ is achievable if there exists a sequence of pairs of encoding schemes satisfying (2) and decoding schemes such that in the limit as n tends to infinity the probability of a decoding error at each receiver tends to 0. Note that each receiver bases its decision on the signal it receives only. Denoting by R_Σ the sum of the rates $R_1,\ldots,R_K$, i.e.,

$$R_\Sigma = \sum_{j=1}^{K} R_j$$

we can define the sum-rate capacity $C_\Sigma(P,\mathsf{H},\{\mathcal{S}_k\})$ as the supremum of the sum-rates over all achievable rate tuples.

In this work we focus on the behavior of the sum-rate capacity $C_\Sigma(P,\mathsf{H},\{\mathcal{S}_k\})$ in the high SNR regime, i.e., in the limit when $P \to \infty$. In particular, the quantity of interest in this regime is the limit of the ratio of the sum-rate capacity to the Gaussian single-user channel capacity when the available power tends to infinity:

$$\eta\left(\mathsf{H},\{\mathcal{S}_k\}\right) \triangleq \varlimsup_{P\to\infty} \frac{C_\Sigma(P,\mathsf{H},\{\mathcal{S}_k\})}{\frac{1}{2}\log\left(1+P\right)}. \quad (3)$$

The limiting ratio $\eta\left(\mathsf{H},\{\mathcal{S}_k\}\right)$ determines the logarithmic growth of the sum-rate capacity at high power, and we will refer to it as the pre-log of the network. Note that the pre-log depends both on the message sets $\{\mathcal{S}_k\}_{k\in\mathcal{K}}$ and on the channel matrix H. The main goal of this work is to examine the influence of the sets $\{\mathcal{S}_k\}_{k\in\mathcal{K}}$ on the pre-log of an interference network with given channel matrix H.

For *full side-information* settings the pre-log is already known to be equal K [1]. However, for *partial side-information* settings and for *no side-information* settings the pre-log is not yet known for general interference networks. But see [2], [3], [5], and [6] for some special networks.

In [2] the two-transmitters/two-receivers interference network with *no side-information* is investigated. The results therein include the result that the pre-log of the setting equals 1 and furthermore even characterize the capacity region of the network to within 1 bit.

The pre-log of the two-transmitters/two-receivers network with *partial side-information* was studied in [5]. There it was shown that for no *partial side-information* setting the pre-log is larger than 1; only *full side-information* yields the "full" pre-log 2.

The more general scenario where both transmitters and both receivers can communicate with multiple antennas is treated in [3].

It should be emphasized that our setting does not include as a special case the X-channel where each transmitter sends independent messages to the *two* receivers [4], [5].

In contrast to the described works in this submission we consider networks with generally more than two transmitters and receivers.

Recently, the authors [6] considered a particular example of an interference network with more than two transmitters and more than two receivers. They showed that in interference networks *partial side-information* settings can exist with a larger pre-log than in the *no side-information* setting. In particular, the authors considered an interference network where the channel matrix is given by the matrix with ones on the diagonal, some constant α on the first lower secondary diagonal, and 0 everywhere else. Thus, in the considered network, Receiver j observes the sum of Transmitter j's input signal, Transmitter $(j-1)$'s input signal scaled by the factor α, and additive white Gaussian noise. For this network it was shown that partial side-information can increase the pre-log significantly and even lead to the "full" pre-log K, the same pre-log as in the *full side-information* setting.

Thus, we see that for the two-transmitters/two-receivers interference network described in [5] and for the interference network described in [6] the impact of *partial side-information* on the pre-log is drastically different. This fact might not seem so surprising to the reader since the two networks have a very different structure. However, it is not clear which properties of a network determine how *partial side-information* influences the pre-log. In fact we will show later in this paper that for the two similar networks with channel matrices

$$\mathsf{H}_1 = \begin{pmatrix} 1 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 0 & 1/2 & 1 \end{pmatrix} \quad (4)$$

and

$$\mathsf{H}_2 = \begin{pmatrix} 1 & 1/2 & 0 \\ 1/2 & 1 & 1/2 \\ 0 & 1/2 & 1 \end{pmatrix} \quad (5)$$

the dependence of the pre-log on the message sets $\{\mathcal{S}_k\}$ is completely different. For networks with channel matrix H_1 in the *no side-information* setting the pre-log equals 1, and there are *partial side-information* settings with pre-log equal 2 and *partial side-information* settings with pre-log 3. In contrast, for networks with channel matrix H_2, in any *partial side-information* setting and in the *no side-information* setting the pre-log equals 2 and only in the *full side-information* setting the pre-log equals 3.

In the next section we will identify which properties of a network determine how *partial side-information* influences the pre-log of a setting.

II. Main Contributions

In this section we state the main results of our work. For proofs we refer to a forthcoming longer version of this paper.

We begin by stating for which interference network settings we can determine the pre-log exactly based on the lower bound and the upper bound derived in the last two subsections. In the second subsection we characterize when *partial side-information* can increase the pre-log, and in the third subsection we give some examples of specific networks to illustrate the results in the previous two subsections. In the subsection before last we describe an encoding scheme—the *linear partial-cancelation* scheme— leading to the lower bound on the pre-log. Finally, in the last subsection we describe how to derive the upper bound on the pre-log.

A. Exact Results

For general interference settings there is a gap between the upper bound and the lower bound obtained with a linear partial-cancelation scheme. Nevertheless, for certain networks the two bounds meet, thus demonstrating the asymptotic optimality of the linear partial-cancelation scheme. Examples of such settings include the setting described in [6] and also—for any given message sets $\{\mathcal{S}_k\}$—the fully connected 2-by-2 interference networks and the networks with channel matrix H_2 given in (5). For *no side-information* settings and for certain *partial side-information* settings the bounds also meet for networks with channel matrix H_1 given in (4). Also, the lower bound and the upper bound also meet for all settings where $p^* = K-1$ and (trivially) where $p^* = K$. Here p^*, which is given ahead in (14), is the best pre-log achieved with a linear partial-cancelation scheme.

From the lower bound and the upper bound we obtain the following results on the pre-log $\eta(\mathsf{H}, \mathcal{S}_k)$ depending on p^*.

Theorem 1: Consider an interference network with channel matrix H and message sets $\{\mathcal{S}_k\}$. Let p^* be defined as in (14). Then:

$$p^* = K \implies \eta(\mathsf{H}, \{\mathcal{S}_k\}) = K, \quad (6)$$

$$p^* = K-1 \implies \eta(\mathsf{H}, \{\mathcal{S}_k\}) = K-1, \quad (7)$$

$$p^* \leq K-2 \implies \eta(\mathsf{H}, \{\mathcal{S}_k\}) < K-1. \quad (8)$$

Since p^* takes on only positive integer values smaller or equal to K, the following corollary can be obtained from Theorem 1.

Corollary 1: For an interference network with channel matrix H and message sets $\{\mathcal{S}_k\}$:

$$\eta(\mathsf{H}, \{\mathcal{S}_k\}) = K \iff p^* = K, \quad (9)$$

and

$$\eta(\mathsf{H}, \{\mathcal{S}_k\}) = K-1 \iff p^* = K-1. \quad (10)$$

Furthermore, the pre-log $\eta(\mathsf{H}, \mathcal{S}_k)$ can never take value in the open interval $(K-1, K)$.

This result is somewhat surprising since for certain interference networks the pre-log can indeed be a non-integer value. An example of an interference network with non-integer pre-log is given in Section II-D.

B. When Partial Side-Information increases the Pre-log

With the results of Theorem 1 in mind we address the following two problems: the problem of identifying the channel matrices H for which *full side-information* is necessary in order to have "full" pre-log K; and the problem of identifying the channel matrices H for which *partial side-information* is beneficial, in the sense that there is a *partial side-information* setting with a pre-log which is larger than the pre-log of the *no side-information* setting. The following theorem answers these questions.

Theorem 2: Consider an interference network with channel matrix H and let $\mathsf{H}_{(j)}^{(k)} \in \mathbb{R}^{(K-1)\times(K-1)}$ denote the matrix obtained when deleting the j-th row and the k-th column from the channel matrix H, and let $h_{j,k}$ denote the element of H in row j and column k. Then

1) The message sets $\{\mathcal{S}\}_{k\in\mathcal{K}}$ have to fulfill the following sufficient and necessary conditions for the pre-log to equal K:

$$(\eta(\mathsf{H}, \{\mathcal{S}_k\}) = K)$$
$$\Longleftrightarrow$$
$$\left(\forall j,k \in \mathcal{K}: \quad \left(\text{rank}\left(\mathsf{H}_{(j)}^{(k)}\right) = K-1 \Longrightarrow j \in \mathcal{S}_k\right)\right).$$

Thus, in particular, *full side-information* is necessary for that the pre-log of a network H is equal K, if and only if, $\text{rank}\left(\mathsf{H}_{(j)}^{(k)}\right) = K-1$ for all $j \neq k$, and $j,k \in \mathcal{K}$.

2) Let $\mathcal{H}$ be the union of the set of all diagonal K by K matrices and of the set of all K by K matrices for which there is an index $k^* \in \mathcal{K}$ such that

$$h_{j,k} = \begin{cases} 0, & \text{if } j \neq k, \quad j \neq k^*, k \neq k^* \\ \text{arbitrary}, & j = k = k^* \\ \neq 0, & \text{else} \end{cases}. \quad (11)$$

Then, for all channel matrices H in $\mathcal{H}$ the pre-log of any *partial side-information* setting equals the pre-log of the *no side-information* setting.

For all channel matrices which are not contained in the set $\mathcal{H}$ there exist *partial side-information* settings with a pre-log which is strictly larger than the pre-log of the *no side-information* setting.

In the remaining of this section we want to have a closer look at Conditions (11). The matrices satisfying these conditions can be illustrated as follows:

$$\mathsf{H} = \begin{pmatrix} \times & 0 & 0 & \dots & 0 & \times & 0 & \dots & 0 & 0 \\ 0 & \times & 0 & \dots & 0 & \times & 0 & \dots & 0 & 0 \\ & & & \dots & & & \dots & & & \\ 0 & 0 & 0 & \dots & \times & \times & 0 & \dots & 0 & 0 \\ \times & \times & \times & \dots & \times & ? & \times & \dots & \times & \times \\ 0 & 0 & 0 & \dots & 0 & \times & \times & \dots & 0 & 0 \\ & & & \dots & & & \dots & & & \\ 0 & 0 & 0 & \dots & 0 & \times & 0 & \dots & \times & 0 \\ 0 & 0 & 0 & \dots & 0 & \times & 0 & \dots & 0 & \times \end{pmatrix} \quad (12)$$

where the index of the row with $K-1$ occurrences of "$\times$" is the same as the index of the column with $K-1$ occurrences of "$\times$". At all positions which are marked by an "$\times$" the matrix H must contain a non-zero element, but these elements do not have to be identical. At the position which is marked by "?" the matrix H can be arbitrary, possibly also 0.

Remark 1: The pre-log of interference networks with channel matrices of the form given in (12) equals $K-1$ in the *no side-information* setting and in any *partial side-information* setting, and the pre-log equals K only in the *full side-information* setting.

C. Examples

1) The fully connected 2-by-2 interference network: The two-transmitters/two-receivers interference network with channel matrix with only non-zero components is of the structure illustrated in (12). Thus with Remark 1 it is possible to reconstruct the results about the interference network in [5], that is, that the pre-log equals 2 only in the *full side-information* setting whereas in the *partial side-information* setting the pre-log equals 1, the same as in the *no side-information* setting.

Remark 2: The fully connected 2-by-2 interference network and trivially the single-user channel are the only fully connected interference networks—i.e., networks with a channel matrix with only non-zero components—for which there is no *partial side-information* setting with pre-log larger than the pre-log of the *no side-information* setting.

2) Networks H_1 *and* H_2*:* Next, let us consider again the channel matrices H_1 and H_2. We see that the channel matrix H_2 is of the form displayed in (12) and therefore, by Remark 1 we can conclude—as announced in Section I—that in any *partial side-information* setting and in the *no side-information* setting the pre-log equals 2 and in the *full side-information* setting the pre-log equals 3.

For the channel matrix H_1 we see that the sub-matrix

$$\mathsf{H}_{1,(3)}^{(1)} = \begin{pmatrix} 1/2 & 1/4 \\ 1 & 1/2 \end{pmatrix}$$

is of rank 1. Therefore, we can conclude that the interference network with channel matrix H_1 and message sets $\mathcal{S}_1 = \{1,2\}$ and $\mathcal{S}_2 = \mathcal{S}_3 = \mathcal{K}$ has pre-log 3. Since all other sub-matrices of the form $\mathsf{H}_{1,(j)}^{(k)}$ for $j \neq k$ and $(j,k) \neq (3,1)$ have rank $K-1$, we can also conclude that in any other *partial side-information* setting the pre-log is at most 2. Furthermore, computing the rates achievable with the linear partial-cancelation scheme one easily finds the message sets $\{\mathcal{S}_k\}$ such that $p^* = 2$ and hence $\eta(\mathsf{H}_1, \{\mathcal{S}_k\}) = 2$. In the *no side-information* setting with channel matrix H_1 the pre-log is given by $\eta\left(\mathsf{H}_1, \{\mathcal{S}_k = \{k\}\}\right) = 1$. This follows from the upper bound in Lemma 3.

3) Wyner's Linear Cellular Interference Model: In [8] Wyner introduced a linear model for cellular wireless communication systems. The network model is a symmetric version of the network considered in [6], this is, a K-by-K interference network where Receiver j observes the sum of Transmitter j's input signal, Transmitter $(j+1)$'s input signal scaled by a factor $\alpha \neq 0$, Transmitter $(j-1)$'s input signal scaled by the same factor α, and additive white Gaussian noise. Thus the channel matrix is given by 1's on the main diagonal, α's on the first upper and lower secondary diagonals and 0 every where else.

In his work Wyner considered the case when all receivers are allowed to cooperate, and hence the setting becomes a multi-access setting. Here, we consider the case where the receivers are not allowed to cooperate, and we also assume that the transmitters have some kind of side-information about the other transmitter's messages. More precisely, let each transmitter beside its own message know the messages of the J previous transmitters and the messages of the J next transmitters for some integer $J \geq 0$.

The pre-log of this setting for given parameters α, J, and K can be shown to be

$$\eta_{\text{Wyner}}(\alpha, J, K) = K - \left\lfloor \frac{K}{J+2} \right\rfloor.$$

Note that the functional dependence of the pre-log for this setting on the parameters α, J, and K is the same as the functional dependence of the pre-log for the asymmetric setting in [6] on these parameters.

D. A Lower Bound

We propose an encoding scheme—the linear partial-cancelation scheme—for an arbitrary interference network with channel matrix H and message sets $\{\mathcal{S}_k\}$ as described in Section I. The encoding scheme is based on random coding arguments.

Prior to transmission, K independent random codebooks $\mathcal{C}_1, \ldots, \mathcal{C}_K$ are generated according to a zero-mean Gaussian distribution of variance P. Here, the codebook $\mathcal{C}_j$ is the set of n-length codewords $\left\{\mathbf{u}_j^n(1), \ldots, \mathbf{u}_j^n\left(\lfloor e^{nR_j}\rfloor\right)\right\}$, and it is used to encode the Message M_j, $j \in \mathcal{K}$. Then, the codebooks are revealed to all transmitters and to all receivers.

For the encoding each transmitter forms a linear combination of the codewords $\mathbf{u}_j^n(M_j)$ where it knows Message M_j and such that the input power constraint (2) is satisfied. Thus, Transmitter k's input sequence is given by

$$\mathbf{X}_k^n = \sum_{j \in \mathcal{S}_k} d_{j,k} \mathbf{u}_j^n(M_j), \quad k \in \mathcal{K},$$

for some real coefficients $d_{j,k}$ satisfying

$$\sum_{j \in \mathcal{S}_k} d_{j,k}^2 \leq 1, \quad k \in \mathcal{K}.$$

For every choice of coefficients $\{d_{j,k}\}_{k \in \mathcal{K}, j \in \mathcal{S}_k}$ we can define the set $\mathcal{R}(\{d_{j,k}\}) \subseteq \mathcal{K}$ of all indices j such that the interference for Receiver j is canceled. More precisely, the set $\mathcal{R}(\{d_{j,k}\})$ is the set of all $j \in \mathcal{K}$ such that the received sequence $\mathbf{Y}_j^n = (Y_j(1), \ldots, Y_j(n))^{\mathsf{T}}$ at Receiver j can be expressed as

$$\mathbf{Y}_j^n = \xi_j \mathbf{u}_j^n(M_j) + \mathbf{Z}_j^n, \quad j \in \mathcal{R}(\{d_{j,k}\}) \tag{13}$$

for $\xi_j \neq 0$. Note that the set $\mathcal{R}(\{d_{j,k}\})$ depends on the channel matrix H, on the message sets $\{\mathcal{S}_k\}_{k \in \mathcal{K}}$, and of course also on the chosen coefficients $\{d_{j,k}\}$.

Let

$$p^*(\mathsf{H}, \{\mathcal{S}_k\}) = \max_{\{d_{j,k}\}} |\mathcal{R}(\{d_{j,k}\})|, \tag{14}$$

where $|\mathcal{A}|$ denotes the cardinality of the set $\mathcal{A}$. If $\{d_{j,k}^*\}$ achieves $p^*(\mathsf{H}, \{\mathcal{S}_k\})$, then by using $\{d_{j,k}^*\}$ the original interference network is transformed into $p^*(\mathsf{H}, \{\mathcal{S}_k\})$ parallel Gaussian single-user channels and a network with K transmitters and $K - p^*(\mathsf{H}, \{\mathcal{S}_k\})$ receivers. Since on the parallel Gaussian single-user channels the rates $\frac{1}{2}\log\left(1 + \xi_j^2 P\right)$ are achievable, $j \in \mathcal{R}(\{d_{j,k}^*\})$, the following lower bound on the sum-rate capacity is obtained

$$C_\Sigma(P, \mathsf{H}, \{\mathcal{S}_k\}) \geq \frac{p^*(\mathsf{H}, \{\mathcal{S}_k\})}{2} \log\left(1 + \min_{j \in \mathcal{R}(\{d_{j,k}^*\})} \xi_j^2 P\right), \tag{15}$$

and hence

$$\eta(\mathsf{H}, \{\mathcal{S}_k\}) \geq p^*(\mathsf{H}, \{\mathcal{S}_k\}). \tag{16}$$

Inspired by [7], we can improve the linear partial-cancelation scheme by extending it over $\mu > 1$ consecutive channel uses. To this end, let the encoder and the decoder group μ consecutive channel uses into a single channel use of a new K-by-K multi-antenna interference network where each transmitter and each receiver consists of μ antennas. Note that any achievable rate tuple for the new network is also achievable, when divided by μ, on the original network. As we next show, we can derive an achievable tuple for the new network by introducing linear processing at the receivers; by converting it to a new μK-by-μK single-antenna interference network; and by then applying the linear partial-cancelation scheme to the resulting network.

We split Message M_j, $j \in \{1, \dots, K\}$, into μ independent Sub-Messages $M_{(j,1)} \dots, M_{(j,\mu)}$ such that there is a one-to-one mapping between M_j and the tuple $(M_{(j,1)}, \dots, M_{(j,\mu)})$.[1]

As in [7] we let Receiver j of the multi-antenna K-by-K interference network linearly process the observed μ antenna outputs by multiplying them with an arbitrarily chosen μ-by-μ matrix A_j. The network is now converted to a single-antenna μK-by-μK network treating the μK receive antennas as separate receivers and by treating each μ-tuple of transmit antennas as corresponding to μ single users that are cognizant of each others messages.

Indexing the transmitters and receivers of the μK-by-μK network by (k, i) and (j, i) respectively where $1 \leq k, j \leq K$ and $1 \leq j \leq \mu$ we can describe the network as follows: The message sets are

$$\mathcal{S}_{(k,i)} = \{(k', i') : k' \in \mathcal{S}_k, 1 \leq i' \leq \mu\}$$

and the channel matrix

$$\mathsf{H}_\mu(\mathsf{H}, \{\mathsf{A}_j\}) = (\mathsf{H} \otimes \mathsf{I}_\nu) \operatorname{diag}(\mathsf{A}_1, \dots, \mathsf{A}_K)$$

where $\otimes$ denotes the Kronecker product and where $\operatorname{diag}(\mathsf{A}_1, \dots, \mathsf{A}_K)$ denotes the block-diagonal matrix with the blocks $\mathsf{A}_1, \dots, \mathsf{A}_K$. If the rate-tuple $(R_{(j,1)}, \dots, R_{(j,\mu)})$ is achievable in the μK-by-μK interference network then the rate

$$R_j = \frac{1}{\mu} \left(R_{(j,1)} + \dots + R_{(j,\mu)} \right) \tag{17}$$

is achievable in the original interference network.

For the described μK-by-μK interference network we can apply the linear partial-cancelation scheme and hence we obtain the achievability of the pre-log: $p^*(\mathsf{H}_\mu(\mathsf{H}, \{\mathsf{A}_j\}), \{\mathcal{S}_{(k,i)}\})$. Combined with (17) this yields a bound on the pre-log of the original network:

$$\eta(\mathsf{H}, \{\mathcal{S}_k\}) \geq \frac{p^*(\mathsf{H}_\mu(\mathsf{H}, \{\mathsf{A}_j\}), \{\mathcal{S}_{(k,i)}\})}{\mu}$$

for any set of processing matrices $\{\mathsf{A}_j\}$. Hence the best lower bound on the pre-log one can obtain by extending the linear partial-cancelation over several channel uses is given by

$$\eta(\mathsf{H}, \{\mathcal{S}_k\}) \geq \sup_{\mu \in \mathbb{Z}^+} \max_{\{\mathsf{A}_j\}_{j \in \mathcal{K}}} \frac{p^*(\mathsf{H}_\mu(\mathsf{H}, \{\mathsf{A}_j\}), \{\mathcal{S}_{(k,i)}\})}{\mu}. \tag{18}$$

That this modification of the linear-partial cancelation scheme indeed leads to an improvement in the achievable rates (and in the lower bound on the pre-log) over the rates (and over the lower bound on the pre-log) achieved in the original linear partial-cancelation scheme can be seen in the following example.

1) Extending the Linear Partial-Cancelation Scheme over several Channel Uses helps: In this section we want to give an example of a network where by extending the linear partial-cancelation scheme over several channel uses leads to a pre-log which is strictly larger than the pre-log achieved with the simple linear partial-cancelation scheme.

Consider the family of channel matrices $\{\mathsf{H}_K\}$ indexed by the number of transmitters and receivers K. For a given $K > 1$ we consider a K-by-K interference network where Receiver j, $j \in \{1, \dots, K\}$ receives a noisy version of the sum of all input signals except for that of Transmitter $(j-1)$ where $j-1$ should be interpreted as K when $j = 1$.

With the result presented in Section II we obtain that the pre-log of the described settings is given by

$$\eta(\mathsf{H}_K, \{\mathcal{S}_k = \{M_k\}\}) = \frac{K}{K-1}, \quad K > 1.$$

To show that this pre-log is indeed achievable the linear partial-cancelation scheme needs to be extended over $K - 1$ channel uses. Extending the scheme to less than $K - 1$ channel uses achieves only a pre-log of 1.

E. An Upper Bound

In this section we provide an upper bound on the sum of the rates (Theorem 3). We do not give a detailed proof of this upper bound but state an auxiliary lemma (Lemma 1) and sketch how this leads to the theorem.

We start by introducing the concept of degradedness for interference networks with K_T transmitters and K_R receivers. Here, we allow the number of transmitters to differ from the number of receivers. Also, in this section we use the concept of multi-antenna interference networks, that is, we assume that Transmitter k consists of t_k transmit antennas and Receiver j consists of r_j receive antennas. We denote Transmitter k's time-t channel input by the vector $\mathbf{X}_k(t) \in \mathbb{R}^{t_k}$ and Receiver j's time-t channel output by the vector $\mathbf{Y}_j(t) \in \mathbb{R}^{r_j}$. The message sets $\{\mathcal{S}_k\}$ are defined as for the K-by-K single-antenna networks. We say that an input distribution is allowed if for any time t the vector $\mathbf{X}_k(t)$, $k \in \mathcal{K}$, depends only on Messages M_j for which $j \in \mathcal{S}_k$.

Definition 1: A K_T-transmitters/K_R-receivers multi-antenna interference network is called *degraded* with respect to the permutation $\boldsymbol{\pi}$ on the set of receivers, $\boldsymbol{\pi} : \{1, \dots, K_R\} \to \{1, \dots, K_R\}$, if any time t

$$\mathbf{Y}_{\boldsymbol{\pi}(1)}(t) \subseteq \mathbf{Y}_{\boldsymbol{\pi}(2)}(t) \subseteq \dots \subseteq \mathbf{Y}_{\boldsymbol{\pi}(K_R-1)}(t) \subseteq \mathbf{Y}_{\boldsymbol{\pi}(K_R)}(t). \tag{19}$$

Note that the definition does not depend on the side-information available at the encoders. It is only a property of the channel.

Lemma 1: Consider a K_T-transmitters/K_R-receivers multi-antenna interference network which is degraded with respect to some permutation $\boldsymbol{\pi} : \{1, \dots, K_R\} \to \{1, \dots, K_R\}$. If for

[1] For example one can think of this splitting as describing the original message M_j by a sequence of bits and then splitting up this sequence into disjoint (not necessarily equally long) bit-sequences and let every sub-message be described by a different sub-sequence.

all time instants t and for any allowed input distribution the channel outputs for $j \in \{2, \ldots, K_R\}$ fulfill

$$\mathbf{Y}_{\boldsymbol{\pi}(j)}(t) = f_j\left(\mathbf{Y}_{\boldsymbol{\pi}(1)}(t), \ldots, \mathbf{Y}_{\boldsymbol{\pi}(j-1)}(t), M_1, \ldots, M_{j-1}\right) \tag{20}$$

for some set of deterministic functions $\{f_j(\cdot)\}$, then the capacity region of the interference network equals the capacity region of a multi-antenna K_T-transmitters/K_R-receivers interference network where at time t all receivers observe only the output $\mathbf{Y}_{\boldsymbol{\pi}(1)}(t)$.

The proof of the lemma is omitted. It relies on the fact that from the channel output sequence $\mathbf{Y}_{\boldsymbol{\pi}(1)}(1), \ldots, \mathbf{Y}_{\boldsymbol{\pi}(1)}(n)$ it is possible to reconstruct the sequences $(\mathbf{Y}_{\boldsymbol{\pi}(2)}(1), \ldots, \mathbf{Y}_{\boldsymbol{\pi}(2)}(n)), \ldots, (\mathbf{Y}_{\boldsymbol{\pi}(K_R)}(1), \ldots, \mathbf{Y}_{\boldsymbol{\pi}(K_R)}(n))$ with probability of error tending to 0 for increasing block-lengths n whenever the rate tuple $(R_1, \ldots, R_{K_R})$ is achievable in the original network.

Lemma 1 is a main tool in the proof of the upper bound in Theorem 3 below. Before stating the theorem we want to give a brief outline of how Lemma 1 is used in the proof.

In a first step we sketch a method on how to obtain an upper bound on the capacity region using Lemma 1 for K-transmitters/K-receivers interference networks fulfilling a certain technical condition, a special case of Condition (21). Then, we outline how this method can be adapted to prove the upper bound in (22) for networks fulfilling Condition (21).

A general interference network can easily be converted into a degraded network by choosing an arbitrary permutation $\boldsymbol{\pi}$ on the set of receivers $\mathcal{K}$ and by letting a genie reveal channel outputs $\mathbf{Y}^n_{\boldsymbol{\pi}(1)}$ through $\mathbf{Y}^n_{\boldsymbol{\pi}(j-1)}$ to Receiver $\boldsymbol{\pi}(j)$, $j \in \mathcal{K}$. Additionally, let a genie reveal linear combinations $\tilde{\mathbf{Z}}^n_1, \ldots, \tilde{\mathbf{Z}}^n_K$ of the Gaussian noise sequences $\mathbf{Z}^n_1, \ldots, \mathbf{Z}^n_K$ to all receivers.

Note that these two steps can only increase the sum-rate capacity. Therefore, any upper bound on the sum-rate capacity of the "genie-aided" network is also an upper bound on the sum-rate capacity of the original network.

Next, we restrict attention to interference networks for which one can choose a permutation $\boldsymbol{\pi}$ and linear combinations $\tilde{\mathbf{Z}}_1, \ldots, \tilde{\mathbf{Z}}_K$ such that for some coefficients $\{\alpha_{j,\ell}\}$ the differences $\mathbf{Y}^n_{\boldsymbol{\pi}(j)} - \sum_{\ell=1}^{j-1} \alpha_{j,\ell} \mathbf{Y}^n_{\boldsymbol{\pi}(\ell)}$, $j = 2, \ldots, K$, are functions of the Messages $M_{\boldsymbol{\pi}(1)}, \ldots, M_{\boldsymbol{\pi}(j-1)}$ and the Gaussian sequences $\tilde{\mathbf{Z}}^n_1, \ldots, \tilde{\mathbf{Z}}^n_K$ only. One can directly verify that the "genie-aided" networks corresponding to the networks under consideration fulfill Condition (20) and thus Lemma 1 can be applied. We conclude that for these networks any upper bound on the resulting network—where all receivers observe $\mathbf{Y}_{\boldsymbol{\pi}(1)}$ and $\tilde{\mathbf{Z}}^n_1, \ldots, \tilde{\mathbf{Z}}^n_K$—is also an upper bound on the sum-rate capacity of the original network. Note that we simplified the problem, since the remaining task of finding the sum-rate capacity (or an upper bound on it) for a network where all receivers observe the same output, is better understood. (In fact it corresponds to a multi-access scenario with partially informed transmitters.)

One can extend this method to a larger class of networks, namely those fulfilling Condition (21). The following modifications are needed: Join Receivers $v_1, \ldots, v_\nu$ into a big common Receiver $v_\mathcal{V}$, thus transforming the K-transmitters/K-receivers network into a K-transmitters/$(K-\nu+1)$-receivers network; Let the genie reveal also Messages M_j, for $j \notin (\{v_1, \ldots, v_\nu\} \cup \{j_1, \ldots, j_q\})$ to Receivers $v_\mathcal{V}, j_1, \ldots, j_q$; Adapt Definition 1 of degradedness to apply for this new setting with informed receivers and also to settings where for a given permutation $\boldsymbol{\pi}$ only a subset of receivers fulfills (19); Choose as the degraded subset Receivers $v_\mathcal{V}, j_1, \ldots, j_q$ with respect to the permutation $\boldsymbol{\pi}$: $\boldsymbol{\pi}(1) = v_\mathcal{V}, \boldsymbol{\pi}(2) = j_1, \ldots, \boldsymbol{\pi}(q+1) = j_q$, and apply Lemma 1 —or actually a slightly modified version of it— to this subset of receivers only.

In the following let $\mathbf{h}_j^\mathsf{T}$ denote the j-th row, $j \in \mathcal{K}$, of the channel matrix H. Also, let $\perp\!\!\!\perp$ denote independence.

Theorem 3: Consider a K-transmitters/K-receivers interference network with channel matrix H and message sets $\{\mathcal{S}_k\}$ where $q+\nu$ distinct rows of the channel matrix $\mathbf{h}^\mathsf{T}_{j_1}, \ldots, \mathbf{h}^\mathsf{T}_{j_q}, \mathbf{h}^\mathsf{T}_{v_1}, \ldots, \mathbf{h}^\mathsf{T}_{v_\nu}$ for any time t fulfill

$$\left(\mathbf{h}^\mathsf{T}_{j_i} - \sum_{\ell=1}^{|\mathcal{V}|} \alpha_{i,\ell} \mathbf{h}^\mathsf{T}_{v_\ell}\right) \mathbf{X}(t) \perp\!\!\!\perp (M_{j_i}, \ldots, M_{j_q}), \quad j = 1, \ldots, q \tag{21}$$

for some coefficients $\{\alpha_{i,\ell}\}_{\substack{i=1,\ldots,q\\ \ell=1,\ldots,\nu}}$ and any allowed input distribution. Then, any rate tuple $(R_1, \ldots, R_K)$ can only be achievable if

$$\sum_{i=1}^{q} R_{j_i} + \sum_{\ell=1}^{|\mathcal{V}|} R_{v_\ell} \leq \frac{|\mathcal{V}|}{2} \log\left(1 + \|\mathsf{H}\|^2 P\right) + \mathsf{c}(\{\alpha_{i,\ell}\}) \tag{22}$$

where $\|\mathsf{H}\|$ denotes the operator norm of the matrix H and $\mathsf{c}(\{\alpha_{i,\ell}\})$ is a constant depending on the coefficients $\{\alpha_{i,\ell}\}$.

1) An Improved Upper Bound: For an improved upper bound that applies our techniques *simultaneously* to subsets of the rate, please see a forthcoming longer version of this paper. Some of the results in this paper rely on this improved upper bound.

REFERENCES

[1] T. Yoo and A. Goldsmith, "On the Optimality of Multi-Antenna Broadcast Scheduling Using Zero Forcing Beam Forming," *IEEE JSAC, Special Issue on 4G Wireless Systems*, Vol. 24, No. 3, pp. 528–541, March 2006.

[2] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian Interference Channel Capacity to Within One Bit", Submitted to *IEEE Trans. of Inform. Theory*, 2007.

[3] S. A. Jafar and M. Fakhereddin, "Degrees of Freedom on the MIMO Interference Channel," To appear in *IEEE Trans. of Inform. Theory*, 2007;

[4] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, " Signaling over MIMO Multi-base Systems: Combination of Multi-access and Broadcast Schemes," ISIT 2006, Seattle, Washington, July 9-14, 2006.

[5] N. Devroye and M. Sharif, "The multiplexing gain of MIMO X-channels with partial transmit side information," ISIT 2007, Nice, June 24-29, 2007.

[6] A. Lapidoth, S. Shamai (Shitz), and M. A. Wigger, "A Linear Interference Network with Local Side-Information," ISIT 2007, Nice, June 24-29, 2007.

[7] H. Weingarten, S. Shamai (Shitz), and Gerhard Kramer, "On the Compound MIMO Broadcast Channel", Proc. 2007 Workshop on Inform. Theory and its App., UCSD Campus, La Jolla, CA, Jan.29-Feb.2, 2007.

[8] A. D. Wyner, "Shannon-Theoretic Approach to a Gaussian Cellular Multiple-Access Channel", *IEEE Trans. of Inform. Theory*, Vol. 40, No. 6, November 1994.