On the Optimality of Randomized Time Division and Superposition Coding for the Broadcast Channel

Chandra Nair*, Hyeji Kim† and Abbas El Gamal†

***Abstract*—This paper shows that the slope at each corner point of the capacity region of the general broadcast channel coincides with that of the randomized time division (hence the Marton) inner bound and the Nair–El Gamal (as well as the Körner–Marton) outer bound. We then show that the optimal superposition coding inner bound by Bandemer, El Gamal, and Kim can be simplified to the convex closure of the union of the Cover–Bergmans UX region and the Cover–van der Meulen UV region. Generalizing a result by Hajek and Pursely on the skewed binary symmetric broadcast channel, we show that for binary input broadcast channels, the UV region reduces to time division further simplifying the superposition coding inner bound. Finally we establish necessary and sufficient conditions for the optimality of the superposition inner bound for skewed binary broadcast channels.**

I. Introduction

Consider the 2-receiver discrete-memoryless broadcast channel $p(y_1, y_2|x)$ depicted in Figure 1. We consider the private message setup in which the sender X wishes to communicate a message $M_1 \in [1 : 2^{nR_1}]$ to receiver Y_1 and a message $M_2 \in [1 : 2^{nR_2}]$ to receiver Y_2 and define the capacity region of this channel as the closure of the set of all achievable rates (R_1, R_2) (see [1, chapter 5] for detailed definitions).

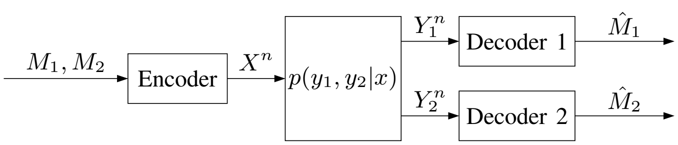

Fig. 1: Broadcast channel with private messages.

The capacity region of this channel is not known in general and there are inner and outer bounds that coincide for several special cases [1]. The first nontrivial inner bound uses superposition coding, which was invented by Cover in his seminal paper on the broadcast channel [2]. The best known inner bound on the capacity region of the broadcast channel is due to Marton [3] and the best known outer bound is due to Nair and El Gamal [4] who showed that it is strictly tighter than the earlier Körner–Marton outer bound [3]. In [5] it was shown that the Nair–El Gamal outer bound is not tight in general. It is not known, however, if the Marton inner bound is optimal in general. Recently, Anantharam, Gohari, and Nair [6] showed that for broadcast channels with binary input, the Marton inner bound reduces to the randomized time division inner bound, which in turn is a special case of the Cover–van der Meulen UVW inner bound [7].

In this paper we establish new optimality results for the randomized time division and superposition coding inner bounds. In the following section, we show that the slope at each corner point of the capacity region of the general irreducible broadcast channel coincides with the slope of the randomized time division (hence the Marton inner bound) inner bound and the Körner–Marton outer bound [3]. This is quite surprising given that randomized time division is in general a special case of the Marton inner bound (and even of the Cover–van der Meulen inner bound).

As pointed out in [8], there are two superposition coding schemes. The first is the UV *scheme*, which is a natural extension of Cover's superposition scheme for the binary symmetric broadcast channel and a special case of the Cover–van der Meulen UVW scheme [7]. The second is the more popular "layered" superposition coding UX (and VX) scheme, also introduced in [2] and formally established by Bergmans [9]. The optimal inner bound on the capacity region of the broadcast channel achieved by the UV scheme was established in [10]. In [8], it was shown that this inner bound includes the inner bound for the UX scheme and that this inclusion can be strict. In Section III, we simplify the superposition coding inner bound in [10]. We then show that for binary input broadcast channels, the UV region reduces to time division, which generalizes the result of Hajek and Pursely's for the skewed binary symmetric broadcast channel [11] and further simplifies the superposition coding inner bound. In Section IV, we establish a simple but general condition under which the superposition coding inner bound is not tight. We then establish necessary and sufficient conditions for the optimality of superposition coding for skewed binary broadcast channels.

II. Optimality of randomized time division inner bound at corner points

In this section, we characterize the slope of the capacity region of the broadcast channel around the corner points. We show that it is equal to the slope of the randomized time

* Chandra Nair is with The Chinese University of Hong Kong (email: chandra@ie.cuhk.edu.hk).

† Hyeji Kim and Abbas El Gamal are with the Department of Electrical Engineering, Stanford University (email: hyejikim@stanford.edu and abbas@ee.stanford.edu). Hyeji Kim is supported by the Alma M. Collins Stanford Graduate Fellowship.

division inner bound and the slope of the Körner–Marton outer bound.

Recall that the *Marton inner bound* [3] is the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &< I(V, W; Y_1), \\ R_2 &< I(U, W; Y_2), \\ R_1 + R_2 &< \min\{I(W; Y_1), I(W; Y_2)\} + I(V; Y_1|W) \\ &\quad + I(U; Y_2|W) - I(U; V|W) \end{aligned} \tag{1}$$

for some pmf $q(u, v, w, x)$. We consider the equivalent weighted sum-rate for $\lambda \in [0, 1]$,

$$\begin{aligned} \lambda R_1 + R_2 = \max_{q(u,v,w,x)} \min_{a\in[0,1]} \Big(&(\lambda - 1)I(V, W; Y_1) \\ &+ (1 - a)I(W; Y_1) + aI(W; Y_2) \\ &+ I(V; Y_1|W) + I(U; Y_2|W) - I(U; V|W)\Big). \end{aligned}$$

We can interchange the max and min operations above (and in similar situations in this paper) by applying Corollary 2 in [5] to yield

$$\begin{aligned} \lambda R_1 + R_2 = \min_{a\in[0,1]} \max_{q(u,v,w,x)} \Big(&(\lambda - 1)I(V, W; Y_1) \\ &+ (1 - a)I(W; Y_1) + aI(W; Y_2) \\ &+ I(V; Y_1|W) + I(U; Y_2|W) - I(U; V|W)\Big). \end{aligned}$$

This characterization can be equivalently represented using the upper concave envelope [12] as

$$\begin{aligned} \lambda R_1 + R_2 = \min_{a\in[0,1]} \max_{q(x)} \Big(&(\lambda - a)I(X; Y_1) + aI(X; Y_2) \\ &+ \mathfrak{C}[-(\lambda - a)I(X; Y_1) - aI(X; Y_2) \\ &+ \max_{q(u,v|x)} \{\lambda I(V; Y_1) + I(U; Y_2) - I(U; V)\}]\Big), \end{aligned}$$

where $\mathfrak{C}[f]$ denotes the upper concave envelope of f (the smallest concave function that is greater than or equal to f).

The *randomized time-division* inner bound [11] can be seen as a restriction of the expression above to the choices of $q(u, v|x)$ governed by $U = X, V = 0$ or $V = X, U = 0$, yielding the achievable weighted sum-rate

$$\begin{aligned} \lambda R_1 + R_2 = \min_{a\in[0,1]} \max_{q(x)} \Big(&(\lambda - a)I(X; Y_1) + aI(X; Y_2) \\ &+ \mathfrak{C}[\max\{aI(X; Y_1) - aI(X; Y_2), \\ &(1 - a)I(X; Y_2) - (\lambda - a)I(X; Y_1)\}]\Big). \end{aligned} \tag{2}$$

Note that for binary input broadcast channels, the randomized time-division region coincides with the Marton's inner bound [6].

The *Nair–El Gamal outer bound* [4] states that if a rate pair (R_1, R_2) is achievable, then it must satisfy the inequalities

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

for some $q(u, v, x)$. This implies that for every achievable (R_1, R_2) and $\lambda \in [0, 1]$,

$$\begin{aligned} \lambda R_1 + R_2 &\le \max_{q(u,x)} \big(I(U; Y_2) + \lambda I(X; Y_1|U)\big) \\ &= \max_{q(x)} \big(I(X; Y_2) + \mathfrak{C}[\lambda I(X; Y_1) - I(X; Y_2)]\big), \end{aligned} \tag{3}$$

and for $\lambda \ge 1$,

$$\begin{aligned} \lambda R_1 + R_2 &\le \max_{q(v,x)} \big(\lambda I(V; Y_1) + I(X; Y_2|V)\big) \\ &\le \max_{q(x)} \big(\lambda I(X; Y_1) + \mathfrak{C}[I(X; Y_2) - \lambda I(X; Y_1)]\big). \end{aligned} \tag{4}$$

Define $C_1 = \max_{q(x)} I(X; Y_1)$ and $C_2 = \max_{q(x)} I(X; Y_2)$. We know that the two rate pairs $(C_1, 0)$ and $(0, C_2)$ lie on the capacity region $\mathcal{C}$. Now define the slope of the capacity region around the corner-point $(C_1, 0)$ as

$$\lambda_1^* = \inf\{\lambda \ge 0 : \lambda R_1 + R_2 \le \lambda C_1, \ \forall (R_1, R_2) \in \mathcal{C}\}.$$

We partition (with respect to the channel $p(y_1|x)$) the set of broadcast channels into two classes:

Class A: There is a capacity achieving pmf q_1^* for the channel $p(y_1|x)$ such that

$$\mathfrak{C}[I(X; Y_1) - I(X; Y_2)]_{q_1^*} = I_{q_1^*}(X; Y_1) - I_{q_1^*}(X; Y_2),$$

Class B: For any capacity achieving pmf q_1^* for the channel $p(y_1|x)$,

$$\mathfrak{C}[I(X; Y_1) - I(X; Y_2)]_{q_1^*} > I_{q_1^*}(X; Y_1) - I_{q_1^*}(X; Y_2),$$

where $\mathfrak{C}[f]_{q_1^*}$ denotes $\mathfrak{C}[f]$ evaluated at q_1^*, and $I_{q_1^*}(X; Y_i)$ denotes the mutual information between X and Y_i for $X \sim q_1^*$, $i = 1, 2$.

Definition 1. A channel $p(y|x)$ is said to be irreducible if any capacity achieving pmf q^* satisfies $q^*(x) > 0$ for all $x \in \mathcal{X}$.

Note that every channel can be expressed as a limit of irreducible channels. Suppose there is an input symbol such that for capacity achieving pmf $q^*(x_0) = 0$. Then introduce a new a output symbol η and change the channel transition matrix to: $p'_{Y|X}(\eta|x_0) = \epsilon, p'_{Y|X}(y|x_0) = (1-\epsilon)p(y|x_0)$ for $y \neq \eta$, leaving other transition probabilities unchanged. Note that any capacity achieving pmf for the new channel must have $q^*(x_0) > 0$ (since, infinite derivative at boundary).

In the following, we characterize the slope of the capacity region λ_1^* around the corner-point $(C_1, 0)$, which is achieved using the randomized time division, under various conditions on the channel.

Theorem 1.

(i) If the broadcast channel belongs to Class A and the channel $p(y_1|x)$ is *irreducible*, then λ_1^* is the smallest $\lambda \in [0, 1]$ such that there exists a capacity achieving pmf q_1^* for the channel $p(y_1|x)$ satisfying

$$\mathfrak{C}[\lambda I(X; Y_1) - I(X; Y_2)]_{q_1^*} = \lambda I_{q_1^*}(X; Y_1) - I_{q_1^*}(X; Y_2).$$

(ii) If the broadcast channel belongs to Class B, then define

$$\hat{\lambda}_1 = \inf\{\lambda \ge 0 : \lambda I(X; Y_1) \ge I(X; Y_2) \ \forall q(x)\}.$$

(In standard fashion, we set $\hat{\lambda}_1 = \infty$ if the set is empty.)

(a) $\lambda_1^* = \hat{\lambda}_1$ if $\hat{\lambda}_1 \geq 1$ and the channel $p(y_1|x)$ is *irreducible*.

(b) If $\hat{\lambda}_1 < 1$, then $\lambda_1^* = 1$. Note that the receiver Y_1 is more capable than the receiver Y_2 and the entire capacity region is characterized in [13].

Proof. We establish the two parts separately.

(i) Assume that the broadcast channel belongs to Class A. Let λ_1' be the smallest $\lambda \in [0,1]$ such that there exists a capacity achieving pmf q_1^* for the channel $p(y_1|x)$ satisfying

$$\mathfrak{C}[\lambda I(X;Y_1) - I(X;Y_2)]_{q_1^*} = \lambda I_{q_1^*}(X;Y_1) - I_{q_1^*}(X;Y_2). \tag{5}$$

To show that $\lambda_1^* \leq \lambda_1'$, we need to show that $\lambda_1' R_1 + R_2 \leq \lambda_1' C_1$ for any $(R_1, R_2) \in \mathcal{C}$. Note that for any $(R_1, R_2) \in \mathcal{C}$ and $\lambda_1' \in [0,1]$, as shown in (3), $\lambda_1' R_1 + R_2 \leq \max_{q(x)}(I(X;Y_2) + \mathfrak{C}[\lambda_1' I(X;Y_1) - I(X;Y_2)])$. Since the right-hand-side is a concave function of $q(x)$, our assumption yields that the outer bound reduces to $\lambda_1' I(X;Y_1)$ at $q(x) = q_1^*$; hence $q(x) = q_1^*$ is a local maximum and thus a global maximum. This implies that for any $(R_1, R_2) \in \mathcal{C}$, one has $\lambda_1' R_1 + R_2 \leq \lambda_1' C_1$.

To show the reverse direction $\lambda_1^* \geq \lambda_1'$ we show that a rate pair (R_1, R_2) such that $\lambda R_1 + R_2 > \lambda C_1$ is achievable for any $\lambda \in [0, \lambda_1')$. *Randomized-time-division*, for $\lambda \in [0,1]$ yields a weighted rate sum (see (2)), given by

$$\begin{aligned}\lambda R_1 + R_2 = \min_{a\in[0,1]} \max_{q(x)} \Big(&(1-\lambda a) I(X;Y_2) + \lambda a I(X;Y_1) \\ &+ \mathfrak{C}[\max\{\lambda a(I(X;Y_2) - I(X;Y_1)), \\ &\lambda(1-a) I(X;Y_1) - (1-\lambda a) I(X;Y_2)\}]\Big).\end{aligned}$$

Suppose the minimum is achieved at $a^* \in [0,1]$ for some $\lambda \in [0, \lambda_1')$, then the weighted rate sum would satisfy

$$\begin{aligned}\lambda R_1 + R_2 &\geq \max_{q(x)} \Big((1-\lambda a^*) I(X;Y_2) + \lambda a^* I(X;Y_1) \\ &\quad + \mathfrak{C}[\lambda(1-a^*) I(X;Y_1) - (1-\lambda a^*) I(X;Y_2)]\Big) \\ &\geq \max_{q(x)} \Big((1-\lambda a^*) I(X;Y_2) + \lambda a^* I(X;Y_1) \\ &\quad + (1-\lambda a^*)\mathfrak{C}[\frac{\lambda(1-a^*)}{1-\lambda a^*} I(X;Y_1) - I(X;Y_2)]\Big) \\ &> (1-\lambda a^*) I_{q_1^*}(X;Y_2) + \lambda a^* I_{q_1^*}(X;Y_1) \\ &\quad + (1-\lambda a^*)\Big(\frac{\lambda(1-a^*)}{1-\lambda a^*} I_{q_1^*}(X;Y_1) - I_{q_1^*}(X;Y_2)\Big) \\ &= \lambda C_1,\end{aligned}$$

where the last inequality holds because for $\lambda \in [0, \lambda_1')$ and for all capacity achieving pmfs q_1^* for the channel $p(y_1|x)$, we have $\mathfrak{C}[\lambda I(X;Y_1) - I(X;Y_2)]_{q_1^*} > \lambda I_{q_1^*}(X;Y_1) - I_{q_1^*}(X;Y_2)$ by the definition of λ_1' in (5).

(ii) Now assume that the broadcast channel belongs to Class B. Let λ_1'' be the smallest $\lambda \in [0, \infty)$ such that

$$\lambda I(X;Y_1) \geq I(X;Y_2) \ \forall q(x).$$

We divide the case into two parts: (a) when $\lambda_1'' \geq 1$ and (b), when when $\lambda_1'' < 1$.

Sub-case (a): We first show that $\lambda_1^* \leq \lambda_1''$. Note that for any $(R_1, R_2) \in \mathcal{C}$ and $\lambda_1'' \geq 1$, as shown in (3), $\lambda_1'' R_1 + R_2 \leq \max_{q(x)} \lambda_1'' I(X;Y_1) + \mathfrak{C}[I(X;Y_2) - \lambda_1'' I(X;Y_1)]$. If $\lambda_1'' I(X;Y_1) \geq I(X;Y_2) \ \forall q(x)$, then clearly $\mathfrak{C}[I(X;Y_2) - \lambda_1'' I(X;Y_1)] = 0$ and $\lambda_1'' R_1 + R_2 \leq \lambda_1'' C_1$.

To show the reverse direction $\lambda_1^* \geq \lambda_1''$, note that since the channel $p(y_1|x)$ is irreducible, q_1^* is an interior point. Hence if there exists some $q(x)$ such that $\lambda I(X;Y_1) < I(X;Y_2)$, then

$$\mathfrak{C}[I(X;Y_2) - \lambda I(X;Y_1)]_{q_1^*} > 0.$$

As stated in (2), randomized-time-division for $\lambda \geq 1$ yields a weighted rate sum

$$\begin{aligned}\lambda R_1 + R_2 = \min_{a\in[0,1]} \max_{q(x)} \Big(&(\lambda - a) I(X;Y_1) + a I(X;Y_2) \\ &+ \mathfrak{C}[\max\{a(I(X;Y_1) - I(X;Y_2)), \\ &(1-a) I(X;Y_2) - (\lambda - a) I(X;Y_1)\}]\Big).\end{aligned}$$

Suppose $a^* \in (0,1]$ is the minimizer; then the weighted sum-rate by randomized-time-division would satisfy

$$\begin{aligned}\lambda R_1 + R_2 &\geq \max_{q(x)} \Big((\lambda - a^*) I(X;Y_1) + a^* I(X;Y_2) \\ &\quad + a^* \mathfrak{C}[I(X;Y_1) - I(X;Y_2)]\Big) \\ &> (\lambda - a^*) I_{q_1^*}(X;Y_1) + a^* I_{q_1^*}(X;Y_2) \\ &\quad + a^*(I_{q_1^*}(X;Y_1) - I_{q_1^*}(X;Y_2)) \\ &= \lambda C_1.\end{aligned}$$

In above the last inequality is due the fact that the broadcast channel belongs to Class B. (The inequality is only strict for $a^* \neq 0$).

If $a^* = 0$ is the minimizer, then the weighted sum-rate by randomized-time-division would satisfy

$$\begin{aligned}\lambda R_1 + R_2 &= \max_{q(x)} \big(\lambda I(X;Y_1) \ + \mathfrak{C}[I(X;Y_2) - \lambda I(X;Y_2)]\big) \\ &> \lambda C_1,\end{aligned}$$

where the last inequality is a consequence of having $\mathfrak{C}[I(X;Y_2) - \lambda I(X;Y_1)]_{q_1^*} > 0$.

Sub-case (b): As mentioned earlier, the entire capacity region is established in [13] and it can be easily seen that for $\lambda \leq 1$

$$\begin{aligned}\max_{(R_1,R_2)\in\mathcal{C}} (\lambda R_1 + R_2) = \min_{a\in[0,1]} \max_{q(x)} \Big(&(1-\lambda a) I(X;Y_2) \\ + \lambda a I(X;Y_1) + \mathfrak{C}[\lambda(1-a) I(X;Y_1) &- (1-\lambda a) I(X;Y_2)]\Big).\end{aligned} \tag{6}$$

Since the channel is in Class B, it is easy to see that for any capacity achieving pmf q_1^* and $\lambda < 1$,

$$\begin{aligned}&\mathfrak{C}[\lambda(1-a) I(X;Y_1) - (1-\lambda a) I(X;Y_2)]_{q_1^*} \\ &\quad > \lambda(1-a) I_{q_1^*}(X;Y_1) - (1-\lambda a) I_{q_1^*}(X;Y_2).\end{aligned}$$

Plugging this into (6), we obtain that for any $\lambda < 1$, $\max_{(R_1,R_2)\in\mathcal{C}} (\lambda R_1 + R_2) > \lambda C_1$. On the other hand, we have $\max_{(R_1,R_2)\in\mathcal{C}} (R_1 + R_2) = C_1$, establishing $\lambda_1^* = 1$. □

Remark 1. The following observations are pertinent here.

- There is two extremal cases, in both the channel $p(y_1|x)$ is reducible, that are not covered by Theorem 1. First is when the broadcast channel belongs to class A, and the other where the broadcast channel belonging to Class B, and $\hat{\lambda}_1 \geq 1$.
- There is an as yet unpublished result by Salman Beigi [14] which characterizes when Time-Division strategy matches the capacity region for broadcast channels. While this condition can be inferred from the slope result mentioned above; the first author is sure that some of the ideas for the slope characterization did permeate during the discussions with Salman Beigi.

III. Simplified characterization of superposition coding inner bound

In this section, we present a simplified characterization of the superposition coding inner bound in [10]. We then simplify this characterization further for binary input broadcast channels.

Consider the UV superposition coding inner bound on the capacity of the broadcast channel. We fix a pmf $q(u, v, x) = \{q(u)q(v),\ x(u, v)\}$. We randomly and independently generate 2^{nR_2} sequences $u^n(m_2) \sim \prod_{i=1}^n p_U(u_i)$, $m_2 \in [1 : 2^{nR_2}]$, and 2^{nR_1} sequences $v^n(m_1) \sim \prod_{i=1}^n p_V(v_i)$, $m_1 \in [1 : 2^{nR_1}]$. To send (m_1, m_2), the sender transmits $x(u_i(m_2), v_i(m_1))$ for $i \in [1 : n]$.

In [10], Bandemer, El Gamal, and Kim showed that simultaneous nonunique decoding is optimal for this scheme. Further, it is shown that the limit of the average probability of error for this scheme approaches zero if (R_1, R_2) is in one of the following four regions:

$\mathcal{R}_{UX}(q)$: The set of rate pairs (R_1, R_2) such that

$$R_2 < I(U;Y_2), R_1 < I(X;Y_1|U), R_1 + R_2 < I(X;Y_1). \tag{7}$$

$\mathcal{R}_{VX}(q)$: The set of rate pairs (R_1, R_2) such that

$$R_1 < I(V;Y_1), R_2 < I(X;Y_2|V), R_1 + R_2 < I(X;Y_2). \tag{8}$$

$\mathcal{R}_{UV}(q)$: The set of rate pairs (R_1, R_2) such that

$$R_1 < I(V;Y_1), R_2 < I(U;Y_2). \tag{9}$$

$\mathcal{R}_{XX}(q)$: Set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 < I(X;Y_2|U), R_2 < I(X;Y_1|V), \\ R_1 + R_2 < \min\{I(X;Y_1), I(X;Y_2)\}. \end{aligned} \tag{10}$$

The optimal superposition coding inner bound in [10] can be characterized as:

$$\mathcal{R}^* = \mathrm{co}\left(\cup_{q(u,v,x)}\left(\mathcal{R}_{UX}(q) \cup \mathcal{R}_{VX}(q) \cup \mathcal{R}_{UV}(q) \cup \mathcal{R}_{XX}(q)\right)\right).$$

Now consider the following four inner bounds

$$\begin{aligned} \mathcal{R}_{UX} &= \cup_{q(u,v,x)}\mathcal{R}_{UX}(q), \quad \mathcal{R}_{VX} = \cup_{q(u,v,x)}\mathcal{R}_{VX}(q), \\ \mathcal{R}_{UV} &= \mathrm{co}(\cup_{q(u,v,x)}\mathcal{R}_{UV}(q)), \quad \mathcal{R}_{XX} = \mathrm{co}(\cup_{q(u,v,x)}\mathcal{R}_{XX}(q)). \end{aligned}$$

Note that $\mathcal{R}_{UX}$ and $\mathcal{R}_{VX}$ are the inner bounds of the "layered" superposition coding scheme established in [2, 9] and $\mathcal{R}_{UV}$ is the Cover–van der Meulen inner bound for the general broadcast channel without the common auxiliary random variable W [7, 15].

It can be easily seen that $\mathcal{R}^*$ can be equivalently written as:

$$\mathcal{R}^* = \mathrm{co}(\mathcal{R}_{UX} \cup \mathcal{R}_{VX} \cup \mathcal{R}_{UV} \cup \mathcal{R}_{XX}). \tag{11}$$

We are now ready to establish our first result, which simplifies the characterization of $\mathcal{R}^*$ in (11).

Theorem 2. The optimal superposition coding inner bound for the broadcast channel with $C_1 \geq C_2$ is

$$\mathcal{R}^* = \mathrm{co}\left(\mathcal{R}_{UX} \cup \mathcal{R}_{UV}\right).$$

Proof. We show that $(\mathcal{R}_{VX} \cup \mathcal{R}_{XX}) \subseteq \mathcal{R}_{UV}$.

Let $(R_1, R_2) \in (\mathcal{R}_{VX} \cup \mathcal{R}_{XX})$. Then (R_1, R_2) satisfies $R_1/C_1 + R_2/C_2 \leq 1$ because for any rate pair $(R_1, R_2) \in \mathcal{R}_{VX} \cup \mathcal{R}_{XX}$, $R_1 + R_2 \leq C_2$ from the inequalities on the sum rate in $\mathcal{R}_{VX}$ and $\mathcal{R}_{XX}$, and $R_1/C_1 + R_2/C_2 \leq R_1/C_2 + R_2/C_2$.

We now show that $\mathcal{R}_{UV}$ contains any rate pair (R_1, R_2) such that $R_1/C_1 + R_2/C_2 \leq 1$. Note that $(R_1, R_2) = (C_1, 0)$ satisfies the inequalities in $\mathcal{R}_{UV}$ for $q_1^* = \arg\max_{q(x)} I(X;Y_1)$ and $(U, V) = (\emptyset, X)$. Similarly, $(R_1, R_2) = (0, C_2)$ satisfies the inequalities in $\mathcal{R}_{UV}$ for $q_2^* = \arg\max_{q(x)} I(X;Y_2)$ and $(U, V) = (X, \emptyset)$. Thus for any $\gamma \in [0, 1]$, $(\gamma C_1, (1 - \gamma)C_2) \in \mathcal{R}_{UV}$, i.e., $\mathcal{R}_{UV}$ includes any rate pair (R_1, R_2) such that $R_1/C_1 + R_2/C_2 \leq 1$. □

It is not difficult to show via examples that $\mathcal{R}_{UX}$ does not always include $\mathcal{R}_{UV}$ nor is always included in it. For example for a binary symmetric broadcast channel (with $0 < p_1 < p_2 < 0.5$) [2], $\mathcal{R}_{UV} \subset \mathcal{R}_{UX}$, since $\mathcal{R}_{UX}$ is the capacity region while $\mathcal{R}_{UV}$ is the time division region. On the other hand, for the simple binary vector broadcast channel in [8], $\mathcal{R}_{UX} \subset \mathcal{R}_{UV}$.

Theorem 3. For binary input broadcast channels,

$$\mathcal{R}_{UV} = \mathcal{R}_{\mathrm{TD}},$$

where $\mathcal{R}_{\mathrm{TD}}$ is the set of rate pairs (R_1, R_2) such that $R_1/C_1 + R_2/C_2 < 1$.

Proof. The key to this is an information inequality, established in [16], that for all random variables satisfying $(U, V) \to X \to (Y_1, Y_2)$ with $|\mathcal{X}| = 2$, the following inequality holds:

$$I(U;Y_1) + I(V;Y_2) - I(U;V) \leq \max\{I(X;Y_1), I(X;Y_2)\}. \tag{12}$$

The details can be found in the full version at [17] □

A direct result of Theorem 3 is that the superposition inner bound in (11) can be further simplified for binary input broadcast channels.

Corollary 1. For binary input broadcast channels with $C_1 \geq C_2$,

$$\mathcal{R}^* = \mathrm{co}\left(\mathcal{R}_{UX} \cup \{(0, C_2)\}\right).$$

IV. Optimality of superposition coding inner bound

It is known that superposition coding is optimal for several classes of broadcast channels, including when it is more capable [18] or effectively less noisy [19]. In this section we first present a general condition under which superposition coding is not optimal, then establish necessary and sufficient conditions under which superposition coding is optimal for the skewed binary broadcast channels.

Proposition 1. For a broadcast channel with $C_1 \geq C_2$, if $(0, C_2) \notin \mathcal{R}_{UX}$ and $I(X;Y_1)$ and $I(X;Y_2)$ are strictly concave in $q(x)$, then superposition coding is suboptimal.

Proof. The proof can be found in the full version [17]. □

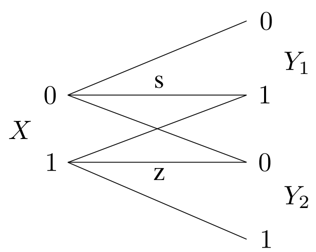

Fig. 2: Skewed binary broadcast channel.

In [19], it was shown that superposition coding is optimal in the shaded area in Figure 3, which corresponds to the case where the superposition rate region coincides with the Körner–Marton outer bound without the sum rate inequality characterized as follows. Theorem 4 shows that superposition coding is only optimal for that region.

Definition 2 (Beyond effectively less noisy [19])**.** For a DM-BC, $p(y_1, y_2|x)$, let $\mathcal{Q}(u,x)$ be a set of pmfs $q(u,x)$ such that for every $\lambda \geq 1$,

$$\max_{q(u,x)} \big(\lambda I(U;Y_2) + I(X;Y_1|U)\big) = \max_{q(u,x)\in\mathcal{Q}(u,x)} \big(\lambda I(U;Y_2) + I(X;Y_1|U)\big). \quad (13)$$

Receiver Y_1 is said to be *beyond effectively less noisy* than receiver Y_2 if there exists a $\mathcal{Q}(u,x)$ such that $I(U;Y_1) \geq I(U;Y_2)$ for every $q(u,x) \in \mathcal{Q}(u,x)$.

Theorem 4. For a skewed binary broadcast channel, superposition coding is optimal if and only if Receiver Y_1 is *beyond effectively less noisy* than receiver Y_2.

Proof. The proof can be found in the full version [17]. □

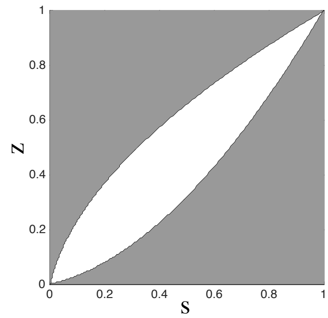

Fig. 3: The shaded area is where superposition coding is optimal for (s, z) SBBC.

References

[1] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge, 2011.

[2] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[3] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, 1979.

[4] 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.

[5] 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.

[6] V. Anantharam, A. Gohari, and C. Nair, "Improved cardinality bounds on the auxiliary random variables in Marton's inner bound," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2013, pp. 1272–1276.

[7] E. C. van der Meulen, "Random coding theorems for the general discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 21, no. 2, pp. 180–190, 1975.

[8] L. Wang, E. Sasoglu, B. Bandemer, and Y.-H. Kim, "A comparison of superposition coding schemes," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2013, pp. 2970–2974.

[9] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 197–207, 1973.

[10] B. Bandemer, A. El Gamal, and Y.-H. Kim, "Optimal achievable rates for interference networks with random codes," *Information Theory, IEEE Transactions on (accepted for publication)*, 2015.

[11] B. E. 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, 1979.

[12] C. Nair, "Upper concave envelopes and auxiliary random variables," *Int. J. Adv. Eng. Sci. Appl. Math*, no. 1, pp. 12–20.

[13] A. El Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 379–381, 1978.

[14] S. Beigi, *Personal communication*.

[15] T. M. Cover, "An achievable rate region for the broadcast channel," *IEEE Trans. Inf. Theory*, vol. 21, no. 4, pp. 399–404, 1975.

[16] Y. Geng, V. Jog, C. Nair, and Z. 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, July 2013.

[17] C. Nair, H. Kim, and A. El Gamal, "On the optimality of randomized time division and superposition coding for the broadcast channel." [Online]. Available: http://stanford.edu/∼hyejikim/optimality2.pdf

[18] A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 166–169, 1979.

[19] H. Kim, B. Nachman, and A. El Gamal, "Superposition coding is almost always optimal for the poisson broadcast channel," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 1782–1794, April 2016.