A Two Auxiliary Receiver Outer Bound to the Capacity Region of a Two-Receiver Discrete Memoryless Broadcast Channel

Amin Gohari, Yi Liu and Chandra Nair
Department of Information Engineering
The Chinese University of Hong Kong
Sha Tin, N.T., Hong Kong
Email: {agohari, ly023, chandra}@ie.cuhk.edu.hk

Abstract

We analyze the inclusion-relationship among existing outer bounds for the sum-rate of the two-receiver broadcast channel, and propose a new outer bound employing two auxiliary receivers. This new bound provides strict improvements upon existing results, yielding tighter upper bounds for certain broadcast channels.

Index Terms

Shannon theory, outer bound, broadcast channel, auxiliary receiver

I. Introduction

We denote a two-receiver memoryless broadcast channel [4] by $(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, T_{Y,Z|X})$, where $\mathcal{X}$ is the input alphabet, $\mathcal{Y}$ and $\mathcal{Z}$ are the output alphabets, and $T_{Y,Z|X}$ is the collection of conditional pmfs $T(y,z|x)$ on $\mathcal{Y} \times \mathcal{Z}$, given each input symbol x. In this communication scenario, a transmitter encodes one common message M_0 and two private messages M_1, M_2 into a codeword X^n, which is then transmitted over the channel. Given channel outputs Y^n and Z^n, receiver Y can reliably decode (M_0, M_1) at rates (R_0, R_1), while receiver Z can reliably decode (M_0, M_2) at rates (R_0, R_2). The model is depicted in Figure 1; a detailed definition can be found in [6].

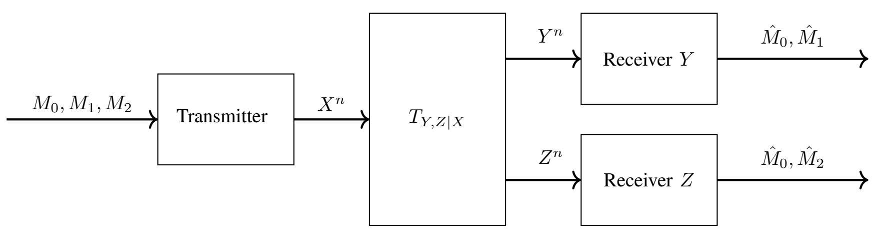

Fig. 1. Two-receiver broadcast communication system.

The capacity region of a general broadcast channel remains unknown. Marton's inner bound [14] is the best achievable region and coincides with the capacity region for a few classes of broadcast channels [7], [8], [15], [17], [18]. See [1], [2], [10], [11] for a series of works on the computability and optimality of Marton's inner bound.

Theorem 1 (Marton '79)**.** *The union of non-negative rate triples (R_0, R_1, R_2) satisfying the constraints*

$$\begin{aligned} R_0 &\le \min\{I(W;Y), I(W;Z)\}, \\ R_0 + R_1 &\le I(U,W;Y), \\ R_0 + R_2 &\le I(V,W;Z), \\ R_0 + R_1 + R_2 &\le \min\{I(W;Y), I(W;Z)\} + I(U;Y|W) + I(V;Z|W) - I(U;V|W), \end{aligned}$$

for any triple of random variables (U,V,W) such that $(U,V,W) \multimap X \multimap (Y,Z)$ is achievable.

Two famous outer bounds are Körner-Marton outer bound [14] and UVW outer bound [16].

Theorem 2 (UVW Outer Bound)**.** *Any achievable rate triple (R_0, R_1, R_2) must satisfy the constraints*

$$\begin{aligned} R_0 &\le \min\{I(W;Y), I(W;Z)\}, \\ R_0 + R_1 &\le \min\{I(W;Y), I(W;Z)\} + I(U;Y|W), \end{aligned}$$

$$
\begin{aligned}
R_0 + R_2 &\le \min\{I(W;Y), I(W;Z)\} + I(V;Z|W),\\
R_0 + R_1 + R_2 &\le \min\{I(W;Y), I(W;Z)\}\\
&\quad + I(U;Y|W) + I(X;Z|U,W),\\
R_0 + R_1 + R_2 &\le \min\{I(W;Y), I(W;Z)\}\\
&\quad + I(V;Z|W) + I(X;Y|V,W),
\end{aligned}
$$

for some triple of random variables (U, V, W) *such that* $(U, V, W) \multimap X \multimap (Y, Z)$. *Further, it suffices to consider* (U, V, W) *satisfying* $|\mathcal{W}| \le |\mathcal{X}| + 5, |\mathcal{U}| \le |\mathcal{X}| + 1, |\mathcal{V}| \le |\mathcal{X}| + 1$.

Recently, a number of outer bounds employing the "auxiliary receiver approach" have been proposed [9], [12]. These bounds utilize "add-and-subtract" and "genie-based" manipulations, which represent two different ways of introducing auxiliary receivers into the outer bounds. In this paper, we examine the relationships among these outer bounds. Additionally, we propose a new outer bound that incorporates two auxiliary receivers, yielding a strictly tighter bound for certain broadcast channels.

Auxiliary receivers are mathematical tools used to study capacity regions in communication problems. They are new artificial receivers that are absent from the original physical model; they function as silent observers that neither communicate nor decode messages. They have been successfully applied to various problems in network information theory, such as the relay channel and the interference channel (see, e.g., [3], [5], [12], [13], [19]). Below, we illustrate how these auxiliary receivers reveal the geometric curvature of the capacity region's surface, facilitating the derivation of stronger outer bounds.

Let $\mathsf{SR}\big(T_{Y,Z|X}\big) \in [0, \infty)$ be the maximum sum-rate $R_1 + R_2$ subject to the given fixed input distribution $p_X(x)$, i.e., the empirical distributions of the codewords converge to $p_X(x)$. Define, $\mathcal{P}_n$ to be all distributions of the form

$$
p_{M_1,M_2,X^n} = p_{M_1} p_{M_2} p_{X^n|M_1,M_2},
$$

where the alphabets of M_1 and M_2 are arbitrary. Moreover, p_{X^n} must satisfy the average input distribution constraint:

$$
(1/n) \sum_{i=1}^{n} p_{X_i}(x) = p_X(x).
$$

Further, let $\mathcal{P} = \cup \mathcal{P}_n$.

The following multi-letter formula for $\mathsf{SR}\big(T_{Y,Z|X}\big)$ is immediate:

$$
\mathsf{SR}\big(T_{Y,Z|X}\big) = \sup_{\mathcal{P}} \frac{1}{n} \left(I(M_1; Y^n) + I(M_2; Z^n) \right).
$$

For the special case of $Y = Z = J$, we know that $\mathsf{SR}\big(T_{J,J|X}\big) = I(X; J)$ since J can decode both messages. In this case, we can write an "apparently" better multi-letter formula:

$$
\mathsf{SR}\big(T_{J,J|X}\big) = \sup_{\text{codes}} \frac{1}{n} \left(I(M_1; J^n) + I(M_2; J^n) + \sum_{i=1}^{n} I(J_i; J^{i-1}) \right).
$$

This is because, for every $P_{M_1,M_2,X^n} \in \mathcal{P}_n$, we have

$$
\begin{aligned}
I(M_1; J^n) + I(M_2; J^n) + \sum_{i=1}^{n} I(J_i; J^{i-1}) &\le I(M_1, M_2; J^n) + \sum_{i=1}^{n} I(J_i; J^{i-1})\\
&\le I(X^n; J^n) + \sum_{i=1}^{n} I(J_i; J^{i-1})\\
&\overset{(a)}{=} \sum_{i=1}^{n} I(X_i; J_i)\\
&\le nI(X; J) = n \cdot \mathsf{SR}\big(T_{J,J|X}\big),
\end{aligned}
$$

where (a) uses the memoryless property of the channel and the last step follows from the constraint $(1/n) \sum_{i=1}^{n} p_{X_i}(x) = p_X(x)$ and the concavity of mutual information in the input distribution. On the other hand, setting X^n to be i.i.d. according to p_X, with $M_1 = X^n$ and M_2 being constant, yields

$$
\sup_{\text{codes}} \frac{1}{n} \left(I(M_1; J^n) + I(M_2; J^n) + \sum_{i=1}^{n} I(J_i; J^{i-1}) \right) \ge I(X; J) = \mathsf{SR}\big(T_{J,J|X}\big).
$$

Next, by applying the "add-and-subtract" idea of the auxiliary receiver approach, we can express

$$
\mathsf{SR}\big(T_{Y,Z|X}\big) = \mathsf{SR}\big(T_{J,J|X}\big) + \mathsf{SR}\big(T_{Y,Z|X}\big) - \mathsf{SR}\big(T_{J,J|X}\big)
$$

$$= I(X;J) + \mathsf{SR}\big(T_{Y,Z|X}\big) - \mathsf{SR}\big(T_{J,J|X}\big).$$

We have

$$\begin{aligned}\mathsf{SR}\big(T_{Y,Z|X}\big) - \mathsf{SR}\big(T_{J,J|X}\big) &= \sup_{\text{codes}} \frac{1}{n}\Big(I(M_1;Y^n) + I(M_2;Z^n)\Big) - \sup_{\text{codes}} \frac{1}{n}\left(I(M_1;J^n) + I(M_2;J^n) + \sum_{i=1}^{n} I(J_i;J^{i-1})\right) \\ &\leq \sup_{\text{codes}} \frac{1}{n}\left(I(M_1;Y^n) + I(M_2;Z^n) - I(M_1;J^n) - I(M_2;J^n) - \sum_{i=1}^{n} I(J_i;J^{i-1})\right)\end{aligned}$$

Observe that

$$\begin{aligned}I(M_1;Y^n) - I(M_1;J^n) &= \sum_{i=1}^{n} I(\hat{U}_i;Y_i|\hat{W}_i) - I(\hat{U}_i;J_i|\hat{W}_i) \\ I(M_2;Z^n) - I(M_2;J^n) &= \sum_{i=1}^{n} I(\tilde{V}_i;Z_i|\tilde{W}_i) - I(\tilde{V}_i;J_i|\tilde{W}_i)\end{aligned}$$

where $\hat{U}_i = \tilde{U}_i = M_1$, $\hat{V}_i = \tilde{V}_i = M_2$, $\hat{W}_i = (Y^{i-1}, J_{i+1}^n)$ and $\tilde{W}_i = (Z^{i-1}, J_{i+1}^n)$. The term $-\sum_{i=1}^n I(J_i;J^{i-1})$ can be manipulated as follows:

$$-\sum_{i=1}^{n} I(J_i;J^{i-1}) \leq -\sum_{i=1}^{n} I(J_i;J^{i-1}) + \sum_{i=1}^{n} I(Y_i;Y^{i-1}) = \sum_{i=1}^{n} I(\hat{W}_i;Y_i) - I(\hat{W}_i;J_i).$$

Similarly, we have:

$$-\sum_{i=1}^{n} I(J_i;J^{i-1}) \leq \sum_{i=1}^{n} I(\tilde{W}_i;Z_i) - I(\tilde{W}_i;J_i).$$

Using a time-sharing random variable Q, and letting $\tilde{W} = (\tilde{W}_Q, Q)$ and $\hat{W} = (\hat{W}_Q, Q)$, etc., we derive the following upper bound:

$$\begin{aligned}\mathsf{SR}\big(T_{Y,Z|X}\big) \leq & I(X;J) + I(\hat{U};Y|\hat{W}) - I(\hat{U};J|\hat{W}) + I(\tilde{V};Z|\tilde{W}) - I(\tilde{V};J|\tilde{W}) \\ & + \min(I(\hat{W};Y) - I(\hat{W};J), I(\tilde{W};Z) - I(\tilde{W};J))\end{aligned} \tag{1}$$

for some $p_{\hat{U},\hat{W},\tilde{U},\tilde{W},X}T_{J,Y,Z|X}$. This upper bound is valid for any arbitrary $T_{J|XYZ}$. Setting $J = Y$ or $J = Z$ yields the two UVW outer bound sum-rate constraints. But one can choose $J \notin \{Y, Z\}$ to obtain strict improvements over the UVW outer bound.

The above manipulation appears in [12, Theorem 7] for the full capacity region, with more constraints on the individual and sum rates, and the auxiliary random variables. Note that in this manipulation (termed a J-bound manipulation), we move from $\mathsf{SR}\big(T_{Y,Z|X}\big)$ to $\mathsf{SR}\big(T_{J,J|X}\big)$. The key new idea of this paper is to move from $\mathsf{SR}\big(T_{Y,Z|X}\big)$ to $\mathsf{SR}\big(T_{G,K|X}\big)$, allowing for even better bounds.

This paper is organized as follows: in Section II, we review existing auxiliary receiver-based outer bounds. Section III focuses on the sum-rate constraints, providing comparisons among these bounds and introducing a new outer bound utilizing two auxiliary receivers. Finally, in Section IV, we present numerical evidence demonstrating that our new bound offers superior performance compared to all previous outer bounds for a binary skew-symmetric channel.

II. Known bounds

In this subsection, we list the three known outer bounds for broadcast channels in the literature that utilize auxiliary receivers. The first result is a bound with one auxiliary receiver J. The core idea is to transition from $T_{Y,Z|X}$ to $T_{J,J|X}$ using "add-and-subtract" manipulations, and subsequently compare the capacity region of $T_{Y,Z|X}$ with that of $T_{J,J|X}$:

Theorem 3 (Theorem 7, [12]). *Given a broadcast channel characterized by $T_{Y,Z|X}$ and any achievable rate triple (R_0, R_1, R_2), one can find some input distribution $p(x)$ such that for any auxiliary channel $T_{J|X,Y,Z}$, the following constraints are satisfied:*

$$R_0 \leq \min\{I(W;Y), I(\hat{W};Y), I(W;Z), I(\tilde{W};Z)\}, \tag{2a}$$

$$R_0 + R_1 \leq \min\{I(W;Y), I(W;Z)\} + I(U;Y|W), \tag{2b}$$

$$\begin{aligned}R_0 + R_1 \leq \min & \left\{ I(\tilde{W};Z) + \min\big[0, I(W;Y) - I(W;Z)\big], \quad I(\tilde{W};J) + I(\hat{W};Y) - I(\hat{W};J) \right\} \\ & + I(\tilde{U};J|\tilde{W}) + I(\hat{U};Y|\hat{W}) - I(\hat{U};J|\hat{W}),\end{aligned} \tag{2c}$$

$$R_0+R_1 \le \min\Big\{I(\hat{W};Y)+\min\big[0, I(W;Z)-I(W;Y)\big],\quad I(\hat{W};J)+I(\tilde{W};Z)-I(\tilde{W};J)\Big\} + I(\hat{U};Y|\hat{W}), \tag{2d}$$

$$R_0+R_2 \le \min\{I(W;Y), I(W;Z)\}+I(V;Z|W), \tag{2e}$$

$$R_0+R_2 \le \min\Big\{I(\hat{W};Y)+\min\big[0, I(W;Z)-I(W;Y)\big],\quad I(\hat{W};J)+I(\tilde{W};Z)-I(\tilde{W};J)\Big\} + I(\hat{V};J|\hat{W})+I(\tilde{V};Z|\tilde{W})-I(\tilde{V};J|\tilde{W}), \tag{2f}$$

$$R_0+R_2 \le \min\Big\{I(\tilde{W};Z)+\min\big[0, I(W;Y)-I(W;Z)\big],\quad I(\tilde{W};J)+I(\hat{W};Y)-I(\hat{W};J)\Big\} + I(\tilde{V};Z|\tilde{W}), \tag{2g}$$

$$R_0+R_1+R_2 \le \min\Big\{I(\hat{W};Y)-I(\hat{W};J),\quad I(\tilde{W};Z)-I(\tilde{W};J)\Big\}+I(X;J) + I(\hat{U};Y|\hat{W})-I(\hat{U};J|\hat{W})+I(\tilde{V};Z|\tilde{W})-I(\tilde{V};J|\tilde{W}), \tag{2h}$$

$$R_0+R_1+R_2 \le \min\Big\{I(W;Y), I(W;Z)\Big\} + \min\Big\{I(V;Z|W)+I(X;Y|V,W), I(U;Y|W)+I(X;Z|U,W)\Big\}, \tag{2i}$$

for some choice of distribution over the variables

$$p_{U,V,W,\tilde{U},\tilde{V},\tilde{W},\hat{U},\hat{V},\hat{W},X,Y,Z,J} = p_{U,V,W,X}\, p_{\tilde{W},\tilde{U},\tilde{V}|X}\, p_{\hat{W},\hat{U},\hat{V}|X}\, T_{Y,Z|X}\, T_{J|X,Y,Z}$$

satisfying

$$I(\tilde{W};Z)-I(\tilde{W};J) + I(\hat{W};J)-I(\hat{W};Y) = I(W;Z)-I(W;Y), \tag{3a}$$

$$I(\tilde{U};Z|\tilde{W})-I(\tilde{U};J|\tilde{W}) + I(\hat{U};J|\hat{W})-I(\hat{U};Y|\hat{W}) = I(U;Z|W)-I(U;Y|W), \tag{3b}$$

$$I(\tilde{V};Z|\tilde{W})-I(\tilde{V};J|\tilde{W}) + I(\hat{V};J|\hat{W})-I(\hat{V};Y|\hat{W}) = I(V;Z|W)-I(V;Y|W), \tag{3c}$$

and

$$0 \le I(X;Z|\tilde{U},\tilde{W})-I(X;J|\tilde{U},\tilde{W}) \le I(\tilde{V};Z|\tilde{W})-I(\tilde{V};J|\tilde{W}), \tag{4a}$$

$$0 \le I(X;Y|\hat{V},\hat{W})-I(X;J|\hat{V},\hat{W}) \le I(\hat{U};Y|\hat{W})-I(\hat{U};J|\hat{W}), \tag{4b}$$

$$I(V;Z|W)+I(X;Y|V,W) = I(U;Y|W)+I(X;Z|U,W). \tag{4c}$$

Remark 1. The outer bound depends on J only through $T_{J|X}$, $T_{Y|X}$ and $T_{Z|X}$. Therefore, without loss of generality, we may assume $T_{Y,Z,J|X}=T_{Y|X}T_{Z|X}T_{J|X}$.

Remark 2. The choice of auxiliary random variables are

$$\hat{W}=(M_0, J^{Q-1}, Y_{Q+1}^n, Q), \tag{5a}$$

$$\tilde{W}=(M_0, Z^{Q-1}, J_{Q+1}^n, Q), \tag{5b}$$

$$W=(M_0, Z^{Q-1}, Y_{Q+1}^n, Q), \tag{5c}$$

$$U=\hat{U}=\tilde{U}=M_1, \tag{5d}$$

$$V=\hat{V}=\tilde{V}=M_2. \tag{5e}$$

Remark 3. Note that equation (1) is the same as (2h). It is also possible to replace $\min\big\{I(\hat{W};Y)-I(\hat{W};J),\quad I(\tilde{W};Z)-I(\tilde{W};J)\big\}$ with $\min\big\{0, I(\hat{W};Y)-I(\hat{W};J),\quad I(\tilde{W};Z)-I(\tilde{W};J)\big\}$ in (2h).

The second outer bound is a (G,K)-bound whose idea is to move from $T_{Y,Z|X}$ to $T_{GY,ZK|X}$. Observe that G and K play the role of free side information (genie).

Theorem 4 (Theorem 8, [12])**.** *Given a broadcast channel $T_{Y,Z|X}$, any achievable non-negative rate triple (R_0,R_1,R_2) must satisfy the following constraints*

$$R_0 \le \min\{I(W_b;G)+I(W_a;Y|G), I(W_b;Z|K)+I(W_a;K)\}, \tag{6a}$$

$$R_0+R_1 \le I(U_b,W_b;G)+I(U_a,W_a;Y|G), \tag{6b}$$

$$R_0+R_1 \le I(W_b;Z|K)+I(W_a,G;K)+I(U_b;G|W_b,K)+I(U_a;Y|W_a,G), \tag{6c}$$

$$R_0 + R_2 \le I(W_b, K; G) + I(W_a; Y|G) + I(V_b; Z|W_b, K) + I(V_a; K|W_a, G), \tag{6d}$$

$$R_0 + R_2 \le I(V_a, W_a; K) + I(V_b, W_b; Z|K), \tag{6e}$$

$$\begin{aligned} R_0 + R_1 + R_2 \le \min\{&I(W_b, K; G) + I(W_a; Y|G),\ I(W_b; Z|K) + I(W_a, G; K)\} \\ &+ I(U_a; Y|W_a, G) + I(X; K|U_a, W_a, G) \\ &+ \min\big\{I(U_b; G|W_b, K) + I(X; Z|U_b, W_b, K),\ I(V_b; Z|W_b, K) + I(X; G|V_b, W_b, K)\big\}, \end{aligned} \tag{6f}$$

$$\begin{aligned} R_0 + R_1 + R_2 \le \min\{&I(W_b, K; G) + I(W_a; Y|G),\ I(W_b; Z|K) + I(W_a, G; K)\} \\ &+ I(V_b; Z|W_b, K) + I(X; G|V_b, W_b, K) \\ &+ \min\big\{I(U_a; Y|W_a, G) + I(X; K|U_a, W_a, G),\ I(V_a; K|W_a, G) + I(X; Y|V_a, W_a, G)\big\}, \end{aligned} \tag{6g}$$

for any $T_{G,K|X,Y,Z}$ and some $p(w_a, v_a, u_a|x)p(w_b, v_b, u_b|x)p(x)$ satisfying $|\mathcal{W}_b|, |\mathcal{W}_a| \le |\mathcal{X}| + 7$, $|\mathcal{U}_b|, |\mathcal{V}_a| \le |\mathcal{X}| + 2$, $|\mathcal{V}_b|, |\mathcal{U}_a| \le |\mathcal{X}| + 1$.

The third outer bound is a J-bound with one auxiliary receiver:

Theorem 5 (Theorem 1, [9])**.** *For any $\mu > 0$ and for any $T_{J|XYZ}$, any achievable rate $(0, R_1, R_2)$ satisfies*

$$\begin{aligned} R_1 + \mu R_2 \le &\max_{p(x)} \min_{T_{J|X}, \sigma\in[0,\min\{1,\mu\}], \gamma\in[0,1]} (1-\sigma\gamma)I(X;Y) + (\mu - \sigma\bar{\gamma})I(X;Z) \\ &+ \mathfrak{C}\big[\mathfrak{C}[\sigma I(X;J) - I(X;Y)] + \sigma\gamma(I(X;Y) - I(X;J))\big] \\ &+ \mathfrak{C}\big[\mathfrak{C}[\sigma I(X;J) - \mu I(X;Z)] + \sigma\bar{\gamma}(I(X;Z) - I(X;J))\big], \end{aligned} \tag{7}$$

where $\bar{\gamma} = 1 - \gamma$, and $\mathfrak{C}[f(x)]$ denotes the upper concave envelope of a function $f(x)$.

III. Main results

Proposition 1. *For every $T_{J|XYZ}$, the outer bound in Theorem 3 implies the outer bound in Theorem 5 with the same choice for the auxiliary receiver.*

Proof. Letting $R_0 = 0$, and using (2d), (2g) and (2h), we obtain

$$R_1 \le I(\hat{U}, \hat{W}; Y), \tag{8a}$$

$$R_2 \le I(\tilde{V}, \tilde{W}; Z), \tag{8b}$$

$$R_1 + R_2 \le I(X; J|\hat{U}, \hat{W}) + I(\hat{U}, \hat{W}; Y) + I(\tilde{V}; Z|\tilde{W}) - I(\tilde{V}; J|\tilde{W}), \tag{8c}$$

$$R_1 + R_2 \le I(X; J|\tilde{V}, \tilde{W}) + I(\hat{U}; Y|\hat{W}) - I(\hat{U}; J|\hat{W}) + I(\tilde{V}, \tilde{W}; Z). \tag{8d}$$

Take some arbitrary $\mu > 0$ and $\sigma \in [0, \min\{1, \mu\}]$, $\gamma \in [0, 1]$. Now, multiply (8a) by ω_1, (8b) by ω_2, (8c) by ω_3 and (8d) by ω_4, where:

$$\omega_1 = 1 - \sigma, \quad \omega_2 = \mu - \sigma, \quad \omega_4 = \sigma\gamma, \quad \omega_3 = \sigma(1-\gamma).$$

This yields the following expression:

$$\begin{aligned} R_1 + \mu R_2 \le\ &(\omega_1 + \omega_3)I(\hat{U}\hat{W}; Y) + (\omega_2 + \omega_4)I(\tilde{W}\tilde{V}; Z) \\ &+ \omega_3(I(X; J|\hat{U}\hat{W}) + I(\tilde{V}; Z|\tilde{W}) - I(\tilde{V}; J|\tilde{W})) \\ &+ \omega_4(I(X; J|\tilde{V}\tilde{W}) + I(\hat{U}; Y|\hat{W}) - I(\hat{U}; J|\hat{W})) \\ =\ &\sigma I(X; J|\hat{U}\hat{W}) - I(X; Y|\hat{U}\hat{W}) + \sigma\gamma(I(X; Y|\hat{W}) - I(X; J|\hat{W})) + (1 - \sigma\gamma)I(X; Y) \\ &+ \sigma I(X; J|\tilde{V}\tilde{W}) - \mu I(X; Z|\tilde{V}\tilde{W}) + (\sigma - \sigma\gamma)(I(X; Z|\tilde{W}) - I(X; J|\tilde{W})) + (\mu - \sigma + \sigma\gamma)I(X; Z). \end{aligned}$$

This is the same expression as in Theorem 5. □

A. On Sum-Rate Constraints of Broadcast Channels

In [9], Geng considers a binary skew-symmetric broadcast channel (BSSC), as depicted in Figure 2, defined by the transition probabilities:

$$T_{Y|X} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}, \qquad T_{Z|X} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

The sum-rate capacity of the BSSC is unknown. The sum-rate of Marton's inner bound numerically evaluates to approximately 0.3616, while UVW outer bound yields a value of around 0.3725, demonstrating a non-trivial gap between the inner bound

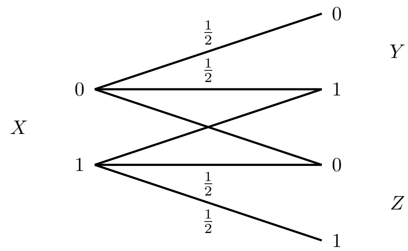

Fig. 2. Binary skew-symmetric broadcast channel.

and the outer bound. Geng evaluates the outer bound in Theorem 5 by taking the channel X to J to be a binary symmetric channel (BSC)

$$T_{J|X} = \begin{bmatrix} 0.810013443141155 & 0.189986556858845 \\ 0.189986556858845 & 0.810013443141155 \end{bmatrix}$$

As noted in Remark 1, the outer bound presented in Theorem 5 relies on J solely through $T_{J|X}$. He obtains an improved upper bound of 0.369316568803963 on the sum-rate. Although the outer bound in Theorem 3 implies the bound in Theorem 5, the latter is more straightforward to evaluate. Conversely, Theorem 3 poses challenges for numerical evaluation due to the numerous auxiliary random variables and the accompanying equality constraints. To address this, we provide a weaker version of the outer bound in Theorem 3 (for the sum-rate) which can be numerically evaluated.

Theorem 6. *For any $T_{J|XYZ}$, any achievable rate tuple must satisfy*

$$\begin{aligned} R_0 + R_1 + R_2 &\le \min\left\{I(\hat{W};Y) - I(\hat{W};J),\quad I(\tilde{W};Z) - I(\tilde{W};J)\right\} + I(X;J) \\ &\quad + I(\hat{U};Y|\hat{W}) - I(\hat{U};J|\hat{W}) + I(\tilde{V};Z|\tilde{W}) - I(\tilde{V};J|\tilde{W}), &\text{(9a)} \\ R_0 + R_1 + R_2 &\le I(X;Y) + I(\tilde{V}\tilde{W};Z) - I(\tilde{V}\tilde{W};J) + I(\hat{V}\hat{W};J) - I(\hat{V}\hat{W};Y), &\text{(9b)} \\ R_0 + R_1 + R_2 &\le I(X;Y) + I(\tilde{V};Z|\tilde{W}) - I(\tilde{V};J|\tilde{W}) + I(\hat{V};J|\hat{W}) - I(\hat{V};Y|\hat{W}), &\text{(9c)} \\ R_0 + R_1 + R_2 &\le I(X;Z) - I(\tilde{U};Z|\tilde{W}) + I(\tilde{U};J|\tilde{W}) - I(\hat{U};J|\hat{W}) + I(\hat{U};Y|\hat{W}), &\text{(9d)} \\ R_0 + R_1 + R_2 &\le I(X;Z) - I(\tilde{U}\tilde{W};Z) + I(\tilde{U}\tilde{W};J) - I(\hat{U}\hat{W};J) + I(\hat{U}\hat{W};Y), &\text{(9e)} \end{aligned}$$

for some choice of distribution over the variables

$$p_{\tilde{U},\tilde{V},\tilde{W},\hat{U},\hat{V},\hat{W},X,Y,Z,J} = p_X p_{\tilde{W},\tilde{U},\tilde{V}|X} p_{\hat{W},\hat{U},\hat{V}|X} T_{Y,Z|X} T_{J|X,Y,Z}.$$

Proof. Recall the sum-rate constraints in Theorem 3 are

$$\begin{aligned} R_0 + R_1 + R_2 &\le \min\left\{I(\hat{W};Y) - I(\hat{W};J),\quad I(\tilde{W};Z) - I(\tilde{W};J)\right\} + I(X;J) \\ &\quad + I(\hat{U};Y|\hat{W}) - I(\hat{U};J|\hat{W}) + I(\tilde{V};Z|\tilde{W}) - I(\tilde{V};J|\tilde{W}), &\text{(10)} \\ R_0 + R_1 + R_2 &\le \min\left\{I(W;Y), I(W;Z)\right\} \\ &\quad + \min\left\{I(V;Z|W) + I(X;Y|V,W), I(U;Y|W) + I(X;Z|U,W)\right\}. &\text{(11)} \end{aligned}$$

They should satisfy the equality constraints

$$\begin{aligned} I(W;Z) - I(W;Y) &= I(\tilde{W};Z) - I(\tilde{W};J) + I(\hat{W};J) - I(\hat{W};Y), &\text{(12a)} \\ I(U;Z|W) - I(U;Y|W) &= I(\tilde{U};Z|\tilde{W}) - I(\tilde{U};J|\tilde{W}) + I(\hat{U};J|\hat{W}) - I(\hat{U};Y|\hat{W}), &\text{(12b)} \\ I(V;Z|W) - I(V;Y|W) &= I(\tilde{V};Z|\tilde{W}) - I(\tilde{V};J|\tilde{W}) + I(\hat{V};J|\hat{W}) - I(\hat{V};Y|\hat{W}), &\text{(12c)} \\ I(V;Z|W) + I(X;Y|V,W) &= I(U;Y|W) + I(X;Z|U,W). &\text{(12d)} \end{aligned}$$

The sum-rate constraint in (11) can be expanded as

$$\begin{aligned} R_0 + R_1 + R_2 &\le I(V;Z|W) - I(V;Y|W) + I(X;Y), \\ R_0 + R_1 + R_2 &\le I(V;Z|W) - I(V;Y|W) + I(X;Y) + I(W;Z) - I(W;Y), \end{aligned}$$

$$R_0 + R_1 + R_2 \leq I(U;Y|W) - I(U;Z|W) + I(X;Z),$$
$$R_0 + R_1 + R_2 \leq I(U;Y|W) - I(U;Z|W) + I(X;Z) + I(W;Y) - I(W;Z).$$

By substituting the quantities $I(W;Y) - I(W;Z)$, $I(U;Z|W) - I(U;Y|W)$ and $I(V;Z|W) - I(V;Y|W)$ with their value from (12a)-(12d), we can obtain the claimed region. $\square$

The sum-rate upper bound in Theorem 6 is less than or equal to that in Theorem 5, primarily due to the constraint in (9a). However, our attempts to utilize this upper bound did not yield an improvement over the sum-rate of 0.369316568803963 for the BSSC channel, based on the choices of auxiliary receiver J that we explored.

Next, we show that the sum-rate outer bound in Theorem 4 is implied by that in Theorem 6. Consequently, it is not feasible to use Theorem 4 to achieve better sum-rate upper bounds for the BSSC channel.

Proposition 2. *For every $T_{G,K|X,Y,Z}$, the sum-rate of the outer bound region in Theorem 4 is greater than or equal to the sum-rate of the outer bound region in Theorem 6 for the choice of $J = (G, K)$. In other words, the sum-rate constraints established by Theorem 6 (Theorem 3) imply the sum-rate constraints set forth in Theorem 4.*

Proof. Take an arbitrary $T_{G,K|X,Y,Z}$. Consider an enhanced broadcast channel $T_{Y',Z'|X}$ with the same input alphabet $\mathcal{X}$, where $Y' = (Y, G)$ and $Z' = (Z, K)$. The upper bound in Theorem 6 for $T_{Y,Z|X}$ is less than or equal to that of the enhanced channel $T_{Y',Z'|X}$, because the rate expressions are non-decreasing under the transformations $Y \to (Y, G)$ and $Z \to (Z, K)$. For instance, in (9b), we have the inequality

$$I(X;Y) - I(\hat{V}\hat{W};Y) \leq I(X;Y') - I(\hat{V}\hat{W};Y').$$

By setting $J = (G, K)$, the sum-rate constraints of Theorem 6 for the enhanced channel $T_{Y',Z'|X}$ become:

$$\begin{aligned}
R_0 + R_1 + R_2 \leq \min\Big\{ & I(\hat{W};Y,G) - I(\hat{W};G,K), \quad I(\tilde{W};Z,K) - I(\tilde{W};G,K)\Big\} + I(X;G,K) \\
& + I(\hat{U};Y,G|\hat{W}) - I(\hat{U};G,K|\hat{W}) + I(\tilde{V};Z,K|\tilde{W}) - I(\tilde{V};G,K|\tilde{W}), \\
R_0 + R_1 + R_2 \leq\ & I(X;Y,G) + I(\tilde{V}\tilde{W};Z,K) - I(\tilde{V}\tilde{W};G,K) + I(\hat{V}\hat{W};G,K) - I(\hat{V}\hat{W};Y,G), \\
R_0 + R_1 + R_2 \leq\ & I(X;Y,G) + I(\tilde{V};Z,K|\tilde{W}) - I(\tilde{V};G,K|\tilde{W}) + I(\hat{V};G,K|\hat{W}) - I(\hat{V};Y,G|\hat{W}), \\
R_0 + R_1 + R_2 \leq\ & I(X;Z,K) - I(\tilde{U};Z,K|\tilde{W}) + I(\tilde{U};G,K|\tilde{W}) - I(\hat{U};G,K|\hat{W}) + I(\hat{U};Y,G|\hat{W}), \\
R_0 + R_1 + R_2 \leq\ & I(X;Z,K) - I(\tilde{U}\tilde{W};Z,K) + I(\tilde{U}\tilde{W};G,K) - I(\hat{U}\hat{W};G,K) + I(\hat{U}\hat{W};Y,G),
\end{aligned}$$

for some $p(\hat{w}, \hat{u}, \hat{v}|x)p(\tilde{w}, \tilde{u}, \tilde{v}|x)p(x)$.

Upon rearranging the terms, the above inequalities become

$$\begin{aligned}
R_0 + R_1 + R_2 \leq\ & \min\{I(\hat{W};Y|G) + I(\tilde{W},K;G), I(\tilde{W};Z|K) + I(\hat{W},G;K)\} \\
& + I(X;G|\tilde{V},\tilde{W},K) + I(\hat{U};Y|\hat{W},G) + I(X;K|\hat{U},\hat{W},G) + I(\tilde{V};Z|\tilde{W},K), \\
R_0 + R_1 + R_2 \leq\ & \min\{I(\hat{W};Y|G) + I(\tilde{W},K;G), I(\tilde{W};Z|K) + I(\hat{W},G;K)\} \\
& + I(X;Y|\hat{V},\hat{W},G) + I(X;G|\tilde{V},\tilde{W},K) + I(\tilde{V};Z|\tilde{W},K) + I(\hat{V};K|\hat{W},G), \\
R_0 + R_1 + R_2 \leq\ & \min\{I(\hat{W};Y|G) + I(\tilde{W},K;G), I(\tilde{W};Z|K) + I(\hat{W},G;K)\} \\
& + I(X;Z|\tilde{U},\tilde{W},K) + I(X;K|\hat{U},\hat{W},G) + I(\tilde{U};G|\tilde{W},K) + I(\hat{U};Y|\hat{W},G).
\end{aligned}$$

Relating the auxiliary random variables $(\hat{W}, \hat{U}, \hat{V}) \leftrightarrow (W_a, U_a, V_a)$ and $(\tilde{W}, \tilde{U}, \tilde{V}) \leftrightarrow (W_b, U_b, V_b)$, the above constraints are identical to the sum-rate constraints in Theorem 4. $\square$

Remark 4. Consider the identification of random variables in Theorem 3 as given in (5a)-(5e). For the choice of $Y' = (Y, G)$, $Z' = (Z, K)$ and $J = (G, K)$, the identifications of the auxiliary random variables become:

$$\hat{W} = ((Y,G)^{i-1}, (G,K)_{i+1}^n, M_0, Q) = (Y^{i-1}, G^{n\setminus i}, K_{i+1}^n, Q); \hat{U} = M_1; \hat{V} = M_2;$$
$$\tilde{W} = ((G,K)^{i-1}, (Z,K)_{i+1}^n, M_0, Q) = (G^{i-1}, K^{n\setminus i}, Z_{i+1}^n, Q); \tilde{U} = M_1; \tilde{V} = M_2.$$

This identification is consistent with the proof of Theorem 4.

B. A New Outer Bound with Two Auxiliary Receivers

As noted earlier, the sum-rate upper bound in Theorem 4 does not exceed that in Theorem 6. Additionally, we could not find examples improving over the upper bound of 0.369316568803963 for the BSSC channel by searching over auxiliary channels with binary and ternary alphabet $\mathcal{J}$. This motivated us to derive new outer bounds on the capacity region, which have the potential to surpass the aforementioned upper bound for the BSSC channel. The key idea is to consider two new auxiliary receivers, G and K, and to compare the sum-rate across the following sequence of broadcast channels:

$$\begin{tikzcd} & & T_{G,G|X} \\ T_{Y,Z|X} \arrow[r] & T_{G,K|X} \arrow[ur] \arrow[dr] & \\ & & T_{K,K|X} \end{tikzcd}$$

As discussed in the introduction, the move from $T_{G,K|X}$ to $T_{G,G|X}$ and $T_{K,K|X}$ means that we write the two sum-rate constraints of the UVW outer bound for the $P_{G,K|X}$ channel. We focus on the sum-rate constraints below. The outer bound to the full capacity region can be found in Appendix B.

Theorem 7 (GK-Bound)**.** *Given a broadcast channel characterized by $T_{Y,Z|X}$ and any achievable rate triple (R_0, R_1, R_2), one can find some input distribution $p(x)$ such that for any auxiliary channel $T_{G,K|X,Y,Z}$, the following constraints regarding sum-rate are satisfied:*

$$\begin{aligned}
R_0 + R_1 + R_2 \leq & \min\{I(W_a; Y), I(W_c; Z) + I(W_a; G) - I(W_b; G) + I(W_b; K) - I(W_c; K)\} \\
& + I(U_c, W_c; K) - I(U_b, W_b; K) + I(U_b, W_b; G) - I(U_a, W_a; G) \\
& + I(U_a; Y|W_a) + I(X; Z|U_c, W_c), \qquad (13a) \\
R_0 + R_1 + R_2 \leq & \min\{I(W_a; Y) + I(W_c; K) - I(W_b; K) + I(W_b; G) - I(W_a; G), I(W_c; Z)\} \\
& + I(V_a, W_a; G) - I(V_b, W_b; G) + I(V_b, W_b; K) - I(V_c, W_c; K) \\
& + I(V_c; Z|W_c) + I(X; Y|V_a, W_a), \qquad (13b) \\
R_0 + R_1 + R_2 \leq & I(W_a; Y) + I(U_a; Y|W_a) + I(V_c; Z|W_c) \\
& + I(U_b, W_b; G) - I(U_a, W_a; G) - I(V_c; K|W_c) + I(X; K|U_b, W_b), \qquad (13c) \\
R_0 + R_1 + R_2 \leq & I(W_a; Y) + I(U_a; Y|W_a) + I(V_c; Z|W_c) \\
& + I(V_b; K|W_b) - I(V_c; K|W_c) - I(V_b; G|W_b) + I(X; G|U_a, W_a), \qquad (13d) \\
R_0 + R_1 + R_2 \leq & I(W_c; Z) + I(U_a; Y|W_a) + I(V_c; Z|W_c) \\
& + I(V_b, W_b; K) - I(V_c, W_c; K) - I(U_a; G|W_a) + I(X; G|V_b, W_b), \qquad (13e) \\
R_0 + R_1 + R_2 \leq & I(W_c; Z) + I(U_a; Y|W_a) + I(V_c; Z|W_c) \\
& + I(U_b; G|W_b) - I(U_a; G|W_a) - I(U_b; K|W_b) + I(X; K|V_c, W_c), \qquad (13f)
\end{aligned}$$

for some choice of distribution over the variables

$$p_{U_a,V_a,W_a,U_b,V_b,W_b,U_c,V_c,W_c,X,Y,Z,G,K} = p_X p_{U_a,V_a,W_a|X} p_{U_b,V_b,W_b|X} p_{U_c,V_c,W_c|X} T_{Y,Z|X} T_{G,K|X,Y,Z}.$$

Remark 5. Additional equality and inequality constraints may be incorporated, analogous to those in (4a)–(4c).

Proof. Take an arbitrary code (n, M_0, M_1, M_2) with error probability $\lim_{n\to\infty} \epsilon_n = 0$. We can build two auxiliary receivers as

$$p_{G^n,K^n|M_0,M_1,M_2,X^n,Y^n,Z^n} = \prod_{i=1}^{n} T_{G_i,K_i|X_i,Y_i,Z_i}.$$

We present the n-letter expansions used to derive the constraints, which follow from Fano's inequality:

$$\begin{aligned}
n(R_0 + R_1 + R_2) \leq & I(M_0, M_1; K^n) + I(M_2; Z^n|M_0, M_1) \\
& + \Big(I(M_0, M_1; Y^n) - I(M_0, M_1; G^n)\Big) + \Big(I(M_0, M_1; G^n) - I(M_0, M_1; K^n)\Big), \qquad (14a) \\
n(R_0 + R_1 + R_2) \leq & I(M_0; Z^n) + I(M_1; K^n|M_0) + I(M_2; Z^n|M_0, M_1)
\end{aligned}$$

$$+\Big(I(M_1;Y^n|M_0)-I(M_1;G^n|M_0)\Big)+\Big(I(M_1;G^n|M_0)-I(M_1;K^n|M_0)\Big), \tag{14b}$$

$$n(R_0+R_1+R_2)\leq I(M_0;Y^n)+I(M_2;G^n|M_0)+I(M_1;Y^n|M_0,M_2)$$
$$+\Big(I(M_2;Z^n|M_0)-I(M_2;K^n|M_0)\Big)+\Big(I(M_2;K^n|M_0)-I(M_2;G^n|M_0)\Big), \tag{14c}$$

$$n(R_0+R_1+R_2)\leq I(M_0,M_2;G^n)+I(M_1;Y^n|M_0,M_2)$$
$$+\Big(I(M_0,M_2;Z^n)-I(M_0,M_2;K^n)\Big)+\Big(I(M_0,M_2;K^n)-I(M_0,M_2;G^n)\Big), \tag{14d}$$

$$n(R_0+R_1+R_2)\leq I(M_0,M_1;G^n)+I(M_2;K^n|M_0,M_1)$$
$$+\Big(I(M_0,M_1;Y^n)-I(M_0,M_1;G^n)\Big)+\Big(I(M_2;Z^n|M_0)-I(M_2;K^n|M_0)\Big), \tag{14e}$$

$$n(R_0+R_1+R_2)\leq I(M_0,M_1;Y^n)+I(M_2;G^n|M_0,M_1)$$
$$+\Big(I(M_2;Z^n|M_0)-I(M_2;K^n|M_0)\Big)+\Big(I(M_2;K^n|M_0)-I(M_2;G^n|M_0)\Big), \tag{14f}$$

$$n(R_0+R_1+R_2)\leq I(M_0,M_2;K^n)+I(M_1;G^n|M_0,M_2)$$
$$+\Big(I(M_1;Y^n|M_0)-I(M_1;G^n|M_0)\Big)+\Big(I(M_0,M_2;Z^n)-I(M_0,M_2;K^n)\Big), \tag{14g}$$

$$n(R_0+R_1+R_2)\leq I(M_0,M_2;Z^n)+I(M_1;K^n|M_0,M_2)$$
$$+\Big(I(M_1;Y^n|M_0)-I(M_1;G^n|M_0)\Big)+\Big(I(M_1;G^n|M_0)-I(M_1;K^n|M_0)\Big). \tag{14h}$$

Let Q be the time-sharing random variable, uniform over $[n]$, and independent of all previously defined random variables. We make the following identification:

$$W_a=(M_0,Y^{Q-1},G_{Q+1}^n,Q);\quad W_b=(M_0,G_{Q+1}^n,K^{Q-1},Q);\quad W_c=(M_0,K^{Q-1},Z_{Q+1}^n,Q);$$
$$U_a=U_b=U_c=M_1;\qquad V_a=V_b=V_c=M_2,$$
$$X=X_Q,\quad Y=Y_Q,\quad G=G_Q,\quad K=K_Q.$$

The terms single-letterized with W_a are marked in blue, those single-letterized with W_c are marked in red, and those single-letterized with W_b are kept in black. Notice that W_a corresponds to the move from receiver Y to G, while W_c is for moving receiver Z to K. The term W_b is used to formulate the UVW outer bound for $T_{KG|X}$.

The above constraints can then be single-letterized in a canonical manner via Csiszár sum identity (Körner-Marton identity), similar to the approach in [12]. In Appendix A, we present two examples that illustrate the single-letterization process.

The decomposition

$$p_{U_a,V_a,W_a,U_b,V_b,W_b,U_c,V_c,W_c,X}=p_X p_{U_a,V_a,W_a|X}p_{U_b,V_b,W_b|X}p_{U_c,V_c,W_c|X}$$

comes from the fact that the expressions in the outer bound depend only on the marginal distributions of (X,U_a,V_a,W_a), (X,U_b,V_b,W_b) and (X,U_c,V_c,W_c).

□

Theorem 8. *The sum-rate constraints in Theorem 7 imply the sum-rate constraints in Theorem 6.*

Proof. By substituting both G and K with J in GK-Bound, the constraints (19k), (19l), (19m), (19o) become:

$$R_0+R_1+R_2\leq I(X;Z)+I(U_a,W_a;Y)-I(U_a,W_a;J)+I(U_c,W_c;J)-I(U_c,W_c;Z),$$
$$R_0+R_1+R_2\leq I(X;Z)+I(U_a;Y|W_a)-I(U_a;J|W_a)+I(U_c;J|W_c)-I(U_c;Z|W_c),$$
$$R_0+R_1+R_2\leq I(X;Y)+I(V_a,W_a;J)-I(V_a,W_a;Y)+I(V_c,W_c;Z)-I(V_c,W_c;J),$$
$$R_0+R_1+R_2\leq I(X;Y)+I(V_a;J|W_a)-I(V_a;Y|W_a)+I(V_c;Z|W_c)-I(V_c;J|W_c),$$
$$R_0+R_1+R_2\leq I(W_a;Y)-I(W_a;J)+I(X;J)$$
$$+I(U_a;Y|W_a)-I(U_a;J|W_a)+I(V_c;Z|W_c)-I(V_c;J|W_c),$$
$$R_0+R_1+R_2\leq I(W_c;Z)-I(W_c;J)+I(X;J)$$
$$+I(U_a;Y|W_a)-I(U_a;J|W_a)+I(V_c;Z|W_c)-I(V_c;J|W_c).$$

Relating the auxiliary random variables $(W_a,U_a,V_a)\leftrightarrow(\hat{W},\hat{U},\hat{V})$ and $(W_c,U_c,V_c)\leftrightarrow(\tilde{W},\tilde{U},\tilde{V})$, the above constraints are identical to those in Theorem 6. □

IV. Numerical Evaluation

Consider the BSSC channel, which is a skew-symmetric channel. The sum-rate from Theorem 5 is evaluated as follows:

$$\begin{aligned} SR_J := \max_{p(x)} \min_{T_{J|X}} & \frac{1}{2} I(X;Y) + \frac{1}{2}\mathfrak{C}\big[I(X;Y) - I(X;J) + 2\mathfrak{C}[I(X;J) - I(X;Y)]\big] \\ & + \frac{1}{2} I(X;Z) + \frac{1}{2}\mathfrak{C}\big[I(X;Z) - I(X;J) + 2\mathfrak{C}[I(X;J) - I(X;Z)]\big], \end{aligned} \tag{15}$$

where $\mathfrak{C}[f]$ stands for the upper concave envelope of a function f.

Using (19m) and (19o), we can deduce a simple upper bound from Theorem 7 in the concave-envelope form:

$$\begin{aligned} SR_{GK} := \max_{p(x)} \min_{T_{G,K|X}} & \frac{1}{2} I(X;Y) + \frac{1}{2}\mathfrak{C}\big[I(X;Y) - I(X;G) + 2\mathfrak{C}[I(X;G) - I(X;Y)]\big] \\ & + \frac{1}{2} I(X;Z) + \frac{1}{2}\mathfrak{C}\big[I(X;Z) - I(X;K) + 2\mathfrak{C}[I(X;K) - I(X;Z)]\big] \\ & + \frac{1}{2}\mathfrak{C}[I(X;G) - I(X;K)] + \frac{1}{2}\mathfrak{C}[I(X;K) - I(X;G)]. \end{aligned} \tag{16}$$

Exploiting the symmetry of the BSSC channel, we also evaluate Marton's inner bound and UVW outer bound sum-rate bounds as follows:

$$\begin{aligned} SR_M = & \frac{1}{2} I(X;Y) + \frac{1}{2} I(X;Z) \\ & + \frac{1}{2} \max\{I(X;Y) - I(X;Z), I(X;Z) - I(X;Y)\}, \end{aligned} \tag{17}$$

$$\begin{aligned} SR_{UVW} = & \frac{1}{2} I(X;Y) + \frac{1}{2} I(X;Z) \\ & + \frac{1}{2}\mathfrak{C}[I(X;Y) - I(X;Z)] + \frac{1}{2}\mathfrak{C}[I(X;Z) - I(X;Y)]. \end{aligned} \tag{18}$$

Due to skew-symmetry, the maximizing input distribution p_X is uniform for SR_M, SR_{UVW}, SR_J, and SR_{GK}. As mentioned previously, after optimizing over $p_{J|X}$, Geng obtained the sum-rate upper bound of $SR_J = 0.369316568803963$. We obtain $SR_{GK} = 0.369296340638082$ for the choice of

$$T_{G|X} = \begin{bmatrix} 0.206404160918957 & 0.793595839081043 \\ 0.826329217311554 & 0.173670782688446 \end{bmatrix},$$

$$T_{K|X} = \begin{bmatrix} 0.173670782688446 & 0.826329217311554 \\ 0.793595839081043 & 0.206404160918957 \end{bmatrix}.$$

Note that $SR_{GK} < SR_J$, indicating a strict improvement for the BSSC example. However, the improvement is small. In the next section, we will consider an example with a larger gap.

A. *Example for a Larger Gap*

Consider a BSSC with transition probability 0.9 (instead of 0.5), defined by:

$$T_{Y|X} = \begin{bmatrix} 0.9 & 0.1 \\ 0 & 1 \end{bmatrix}, \qquad T_{Z|X} = \begin{bmatrix} 1 & 0 \\ 0.1 & 0.9 \end{bmatrix}.$$

In this case, we have:

$$\begin{aligned} SR_M = & 0.813077267246467, \\ SR_{GK} = & 0.823617442275509, \\ SR_J = & 0.824265223964071, \\ SR_{UVW} = & 0.825057863680834. \end{aligned}$$

The minimizing auxiliary receiver J is a binary symmetric channel

$$T_{J|X} = \begin{bmatrix} 0.958133054922025 & 0.041866945077975 \\ 0.041866945077975 & 0.958133054922025 \end{bmatrix},$$

while the minimizing auxiliary receiver pair (G, K) appear to be a skew-symmetric broadcast channel

$$T_{G|X} = \begin{bmatrix} 0.064116169833244 & 0.935883830166756 \\ 0.977452865531223 & 0.022547134468777 \end{bmatrix},$$

$$T_{K|X} = \begin{bmatrix} 0.022547134468777 & 0.977452865531223 \\ 0.935883830166756 & 0.064116169833244 \end{bmatrix}.$$

Figure 3 further illustrates the difference between $SR_J - SR_M$ and $SR_{GK} - SR_M$ as a function of the transition probability $T_{Y|X}(1|0)$ of the BSSC.

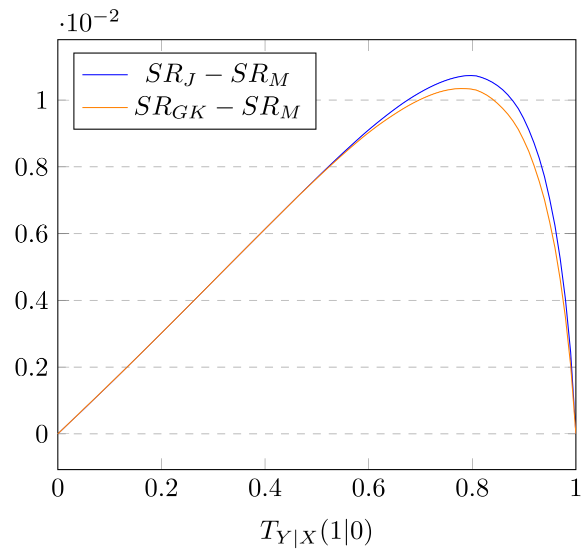

Fig. 3. Difference between SR_J and SR_{GK} evaluated at BSSC with different transition probability

V. Conclusion

We proposed a new outer bound with two auxiliary receivers that strictly improves upon existing bounds for certain broadcast channels. The core idea involves transforming the channel $T_{Y,Z|X}$ into $T_{G,K|X}$ and comparing the mutual information terms that define the n-letter characterization of the capacity region of $T_{Y,Z|X}$ (i.e., $I(M_1; Y^n)$, $I(M_2; Z^n)$, $I(M_1; Y^n|M_2)$, etc) with their counterparts for $T_{G,K|X}$. We note that instead of transforming $T_{Y,Z|X}$ into $T_{G,K|X}$, one may proceed through a series of intermediary steps involving multiple auxiliary broadcast channels.

Acknowledgment

The authors would like to thank Abbas El Gamal for interesting discussions, suggestions, and comments.

References

[1] V. Anantharam, A. Gohari, and C. Nair, *On the evaluation of Marton's inner bound for two-receiver broadcast channels*, IEEE Transactions on Information Theory **65** (2019), no. 3, 1361–1371.
[2] Venkat Anantharam, Amin Aminzadeh Gohari, and Chandra Nair, *Improved cardinality bounds on the auxiliary random variables in Marton's inner bound*, 2013 IEEE International Symposium on Information Theory (ISIT'2013) (Istanbul, Turkey), July 2013.
[3] Zijie Chen, Amin Gohari, and Chandra Nair, *A differential equation approach to the most-informative boolean function conjecture*, arXiv preprint arXiv:2502.10019 (2025).
[4] T. Cover, *Broadcast channels*, IEEE Transactions on Information Theory **18** (1972), no. 1, 2–14.
[5] Abbas El Gamal, Amin Gohari, and Chandra Nair, *A strengthened cutset upper bound on the capacity of the relay channel and applications*, IEEE Trans. Inf. Theory **68** (2022), no. 8, 5013–5043.
[6] Abbas El Gamal and Young-Han Kim, *Network Information Theory*, Cambridge University Press, 2012.
[7] Y. Geng, A. Gohari, C. Nair, and Y. Yu, *On Marton's inner bound and its optimality for classes of product broadcast channels*, IEEE Transactions on Information Theory **60** (2014), no. 1, 22–41.
[8] Y. Geng and C. Nair, *The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages*, IEEE Transactions on Information Theory **60** (2014), no. 4, 2087–2104.
[9] Yanlin Geng, *A simple auxiliary receiver outer bound for broadcast channels*, 2025 IEEE International Symposium on Information Theory (ISIT), 2025, pp. 1–6.
[10] A. Gohari, C. Nair, and V. Anantharam, *On Marton's inner bound for broadcast channels*, 2012 IEEE International Symposium on Information Theory, July 2012, pp. 581–585.
[11] Amin Gohari, Yi Liu, and Chandra Nair, *A conjecture regarding the optimizers of marton's inner bound for the two-receiver broadcast channel*, 2025 IEEE International Symposium on Information Theory (ISIT), 2025, pp. 1–6.
[12] Amin Gohari and Chandra Nair, *Outer bounds for multiuser settings: The auxiliary receiver approach*, IEEE Transactions on Information Theory **68** (2022), no. 2, 701–736.
[13] Amin Gohari, Chandra Nair, and Jinpei Zhao, *On the capacity region of some classes of interference channels*, 2024 IEEE International Symposium on Information Theory (ISIT), 2024, pp. 3136–3141.
[14] K. Marton, *A coding theorem for the discrete memoryless broadcast channel*, IEEE Transactions on Information Theory **25** (1979), no. 3, 306–311.
[15] C. Nair, *Capacity regions of two new classes of two-receiver broadcast channels*, IEEE Transactions on Information Theory **56** (2010), no. 9, 4207–4214.

[16] C Nair and A El Gamal, *An outer bound to the capacity region of the broadcast channel*, IEEE Trans. Info. Theory **IT-53** (January, 2007), 350–355.
[17] C. Nair, H. Kim, and A. El Gamal, *On the optimality of randomized time division and superposition coding for the broadcast channel*, 2016 IEEE Information Theory Workshop (ITW), Sept 2016, pp. 131–135.
[18] H. Weingarten, Y. Steinberg, and S. S. Shamai, *The capacity region of the Gaussian multiple-input multiple-output broadcast channel*, IEEE Transactions on Information Theory **52** (2006), no. 9, 3936–3964.
[19] Zhenduo Wen and Amin Gohari, *A new upper bound for distributed hypothesis testing using the auxiliary receiver approach*, arXiv preprint arXiv:2409.14148 (2024).

Appendix A
Single-letterization Proof

We state the following lemma without proof.

Lemma 1. *For any random variable tuple* (A, B, Y^n, Z^n)*,*

$$
\begin{aligned}
I(A;Y^n|B) - I(A;Z^n|B) &= \sum_{i=1}^{n} \Big(I(A;Y_i|B,Y^{i-1},Z_{i+1}^n) - I(A;Z_i|B,Y^{i-1},Z_{i+1}^n) \Big) \\
&= \sum_{i=1}^{n} \Big(I(A,Y^{i-1},Z_{i+1}^n;Y_i|B) - I(A,Y^{i-1},Z_{i+1}^n;Z_i|B) \Big) \\
&\quad + \sum_{i=1}^{n} \Big(I(Z_i;Z^{i-1}|B) - I(Y_i;Y^{i-1}|B) \Big).
\end{aligned}
$$

Moreover,

$$
I(X^n;Y^n) = \sum_{i=1}^{n} I(X_i;Y_i) - \sum_{i=1}^{n} I(Y_i;Y^{i-1}).
$$

Recall that the following identification is made, with Q be the time-sharing variable.

$$
\begin{aligned}
&W_a = (M_0, Y^{Q-1}, G_{Q+1}^n, Q); \quad W_b = (M_0, G_{Q+1}^n, K^{Q-1}, Q); \quad W_c = (M_0, K^{Q-1}, Z_{Q+1}^n, Q); \\
&U_a = U_b = U_c = M_1; \qquad V_a = V_b = V_c = M_2, \\
&X = X_Q, \quad Y = Y_Q, \quad G = G_Q, \quad K = K_Q.
\end{aligned}
$$

For (14a)**:** for the red term, we have

$$
\begin{aligned}
&I(M_0, M_1; K^n) + I(M_2; Z^n|M_0, M_1) \\
&= I(M_0, M_1; K^n) + I(X^n; Z^n|M_0, M_1) \\
&= I(M_0, M_1; K^n) - I(M_1, M_2; Z^n) + I(X^n; Z^n) \\
&= \sum_{i=1}^{n} \Big(I(M_0, M_1, K^{i-1}, Z_{i+1}^n; K_i) + I(X_i; Z_i|M_0, M_1, K^{i-1}, Z_{i+1}^n) - I(K^{i-1}; K_i) \Big) \\
&= n \cdot \big(I(U_c, W_c; K) + I(X; Z|U_c, W_c) \big) - nI(Q;K) - \sum_{i=1}^{n} I(K^{i-1}; K_i).
\end{aligned}
$$

Observe that we get an extra negative term $-nI(Q;K) - \sum_{i=1}^{n} I(K^{i-1};K_i)$ after single-letterization, where Q is the time-sharing variable.

Next, consider the blue term

$$
\begin{aligned}
&I(M_0, M_1; Y^n) - I(M_0, M_1; G^n) \\
=&\sum_{i} \Big(I(M_0, M_1, Y^{i-1}, G_{i+1}^n; Y_i) - I(M_0, M_1, Y^{i-1}, G_{i+1}^n; G_i) - I(Y^{i-1}; Y_i) + I(G_{i+1}^n; G_i) \Big) \\
=&n \cdot \big(I(U_a, W_a; Y) - I(U_a, W_a; G) \big) - nI(Q;Y) + nI(Q;G) - \sum_{i=1}^{n} I(Y^{i-1}; Y_i) + \sum_{i=1}^{n} I(G_{i+1}^n; G_i)
\end{aligned}
$$

with the extra term $-nI(Q;Y) + nI(Q;G) - \sum_{i=1}^{n} I(Y^{i-1};Y_i) + \sum_{i=1}^{n} I(G_{i+1}^n;G_i)$. Similarly, for the black term

$$
\begin{aligned}
&I(M_0, M_1; G^n) - I(M_0, M_1; K^n) \\
=&n\big(I(U_b, W_b; G) - I(U_b, W_b; K) \big) - nI(Q;G) + nI(Q;K) - \sum_{i=1}^{n} I(G^{i-1}; G_i) + \sum_{i=1}^{n} I(K^{i-1}; K_i),
\end{aligned}
$$

with the extra term $-nI(Q;G) + nI(Q;K) - \sum_{i=1}^{n} I(G_{i+1}^n;G_i) + \sum_{i=1}^{n} I(K^{i-1};K_i)$. Observe that if we add up the extra terms, we get

$$
- nI(Q;K) - \sum_{i=1}^{n} I(K^{i-1}; K_i) - nI(Q;Y) + nI(Q;G) - \sum_{i=1}^{n} I(Y^{i-1}; Y_i) + \sum_{i=1}^{n} I(G_{i+1}^n; G_i)
$$

$$-nI(Q;G)+nI(Q;K)-\sum_{i=1}^{n} I(G_{i+1}^n;G_i)+\sum_{i=1}^{n} I(K^{i-1};K_i)\le 0.$$

For (14g)**:** for the black term,

$$\begin{aligned}
&I(M_0,M_2;K^n)+I(M_1;G^n|M_0,M_2)\\
=&I(M_0,M_2;K^n)+I(X^n;G^n|M_0,M_2)\\
=&I(M_0,M_2;K^n)-I(M_0,M_2;G^n)+I(X^n;G^n)\\
=&\sum_{i=1}^{n}\Big(I(M_0,M_2,G_{i+1}^n,K^{i-1};K_i)-I(M_0,M_2,G_{i+1}^n,K^{i-1};G_i)+I(X_i;G_i)-I(K^{i-1};K_i)\Big)\\
=&n\cdot\big(I(V_b,W_b;K)+I(X;G|V_b,W_b)\big)-nI(Q;K)-\sum_{i=1}^{n} I(K^{i-1};K_i).
\end{aligned}$$

We obtain the extra term $-nI(Q;K)-\sum_{i=1}^{n} I(K^{i-1};K_i)$ where Q is the time-sharing variable. For blue term,

$$\begin{aligned}
&\Big(I(M_1;Y^n|M_0)-I(M_1;G^n|M_0)\Big)\\
=&\Big(I(M_0,M_1;Y^n)-I(M_0,M_1;G^n)+I(M_0;G^n)-I(M_0;Y^n)\Big)\\
=&\sum_{i=1}^{n}\Big(I(M_0,M_1,Y^{i-1},G_{i+1}^n;Y_i)-I(M_0,M_1,Y^{i-1},G_{i+1}^n;G_i)\\
&-I(M_0,Y^{i-1},G_{i+1}^n;Y_i)+I(M_0,Y^{i-1},G_{i+1}^n;G_i)\Big)\\
=&n\cdot\big(I(U_a;Y|W_a)-I(U_a;G|W_a)\big),
\end{aligned}$$

with no extra terms.

And similarly for the red term,

$$\begin{aligned}
&\Big(I(M_0,M_2;Z^n)-I(M_0,M_2;K^n)\Big)\\
=&\sum_{i=1}^{n}\Big(I(M_0,M_2,Z_{i+1}^n,K^{i-1};Z_i)-I(M_0,M_2,Z_{i+1}^n,K^{i-1};K_i)-I(Z_i;Z_{i+1}^n)+I(K_i;K^{i-1})\Big)\\
=&n\cdot\big(I(W_c,V_c;Z)-I(W_c,V_c;K)\big)+n(I(Q;K)-I(Q;Z))-\sum_{i=1}^{n} I(Z_i;Z_{i+1}^n)+\sum_{i=1}^{n} I(K_i;K^{i-1}),
\end{aligned}$$

with the extra term $n(I(Q;K)-I(Q;Z))-\sum_{i=1}^{n} I(Z_i;Z_{i+1}^n)+\sum_{i=1}^{n} I(K_i;K^{i-1})$ where Q is the time-sharing variable. Observe that if we add up the extra terms, we get

$$-nI(Q;K)-\sum_{i=1}^{n} I(K^{i-1};K_i)+n(I(Q;K)-I(Q;Z))-\sum_{i=1}^{n} I(Z_i;Z_{i+1}^n)+\sum_{i=1}^{n} I(K_i;K^{i-1})\le 0.$$

Appendix B
GK-Bound to the full capacity region

Theorem 9 (GK-Bound (Full Version))**.** *Given a broadcast channel characterized by $T_{Y,Z|X}$ and any achievable rate triple (R_0,R_1,R_2), one can find some input distribution $p(x)$ such that for any auxiliary channel $T_{G,K|X,Y,Z}$, the following constraints regarding sum-rate are satisfied:*

$$R_0\le I(W_a;Y)+\min\{0,I(W_b;G)-I(W_a;G),I(W_b;G)-I(W_a;G)+I(W_c;K)-I(W_b;K)\}, \tag{19a}$$

$$R_0\le I(W_c;Z)+\min\{0,I(W_b;K)-I(W_c;K),I(W_b;K)-I(W_c;K)+I(W_a;G)-I(W_b;G)\}, \tag{19b}$$

$$\begin{aligned}R_0+R_1\le\ & I(W_a;Y)+I(U_a;Y|W_a) && (19c)\\ &+\min\{0,I(U_b,W_b;G)-I(U_a,W_a;G),I(U_b,W_b;G)-I(U_a,W_a;G)+I(U_c,W_c;K)-I(U_b,W_b;K)\}, && (19d)\end{aligned}$$

$$\begin{aligned}R_0+R_1\le\ & I(W_c;Z)+I(U_a;Y|W_a)+I(W_a;G)-I(W_b;G)+I(W_b;K)-I(W_c;K) && (19e)\\ &+\min\{0,I(U_b,W_b;G)-I(U_a,W_a;G),I(U_b,W_b;G)-I(U_a,W_a;G)+I(U_c,W_c;K)-I(U_b,W_b;K)\}, && (19f)\end{aligned}$$

$$R_0+R_2\le I(W_a;Y)+I(V_c;Z|W_c)+I(W_c;K)-I(W_b;K)+I(W_b;G)-I(W_a;G) \tag{19g}$$

$$+\min\{0, I(V_b,W_b;K)-I(V_c,W_c;K), I(V_b,W_b;K)-I(V_c,W_c;K)+I(V_a,W_a;G)-I(V_b,W_b;G)\}, \tag{19h}$$

$$R_0+R_2 \le I(W_c;Z)+I(V_c;Z|W_c) \tag{19i}$$

$$+\min\{0, I(V_b,W_b;K)-I(V_c,W_c;K), I(V_b,W_b;K)-I(V_c,W_c;K)+I(V_a,W_a;G)-I(V_b,W_b;G)\}, \tag{19j}$$

$$\begin{aligned} R_0+R_1+R_2 \le& \min\{I(W_a;Y), I(W_c;Z)+I(W_a;G)-I(W_b;G)+I(W_b;K)-I(W_c;K)\} \\ &+I(U_c,W_c;K)-I(U_b,W_b;K)+I(U_b,W_b;G)-I(U_a,W_a;G) \\ &+I(U_a;Y|W_a)+I(X;Z|U_c,W_c), \end{aligned} \tag{19k}$$

$$\begin{aligned} R_0+R_1+R_2 \le& \min\{I(W_a;Y)+I(W_c;K)-I(W_b;K)+I(W_b;G)-I(W_a;G), I(W_c;Z)\} \\ &+I(V_a,W_a;G)-I(V_b,W_b;G)+I(V_b,W_b;K)-I(V_c,W_c;K) \\ &+I(V_c;Z|W_c)+I(X;Y|V_a,W_a), \end{aligned} \tag{19l}$$

$$\begin{aligned} R_0+R_1+R_2 \le& I(W_a;Y)+I(U_a;Y|W_a)+I(V_c;Z|W_c) \\ &+I(U_b,W_b;G)-I(U_a,W_a;G)-I(V_c;K|W_c)+I(X;K|U_b,W_b), \end{aligned} \tag{19m}$$

$$\begin{aligned} R_0+R_1+R_2 \le& I(W_a;Y)+I(U_a;Y|W_a)+I(V_c;Z|W_c) \\ &+I(V_b;K|W_b)-I(V_c;K|W_c)-I(V_b;G|W_b)+I(X;G|U_a,W_a), \end{aligned} \tag{19n}$$

$$\begin{aligned} R_0+R_1+R_2 \le& I(W_c;Z)+I(U_a;Y|W_a)+I(V_c;Z|W_c) \\ &+I(V_b,W_b;K)-I(V_c,W_c;K)-I(U_a;G|W_a)+I(X;G|V_b,W_b), \end{aligned} \tag{19o}$$

$$\begin{aligned} R_0+R_1+R_2 \le& I(W_c;Z)+I(U_a;Y|W_a)+I(V_c;Z|W_c) \\ &+I(U_b;G|W_b)-I(U_a;G|W_a)-I(U_b;K|W_b)+I(X;K|V_c,W_c), \end{aligned} \tag{19p}$$

for some choice of distribution over the variables

$$p_{U_a,V_a,W_a,U_b,V_b,W_b,U_c,V_c,W_c,X,Y,Z,G,K} = p_X p_{U_a,V_a,W_a|X} p_{U_b,V_b,W_b|X} p_{U_c,V_c,W_c|X} T_{Y,Z|X} T_{G,K|X,Y,Z}$$

satisfying

$$\begin{aligned} 0 &\le I(X;Z|U_c,W_c)-I(X;K|U_c,W_c) \le I(V_c;Z|W_c)-I(V_c;K|W_c), \\ 0 &\le I(X;Y|V_a,W_a)-I(X;G|V_a,W_a) \le I(U_a;Y|W_a)-I(U_a;G|W_a). \end{aligned}$$