On degraded two message set broadcast

S. Saeedi S. Diggavi C. Fragouli V. Prabhakaran

Abstract

We consider the two message set problem, where a source broadcasts a common message W_1 to an arbitrary set of receivers $\mathcal{U}$ and a private message W_2 to a subset of the receivers $\mathcal{P} \subseteq \mathcal{U}$. Transmissions occur over linear deterministic channels. For the case where at most two receivers do not require the private message, we give an exact characterization of the capacity region, where achievability is through linear coding.

1 Introduction

In this paper we study the problem of degraded two-message broadcasting over linear deterministic channels. More specifically, the question we study is the reliable rates at which we can deliver a common message to all users and a private message to a subset of the users, over linear deterministic broadcast channels. This is a special case of a long-standing open question in multi-user informatoin theory of delivering a set degraded messages over a general broadcast channel. The degraded message set problem was first studied by Cover, in the context of the general problem of broadcast channels, in his celebrated paper on broadcast channels [5]. The solution for the case where there is a degradation order between the users' channels was given in [3, 7]. The problem of general two-user broadcast channel with a degraded two message set requirement was solved by Korner and Marton in [8]. However there is comparitively little understanding when there are either more than two users, and/or more than two degraded messages. Recent progress on a special case of this question has been made in [10].

The linear deterministic channel model, introduced in [1], was motivated by its intimate connection to linear Gaussian models. Many insights gained from the study of such deterministic channels have carried over to the noisy Gaussian case in many situations including the wireless relay networks [2], interference channel [4], and relay-interference networks [9] and Therefore, this motivates the study of this special broadcast model in this paper. Recently [11] solved a three-user, degraded three (nested) message set problem over linear deterministic broadcast

channels. This paper builds on these results to an arbitrary number of users, but with the restriction that at most two users do not need all the messages. The main result is summarized in Theorem 2.1. The primary difficulty in this problem is the tension between delivering a common message to all the users (akin to a compound channel problem) and delivering a private message to a subset of users. We show that this tension can be optimally resolved by carefully selecting a structured linear transmission code, which is discovered by solving a matrix completion problem. The solution to this problem shows an intimate connection between our problem and network coding, since we need to judiciously mix independent messages. Another ingredient used is that we reveal some information about the private message even to the users only interested in only the common message. This has some connection to indirect decoding proposed in [10]. An extension of our work to three nested messages, is straightforward, and is summarized in Theorem 2.2.

The paper is organized as follows. In Section 2, we formally define the problem and give the main results of the paper. The rest of the paper is devoted to the proof of the main result, starting with the outer bound in Section 3. The construction of the structured linear code achieving the outer bound is given in Section 4.

2 Problem Formulation and Results

2.1 Model

The problem of interest is communication of a common message and a private message to a set of receivers $\mathcal{U} = \{1, \cdots, K\}$ through a *linear deterministic broadcast channel* [1]. The common message W_1 of rate R_1 is required at all the receivers while the private message W_2 of rate R_2 is required only at receivers $i \in \mathcal{P}$, $\mathcal{P}$ being a subset of $\mathcal{U}$. We call this scenario, the *two-message set* scenario.

The underlying channel model is essentially the same as studied in [11]. The input X to the channel lies in an m dimensional vector space $\mathbb{F}^m$, where $\mathbb{F}$ is a finite field. The received signal $Y_i \in \mathbb{F}^{n_i}$ at each receiver i is

$$Y_i = \mathbf{H}_i X, \tag{1}$$

where the channel matrix $\mathbf{H}_i$ is an $n_i \times m$ matrix in $\mathbb{F}$ of rank r_i.

We denote with $\mathcal{N}_i$ the nullspace of $\mathbf{H}_i$. Furthermore, for any subset $\mathcal{S}$ of $\mathcal{U}$, $\mathcal{S} = \{i_1, \cdots, i_{|\mathcal{S}|}\}$, we denote the rank of the matrix that collects the corresponding

channels as

$$\text{rank}\begin{bmatrix} \mathbf{H}_{i_1} \\ \vdots \\ \mathbf{H}_{i_{|\mathcal{S}|}} \end{bmatrix} \triangleq r_{i_1,\cdots i_{|\mathcal{S}|}}, \tag{2}$$

and the nullspace of this augmented matrix as $\mathcal{N}_{i_1,\cdots i_{|\mathcal{S}|}}$.

2.2 Main Result

Theorem 2.1 *The capacity region $\mathcal{R}$ of the two-message set broadcasting over linear deterministic channels, with $\mathcal{U} = \{1, \cdots, K\}$ and $\mathcal{P} = \{3, \cdots, K\}$, is given by*

$$\begin{aligned} R_1 &\leq \min_{i\in\mathcal{U}}\{r_i\} && (3)\\ R_1 + R_2 &\leq \min_{i\in\mathcal{P}}\{r_i\} && (4)\\ 2R_1 + R_2 &\leq \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}, && (5) \end{aligned}$$

where $|\mathbb{F}|$ is greater than K. The rates given above are expressed in $\log_{|\mathbb{F}|}(\cdot)$.

■

This result can easily be extended to the *three-message set* problem, where all receivers are interested in the common message W_1, users $\{2, \ldots, K\}$ want messages (W_1, W_2) and users $\{3, \ldots, K\}$ want all three messages (W_1, W_2, W_3).

Theorem 2.2 *The capacity region $\mathcal{R}$ of the three message set broadcast over linear deterministic channels, is given by*

$$\begin{aligned} R_1 &\leq \min_{i}\{r_i\}, && (6)\\ R_1 + R_2 &\leq \min_{i\geq 2}\{r_i\}, && (7)\\ R_1 + R_2 + R_3 &\leq \min_{i\geq 3}\{r_i\}, \ \textit{and} && (8)\\ 2R_1 + R_2 + R_3 &\leq \min_{i\geq 3}\{r_1 + r_2 + r_{1,2,i} - r_{1,2}\}. && (9) \end{aligned}$$

where $|\mathbb{F}|$ is greater than K. The rates given above are expressed in $\log_{|\mathbb{F}|}(\cdot)$.

■

3 Outer bound

In this section we prove an outer bound to the more general problem; i.e., when $\mathcal{P}$ can be any subset of $\mathcal{U}$. For $\mathcal{P} = \{3, \cdots, K\}$, the converse to Theorem 2.1 follows.

Theorem 3.1 *The capacity region of the linear deterministic broadcast channel in the two-message set scenario with $\mathcal{U} = \{1, \cdots, K\}$ and $\mathcal{P} \subseteq \mathcal{U}$ is inside the polytope characterized by*

$$R_1 \leq \min_{i \in \mathcal{U}} \{r_i\} \tag{10}$$

$$R_1 + R_2 \leq \min_{i \in \mathcal{P}} \{r_i\} \tag{11}$$

$$\forall k \leq |\mathcal{P}^c| :$$

$$kR_1 + R_2 \leq \min_{i \in \mathcal{P}, j_1, \cdots j_k \notin \mathcal{P}^c} \{\sum_{l=1}^{k} r_{j_l} + r_{j_1, j_2, \cdots, j_k, i} - r_{j_1, j_2, \cdots, j_k}\}. \tag{12}$$

Proof Assume communication using blocks of an arbitrary length n, and denote the received signal at each receiver i by Y_i^n. Then (10) and (11) follow from:

$$\forall i \in \mathcal{U} : \; n(R_1) \leq I(W_1; Y_i^n) \leq H(Y_i^n) - H(Y_i^n | W_1) \leq n r_i. \tag{13}$$

$$\begin{aligned} \forall i \in \mathcal{P} : \quad n(R_1 + R_2) &\leq I(W_1, W_2; Y_i^n) && (14)\\ &\leq H(Y_i^n) - H(Y_i^n | W_1, W_2) && (15)\\ &\leq n r_i. && (16) \end{aligned}$$

From (13), it follows that

$$H(Y_i^n | W_1) \leq n(r_i - R_1). \tag{17}$$

To obtain (12), we use the approach in [11]. Each time, we give the received signal at receivers $j_1 \cdots j_k \in \mathcal{P}^c$ to receiver $i \in \mathcal{P}$:

$$\begin{aligned} n(R_2) &\leq I(W_2; Y_i^n)\\ &\leq I(W_2; Y_i^n | W_1)\\ &\leq I(W_2; Y_{j_1}^n, Y_{j_2}^n, \cdots, Y_{j_k}^n, Y_i^n | W_1)\\ &\overset{(a)}{=} H(Y_{j_1}^n, Y_{j_2}^n, \cdots, Y_{j_k}^n, Y_i^n | W_1)\\ &= \sum_{l=1}^{k} H(Y_{j_l}^n | W_1, Y_{j_1}^n, \cdots, Y_{j_{l-1}}^n) + H(Y_i^n | Y_{j_1}^n, \cdots, Y_{j_k}^n, W_1)\\ &\leq \sum_{l=1}^{k} H(Y_{j_l}^n | W_1) + H(Y_i^n | Y_{j_1}^n, \cdots, Y_{j_k}^n, W_1)\\ &\overset{(b)}{\leq} \sum_{l=1}^{k} n(r_{j_l} - R_1) + n(r_{j_1, j_2, \cdots, j_k, i} - r_{j_1, j_2, \cdots, j_k}). \end{aligned}$$

Equality (a) is the result of the deterministic assumption and inequality (b) is obtained by using (17) and upper bounding $H(Y_i^n|Y_1^n, Y_2^n)$ by $n(r_{1,2,i} - r_{1,2})$ as in [11]. ■

4 Achievability Proof

The challenge in the achievability scheme design for the two-message problem stems from the fact that, although the first two receivers are only interested in the common message of rate R_1, they might nevertheless also need decode additional partial information, to allow the reception of the private message by the remaining receivers. For example, if the common message is represented by variable w_1 and the private message is represented by variables $[w_2 \ w_3]$, the first receiver might decode w_1 and $w_2 + w_3$, while the second receiver w_1 and $w_2 + 2w_3$. Instead of specifying in advance what the first two receivers decode, we will instead derive conditions on the structure of the matrices they observe, that guarantee they can decode the common information. We will then essentially reduce our problem to a set of matrix completion problems, where we will now require some of the involved matrices to have full rank, and some submatrices to satisfy some rank conditions (which arise from the need for some users to only decode some variables). We will finally show that such matrix completion problems can be simultaneously satisfied with a single matrix by applying the sparse-zero lemma [6].

The technical steps can be described as follows:

- We will design in section 4.1 a basis for $\mathbb{F}^m$ which depends on the channel matrices $\mathbf{H}_1, \mathbf{H}_2$. This is used to design a linear encoding scheme which depends on a matrix of indeterminates $\tilde{\mathbf{A}}$, which we will attempt to fill (complete) so that the decoding requirements are fulfilled. The basis is chosen such that the first two receivers can directly obtain linear combinations of specific subsets of the rows of $\tilde{\mathbf{A}}$, while the remaining receivers can potentially observe some linear transformation of $\tilde{\mathbf{A}}$. We impose a structure on $\tilde{\mathbf{A}}$. Given this structure, the decoding requirements of users 1 and 2 constrains some entries in $\tilde{\mathbf{A}}$. The structure imposed on the indeterminates is parametrized by, a_1, a_2, and b. The matrix completion problem is to fill the rest of $\tilde{\mathbf{A}}$ appropriately.

- In section 4.2, we derive necessary conditions that allow decodability for all receivers. For the first two receivers these conditions require specific submatrices of $\tilde{\mathbf{A}}$ to have given ranks, as well as relationships between column spaces of specific submatrices. These imposed constraints will need to be respected while completing $\tilde{\mathbf{A}}$ so that any other user, with appropriate rank

requirements, is able to decode all the messages. We will show that these rank requirements match the bounds given in Theorem 2.1.

- We will then show in section 4.3, that there is a universal choice for $\tilde{\mathbf{A}}$ which will satisfy all the users. This is done by applying the sparse zeros lemma [6] to the new set of matrix completions to show that there exist variable choices that satisfy all the decodability conditions. To apply the sparse zero lemma, we will make some judicious choice for the structure parameters a_1, a_2, and b.

The example given in Section 4.4 will illustrate some of the ideas outlined above.

4.1 Problem Reduction

Let $w_{1,1}, \cdots w_{1,R_1}$ and $w_{2,1}, \cdots, w_{2,R_2}$ be the variables in $\mathbb{F}$ for messages W_1 and W_2 respectively, and W in $\mathbb{F}^{R_1+R_2}$ the vector with coordinates in the standard basis

$$W = [w_{1,1} \ \ldots \ w_{1,R_1} \ w_{2,1} \ \ldots \ w_{2,R_2}]^T. \tag{18}$$

We will use linear coding as our encoding scheme and broadcast a signal in the form

$$X = \mathbf{A}W. \tag{19}$$

$\mathbf{A}$ is the $m \times (R_1 + R_2)$ matrix over the finite field $\mathbb{F}$ that we need to design so that the first two receivers decode W_1 and all the remaining both W_1 and W_2.

We choose a new basis, $\mathcal{B}$, for $\mathbb{F}^m$ in the following manner (see Fig. 1): First select a set of vectors $\mathcal{B}_\Phi$ such that $\langle \mathcal{B}_\Phi \rangle = \mathcal{N}_{12}$. Then select vectors $\mathcal{B}_1$ and $\mathcal{B}_2$ such that $\langle \mathcal{B}_\Phi \rangle \oplus \langle \mathcal{B}_1 \rangle = \mathcal{N}_2$, and $\langle \mathcal{B}_\Phi \rangle \oplus \langle \mathcal{B}_2 \rangle = \mathcal{N}_1$. Form, finally, $\mathcal{B}_{12}$ such that $\langle \mathcal{B}_\Phi \rangle \oplus \langle \mathcal{B}_1 \rangle \oplus \langle \mathcal{B}_2 \rangle \oplus \langle \mathcal{B}_{12} \rangle = \mathbb{F}^m$. Let $\mathcal{B} = \mathcal{B}_\Phi \cup \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_{12}$. Let the associated transformation matrix be

$$\mathbf{V} = \left[\ \mathbf{V}_{12} \ \middle| \ \mathbf{V}_2 \ \middle| \ \mathbf{V}_1 \ \middle| \ \mathbf{V}_\Phi \ \right],$$

where the column vectors of $\mathbf{V}_{12}$ are the vectors in $\mathcal{B}_{12}$ and so on. Note that

$$\begin{aligned} |\mathcal{B}_\Phi| &= m - r_{12}, \\ |\mathcal{B}_1| &= r_{12} - r_2, \\ |\mathcal{B}_2| &= r_{12} - r_1, \\ |\mathcal{B}_{12}| &= r_1 + r_2 - r_{12}. \end{aligned}$$

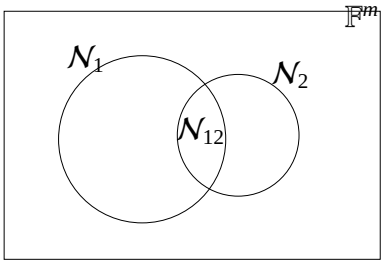


Figure 1: Venn diagram of the null spaces of the 2 receivers requiring only W_1.

Then we may expand the input X to the channel using this basis $\mathcal{B}$ as follows

$$X = \mathbf{V}\tilde{X} = \left[\begin{array}{c|c|c|c} \mathbf{V}_{12} & \mathbf{V}_2 & \mathbf{V}_1 & \mathbf{V}_\Phi \end{array}\right] \left[\begin{array}{c} \tilde{X}_{12} \\ \tilde{X}_2 \\ \tilde{X}_1 \\ \tilde{X}_\Phi \end{array}\right],$$

where $\tilde{X} \in \mathbb{F}^m$ is the vector of coefficients of the basis vectors under this basis expansion. Further, we defined $\tilde{X}_{12}$ to be the first $|\mathcal{B}_{12}|$ coefficients of $\tilde{X}$ corresponding to the column vectors of $\mathbf{V}_{12}$, and $\tilde{X}_2$ to be the next $|\mathcal{B}_2|$ coefficients and so on. It is clear that we may take $\tilde{X} \in \mathbb{F}^m$ to be the input of an equivalent channel in which the channel output at receiver-i is

$$Y_i = \mathbf{H}_i\mathbf{V}\tilde{X}.$$

For user-1, the resulting channel matrix is

$$\begin{aligned} \mathbf{H}_1\mathbf{V} &= \mathbf{H}_1\left[\begin{array}{c|c|c|c} \mathbf{V}_{12} & \mathbf{V}_2 & \mathbf{V}_1 & \mathbf{V}_\Phi \end{array}\right] \\ &= \left[\begin{array}{c|c|c|c} \mathbf{H}_1\mathbf{V}_{12} & \mathbf{0} & \mathbf{H}_1\mathbf{V}_1 & \mathbf{0} \end{array}\right] \end{aligned}$$

Hence,

$$Y_1 = \left[\begin{array}{c|c} \mathbf{H}_1\mathbf{V}_{12} & \mathbf{H}_1\mathbf{V}_1 \end{array}\right] \left[\begin{array}{c} \tilde{X}_{12} \\ \tilde{X}_1 \end{array}\right].$$

Moreover, by the manner in which $\mathcal{B}$ was formed, the matrix $\left[\begin{array}{c|c} \mathbf{H}_1\mathbf{V}_{12} & \mathbf{H}_1\mathbf{V}_1 \end{array}\right]$ has full (column) rank. Hence, we may replace the output at user-1 without loss of generality with

$$\tilde{Y}_1 = \left[\begin{array}{c} \tilde{X}_{12} \\ \tilde{X}_1 \end{array}\right] = \left[\begin{array}{cccc} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{array}\right] \tilde{X} =: \tilde{\mathbf{H}}_1\tilde{X}. \tag{20}$$

Similarly,

$$\tilde{Y}_2 = \begin{bmatrix} \tilde{X}_{12} \\ \tilde{X}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tilde{X} =: \tilde{\mathbf{H}}_2 \tilde{X}. \tag{21}$$

For the rest of the users, we simply set

$$\tilde{Y}_k = Y_k = \mathbf{H}_k \mathbf{V} \tilde{X} =: \tilde{\mathbf{H}}_k \tilde{X}, \quad k \in \mathcal{P} = 3, 4, \ldots, K, \tag{22}$$

where

$$\begin{aligned} \tilde{\mathbf{H}}_k &= \left[\begin{array}{c|c|c|c} \mathbf{H}_k \mathbf{V}_{12} & \mathbf{H}_k \mathbf{V}_2 & \mathbf{H}_k \mathbf{V}_1 & \mathbf{H}_k \mathbf{V}_\Phi \end{array} \right] \\ &=: \left[\begin{array}{c|c|c|c} \tilde{\mathbf{H}}_k^{12} & \tilde{\mathbf{H}}_k^{2} & \tilde{\mathbf{H}}_k^{1} & \tilde{\mathbf{H}}_k^{\Phi} \end{array} \right]. \end{aligned} \tag{23}$$

We have now an equivalent problem in which the input to the channel is $\tilde{X} \in \mathbb{F}^m$, and the received signal at user-i is

$$\tilde{Y}_i = \tilde{\mathbf{H}}_i \tilde{X}, \quad i \in \mathcal{U}, \tag{24}$$

where $\tilde{\mathbf{H}}_i$ are given by (20)-(22). The following lemma calculates the ranks of certain submatrices of $\tilde{\mathbf{H}}_i$ and will be used in 4.3 to prove the achievability of our coding theorem.

Lemma 4.1 *For $k \in \mathcal{P}$,*

$$\begin{aligned} &rank\left(\tilde{\mathbf{H}}_k^{\Phi}\right) = r_{12k} - r_{12}, \\ &rank\left(\left[\begin{array}{c|c} \tilde{\mathbf{H}}_k^{1} & \tilde{\mathbf{H}}_k^{\Phi} \end{array} \right]\right) = r_{2k} - r_2, \\ &rank\left(\left[\begin{array}{c|c} \tilde{\mathbf{H}}_k^{2} & \tilde{\mathbf{H}}_k^{\Phi} \end{array} \right]\right) = r_{1k} - r_1, \\ &rank\left(\left[\begin{array}{c|c|c} \tilde{\mathbf{H}}_k^{2} & \tilde{\mathbf{H}}_k^{1} & \tilde{\mathbf{H}}_k^{\Phi} \end{array} \right]\right) \geq \max \left\{ \begin{array}{c} r_{1k} - r_1, \\ r_{2k} - r_2, \\ r_k - r_1 - r_2 + r_{12} \end{array} \right\}, \\ &rank\left(\left[\begin{array}{c|c|c|c} \tilde{\mathbf{H}}_k^{12} & \tilde{\mathbf{H}}_k^{2} & \tilde{\mathbf{H}}_k^{1} & \tilde{\mathbf{H}}_k^{\Phi} \end{array} \right]\right) = r_k. \end{aligned}$$

Proof The key point of the proof is to note that rank $(\mathbf{H}_k \mathbf{V})$ is the same for all matrices $\mathbf{V}$ with the same column space. Thus, without loss of generality we assume in this proof that $\mathcal{B}_\Phi = \mathcal{B}_{\Phi,\bar{k}} \cup \mathcal{B}_{\Phi,k}$ such that $\left\langle \mathcal{B}_{\Phi,\bar{k}} \right\rangle = \mathcal{N}_{12k}$, $\mathcal{B}_2 \cup \mathcal{B}_\Phi = \mathcal{B}_{2,\bar{k}} \cup \mathcal{B}_{2,k}$, such that $\left\langle \mathcal{B}_{2,\bar{k}} \right\rangle = \mathcal{N}_{1,k}$, and $\mathcal{B}_1 \cup \mathcal{B}_\Phi = \mathcal{B}_{1,\bar{k}} \cup \mathcal{B}_{1,k}$, such that $\left\langle \mathcal{B}_{1,\bar{k}} \right\rangle = \mathcal{N}_{2k}$. This way, we have the following:

1. $\text{rank}\left(\tilde{\mathbf{H}}_k^{\Phi}\right)$:
   $\tilde{\mathbf{H}}_k^{\Phi} = \mathbf{H}_k\mathbf{V}_{\Phi} = [\mathbf{0}\ \mathbf{H}_k\mathbf{V}_{\Phi,k}]$, where $\mathbf{V}_{\Phi,k}$ denotes the matrix that has as its columns the vectors in $\mathcal{B}_{\Phi,k}$. Because the vectors in $\mathcal{B}_{\Phi,k}$ are linearly independent and do not belong in the null space of $\mathbf{H}_k$, the columns $\mathbf{H}_k\mathbf{V}_{\Phi,k}$ are also linearly independent, and thus $\text{rank}\mathbf{H}_k\mathbf{V}_{\Phi} = |\mathcal{B}_{\Phi,k}| = |\mathcal{N}_{12}| - |\mathcal{N}_{12k}| = r_{12k} - r_{12}$.

2. $\text{rank}\left(\left[\ \tilde{\mathbf{H}}_k^1 \mid \tilde{\mathbf{H}}_k^{\Phi}\ \right]\right)$:
   Similarly, $\text{rank}\left(\left[\ \tilde{\mathbf{H}}_k^1 \mid \tilde{\mathbf{H}}_k^{\Phi}\ \right]\right) = \text{rank}\left(\mathbf{H}_k\left[\ \mathbf{V}_1 \mid \mathbf{V}_{\Phi}\ \right]\right) = |\mathcal{B}_{1,k}| = |\mathcal{N}_2| - |\mathcal{N}_{2k}| = r_{2k} - r_2$.

3. $\text{rank}\left(\left[\ \tilde{\mathbf{H}}_k^2 \mid \tilde{\mathbf{H}}_k^{\Phi}\ \right]\right)$:
   This rank is calculated similarly and found to be $r_{1k} - r_1$.

4. $\text{rank}\left(\left[\ \tilde{\mathbf{H}}_k^2 \mid \tilde{\mathbf{H}}_k^1 \mid \tilde{\mathbf{H}}_k^{\Phi}\ \right]\right)$:
   Though it turns out that calculating this rank is not trivial, the following bounds prove sufficient for our achievability conclusion.

$$\begin{aligned} & \text{rank}\left(\left[\ \tilde{\mathbf{H}}_k^2 \mid \tilde{\mathbf{H}}_k^1 \mid \tilde{\mathbf{H}}_k^{\Phi}\ \right]\right) && (25)\\ = \ & \text{rank}\left(\mathbf{H}_k[\mathbf{V}_1|\mathbf{V}_2|\mathbf{V}_{\Phi}]\right) \\ = \ & |\langle\mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_{\Phi}\rangle| - |\langle\mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_{\Phi}\rangle \cap \mathcal{N}_k| \\ \geq \ & |\mathcal{N}_1| + |\mathcal{N}_2| - |\mathcal{N}_{12}| - |\mathcal{N}_k| \\ = \ & r_k - r_1 - r_2 + r_{12}. && (26)\end{aligned}$$

   Furthermore,

$$\text{rank}\left(\left[\ \tilde{\mathbf{H}}_k^2 \mid \tilde{\mathbf{H}}_k^1 \mid \tilde{\mathbf{H}}_k^{\Phi}\ \right]\right) \geq r_{1k} - r_1 \quad (27)$$

$$\text{rank}\left(\left[\ \tilde{\mathbf{H}}_k^2 \mid \tilde{\mathbf{H}}_k^1 \mid \tilde{\mathbf{H}}_k^{\Phi}\ \right]\right) \geq r_{2k} - r_2. \quad (28)$$

5. $\text{rank}\left(\left[\ \tilde{\mathbf{H}}_k^{12} \mid \tilde{\mathbf{H}}_k^2 \mid \tilde{\mathbf{H}}_k^1 \mid \tilde{\mathbf{H}}_k^{\Phi}\ \right]\right)$:
   This rank is immediately r_k. ■

4.2 Decodability Basic Lemmas

To argue decodability of W_1 at receiver 1 and 2, and decodability of W_1, W_2 at receivers $k \in \{3, \cdots, K\}$, we need the following lemmas.

Lemma 4.2 *Consider* $\mathbf{G} \in \mathbb{F}^{n\times m}$ *and* $W = [\ w_1\ \cdots\ w_m\ ]^T$. $[\ w_1\ \cdots\ w_d\ ]^T$, $d \leq m$, *can be decoded uniquely from* $\mathbf{G}W$ *iff*

- $\left\langle \underline{g}_1, \cdots, \underline{g}_d \right\rangle \cap \left\langle \underline{g}_{d+1}, \cdots, \underline{g}_m \right\rangle = \phi$,
- $\{\underline{g}_i\}_{i=1}^{d}$ *are linearly independent,*

where $\{\underline{g}_i\}_{i=1}^{m}$ *are the columns of* $\mathbf{G}$.

Proof $w_1 \cdots w_d$ is uniquely decodable from $\mathbf{G}\underline{W}$ *iff*

$$\mathbf{G}\underline{W} \neq \mathbf{G}\underline{V} \tag{29}$$

for all $\underline{V}$ that differ from $\underline{W}$ in the first d entries. Equivalently,

$$\sum_{i=1}^{d} w_i \underline{g}_i + \sum_{i=d+1}^{n} w_i \underline{g}_i \neq \sum_{i=1}^{d} v_i \underline{g}_i + \sum_{i=d+1}^{n} v_i \underline{g}_i, \tag{30}$$

or,

$$\sum_{i=1}^{d} \delta_i \underline{g}_i \neq \sum_{i=d+1}^{n} \beta_i \underline{g}_i \quad \forall \delta_i, \beta_i \in \mathbb{F},\ \delta_i \neq 0. \tag{31}$$

This concludes the proof. ■

Lemma 4.3 *Consider a matrix* $\mathbf{B} = \left[\ \mathbf{B}_1 \ \middle|\ \mathbf{B}_2\ \right]$, *where* $\mathbf{B}_1 \in \mathbb{F}^{n\times d}$, $\mathbf{B}_2 \in \mathbb{F}^{n\times(m-d)}$, *and* $d \leq \min\{n, m\}$. *Form the matrix* $\mathbf{G} = \left[\ \mathbf{B}_1 \ \middle|\ \mathbf{L}_1\ \right]$, *where* $\mathbf{L}_1 \in \mathbb{F}^{n\times l}$ *is the first component of* $\mathbf{B}_2 = \mathbf{L}_1\mathbf{L}_2$. *If* $l \leq n - d$, *then* $\mathbf{G}$ *being full-rank guarantees*

- $\left\langle \underline{b}_1, \cdots, \underline{b}_d \right\rangle \cap \left\langle \underline{b}_{d+1}, \cdots, \underline{b}_m \right\rangle = \phi$,
- $\{\underline{b}_i\}_{i=1}^{d}$ *are linearly independent,*

where $\{\underline{b}_i\}_{i=1}^{m}$ *are the columns of* $\mathbf{B}$.

Proof Let $\text{rank}(\mathbf{B}_2) = l$. $\mathbf{B}_2$ can thus be written as

$$\mathbf{B}_2 = \mathbf{L}_1\mathbf{L}_2, \tag{32}$$

where $\mathbf{L}_1$ is a full rank matrix of dimension $n \times l$ and $\mathbf{L}_2$ a full rank matrix of dimension $l\times(m-d)$. $\mathbf{L}_1$ is essentially just a set of linearly independent columns of $\mathbf{B}_2$ spanning its columns space. Now form $\mathbf{G} = \left[\ \mathbf{B}_1 \ \middle|\ \mathbf{L}_1\ \right]$ which is of dimension $n \times (d + l)$. since $(d + l) \leq n$, $\mathbf{G}$ being full-rank guarantees

1. $\left\langle \underline{b}_1, \cdots, \underline{b}_d \right\rangle \cap \left\langle \underline{l}_1, \cdots, \underline{l}_l \right\rangle = \phi$, where $\{\underline{l}_i\}_{i=1}^{l}$ are the column vectors of $\mathbf{L}_1$.
2. $\{\underline{b}_i\}_{i=1}^{d}$ are linearly independent.

The fact that

$$\left\langle \underline{l}_1, \cdots, \underline{l}_r \right\rangle = \left\langle \underline{b}_{d+1}, \cdots, \underline{b}_m \right\rangle \tag{33}$$

concludes the proof. ■

To summarize lemma 4.2 and 4.3 in a more intuitive way, let $W = [w_1 \ \cdots \ w_m]^T$ and for $i \leq j$, let $W_i^j = [w_i \ w_{i+1} \ \cdots \ w_j]^T$. Then

$$\begin{aligned} \mathbf{B}W &= \left[\begin{array}{c|c|c} \mathbf{B}_1 & \mathbf{L}_1\mathbf{L}_2 \end{array} \right] W & (34) \\ &= \left[\begin{array}{c|c} \mathbf{B}_1 & \mathbf{L}_1 \end{array} \right] \left[\begin{array}{c} W_1^d \\ \mathbf{L}_2 W_{d+1}^m \end{array} \right] & (35) \\ &= \mathbf{G} \left[\begin{array}{c} W_1^d \\ \mathbf{L}_2 W_{d+1}^m \end{array} \right]. & (36) \end{aligned}$$

One should note now that $\mathbf{G}$ of dimension $n \times (d + l)$ $(d + l \leq n)$ being full-rank guarantees decodability of $[\ w_1 \ \cdots \ w_d \ \ W_{d+1}^m \mathbf{L}_2^T \ ]^T$.

Lemma 4.4 *Consider a matrix* $\mathbf{T}$ *over the finite field* $\mathbb{F}$ *of the form*

$$\mathbf{T} = \left[\begin{array}{c|c|c|c} \mathbf{T}_1 & \mathbf{T}_2 & \mathbf{T}_3 & \mathbf{T}_4 \end{array} \right]. \tag{37}$$

Let t_1, t_2, t_3, *and* t_4 *be non-negative integers such that*

$$\begin{aligned} &rank\left(\mathbf{T}_4 \right) \geq t_4, & (38) \\ &rank\left(\left[\begin{array}{c|c} \mathbf{T}_3 & \mathbf{T}_4 \end{array} \right] \right) \geq t_3 + t_4, & (39) \\ &rank\left(\left[\begin{array}{c|c} \mathbf{T}_2 & \mathbf{T}_4 \end{array} \right] \right) \geq t_2 + t_4, & (40) \\ &rank\left(\left[\begin{array}{c|c|c} \mathbf{T}_2 & \mathbf{T}_3 & \mathbf{T}_4 \end{array} \right] \right) \geq t_2 + t_3 + t_4, \ \textit{and} & (41) \\ &rank\left(\left[\begin{array}{c|c|c|c} \mathbf{T}_1 & \mathbf{T}_2 & \mathbf{T}_3 & \mathbf{T}_4 \end{array} \right] \right) \geq t_1 + t_2 + t_3 + t_4. & (42) \end{aligned}$$

Then, there are matrices $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{U}_3$, *and* $\mathbf{U}_4$ *such that the columns of* $\mathbf{U}_4$ *are drawn from the columns of* $\mathbf{T}_4$, *the columns of* $\mathbf{U}_3$ *from the columns of* $\mathbf{T}_3$ *and* $\mathbf{T}_4$, *the columns of* $\mathbf{U}_2$ *from the columns of* $\mathbf{T}_2$ *and* $\mathbf{T}_4$, *and, finally, the columns of* $\mathbf{U}_1$ *are taken from the columns of* $\mathbf{T}_1$, $\mathbf{T}_2$, $\mathbf{T}_3$, *and* $\mathbf{T}_4$ *such that they satisfy*

- $rank\,(\mathbf{U}_i) = t_i, \quad i \in \{1, 2, 3, 4\}$,
- $\left[\begin{array}{c|c|c|c} \mathbf{U}_1 & \mathbf{U}_2 & \mathbf{U}_3 & \mathbf{U}_4 \end{array} \right]$ *has linearly independent columns.* ■

Proof We form a basis $\mathcal{E}$ for the column space of $\mathbf{T}$ as follows: We pick exactly rank($\mathbf{T}_4$) linearly independent vectors from $\mathbf{T}_4$ and denote the set of these vectors by $\mathcal{T}_4$. Then, we find a set $\mathcal{T}_3$ of rank($[\mathbf{T}_3|\mathbf{T}_4]$) − rank($\mathbf{T}_4$) linearly independent vectors from $\mathbf{T}_3$ such that $\mathcal{T}_3 \cup \mathcal{T}_4$ is a linearly independent set. Similarly, we proceed to find a set of vectors $\mathcal{T}_2$ from the columns of $\mathbf{T}_2$ and $\mathcal{T}_1$ from the columns of $\mathbf{T}_1$ such that $|\mathcal{T}_2|$ = rank($[\mathbf{T}_2|\mathbf{T}_3|\mathbf{T}_4]$) − rank($[\mathbf{T}_3|\mathbf{T}_4]$), $|\mathcal{T}_1|$ = rank(T) − rank($[\mathbf{T}_2|\mathbf{T}_3|\mathbf{T}_4]$), and the vectors of $\mathcal{T}_4 \cup \mathcal{T}_3 \cup \mathcal{T}_2 \cup \mathcal{T}_1$ are linearly independent. Clearly,

$$\mathcal{E} = \mathcal{T}_4 \cup \mathcal{T}_3 \cup \mathcal{T}_2 \cup \mathcal{T}_1$$

is a basis for the column space of $\mathbf{T}$. We will now attempt to choose the $\mathbf{U}$ matrices such that

- $\mathbf{U}_4$ has t_4 distinct columns from $\mathcal{T}_4$,
- $\mathbf{U}_3$ has t_3 distinct columns from $\mathcal{T}_3 \cup \mathcal{T}_4$,
- $\mathbf{U}_2$ has t_2 distinct columns from $\mathcal{T}_2 \cup \mathcal{T}_4$,
- $\mathbf{U}_1$ has t_1 distinct columns from $\mathcal{E}$, and
- no two $\mathbf{U}$ matrices share a column.

It is clear that the lemma is proved if we can find such an assignment. We proceed in steps.

1) Let us choose any t_4 vectors from $\mathcal{T}_4$ to form the $\mathbf{U}_4$ matrix. We may do this since

   $$t_4 \leq \text{rank}(\mathbf{T}_4) = |\mathcal{T}_4|.$$

   After this, we have $|\mathcal{T}_4| - t_4$ vectors in $\mathcal{T}_4$ which could be assigned to other $\mathbf{U}$ matrices.

2) To form $\mathbf{U}_3$, we choose any $\min(t_3, |\mathcal{T}_3|)$ vectors from $\mathcal{T}_3$, and an additional $(t_3 - \mathcal{T}_3)^+$ vectors from the unassigned vectors in $\mathcal{T}_4$. We may do this if the number of unassigned vectors available in $\mathcal{T}_4$ is at least equal to the number of vectors we need, i.e.,

   $$|\mathcal{T}_4| - t_4 \geq (t_3 - |\mathcal{T}_3|)^+,$$

   where $(x)^+$ stands for $\max(0, x)$. But, this holds since

   $$|\mathcal{T}_4| + |\mathcal{T}_3| = \text{rank}([\mathbf{T}_3|\mathbf{T}_4]) \geq t_3 + t_4.$$

At the end of this step, we have

$$|\mathcal{T}_4| - t_4 - (t_3 - |\mathcal{T}_3|)^+$$

unassigned vectors in $\mathcal{T}_4$.

3) We now form $\mathbf{U}_2$ by choosing any $\min(t_2, |\mathcal{T}_2|)$ vectors from $\mathcal{T}_2$, and an additional $(t_2 - |\mathcal{T}_2|)^+$ vectors from among the unassigned vectors in $\mathcal{T}_4$ when available. This assignment fails only if we fall short of unassigned vectors in $\mathcal{T}_4$, and this happens when

$$t_2 - |\mathcal{T}_2| > |\mathcal{T}_4| - t_4 - (t_3 - |\mathcal{T}_3|)^+.$$

However, if $t_3 \geq |\mathcal{T}_3|$, the above condition reduces to

$$t_2 + t_3 + t_4 > |\mathcal{T}_2| + |\mathcal{T}_3| + |\mathcal{T}_4| = \text{rank}([\mathbf{T}_2|\mathbf{T}_3|\mathbf{T}_4])$$

which violates our hypothesis (41). Thus, we will fail to find a $\mathbf{U}_2$ only if

$$t_2 - |\mathcal{T}_2| > |\mathcal{T}_4| - t_4. \tag{43}$$

This case will be handled separately below. For now, we will proceed assuming that the above is not the case.

4) We choose t_4 vectors from among the vectors in $\mathcal{E}$ which have not been assigned. Since, $|\mathcal{E}| = \text{rank}(\mathbf{T})$ and

$$t_1 + t_2 + t_3 + t_4 \leq \text{rank}(\mathbf{T}),$$

we are guaranteed to find such vectors.

It only remains to prove the lemma when (43) holds. With this end, we will form a new basis $\mathcal{E}'$ for the column space of $\mathbf{T}$. We first pick the same $\mathcal{T}_4$ as before. Then we pick a $\mathcal{T}_2' \supseteq \mathcal{T}_2$ such that it has $t_2 - (|\mathcal{T}_4| - t_4)$ linearly independent vectors from the columns of $\mathbf{T}_2$ and $\mathcal{T}_2' \cup \mathcal{T}_4$ forms a basis for the column space of $[\mathbf{T}_2|\mathbf{T}_4]$. We may do this since (i) by (43), the required size of $\mathcal{T}_2'$ satisfies

$$\mathcal{T}_2' = t_2 - (|\mathcal{T}_4| - t_4) > \mathcal{T}_2,$$

and (ii) the required size of $\mathcal{T}_2'$ is not larger than the number of linearly independent vectors available in $\mathbf{T}_2$ such that $\mathcal{T}_2' \cup \mathcal{T}_4$ is a linearly independent set. This second fact can be seen from

$$(t_2 - (|\mathcal{T}_4| - t_4)) + |\mathcal{T}_4| = t_2 + t_4 \leq \text{rank}([\mathbf{T}_2|\mathbf{T}_4]).$$

Since $\mathcal{T}_3 \cup \mathcal{T}_2 \cup \mathcal{T}_4$ is a basis for the column space of $[\mathbf{T}_2|\mathbf{T}_3|\mathbf{T}_4]$, and the basis vectors of the new basis $\mathcal{E}'$ that we have picked so far includes all the vectors in $\mathcal{T}_4$ and $\mathcal{T}_2$, it is clear that there is a $\mathcal{T}_3' \subseteq \mathcal{T}_3$ such that $\mathcal{T}_3' \cup \mathcal{T}_2' \cup \mathcal{T}_4$ is also a basis for the column space of $[\mathbf{T}_2|\mathbf{T}_3|\mathbf{T}_4]$. Moreover, the size of this $\mathcal{T}_3'$ is

$$\begin{aligned} |\mathcal{T}_3'| &= |\mathcal{T}_2| + |\mathcal{T}_3| - |\mathcal{T}_2'| \\ &= |\mathcal{T}_2| + |\mathcal{T}_3| + |\mathcal{T}_4| - t_2 - t_4 \\ &= \text{rank}([\mathbf{T}_2|\mathbf{T}_3|\mathbf{T}_4]) - (t_2 + t_4) \\ &\geq t_3, \end{aligned} \tag{44}$$

where the last step follows from our hypothesis (41). We may now pick $\mathcal{T}_1$ to complete the new basis

$$\mathcal{E}' = \mathcal{T}_4 \cup \mathcal{T}_3' \cup \mathcal{T}_2' \cup \mathcal{T}_1.$$

We now proceed to pick the $\mathbf{U}$ matrices from this new basis following the same steps as before. Step 1 remains unchanged, step 2 goes through since (44) implies that we have sufficient number of vectors in $\mathcal{T}_3'$ from which to pick $\mathbf{U}_3$. Step 3 also succeeds since the number of unassigned vectors in $\mathcal{T}_4$ is exactly equal to the number of additional vectors needed to form $\mathbf{U}_2$, i.e.,

$$t_2 - |\mathcal{T}_2'| = |\mathcal{T}_4| - t_4.$$

Finally, a choice for $\mathbf{U}_4$ exists for the same reason as earlier. This completes the proof. ■

Lemma 4.5 *Consider a matrix* $\mathbf{G}$ *of the form*

$$\underbrace{\overset{m_1}{\leftrightarrow}\left[\mathbf{T}_1\right]\overset{m_2}{\leftrightarrow}\left[\mathbf{T}_2\right]\overset{m_3}{\leftrightarrow}\left[\mathbf{T}_3\right]\overset{m_4}{\leftrightarrow}\left[\mathbf{T}_4\right]}_{\mathbf{T}_{n\times m}} \underbrace{\begin{array}{|c|c|c|c|} \hline & \overset{t_1}{\longleftrightarrow} & \overset{t_2}{\leftrightarrow} & \overset{t_3}{\leftrightarrow} & \overset{t_4}{\leftrightarrow} \\ \hline & & 0 & 0 & 0 \\ \hline & & 0 & & 0 \\ \hline & & & 0 & 0 \\ \hline & & & & \\ \hline \end{array}}_{\mathbf{\Lambda}_{m\times p}} \begin{array}{l} \updownarrow m_1 \\ \updownarrow m_2 \\ \updownarrow m_3 \\ \updownarrow m_4 \end{array}$$

where the matrix $\mathbf{T}$ *is a fixed matrix and matrix* $\mathbf{\Lambda}$ *can be any matrix in* $\mathbb{F}^{m\times p}$ *in the given structure, and we have* $p \leq \min\{m, n\}$. $\mathbf{G}$ *can be made full-rank iff*

- $t_4 \leq rank\mathbf{T}_4$
- $t_2 + t_4 \leq rank[\mathbf{T}_2|\mathbf{T}_4]$
- $t_3 + t_4 \leq rank[\mathbf{T}_3|\mathbf{T}_4]$

- $t_2 + t_3 + t_4 \leq rank[\mathbf{T}_2|\mathbf{T}_3|\mathbf{T}_4]$
- $t_1 + t_2 + t_3 + t_4 \leq rank[\mathbf{T}_1|\mathbf{T}_2|\mathbf{T}_3|\mathbf{T}_4]$

Proof Let $\mathbf{\Lambda}$ be a $0-1$ matrix so that there is one 1 in each column. Thus multiplying $\mathbf{T}$ with $\mathbf{\Lambda}$ gives a column collection of $\mathbf{T}$ in the form of $[\mathbf{U}_1|\mathbf{U}_2|\mathbf{U}_3|\mathbf{U}_4]$ where the zero-one structure of $\mathbf{\Lambda}$ forces $\mathbf{U}_4$'s columns to be drawn from the set of columns of $\mathbf{T}_4$, $\mathbf{U}_3$'s columns to be drawn from the set of columns of $\mathbf{T}_3$ and $\mathbf{T}_4$, $\mathbf{U}_2$'s columns to be drawn from the set of columns of $\mathbf{T}_2$ and $\mathbf{T}_4$, and finally $\mathbf{U}_1$'s columns be drawn from all the $\mathbf{T}_i$'s. From lemma 4.4, we know that such $\mathbf{U}_i$'s exist that satisfy rank$\mathbf{U}_i = t_i$ and $[\mathbf{U}_1|\mathbf{U}_2|\mathbf{U}_3|\mathbf{U}_4]$ having linearly independent columns. So the lemma is proved by designing $\mathbf{\Lambda}$ so that the t_i columns pick $\mathbf{U}_i$'s columns, letting $\mathbf{G} = [\mathbf{U}_1|\mathbf{U}_2|\mathbf{U}_3|\mathbf{U}_4]$ be full rank. ■

4.3 Structured Linear Code

We will now prove the achievability part of our coding theorem for the equivalent channel defined in section 4.1. We will use linear coding as our encoding scheme and broadcast a signal in the form

$$\tilde{X} = \tilde{\mathbf{A}}W, \tag{45}$$

where $\tilde{\mathbf{A}}$ maps the vector of messages $W \in \mathbb{F}^{R_1+R_2}$ to $\tilde{X} \in \mathbb{F}^m$, the input to the channel. The message vector W consists of two parts W_1 and W_2. We select the following specific structure for the matrix $\tilde{\mathbf{A}}$

$$\tilde{\mathbf{A}} = \begin{array}{c} \overset{a_1-b}{\longleftrightarrow}\ \overset{a_2-b}{\longleftrightarrow}\ \overset{b}{\longleftrightarrow} \\ \left[\begin{array}{c|c|c} & \mathbf{0} & \mathbf{0} \;\; \mathbf{0} \\ \hline & \mathbf{0} & \mathbf{0} \\ \hline & & \mathbf{0} \;\; \mathbf{0} \\ \hline & & \end{array}\right] \end{array} \begin{array}{l} \updownarrow |\mathcal{B}_{12}| \\ \updownarrow |\mathcal{B}_2| \\ \updownarrow |\mathcal{B}_1| \\ \updownarrow |\mathcal{B}_\phi| \end{array} \tag{46}$$

where a_1, a_2 and b are size parameters to be decided, and satisfy $a_1 + a_2 - b \leq R_2$.

In the rest of this section, we first construct matrices $\mathbf{G}^{(k)}$ such that (1) For each $k \in \{1, 2\}$, if $\mathbf{G}^{(k)}$ is full-rank, then receiver k can decode W_1 from $\tilde{Y}_k$, and (2) For each $k \in \mathcal{P}$, if $\mathbf{G}^{(k)}$ is full-rank, then receiver k can decode W_1, W_2 from $\tilde{Y}_k$. We then find conditions on a_1, a_2, and b so that such $\mathbf{G}^{(k)}$ exist, and could be made fullrank for each $k \in \mathcal{U}$. Finally, we find a universal choice of a_1, a_2, and b and, using the sparse zeros lemma, an assignment of values to $\tilde{\mathbf{A}}$.

From (24), receiver $k \in \{1, 2\}$ can decode W_1, if it can decode it from $\tilde{Y}_k = \tilde{\mathbf{H}}_k\tilde{\mathbf{A}}W$. Let

$$\tilde{\mathbf{H}}_k\tilde{\mathbf{A}} = [\ \mathbf{B}_1^{(k)} \ \big|\ \mathbf{B}_2^{(k)}\ ], \tag{47}$$

where $\mathbf{B}_1^{(k)} \in \mathbb{F}^{r_k \times R_1}$, $\mathbf{B}_2^{(k)} = \mathbf{L}_1^{(k)}\mathbf{L}_2^{(k)}$, $l_1^k = \text{rank}(\mathbf{B}_2^{(k)})$, and $\mathbf{L}_1^{(k)} \in \mathbb{F}^{r_k \times l_1^k}$. Note that given the structure (20) and (21) of $\tilde{\mathbf{H}}_k$ and the structure (46) of $\tilde{\mathbf{A}}$,

(i) $\text{rank}\mathbf{B}_2^{(k)} \leq R_2 - a_k$,

(ii) $\tilde{\mathbf{H}}_1\tilde{\mathbf{A}}$ (resp. $\tilde{\mathbf{H}}_2\tilde{\mathbf{A}}$) is just a collection of the first $|\mathcal{B}_{12}|$ and the third $|\mathcal{B}_1|$ (resp. second $|\mathcal{B}_2|$) rows of $\tilde{\mathbf{A}}$.

From lemma 4.2 and lemma 4.3, we know that receiver $k \in \{1, 2\}$ can decode W_1 if $\text{rank}\mathbf{B}_2^{(k)} \leq r_k - R_1$ and if $\mathbf{G}^{(k)} = [\mathbf{B}_1^{(k)}|\mathbf{L}_1^{(k)}]$ is full-rank. (Recall from (10) that $R_1 \leq \min(r_k, R_1 + R_2)$ as required by lemma 4.3.) We have thus proved the following lemma.

Lemma 4.6 *Assuming that for $k \in \{1, 2\}$*

$$a_k \geq R_1 + R_2 - r_k, \tag{48}$$

receiver k can decode W_1 if $\mathbf{G}^{(k)}$ as defined above is full-rank.

Lemma 4.7 *For each $k \in \{1, 2\}$, there exists an assignment of values to $\tilde{\mathbf{A}}$ such that $\mathbf{G}^{(k)}$ is full-rank.*

Proof In the matrix $\mathbf{G}^{(k)} = [\mathbf{B}_1^{(k)}|\mathbf{L}_1^{(k)}]$, first select an assignment for the columns of $\mathbf{B}_2^{(k)}$ in $\mathbf{L}_1^{(k)}$ that makes them linearly independent (such an assignment exists from construction). Since the indeterminants in $\mathbf{B}_2^{(k)}$ are independent of those in $\mathbf{B}_1^{(k)}$, $\mathbf{G}^{(k)}$ can be made full-rank just by picking R_1 vectors from $\mathbb{F}^{r_k}$ linearly independent from the columns of $\mathbf{L}_1^{(k)}$. This is possible since $\text{rank}\mathbf{B}_2^{(k)} + R_1 \leq r_k$. The constructed $\mathbf{G}^{(k)}$ then uniquely defines $\tilde{\mathbf{A}}_k$ and can be completed arbitrarily to give a corresponding $\tilde{\mathbf{A}}$. ■

From (24), receiver $k \in \mathcal{P}$ can decode W_1 and W_2, if $\mathbf{G}^{(k)} = \tilde{\mathbf{H}}_k\tilde{\mathbf{A}}$ is full-rank. Lemma 4.5 translates this in conditions on a_1, a_2, and b such that there exists an assignment of the structured $\tilde{\mathbf{A}}$ that makes $\mathbf{G}^{(k)}$ full-rank:

$$b \leq \text{rank}\mathbf{H}_k\mathbf{V}_{B_\Phi} \tag{49}$$
$$a_1 \leq \text{rank}\mathbf{H}_k[\mathbf{V}_{B_2}|\mathbf{V}_{B_\Phi}] \tag{50}$$
$$a_2 \leq \text{rank}\mathbf{H}_k[\mathbf{V}_{B_1}|\mathbf{V}_{B_\Phi}] \tag{51}$$
$$a_1 + a_2 - b \leq \text{rank}\mathbf{H}_k[\mathbf{V}_{B_1}|\mathbf{V}_{B_2}|\mathbf{V}_{B_\Phi}] \tag{52}$$
$$R_1 + R_2 \leq \text{rank}\mathbf{H}_k[\mathbf{V}_{B_{12}}|\mathbf{V}_{B_1}|\mathbf{V}_{B_2}|\mathbf{V}_{B_\Phi}] \tag{53}$$

We then have the following lemma.

Lemma 4.8 *For $k \in \mathcal{P}$, there exists an assignment of $\tilde{\mathbf{A}}$, such that $\mathbf{G}^{(k)}$ is full-rank if*

$$b \leq r_{12k} - r_{12} \tag{54}$$
$$a_1 \leq r_{1k} - r_1 \tag{55}$$
$$a_2 \leq r_{2k} - r_2 \tag{56}$$
$$a_1 + a_2 - b \leq \max \left\{ \begin{array}{c} r_{1k} - r_1, \\ r_{2k} - r_2, \\ r_k - r_1 - r_2 + r_{12} \end{array} \right\} \tag{57}$$
$$R_1 + R_2 \leq r_k. \tag{58}$$

So the questions of interest become:
(*i*) Whether there exists a_1, a_2, and b such that they satisfy the structural constraints

$$a_1 - b, a_2 - b, b \geq 0 \tag{59}$$
$$a_1 + a_2 - b \geq R_2 \tag{60}$$

along with the requirement (48) for all $k \in \{1, 2\}$, and requirements (54) to (58), for all $k \in \{3, \cdots, K\}$.
(*ii*) If such a universal tuple (a_1, a_2, b) exists, whether an assignment of $\tilde{\mathbf{A}}$ within the structure of (46) exists such that all $\mathbf{G}^{(k)}$ are full-rank simultaneously for all $k \in \mathcal{U}$.

Item (*i*) can be answered in different manners. One way is to solve a feasibility problem on a_1, a_2, b; one can show that there always exists an integer tuple (a_1, a_2, b) satisfying all the requirements in (*i*). This method, nevertheless, does not necessarily give us a unique good choice of a_1, a_2, b for all rate pairs of the region $\mathcal{R}$.

We are instead proposing to use the specific universal choice:

$$\begin{aligned} a_1 &= (R_1 + R_2 - r_1)^+ \\ a_2 &= (R_1 + R_2 - r_2)^+ \\ b &= (a_1 + a_2 - R_2)^+ \end{aligned} \tag{61}$$

To show that this is a valid choice, it is sufficient to prove the achievability for the rates on the facet $2R_1 + R_2 = \min_{i \in \mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}$ when[1] $r_1 + \min_{i \in \mathcal{P}} r_i \geq \min_{i \in \mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}$ (i.e., when this facet exists) and otherwise, when $r_1 + \min_{i \in \mathcal{P}} r_i \leq \min_{i \in \mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}$, for the rate pair $(r_1, \min_{i \in \mathcal{P}} r_i - r_1)$. (We assume without loss of generality that $r_1 \leq r_2$.) It is sufficient to do so, because, for the choice of values that we make in (61), the rest of the rate pairs

[1]Here we have for notational convenience assumed $r_1 \leq r_2$.

in $\mathcal{R}$ will be "redundant". By this, we mean that they are either dominated by the rate pairs we study, or can be achieved from them by a rate transfer. To be more specific, consider the rate pairs (R_1, R_2) on the facet $R_1 + R_2 = \min_{i\in\mathcal{P}} r_i$. We argue here that all those rate pairs can be achieved as the corner point (R_1^B, R_2^B) is achieved (see Fig. 2), simply by a rate transfer of an amount of $R_2^B - R_2$ part of the private message to the common message. This is possible because, for the choice of values in (61), both the first two receivers decode enough variables from W_2 for the rate transfer to occur.

We show in the following that a_1, a_2, and b selected as in (61) satisfy all the requirements mentioned in (i) for the non-redundant rate pairs discussed. Clearly, the structural constraints are satisfied by definition. (48) also holds for $k = 1, 2$. (54) holds for all $k \in \{3, \cdots, K\}$ by positivity of $r_{12k} - r_{12}$ and by the characterization (12) of the rate region $\mathcal{R}$. (55) and (56) hold for all $k \in \{3, \cdots, K\}$ by positivity of $r_{1k} - r_1$ and $r_{2k} - r_2$ and (11) characterization of $\mathcal{R}$. (58) being true for all $k \in \{3, \cdots, K\}$ is also a result of (11) characterization of $\mathcal{R}$. (57) holds for the non-redundant pairs under study as follows. We first present the case where $r_1 + \min_{i\in\mathcal{P}} r_i \geq \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}$.

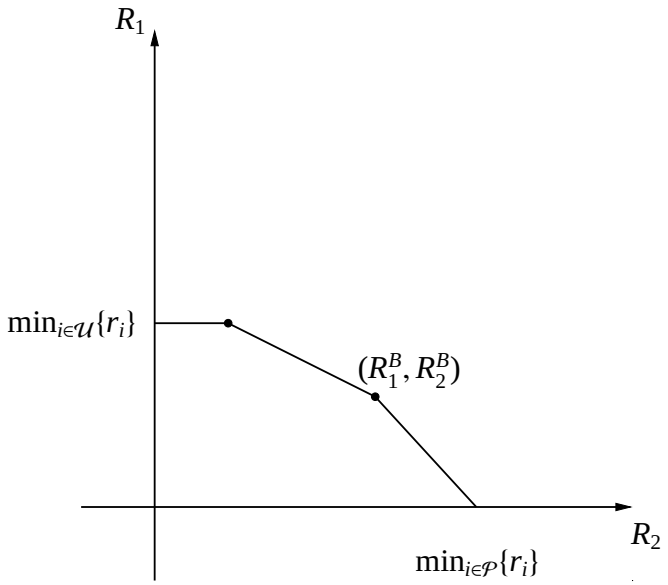
$$\begin{aligned}
a_1 + a_2 - b &= \min\{R_2, a_1 + a_2\} && (62)\\
&\leq R_2 && (63)\\
&\overset{(a)}{=} 2R_1 + 2R_2 - \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\} && (64)\\
&\leq r_k + \min_{i\in\mathcal{P}} r_i \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\} && (65)\\
&\overset{(b)}{\leq} r_k - r_1 - r_2 + r_{12}. && (66)
\end{aligned}$$

Step (a) follows by the assumption that the rate pairs (R_1, R_2) are on the facet of $2R_1 + R_2 = \min_{i\in\mathcal{P}} r_1 + r_2 + r_{12i} - r_{12}$ and step (b) follows from $\min_{i\in\mathcal{P}} r_i \leq \min_{i\in\mathcal{P}} r_{12i}$. Similar arguments hold for the other case when $r_1 + \min_{i\in\mathcal{P}} r_i \leq \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}$, namely

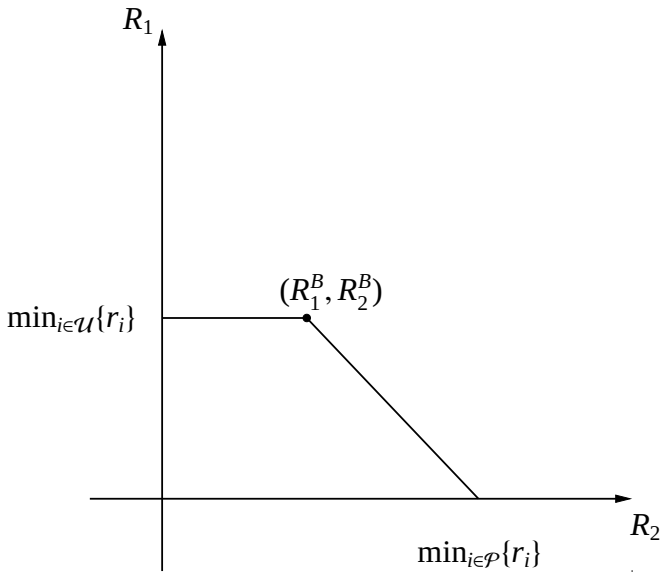
$$\begin{aligned}
a_1 + a_2 - b &= \min\{R_2, a_1 + a_2\} && (67)\\
&\leq R_2 && (68)\\
&\overset{(a)}{=} R_1 + R_2 - r_1 && (69)\\
&\leq r_k - r_1 && (70)\\
&\leq r_{1k} - r_1. && (71)
\end{aligned}$$

Step (a) follows by the assumption of the non-redundant rate pair (R_1, R_2) being $(r_1, \min_{i\in\mathcal{P}} r_i - r_1)$ in this case. Finally, (58) holds as a result of characterization (11) of $\mathcal{R}$.

Now that we proved such a universal tuple (a_1, a_2, b) exists, we answer (ii) by showing that an assignment of $\tilde{\mathbf{A}}$ within the structure of (46) exists such that



(a) $r_1 + \min_{i\in\mathcal{P}} r_i \geq \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}$



(b) $r_1 + \min_{i\in\mathcal{P}} r_i \leq \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}$

Figure 2: Rate region $\mathcal{R}$

all $\mathbf{G}^{(k)}$ are full-rank simultaneously for all $k \in \mathcal{U}$; i.e., an assignment of $\tilde{\mathbf{A}}$ such that linearly encoding the messages W_1, and W_2 with it lets all receivers $k \in \{1, 2\}$ decode W_1 and all receivers $k \in \mathcal{P}$ decode W_1 and W_2.

We will use the sparse zeros lemma to this end. From lemma 4.7 and 4.8, we have shown that for each $k \in \mathcal{U}$, there exists an assignment of $\tilde{\mathbf{A}}$ in the structure of (46) with (a_1, a_2, b) of (61) such that $\mathbf{G}^{(k)}$ is full-rank. This implies that there exists a full rank square submatrix of $\mathbf{G}^{(k)}$, say $\mathbf{G}_s^{(k)}$. Let $\mathcal{P}^{(k)}$ be the polynomial corresponding to the determinant of $\mathbf{G}_s^{(k)}$, and $\mathcal{G} = \prod_k \mathcal{P}^{(k)}$. Given that there exists an assignment for the variables such that each individual polynomial $\mathcal{P}^{(k)}$ is nonzero, we can conclude from the sparse zero lemma that there exists an assignment of $\tilde{A}$ over any field $\mathbf{F}'$, $|\mathbf{F}'|$ being larger than the maximum degree of $\mathcal{G}$ in its variables, such that all polynomials are simultaneously nonzero. With this assignment, all users can simultaneously receive their required messages.

The following lemma, provides an upper bound on the required size for $\mathbb{F}$. Note that operation over smaller fields is also possible, by using vector coding.

Lemma 4.9 *The two-message set problem with K receivers has always a solution over a field of size* $|\mathbb{F}| > K$.

Proof What should be proved here is that the maximum degree of the polynomial $\mathcal{G} = \prod_k \mathcal{P}^{(k)}$ is at most K. We first bound the degree of each of $\mathcal{P}^{(k)}$ by 1 and then conclude the proof.
(*i*) Consider a matrix $\mathbf{G}_{l\times l}$ that has $l \times l$ independent variables $g_{i,j}$. Call $\sqrt{\ } = \det \mathbf{G}$.Obviously, degree of $\mathcal{P}$ in $g_{i,j}$ is at most 1.
(*ii*)Consider a matrix $\mathbf{G}_{l\times l} = \mathbf{T}_{l\times m}\mathbf{A}_{m\times l}$, where $\mathbf{T}$ is a fixed matrix and $\mathbf{A}_{m\times l}$ composed of $m \times l$ independent variables $a_{i,j}$. Then each $g_{i,j} = \sum_l t_{i,l}a_{l,j}$. Using the Laplace expansion to calculate $\mathcal{P} = \det \mathbf{G}$, we have

$$\mathcal{P} = \sum_i (-1)^{i+j} g_{i,j} \det \mathbf{G}_{i,j}. \tag{72}$$

where $\det \mathbf{G}_{i,j}$ is not a function of $a_{i,j}$ since $a_{i,j}$ shows up only in column j of $\mathbf{G}$. Thus again, degree of $\det \mathbf{G}$ in $a_{i,j}$ is at most 1. $G^{(1)}$ and $G^{(2)}$ are of the form (*i*) and all the other $G^{(k)}$'s are of the form (*ii*):

- $k \in \{1, 2\}$:
  $\mathbf{G}^{(k)}$ is basically a submatrix of $\tilde{\mathbf{A}}$, and has thus all independent variables $\tilde{a}_{i,j}$. From (*i*) degree of $\mathcal{P}^{(k)}$ in $\tilde{a}_{i,j}$ is at most 1.

- $k \in \mathcal{P}$:
  $\mathbf{G}^{(k)} = \tilde{\mathbf{H}}_k \tilde{A}$ and $\tilde{\mathbf{A}}$ has all independent variables $\tilde{a}_{i,j}$. From (*ii*), degree of $\mathcal{P}^{(k)}$ in $\tilde{a}_{i,j}$ is at most 1.

Constructing $\mathcal{G} = \prod_{k\in\mathcal{U}} \mathcal{P}^{(k)}$, where each $\mathcal{P}^{(k)}$ is of degree at most 1 in $\tilde{a}_{i,j}$ results in the degree of $\mathcal{G}$ being at most K in each $\tilde{a}_{i,j}$ and this concludes the proof. ■

4.4 Example

We conclude with an example that illustrates our code design for a specific instantiation. Consider the following channel matrices.

$$\mathbf{H}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \tag{73}$$

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{74}$$

$$\mathbf{H}_3 = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{75}$$

$$\mathbf{H}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \tag{76}$$

$$\tag{77}$$

One could verify that the rate region $\mathcal{R}$ for this example would be characterized by

$$R_1, R_2 \geq 0 \tag{78}$$

$$R_1 \leq 2 \tag{79}$$

$$R_1 + R_2 \leq 3 \tag{80}$$

$$2R_1 + R_2 \leq 4 \tag{81}$$

In this example, we design $\tilde{\mathbf{A}}$ to achieve the corner point $(R_1^B, R_2^B) = (1, 2)$.

Basis $\mathcal{B}$ is formed by $\mathcal{B}_\Phi = \left\{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}\right\}$, $\mathcal{B}_2 = \left\{\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}\right\}$, $\mathcal{B}_1 = \left\{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}\right\}$, and $\mathcal{B}_{12} = \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right\}$ as explained in 4.1. This gives

$$\mathbf{V} = \left[\begin{array}{c|c|c|c} \mathbf{V}_{12} & \mathbf{V}_2 & \mathbf{V}_1 & \mathbf{V}_\Phi \end{array}\right] = \left[\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Our code design, using the size parameters $(a_1, a_2, b) = (1, 1, 0)$ from 61 makes matrix $\tilde{\mathbf{A}}$ to be strutured as

$$\tilde{\mathbf{A}} = \begin{bmatrix} x_1 & 0 & 0 \\ x_2 & 0 & x_5 \\ x_3 & x_4 & 0 \\ x_6 & x_7 & x_8 \end{bmatrix}. \tag{82}$$

Furthermore, $\tilde{Y}_k$, $k \in \{1, 2, 3, 4\}$ from which each receiver k decodes its required message is as follows:

$$\tilde{Y}_1 = \tilde{\mathbf{H}}_1 \tilde{\mathbf{A}} W = \begin{bmatrix} x_1 & 0 & 0 \\ x_3 & x_4 & 0 \end{bmatrix} \begin{bmatrix} w_{1,1} \\ w_{2,1} \\ w_{2,2} \end{bmatrix} \tag{83}$$

$$\tilde{Y}_2 = \tilde{\mathbf{H}}_2 \tilde{\mathbf{A}} W = \begin{bmatrix} x_1 & 0 & 0 \\ x_2 & 0 & x_5 \end{bmatrix} \begin{bmatrix} w_{1,1} \\ w_{2,1} \\ w_{2,2} \end{bmatrix} \tag{84}$$

$$Y_3 = \tilde{\mathbf{H}}_3 \tilde{\mathbf{A}} W = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & 0 & 0 \\ x_2 & 0 & x_5 \\ x_3 & x_4 & 0 \\ x_6 & x_7 & x_8 \end{bmatrix} \begin{bmatrix} w_{1,1} \\ w_{2,1} \\ w_{2,2} \end{bmatrix} \tag{85}$$

$$Y_4 = \tilde{\mathbf{H}}_4 \tilde{\mathbf{A}} W = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 & 0 & 0 \\ x_2 & 0 & x_5 \\ x_3 & x_4 & 0 \\ x_6 & x_7 & x_8 \end{bmatrix} \begin{bmatrix} w_{1,1} \\ w_{2,1} \\ w_{2,2} \end{bmatrix}. \tag{86}$$

One can readily verify that $\mathbf{G}^{(3)}$ is individually full-rank for $x_1 = 1, x_4 = 1, x_8 = 1$, and the remaining x_i's being zero, and $\mathbf{G}^{(4)}$ is individually full-rank for $x_1 = 1, x_4 = 1, x_5 = 1$, and the remaining x_i's being zero. One should note that either of x_4 and x_5 being zero makes it impossible for receiver 4 to decode W_2. Finally, the following $\tilde{\mathbf{A}}$ allows receivers 1 and 2 to decode W_1 and receivers 3 and 4 to decode W_1 and W_2:

$$\tilde{\mathbf{A}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{87}$$

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] S. Avestimehr, S. Diggavi, and D N C. Tse. Wireless network information flow: A deterministic approach submitted to IEEE Transactions on Information Theory, July 2009, available from ArXiv at http://arxiv.org/abs/0906.5394

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

[4] G. Bresler and D N C. Tse, "The two-user Gaussian interference channel: a deterministic view", European Transactions on Telecommunications, vol 19, issue 4, pp. 333-354, June, 2008.

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

[6] C. Fragouli and E. Soljanin, "Network Coding Fundamentals", , Monograph in Series, Foundations and Trends in Networking, 2007.

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

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

[9] S. Mohajer, S N. Diggavi and D. Tse, "Approximate Capacity of a Class of Gaussian Relay-Interference Networks," IEEE International Symposium on Information Theory, Seoul, Korea, June 2009.

[10] C. Nair and A. El Gamal, "The Capacity Region of a Class of 3-Receiver Broadcast Channels with Degraded Message Sets", Proceedings of the International Symposium on Information Theory, pp. 1706-1710, Toronto, June 2008.

[11] V. Prabhakaran, S. Diggavi, and D. Tse, "Broadcasting with degraded message sets: A deterministic approach," Proceedings of the 45th Annual Allerton Conference on Communication, Control and Computing, 2007.