Broadcasting with Degraded Message Sets: A Deterministic Approach

Vinod Prabhakaran †, Suhas Diggavi ‡, and David Tse †
† Wireless Foundations, Dept. of EECS, University of California, Berkeley
‡ Laboratoire des systémes d'information et de communication, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Email: vinodmp@eecs.berkeley.edu, suhas.diggavi@epfl.ch, dtse@eecs.berkeley.edu

***Abstract*— In this paper we study the nested (degraded) message set problem, where user $i \in \{1, \dots, K\}$ requires messages $W_1, \dots, W_i$. We study this for a MIMO linear deterministic broadcast channel model, which is motivated by some recent successes in using such models to obtain insights into approximate characterizations for the Gaussian relay and interference channels. We establish the complete solution for the $K = 3$ user nested message problem for the MIMO linear deterministic broadcast channel. We also establish some extremal points for the general K-user case, where there are only two messages of interest.**

I. Introduction

One of the classical problems in multi-user information theory is that of broadcast, where a set of messages is to be conveyed reliably from a single source to many receivers. This problem was formulated by Cover [4] and the solution to this problem is still open. The complete solution to this problem has been obtained for the case where there is a degradation order between the users' channels [2], [10]. For example, the scalar Gaussian broadcast channel has such an ordering and therefore one can have a complete solution for such a case. More recently the MIMO broadcast channel has been solved for the case where only individual (private) messages are needed by the users [17]. However, a complete answer for this case is unknown if some of the users require common subsets of messages.

Consider a two user broadcast channel, where both user $1, 2$ are interested in message W_1 and only user 2 is interested in message W_2. This means that the messages required by the users are *degraded*, though the channels may not be so. This case has been completely solved in [12]. Given this, one would hope that perhaps by asking for less, *i.e.,* giving structure to the message sets needed by the users, we could make progress on the broadcast problem. However, even the three user problem where all users are interested in W_1, and only user 3 is interested in W_3 is still unresolved. Given this dismal situation, we would like to make progress by perhaps further simplifying the problem. The best known coding theorem for the general broadcast channels is by Marton [14], and the idea behind the coding scheme becomes transparent if we study the *deterministic* (noiseless) version of the problem (see [11] and references therein). Motivated by this, we study a deterministic version of the broadcast channel with degraded message sets.

We study a particular deterministic model which is motivated by the MIMO Gaussian broadcast problem in the high SNR regime. A similar line of research has led to capacity characterizations within a constant number of bits for many problems like the interference channel [9], [3] and the relay channel [1]. However, we would like to point out that in this paper we do not prove any results for the MIMO Gaussian broadcast with degraded message set problem which still remains open in general. However, the hope is that given the ideas for the deterministic case, we can make progress towards at least an approximate characterization for the MIMO Gaussian degraded message set problem. Our approach is to reduce the degraded message set broadcast problem into many virtual multicast problems which can then be solved using compound channel codes or more explicitly by linear network codes. Another related recent paper is [19] which adopts a deterministic network coding based approach for designing universal codes for Gaussian channels.

Another motivation for studying the problem from this viewpoint comes from a broadcast strategy for transmitting over an uncertain point-to-point wireless link. Given the uncertainty in wireless channel, the classical approach is to ensure a desired level of reliability against adverse channels while maximizing the transmission rate. This tradeoff between rate and reliability is fundamental [20]. However, the classical approach, while ensuring reliability against adverse channels (deep fades), does not take advantage of good channel realizations. An alternate strategy was proposed in [5], where a high-reliability code was embedded within a higher rate code. This allowed opportunistic communication where the high rate code opportunistically takes advantage of good channel realizations whereas the embedded high-diversity code ensures that at least part of the information is received reliably.

A natural way to formalize this strategy is using the notion of outage. Given the space of channel realizations, we divide it into outage set $\mathcal{O}$ and a non-outage set $\bar{\mathcal{O}}$. We require a code which can be decoded reliably (with arbitrarily small error probability) for all channels in the set $\bar{\mathcal{O}}$. In terms of outages, diversity embedded codes (or opportunistic codes) can be viewed as sending two messages m_H (higher priority) and m_L (lower priority) such that m_H is recovered reliably for all channels in the non-outage set $\bar{\mathcal{O}}_H$, but in addition m_L is to be recovered if we get a better channel in $\mathcal{G} \subset \bar{\mathcal{O}}_H$. We can connect this idea to broadcast codes with two (degraded) messages by thinking of the set of users as identified by the different channel realizations in $\bar{\mathcal{O}}_H$. We identify m_H

as the common message to be delivered to all the users and m_L as a private message to be delivered to a subset of the users $\mathcal{G} \subset \bar{\mathcal{O}}_H$. Therefore, in this setting we would like to characterize the performance in the high-SNR regime, and therefore even an *approximate* characterization of this problem would be sufficient for this problem. This was formalized in [6], where it was shown that for parallel channels, there was a coarse characterization when we needed to ensure that the high-priority message m_H got performance as if it were on its own (*i.e.,* its reliability-rate was on the diversity-multiplexing trade-off). This result was obtained by solving the underlying degraded message set problem for Gaussian parallel channels [6]. In this case, we do not get successive refinable property of the diversity-multiplexing trade-off as was seen in the case of single transmit antenna (SIMO) and single receive antenna (MISO) under Rayleigh fading [7]. In order to characterize all possible rate-diversity tuples for the diversity embedded codes, we would need to approximately solve the MIMO Gaussian degraded message set problem. This motivates understanding the MIMO case using the specific linear deterministic model and formulation studied in this paper.

In this paper, the main focus is on the 3-user problem where we provide a complete characterization for the rate region for the nested (degraded) message set problem for linear deterministic broadcast channels. The message degradation in this case is that W_1 is interesting to all users, W_2 is for users $2, 3$ and finally W_3 is the private message for user 3. We also have a partial characterization for the K-user problem for specific rate tuples; when there are only two messages, one common and other private for user K; also when one message is common and the other is interesting to all except user 1. However, the general K-user characterization for the degraded message set problem, even for the linear deterministic broadcast model, is open.

The paper is organized as follows. In Section II, we give the specific deterministic model and motivate its use in our problem. We also state the main results obtained in this paper. In Section III, we prove the outer bounds associated with the main results and in Section IV, provide the matching achievability results. We conclude with a brief discussion in Section V.

II. Problem statement and results

A. Model

We consider the following *linear deterministic broadcast channel* which is motivated by some connections to the Gaussian case (see [9], [3], [1]). The input X to the channel lies in an m-dimensional vector space $\mathbb{F}^m$, where $\mathbb{F}$ is a finite field. Let the set of users be represented by $\mathcal{U} = \{1, 2, \ldots, K\}$. Then, for every $k \in \mathcal{U}$, user-k receives the n_k dimensional vector,

$$Y_k = \boldsymbol{H}_k X,$$

where $\boldsymbol{H}_k$ is an $n_k \times m$-matrix in $\mathbb{F}$. There are K messages $W_1, W_2, \ldots, W_K$ of rates $R_1, R_2, \ldots, R_K$ respectively. The k-th user is interested in receiving all the messages up to and including message-k.

We will express all rates in terms of $\log_2 |\mathbb{F}|$ bits/sample. Let us denote the point-to-point capacity of user-k by $r_k = \text{rank}(\boldsymbol{H}_k)$. Also, $r_{1,2} = \text{rank}\left(\begin{bmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \end{bmatrix}\right)$, the point-to-point capacity when users 1 and 2 co-operate, and in general when the set $\mathcal{S}$ users co-operate, the rate is denoted by $r_{\mathcal{S}}$. Also, we will denote the dimension of the null-space of $\boldsymbol{H}_k$ by $\mathcal{N}_k$, null-space of $\begin{bmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \end{bmatrix}$ by $\mathcal{N}_{1,2}$ and in general the dimension of the null space of the matrix containing $\{\boldsymbol{H}_k\}_{k \in \mathcal{S}}$ is denoted by $\mathcal{N}_{\mathcal{S}}$.

Before we present our results for $K = 3$ and more, let us consider the case of $K = 2$. The capacity region has a particularly simple form and can be achieved using a simple blocklength-1 strategy as we show below.

Theorem II.1. *For $K = 2$ users, the capacity region of the linear deterministic broadcast channel, besides the non-negativity constraints on the rates is given by*

$$\begin{aligned} R_1 &\leq \min\{r_1, r_2\}, & (1) \\ R_1 + R_2 &\leq r_2. & (2) \end{aligned}$$

The converse readily follows from point-to-point capacity considerations. The achievability scheme is also simple. We will see how the corner point $(R_1 = \min\{r_1, r_2\}, R_2 = r_2 - \min\{r_1, r_2\})$ can be achieved using a blocklength-1 strategy. Supposed $w_{1,1}, \ldots, w_{1,R_1}$ are arbitrary elements in $\mathbb{F}$ representing the common message, and $w_{2,1}, \ldots, w_{2,R_2}$ represent the private message. To see the explicit coding strategy, we first we note that there is a basis $\mathcal{B} = \mathcal{B}_{1,2} \cup \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_0$ such that $\mathcal{B}_{1,2}$ is a basis for $\mathcal{N}_1 \cap \mathcal{N}_2$, and $\mathcal{B}_{1,2} \cup \mathcal{B}_k$ is a basis for $\mathcal{N}_k$, for $k = 1, 2$. The input vector X is chosen as a linear combination of these basis vectors with the co-efficients being the messages W's as follows. We first pick the set $\mathcal{B}_0$ which is not in either of the null-spaces and weight these vectors with distinct common messages. If $\boldsymbol{B}_0$ is a matrix whose columns are the elements of $\mathcal{B}_0$ and $W_1^{(1)}$ is the column vector of the first $w_{1,k}$ messages chosen such that the length of $W_1^{(1)}$, then we get $\boldsymbol{B}_0 W_1^{(1)}$. Since the rest of the basis vectors are in at least one of the null spaces, the rest of the common message has to be repeated. For user 1, we form $\boldsymbol{B}_2$ by choosing column vectors from $\mathcal{B}_2$ and multiply with the rest of the common messages $W_1^{(2)}$ to get $\boldsymbol{B}_2 W_1^{(2)}$. Note that the size of $\mathcal{B}_0 \cup \mathcal{B}_2$ is r_1 and this ensures that we will not run out of basis vectors to accommodate all of $W_1^{(2)}$. Similarly, we form a $\boldsymbol{B}_1^{(1)}$ from $\mathcal{B}_1$ for user 2 to get $W_1^{(2)}$. Since $|\mathcal{B}_0 \cup \mathcal{B}_1| = r_2$, we should have R_2 vectors in $\mathcal{B}_1$ left over after forming $\boldsymbol{B}_1^{(1)}$. We form a matrix $\boldsymbol{B}_1^{(2)}$ from them and use it to send the private message vector W_2. In summary, we have

$$X = \begin{bmatrix} \boldsymbol{B}_0 & \boldsymbol{B}_2 & \boldsymbol{B}_1^{(1)} & \boldsymbol{B}_1^{(2)} \end{bmatrix} \begin{bmatrix} W_1^{(1)} \\ W_1^{(2)} \\ W_1^{(2)} \\ W_2 \end{bmatrix}.$$

From the choices we made, it is clear that both users can recover the common message and user 2 can recover, in addition, the private message by matrix inversion.

B. Results

For the 3-user problem, besides inequalities of the form (1) and (2), the characterization of the capacity region involves an additional inequality. The result is

Theorem II.2. *For $K = 3$, the capacity region of the linear deterministic broadcast channel, besides the non-negativity constraints on the rates is given by*

$$R_1 \leq \min\{r_1, r_2, r_3\}, \tag{3}$$

$$R_1 + R_2 \leq \min\{r_2, r_3\}, \tag{4}$$

$$R_1 + R_2 + R_3 \leq r_3, \text{ and} \tag{5}$$

$$2R_1 + R_2 + R_3 \leq r_1 + r_2 + r_{1,2,3} - r_{1,2}. \tag{6}$$

For the 3-user problem, a blocklength-1 scheme like the one for the 2-user problem is not available. We present our achievable scheme in Section IV. Our scheme reduces the problem to many virtual multicast problems. These multicast problems can then be solved using compound channel codes, or somewhat more explicitly using ideas from network coding. We prove the converse in the next section.

For $K > 3$ users, we can characterize two cases: (i) when there are only two messages, one common and the other private for user K; and (ii), when one message is common and the other is interesting to all except user 1.

Theorem II.3. *Let $L > 3$.*
(i) For $R_2 = R_3 = \ldots = R_{L-1} = 0$, the capacity region is

$$R_1 \leq \min\{r_1, r_2, \ldots, r_K\}, \text{ and} \tag{7}$$

$$\begin{aligned} & mR_1 + R_K \\ & \leq \min\{r_{a_1} + r_{a_2} + \ldots + r_{a_m} + r_{a_1,a_2,\ldots,a_m,K} - r_{a_1,a_2,\ldots,a_m} : \\ & \qquad 1 \leq a_1 < a_2 < \ldots < a_m \leq K-1, \\ & \qquad m = 1, 2, \ldots, K-1\}. \end{aligned} \tag{8}$$

(ii) For $R_3 = \ldots = R_K = 0$, the capacity region is

$$R_1 \leq \min\{r_1, r_2, \ldots, r_K\}, \text{ and} \tag{9}$$

$$R_1 + R_2 \leq \min\{r_2, \ldots, r_K\}. \tag{10}$$

III. Outerbounds

We first prove the converses for all the theorems. The conditions (3) - (5), (7), (9), and (10) follow trivially from the point-to-point capacity considerations. The idea behind showing (6) is to give the outputs of users 1 and 2 (i.e., Y_1 and Y_2) to user 3. The intuition is that this inequality captures the constraint on the rates of messages W_2 and W_3 imposed by the fact that the common message W_1 needs to be served to both users 1 and 2 assuming that user 3 is not the bottleneck. When user 3 is the bottleneck, the other inequalities are enough as we will see. Using Fano's inequality for user 1 and message W_1 (and dropping $o(n)$ terms), we can write

$$\begin{aligned} nR_1 &\leq I(W_1; Y_1^n) \\ &= H(Y_1^n) - H(Y_1^n|W_1) \\ &\leq nr_1 - H(Y_1^n|W_1) \end{aligned}$$

We can re-write this as

$$H(Y_1^n|W_1) \leq n(r_1 - R_1). \tag{11}$$

$$\text{Similarly,} \qquad H(Y_2^n|W_1) \leq n(r_2 - R_2). \tag{12}$$

Applying Fano's inequality again, but now for user 3 and messages W_2 and W_3 (and dropping $o(n)$ terms)

$$\begin{aligned} n(R_2 + R_3) &\leq I(W_2, W_3; Y_3^n) \\ &\leq I(W_2, W_3; Y_3^n|W_1) \\ &\leq I(W_2, W_3; Y_1^n, Y_2^n, Y_3^n|W_1) \\ &= H(Y_1^n, Y_2^n, Y_3^n|W_1) \\ &\leq H(Y_1^n|W_1) + H(Y_2^n|W_1) + H(Y_3^n|Y_1^n, Y_2^n) \\ &\leq n(r_1 - R_1) + n(r_2 - R_1) + H(Y_3^n|Y_1^n, Y_2^n), \end{aligned}$$

where we used the determinism of the channel in step 4 and the two numbered inequalities above in the last step. To upper-bound $H(Y_3^n|Y_1^n, Y_2^n)$, we will bound for any (consistent) realization for the outputs of users 1 and 2 at time k, the number of possibilities the output $Y_{3,k}$ of user 3 at time k can take. Let $\boldsymbol{U}_1$ be a matrix whose column vectors form a basis for $\mathcal{N}_{1,2,3}$. Then we can find a matrix $\boldsymbol{U}_2$ such that the column vectors of $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ form a basis for $\mathcal{N}_{1,2}$. Also we can find a matrix $\boldsymbol{U}_3$ such that the column vectors of all these three $\boldsymbol{U}$-matrices is a basis $\mathcal{B}$ for the input space $\mathbb{F}^m$. Given $Y_{1,k}$ and $Y_{2,k}$, the coefficients of the basis vectors in $\boldsymbol{U}_3$ for the expansion of X_k in the above basis $\mathcal{B}$ is specified. Also, irrespective of what the coefficients for the basis vectors in $\boldsymbol{U}_1$ are, $Y_{3,k}$ remains the same. The possible values of $Y_{3,k}$ are determined by the choice of coefficients for vectors in $\boldsymbol{U}_2$. Since there are

$$\text{rank}(\mathcal{N}_{1,2}) - \text{rank}(\mathcal{N}_{1,2,3}) = r_{1,2,3} - r_{1,2}$$

such vectors, we have $H(Y_3^n|Y_1^n, Y_2^n) \leq n(r_{1,2,3} - r_{1,2})$. Thus

$$2R_1 + R_2 + R_3 \leq r_1 + r_2 + r_{1,2,3} - r_{1,2}.$$

To prove (8), we give the outputs of users $a_1, a_2, \ldots, a_m$ to user L and follow the same argument as above.

IV. Achievability

To show the achievability in Theorem II.2, it is enough to show how the six marked corner points in Figure II-B can be achieved since the rest of the boundary of the region can be obtained by time-sharing between these corner points. To show the achievability of a corner point (R_1, R_2, R_3), our strategy

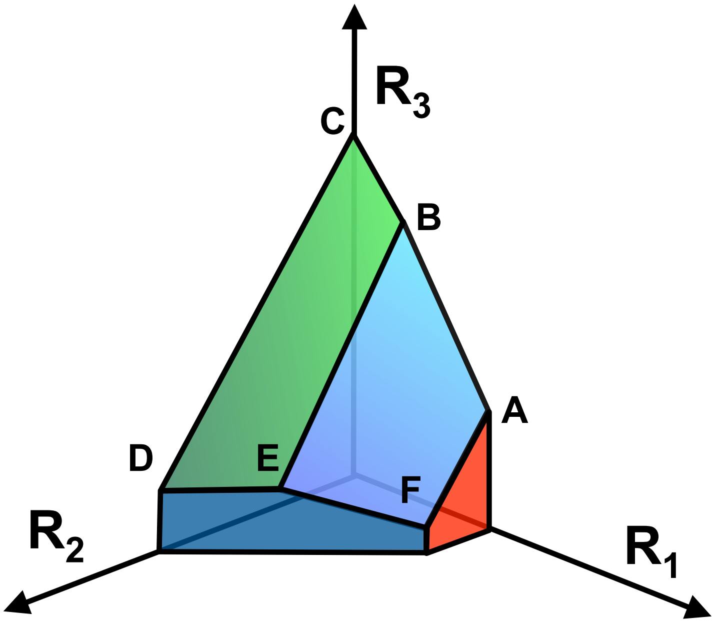


Fig. 1. The degraded message set capacity region of the 3-user linear deterministic broadcast channel.

is to reduce the problem into three virtual parallel multicast problems (where a single message needs to be conveyed to many receivers), one for each of the messages. In particular, the input X will be of the form

$$X = X_1 + X_2 + X_3,$$

where $X_k \in \mathbb{G}_k,\ k = 1,2,3$, such that $\mathbb{G}_k$'s are subspaces of $\mathbb{F}^m$. We treat X_k as the input to the virtual channel for the multicast problem carrying message W_k. The outputs of the virtual channel carrying message W_1 are defined as follows by means of matrices $\boldsymbol{P}_{1,j}$'s

$$\begin{aligned} Y_{1,1} &= \boldsymbol{P}_{1,1} Y_1 = \boldsymbol{P}_{1,1}\boldsymbol{H}_1(X_1 + X_2 + X_3),\\ Y_{1,2} &= \boldsymbol{P}_{1,2} Y_2 = \boldsymbol{P}_{1,2}\boldsymbol{H}_2(X_1 + X_2 + X_3),\\ Y_{1,3} &= \boldsymbol{P}_{1,3} Y_3 = \boldsymbol{P}_{1,3}\boldsymbol{H}_3(X_1 + X_2 + X_3). \end{aligned}$$

The matrices $\boldsymbol{P}'_{1,j}s$ will be chosen such that the following two conditions hold for each $j = 1,2,3$:

(1a) $\mathbb{G}_2 \oplus \mathbb{G}_3$ is in the nullspace of $\boldsymbol{P}_{1,j}\boldsymbol{H}_j$. This will ensure that $Y_{1,j} = \boldsymbol{P}_{1,j}\boldsymbol{H}_j X_1$.

(1b) The dimension of $(\boldsymbol{P}_{1,j}\boldsymbol{H}_j)$-image of $\mathbb{G}_1$ is at least R_1.

The first condition eliminates the interference from messages W_2 and W_3 on the virtual channel outputs $Y_{1,j}$'s for message W_1. The second condition ensures that the resulting virtual channel can support a multicast rate R_1 for the message W_1. It is easy to see from results on compound channel [13] that a random code built with a uniform input distribution over $\mathbb{G}_1$ will achieve this. In fact, a linear code can be used to achieve this by using ideas along the lines of results in network coding [18].

Once all the users have decoded message W_1, they can eliminate the effect of this message on the outputs. We define the outputs of the virtual channel for message W_2 as follows

$$\begin{aligned} Y_{2,2} &= \boldsymbol{P}_{2,2}(Y_2 - \boldsymbol{H}_2 X_1) = \boldsymbol{P}_{2,2}\boldsymbol{H}_2(X_2 + X_3),\\ Y_{2,3} &= \boldsymbol{P}_{2,3}(Y_3 - \boldsymbol{H}_3 X_1) = \boldsymbol{P}_{2,3}\boldsymbol{H}_3(X_2 + X_3). \end{aligned}$$

Again, we choose the matrices to satisfy the following conditions for $j = 2,3$

(2a) $\mathbb{G}_3$ is in the nullspace of $\boldsymbol{P}_{2,j}\boldsymbol{H}_j$. This will ensure that $Y_{2,j} = \boldsymbol{P}_{2,j}\boldsymbol{H}_j X_2$.

(2b) The dimension of $(\boldsymbol{P}_{2,j}\boldsymbol{H}_j)$-image of $\mathbb{G}_2$ is at least R_2.

These conditions are enough to make sure that we have a virtual channel for the multicast problem with message W_2 and users 2 and 3 as receivers, and a code for this channel exists which can support rate R_2.

Finally, user 3 can eliminate the effects of W_1 and W_2 after successfully decoding these messages. This creates a virtual

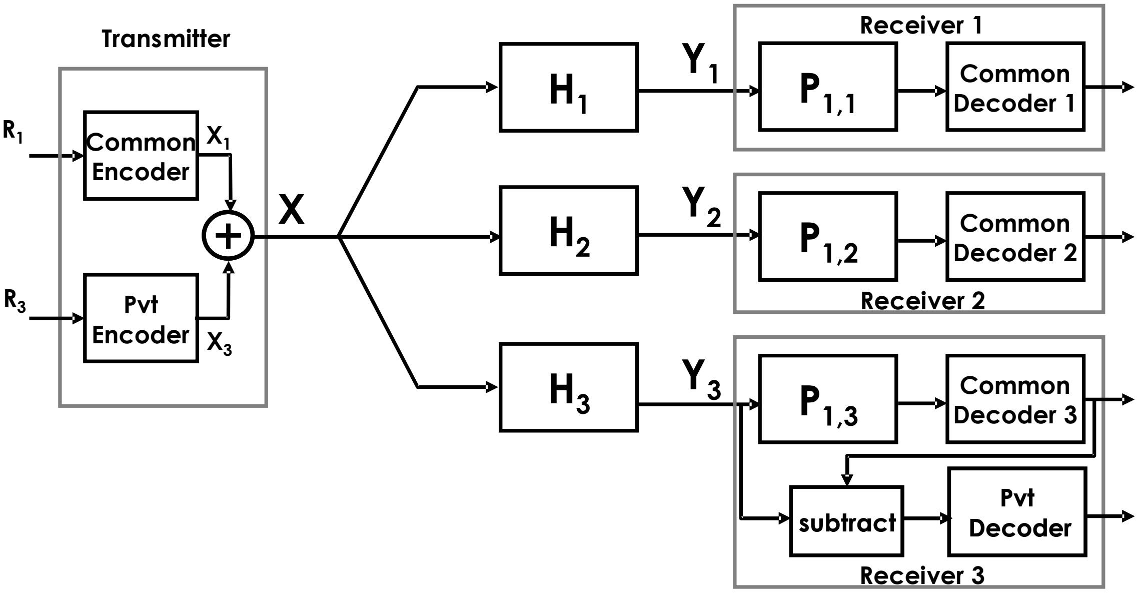


Fig. 2. Block diagram showing the reduction to two multicast problems when message 2 is absent ($R_2 = 0$).

channel for delivering message W_3.

$$Y_{3,3} = (Y_3 - \boldsymbol{H}_3(X_1 + X_2)) = \boldsymbol{H}_3 X_3.$$

We should make sure the following condition is satisfied in order to deliver message W_3 at rate R_3 to user 3:

(3b) The dimension of $\boldsymbol{H}_3$-image of $\mathbb{G}_3$ is at least R_3.

Thus to demonstrate the achievability, it is enough to show that for each of the corner points in Figure II-B, we can choose $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3, \boldsymbol{P}_{1,1}, \boldsymbol{P}_{1,2}, \boldsymbol{P}_{1,3}, \boldsymbol{P}_{2,2}$, and $\boldsymbol{P}_{2,3}$ satisfying the above conditions.

Point A: If $r_3 \leq \min(r_1, r_2)$, $R_1 = r_3, R_2 = R_3 = 0$. Then, the choices are clear $\mathbb{G}_1 = \mathbb{F}^m, \mathbb{G}_2 = \mathbb{G}_3 = \{\phi\}$, where ϕ is the zero-vector in $\mathbb{F}^m$, and $\boldsymbol{P}_{1,1} = \boldsymbol{I}, \boldsymbol{P}_{1,2} = \boldsymbol{I}, \boldsymbol{P}_{1,3} = \boldsymbol{I}, \boldsymbol{P}_{2,2} = []$, and $\boldsymbol{P}_{2,3} = []$.

If $r_3 \geq \min(r_1, r_2)$, without loss of generality for point A, we can assume $r_1 \leq r_2$. Then $R_1 = r_1, R_2 = 0$, and

$$R_3 = \min(r_3 - r_1, r_2 - r_1 + r_{1,2,3} - r_{1,2}).$$

We set $\mathbb{G}_2 = \{\phi\}$, $\boldsymbol{P}_{2,j} = []$ and satisfy conditions (2a) and (2b) trivially. Since $R_1 = r_1$, to satisfy condition (1b), we must choose $\mathbb{G}_1 = \mathbb{F}^m$ and $\boldsymbol{P}_{1,1} = I$. In addition, to satisfy condition (1a) we should choose $\mathbb{G}_3$ to be a subspace of the nullspace $\mathcal{N}_1$ of user 1's channel matrix $\boldsymbol{H}_1$. We will do this next for which we need the following simple fact from linear algebra.

Fact 1: There are sets U, V, W, Z of linearly independent vectors in $\mathbb{F}^m$ such that $U \subseteq \mathcal{N}_{1,2,3},\ V \subseteq \mathcal{N}_{1,2} \backslash \mathcal{N}_{1,2,3},\ W \subseteq \mathcal{N}_{1,3} \backslash \mathcal{N}_{1,2,3},\ Z \subseteq \mathcal{N}_1 \backslash (\mathcal{N}_2 \cup \mathcal{N}_3)$, and the following are true

- U is a basis for $\mathcal{N}_{1,2,3}$,
- $U \cup V$ for $\mathcal{N}_{1,2}$,
- $U \cup W$ for $\mathcal{N}_{1,3}$, and
- $U \cup V \cup W \cup Z$ for $\mathcal{N}_1$.

Then the cardinality of these sets are $|U| = m - r_{1,2,3},\ |V| = r_{1,2,3} - r_{1,2},\ |W| = r_{1,2,3} - r_{1,3},\ |Z| = -r_1 + r_{1,2} + r_{1,3} - r_{1,2,3}$.

We construct $\mathbb{G}_3$ by picking basis vectors for it starting from V and if required we will continue picking from Z if V is exhausted until R_3 vectors have been picked. Since $|U \cup Z| = r_{1,3} - r_1 \geq r_3 - r_1 \geq R_3$, we are guaranteed to find enough basis vectors to construct $\mathbb{G}_3$. Since $\mathbb{G}_3 \cap \mathcal{N}_3 = \{\phi\}$, condition (3b) is satisfied. To choose $\boldsymbol{P}_{1,2}$ and $\boldsymbol{P}_{1,3}$ so that condition (1a) is satisfied, we will use the following fact.

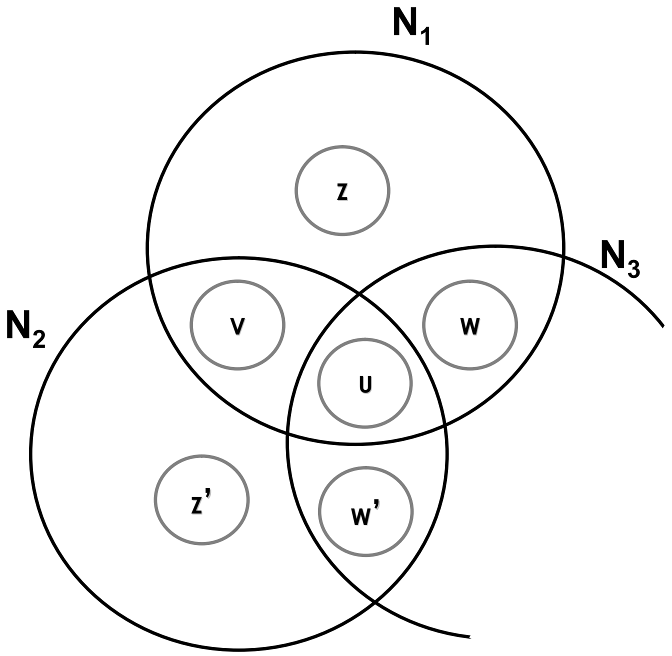


Fig. 3. Diagram illustrating Fact 1 and Fact 3.

Fact 2: Let N be the nullspace of $\boldsymbol{H} \in \mathbb{F}^{n \times m}$. If $\tilde{\mathcal{N}} \supseteq N$ be a subspace of $\mathbb{F}^m$, then there is a matrix $\boldsymbol{P} \in \mathbb{F}^{n \times n}$ such that $\tilde{\mathcal{N}}$ is the nullspace of $\boldsymbol{PH}$.

By defining $\tilde{\mathcal{N}}_2 = \mathcal{N}_2 \oplus \mathbb{G}_3 \supseteq \mathbb{G}_3$, the above fact allows us to find a $\boldsymbol{P}_{1,2}$ such that $\mathbb{G}_2 \oplus \mathbb{G}_3 = \mathbb{G}_3$ is in the nullspace of $\boldsymbol{P}_{1,2}\boldsymbol{H}_2$ as required in condition (1a). Also, the dimension of $(\boldsymbol{P}_{1,2}\boldsymbol{H}_2$ image of $\mathbb{G}_1 = \mathbb{F}^m$ is just

$$
\begin{aligned}
m - \dim(\tilde{\mathcal{N}}_2) &= m - (\dim(\mathcal{N}_2) + [R_3 - |V|]_+) \\
&= r_2 - [R_3 - r_{1,2,3} + r_{1,2}]_+ \geq r_1 = R_1,
\end{aligned}
$$

where $[x]_+ = \max(0, x)$. This ensures that our choice of $\boldsymbol{P}_{1,2}$ also satisfies condition (1b). Similarly, defining $\tilde{\mathcal{N}}_3 = \mathcal{N}_3 \oplus \mathbb{G}_3$ and noting that

$$
m - \dim(\tilde{\mathcal{N}}_3) = m - (\dim(\mathcal{N}_3) + R_3) = r_3 - R_3 \geq r_1 = R_1,
$$

we can find $\boldsymbol{P}_{1,3}$ to satisfy conditions (1a) and (1b).

Point *B*: This point is distinct from point A only when

$$
r_3 \geq \max(r_1, r_2) + r_{1,2,3} - r_{1,2}.
$$

Again, without loss of generality, we will assume that in addition $r_1 \leq r_2$. Then point B is

$$
\begin{aligned}
R_1 &= r_1 + r_2 - r_3 + r_{1,2,3} - r_{1,2} \\
R_2 &= 0 \\
R_3 &= 2r_3 - r_1 - r_2 - r_{1,2,3} + r_{1,2}
\end{aligned}
$$

To see how to achieve point B we need the following fact.

Fact 3: There are sets $W' \subseteq \mathcal{N}_{2,3} \backslash \mathcal{N}_{1,2,3}$ and $Z' \subseteq \mathcal{N}_2 \backslash (\mathcal{N}_1 \cup \mathcal{N}_3)$ of linearly independent vectors such that for the linearly independent sets in Fact 1, along with the properties stated there, the following are true

- $U \cup W'$ is a basis for $\mathcal{N}_{2,3}$,
- $U \cup V \cup W' \cup Z'$ for $\mathcal{N}_2$, and
- $U \cup V \cup W \cup W' \cup Z \cup Z'$ for $\mathcal{N}_1 \oplus \mathcal{N}_2$.

Also, $|W'| = \mathcal{N}_{1,2} - \mathcal{N}_{1,2,3}$ and $|Z'| = -r_2 + r_{1,2} + r_{2,3} - r_{1,2,3}$. We will now identify a subspace $\tilde{\mathcal{N}}_1 \subseteq \mathcal{N}_1$ and use Fact 2 to find a $\boldsymbol{P}_{1,1}$ such that $\tilde{\mathcal{N}}_1$ is the nullspace of $\boldsymbol{P}_{1,1}\boldsymbol{H}_1$. We can treat this as the channel matrix of a user, say user 1' whose output can be derived from that of user 1. We will show that the point A for the problem with user 1 replaced by user 1' coincides with the point B we are trying to achieve here and thus complete the proof of achievability of point B. We define $\tilde{\mathcal{N}}_1 = \mathcal{N}_1 \oplus \tilde{\mathbb{G}}_1$ where $\tilde{\mathbb{G}}_1$ is the subspace spanned by picking some $r_1 - R_1$ vectors from Z'. It is possible to pick these vectors since $|Z'| - (r_1 - R_1) = r_{2,3} - r_3 \geq 0$. Point A with user 1 replaced by user 1' can be easily checked to be

$$
(r_{1'}, 0, \min(r_3 - r_{1'}, r_2 - r_{1'} + r_{1',2,3} - r_{1',2}) = (R_1, 0, R_3).
$$

Point *C*: Here $R_1 = R_2 = 0$ and $R_3 = r_3$. This is easy to achieve by transmitting only the message W_3 to user 3 at the point-to-point capacity r_3 of this user.

Point *D*: Only users 2 and 3 are being sent to at this point. Thus, it is a two-user problem. We can treat it as point A with W_2 playing the role of W_1 there, and effectively removing user 1 by setting $H_1 = I$. A more concrete way of seeing the achievability follows. If $r_2 \geq r_3$, point D is $R_2 = r_2,\ R_1 = R_3 = 0$. This can be achieved by choosing $\mathbb{G}_2 = \mathbb{F}^m,\ \boldsymbol{P}_{2,j} = I$, and all the other $\boldsymbol{P} = []$ and $\mathbb{G}_1 = \mathbb{G}_3 = \{\phi\}$. When $r_2 < r_3$, point D i $R_1 = 0, R_2 = r_2, R_3 = r_3 - r_2$. Then we use the following fact

Fact 4: There are sets R, S, T of linearly independent vectors such that $R \subseteq \mathcal{N}_{2,3}$, $S \subseteq \mathcal{N}_2 \backslash \mathcal{N}_{2,3}$, and $T \subseteq \mathcal{N}_3 \backslash \mathcal{N}_{2,3}$ satisfying the following:

- R is a basis for $\mathcal{N}_{2,3}$,
- $R \cup S$ is a basis for $\mathcal{N}_2$,
- $R \cup T$ is a basis for $\mathcal{N}_3$, and
- $R \cup S \cup T$ is a basis for $\mathcal{N}_2 \oplus \mathcal{N}_3$.

We will choose $\mathbb{G}_3$ as a subspace spanned by $R_3 = r_3 - r_2$ linearly independent vectors from S which has size $|S| = r_{2,3} - r_2 \geq R_3$. We also set $\mathbb{G}_2 = \mathbb{F}^m$ and $\mathbb{G}_1 = \{\phi\}$. Fact 2 will allow us to choose a $\boldsymbol{P}_{2,3}$. We also set $\boldsymbol{P}_{2,2} = I$, and $\boldsymbol{P}_{1,j} = []$. It is easy to check that all the conditions are satisfied.

Point *F*: We can assume without loss of generality, $r_1 \leq r_2 \leq r_3$, since otherwise this point does not appear as a distinct corner point. [If $r_3 \leq \min(r_1, r_2)$ F coincides with A=$(r_3, 0, 0)$. If $r_3 \leq \min(r_1, r_2)$, but $r_2 \leq r_1$, F again coincides with A which will now be $(r_2, 0, \min(r_3 - r_2, r_1 - r_2 + r_{1,2,3} - r_{1,2}))$]. Then point F is

$$R_1 = r_1,\ R_2 = r_2 - r_1\ R_3 = \min(r_3 - r_2, [r_{1,2,3} - r_{1,2} - r_1]_+).$$

Using Fact 1, we make the following choices.

$$\mathbb{G}_1 = \mathbb{F}^m,\ \boldsymbol{P}_{1,1} = I.$$

$\mathbb{G}_2$ is formed by picking basis vectors, first from Z. If $|Z| \geq r_2 - r_1 = R_2$, then upon picking R_2 basis vectors we stop. However, if $|Z| < r_2 - r_1$, we will continue picking more basis vectors for $\mathbb{G}_2$. Now, we pick $[R_2 - |Z|]_+$ pairs of vectors, each pair containing one vector from W and one from V. It is easy to see that there are enough vectors to pick since

$$\begin{aligned} |Z| + |W| &= r_{1,2} - r_1 \geq r_2 - r_1 = R_2, \text{ and} \\ |Z| + |V| &= r_{1,3} - r_1 \geq r_3 - r_1 \geq R_2. \end{aligned}$$

We choose $\mathbb{G}_3$ as the subspace spanned by picking R_3 vectors from V which were not selected for forming the basis for $\mathbb{G}_2$. The number of such vectors available to be picked is

$$\begin{aligned} |V| - [R_2 - |Z|]_+ &= r_{1,2,3} - r_{1,2} - [r_2 + r_{1,2,3} - r_{1,2} - r_{1,3}]_+ \\ &\geq \min(r_{1,3} - r_2, r_{1,2,3} - r_{1,2}) \\ &\geq \min(r_3 - r_2, [r_{1,2,3} - r_{1,2} - r_1]_+) = R_3. \end{aligned}$$

We can now use Fact 2 to find $\boldsymbol{P}$'s to satisfy the conditions (1a) and (2a). It is also not hard to check that the resulting $\boldsymbol{P}$'s satisfy conditions (1b) and (2b). By design, we satisfy condition (3b).

Point *E*: This point is distinct from point F only when (6) is active. Then point E is

$$\begin{aligned} R_1 &= r_1 + r_2 + r_{1,2,3} - r_{1,2} - r_3, \\ R_2 &= \min(r_2, r_3) - R_1, \\ R_3 &= r_3 - \min(r_2, r_3). \end{aligned}$$

If $r_2 \geq r_3$, this coincides with point D. Hence, we can assume that $r_2 < r_3$, and point E becomes

$$\begin{aligned} R_1 &= r_1 + r_2 + r_{1,2,3} - r_{1,2} - r_3, \\ R_2 &= r_3 - r_1 - r_{1,2,3} - r_{1,2}, \\ R_3 &= r_3 - r_2. \end{aligned}$$

Acheivability of this point can be shown by considering the user 1' we created to prove the achievability of point B. It is not hard to show that the point E under consideration is in fact point F of the problem where user 1 is replaced by user 1'. This concludes the proof of achievability of Theorem II.2. The achievability of Theorem II.3 uses similar ideas and is omitted.

V. Conclusion

In this paper we studied the degraded message set problem for linear deterministic broadcast channel. The rate-region for the two-message problem is illustrated in Figure 4. The interesting region is that shown in Figure 4(a), which shows a piecewise linear characterization. Though we have a characterization for the 3-user problem, the general characterization is elusive. Also, as part of ongoing work, we are connecting these results to an *approximate* characterization of the corresponding Gaussian problem. There seems to be an intimate connection between the degrees-of-freedom (high-SNR) characterization and the deterministic channel. We hope that this connection would shed more insight to long-standing open questions on broadcast channels.

References

[1] S. Avestimehr, S. Diggavi, and D N C. Tse. Wireless network information flow. in *Proc. Allerton Conf. Commun., Contr., Computing*, Monticello, IL, Sep. 2007.

[2] P. Bergmans. A simple converse for broadcast channels with additive white Gaussian noise. *IEEE Transactions on Information Theory*, 20:279–280, March 1974.

[3] G. Bresler, A. Parekh, and D N C. Tse. The Approximate Capacity of an One-Sided Gaussian Interference Channel. in *Proc. Allerton Conf. Commun., Contr., Computing*, Monticello, IL, Sep. 2007.

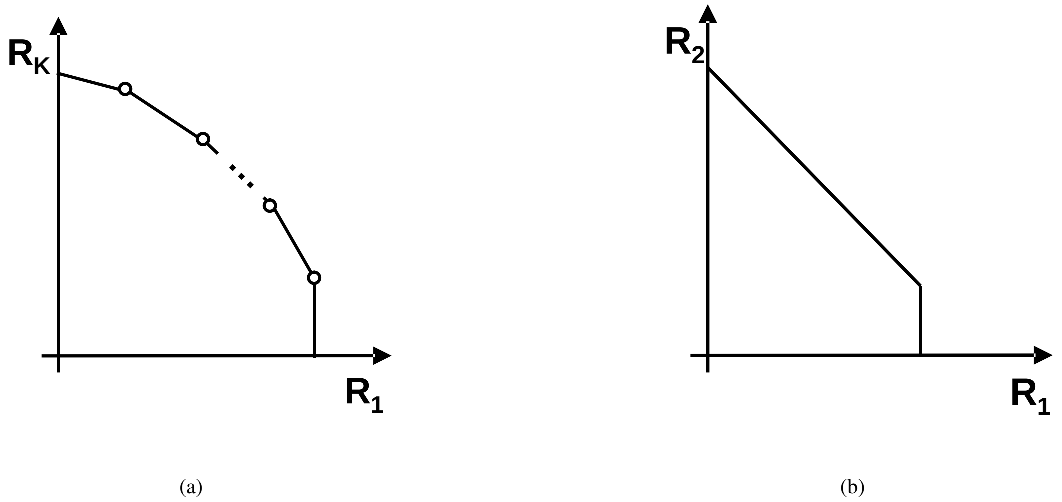
[4] T M. Cover. Broadcast channels. *IEEE Transactions on Information Theory*, 18:2–14, January 1972.

[5] S. N. Diggavi, N. Al-Dhahir, and A. R. Calderbank, *Diversity embedding in multiple antenna communications*, in DIMACS workshop on Network Information Theory, March 2003. Also part of AMS edited volume “Network Information Theory”, edited by P. Gupta, G. Kramer and A J. van Wijngaarden, AMS volume 66, Series on Discrete Mathematics and Theoretical Computer Science, pp 285–302, 2004.

[6] S. N Diggavi and D N C. Tse, “On opportunistic codes and broadcast codes with degraded message sets”, *IEEE Information Theory Workshop (ITW)*, Punta del Este, Uruguay, pp –, March, 2006.

[7] S. N Diggavi and D N C. Tse, *Fundamental Limits of Diversity-Embedded Codes over Fading Channels*, IEEE ISIT, Adelaide, Australia, pp 510–514, September, 2005.

[8] A A. El-Gamal. Capacity of the product and sum of two unmatched broadcast channels. *Problemy Peredachi Informatsii*, 17(1):3–23, January-March 1980 (English translation *Problems in Information Transmission*, 16(1):1–16, January-March 1980).



(a) (b)

Fig. 4. Figure (a) shows capacity region for K-user problem with two messages, one common and one private for user K only as given in Theorem II.3 (i). Figure (b) shows capacity region for K-user problem with two messages, one common and another for all users except user 1 as given in Theorem II.3 (ii).

[9] R. Etkin, D N C. Tse and H. Wang. Gaussian Interference Channel Capacity to Within One Bit submitted to *IEEE Transactions on Information Theory*.

[10] R G. Gallager. Capacity and coding for degraded broadcast channels. *Problemy Peredachi Informatsii*, 10(3):3–14, 1974.

[11] T S. Han, “The capacity region of the deterministic broadcast channel with common message”, *IEEE Transactions on Information Theory*, 27(1):122–125, January 1980.

[12] J. Korner and K. Marton. General broadcast channels with degraded message sets. *IEEE Trans. IT*, 23(1):60–64, January 1977.

[13] A. Lapidoth and P. Narayan. Reliable communication under channel uncertainty. *IEEE Trans. IT*, 44(6):2148-2177, 1988.

[14] K. Marton. “A coding theorem for the discrete memoryless broadcast channel”, *IEEE Transactions on Information Theory*, 25(3):306–311, May 1979.

[15] S Shamai (Shitz). A broadcast strategy for the Gaussian slowly fading channel. In *IEEE ISIT*, page 150, June 1997.

[16] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.

[17] H. Weingarten, Y. Steinberg, and S Shamai (Shitz)., *On the Capacity Region of the Multi-Antenna Broadcast Channel with Common Messages*, IEEE ISIT, pp 2195-2199, 2006.

[18] R W. Yeung and N. Cai. *Network Coding Theory*, Now Publishers Inc., 2006.

[19] L. Zhao and S-Y. Chung. Perfect rate-compatible codes for the Gaussian channel. in *Proc. Allerton Conf. Commun., Contr., Computing*, Monticello, IL, Sep. 2007.

[20] L Zheng and D N C. Tse. Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels. *IEEE Trans. IT*, 49(5):1073–1096, May 2003.