On degraded two message set broadcasting

Shirin Saeedi
EPFL,
Switzerland
Email: shirin.saeedi@epfl.ch

Suhas Diggavi
EPFL,
Switzerland
suhas.diggavi@epfl.ch

Christina Fragouli
EPFL,
Switzerland
christina.fragouli@epfl.ch

Vinod Prabhakaran
University of Illinois,
Urbana-Champaign
vinodmp@uiuc.edu

***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.**

I. 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 information 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 [4]. The solution for the case where there is a degradation order between the users' channels was given in [3], [6]. The problem of general two-user broadcast channel with a degraded two message set requirement was solved by Korner and Marton in [7]. However there is comparatively 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 [9].

The linear deterministic channel model, introduced in [1], was motivated by its intimate connection to linear Gaussian models [2], [8]. Recently [10] solved a three-user, degraded three (nested) message set problem over linear deterministic broadcast channels. This paper builds on these results for 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 also shows an intimate connection between our problem and network coding techniques, 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 [9].

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

II. Problem Formulation and Results

A. 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 [10]. 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 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 $\mathcal{N}_{i_1,\cdots i_{|\mathcal{S}|}}$.

B. Main Result

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

$$R_1 \leq \min_{i \in \mathcal{U}} r_i \tag{3}$$

$$R_1 + R_2 \leq \min_{i \in \mathcal{P}} r_i \tag{4}$$

$$2R_1 + R_2 \leq \min_{i \in \mathcal{P}} \{r_1 + r_2 + r_{12i} - r_{12}\}, \tag{5}$$

where the size of $\mathbb{F}$ is larger than K. The rates given above are expressed in $\log_{|\mathbb{F}|}(\cdot)$. ■

III. 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

$$R_1 \leq \min_{i \in \mathcal{U}} r_i \tag{6}$$

$$R_1 + R_2 \leq \min_{i \in \mathcal{P}} r_i \tag{7}$$

$$\forall k \leq |\mathcal{P}^c| : \quad kR_1 + R_2 \leq \min_{i \in \mathcal{P}, j_1, \cdots j_k \notin \mathcal{P}^c} \{\sum_{l=1}^{k} r_{j_l} + r_{j_1, j_2, \cdots, j_k, i} - r_{j_1, j_2, \cdots, j_k}\} \tag{8}$$

Proof: Assume communication using blocks of an arbitrary length n, and denote the received signal at each receiver i by Y_i^n. Then (3) and (4) follow from:

$$\forall i \in \mathcal{U} : \; n(R_1) \leq I(W_1; Y_i^n) \leq H(Y_i^n) - H(Y_i^n | W_1) \leq n r_i. \tag{9}$$

$$\begin{aligned} \forall i \in \mathcal{P} : \quad n(R_1 + R_2) &\leq I(W_1, W_2; Y_i^n) && (10)\\ &\leq H(Y_i^n) - H(Y_i^n | W_1, W_2) && (11)\\ &\leq n r_i. && (12) \end{aligned}$$

From (9), it follows that

$$H(Y_i^n | W_1) \leq n(r_i - R_1). \tag{13}$$

To obtain (5), we use the approach in [10]. 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 (13) and upper bounding $H(Y_i^n | Y_{j_1}^n, \cdots, Y_{j_k}^n)$ by $n(r_{j_1, j_2, \cdots, j_k, i} - r_{j_1, j_2, \cdots, j_k})$ as in [10]. ■

IV. 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, while the second receiver w_1 and w_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 zeros lemma [5].

The technical steps can be described as follows:

- We will design in section IV-A a new 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 $\tilde{\mathbf{A}}$ of indeterminates, which we will attempt to fill (complete) so that the decoding requirements are fulfilled. The basis is chosen such that so 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 $\tilde{\mathbf{A}}$. The matrix completion problem is to fill $\tilde{\mathbf{A}}$ appropriately.
- In section IV-B, 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 impose in section IV-C, a specific structure to $\tilde{\mathbf{A}}$, parametrized by structure parameters a_1, a_2, and b. We will show that there exists a universal choice for the parameters a_1, a_2, and b that allows to satisfy the decodability requirements for each receiver separately. We will then apply the sparse zeros lemma (see for example [5]) to show that there exist variable choices that satisfy all the decodability conditions simultaneously.

A. Problem Reduction

We choose a 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

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

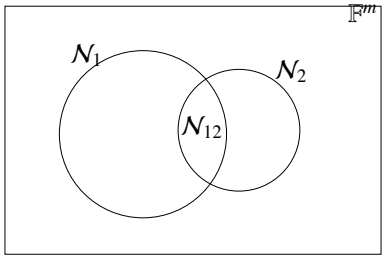

Fig. 1: Venn diagram of the null spaces of the 2 receivers requiring only W_1.

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

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

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

$$\tilde{Y}_1 = \left[\begin{array}{c} \tilde{X}_{12} \\ \tilde{X}_1 \end{array}\right] = \left[\begin{array}{cccc} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{array}\right]\tilde{X} =: \tilde{\mathbf{H}}_1\tilde{X}. \tag{14}$$

Similarly,

$$\tilde{Y}_2 = \left[\begin{array}{c} \tilde{X}_{12} \\ \tilde{X}_2 \end{array}\right] = \left[\begin{array}{cccc} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \end{array}\right]\tilde{X} =: \tilde{\mathbf{H}}_2\tilde{X}. \tag{15}$$

For the rest of the users, we simply set

$$\tilde{Y}_k = Y_k = \mathbf{H}_k\mathbf{V}\tilde{X} =: \tilde{\mathbf{H}}_k\tilde{X}, \quad k \in \mathcal{P} = 3, 4, \ldots, K, \tag{16}$$

where

$$\begin{aligned}\tilde{\mathbf{H}}_k &= \left[\begin{array}{c|c|c|c} \mathbf{H}_k\mathbf{V}_{12} & \mathbf{H}_k\mathbf{V}_2 & \mathbf{H}_k\mathbf{V}_1 & \mathbf{H}_k\mathbf{V}_\Phi \end{array}\right]\\ &=: \left[\begin{array}{c|c|c|c} \tilde{\mathbf{H}}_k^{12} & \tilde{\mathbf{H}}_k^2 & \tilde{\mathbf{H}}_k^1 & \tilde{\mathbf{H}}_k^\Phi \end{array}\right].\end{aligned} \tag{17}$$

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

$$\tilde{Y}_i = \tilde{\mathbf{H}}_i\tilde{X}, \quad i \in \mathcal{U}, \tag{18}$$

where $\tilde{\mathbf{H}}_i$ are given by (14)-(16). The following lemma calculates the ranks of certain submatrices of $\tilde{\mathbf{H}}_i$ and will be used in IV-C to prove the achievability of our coding theorem.

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

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

The proof is provided in [11].

B. 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. The proofs are provided in [11].

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 $\mathbf{G}$.

Lemma 4.3: Consider a matrix $\mathbf{B} = \left[\begin{array}{c|c} \mathbf{B}_1 & \mathbf{B}_2 \end{array}\right]$, 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 $\mathbf{G} = \left[\begin{array}{c|c} \mathbf{B}_1 & \mathbf{L}_1 \end{array}\right]$, 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 $l = \operatorname{rank}(\mathbf{B}_2) \leq n - d$, 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 $\mathbf{B}$.

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

$$\begin{aligned}\mathbf{B}W &= \left[\begin{array}{c|c} \mathbf{B}_1 & \mathbf{L}_1\mathbf{L}_2 \end{array}\right]W &(19)\\ &= \left[\begin{array}{c|c} \mathbf{B}_1 & \mathbf{L}_1 \end{array}\right]\left[\begin{array}{c} W_1^d \\ \mathbf{L}_2 W_{d+1}^m \end{array}\right] &(20)\\ &= \mathbf{G}\left[\begin{array}{c} W_1^d \\ \mathbf{L}_2 W_{d+1}^m \end{array}\right]. &(21)\end{aligned}$$

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

$$\mathbf{T} = \left[\begin{array}{c|c|c|c} \mathbf{T}_1 & \mathbf{T}_2 & \mathbf{T}_3 & \mathbf{T}_4 \end{array}\right]. \tag{22}$$

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

$$\text{rank}\left(\begin{array}{c} \mathbf{T}_4 \end{array}\right) \geq t_4, \tag{23}$$

$$\text{rank}\left(\left[\begin{array}{c|c} \mathbf{T}_3 & \mathbf{T}_4 \end{array}\right]\right) \geq t_3 + t_4, \tag{24}$$

$$\text{rank}\left(\left[\begin{array}{c|c} \mathbf{T}_2 & \mathbf{T}_4 \end{array}\right]\right) \geq t_2 + t_4, \tag{25}$$

$$\text{rank}\left(\left[\begin{array}{c|c|c} \mathbf{T}_2 & \mathbf{T}_3 & \mathbf{T}_4 \end{array}\right]\right) \geq t_2 + t_3 + t_4, \text{ and} \tag{26}$$

$$\text{rank}\left(\left[\begin{array}{c|c|c|c} \mathbf{T}_1 & \mathbf{T}_2 & \mathbf{T}_3 & \mathbf{T}_4 \end{array}\right]\right) \geq t_1 + t_2 + t_3 + t_4. \tag{27}$$

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

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

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

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

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*

- $t_4 \leq \text{rank}\left(\begin{array}{c} \mathbf{T}_4 \end{array}\right)$,
- $t_2 + t_4 \leq \text{rank}\left(\left[\begin{array}{c|c} \mathbf{T}_3 & \mathbf{T}_4 \end{array}\right]\right)$,
- $t_3 + t_4 \leq \text{rank}\left(\left[\begin{array}{c|c} \mathbf{T}_2 & \mathbf{T}_4 \end{array}\right]\right)$,
- $t_2 + t_3 + t_4 \leq \text{rank}\left(\left[\begin{array}{c|c|c} \mathbf{T}_2 & \mathbf{T}_3 & \mathbf{T}_4 \end{array}\right]\right)$,
- $t_1 + t_2 + t_3 + t_4 \leq \text{rank}\left(\left[\begin{array}{c|c|c|c} \mathbf{T}_1 & \mathbf{T}_2 & \mathbf{T}_3 & \mathbf{T}_4 \end{array}\right]\right)$.

C. Structured Linear Code

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

$$\tilde{X} = \tilde{\mathbf{A}} W, \tag{28}$$

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 $\tilde{\mathbf{A}}$

$$\tilde{\mathbf{A}} = \begin{array}{c} \overset{a_1-b}{\longleftrightarrow} \\ \left[\begin{array}{c|c|c|c} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline & \mathbf{0} & & \mathbf{0} \\ \hline & & \mathbf{0} & \mathbf{0} \\ \hline & & & \end{array}\right] \end{array} \begin{array}{l} \overset{a_2-b}{\longleftrightarrow} \\ \overset{b}{\longleftrightarrow} \\ \updownarrow |\mathcal{B}_{12}| \\ \updownarrow |\mathcal{B}_2| \\ \updownarrow |\mathcal{B}_1| \\ \updownarrow |\mathcal{B}_\Phi| \end{array} \tag{29}$$

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 (18), 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

$$\tilde{\mathbf{H}}_k \tilde{\mathbf{A}} = \left[\begin{array}{c|c} \mathbf{B}_1^{(k)} & \mathbf{B}_2^{(k)} \end{array}\right], \tag{30}$$

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 $\mathbf{L}_1^{(k)} \in \mathbb{F}^{r_k \times l_l^k}$. Note that given the structure (14) and (15) of $\tilde{\mathbf{H}}_k$ and the structure (29) 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 (3) 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 $k \in \{1, 2\}$

$$a_k \geq R_1 + R_2 - r_k, \tag{31}$$

receiver k can decode W_1 *if* $\mathbf{G}^{(k)}$ as defined above is full-rank. The proof of the following lemmas is provided in [11].

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.

From (18), receiver $k \in \mathcal{P}$ can decode W_1 and W_2 from $\tilde{Y}_k$, if $\mathbf{G}^{(k)} = \tilde{\mathbf{H}}_k \tilde{\mathbf{A}}$ is full-rank. Lemma 4.5 translates the existence of an assignment of $\tilde{\mathbf{A}}$ that makes $\mathbf{G}^{(k)}$ full-rank in conditions on a_1, a_2, and b in terms of $\text{rank}\left(\tilde{\mathbf{H}}_k^\Phi\right)$, $\text{rank}\left(\left[\begin{array}{c|c} \tilde{\mathbf{H}}_k^1 & \tilde{\mathbf{H}}_k^\Phi \end{array}\right]\right)$, $\text{rank}\left(\left[\begin{array}{c|c} \tilde{\mathbf{H}}_k^2 & \tilde{\mathbf{H}}_k^\Phi \end{array}\right]\right)$, $\text{rank}\left(\left[\begin{array}{c|c|c} \tilde{\mathbf{H}}_k^2 & \tilde{\mathbf{H}}_k^1 & \tilde{\mathbf{H}}_k^\Phi \end{array}\right]\right)$ and $\text{rank}\left(\left[\begin{array}{c|c|c|c} \tilde{\mathbf{H}}_k^{12} & \tilde{\mathbf{H}}_k^2 & \tilde{\mathbf{H}}_k^1 & \tilde{\mathbf{H}}_k^\Phi \end{array}\right]\right)$. Applying lemma 4.1 to the ranks of these submatrices of $\tilde{\mathbf{H}}_k$, we 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*

$$b \leq r_{12k} - r_{12} \tag{32}$$

$$a_1 \leq r_{1k} - r_1 \tag{33}$$

$$a_2 \leq r_{2k} - r_2 \tag{34}$$

$$a_1 + a_2 - b \leq \max\left\{\begin{array}{c} r_{1k} - r_1, \\ r_{2k} - r_2, \\ r_k - r_1 - r_2 + r_{12} \end{array}\right\} \tag{35}$$

$$R_1 + R_2 \leq r_k. \tag{36}$$

So the question of interest becomes if there exists a universal choice of a_1, a_2, and b such that they satisfy the structural

constraints

$$a_1 - b,\ a_2 - b,\ b \geq 0, \tag{37}$$

$$a_1 + a_2 - b \geq R_2, \tag{38}$$

along with the requirement (31) for all $k \in \{1,2\}$, and requirements (32) to (36), for all $k \in \mathcal{P}$.
We provide the universal choice:

$$\begin{aligned} a_1 &= (R_1 + R_2 - r_1)^+, \\ a_2 &= (R_1 + R_2 - r_2)^+, \\ b &= (a_1 + a_2 - R_2)^+. \end{aligned} \tag{39}$$

To show that this is a valid choice, we assume without loss of generality that $r_1 \leq r_2$ and argue in [11] that 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 $r_1 + \min_{i\in\mathcal{P}} r_i \leq \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}$, on the rate pair $(r_1, \min_{i\in\mathcal{P}} r_i - r_1)$. It is sufficient to do so, because, for the choice of values that we make in (39), the rest of the rate pairs 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.

We show in the following that a_1, a_2, and b selected as in (39) satisfy all the requirements mentioned in (i) for the non-redundant rate pairs discussed. Clearly, the structural constraints are satisfied by definition. (31) also holds for $k = 1, 2$. (32) holds for all $k \in \mathcal{P}$ by positivity of $r_{12k} - r_{12}$ and by the characterization (5) of the rate region $\mathcal{R}$. (33) and (34) hold for all $k \in \mathcal{P}$ by positivity of $r_{1k} - r_1$ and $r_{2k} - r_2$ and characterization (4) of $\mathcal{R}$. (35) holds for the non-redundant pairs under study as follows. We first present the case where $r_1 + \min_{i\in\mathcal{P}} r_i \geq \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\}$.

$$\begin{aligned} a_1 + a_2 - b &= \min\{R_2, a_1 + a_2\} && (40)\\ &\leq R_2 && (41)\\ &\overset{(a)}{=} 2R_1 + 2R_2 - \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\} && (42)\\ &\leq r_k + \min_{i\in\mathcal{P}} r_i \min_{i\in\mathcal{P}}\{r_1 + r_2 + r_{12i} - r_{12}\} && (43)\\ &\overset{(b)}{\leq} r_k - r_1 - r_2 + r_{12}. && (44) \end{aligned}$$

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

$$\begin{aligned} a_1 + a_2 - b &= \min\{R_2, a_1 + a_2\} && (45)\\ &\leq R_2 && (46)\\ &\overset{(a)}{=} R_1 + R_2 - r_1 && (47)\\ &\leq r_k - r_1 && (48)\\ &\leq r_{1k} - r_1. && (49) \end{aligned}$$

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

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, (36) holds as a result of characterization (4) of $\mathcal{R}$.

Now that we proved such a universal tuple (a_1, a_2, b) exists, we show that an assignment of $\tilde{\mathbf{A}}$ within the structure of (29) exists such that 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 (29) with (a_1, a_2, b) of (39) 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 such that all polynomials are simultaneously nonzero. With this assignment, all users can simultaneously receive their required messages.

The following lemma, proved in [11], 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 $|\mathbb{F}| > K$.

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

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

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

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

[8] 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.

[9] 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.

[10] 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.

[11] S. Saeedi, S. Diggavi, C. Fragouli, and V. Prabhakaran, "On degraded two message set broadcasting," EPFL technical report, August 2009, available from http://infoscience.epfl.ch.