

Capacity Theorems for Broadcast Channels With Two Channel State Components Known at the Receivers

Hyeji Kim and Abbas El Gamal, *Fellow, IEEE*

Abstract—We establish the capacity region of several classes of broadcast channels with random state in which the channel to each user is selected from two possible channel state components and the state is known only at the receivers. When the channel components are deterministic, we show that the capacity region is achieved via Marton coding. This channel model does not belong to any class of broadcast channels for which the capacity region was previously known and is useful in studying wireless communication channels when the fading state is known only at the receivers. We then establish the capacity region when the channel components are ordered, e.g., degraded. In particular, we show that the capacity region for the broadcast channel with degraded Gaussian vector channel components is attained via Gaussian input distribution. Finally, we extend the results on ordered channels to two broadcast channel examples with more than two channel components, but show that these extensions do not hold in general.

Index Terms—Broadcast channel, Broadcast channel with state, superposition coding, Marton's inner bound, Nair—El Gamal outer bound.

I. INTRODUCTION

CONSIDER the discrete memoryless broadcast channel (DM-BC) with random (IID) state $(\mathcal{X} \times \mathcal{S}, p(y_1, y_2|x, s)p(s), \mathcal{Y}_1 \times \mathcal{Y}_2)$ with the state S known only at the receivers. Assume the setup in which the sender wishes to transmit a common message $M_0 \in [1 : 2^{nR_0}]$ to both receivers and private messages $M_j \in [1 : 2^{nR_j}]$ to receiver $j \in \{1, 2\}$ as depicted in Figure 1.

It is well known that this broadcast channel with state setup can be viewed as a general DM-BC with input X and outputs (Y_1, S) and (Y_2, S) . Hence the definitions of a $(2^{nR_0}, 2^{nR_1}, 2^{nR_2})$ code, achievability and the capacity region \mathcal{C} are the same as for the general broadcast channels [1]. Moreover, the capacity region for this broadcast channel with state setup is not known in general. The Marton inner bound and the UV outer bound on the general broadcast channel hold for this channel and they coincide when the channel

$X \rightarrow (Y_1, S), (Y_2, S)$ falls into any of the classes of the broadcast channel for which the capacity region is known (see [2] for examples of these classes). Beyond these classes, there have been some efforts on evaluating inner bounds on the capacity region of the Gaussian fading BC model, including superposition coding by Jafarian and Vishwanath [3], time division with power control by Liang and Goldsmith [4], and superposition of binary inputs motivated by a capacity achieving strategy for a layered erasure broadcast channel by Tse and Yates [5].

In this paper we focus on the special class of the broadcast channel with state in Figure 1 in which there are only two channel state components as depicted in Figure 2. In this model, which we refer to as the *broadcast channel with two channel states* (BC-TCS), the state $S = (S_1, S_2) \in \{1, 2\}^2$ with $p_{S_1}(1) = p_1, p_{S_1}(2) = 1 - p_1 = \bar{p}_1$ and $p_{S_2}(1) = p_2, p_{S_2}(2) = \bar{p}_2$, and the two possible channel components are denoted by $\tilde{Y}_1 \sim p(\tilde{y}_1|x)$ and $\tilde{Y}_2 \sim p(\tilde{y}_2|x)$. The outputs of the BC-TCS is

$$\begin{aligned} Y_1 &= \begin{cases} \tilde{Y}_1 & \text{if } S_1 = 1, \\ \tilde{Y}_2 & \text{if } S_1 = 2, \end{cases} \\ Y_2 &= \begin{cases} \tilde{Y}_1 & \text{if } S_2 = 1, \\ \tilde{Y}_2 & \text{if } S_2 = 2. \end{cases} \end{aligned} \quad (1)$$

Without loss of generality, we assume throughout that $p_1 \geq p_2$ and that receiver $j = 1, 2$ knows the state sequence S^n but the sender does not.¹

In [6], we established the private message capacity region for the special case of the BC-TCS in which the state components are deterministic functions, i.e., $\tilde{Y}_1 = f_1(X)$ and $\tilde{Y}_2 = f_2(X)$. Achievability is established using Marton coding [7]. The key observation is that the auxiliary random variables in the Marton region characterization, U_1 and U_2 , are always set to f_1, f_2, X , or \emptyset . In particular if the channel $p(y_1|x)$ is more likely to be f_1 than the channel $p(y_2|x)$, then (U_1, U_2) are set to $(X, \emptyset), (\emptyset, X)$, or (f_1, f_2) . The converse is established by showing that the Marton inner bound with these extreme choices of auxiliary random variables coincides with the UV outer bound [8]. It is important to note that this class of broadcast channels with two deterministic channel components (BC-TDCS) does not belong to any class of broadcast channels for which the capacity region is known. It also provides yet

¹Since the capacity region of the broadcast channel depends only on its marginal distributions [1], we only need to specify the marginal pmfs of S_1 and S_2 . Moreover, it suffices to assume that receiver $j = 1, 2$ knows only its state sequence S_j^n .

Manuscript received September 14, 2015; accepted August 30, 2016. Date of publication October 21, 2016; date of current version November 18, 2016. This work was supported by Air Force under Grant FA9550-10-1-0124. This paper was presented at the 2014 Proceedings IEEE International Symposium on Information Theory.

H. Kim was with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA. She is now with the University of Illinois at Urbana-Champaign, Champaign, IL 61820 USA (e-mail: hyejkim@illinois.edu).

A. El Gamal is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: abbas@ee.stanford.edu).

Communicated by T. Liu, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2016.2620159

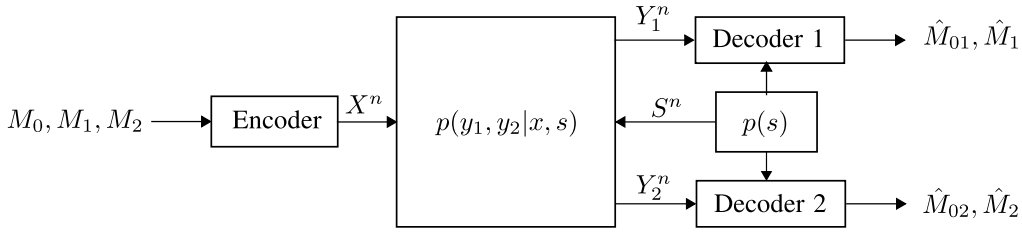
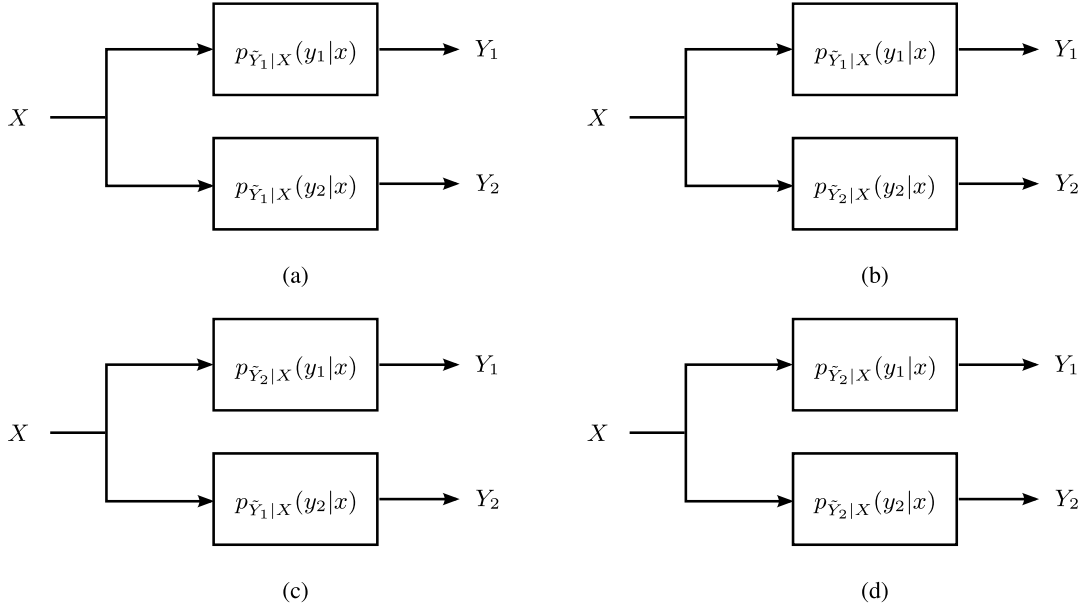


Fig. 1. Broadcast channel with state known only at the receivers

Fig. 2. Broadcast channel with two channel state components. (a) $(S_1, S_2) = (1, 1)$, $p_{S_1}(1) = p_1$, $p_{S_2}(1) = p_2$. (b) $(S_1, S_2) = (1, 2)$, $p_{S_1}(1) = p_1$, $p_{S_2}(2) = \bar{p}_2$. (c) $(S_1, S_2) = (2, 1)$, $p_{S_1}(2) = \bar{p}_1$, $p_{S_2}(1) = p_2$. (d) $(S_1, S_2) = (2, 2)$, $p_{S_1}(2) = \bar{p}_1$, $p_{S_2}(2) = \bar{p}_2$.

another class of broadcast channels for which Marton coding is optimal. Moreover, the BC-TDCS model can be used to approximate certain fading broadcast channels in high SNR (see Example 2 in Section II).

In this paper we provide a complete proof for the result in [6] and extend it to the case with common message (see Section II). In addition, we include several new results on the capacity region of the BC-TCS. In Section III, we study the case when the channel components are ordered, which models, for example, a wireless downlink channel in which the channel to each user can be either “strong” or “weak”. We show that if the BC $p(\tilde{y}_1, \tilde{y}_2|x)$ is degraded, less noisy, more capable, or dominantly c -symmetric, then the corresponding BC-TCS $p(y_1, y_2, s|x)$ is degraded, less noisy, more capable, or dominantly c -symmetric, and the capacity region is achieved via superposition coding. This is surprising (and as we will show does not extend to more than two components in general) because the sender does not know the state, hence does not know which of the two channels $p(y_1, s|x)$ or $p(y_2, s|x)$ is stronger. We further show that the capacity region of the BC-TCS with degraded Gaussian vector channel components, which is a special case of the BC-TCS with degraded channel components, is attained by Gaussian channel input. This is again unexpected because for

the general degraded fading Gaussian BC (where we know that one channel is always a degraded version of the other), the optimizing input distribution is not always Gaussian [9]. In Section IV, we present results on the broadcast channel with more than two channel components. We establish the capacity region when there are three BEC or BSC channel components and show that there is a gap between the Marton inner bound and the UV outer bound when there are four BSC channel components. Hence our results for the two channel state components do not extend to more than two state components in general.

II. DETERMINISTIC CHANNEL STATE COMPONENTS

In this section, we consider the BC-TCS with two deterministic channel components $\tilde{Y}_1 = f_1(X)$ and $\tilde{Y}_2 = f_2(X)$, henceforth referred to as BC-TDCS. We show that the capacity region of the BC-TDCS is achieved using Marton coding.

Theorem 1 (Private Message Capacity Region for BC-TDCS [6]): The private message capacity region of the BC-TDCS $(\mathcal{X} \times \mathcal{S}, p(s)p(y_1, y_2|x, s), \mathcal{Y}_1 \times \mathcal{Y}_2)$ with the state known only at the receivers is

$$\mathcal{C} = \text{co}\{\mathcal{R}_1 \cup \{(C_1, 0)\} \cup \{(0, C_2)\}\}, \quad (2)$$

where $C_j = \max_{p(x)} I(X; Y_j|S)$ for $j = 1, 2$, and

$$\begin{aligned} \mathcal{R}_1 = \{ & (R_1, R_2): R_1 \leq I(f_1; Y_1|S), \\ & R_2 \leq I(f_2; Y_2|S), \\ & R_1 + R_2 \leq I(f_1; Y_1|S) + I(f_2; Y_2|S) - I(f_1; f_2) \\ & \text{for some } p(x) \in \mathcal{P}, \end{aligned}$$

where

$$\begin{aligned} \mathcal{P} = \{ & \arg \max_{p(x)} (I(f_1; Y_1|S) + \lambda(I(f_2; Y_2|S) - I(f_1; f_2))) \\ & \text{for } \bar{p}_1/\bar{p}_2 \leq \lambda \leq 1, \\ & \arg \max_{p(x)} (I(f_1; Y_1|S) - I(f_1; f_2) + \nu I(f_2; Y_2|S)) \\ & \text{for } 1 \leq \nu \leq p_1/p_2 \}. \end{aligned}$$

Proof: For achievability we use Marton coding which achieves the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 & \leq I(U_1; Y_1|S), \\ R_2 & \leq I(U_2; Y_2|S), \\ R_1 + R_2 & \leq I(U_1; Y_1|S) + I(U_2; Y_2|S) - I(U_1; U_2) \end{aligned} \quad (3)$$

for some pmf $p(u_1, u_2, x)$.

Note that the rate pair $(C_1, 0)$ satisfies the inequalities (3) for $p(x) = \arg \max I(X; Y_1|S)$ and $(U_1, U_2) = (X, \emptyset)$. Similarly, the rate pair $(0, C_2)$ satisfies the inequalities (3) for $p(x) = \arg \max I(X; Y_2|S)$ and $(U_1, U_2) = (\emptyset, X)$. Thus $(C_1, 0)$ and $(0, C_2)$ are achievable. Now let \mathcal{R}'_1 be the set of rate pairs that satisfy (3) for some $p(x)$ and $(U_1, U_2) = (f_1, f_2)$. We can easily see that $\mathcal{R}_1 \subseteq \mathcal{R}'_1$. Thus \mathcal{C} is achievable via Marton coding and time-sharing.

To establish the converse, we show that \mathcal{C} coincides with the UV outer bound. The UV outer bound for the broadcast channel with state known at the receivers states that if a rate pair (R_1, R_2) is achievable, then it must satisfy the inequalities

$$\begin{aligned} R_1 & \leq I(U; Y_1|S), \\ R_2 & \leq I(V; Y_2|S), \\ R_1 + R_2 & \leq I(U; Y_1|S) + I(X; Y_2|U, S), \\ R_1 + R_2 & \leq I(V; Y_2|S) + I(X; Y_1|V, S) \end{aligned} \quad (4)$$

for some pmf $p(u, v, x)$. Let this outer bound be denoted by $\tilde{\mathcal{R}}$. Clearly $\mathcal{C} \subseteq \tilde{\mathcal{R}}$. We now show that every supporting hyperplane of $\tilde{\mathcal{R}}$ intersects \mathcal{C} , i.e., for all $\lambda \geq 0$,

$$\max_{(R_1, R_2) \in \tilde{\mathcal{R}}} (R_1 + \lambda R_2) \leq \max_{(r_1, r_2) \in \mathcal{C}} (r_1 + \lambda r_2). \quad (5)$$

We first show that inequality (5) holds for $0 \leq \lambda \leq 1$. Consider

$$\begin{aligned} \max_{(R_1, R_2) \in \tilde{\mathcal{R}}} (R_1 + \lambda R_2) & \leq \max_{p(u, x)} (I(U; Y_1|S) + \lambda H(Y_2|U, S)) \\ & = \max_{p(x)} \left(H(Y_1|S) + \max_{p(u|x)} (\lambda H(Y_2|U, S) - H(Y_1|U, S)) \right) \\ & = \max_{p(x)} \left(p_1 H(f_1) + \bar{p}_1 H(f_2) + \max_{p(u|x)} ((\lambda \bar{p}_2 - \bar{p}_1) H(f_2|U) \right. \\ & \quad \left. + (\lambda p_2 - p_1) H(f_1|U)) \right). \end{aligned}$$

We now consider different ranges of $0 \leq \lambda \leq 1$.

- For $0 \leq \lambda \leq \bar{p}_1/\bar{p}_2$, $(\lambda \bar{p}_2 - \bar{p}_1) H(f_2|U) + (\lambda p_2 - p_1) H(f_1|U) \leq 0$ for any fixed $p(x)$ with equality if $U = X$. Thus,

$$\max_{(R_1, R_2) \in \tilde{\mathcal{R}}} (R_1 + \lambda R_2) \leq \max_{p(x)} (p_1 H(f_1) + \bar{p}_1 H(f_2)).$$

Since $(C_1, 0) = (\max_{p(x)} (p_1 H(f_1) + \bar{p}_1 H(f_2)), 0) \in \mathcal{C}$, the inequality (5) holds.

- For $\bar{p}_1/\bar{p}_2 < \lambda \leq 1$, consider

$$\begin{aligned} & (\lambda \bar{p}_2 - \bar{p}_1) H(f_2|U) + (\lambda p_2 - p_1) H(f_1|U) \\ & = (\lambda - 1) H(f_1|U) \\ & \quad + (\lambda \bar{p}_2 - \bar{p}_1) (H(f_2|f_1, U) - H(f_1|f_2, U)) \\ & \leq (\lambda \bar{p}_2 - \bar{p}_1) H(f_2|f_1) \end{aligned}$$

for any fixed $p(x)$ with equality if $U = f_1$. Thus,

$$\begin{aligned} \max_{(R_1, R_2) \in \tilde{\mathcal{R}}} (R_1 + \lambda R_2) & \leq \max_{p(x)} (p_1 H(f_1) + \bar{p}_1 H(f_2) + (\lambda \bar{p}_2 - \bar{p}_1) H(f_2|f_1)) \\ & = \max_{p(x)} (I(f_1; Y_1|S) + \lambda (I(f_2; Y_2|S) - I(f_1; f_2))) \\ & = \max_{p(x) \in \mathcal{P}} (I(f_1; Y_1|S) + \lambda (I(f_2; Y_2|S) - I(f_1; f_2))) \end{aligned}$$

Finally, since $(I(f_1; Y_1|S), I(f_2; Y_2|S) - I(f_1; f_2)) \in \mathcal{C}$ for $p(x) \in \mathcal{P}$, the inequality (5) holds.

We now prove the inequality (5) for $\lambda > 1$. We consider the equivalent maximization problem: $\max_{(R_1, R_2) \in \tilde{\mathcal{R}}} (\lambda^{-1} R_1 + R_2)$. Consider

$$\begin{aligned} \max_{(R_1, R_2) \in \tilde{\mathcal{R}}} (\lambda^{-1} R_1 + R_2) & \leq \max_{p(v, x)} (\lambda^{-1} H(Y_1|V, S) + I(V; Y_2|S)) \\ & = \max_{p(x)} \left(H(Y_2|S) + \max_{p(v|x)} (\lambda^{-1} H(Y_1|V, S) - H(Y_2|V, S)) \right) \\ & = \max_{p(x)} \left(p_2 H(f_1) + \bar{p}_2 H(f_2) \right. \\ & \quad \left. + \max_{p(v|x)} ((\lambda^{-1} \bar{p}_1 - \bar{p}_2) H(f_2|V) + (\lambda^{-1} p_1 - p_2) H(f_1|V)) \right). \end{aligned}$$

We now consider different ranges of $\lambda > 1$.

- For $\lambda > p_1/p_2$, $(\lambda^{-1} \bar{p}_1 - \bar{p}_2) H(f_2|V) + (\lambda^{-1} p_1 - p_2) H(f_1|V) \leq 0$ for any fixed $p(x)$ with equality if $V = X$. Thus,

$$\max_{(R_1, R_2) \in \tilde{\mathcal{R}}} (R_1 + \lambda R_2) \leq \max_{p(x)} (\lambda p_2 H(f_1) + \lambda \bar{p}_2 H(f_2)).$$

Since $(0, C_2) = (0, \max_{p(x)} (\lambda p_2 H(f_1) + \lambda \bar{p}_2 H(f_2))) \in \mathcal{C}$, the inequality (5) holds.

- For $1 < \lambda \leq p_1/p_2$, consider

$$\begin{aligned} & (\lambda^{-1} \bar{p}_1 - \bar{p}_2) H(f_2|V) + (\lambda^{-1} p_1 - p_2) H(f_1|V) \\ & = (\lambda^{-1} - 1) H(f_2|V) \\ & \quad + (\lambda^{-1} p_1 - p_2) (H(f_1|f_2, V) - H(f_2|f_1, V)) \\ & \leq (\lambda^{-1} p_1 - p_2) H(f_1|f_2) \end{aligned}$$

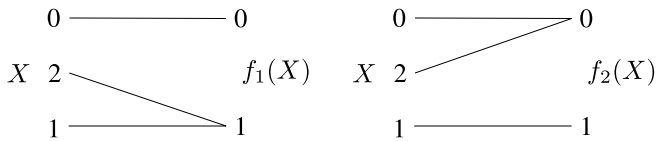


Fig. 3. The deterministic components of the Blackwell channel with state.

for any fixed $p(x)$ with equality if $V = f_2$. Thus,

$$\begin{aligned}
 & \max_{(R_1, R_2) \in \mathcal{R}} (R_1 + \lambda R_2) \\
 & \leq \max_{p(x)} (\lambda p_2 H(f_1) + \lambda \bar{p}_2 H(f_2) + (p_1 - \lambda p_2) H(f_1|f_2)) \\
 & = \max_{p(x)} (I(f_1; Y_1|S) - I(f_1; f_2) + \lambda I(f_2; Y_2|S)) \\
 & = \max_{p(x) \in \mathcal{P}} (I(f_1; Y_1|S) - I(f_1; f_2) + \lambda I(f_2; Y_2|S)).
 \end{aligned}$$

Finally since $(I(f_1; Y_1|S) - I(f_1; f_2), I(f_2; Y_2|S)) \in \mathcal{C}$ for $p(x) \in \mathcal{P}$, the inequality (5) holds.

The proof of the converse is completed using the following lemma. \blacksquare

Lemma 1 [10]: Let $\mathcal{R} \in \mathbb{R}^d$ be convex and $\mathcal{R}_1 \subseteq \mathcal{R}_2$ be two bounded convex subsets of \mathcal{R} , closed relative to \mathcal{R} . If every supporting hyperplane of \mathcal{R}_2 intersects \mathcal{R}_1 , then $\mathcal{R}_1 = \mathcal{R}_2$.

As an example of a BC-TDCS, consider the following.

Example 1 (Blackwell Channel With State [11]): The functions f_1 and f_2 for this example are depicted in Figure 3.

The private message capacity region of the Blackwell channel with state known only to the receivers is the convex hull of

$$\begin{aligned}
 \mathcal{R}'_1 = \{ & (R_1, R_2): R_1 \leq H(\alpha_0) - \bar{p}_1 \bar{\alpha}_1 H(\alpha_0/\bar{\alpha}_1), \\
 & R_2 \leq H(\alpha_1) - p_2 \bar{\alpha}_0 H(\alpha_1/\bar{\alpha}_0), \\
 & R_1 + R_2 \leq H(\alpha_0) - \bar{p}_1 \bar{\alpha}_1 H(\alpha_0/\bar{\alpha}_1) \\
 & \quad + \bar{p}_2 \bar{\alpha}_0 H(\alpha_1/\bar{\alpha}_0) \\
 & \text{for some } \alpha_0, \alpha_1 \geq 0, \alpha_0 + \alpha_1 \leq 1\}.
 \end{aligned}$$

where $H(a)$, $a \in [0, 1]$ is the binary entropy function. Note that \mathcal{R}'_1 , defined in the proof of Theorem 1, is the Marton rate region with $(U_1, U_2) = (f_1, f_2)$ and $X \in \{0, 1, 2\}$ for $p_X(0) = \alpha_0$, $p_X(1) = \alpha_1$, $p_X(2) = 1 - \alpha_0 - \alpha_1$ for $\alpha_0, \alpha_1 \geq 0$, $\alpha_0 + \alpha_1 \leq 1$. Also, since the rate pairs $(C_1, 0) = (1, 0) \in \mathcal{R}'_1$ and $(0, C_2) = (0, 1) \in \mathcal{R}'_1$, \mathcal{C} is the convex hull of \mathcal{R}'_1 . The capacity region with state for $(p_1, p_2) = (0.5, 0.5)$, $(0.7, 0.3)$, and $(1, 0)$ is plotted in Figure 4. For $(p_1, p_2) = (0.5, 0.5)$, the two channels are statistically identical, hence the capacity region coincides with the time-division region. For $(p_1, p_2) = (1, 0)$, the channel reduces to the Blackwell channel with no state [12]. For (p_1, p_2) in between these two extreme cases, the capacity region is established by our theorem.

Next consider the following example which is motivated by deterministic approximations of wireless channels.

Example 2 (Finite-Field BC-TDCS): Consider the BC-TDCS with the state known only to the receivers with

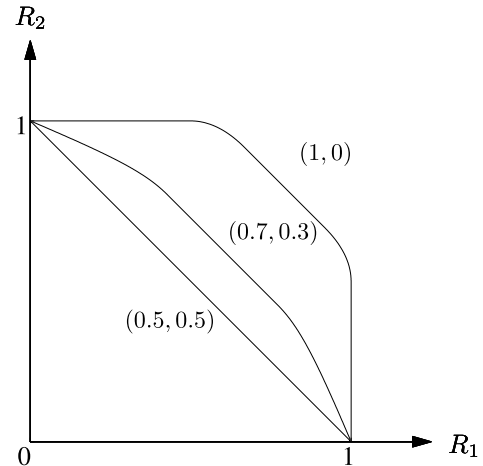


Fig. 4. Capacity region of the Blackwell channel with the state.

$$\mathbf{X} = [X_1 \ X_2]^T:$$

$$\begin{aligned}
 Y_1 &= \begin{cases} h_{11}X_1 + h_{12}X_2 & \text{if } S_1 = 1, \\ h_{21}X_1 + h_{22}X_2 & \text{if } S_1 = 2, \end{cases} \\
 Y_2 &= \begin{cases} h_{11}X_1 + h_{12}X_2 & \text{if } S_2 = 1, \\ h_{21}X_1 + h_{22}X_2 & \text{if } S_2 = 2, \end{cases} \quad (6)
 \end{aligned}$$

where the channel matrix is full-rank, $\mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{X}_1 = \mathcal{X}_2 = [0 : K - 1]$, and the arithmetic is over the finite field.

To compute the private message capacity region, note that $C_1 = \log K$ and $C_2 = \log K$. To evaluate \mathcal{R}_1 , we compute \mathcal{P} . Consider

$$\begin{aligned}
 & p_1 H(f_1) + \bar{p}_1 I(f_1; f_2) + \lambda \bar{p}_2 H(f_2|f_1) \\
 & = p_1 H(f_1) + \bar{p}_1 H(f_2) + (\lambda \bar{p}_2 - \bar{p}_1) H(f_2|f_1) \\
 & \leq (p_1 + \lambda \bar{p}_2) \log K
 \end{aligned}$$

for $\bar{p}_1/\bar{p}_2 \leq \lambda \leq 1$ with equality if $\mathbf{X} \sim \text{Unif}([0 : K - 1]^2)$. Similarly,

$$\begin{aligned}
 & p_1 H(f_1) + \bar{p}_1 I(f_1; f_2) + \lambda \bar{p}_2 H(f_2|f_1) \\
 & = p_1 H(f_1) + \bar{p}_1 H(f_2) + (\lambda \bar{p}_2 - \bar{p}_1) H(f_2|f_1) \\
 & \leq (p_1 + \lambda \bar{p}_2) \log K
 \end{aligned}$$

for $1 \leq \lambda \leq p_1/p_2$ with equality if $\mathbf{X} \sim \text{Unif}([0 : K - 1]^2)$. Thus, $\mathcal{P} = \{\text{Unif}([0 : K - 1]^2)\}$. Note that when \mathbf{X} is uniform, $H(f_1) = H(f_2) = H(f_1|f_2) = H(f_2|f_1) = \log K$. Hence,

$$\mathcal{R}_1 = \{(R_1, R_2): R_1 \leq p_1 \log K, R_2 \leq \bar{p}_2 \log K\},$$

and the capacity region is

$$\begin{aligned}
 \mathcal{C} &= \text{co}\{\mathcal{R}_1 \cup \{(\log K, 0)\} \cup \{(0, \log K)\}\} \\
 &= \text{co}\{(0, 0) \cup (\log K, 0) \cup (0, \log K) \cup (p_1 \log K, \bar{p}_2 \log K)\}.
 \end{aligned}$$

Figure 5 plots the capacity region for $(p_1, p_2) = (0.5, 0.5)$, $(0.7, 0.4)$, and $(1, 0)$. For $(p_1, p_2) = (0.5, 0.5)$, the two channels are statistically identical and the capacity region coincides with the time-division region. For $(p_1, p_2) = (1, 0)$, the capacity region is $\{(R_1, R_2): R_1 \leq \log K, R_2 \leq \log K\}$ because the channel matrix is full-rank. For (p_1, p_2)

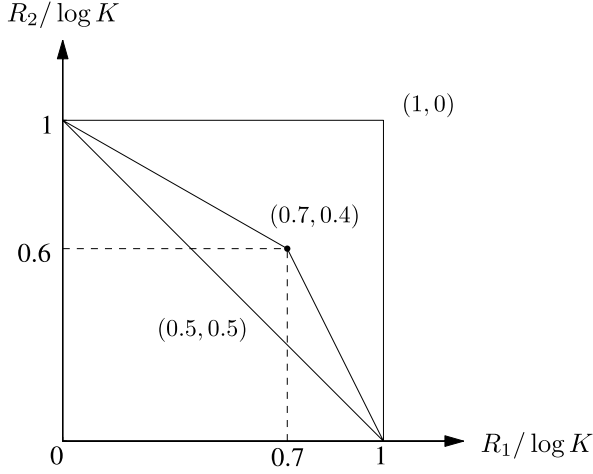


Fig. 5. Capacity region of the Finite Field BC-TDCS.

in between these two extreme cases, the capacity region is established by our theorem.

Remark 1 (Connection to Wireless Channels): Consider the following fading broadcast channel

$$Y_j = \mathbf{H}_j^\dagger \mathbf{X} + Z_j \text{ for } j = 1, 2, \quad (7)$$

where \dagger denotes the conjugate-transpose, $\mathbf{X} = [X_1 \ X_2]^T \in \mathbb{C}^{2 \times 1}$, $\mathbf{E}[\mathbf{X}^\dagger \mathbf{X}] \leq P$, $Z_j \sim \mathcal{CN}(0, 1)$ and the noise sequences Z_{ji} , $j = 1, 2$ and $i \in [1 : n]$, are i.i.d. In addition, for $j = 1, 2$,

$$\mathbf{H}_j^\dagger = \begin{cases} [h_{11} & h_{12}] & \text{if } S_j = 1 \text{ w.p. } p_j, \\ [h_{21} & h_{22}] & \text{if } S_j = 2 \text{ w.p. } \bar{p}_j, \end{cases}$$

where the channel matrix is in $\mathbb{C}^{2 \times 2}$ and is full rank.

We now show that the degrees of freedom (DoF) of this fading Gaussian broadcast channel, obtained by dividing the maximum sum-rate by $\log P$ and taking the limit, is $p_1 + \bar{p}_2$.

Since the variance of the noise Z_j is bounded, the DoF of channel in (7) is equal to that of the BC-TDCS with $Y_j = \mathbf{H}_j^\dagger \mathbf{X}$ for $j = 1, 2$ [13]. We show that the DoF is achieved when $U_1 = f_1$ and $U_2 = f_2$ are independent and Gaussian with variances αP and βP for some $\alpha, \beta > 0$ such that

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}^{-1} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$$

satisfy the power constraint. First note that for $(R_1, R_2) \in \mathcal{C}$,

$$\begin{aligned} & \max_{p(\mathbf{X})} \lim_{P \rightarrow \infty} \frac{R_1 + R_2}{\log P} \\ &= \max_{p(\mathbf{X})} \lim_{P \rightarrow \infty} \frac{p_1 H(f_1) + \bar{p}_2 H(f_2) + (\bar{p}_1 - \bar{p}_2) I(f_1; f_2)}{\log P}. \end{aligned} \quad (8)$$

Now we show that each term in (8) is maximized with the chosen input. First, $\lim_{P \rightarrow \infty} p_1 H(f_1) / \log P = \lim_{P \rightarrow \infty} p_1 \log(\alpha P) / \log P = p_1$. Now we show that $p_1 = \max \lim_{P \rightarrow \infty} p_1 H(f_1) / \log P$. Since $\text{Var}(f_1) = \text{Var}(h_{11} X_1 + h_{12} X_2) = |h_{11}|^2 \gamma P + |h_{12}|^2 \bar{\gamma} P + (h_{11}^* h_{12} + h_{12}^* h_{11}) \rho \sqrt{\gamma \bar{\gamma}} P$ for some $0 \leq \gamma, \rho \leq 1$ due to the power constraint,

$H(f_1) \leq \log(|h_{11}|^2 \gamma + |h_{12}|^2 \bar{\gamma} + (h_{11}^* h_{12} + h_{12}^* h_{11}) \rho \sqrt{\gamma \bar{\gamma}}) + \log P$. Hence, $\lim_{P \rightarrow \infty} p_1 H(f_1) / \log P \leq p_1$. Similarly, $\lim_{P \rightarrow \infty} \bar{p}_2 H(f_2) / \log P$ is maximized and is equal to \bar{p}_2 , and $\lim_{P \rightarrow \infty} (\bar{p}_1 - \bar{p}_2) I(f_1; f_2) / \log P$ is maximized and is equal to 0. Thus, the following holds:

$$\begin{aligned} & \max_{p(\mathbf{X})} \lim_{P \rightarrow \infty} \frac{p_1 H(f_1) + \bar{p}_2 H(f_2) + (\bar{p}_1 - \bar{p}_2) I(f_1; f_2)}{\log P} \\ &= p_1 + \bar{p}_2, \end{aligned}$$

and the DoF of the fading Gaussian BC in (7) is $p_1 + \bar{p}_2$.

The capacity region result can be readily extended to the case with common message ($R_0 \neq 0$).

Theorem 2: The capacity region of a BC-TDCS ($\mathcal{X} \times \mathcal{S}$, $p(s)p(y_1, y_2|x, s)$, $\mathcal{Y}_1 \times \mathcal{Y}_2$) with the state known only at the receivers is the convex hull of the set of all rate pairs (R_0, R_1, R_2) such that

$$\begin{aligned} R_0 &\leq \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ R_0 + R_1 &\leq I(U_0; Y_1|S) + I(U_1; Y_1|U_0, S) \\ R_0 + R_2 &\leq I(U_0; Y_2|S) + I(U_2; Y_2|U_0, S) \\ R_0 + R_1 + R_2 &\leq \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ &\quad + I(U_1; Y_1|U_0, S) + I(U_2; Y_2|U_0, S) \\ &\quad - I(U_1; U_2|U_0) \end{aligned}$$

for some $p(u_0, x)$ with $|u_0| \leq |\mathcal{X}| + 4$, and either $(U_1, U_2) = (X, \emptyset)$, $(U_1, U_2) = (f_1, f_2)$, or $(U_1, U_2) = (\emptyset, X)$.

The proof is in Appendix A.

III. ORDERED CHANNEL STATE COMPONENTS

Recall the definitions of the following classes of BC for which superposition coding was shown to be optimal.

Definition 1 (Degraded BC [14]): For a DM-BC (\mathcal{X} , $p(\tilde{y}_1, \tilde{y}_2|x)$, $\tilde{\mathcal{Y}}_1 \times \tilde{\mathcal{Y}}_2$) receiver \tilde{Y}_2 is said to be a degraded version of \tilde{Y}_1 if there exists Z such that $Z|X=x \sim p_{\tilde{Y}_1|X}(z|x)$, i.e., Z has the same conditional pmf as \tilde{Y}_1 (given X), and $X \rightarrow Z \rightarrow \tilde{Y}_2$ form a Markov chain.

Definition 2 (Less Noisy BC [15]): For a DM-BC (\mathcal{X} , $p(\tilde{y}_1, \tilde{y}_2|x)$, $\tilde{\mathcal{Y}}_1 \times \tilde{\mathcal{Y}}_2$) receiver \tilde{Y}_1 is said to be less noisy than \tilde{Y}_2 if $I(U; \tilde{Y}_1) \geq I(U; \tilde{Y}_2)$ for all $p(u, x)$.

van Dijk [16] showed that receiver \tilde{Y}_1 is less noisy than receiver \tilde{Y}_2 if $I(X; \tilde{Y}_1) - I(X; \tilde{Y}_2)$ is concave in $p(x)$, or equivalently, $I(X; \tilde{Y}_1) - I(X; \tilde{Y}_2)$ is equal to its upper concave envelope $\mathfrak{C}[I(X; \tilde{Y}_1) - I(X; \tilde{Y}_2)]$ (the smallest concave function that is greater than or equal to $I(X; \tilde{Y}_1) - I(X; \tilde{Y}_2)$).

Definition 3 (More Capable BC [15]): For a DM-BC (\mathcal{X} , $p(\tilde{y}_1, \tilde{y}_2|x)$, $\tilde{\mathcal{Y}}_1 \times \tilde{\mathcal{Y}}_2$) receiver \tilde{Y}_1 is said to be more capable than \tilde{Y}_2 if $I(X; \tilde{Y}_1) \geq I(X; \tilde{Y}_2)$ for all $p(x)$.

The more capable condition can also be recast in terms of the concave envelope: Receiver \tilde{Y}_1 is more capable than \tilde{Y}_2 if $\mathfrak{C}[I(X; \tilde{Y}_2) - I(X; \tilde{Y}_1)] = 0$ for every $p(x)$.

Definition 4 (Dominantly c-Symmetric BC [17]):

A discrete memoryless channel (DMC) with input alphabet $\mathcal{X} = \{0, 1, \dots, m-1\}$ and output alphabet \mathcal{Y} of size n is said to be *c-symmetric* if, for each $j = 0, \dots, m-1$, there is a permutation $\pi_j(\cdot)$ of \mathcal{Y} such that $p_{Y|X}(\pi_j(y)|(i+j)_m) = p_{Y|X}(y|i)$ for

all i , where $(i + j)_m = (i + j) \bmod m$. A DM-BC $(\mathcal{X}, p(\tilde{y}_1, \tilde{y}_2|x), \tilde{\mathcal{Y}}_1 \times \tilde{\mathcal{Y}}_2)$ is said to be c-symmetric if both channel components X to \tilde{Y}_1 and X to \tilde{Y}_2 are c-symmetric. A c-symmetric DM-BC is said to be *dominantly c-symmetric* if

$$I(X; \tilde{Y}_1)_p - I(X; \tilde{Y}_2)_p \leq I(X; \tilde{Y}_1)_u - I(X; \tilde{Y}_2)_u \quad (9)$$

for every $p(x)$, where $u(x)$ is the uniform pmf and $I(X; \tilde{Y}_1)_p$ denotes the mutual information between X and \tilde{Y}_1 for $X \sim p(x)$.

The definitions of degraded, less noisy, more capable, and dominantly c-symmetric channels can be readily extended to the BC-TCS with state known only at the receivers.

Definition 5: A broadcast channel with state $(\mathcal{X} \times \mathcal{S}, p(s)p(y_1, y_2|x, s), \mathcal{Y}_1 \times \mathcal{Y}_2)$ with the state known only at the receivers is degraded, less noisy, more capable, or dominantly c-symmetric if the corresponding DM-BC $(\mathcal{X}, p(y_1, y_2, s|x), (\mathcal{Y}_1, \mathcal{S}) \times (\mathcal{Y}_2, \mathcal{S}))$ is degraded, less noisy, more capable, or dominantly c-symmetric, respectively.

In the following we show the surprising fact that if the DM-BC satisfies any of the conditions in Definitions 1–4, then the corresponding BC-TCS with the state known at the receivers also satisfies the same condition. Hence, the capacity regions for these corresponding BC-TCS are achieved using superposition coding.

Theorem 3: A BC-TCS $(\mathcal{X} \times \mathcal{S}, p(s)p(y_1, y_2|x, s), \mathcal{Y}_1 \times \mathcal{Y}_2)$ with state S known only at the receivers is

- (i) degraded if the DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$ is degraded,
- (ii) less noisy if the DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$ is less noisy,
- (iii) more capable if the DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$ is more capable,
- (iv) dominantly c-symmetric if the DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$ is dominantly c-symmetric.

Proof: We prove (i). The proof of the rest of this theorem is in Appendix B.

- (i) For a degraded DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$, there exists Z such that $Z|X = x \sim p_{\tilde{Y}_1|X}(z|x)$ and $X \rightarrow Z \rightarrow \tilde{Y}_2$ forms a Markov chain. We show that there exists (Y'_1, S'_1, S'_2) such that $(Y'_1, S'_1, S'_2)|X = x \sim p_{Y_1, S_1, S_2|X}(y'_1, s'_1, s'_2|x)$ and $X \rightarrow (Y'_1, S'_1, S'_2) \rightarrow (Y_2, S_1, S_2)$ forms a Markov chain.

Let (S'_1, S'_2) be distributed according to

$$p_{S'_1|S_1, S_2}(s'_1|s_1, s_2) = \begin{cases} p_2 & \text{if } (s'_1, s_2) = (1, 1), \\ p_1 - p_2 & \text{if } (s'_1, s_2) = (1, 2), \\ 1 - p_1 & \text{if } (s'_1, s_2) = (2, 2), \end{cases}$$

and $p_{S'_2|S'_1, S_1, S_2}(s'_2|s'_1, s_1, s_2) = p_{S_2|S_1}(s'_2|s'_1)$. Thus $(S'_1, S'_2)|X = x \sim p_{S_1, S_2|X}(s'_1, s'_2|x)$ and

$$p(s_1, s_2|y'_1, s'_1, s'_2, x) = p(s_1, s_2|y'_1, s'_1, s'_2). \quad (10)$$

Let Y'_1 be distributed according to

$$Y'_1 = \begin{cases} Y_2 & \text{if } (S'_1, S_2) = (1, 1) \text{ or } (2, 2), \\ Z & \text{if } (S'_1, S_2) = (1, 2) \end{cases}$$

where $X \rightarrow Z \rightarrow \tilde{Y}_2$. Then $Y'_1|{(S'_1, S'_2, X) = (s'_1, s'_2, x)} \sim p_{Y_1|S_1, S_2, X}(y'_1|s'_1, s'_2, x)$ and

$$p(y_2, |s_1, s_2, y'_1, s'_1, s'_2, x) = p(y_2|s_1, s_2, y'_1, s'_1, s'_2). \quad (11)$$

By (10) and (11), It follows that $(Y'_1, S'_1, S'_2)|X = x \sim p_{Y_1, S_1, S_2|X}(y'_1, s'_1, s'_2|x)$ and $X \rightarrow (Y'_1, S'_1, S'_2) \rightarrow (Y_2, S_1, S_2)$ forms a Markov chain. ■

Remark 2: A BC-TCS $(\mathcal{X} \times \mathcal{S}, p(s)p(y_1, y_2|x, s), \mathcal{Y}_1 \times \mathcal{Y}_2)$ with state S known only at the receivers is degraded, less noisy, or more capable if and only if the DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$ is degraded, less noisy, or more capable, respectively (assuming $p_1 > p_2$). If the DM-BC $(\mathcal{X}, p(y_1, y_2, s|x), (\mathcal{Y}_1, \mathcal{S}) \times (\mathcal{Y}_2, \mathcal{S}))$ is degraded, there exists (Y'_1, S'_1, S'_2) such that $(Y'_1, S'_1, S'_2)|X = x \sim p_{Y_1, S_1, S_2|X}(y'_1, s'_1, s'_2|x)$ and $X \rightarrow (Y'_1, S'_1, S'_2) \rightarrow (Y_2, S_1, S_2)$ forms a Markov chain. Let Z be distributed according to $p_{Z|X}(z|x) = p_{Y'_1|S'_1, S'_2}(z|1, x)$ and $p_{\tilde{Y}_2|Z, X}(\tilde{y}_2|z, x) = p_{Y_2|S_2, S'_1, Y'_1, X}(\tilde{y}_2|2, 1, z, x)$. Then $Z|X = x \sim p_{\tilde{Y}_1|X}(z|x)$ and $\tilde{Y}_2|X = x \sim p_{\tilde{Y}_2|X}(\tilde{y}_2|x)$. Also $X \rightarrow Z \rightarrow \tilde{Y}_2$ because $p_{Y_2|S_2, S'_1, Y'_1, X}(\tilde{y}_2|2, 1, z, x) = p_{Y_2|S_2, S'_1, Y'_1}(\tilde{y}_2|2, 1, z)$. Therefore the DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$ is degraded. The proofs for less noisy and more capable DM-BC follow directly from the proof of part (ii) and (iii) of Theorem 3. We do not know however if the BC-TCS $(\mathcal{X} \times \mathcal{S}, p(s)p(y_1, y_2|x, s), \mathcal{Y}_1 \times \mathcal{Y}_2)$ with state S known only at the receivers is dominantly c-symmetric if and only if the DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$ is dominantly c-symmetric.

It follows from Theorem 3 that the capacity region of the BC-TCS satisfying the conditions in Theorem 3 is the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq I(X; Y_1|U, S), \\ R_2 &\leq I(U; Y_2|S), \\ R_1 + R_2 &\leq I(X; Y_1|S) \end{aligned} \quad (12)$$

for some $p(u, x)$ with $|\mathcal{U}| \leq |\mathcal{X}| + 1$.

Remark 3: Using superposition coding, receiver \tilde{Y}_1 can recover receiver \tilde{Y}_2 's message. Hence when there is common message ($R_0 \neq 0$), the capacity region is obtained by replacing R_1 with $R_0 + R_1$.

As an example of BC-TCS with more capable or dominantly c-symmetric components, consider the following.

Example 3 (A BC-TCS With a BSC and a BEC Channel Components): A BC-TCS with a BSC and a BEC channel components has input $\mathcal{X} = \{0, 1\}$ and channel components BSC(p) and BEC(e). Without loss of generality, we assume $0 \leq p \leq 1/2$ and $0 \leq e \leq 1$. In [17], it is shown that for the DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$,

- 1) \tilde{Y}_1 is a degraded version of \tilde{Y}_2 if and only if $0 \leq e \leq 2p$.
- 2) \tilde{Y}_2 is less noisy than \tilde{Y}_1 if and only if $0 \leq e \leq 4p(1-p)$.
- 3) \tilde{Y}_2 is more capable than \tilde{Y}_1 if and only if $0 \leq e \leq H(p)$.
- 4) \tilde{Y}_1 is dominantly c-symmetric if $H(p) \leq e \leq 1$.

Hence, by Theorem 3, the corresponding BC-TCS with BSC(p) and BEC(e) channel components is degraded, less noisy, more capable, or dominantly c-symmetric for the above channel parameter ranges.

A. A Product of Reversely More Capable Channel Components

Another class of broadcast channel for which superposition coding is shown to be optimal for each component is the product of reversely more capable broadcast channels [2].

Definition 6 (Product of Reversely More Capable BCs): A DM-BC $(\mathcal{X}, p(\tilde{y}_1, \tilde{y}_2|x), \tilde{\mathcal{Y}}_1 \times \tilde{\mathcal{Y}}_2)$ is said to be a product of reversely more capable DM-BC if $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2)$, $\tilde{\mathcal{Y}}_1 = (\tilde{\mathcal{Y}}_{11}, \tilde{\mathcal{Y}}_{12})$, $\tilde{\mathcal{Y}}_2 = (\tilde{\mathcal{Y}}_{21}, \tilde{\mathcal{Y}}_{22})$, and $p(\tilde{y}_{11}, \tilde{y}_{12}, \tilde{y}_{21}, \tilde{y}_{22}|x_1, x_2) = p(\tilde{y}_{11}, \tilde{y}_{12}|x_1)p(\tilde{y}_{21}, \tilde{y}_{22}|x_2)$, and $I(X_1; \tilde{Y}_{11}) \geq I(X_1; \tilde{Y}_{21})$ for all $p(x_1)$ and $I(X_2; \tilde{Y}_{12}) \leq I(X_2; \tilde{Y}_{22})$ for all $p(x_2)$.

We extend this definition to the broadcast channel with two channel state components as follows.

Definition 7 (A Product BC-TCS): A 2-receiver product broadcast channel with two channel state components is a DM-BC with random state $(\mathcal{X} \times \mathcal{S}, p(s)p(y_1, y_2|x, s), \mathcal{Y}_1 \times \mathcal{Y}_2)$, where $X = [X_1, X_2]$ and $S = (S_1, S_2)$ for

$$\begin{aligned} Y_j &= [Y_{j1}, Y_{j2}], \\ S_j &= [S_{j1}, S_{j2}], \\ Y_{j1} &= \begin{cases} \tilde{Y}_{11} & \text{if } S_{j1} = 1, \\ \tilde{Y}_{21} & \text{if } S_{j1} = 2 \end{cases} \\ Y_{j2} &= \begin{cases} \tilde{Y}_{12} & \text{if } S_{j2} = 1, \\ \tilde{Y}_{22} & \text{if } S_{j2} = 2 \end{cases} \end{aligned}$$

for $j = 1, 2$ and $p(\tilde{y}_{11}, \tilde{y}_{12}|x_1)p(\tilde{y}_{21}, \tilde{y}_{22}|x_2)$.

Let $p_{S_{j1}}(1) = p_{j1}$ and $p_{S_{j2}}(1) = p_{j2}$ for $j = 1, 2$. Without loss of generality, we assume $p_{11} \geq p_{12}$ and $p_{21} \geq p_{22}$. In the following we establish the capacity region of BC-TCS with reversely more capable components.

Theorem 4: A 2-receiver product BC-TCS $(\mathcal{X} \times \mathcal{S}, p(s)p(y_1, y_2|x, s), \mathcal{Y}_1 \times \mathcal{Y}_2)$ is more capable if the product DM-BC $(\mathcal{X}, p(\tilde{y}_1, \tilde{y}_2|x), \tilde{\mathcal{Y}}_1 \times \tilde{\mathcal{Y}}_2)$ for $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2)$, $\tilde{\mathcal{Y}}_1 = (\tilde{\mathcal{Y}}_{11}, \tilde{\mathcal{Y}}_{12})$, $\tilde{\mathcal{Y}}_2 = (\tilde{\mathcal{Y}}_{21}, \tilde{\mathcal{Y}}_{22})$ is reversely more capable.

Proof: We show that the product DM-BC $(\mathcal{X}, p(y_1, y_2|x), \mathcal{Y}_1 \times \mathcal{Y}_2)$ for $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2)$, $\mathcal{Y}_1 = ((\mathcal{Y}_{11}, \mathcal{S}_{11}), (\mathcal{Y}_{12}, \mathcal{S}_{12}))$, $\mathcal{Y}_2 = ((\mathcal{Y}_{21}, \mathcal{S}_{21}), (\mathcal{Y}_{22}, \mathcal{S}_{22}))$ is reversely more capable if the product DM-BC $(\mathcal{X}, p(\tilde{y}_1, \tilde{y}_2|x), \tilde{\mathcal{Y}}_1 \times \tilde{\mathcal{Y}}_2)$ for $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2)$, $\tilde{\mathcal{Y}}_1 = (\tilde{\mathcal{Y}}_{11}, \tilde{\mathcal{Y}}_{12})$, $\tilde{\mathcal{Y}}_2 = (\tilde{\mathcal{Y}}_{21}, \tilde{\mathcal{Y}}_{22})$ is reversely more capable. Consider

$$\begin{aligned} I(X_1; Y_{11}, S_{11}) &= p_{11}I(X_1; \tilde{Y}_{11}) + \bar{p}_{11}I(X_1; \tilde{Y}_{21}) \\ &\geq p_{12}I(X_1; \tilde{Y}_{11}) + \bar{p}_{12}I(X_1; \tilde{Y}_{21}) \\ &= I(X_1; Y_{21}, S_{21}). \end{aligned}$$

Similarly, it can be shown that $I(X_2; Y_{12}, S_{12}) \leq I(X_2; Y_{22}, S_{22})$. ■

An immediate consequence of Theorem 4 is that superposition coding for each component is optimal. The capacity region is the region shown in [2, Th. 3] by replacing Y_{ji} with (Y_{ji}, S) for $j, i \in [1 : 2]$.

B. Gaussian Vector Channel Components

Consider the BC-TCS with degraded vector Gaussian channel components

$$\begin{aligned} \tilde{Y}_1 &= \mathbf{G}\mathbf{X} + \mathbf{Z}_1, \\ \tilde{Y}_2 &= \mathbf{G}\mathbf{X} + \mathbf{Z}_2, \end{aligned} \quad (13)$$

where $\mathbf{X}, \mathbf{Z}_1, \mathbf{Z}_2 \in \mathbb{R}^t$ and \mathbf{X} and \mathbf{Z}_j are independent for $j = 1, 2$. The channel gain matrix is $G \in \mathbb{R}^{t \times t}$, and $\mathbf{Z}_1 \sim \mathcal{N}(0, N_1)$, and $\mathbf{Z}_2 \sim \mathcal{N}(0, N_2)$ for some

$N_2 - N_1 \geq 0$. Assume the average transmission power constraint $\sum_{i=1}^n \mathbf{x}^T(m_1, m_2, i)\mathbf{x}(m_1, m_2, i) \leq nP$ for $(m_1, m_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$.

By Theorem 3, the BC-TCS with degraded vector Gaussian channel components is degraded and its capacity region is achieved via superposition coding. In the following, we show that it suffices to consider only Gaussian (U, X) .

Proposition 1: The capacity region of a BC-TCS with degraded vector Gaussian components is the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq p_1 \log \frac{|GK_1G^T + N_1|}{|N_1|} + \bar{p}_1 \log \frac{|GK_1G^T + N_2|}{|N_2|}, \\ R_2 &\leq p_2 \log \frac{|GKG^T + N_1|}{|GK_1G^T + N_1|} + \bar{p}_2 \log \frac{|GKG^T + N_2|}{|GK_1G^T + N_2|} \end{aligned} \quad (14)$$

for some $K \geq 0$ for $\text{tr}(K) \leq P$ and $K \geq K_1 \geq 0$.

Proof: By Theorem 3, the capacity region is set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq I(\mathbf{X}; \mathbf{Y}_1|\mathbf{U}, S), \\ R_2 &\leq I(\mathbf{U}; \mathbf{Y}_2|S) \end{aligned} \quad (15)$$

for some pmf $p(\mathbf{u}, \mathbf{x})$.

Let \mathcal{C}_G denote the set of rate pairs (R_1, R_2) that satisfy the inequalities in (15) for some $\mathbf{U} \sim \mathcal{N}(0, K_1)$ and $\mathbf{V} \sim \mathcal{N}(0, K - K_1)$, independent of each other, and $\mathbf{X} = \mathbf{U} + \mathbf{V}$ for some $K \geq K_1 \geq 0$ and $\text{tr}(K) \leq P$. It can be easily shown that \mathcal{C}_G is the set of rate pairs that satisfy inequalities in (14). To show that \mathcal{C}_G is the capacity region, we show the following.

Lemma 2: For all $\lambda \geq 0$,

$$\begin{aligned} \max_{(R_1, R_2) \in \mathcal{C}_G} (R_1 + \lambda R_2) \\ \geq \max_{p(\mathbf{u}, \mathbf{x}): \mathbb{E}[\mathbf{X}^T \mathbf{X}] \leq P} (I(\mathbf{X}; \mathbf{Y}_1|\mathbf{U}, S) + \lambda I(\mathbf{U}; \mathbf{Y}_2|S)). \end{aligned}$$

The proof of this lemma is in Appendix C. The proof of Proposition 1 is completed using Lemma 1. ■

Remark 4: Recall that Gaussian superposition coding and dirty paper coding both achieve the capacity region of Gaussian BC-TCS when $(p_1, p_2) = (1, 0)$, i.e., when the channel gain is fixed. For general (p_1, p_2) , Gaussian superposition coding achieves the capacity region, but dirty paper coding does not. See Appendix D for the proof.

IV. MORE THAN TWO CHANNEL STATE COMPONENTS

In this section we consider the BC with more than two channel state components. Consider a DM-BC with random state, where the state $S = (S_1, S_2) \in [1 : k]^2$, $p_{S_1}(i) = p_i$ and $p_{S_2}(i) = q_i$, channel components $p(\tilde{y}_i|x)$ for $i \in [1 : k]$, and outputs $Y_1 = \tilde{Y}_i$ if $S_1 = i$ and $Y_2 = \tilde{Y}_i$ if $S_2 = i$ for $i \in [1 : k]$.

In the following we establish several results when $k > 2$.

A. Binary Erasure Broadcast Channel With k Channel Components

Consider a BC with k state components where the channel $p(\tilde{y}_i|x)$ is a BEC(ϵ_i), $0 \leq \epsilon_i \leq 1$, for $i \in [1 : k]$. We show that this channel is always less noisy.

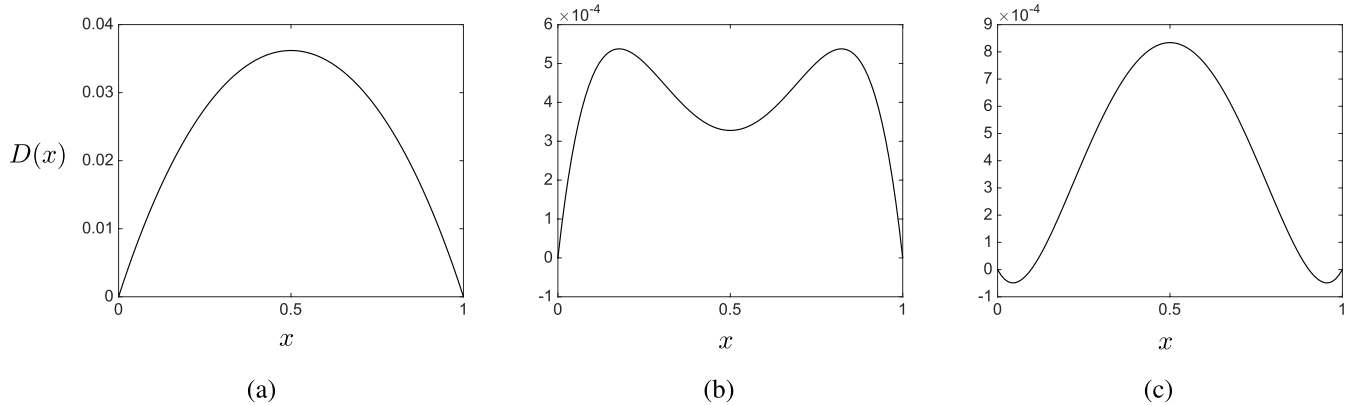


Fig. 6. Examples of $D(x)$ vs x for $\alpha = [0.2, 0.3, 0.4]$, $p = [1/3, 1/3, 1/3]$ and (a) $q = [0.2, 0.3, 0.5]$ (b) $q = [0.2, 0.7, 0.1]$ (c) $q = [0.45, 0, 0.55]$.

Theorem 5: The binary erasure broadcast channel with k channel state components with the state known only at the receivers is always less noisy.

Proof: Without loss of generality, assume that the capacity of channel $p(y_1, s|x)$, C_1 , is larger than the capacity of the channel $p(y_2, s|x)$, C_2 . Then for any $p(u, x)$,

$$\begin{aligned} I(U; Y_1, S) &= H(U) - \sum_{i=1}^k p_i (\epsilon_i H(U) + (1 - \epsilon_i) H(U|X)) \\ &= C_1 I(U; X) \\ &\geq C_2 I(U; X) \\ &= I(U; Y_2, S). \end{aligned}$$

An immediate consequence of this theorem is that the capacity region is achieved via superposition coding. Since $I(U; Y_1|S) = C_1 I(U; X)$ and $I(X; Y_2|S) = H(X) - \sum_{i=1}^k p_i \epsilon_i H(X) = C_2 H(X)$, superposition coding inner bound in (12) is equivalent to the set of rate pairs that satisfy

$$\begin{aligned} R_1 &\leq C_1 I(U; X), \\ R_2 &\leq C_2 H(X|U), \\ R_1 + R_2 &\leq C_2 H(X) \end{aligned}$$

for some $p(u, x)$. It can be easily seen that any achievable rate pair (R_1, R_2) satisfies $R_1/C_1 + R_2/C_2 \leq H(X) \leq 1$, and the rate pairs $(C_1, 0)$ and $(0, C_2)$ are achievable. Thus capacity region is the set of rate pairs (R_1, R_2) such that

$$\frac{R_1}{C_1} + \frac{R_2}{C_2} \leq 1.$$

B. Binary Symmetric Broadcast Channel With Three Channel Components

Consider a BC with three channel state components where the channel $p(\tilde{y}_i|x)$ is a BSC(α_i), $0 \leq \alpha_i \leq 1$, for $i \in [1 : 3]$. We can show that superposition coding is optimal for this channel.

Theorem 6: The BC with three binary symmetric channel state components is more capable or dominantly c-symmetric.

Proof: Let $D(x) = I(X; Y_1|S) - I(X; Y_2|S)$ for $X \sim \text{Bern}(x)$, i.e.,

$$\begin{aligned} D(x) &= \sum_{i=1}^3 p_i H(x * \alpha_i) - \sum_{i=1}^3 p_i H(\alpha_i) \\ &\quad - \left(\sum_{i=1}^3 q_i H(x * \alpha_i) - \sum_{i=1}^3 q_i H(\alpha_i) \right), \end{aligned}$$

where $a * b = a(1 - b) + b(1 - a)$ for $a, b \in [0, 1]$. Without loss of generality, we assume $D(0.5) = C_1 - C_2 \geq 0$.

The DM-BC $(\mathcal{X}, p(y_1, y_2, s|x), (\mathcal{Y}_1, \mathcal{S}) \times (\mathcal{Y}_2, \mathcal{S}))$ is dominantly c-symmetric if $X \rightarrow (Y_1, S)$ and $X \rightarrow (Y_2, S)$ are c-symmetric and $I(X; Y_1|S)_p - I(X; Y_2|S)_p \leq I(X; Y_1|S)_u - I(X; Y_2|S)_u$.

Note that the proof of part (iv) of Theorem 3 which shows that $X \rightarrow (Y_j, S)$ are c-symmetric if $X \rightarrow \tilde{Y}_j$ are symmetric for $j = 1, 2$ does not rely on the cardinality of \mathcal{S} . Thus the proof can be extended to show that $X \rightarrow (Y_j, S)$ for $j = 1, 2$ are c-symmetric for BC with three channel state components. In order to show the DM-BC $(\mathcal{X}, p(y_1, y_2, s|x), (\mathcal{Y}_1, \mathcal{S}) \times (\mathcal{Y}_2, \mathcal{S}))$ is more capable or dominantly c-symmetric, we now show that $D(0.5) \geq D(x)$ for every $x \in [0, 1]$ or $D(x) \geq 0$ for every $x \in [0, 1]$. After some computation, we obtain

$$\begin{aligned} D''(x) &= \frac{(p_1 - q_1)((1 - 2\alpha_3)^2 \alpha_1 \bar{\alpha}_1 - (1 - 2\alpha_1)^2 \alpha_3 \bar{\alpha}_3)}{(x * \alpha_1)(1 - x * \alpha_1)(x * \alpha_3)(1 - x * \alpha_3)} \\ &\quad + \frac{(p_2 - q_2)((1 - 2\alpha_3)^2 \alpha_2 \bar{\alpha}_2 - (1 - 2\alpha_2)^2 \alpha_3 \bar{\alpha}_3)}{(x * \alpha_2)(1 - x * \alpha_2)(x * \alpha_3)(1 - x * \alpha_3)}. \end{aligned}$$

Note that $D''(x) = 0$ if

$$\begin{aligned} &\frac{(p_1 - q_1)((1 - 2\alpha_3)^2 \alpha_1 \bar{\alpha}_1 - (1 - 2\alpha_1)^2 \alpha_3 \bar{\alpha}_3)}{(x * \alpha_1)(1 - x * \alpha_1)} \\ &\quad + \frac{(p_2 - q_2)((1 - 2\alpha_3)^2 \alpha_2 \bar{\alpha}_2 - (1 - 2\alpha_2)^2 \alpha_3 \bar{\alpha}_3)}{(x * \alpha_2)(1 - x * \alpha_2)} = 0. \end{aligned}$$

Since $D''(x) = 0$ has at most two solutions in $(0, 1)$, $D'(x) = 0$ has at most three solutions in $(0, 1)$. Since $D'(0.5) = 0$ and $D(x) = D(1 - x)$, i.e., symmetric with respect to $x = 0.5$, $D'(x) = 0$ has one solution or three solutions. If it has one solution, $D(x)$ is concave (see Figure 6-(a) for an example). If it has three solutions, $D(x) \geq 0$ or $D(0.5) \geq D(x)$ for $x \in [0, 1]$ as illustrated in Figure 6-(b) and 6-(c). ■

An immediate consequence of this theorem is that the capacity region of the BC with three BSC state components when the state is known only at the receivers is the set of rate pairs (R_1, R_2) that satisfy the inequalities in (12). We now show that this region is reduced to the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq \sum_{i=1}^3 p_i \left(\sum_{j=1}^2 \gamma_j H(\beta_j * \alpha_i) - H(\alpha_i) \right), \\ R_2 &\leq 1 - \sum_{i=1}^3 q_i \sum_{j=1}^2 \gamma_j H(\beta_j * \alpha_i), \\ R_1 + R_2 &\leq 1 - \sum_{i=1}^3 p_i H(\alpha_i) \end{aligned} \quad (16)$$

for some $0 \leq \gamma_j, \beta_j \leq 1$ for $j \in [1 : 2]$ such that $\gamma_1 + \gamma_2 = 1$.

Suppose a rate pair (R_1, R_2) satisfies the inequalities in (12) for some $p(u, x)$ such that $|\mathcal{U}| = 3$. Then, this rate pair is also achievable with the following (U', X') such that $U' \in \{-3, -2, -1, 1, 2, 3\}$ and

$$\begin{aligned} p_{U'}(u) &= p_{U'}(-u) = \frac{1}{2} p_U(u), \quad u \in \{1, 2, 3\}, \\ p_{X'|U'}(x|u) &= p_{X'|U'}(1-x|-u) = p_{X|U}(x|u), \\ (u, x) &\in \{1, 2, 3\} \times \{0, 1\}. \end{aligned}$$

Further let (Y'_1, Y'_2) be the output when the input is X' . It can be easily seen that $H(Y'_1|U', S) = H(Y_1|U, S)$ and $H(Y'_2|U', S) = H(Y_2|U, S)$. Also note that $H(Y'_1|S) = H(Y'_2|S) = 1$ because $X' \sim \text{Bern}(1/2)$. Thus, $I(U'; Y'_1|S) \leq I(U'; Y_1|S)$, $I(X'; Y_2|U, S) \leq I(X'; Y'_2|U', S)$, and $I(X'; Y_2|S) \leq I(X'; Y'_2|S)$.

Therefore, it suffices to evaluate the superposition rate region with the above symmetric input pmfs $p(u', x')$, and the capacity region is the set of rate pairs (R_1, R_2) that satisfy

$$\begin{aligned} R_1 &\leq \sum_{i=1}^3 p_i \left(\sum_{j=1}^3 \gamma_j H(\beta_j * \alpha_i) - H(\alpha_i) \right), \\ R_2 &\leq 1 - \sum_{i=1}^3 q_i \sum_{j=1}^3 \gamma_j H(\beta_j * \alpha_i), \\ R_1 + R_2 &\leq 1 - \sum_{i=1}^3 p_i H(\alpha_i) \end{aligned}$$

for some $0 \leq \gamma_j, \beta_j \leq 1$ for $j \in [1 : 3]$ such that $\gamma_1 + \gamma_2 + \gamma_3 = 1$. Note that this rate region can be written as an intersection of two rate regions, $\mathcal{R}_1 \cap \mathcal{R}_2$, where $\mathcal{R}_1 = \{(R_1, R_2) : R_1 + R_2 \leq 1 - \sum_{i=1}^3 p_i H(\alpha_i)\}$ and

$$\begin{aligned} \mathcal{R}_2 &= \left\{ (R_1, R_2) : R_1 \leq \sum_{i=1}^3 p_i \left(\sum_{j=1}^3 \gamma_j H(\beta_j * \alpha_i) - H(\alpha_i) \right), \right. \\ R_2 &\leq 1 - \sum_{i=1}^3 q_i \sum_{j=1}^3 \gamma_j H(\beta_j * \alpha_i) \\ &\left. \text{for } 0 \leq \gamma_j, \beta_j \leq 1 \text{ s.t. } \sum_{j=1}^3 \gamma_j = 1, j \in [1 : 3] \right\}. \end{aligned}$$

Let $\bar{\mathcal{R}}_2$ denote the *convex hull* of the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq \sum_{i=1}^3 p_i (H(\beta * \alpha_i) - H(\alpha_i)), \\ R_2 &\leq 1 - \sum_{i=1}^3 q_i H(\beta * \alpha_i) \end{aligned}$$

for some $0 \leq \beta \leq 1$. Note that since $\bar{\mathcal{R}}_2$ is a convex set in 2-dimension, all rate pairs in $\bar{\mathcal{R}}_2$ is a convex combination of two rate pairs included in $\bar{\mathcal{R}}_2$. Thus,

$$\begin{aligned} \bar{\mathcal{R}}_2 &= \left\{ (R_1, R_2) : R_1 \leq \sum_{i=1}^3 p_i \left(\sum_{j=1}^2 \gamma_j H(\beta_j * \alpha_i) - H(\alpha_i) \right), \right. \\ R_2 &\leq 1 - \sum_{i=1}^3 q_i \sum_{j=1}^2 \gamma_j H(\beta_j * \alpha_i) \\ &\left. \text{for } 0 \leq \gamma_j, \beta_j \leq 1 \text{ s.t. } \sum_{j=1}^2 \gamma_j = 1, j \in [1 : 2] \right\}. \end{aligned}$$

Note that this rate region is a subset of \mathcal{R}_2 , so $\bar{\mathcal{R}}_2 \subseteq \mathcal{R}_2$. Also it can be easily seen that $\mathcal{R}_2 \subseteq \bar{\mathcal{R}}_2$, and so $\mathcal{R}_2 = \bar{\mathcal{R}}_2$. Therefore, the capacity region for BC-TCS with three BSC channel components is $\mathcal{R}_1 \cap \bar{\mathcal{R}}_2$, the region shown in (16).

In the following we show that superposition coding is not in general optimal for BC with more than three BSC state components.

C. Binary Symmetric Broadcast Channel With Four Channel Components

Consider a BC-TCS with BSC components with $\alpha_1 = 0.28, \alpha_2 = 0.04, \alpha_3 = 0.02, \alpha_4 = 0.18$, and $p = [0.38, 0.62, 0, 0]$ and $q = [0, 0, 0.38, 0.62]$. Thus $C_1 = 0.5247, C_2 = 0.5246$, and the maximum sum rate for superposition coding is $\max(C_1, C_2) = 0.5247$.

Now we consider the Marton inner bound. In [18], Geng, Jog, Nair and Wang showed that for binary input broadcast channels, Marton's inner bound reduces to the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< I(W; Y_1) + \sum_{j=1}^k \beta_j I(X; Y_1|W = j), \\ R_2 &< I(W; Y_2) + \sum_{j=k+1}^5 \beta_j I(X; Y_2|W = j), \\ R_1 + R_2 &< \min\{I(W; Y_1), I(W; Y_2)\} \\ &\quad + \sum_{j=1}^k \beta_j I(X; Y_1|W = j) \\ &\quad + \sum_{j=k+1}^5 \beta_j I(X; Y_2|W = j) \end{aligned} \quad (17)$$

for some $p_W(j) = \beta_j, j \in [1 : 5]$, and $p(x|w)$. This region is achieved using *randomized time-division* [19]. This ingenious insight helps simplify the computation of Marton's

inner bound for BC-TCS with BSC components. In this case, the maximum sum rate is 0.5250 and is strictly greater than maximum sum rate for superposition coding. Thus, superposition coding is suboptimal. It is not known whether Marton coding is optimal, however, because there is a gap between the Marton maximum sum rate and the sum rate for the UV outer bound, which in this case is at least 0.5256.

V. CONCLUSION

We established the capacity region of several classes of BC-TCS channel when the state is known only at the receivers. When the channel state components are deterministic, the capacity region is achieved via Marton coding. This is an interesting result because this channel model does not belong to any class of broadcast channels for which the capacity was previously known. When the channel state components are ordered, the BC-TCS is also ordered and the capacity region is achieved via superposition coding. We showed that when the BC-TCS has degraded vector Gaussian channel components, the capacity region is attained via Gaussian input and auxiliary random variables. We extended our results on ordered channel components to two example channels with more than two channel components, but showed that this extension does not hold in general.

APPENDIX A PROOF OF THEOREM 2

Let \mathcal{C}_0 denote the region shown in Theorem 2. Achievability follows immediately since \mathcal{C}_0 is included in Marton's inner bound with common message.

To establish the converse, we show that the capacity region coincides with the UVW outer bound. The UVW outer bound for the broadcast channel with state known at the receivers states that if a rate tuple (R_0, R_1, R_2) is achievable, then it must satisfy the inequalities

$$\begin{aligned} R_0 &\leq \min\{I(U_0; Y_1|S), U(U_0; Y_2|S)\}, \\ R_0 + R_1 &\leq I(U_1; Y_1|U_0, S) \\ &\quad + \min\{I(U_0; Y_1|S), U(U_0; Y_2|S)\}, \\ R_0 + R_2 &\leq I(U_2; Y_2|U_0, S) \\ &\quad + \min\{I(U_0; Y_1|S), U(U_0; Y_2|S)\}, \\ R_0 + R_1 + R_2 &\leq \min\{I(U_0; Y_1|S), U(U_0; Y_2|S)\} \\ &\quad + I(U_1; Y_1|U_2, U_0, S) + I(U_2; Y_2|U_0, S), \\ R_0 + R_1 + R_2 &\leq \min\{I(U_0; Y_1|S), U(U_0; Y_2|S)\} \\ &\quad + I(U_1; Y_1|U_0, S) + I(U_2; Y_2|U_1, U_0, S) \end{aligned}$$

for some pmf $p(u_0, u_1, u_2, x)$. Let this outer bound be denoted by $\bar{\mathcal{R}}_0$. We now show that every supporting hyperplane of $\bar{\mathcal{R}}_0$ intersects \mathcal{C}_0 , i.e.

$$\begin{aligned} &\max_{(R_0, R_1, R_2) \in \bar{\mathcal{R}}_0} (\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2) \\ &\leq \max_{(r_0, r_1, r_2) \in \mathcal{C}_0} (\lambda_0 r_0 + \lambda_1 r_1 + \lambda_2 r_2). \end{aligned} \quad (18)$$

We consider different ranges of $(\lambda_0, \lambda_1, \lambda_2)$ and show that the inequality (18) always holds.

- (1) If $\lambda_2 \leq \lambda_0 \leq \lambda_1$ or $\lambda_0 \leq \lambda_2 \leq \lambda_1$, note that for any $(R_0, R_1, R_2) \in \bar{\mathcal{R}}_0$,

$$\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \leq \lambda_1 (R_0 + R_1) + \lambda_2 R_2.$$

Thus

$$\begin{aligned} &\max_{(R_0, R_1, R_2) \in \bar{\mathcal{R}}_0} (\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2) \\ &\leq \max_{(r_1, r_2) \in \bar{\mathcal{R}}} (\lambda_1 r_1 + \lambda_2 r_2) \\ &\leq \max_{(r_0, r_1, r_2) \in \mathcal{C}_0} (\lambda_0 r_0 + \lambda_1 r_1 + \lambda_2 r_2), \end{aligned}$$

where $\bar{\mathcal{R}}$ denotes the UV outer bound in (4). The last inequality follows because \mathcal{C}_0 includes the private message capacity region \mathcal{C} .

- (2) If $\lambda_1 \leq \lambda_0 \leq \lambda_2$ or $\lambda_0 \leq \lambda_1 \leq \lambda_2$, note that for any $(R_0, R_1, R_2) \in \bar{\mathcal{R}}_0$,

$$\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \leq \lambda_1 R_1 + \lambda_2 (R_0 + R_2).$$

Thus

$$\begin{aligned} &\max_{(R_0, R_1, R_2) \in \bar{\mathcal{R}}_0} (\lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2) \\ &\leq \max_{(r_1, r_2) \in \bar{\mathcal{R}}} (\lambda_1 r_1 + \lambda_2 r_2) \\ &\leq \max_{(r_0, r_1, r_2) \in \mathcal{C}_0} (\lambda_0 r_0 + \lambda_1 r_1 + \lambda_2 r_2), \end{aligned}$$

where $\bar{\mathcal{R}}$ denotes the UV outer bound in (4). The last inequality follows because \mathcal{C}_0 includes the private message capacity region \mathcal{C} .

- (3) If $\lambda_1 \leq \lambda_2 \leq \lambda_0$, note that for any $(R_0, R_1, R_2) \in \bar{\mathcal{R}}_0$,

$$\begin{aligned} &\lambda_0 R_0 + \lambda_2 R_2 + \lambda_1 R_1 \\ &\leq \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ &\quad + \lambda_2 I(U_2; Y_2|U_0, S) + \lambda_1 I(X; Y_1|U_2, U_0, S) \\ &= \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ &\quad + \lambda_2 I(U_2; Y_2|U_0, S) + \lambda_1 H(Y_1|U_2, U_0, S) \\ &= \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} + \lambda_2 H(Y_2|U_0, S) \\ &\quad + \lambda_1 H(Y_1|U_2, U_0, S) - \lambda_2 H(Y_2|U_2, U_0, S) \\ &= \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} + \lambda_2 p_2 H(f_1|U_0) \\ &\quad + \lambda_2 \bar{p}_2 H(f_2|U_0) + (\bar{p}_1 \lambda_1 - \bar{p}_2 \lambda_2) H(f_2|U_2, U_0) \\ &\quad + (p_1 \lambda_1 - p_2 \lambda_2) H(f_1|U_2, U_0) \end{aligned}$$

For a fixed $p(u_0, x)$, only the last two terms depend on $p(u_2|u_0, x)$. We now consider different ranges of (λ_1, λ_2) .

- If $\lambda_2 \geq p_1 \lambda_1 / p_2$, then for any fixed $p(u_0, x)$,

$$\begin{aligned} &(\bar{p}_1 \lambda_1 - \bar{p}_2 \lambda_2) H(f_2|U_2, U_0) \\ &\quad + (p_1 \lambda_1 - p_2 \lambda_2) H(f_1|U_2, U_0) \leq 0 \end{aligned}$$

with equality if $U_2 = X$. Thus,

$$\begin{aligned} &\lambda_0 R_0 + \lambda_2 R_2 + \lambda_1 R_1 \\ &\leq \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ &\quad + \lambda_2 p_2 H(f_1|U_0) + \lambda_2 \bar{p}_2 H(f_2|U_0) \\ &\leq \max_{(r_0, r_1, r_2) \in \mathcal{C}_0} (\lambda_0 r_0 + \lambda_2 r_2 + \lambda_1 r_1). \end{aligned}$$

- If $\lambda_2 < p_1\lambda_1/p_2$, then for any fixed $p(u_0, x)$,

$$\begin{aligned} & (\bar{p}_1\lambda_1 - \bar{p}_2\lambda_2)H(f_2|U_2, U_0) \\ & + (p_1\lambda_1 - p_2\lambda_2)H(f_1|U_2, U_0) \\ & = (\lambda_1 - \lambda_2)H(f_2|U_2, U_0) \\ & + (p_1\lambda_1 - p_2\lambda_2)\{H(f_1|U_2, U_0) - H(f_2|U_2, U_0)\} \\ & = (\lambda_1 - \lambda_2)H(f_2|U_2, U_0) \\ & + (p_1\lambda_1 - p_2\lambda_2)H(f_1|f_2, U_2, U_0) \\ & - (p_1\lambda_1 - p_2\lambda_2)H(f_2|f_1, U_2, U_0) \\ & \leq (p_1\lambda_1 - p_2\lambda_2)H(f_1|f_2, U_0) \end{aligned}$$

with equality if $U_2 = f_2$. Thus,

$$\begin{aligned} & \lambda_0 R_0 + \lambda_2 R_2 + \lambda_1 R_1 \\ & \leq \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_2 H(Y_2|U_0, S) + (p_1\lambda_1 - p_2\lambda_2)H(f_1|f_2, U_0) \\ & = \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_2 p_2 H(f_1|U_0) + \lambda_2 \bar{p}_2 H(f_2|U_0) \\ & + (p_1\lambda_1 - p_2\lambda_2)H(f_1|f_2, U_0) \\ & = \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_2 I(f_2; Y_2|U_0, S) + \lambda_1 p_1 H(f_1|f_2, U_0) \\ & = \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_2 I(f_2; Y_2|U_0, S) + \lambda_1 I(f_1; Y_1|U_0, S) \\ & - \lambda_1 I(f_1; f_2|U_0, S) \\ & \leq \max_{(r_0, r_1, r_2) \in \mathcal{C}_0} (\lambda_0 r_0 + \lambda_2 r_2 + \lambda_1 r_1). \end{aligned}$$

- (4) If $\lambda_2 \leq \lambda_1 \leq \lambda_0$, note that for any $(R_0, R_1, R_2) \in \tilde{\mathcal{R}}_0$,

$$\begin{aligned} & \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\ & \leq \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_1 I(U_1; Y_1|U_0, S) + \lambda_2 I(X; Y_2|U_1, U_0, S) \\ & = \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_1 I(U_1; Y_1|U_0, S) + \lambda_2 H(Y_2|U_1, U_0, S) \\ & = \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} + \lambda_1 H(Y_1|U_0, S) \\ & + \lambda_2 H(Y_2|U_1, U_0, S) - \lambda_1 H(Y_1|U_1, U_0, S) \\ & = \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} + \lambda_1 p_1 H(f_1|U_0) \\ & + \bar{p}_1 H(f_2|U_0) + (p_2\lambda_2 - p_1\lambda_1)H(f_1|U_1, U_0) \\ & + (\bar{p}_2\lambda_2 - \bar{p}_1\lambda_1)H(f_2|U_1, U_0) \end{aligned}$$

For a fixed $p(u_0, x)$, only the last two terms depend on $p(u_1|u_0, x)$. We now consider different ranges of (λ_1, λ_2) .

- If $\lambda_1 \geq \bar{p}_2\lambda_2/\bar{p}_1$, then for any fixed $p(u_0, x)$,

$$\begin{aligned} & (p_2\lambda_2 - p_1\lambda_1)H(f_1|U_1, U_0) \\ & + (\bar{p}_2\lambda_2 - \bar{p}_1\lambda_1)H(f_2|U_1, U_0) \leq 0 \end{aligned}$$

with equality if $U_1 = X$. Thus,

$$\begin{aligned} & \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\ & \leq \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_1 p_1 H(f_1|U_0) + \bar{p}_1 H(f_2|U_0) \\ & \leq \max_{(r_0, r_1, r_2) \in \mathcal{C}_0} (\lambda_0 r_0 + \lambda_2 r_2 + \lambda_1 r_1). \end{aligned}$$

- If $\lambda_1 < \bar{p}_2\lambda_2/\bar{p}_1$, then for any fixed $p(u_0, x)$,

$$\begin{aligned} & (p_2\lambda_2 - p_1\lambda_1)H(f_1|U_1, U_0) \\ & + (\bar{p}_2\lambda_2 - \bar{p}_1\lambda_1)H(f_2|U_1, U_0) \\ & = (\lambda_2 - \lambda_1)H(f_1|U_1, U_0) \\ & + ((\bar{p}_2\lambda_2 - \bar{p}_1\lambda_1)\{H(f_2|U_1, U_0) \\ & - H(f_1|U_1, U_0)\}) \\ & = (\lambda_2 - \lambda_1)H(f_1|U_1, U_0) \\ & + (\bar{p}_2\lambda_2 - \bar{p}_1\lambda_1)H(f_2|f_1, U_1, U_0) \\ & - (\bar{p}_2\lambda_2 - \bar{p}_1\lambda_1)H(f_1|f_2, U_1, U_0) \\ & \leq (\bar{p}_2\lambda_2 - \bar{p}_1\lambda_1)H(f_2|f_1, U_0) \end{aligned}$$

with equality if $U_1 = f_1$. Thus,

$$\begin{aligned} & \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2 \\ & \leq \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_1 H(Y_1|U_0, S) + (\bar{p}_2\lambda_2 - \bar{p}_1\lambda_1)H(f_2|f_1, U_0) \\ & = \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_1 p_1 H(f_1|U_0) + \lambda_1 \bar{p}_1 H(f_2|U_0) \\ & + (\bar{p}_2\lambda_2 - \bar{p}_1\lambda_1)H(f_2|f_1, U_0) \\ & = \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_1 I(f_1; Y_1|U_0, S) + \lambda_2 \bar{p}_2 H(f_2|f_1, U_0) \\ & = \lambda_0 \min\{I(U_0; Y_1|S), I(U_0; Y_2|S)\} \\ & + \lambda_1 I(f_1; Y_1|U_0, S) + \lambda_2 I(X; Y_2|f_1, U_0, S) \\ & \leq \max_{(r_0, r_1, r_2) \in \mathcal{C}_0} (\lambda_0 r_0 + \lambda_2 r_2 + \lambda_1 r_1). \end{aligned}$$

The proof of the converse is completed using Lemma 1.

APPENDIX B

PROOF OF (II) - (IV) OF THEOREM 3

We show that if a DM-BC $(\mathcal{X}, p(\tilde{y}_1, \tilde{y}_2), \tilde{\mathcal{Y}}_1 \times \tilde{\mathcal{Y}}_2)$ is less noisy, more capable, or dominantly c-symmetric, then the DM-BC $(\mathcal{X}, p(y_1, y_2, s|x), (\mathcal{Y}_1, S) \times (\mathcal{Y}_2, S))$ is also less noisy, more capable, or dominantly c-symmetric, respectively.

- (ii) For a less noisy DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$, $I(U; \tilde{Y}_1) \geq I(U; \tilde{Y}_2)$ for every $p(u, x)$. Consider

$$\begin{aligned} & I(U; Y_1, S) - I(U; Y_2, S) \\ & = I(U; Y_1|S) - I(U; Y_2|S) \\ & = p_1 I(U; \tilde{Y}_1) + \bar{p}_1 I(U; \tilde{Y}_2) - p_2 I(U; \tilde{Y}_1) \\ & \quad - \bar{p}_2 I(U; \tilde{Y}_2) \\ & = (p_1 - p_2)(I(U; \tilde{Y}_1) - I(U; \tilde{Y}_2)) \\ & \geq 0. \end{aligned}$$

Thus the DM-BC $(\mathcal{X}, p(y_1, s_1, y_2, s_2|x), (\mathcal{Y}_1, S) \times (\mathcal{Y}_2, S))$ is less noisy. Note that if $p_1 > p_2$, $I(U; Y_1, S) - I(U; Y_2, S) \geq 0$ if and only if $I(U; \tilde{Y}_1) - I(U; \tilde{Y}_2) \geq 0$.

- (iii) For a more capable DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$, $I(X; \tilde{Y}_1) \geq I(X; \tilde{Y}_2)$ for every $p(u, x)$. Consider

$$\begin{aligned} & I(X; Y_1, S) - I(X; Y_2, S) \\ &= I(X; Y_1|S) - I(X; Y_2|S) \\ &= p_1 I(X; \tilde{Y}_1) + \bar{p}_1 I(X; \tilde{Y}_2) - p_2 I(X; \tilde{Y}_1) \\ &\quad - \bar{p}_2 I(X; \tilde{Y}_2) \\ &= (p_1 - p_2)(I(X; \tilde{Y}_1) - I(X; \tilde{Y}_2)) \\ &\geq 0. \end{aligned}$$

Thus the DM-BC $(\mathcal{X}, p(y_1, s_1, y_2, s_2|x), (\mathcal{Y}_1, S) \times (\mathcal{Y}_2, S))$ is more capable. Note that if $p_1 > p_2$, $I(X; Y_1, S) - I(X; Y_2, S) \geq 0$ if and only if $I(X; \tilde{Y}_1) - I(X; \tilde{Y}_2) \geq 0$.

- (iv) For a dominantly c-symmetric DM-BC $p(\tilde{y}_1, \tilde{y}_2|x)$, let $\pi_j^1(y)$ and $\pi_j^2(y)$ be functions that satisfy

$$\begin{aligned} p_{\tilde{Y}_1|X}(\pi_j^1(\tilde{y}_1)|(i+j)_m) &= p_{\tilde{Y}_1|X}(\tilde{y}_1|i), \\ p_{\tilde{Y}_2|X}(\pi_j^2(\tilde{y}_2)|(i+j)_m) &= p_{\tilde{Y}_2|X}(\tilde{y}_2|i) \end{aligned}$$

for $i \in [0 : m-1]$, where $(i+j)_m$ denotes $(i+j) \bmod m$. To show that the DM-BC $(\mathcal{X}, p(y_1, s_1, y_2, s_2|x), (\mathcal{Y}_1, S) \times (\mathcal{Y}_2, S))$ is dominantly c-symmetric, we first show that $X \rightarrow (Y_1, S)$ and $X \rightarrow (Y_2, S)$ are c-symmetric channels. Let

$$\pi_j(y, s) = (\pi_j^s(y), s).$$

Consider

$$\begin{aligned} p_{(Y_1, S)|X}(y, s|i) &= p_S(s) p_{Y_1|X, S}(y|i, s) \\ &= p_S(s) p_{\tilde{Y}_1|X}(y|i) \\ &= p_S(s) p_{\tilde{Y}_1|X}(\pi_j^1(y)|(i+j)_m) \\ &= p_S(s) p_{Y_1|X, S}(\pi_j^1(y)|(i+j)_m, s) \\ &= p_{(Y_1, S)|X}((\pi_j^1(y), s)|(i+j)_m) \\ &= p_{(Y_1, S)|X}(\pi_j(y, s)|(i+j)_m). \end{aligned}$$

Thus $X \rightarrow (Y_1, S)$ is c-symmetric. Similarly we can show that $X \rightarrow (Y_2, S)$ is c-symmetric. To complete the proof we show that the inequality in (9) holds. Consider

$$\begin{aligned} & I(X; Y_1, S)_p - I(X; Y_2, S)_p \\ &= I(X; Y_1|S)_p - I(X; Y_2|S)_p \\ &= (p_1 - p_2)(I(X; \tilde{Y}_1)_p - I(X; \tilde{Y}_2)_p) \\ &\leq (p_1 - p_2)(I(X; \tilde{Y}_1)_u - I(X; \tilde{Y}_2)_u) \\ &= I(X; Y_1, S)_u - I(X; Y_2, S)_u. \end{aligned}$$

Thus the DM-BC $(\mathcal{X}, p(y_1, s_1, y_2, s_2|x), (\mathcal{Y}_1, S) \times (\mathcal{Y}_2, S))$ is dominantly c-symmetric.

APPENDIX C PROOF OF LEMMA 2

We first prove the lemma for $\lambda \geq 1$. For $\lambda \geq 1$, consider

$$\begin{aligned} & \max_{\substack{p(u, x): \\ \mathbb{E}[\mathbf{X}^T \mathbf{X}] \leq P}} (I(\mathbf{X}; \mathbf{Y}_1|U, S) + \lambda I(U; \mathbf{Y}_2|S)) \\ &= \max_{\substack{p(\mathbf{x}): \\ \mathbb{E}[\mathbf{X}^T \mathbf{X}] \leq P}} (\lambda I(\mathbf{X}; \mathbf{Y}_2|S) + \mathfrak{C}[I(\mathbf{X}; \mathbf{Y}_1|S) - \lambda I(\mathbf{X}; \mathbf{Y}_2|S)]) \\ &= \max_{\substack{K \geq 0: \\ \text{tr}(K) \leq P}} \max_{\substack{p(\mathbf{x}): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^T] = K}} (\lambda I(\mathbf{X}; \mathbf{Y}_2|S) \\ &\quad + \mathfrak{C}[I(\mathbf{X}; \mathbf{Y}_1|S) - \lambda I(\mathbf{X}; \mathbf{Y}_2|S)]) \\ &= \max_{\substack{K \geq 0: \\ \text{tr}(K) \leq P}} \max_{p(\mathbf{x}): \mathbb{E}[\mathbf{X}\mathbf{X}^T] = K} (\lambda I(\mathbf{X}; \mathbf{Y}_2|S) \\ &\quad + \mathfrak{C}[I(\mathbf{X}; \mathbf{Y}_1|S) - \lambda I(\mathbf{X}; \mathbf{Y}_2|S)]) \\ &= \max_{\substack{K \geq 0: \\ \text{tr}(K) \leq P}} \max_{p(\mathbf{x}): \mathbb{E}[\mathbf{X}\mathbf{X}^T] = K} (\lambda p_2 I(\mathbf{X}; \tilde{\mathbf{Y}}_1) + \lambda \bar{p}_2 I(\mathbf{X}; \tilde{\mathbf{Y}}_2) \\ &\quad + \mathfrak{C}[(p_1 - \lambda p_2) I(\mathbf{X}; \tilde{\mathbf{Y}}_1) + (\bar{p}_1 - \lambda \bar{p}_2) I(\mathbf{X}; \tilde{\mathbf{Y}}_2)]) \\ &\leq \max_{\substack{K \geq 0: \\ \text{tr}(K) \leq P}} \left(\max_{p(\mathbf{x}): \mathbb{E}[\mathbf{X}\mathbf{X}^T] = K} \lambda p_2 I(\mathbf{X}; \tilde{\mathbf{Y}}_1) \right. \\ &\quad \left. + \max_{p(\mathbf{x}): \mathbb{E}[\mathbf{X}\mathbf{X}^T] = K} \lambda \bar{p}_2 I(\mathbf{X}; \tilde{\mathbf{Y}}_2) \right. \\ &\quad \left. + \max_{p(\mathbf{x}): \mathbb{E}[\mathbf{X}\mathbf{X}^T] = K} \mathfrak{C}[(p_1 - \lambda p_2) I(\mathbf{X}; \tilde{\mathbf{Y}}_1) \right. \\ &\quad \left. + (\bar{p}_1 - \lambda \bar{p}_2) I(\mathbf{X}; \tilde{\mathbf{Y}}_2)] \right) \\ &\stackrel{(a)}{=} \max_{\substack{K \geq 0: \\ \text{tr}(K) \leq P}} \left(\lambda p_2 \log \frac{|GK G^T + N_1|}{|N_1|} \right. \\ &\quad \left. + \lambda \bar{p}_2 \log \frac{|GK G^T + N_2|}{|N_2|} \right. \\ &\quad \left. + \max_{\substack{K_1: \\ 0 \leq K_1 \leq K}} \left((p_1 - \lambda p_2) \log \frac{|GK_1 G^T + N_1|}{|N_1|} \right. \right. \\ &\quad \left. \left. + (\bar{p}_1 - \lambda \bar{p}_2) \log \frac{|GK_1 G^T + N_2|}{|N_2|} \right) \right) \\ &= \max_{\substack{K \geq 0: \\ \text{tr}(K) \leq P}} \max_{0 \leq K_1 \leq K} \left(\lambda p_2 \log \frac{|GK G^T + N_1|}{|GK_1 G^T + N_1|} \right. \\ &\quad \left. + \lambda \bar{p}_2 \log \frac{|GK G^T + N_2|}{|GK_1 G^T + N_2|} \right. \\ &\quad \left. + p_1 \log \frac{|GK_1 G^T + N_1|}{|N_1|} + \bar{p}_1 \log \frac{|GK_1 G^T + N_2|}{|N_2|} \right) \\ &= \max_{(R_1, R_2) \in \mathcal{C}_G} (R_1 + \lambda R_2). \end{aligned}$$

To show step (a) for $\lambda \geq p_1/p_2$, note that $p_1 - \lambda p_2 \leq 0$ and $\bar{p}_1 - \lambda \bar{p}_2 \leq 0$, and thus

$$\max_{\substack{p(\mathbf{x}): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^T] = K}} \mathfrak{C}[(p_1 - \lambda p_2) I(\mathbf{X}; \tilde{\mathbf{Y}}_1) + (\bar{p}_1 - \lambda \bar{p}_2) I(\mathbf{X}; \tilde{\mathbf{Y}}_2)] = 0.$$

To show step (a) for $1 \leq \lambda < p_1/p_2$, note that $(p_1 - \lambda p_2) I(\mathbf{X}; \tilde{\mathbf{Y}}_1) + (\bar{p}_1 - \lambda \bar{p}_2) I(\mathbf{X}; \tilde{\mathbf{Y}}_2) = (p_1 - \lambda p_2) (I(\mathbf{X}; \tilde{\mathbf{Y}}_1) - \mu I(\mathbf{X}; \tilde{\mathbf{Y}}_2))$ where $\mu = 1 + (\lambda - 1)/(p_1 - \lambda p_2) > 1$,

and for $\mu > 1$,

$$\begin{aligned} & \max_{\substack{p(\mathbf{x}): \\ \mathbb{E}[\mathbf{X}\mathbf{X}^T] = \mathbf{K}}} \mathfrak{C}[I(\mathbf{X}; \tilde{Y}_1) - \mu I(\mathbf{X}; \tilde{Y}_2)] \\ &= \max_{\substack{K_1: \\ K \geq K_1}} \left(\log \frac{|GK_1G^T + N_1|}{|N_1|} - \mu \log \frac{|GK_1G^T + N_2|}{|N_2|} \right). \end{aligned}$$

We now prove the lemma for $\lambda < 1$. Since $I(\mathbf{U}; \mathbf{Y}_2|S) \leq I(\mathbf{U}; \mathbf{Y}_1|S)$ for any $p(\mathbf{u}, \mathbf{x})$, it follows that

$$\begin{aligned} & \max_{\substack{p(\mathbf{u}, \mathbf{x}): \\ \mathbb{E}[\mathbf{X}^T \mathbf{X}] \leq P}} (I(\mathbf{X}; \mathbf{Y}_1|\mathbf{U}, S) + \lambda I(\mathbf{U}; \mathbf{Y}_2|S)) \\ & \leq \max_{\substack{p(\mathbf{x}): \\ \mathbb{E}[\mathbf{X}^T \mathbf{X}] \leq P}} I(\mathbf{X}; \mathbf{Y}_1|S) \\ &= \max_{\substack{K \geq 0: \\ \text{tr}(\mathbf{K}) \leq P}} \left(p_1 \log \frac{|GKG^T + N_1|}{|N_1|} + \bar{p}_1 \log \frac{|GKG^T + N_2|}{|N_2|} \right) \\ & \leq \max_{(R_1, R_2) \in \mathcal{C}_G} (R_1 + \lambda R_2). \end{aligned}$$

APPENDIX D

SUBOPTIMALITY OF DIRTY PAPER CODING

Consider a BC-TCS with scalar Gaussian channel components, i.e., $t = 1$ in (13). Let $U_1 \sim \mathcal{N}(0, 1)$, $U_2 \sim \mathcal{N}(0, 1)$, and $\mathbb{E}[U_1 U_2] = \rho$, $X = aU_1 + bU_2$, where (a, b, ρ) satisfies the power constraint, $\mathbb{E}[X^2] = a^2 + b^2 + 2ab\rho = T \leq P$. Using dirty paper coding, (R_1, R_2) is achievable if

$$\begin{aligned} R_1 &< p_1 \log \left(1 + \frac{(a + b\rho)^2}{b^2(1 - \rho^2) + N_1} \right) \\ &\quad + \bar{p}_1 \log \left(1 + \frac{(a + b\rho)^2}{b^2(1 - \rho^2) + N_2} \right), \\ R_2 &< p_2 \log \left(1 + \frac{(b + a\rho)^2}{a^2(1 - \rho^2) + N_1} \right) \\ &\quad + \bar{p}_2 \log \left(1 + \frac{(b + a\rho)^2}{a^2(1 - \rho^2) + N_2} \right), \\ R_1 + R_2 &< p_1 \log \left(1 + \frac{(a + b\rho)^2}{b^2(1 - \rho^2) + N_1} \right) \\ &\quad + \bar{p}_1 \log \left(1 + \frac{(a + b\rho)^2}{b^2(1 - \rho^2) + N_2} \right) \\ &\quad + p_2 \log \left(1 + \frac{(b + a\rho)^2}{a^2(1 - \rho^2) + N_1} \right) \\ &\quad + \bar{p}_2 \log \left(1 + \frac{(b + a\rho)^2}{a^2(1 - \rho^2) + N_2} \right) \\ &\quad - \log \left(\frac{1}{1 - \rho^2} \right). \end{aligned}$$

Let this region be denoted by \mathcal{R}_D . To show that dirty paper coding is suboptimal, we show that $\max_{(r_1, r_2) \in \mathcal{R}_D} (r_1 + \lambda r_2) < \max_{(R_1, R_2) \in \mathcal{C}_G} (R_1 + \lambda R_2)$ for some $\lambda > 1$. Note that for $\lambda > 1$,

$$\begin{aligned} & \max_{(R_1, R_2) \in \mathcal{R}_D} (R_1 + \lambda R_2) \\ &= \max_{a, b, \rho: a^2 + b^2 + 2ab\rho = T \leq P} \left(p_1 \log \left(\frac{T + N_1}{b^2(1 - \rho^2) + N_1} \right) \right. \\ &\quad \left. + \bar{p}_1 \log \left(\frac{T + N_2}{b^2(1 - \rho^2) + N_2} \right) - \log \left(\frac{1}{1 - \rho^2} \right) \right) \end{aligned}$$

$$\begin{aligned} & + \lambda p_2 \log \left(\frac{T + N_1}{a^2(1 - \rho^2) + N_1} \right) \\ & + \lambda \bar{p}_2 \log \left(\frac{T + N_2}{a^2(1 - \rho^2) + N_2} \right) \Big), \\ & \max_{(R_1, R_2) \in \mathcal{C}_G} (R_1 + \lambda R_2) \\ &= \max_{\alpha \in [0, 1], T \leq P} \left(p_1 \log \left(\frac{\alpha T + N_1}{N_1} \right) + \bar{p}_1 \log \left(\frac{\alpha T + N_2}{N_2} \right) \right. \\ &\quad \left. + \lambda p_2 \log \left(\frac{T + N_1}{\alpha T + N_1} \right) + \lambda \bar{p}_2 \log \left(\frac{T + N_2}{\alpha T + N_2} \right) \right). \quad (19) \end{aligned}$$

Consider

$$\begin{aligned} & \max_{(R_1, R_2) \in \mathcal{C}_G} (R_1 + \lambda R_2) \\ & \stackrel{(a)}{\geq} p_1 \log \left(\frac{a^2(1 - \rho^2) + N_1}{N_1} \right) + \bar{p}_1 \log \left(\frac{a^2(1 - \rho^2) + N_2}{N_2} \right) \\ &\quad + \lambda p_2 \log \left(\frac{T + N_1}{a^2(1 - \rho^2) + N_1} \right) \\ &\quad + \lambda \bar{p}_2 \log \left(\frac{T + N_2}{a^2(1 - \rho^2) + N_2} \right) \\ & \stackrel{(b)}{\geq} p_1 \log \left(\frac{T + N_1}{b^2(1 - \rho^2) + N_1} \right) + \bar{p}_1 \log \left(\frac{T + N_2}{b^2(1 - \rho^2) + N_2} \right) \\ &\quad - \log \left(\frac{1}{1 - \rho^2} \right) \\ &\quad + \lambda p_2 \log \left(\frac{T + N_1}{a^2(1 - \rho^2) + N_1} \right) \\ &\quad + \lambda \bar{p}_2 \log \left(\frac{T + N_2}{a^2(1 - \rho^2) + N_2} \right). \end{aligned}$$

Step (a) follows by plugging in $\alpha = a^2(1 - \rho^2)/T$ in (19). To show step (b), note that the difference between the LHS and RHS is

$$\begin{aligned} & p_1 \log \frac{(a^2(1 - \rho^2) + N_1)(b^2(1 - \rho^2) + N_1)}{N_1(1 - \rho^2)(a^2 + b^2 + 2ab\rho + N_1)} \\ &\quad + \bar{p}_1 \log \frac{(a^2(1 - \rho^2) + N_2)(b^2(1 - \rho^2) + N_2)}{N_2(1 - \rho^2)(a^2 + b^2 + 2ab\rho + N_2)} \geq 0 \end{aligned} \quad (20)$$

because $(a^2(1 - \rho^2) + N_j)(b^2(1 - \rho^2) + N_j) - N_j(1 - \rho^2)(a^2 + b^2 + 2ab\rho + N_j) = (ab(1 - \rho^2) - \rho N_j)^2 \geq 0$ for $j = 1, 2$. Equality holds for (20) if and only if $(\rho, a) = (0, 0)$ or $(\rho, b) = (0, 0)$. Thus equality in step (b) holds if and only if $\max_{(R_1, R_2) \in \mathcal{C}_G} (R_1 + \lambda R_2) = \max(C_1, \lambda C_2)$, which is in general not true. If $\max_{(R_1, R_2) \in \mathcal{C}_G} (R_1 + \lambda R_2) \neq \max(C_1, \lambda C_2)$, it follows that

$$\max_{(R_1, R_2) \in \mathcal{C}_G} (R_1 + \lambda R_2) > \max_{(R_1, R_2) \in \mathcal{R}_D} (R_1 + \lambda R_2).$$

REFERENCES

- [1] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [2] Y. Geng, A. Gohari, C. Nair, and Y. Yu, "On Marton's inner bound and its optimality for classes of product broadcast channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 22–41, Jan. 2014.
- [3] A. Jafarian and S. Vishwanath, "The two-user Gaussian fading broadcast channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul./Aug. 2011, pp. 2964–2968.

- [4] Y. Liang and A. Goldsmith, "Rate regions and optimal power allocation for TD fading broadcast channels without CSIT," in *Proc. 43rd Ann. Allerton Conf. Commun. Control Comput.*, Sep. 2005, pp. 1136–1145.
- [5] D. N. C. Tse and R. D. Yates, "Fading broadcast channels with state information at the receivers," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3453–3471, Jun. 2012.
- [6] H. Kim and A. El Gamal, "Capacity region of the broadcast channel with two deterministic channel state components," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 1952–1956.
- [7] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.
- [8] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 350–355, Jan. 2007.
- [9] E. A. Abbe and L. Zheng, "Coding along Hermite polynomials for Gaussian noise channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun./Jul. 2009, pp. 1644–1648.
- [10] H. G. Eggleston, *Convexity*. Cambridge, U.K.: Cambridge Univ. Press, 1958.
- [11] H. Kim, Y.-K. Chia, and A. El Gamal, "A note on the broadcast channel with stale state information at the transmitter," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3622–3631, Jul. 2015.
- [12] D. Blackwell, L. Breiman, and A. J. Thomasian, "Proof of Shannon's transmission theorem for finite-state indecomposable channels," *Ann. Math. Statist.*, vol. 29, no. 4, pp. 1209–1220, 1958.
- [13] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, Jul. 2012.
- [14] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.
- [15] J. Körner and K. Marton, "Comparison of two noisy channels," in *Topics in Information Theory (Colloquia Mathematica Societatis Janos Bolyai, Keszthely, Hungary)*, I. Csiszár and P. Elias, Eds. Amsterdam, The Netherlands: 1977, pp. 411–423.
- [16] M. van Dijk, "On a special class of broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 712–714, Mar. 1997.
- [17] C. Nair, "Capacity regions of two new classes of two-receiver broadcast channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4207–4214, Sep. 2010.
- [18] Y. Geng, V. Jog, C. Nair, and Z. V. Wang, "An information inequality and evaluation of Marton's inner bound for binary input broadcast channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4095–4105, Jul. 2013.
- [19] B. Hajek and M. B. Pursley, "Evaluation of an achievable rate region for the broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 1, pp. 36–46, Jan. 1979.

Hyeji Kim is a postdoctoral research associate with the Coordinated Science Laboratory at University of Illinois at Urbana-Champaign. She received her B.S. degree with honors in Electrical Engineering from the Korea Advanced Institute of Science and Technology (KAIST) in 2011 and her M.S. and Ph.D. degrees in Electrical Engineering from Stanford University in 2013 and 2016, respectively. She was a recipient of the Stanford Graduate Fellowship (SGF). Her research interest include information theory and statistical learning.

Abbas El Gamal (S'71–M'73–SM'83–F'00) is the Fortinet Founders Chair of the Department of Electrical Engineering and the Hitachi America Professor in the School of Engineering at Stanford University. He received his B.Sc. Honors degree from Cairo University in 1972, and his M.S. in Statistics and Ph.D. in Electrical Engineering both from Stanford University in 1977 and 1978, respectively. From 1978 to 1980, he was an Assistant Professor of Electrical Engineering at USC. From 2003 to 2012, he was the Director of the Information Systems Laboratory at Stanford University. His research contributions have been in network information theory, FPGAs, and digital imaging devices and systems. He has authored or coauthored over 230 papers and holds 35 patents in these areas. He is coauthor of the book *Network Information Theory* (Cambridge Press 2011). He is a member of the US National Academy of Engineering and a Fellow of the IEEE. He received several honors and awards for his research contributions, including the 2016 IEEE Richard Hamming Medal, the 2014 Viterbi Lecture, the 2013 Shannon Memorial Lecture, the 2012 Claude E. Shannon Award, the inaugural Padovani Lecture, and the 2004 INFOCOM Paper Award. He has been serving on the Board of Governors of the Information Theory Society since 2009 and is currently Senior Past President. He has played key roles in several semiconductor, EDA, and biotechnology startup companies.