Outer Bounds on the Admissible Source Region for Broadcast Channels With Correlated Sources

Kia Khezeli and Jun Chen, *Member, IEEE*

Abstract—Two outer bounds on the admissible source region for broadcast channels with correlated sources are presented: the first one is strictly tighter than the existing outer bound by Gohari and Anantharam, while the second one provides a complete characterization of the admissible source region in the case where the two sources are conditionally independent given the common part. These outer bounds are deduced from the general necessary conditions established for the lossy source broadcast problem via suitable comparisons between the virtual broadcast channel (induced by the source and the reconstructions) and the physical broadcast channel.

Index Terms—Bandwidth mismatch, broadcast channel, capacity region, deterministic channel, joint source-channel coding.

I. INTRODUCTION

LET $\{S(t)\}_{t=1}^{\infty}$ be an i.i.d. random process with marginal distribution p_S over alphabet $\mathcal{S}$. In the lossy source broadcast problem (see Fig. 1), an encoding function $f^{(m,\rho m)}$: $\mathcal{S}^m \to \mathcal{X}^{\rho m}$ maps a source block $S^m \triangleq (S(1), \cdots, S(m))$ to a channel input block $X^{\rho m} \triangleq (X(1), \cdots, X(\rho m))$, which is sent over a discrete memoryless broadcast channel $p_{Y_1, Y_2|X}$ with input alphabet $\mathcal{X}$ and output alphabets $\mathcal{Y}_i$, $i = 1, 2$; at receiver i, a decoding function $g_i^{(\rho m,m)}$: $\mathcal{Y}_i^{\rho m} \to \hat{\mathcal{S}}_i^m$ maps the channel output block $Y_i^{\rho m} \triangleq (Y_i(1), \cdots, Y_i(\rho m))$ (generated by $X^{\rho m}$) to a source reconstruction block $\hat{S}_i^m \triangleq (\hat{S}_i(1), \cdots, \hat{S}_i(m))$, $i = 1, 2$. The number of channel uses per source sample, i.e., ρ, is referred to as the bandwidth expansion ratio. We assume that $\mathcal{S}, \hat{\mathcal{S}}_1, \hat{\mathcal{S}}_2, \mathcal{X}, \mathcal{Y}_1$, and $\mathcal{Y}_2$ are finite sets throughout the paper.

Definition 1: Let $w_i : \mathcal{S} \times \hat{\mathcal{S}}_i \to [0, \infty)$, $i = 1, 2$, be two distortion measures. We say distortion pair (d_1, d_2) is achievable under distortion measures w_1 and w_2 subject to bandwidth expansion constraint κ if, for every $\epsilon > 0$, there exist encoding function $f^{(m,\rho m)} : \mathcal{S}^m \to \mathcal{X}^{\rho m}$ and decoding

functions $g_i^{(\rho m,m)} : \mathcal{Y}_i^{\rho m} \to \hat{\mathcal{S}}_i^m$, $i = 1, 2$, with $\rho \leq \kappa + \epsilon$, such that

$$\frac{1}{m} \sum_{t=1}^{m} \mathbb{E}[w_i(S(t), \hat{S}_i(t))] \leq d_i + \epsilon, \quad i = 1, 2.$$

A special case of the lossy source broadcast problem, sometimes referred to as broadcasting correlated sources (see Fig. 2), has received particular attention. In this case, $S(t) = (S_1(t), S_2(t))$ with $S_1(t)$ and $S_2(t)$ jointly distributed according to $p_{(S_1, S_2)}$ over alphabet $\mathcal{S}_1 \times \mathcal{S}_2$, $t = 1, 2, \cdots$, and receiver i wishes to reconstruct $\{S_i(t)\}_{t=1}^{\infty}$ almost losslessly, $i = 1, 2$.

Definition 2: A source distribution $p_{(S_1, S_2)}$ is said to be admissible for broadcast channel $p_{Y_1, Y_2|X}$ subject to bandwidth expansion constraint κ if, for every $\epsilon > 0$, there exist encoding function $f^{(m,\rho m)} : \mathcal{S}_1^m \times \mathcal{S}_2^m \to \mathcal{X}^n$ and decoding functions $g_i^{(\rho m,m)} : \mathcal{Y}_i^{\rho m} \to \mathcal{S}_i^m$, $i = 1, 2$, with $\rho \leq \kappa + \epsilon$, such that

$$\frac{1}{m} \sum_{t=1}^{m} \Pr(S_i(t) \neq \hat{S}_i(t)) \leq \epsilon, \quad i = 1, 2.$$

The set of all such $p_{(S_1, S_2)}$ is referred to as the admissible source region for broadcast channel $p_{Y_1, Y_2|X}$ subject to bandwidth expansion constraint κ.

Remark: Definition 2 is a special case of Definition 1 with $d_1 = d_2 = 0$ and $w_i : \mathcal{S} \times \mathcal{S}_i \to \{0, 1\}$ given by

$$w_i((s_1, s_2), \hat{s}_i) = \begin{cases} 0, & s_i = \hat{s}_i \\ 1, & \text{otherwise} \end{cases}, \quad i = 1, 2. \quad (1)$$

It is worth mentioning that, for the problem of broadcasting correlated sources, typically the more restrictive block error probability constraints are adopted. However, it is clear that outer bounds derived under average symbol error probability constraints automatically hold under block error probability constraints.

Han and Costa [1] derived an inner bound on the admissible source region; see [2] for a minor correction and [3] for an alternative characterization. Outer bounds were established by Gohari and Anantharam and by Kramer, Liang, and Shamai. Note that, due to the lack of cardinality bounds on the auxiliary random variables, neither the original version of the Gohari-Anantharam outer bound [4], [5] nor the Kramer-Liang-Shamai outer bound [6] is directly computable. Thanks to [7], a computable characterization of the Gohari-Anantharam outer bound has been found recently [8], [9]. On the other hand, it is difficult, if not impossible, to express the Kramer-Liang-Shamai outer bound
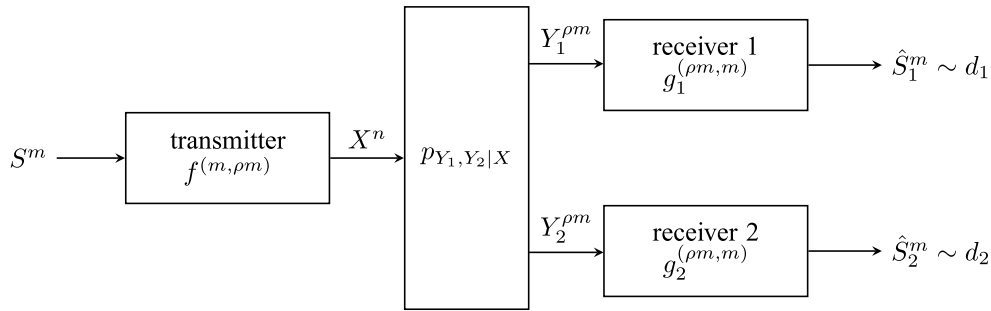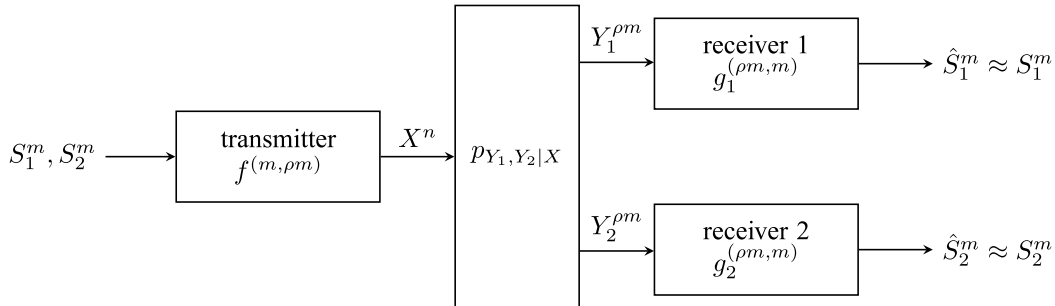
Fig. 1. The lossy broadcast problem.

Fig. 2. Broadcasting correlated sources.

in a computable form because of the fact that certain auxiliary random variables involved in this bound are constrained[1] to be i.i.d. copies of the source variables. For this reason, only the Gohari-Anantharam outer bound is considered in the present work.

In this paper we establish two necessary conditions for the lossy source broadcast problem. Both conditions are built upon the intuition that the virtual broadcast channel (induced by the source and the reconstructions) is dominated by the physical broadcast channel. Our effort is largely devoted to seeking mathematical formulations that can capture, to a certain extent, this vague intuition. It will be seen that the notion of dominance, which has a precise definition in the point-to-point case due to the source-channel separation theorem, permits several possible generalizations to the broadcast channel setting, and each generalization gives rise to a necessary condition for the lossy source broadcast problem. These necessary conditions, when specialized to the problem of broadcasting correlated sources, yield two outer bounds on the admissible source region: the first one is strictly tighter than the Gohari-Anantharam outer bound while the second one provides a complete characterization of the admissible source region in the case where the two sources are conditionally independent given the common part.

The rest of this paper is organized as follows. We explain our general approach in Section II. Section III contains a short review of the relevant capacity results for broadcast channels. The necessary conditions for the lossy source broadcast problem and the induced outer bounds on the source admissible region are presented in Sections IV and V. We conclude the paper in Section VI.

[1]If such a constraint is removed, then the Gohari-Antharam outer bound is at least as tight as the Kramer-Liang-Shamai outer bound.

II. VIRTUAL CHANNEL VERSUS PHYSICAL CHANNEL

For the purpose of illustrating our general approach, it is instructive to first consider the point-to-point communication problem. Specifically, in the point-to-point setting, an encoding function $f^{(m,\rho m)} : \mathcal{S}^m \to \mathcal{X}^{\rho m}$ maps a source block S^m to a channel input block $X^{\rho m}$, which is sent over a discrete memoryless channel $p_{Y|X}$ with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$; at the receiver end, a decoding function $g^{(\rho m,m)} : \mathcal{Y}^{\rho m} \to \hat{\mathcal{S}}^m$ maps the channel output block $Y^{\rho m}$ (generated by $X^{\rho m}$) to a source reconstruction block $\hat{S}^m$. For any conditional distribution $p_{\hat{S}^m|S^m}$, let $p_{\hat{S}|S}$ be its single-letterized version defined as

$$p_{\hat{S}|S}(\hat{s}|s) = \frac{1}{m} \sum_{t=1}^{m} p_{\hat{S}(t)|S(t)}(\hat{s}|s),$$

where

$$p_{\hat{S}(t)|S(t)}(\hat{s}|s) = \sum_{\substack{s^m:s(t)=s \\ \hat{s}^m:\hat{s}(t)=\hat{s}}} p_{\hat{S}^m|S^m}(\hat{s}^m|s^m) \prod_{t':t'\neq t} p_S(s(t')).$$

One can readily verify that

$$\mathbb{E}[w(S, \hat{S})] = \frac{1}{m} \sum_{t=1}^{m} \mathbb{E}[w(S(t), \hat{S}(t))] \qquad (2)$$

for any distortion measure $w : \mathcal{S} \times \hat{\mathcal{S}} \to [0, \infty)$. We say that $p_{\hat{S}^m|S^m}$ is degraded with respect to $p_{Y^{\rho m}|X^{\rho m}}$ (where $p_{Y^{\rho m}|X^{\rho m}}(y^{\rho m}|x^{\rho m}) = \prod_{q=1}^{\rho m} p_{Y|X}(y(q)|x(q))$) if

$$p_{\hat{S}^m|S^m}(\hat{s}^m|s^m)$$
$$= \sum_{x^{\rho m}, y^{\rho m}} p_{X^{\rho m}|S^m}(x^{\rho m}|s^m) p_{Y^{\rho m}|X^{\rho m}}(y^{\rho m}|x^{\rho m})$$
$$\times p_{\hat{S}^m|Y^{\rho m}}(\hat{s}^m|y^{\rho m})$$

for some conditional distributions $p_{X^{\rho m}|S^m}$ and $p_{\hat{S}^m|Y^{\rho m}}$; note that, in the point-to-point communication problem, we have

$$p_{X^{\rho m}|S^m}(x^{\rho m}|s^m) = \mathbb{I}(x^{\rho m} = f^{(m,\rho m)}(s^m)),$$
$$p_{\hat{S}^m|Y^{\rho m}}(\hat{s}^m|y^{\rho m}) = \mathbb{I}(\hat{s}^m = g^{(\rho m,m)}(y^{\rho m})),$$

where $\mathbb{I}(\cdot)$ is the indicator function. We shall refer to an arbitrary conditional distribution $p_{\hat{S}|S}$ as a virtual channel, and say that it is realizable through the physical channel $p_{Y|X}$ with bandwidth expansion ratio ρ if it can be obtained, via single-letterization, from certain $p_{\hat{S}^m|S^m}$ degraded[2] with respect to $p_{Y^{\rho m}|X^{\rho m}}$. It is worth emphasizing that a realizable $p_{\hat{S}|S}$ is not necessarily degraded with respect to $p_{Y|X}$. Indeed, even in the bandwidth-matched case (i.e., $\rho = 1$), there might not exist conditional distributions $p_{X|S}$ and $p_{\hat{S}|Y}$ such that

$$p_{\hat{S}|S}(\hat{s}|s) = \sum_{x,y} p_{X|S}(x|s)p_{Y|X}(y|x)p_{\hat{S}|Y}(\hat{s}|y)$$

since otherwise the end-to-end distortion $\frac{1}{m}\sum_{t=1}^m \mathbb{E}[w(S(t), \hat{S}(t))]$ could be achieved[3] without coding. Nevertheless, it can be shown that every realizable $p_{\hat{S}|S}$ is dominated by $p_{Y|X}$ in the sense that

$$I(S; \hat{S}) \leq \rho I(X; Y) \tag{3}$$

for some input distribution p_X. In a certain sense, (3) is the only connection between a generic realizable virtual channel and the physical channel. Indeed, the source-channel separation theorem essentially asserts that a virtual channel $p_{\hat{S}|S}$ is (asymptotically) realizable through the physical channel $p_{Y|X}$ with bandwidth expansion ratio ρ if and only if $p_{\hat{S}|S}$ is dominated by $p_{Y|X}$ in the sense of (3). Note that, given bandwidth expansion ratio ρ, the achievability of end-to-end distortion d is equivalent to the existence of a realizable virtual channel $p_{\hat{S}|S}$ satisfying $\mathbb{E}[w(S, \hat{S})] \leq d$. More generally, one can formulate the question of determining whether there exists a realizable virtual channel in a prescribed set; imposing distortion constraints can be viewed as a specific way of choosing such sets.

This perspective can also be adopted in the broadcast channel setting. For any conditional distribution $p_{\hat{S}_1^m, \hat{S}_2^m|S^m}$, we define its single-letterized version $p_{\hat{S}_1, \hat{S}_2|S}$ as

$$p_{\hat{S}_1, \hat{S}_2|S}(\hat{s}_1, \hat{s}_2|s) = \frac{1}{m}\sum_{t=1}^m p_{\hat{S}_1(t), \hat{S}_2(t)|S(t)}(\hat{s}_1, \hat{s}_2|s),$$

where

$$p_{\hat{S}_1(t), \hat{S}_2(t)|S(t)}(\hat{s}_1, \hat{s}_2|s)$$
$$= \sum_{\substack{s^m:s(t)=s \\ \hat{s}_i^m:\hat{s}_i(t)=\hat{s}_i, i=1,2}} p_{\hat{S}_1^m, \hat{S}_2^m|S^m}(\hat{s}_1^m, \hat{s}_2^m|s^m)\prod_{t':t'\neq t} p_S(s(t')).$$

---

[2]Since $p_{\hat{S}^m|S^m}$ is only required to be degraded with respect to $p_{Y^{\rho m}|X^{\rho m}}$, we essentially allow non-deterministic encoding and decoding functions. However, it can be shown via a standard derandomization argument that restricting encoding and decoding functions to deterministic ones does not affect the set of (asymptotically) realizable virtual channels.

[3]One could simply use source variable S to generate channel input X via $p_{X|S}$, and use channel output Y to generate reconstruction variable $\hat{S}$ via $p_{\hat{S}|Y}$. In light of (2), the resulting distortion $\mathbb{E}[w(S, \hat{S})]$ is the same as $\frac{1}{m}\sum_{t=1}^m \mathbb{E}[w(S(t), \hat{S}(t))]$.

It can be verified that

$$\mathbb{E}[w_i(S, \hat{S}_i)] = \frac{1}{m}\sum_{t=1}^m \mathbb{E}[w_i(S(t), \hat{S}_i(t))]$$

for any distortion measure $w_i : \mathcal{S} \times \hat{\mathcal{S}} \to [0, \infty)$, $i = 1, 2$. We say that $p_{\hat{S}_1^m, \hat{S}_2^m|S^m}$ is degraded with respect to $p_{Y_1^{\rho m}, Y_2^{\rho m}|X^{\rho m}}$ (where $p_{Y_1^{\rho m}, Y_2^{\rho m}|X^{\rho m}}(y_1^{\rho m}, y_2^{\rho m}|x^{\rho m}) = \prod_{q=1}^{\rho m} p_{Y|X}(y_1(q), y_2(q)|x(q))$) if

$$p_{\hat{S}_1^m, \hat{S}_2^m|S^m}(\hat{s}_1^m, \hat{s}_2^m|s^m)$$
$$= \sum_{x^{\rho m}, y_1^{\rho m}, y_2^{\rho m}} p_{X^{\rho m}|S^m}(x^{\rho m}|s^m)$$
$$\times p_{Y_1^{\rho m}, Y_2^{\rho m}|X^{\rho m}}(y_1^{\rho m}, y_2^{\rho m}|x^{\rho m})\prod_{i=1}^2 p_{\hat{S}_i^m|Y_i^{\rho m}}(\hat{s}_i^m|y_i^{\rho m}) \tag{4}$$

for some conditional distributions $p_{X^{\rho m}|S^m}$ and $p_{\hat{S}_i^m|Y_i^{\rho m}}$, $i = 1, 2$; note that

$$p_{X^{\rho m}|S^m}(x^{\rho m}|s^m) = \mathbb{I}(x^{\rho m} = f^{(m,\rho m)}(s^m)),$$
$$p_{\hat{S}_i^m|Y_i^{\rho m}}(\hat{s}_i^m|y_i^{\rho m}) = \mathbb{I}(\hat{s}_i^m = g_i^{(\rho m,m)}(y_i^{\rho m})), \quad i = 1, 2,$$

in the lossy source broadcast problem. We shall refer to an arbitrary conditional distribution $p_{\hat{S}_1, \hat{S}_2|S}$ as a virtual broadcast channel,[4] and say that it is realizable through the physical broadcast channel $p_{Y_1, Y_2|X}$ with bandwidth expansion ratio ρ if it can be obtained, via single-letterization, from certain $p_{\hat{S}_1^m, \hat{S}_2^m|S^m}$ degraded with respect to $p_{Y_1^{\rho m}, Y_2^{\rho m}|X^{\rho m}}$. The fundamental problem here is to determine whether there exists a realizable virtual broadcast channel satisfying the distortion constraints (or more generally, whether there exists a realizable virtual broadcast channel in a prescribed set). It is conceivable that every realizable virtual broadcast channel must be dominated, in a certain sense, by the physical broadcast channel. However, it is apparently a formidable task to develop a computable notion of dominance that can completely characterize the set of realizable virtual broadcast channels (since that would solve several long-standing open problems in network information theory). Nevertheless, if one does not insist on such a complete characterization, then it is indeed possible to establish certain connections between a generic realizable virtual broadcast channel and the physical broadcast channel by suitably generalizing (3). For example, it is straightforward to show that, if $p_{\hat{S}|S}$ is realizable through $p_{Y_1, Y_2|X}$ with bandwidth expansion ratio ρ, then $p_{\hat{S}|S}$ must be dominated by $p_{Y_1, Y_2|X}$ in the sense that

$$I(S; \hat{S}_i) \leq \rho I(X; Y_i), \quad i = 1, 2,$$

for some input distribution p_X. This is by no means the only possible generalization of (3) to the broadcast channel setting, and two stronger notions of dominance will be presented in Sections IV and V. Note that each notion of dominance implicitly provides an outer bound on the set of realizable virtual broadcast channels, which in turn yields a necessary condition for the achievability of any given distortion pair.

---

[4]It is worth mentioning that the idea of viewing the conditional distribution of the reconstructions given the source as a virtual broadcast channel was exploited earlier in [10] and [11] through a different angle.
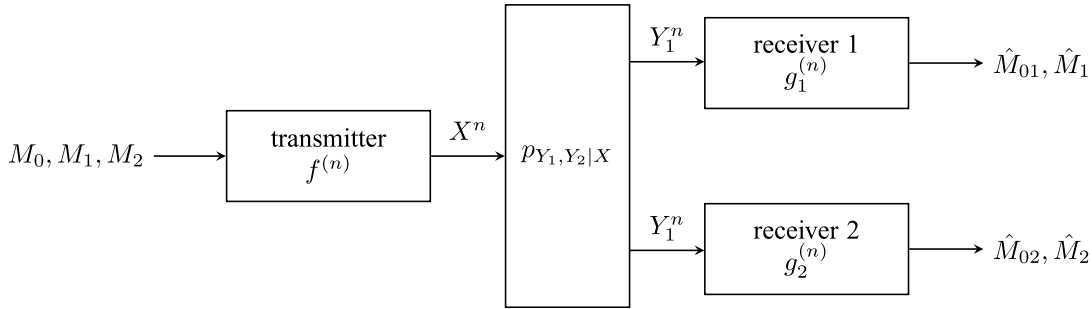
Fig. 3.   Broadcast channel with common and private messages.

III. REVIEW OF CAPACITY RESULTS FOR BROADCAST CHANNELS

We shall give a brief review of certain capacity results for broadcast channels that are relevant to the notions of dominance developed in Sections IV and V. Let $p_{Y_1,Y_2|X}$ be a discrete memoryless broadcast channel with input alphabet $\mathcal{X}$ and output alphabets $\mathcal{Y}_i$, $i = 1, 2$. A length-n coding scheme (see Fig. 3) for $p_{Y_1,Y_2|X}$ consists of

- a common message M_0 and two private messages M_i, $i = 1, 2$, where (M_0, M_1, M_2) is uniformly distributed over $\mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2$,
- an encoding function $f^{(n)} : \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{X}^n$ that maps (M_0, M_1, M_2) to a channel input block X^n,
- two decoding functions $g_i^{(n)} : \mathcal{Y}_i^n \to \mathcal{M}_0 \times \mathcal{M}_i$, $i = 1, 2$, where $g_i^{(n)}$ maps the channel output block at receiver i, i.e., $Y_{i,1}^n$, to $(\hat{M}_{0i}, \hat{M}_i)$, $i = 1, 2$.

Definition 3: A rate triple $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ is said to be achievable for broadcast channel $p_{Y_1,Y_2|X}$ if there exists a sequence of encoding functions $f^{(n)} : \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{X}^n$ with $\frac{1}{n} \log |\mathcal{M}_i| \geq R_i$, $i = 0, 1, 2$, and decoding functions $g_i^{(n)} : \mathcal{Y}_i^n \to \mathcal{M}_0 \times \mathcal{M}_i$, $i = 1, 2$, such that

$$\lim_{n \to \infty} \Pr\{(\hat{M}_{01}, \hat{M}_1) \neq (M_0, M_1)$$
$$\text{or } (\hat{M}_{02}, \hat{M}_2) \neq (M_0, M_2)\} = 0.$$

The capacity region $\mathcal{C}(p_{Y_1,Y_2|X})$ is the closure of the set of all achievable (R_0, R_1, R_2) for broadcast channel $p_{Y_1,Y_2|X}$.

Let $\mathcal{C}_{\text{in}}(p_X, p_{Y_1,Y_2|X})$ denote the set of $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ satisfying

$$R_0 \leq \min\{I(V_0; Y_1), I(V_0; Y_2)\},$$
$$R_0 + R_i \leq I(V_0, V_i; Y_i), \quad i = 1, 2,$$
$$R_0 + R_1 + R_2 \leq \min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_1; Y_1|V_0)$$
$$+ I(V_2; Y_2|V_0) - I(V_1; V_2|V_0)$$

for some $p_{V_0,V_1,V_2,X,Y_1,Y_2} = p_{V_0,V_1,V_2|X} p_X p_{Y_1,Y_2|X}$. Here it suffices to consider $|\mathcal{V}_0| \leq |\mathcal{X}| + 4$ and $|\mathcal{V}_i| \leq |\mathcal{X}|$, $i = 1, 2$; moreover, there is no loss of generality in assuming that X is a deterministic function of (V_0, V_1, V_2) [12, Th. 1].[5] Define

$$\mathcal{C}_{\text{in}}(p_{Y_1,Y_2|X}) = \bigcup_{p_X} \mathcal{C}_{\text{in}}(p_X, p_{Y_1,Y_2|X}).$$

[5]It is expected [13] that one can further improve the cardinality bounds on $\mathcal{V}_i$, $i = 1, 2$, to $|\mathcal{V}_1| + |\mathcal{V}_2| \leq |\mathcal{X}| + 1$ by leveraging the techniques developed in [14].

We have [15, Th. 1] [16, p. 391, Problem 10(c)]

$$\mathcal{C}_{\text{in}}(p_{Y_1,Y_2|X}) \subseteq \mathcal{C}(p_{Y_1,Y_2|X}).$$

Note that $\mathcal{C}_{\text{in}}(p_{Y_1,Y_2|X})$ is widely known as Marton's inner bound (see [17, Th. 2] for the case $R_0 = 0$).

Let $\mathcal{C}_{\text{out}}(p_X, p_{Y_1,Y_2|X})$ denote the set of $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ satisfying

$$R_0 \leq \min\{I(V_0; Y_1), I(V_0; Y_2)\},$$
$$R_0 + R_i \leq \min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_i; Y_i|V_0),$$
$$i = 1, 2,$$
$$R_0 + R_1 + R_2 \leq \min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_1; Y_1|V_0)$$
$$+ I(X; Y_2|V_0, V_1),$$
$$R_0 + R_1 + R_2 \leq \min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_2; Y_2|V_0)$$
$$+ I(X; Y_1|V_0, V_2)$$

for some $p_{V_0,V_1,V_2,X,Y_1,Y_2} = p_{V_0,V_1,V_2|X} p_X p_{Y_1,Y_2|X}$. Here it suffices to consider $|\mathcal{V}_0| \leq |\mathcal{X}| + 5$ and $|\mathcal{V}_i| \leq |\mathcal{X}| + 1$, $i = 1, 2$ [7].[6] Define

$$\mathcal{C}_{\text{out}}(p_{Y_1,Y_2|X}) = \bigcup_{p_X} \mathcal{C}_{\text{out}}(p_X, p_{Y_1,Y_2|X}).$$

We have [7]

$$\mathcal{C}(p_{Y_1,Y_2|X}) \subseteq \mathcal{C}_{\text{out}}(p_{Y_1,Y_2|X}). \tag{5}$$

It is worth noting [15], [18] that $\mathcal{C}_{\text{in}}(p_X, p_{Y_1,Y_2|X})$ can be defined equivalently as the set of $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ satisfying

$$R_0 \leq \min\{I(V_0; Y_1), I(V_0; Y_2)\},$$
$$R_0 + R_i \leq \min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_i; Y_i|V_0),$$
$$i = 1, 2,$$
$$R_0 + R_1 + R_2 \leq \min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_1; Y_1|V_0)$$
$$+ I(V_2; Y_2|V_0) - I(V_1; V_2|V_0)$$

for some $p_{V_0,V_1,V_2,X,Y_1,Y_2} = p_{V_0,V_1,V_2|X} p_X p_{Y_1,Y_2|X}$. With this equivalent definition of $\mathcal{C}_{\text{in}}(p_X, p_{Y_1,Y_2|X})$, one can readily show

$$\mathcal{C}_{\text{in}}(p_X, p_{Y_1,Y_2|X}) \subseteq \mathcal{C}_{\text{out}}(p_X, p_{Y_1,Y_2|X})$$

[6]In fact, the cardinality bounds on $\mathcal{V}_i$, $i = 1, 2$, can be further improved to $|\mathcal{V}_i| \leq |\mathcal{X}|$, $i = 1, 2$ [18].

by invoking the fact that, for any $p_{V_0,V_1,V_2,X,Y_1,Y_2} = p_{V_0,V_1,V_2|X} p_X p_{Y_1,Y_2|X}$,

$$I(X; Y_2|V_0, V_1) \geq I(V_2; Y_2|V_0, V_1)$$
$$= I(V_2; Y_2, V_1|V_0) - I(V_1; V_2|V_0)$$
$$\geq I(V_2; Y_2|V_0) - I(V_1; V_2|V_0)$$

and similarly

$$I(X; Y_1|V_0, V_2) \geq I(V_1; Y_1|V_0) - I(V_1; V_2|V_0).$$

IV. APPROACH I: COMPARISON OF CERTAIN MEASUREMENTS INDUCED BY TEST DISTRIBUTIONS

A. The Lossy Source Broadcast Problem

As explained in Section II, our goal is to develop suitable notions of dominance that can (partially) characterize the set of realizable virtual broadcast channels. One such notion is established in the following Lemma (with its proof relegated to Appendix A). Roughly speaking, it shows that every realizable virtual broadcast channel $p_{\hat{S}_1,\hat{S}_2|S}$ (with input distribution p_S) is dominated by the physical broadcast channel $p_{Y_1,Y_2|X}$ (with certain input distribution p_X) in the sense that, given any test distribution $p_{U_0,\cdots,U_L|S}$ for $p_{\hat{S}_1,\hat{S}_2|S}$, one can find a corresponding test distribution $p_{V_0,\cdots,V_L|X}$ for $p_{Y_1,Y_2|X}$ such that certain measurements based on $p_{U_0,\cdots,U_L|S} p_S p_{\hat{S}_1,\hat{S}_2|S}$ are less than or equal to those based on $p_{V_0,\cdots,V_L|X} p_X p_{Y_1,Y_2|X}$ multiplied by bandwidth expansion ratio ρ.

Lemma 1: If a virtual broadcast channel $p_{\hat{S}_1,\hat{S}_2|S}$ is realizable through the physical broadcast channel $p_{Y_1,Y_2|X}$ with bandwidth expansion ratio ρ, then there exists an input distribution p_X such that, for any $p_{U_0,\cdots,U_L,S,\hat{S}_1,\hat{S}_2} = p_{U_0,\cdots,U_L|S} p_S p_{\hat{S}_1,\hat{S}_2|S}$, one can find $p_{V_0,\cdots,V_L,X,Y_1,Y_2} = p_{V_0,\cdots,V_L|X} p_X p_{Y_1,Y_2|X}$ satisfying[7]

$$\sum_{i=1}^{k} I(U_{\mathcal{A}_i}; \hat{S}_{a(i)}|U_{\cup_{j=1}^{i-1}\mathcal{A}_j}) \leq \rho \sum_{i=1}^{k} I(V_{\mathcal{A}_i}; Y_{a(i)}|V_{\cup_{j=1}^{i-1}\mathcal{A}_j})$$

for any $\mathcal{A}_i \subseteq \{0,\cdots,L\}$ and $a(i) \in \{1,2\}$, $i = 1,\cdots,k$.

The following result, which gives a general necessary condition for the lossy source broadcast problem, is a simple consequence of Lemma 1. Its proof can be found in Appendix B.

Theorem 1: If distortion pair (d_1, d_2) is achievable under distortion measures w_1 and w_2 subject to bandwidth expansion constraint κ, then there exists a virtual broadcast channel $p_{\hat{S}_1,\hat{S}_2|S}$ with $\mathbb{E}[w_i(S, \hat{S}_i)] \leq d_i$, $i = 1, 2$, and an input distribution p_X such that, for any $p_{U_0,U_1,U_2,S,\hat{S}_1,\hat{S}_2} = p_{U_0,U_1,U_2|S} p_S p_{\hat{S}_1,\hat{S}_2|S}$, one can find $p_{V_0,V_1,V_2,X,Y_1,Y_2} = p_{V_0,V_1,V_2|X} p_X p_{Y_1,Y_2|X}$ satisfying

$$I(U_0; \hat{S}_i) \leq \kappa I(V_0; Y_i), \quad i = 1, 2,$$
$$I(U_0, U_i; \hat{S}_i) \leq \kappa I(V_0, V_i; Y_i), \quad i = 1, 2,$$
$$I(U_0; \hat{S}_1) + I(U_2; \hat{S}_2|U_0) \leq \kappa[I(V_0; Y_1) + I(V_2; Y_2|V_0)],$$
$$I(U_0; \hat{S}_2) + I(U_1; \hat{S}_1|U_0) \leq \kappa[I(V_0; Y_2) + I(V_1; Y_1|V_0)],$$

---

[7]Here L and k are arbitrary positive integers. We define $U_{\mathcal{A}} = (U_i)_{i\in\mathcal{A}}$ when $\mathcal{A}$ is a non-empty subset of $\{0,\cdots,L\}$ and define $U_{\mathcal{A}} = 0$ otherwise; $V_{\mathcal{A}}$ is defined analogously.

$$I(U_0, U_1; \hat{S}_1) + I(S; \hat{S}_2|U_0, U_1)$$
$$\leq \kappa[I(V_0, V_1; Y_1) + I(X; Y_2|V_0, V_1)],$$
$$I(U_0, U_2; \hat{S}_2) + I(S; \hat{S}_1|U_0, U_2)$$
$$\leq \kappa[I(V_0, V_2; Y_2) + I(X; Y_1|V_0, V_2)],$$
$$I(U_0; \hat{S}_1) + I(U_2; \hat{S}_2|U_0) + I(S; \hat{S}_1|U_0, U_2)$$
$$\leq \kappa[I(V_0; Y_1) + I(V_2; Y_2|V_0) + I(X; Y_1|V_0, V_2)],$$
$$I(U_0; \hat{S}_2) + I(U_1; \hat{S}_1|U_0) + I(S; \hat{S}_2|U_0, U_1)$$
$$\leq \kappa[I(V_0; Y_2) + I(V_1; Y_1|V_0) + I(X; Y_2|V_0, V_1)].$$

Here it suffices to consider $|\mathcal{U}_0| \leq |\mathcal{S}|$, $|\mathcal{V}_0| \leq |\mathcal{X}| + 5$, $|\mathcal{U}_i| \leq |\mathcal{S}|$, and $|\mathcal{V}_i| \leq |\mathcal{X}|$, $i = 1, 2$.

Remark: Gohari and Anantharam independently obtained a necessary condition for the lossy source broadcast channel [8, Th. 2], which is, roughly speaking, a special case of Theorem 1 with $\kappa = 1$, $S = (S_1, S_2)$, and $U_i = S_i$, $i = 1, 2$.

The following result is a direct consequence of Theorem 1.

Corollary 1: If distortion pair (d_1, d_2) is achievable under distortion measures w_1 and w_2 subject to bandwidth expansion constraint κ, then there exists a virtual broadcast channel $p_{\hat{S}_1,\hat{S}_2|S}$ with $\mathbb{E}[w_i(S, \hat{S}_i)] \leq d_i$, $i = 1, 2$ and an input distribution p_X such that

$$\mathcal{C}_{\mathsf{out}}(p_S, p_{\hat{S}_1,\hat{S}_2|S}) \subseteq \kappa \mathcal{C}_{\mathsf{out}}(p_X, p_{Y_1,Y_2|X}).$$

B. An Improved Outer Bound on the Source Admissible Region

The following outer bound on the admissible source region was established by Gohari and Anantharam [8, Corollary 2] (see also [9, Th. 2]).

Let S_0 denote the common part between S_1 and S_2 in the sense of [19] and [20].

Theorem 2: If $p_{(S_1,S_2)}$ is admissible for broadcast channel $p_{Y_1,Y_2|X}$ subject to bandwidth expansion constraint κ, then there exists $p_{V_0,V_1,V_2,X,Y_1,Y_2} = p_{V_0,V_1,V_2|X} p_X p_{Y_1,Y_2|X}$ such that

$$H(S_0) \leq \kappa \min\{I(V_0; Y_1), I(V_0; Y_2)\},$$
$$H(S_i) \leq \kappa[\min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_i; Y_i|V_0)],$$
$$i = 1, 2,$$
$$H(S_1, S_2) \leq \kappa[\min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_1; Y_1|V_0) + I(X; Y_2|V_0, V_1)],$$
$$H(S_1, S_2) \leq \kappa[\min\{I(V_0; Y_1), I(V_0; Y_2)\} + I(V_2; Y_2|V_0) + I(X; Y_1|V_0, V_2)].$$

It is easy to observe that the inequalities in the statement of Theorem 2 closely resemble those in the definition of $\mathcal{C}_{\mathsf{out}}(p_{Y_1,Y_2|X})$. In fact, one can readily establish Theorem 2 by interpreting S_0 as the common message and S_i as the message (including both the private message M_i and the common message M_0) intended for receiver i, $i = 1, 2$, and then following the proof of (5). However, this approach is not completely satisfactory. Note that M_0, M_1, and M_2 are assumed to be independent. If the correspondence between sources and messages is exact, then $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ must form a Markov chain. That is to say, the source-message

correspondence does not fully capture the dependence structure between S_1 and S_2.

We shall show that one can obtain a tighter outer bound on the admissible source region by specializing Theorem 1 to the case of broadcasting correlated sources. Note that, if $\sum_{s,\hat{s}_i} p_S(s) p_{\hat{S}_i|S}(\hat{s}_i|s) w_i(s, \hat{s}_i) = 0$ for w_i given by (1), $i = 1, 2$, then we must have[8] $p_{\hat{S}_1,\hat{S}_2|S} = p_{S_1,S_2|(S_1,S_2)}$, which is a deterministic broadcast channel; moreover, in this case, there is no loss of optimality in choosing $U_i = S_i$, $i = 1, 2$. As a consequence, we obtain the following outer bound on the admissible source region.

Theorem 3: If $p_{(S_1,S_2)}$ is admissible for broadcast channel $p_{Y_1,Y_2|X}$ subject to bandwidth expansion constraint κ, then there exists an input distribution p_X such that, for any $p_{U,(S_1,S_2)} = p_{U|(S_1,S_2)} p_{(S_1,S_2)}$, one can find $p_{V_0,V_1,V_2,X,Y_1,Y_2} = p_{V_0,V_1,V_2|X} p_X p_{Y_1,Y_2|X}$ satisfying

$$I(U; S_i) \leq \kappa I(V_0; Y_i), \quad i = 1, 2,$$
$$H(S_i) \leq \kappa I(V_0, V_i; Y_i), \quad i = 1, 2,$$
$$I(U; S_1) + H(S_2|U) \leq \kappa[I(V_0; Y_1) + I(V_2; Y_2|V_0)],$$
$$I(U; S_2) + H(S_1|U) \leq \kappa[I(V_0; Y_2) + I(V_1; Y_1|V_0)],$$
$$I(U; S_1) + H(S_1, S_2|U)$$
$$\leq \kappa[I(V_0, V_1; Y_1) + I(X; Y_2|V_0, V_1)],$$
$$I(U; S_2) + H(S_1, S_2|U)$$
$$\leq \kappa[I(V_0, V_2; Y_2) + I(X; Y_1|V_0, V_2)],$$
$$I(U; S_1) + H(S_1, S_2|U)$$
$$\leq \kappa[I(V_0; Y_1) + I(V_2; Y_2|V_0) + I(X; Y_1|V_0, V_2)],$$
$$I(U; S_2) + H(S_1, S_2|U)$$
$$\leq \kappa[I(V_0; Y_2) + I(V_1; Y_1|V_0) + I(X; Y_2|V_0, V_1)].$$

Here it suffices to consider $|\mathcal{U}| \leq |\mathcal{S}_1| \times |\mathcal{S}_2|$, $|\mathcal{V}_0| \leq |\mathcal{X}| + 5$, and $|\mathcal{V}_i| \leq |\mathcal{X}|$, $i = 1, 2$.

It is easy to verify that

$$\mathcal{C}_{\text{in}}(p_{(S_1,S_2)}, p_{S_1,S_2|(S_1,S_2)}) = \mathcal{C}_{\text{out}}(p_{(S_1,S_2)}, p_{S_1,S_2|(S_1,S_2)})$$
$$= \mathcal{C}(p_{(S_1,S_2)}), \quad (6)$$

where $\mathcal{C}(p_{(S_1,S_2)})$ denotes the set of $(R_0, R_1, R_2) \in \mathbb{R}^3_+$ satisfying

$$R_0 \leq \min\{I(U; S_1), I(U; S_2)\},$$
$$R_0 + R_i \leq \min\{I(U; S_1), I(U; S_2)\} + H(S_i|U),$$
$$i = 1, 2,$$
$$R_0 + R_1 + R_2 \leq \min\{I(U; S_1), I(U; S_2)\} + H(S_1, S_2|U)$$

for some $p_{U,(S_1,S_2)} = p_{U|(S_1,S_2)} p_{(S_1,S_2)}$ with $|\mathcal{U}| \leq |\mathcal{S}_1| \times |\mathcal{S}_2| + 2$. One can deduce the following result from Theorem 3 (or Corollary 1).

Corollary 2: If $p_{(S_1,S_2)}$ is admissible for broadcast channel $p_{Y_1,Y_2|X}$ subject to bandwidth expansion constraint κ, then there exists an input distribution p_X such that

$$\mathcal{C}(p_{(S_1,S_2)}) \subseteq \kappa \mathcal{C}_{\text{out}}(p_X, p_{Y_1,Y_2|X}). \quad (7)$$

[8]More precisely, we have $p_{\hat{S}_1,\hat{S}_2|S} = p_{S_1,S_2|(S_1,S_2)}$ when the input alphabet restricted to $\{s \in \mathcal{S} : p_S(s) > 0\}$.

The necessary condition in Theorem 2 can be written compactly as

$$(H(S_0), H(S_1|S_0), H(S_2|S_0)) \in \kappa \mathcal{C}_{\text{out}}(p_{Y_1,Y_2|X}) \quad (8)$$

when $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ form a Markov chain. The following result shows that the same simplification is possible for Theorem 3 and, as a consequence, these two theorems are equivalent in this special case.

Theorem 4: The necessary condition in Theorem 3 is equivalent to (8) when $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ form a Markov chain.

Proof: See Appendix C. ∎

It is clear that one can recover Theorem 2 from Theorem 3 by choosing $U = S_0$. Therefore, the new outer bound is at least as tight as the Gohari-Anantharam outer bound. We shall give an example to show that the improvement can be strict. Our example is motivated by the observation that the characterization of the capacity region of the deterministic broadcast channel with a common message involves an auxiliary random variable which is not necessarily a function of the channel input [21] as well as the observation that $H(S_0)$ is not a continuous function of $p_{(S_1,S_2)}$.

Now consider the example where the physical broadcast channel is the Blackwell channel $p^B_{Y_1,Y_2|X}$, where

$$p^B_{Y_1,Y_2|X}(y_1, y_2|x)$$
$$= \begin{cases} 1, & (x, y_1, y_2) = (0, 0, 0), (1, 1, 1), (2, 0, 1) \\ 0, & \text{otherwise} \end{cases}$$

with $x \in \{0, 1, 2\}$ and $y_i \in \{0, 1\}$, $i = 1, 2$; moreover, let $S_1 = (\tilde{S}_0(\alpha), \tilde{S}_1)$ and $S_2 = (\bar{S}_0(\alpha), \tilde{S}_2)$ with $\alpha \in [H_b^{-1}(\frac{1}{2}\log_2 3 - \frac{2}{3}), H_b^{-1}(\log_2 3 - \frac{4}{3}))$, where $\tilde{S}_0(\alpha)$, $\bar{S}_0(\alpha)$, $\tilde{S}_1$, and $\tilde{S}_2$ are binary random variables defined over $\{0, 1\}$, and $H_b^{-1}(\cdot) : [0, 1] \to [0, \frac{1}{2}]$ is the inverse of the binary entropy function $H_b(\cdot)$. Specifically, we assume that $(\tilde{S}_0(\alpha), \bar{S}_0(\alpha))$, $\tilde{S}_1$, and $\tilde{S}_2$ are mutually independent with

$$p_{\tilde{S}_1}(0) = p_{\tilde{S}_2}(0) = H_b^{-1}(\frac{2}{3}),$$
$$p_{\tilde{S}_0(\alpha)}(0) = p_{\bar{S}_0(\alpha)}(0) = \alpha,$$
$$p_{\bar{S}_0(\alpha)|\tilde{S}_0(\alpha)}(1|0) = p_{\tilde{S}_0(\alpha)|\bar{S}_0(\alpha)}(1|0) = \beta(\alpha),$$
$$p_{\bar{S}_0(\alpha)|\tilde{S}_0(\alpha)}(0|1) = p_{\tilde{S}_0(\alpha)|\bar{S}_0(\alpha)}(0|1) = 1 - \frac{\alpha\beta(\alpha)}{1-\alpha},$$

where $\beta(\alpha)$ is the unique solution in $(0, 1-\alpha]$ of the following equation

$$H(\tilde{S}_0(\alpha), \bar{S}_0(\alpha)) = \log_2 3 - \frac{4}{3}.$$

Note that such S_1 and S_2 have no non-trivial common part, i.e., $H(S_0) = 0$. By setting $p_X(0) = p_X(1) = p_X(2) = \frac{1}{3}$, $V_0 = 0$, and $V_i = Y_i$, $i = 1, 2$, one can readily verify $p_{(S_1,S_2)}$ satisfies the necessary condition in Theorem 2 with $\kappa = 1$ for any $\alpha \in [H_b^{-1}(\frac{1}{2}\log_2 3 - \frac{2}{3}), H_b^{-1}(\log_2 3 - \frac{4}{3}))$. However, we shall show that this is not the case for Theorem 3. It is easy to see that, if the afore-described $p_{(S_1,S_2)}$ is admissible for the Blackwell channel $p^B_{Y_1,Y_2|X}$ subject to bandwidth expansion constraint κ, then, by Corollary 2 as well as the fact that $\mathcal{C}_{\text{out}}(p^B_{Y_1,Y_2|X}) = \mathcal{C}(p^B_{Y_1,Y_2|X})$, we must have

$$\mathcal{C}(p_{(S_1,S_2)}) \subseteq \kappa \mathcal{C}(p^B_{Y_1,Y_2|X}). \quad (9)$$

By choosing $U = (\tilde{S}_0(\alpha), \bar{S}_0(\alpha))$, one can readily verify that $(H_b(\alpha), \frac{2}{3}, \frac{2}{3})$ is contained in $\mathcal{C}(p_{(S_1, S_2)})$ for any $\alpha \in [H_b^{-1}(\frac{1}{2} \log_2 3 - \frac{2}{3}), H_b^{-1}(\log_2 3 - \frac{4}{3}))$. On the other hand, it follows from [2, Lemma 1] that $(\log_2 3 - \frac{4}{3}, \frac{2}{3}, \frac{2}{3})$ is not contained in $\mathcal{C}(p_{Y_1, Y_2|X}^B)$. Note that $(H_b(\alpha), \frac{2}{3}, \frac{2}{3})$ converges to $(\log_2 3 - \frac{4}{3}, \frac{2}{3}, \frac{2}{3})$ as $\alpha \to H_b^{-1}(\log_2 3 - \frac{4}{3})$. Since $\mathcal{C}(p_{Y_1, Y_2|X})$ is closed, it follows that $p_{(S_1, S_2)}$ violates (9) with $\kappa = 1$ (and consequently the necessary condition in Theorem 3 with $\kappa = 1$) when α is sufficiently close to $H_b^{-1}(\log_2 3 - \frac{4}{3})$.

This example indicates that choosing $U = S_0$ in Theorem 3 is not always optimal. In this sense, the common part between S_1 and S_2 does not play a fundamental role in the new outer bound; see [3] for a related observation.

V. APPROACH II: COMPARISON OF CAPACITY REGIONS

A. The Lossy Source Broadcast Problem

In a certain sense, the notion of dominance in Section IV hinges upon the converse results for broadcast channels. In this section we shall develop a different notion of dominance that is mainly based on the achievability results for broadcast channels. This notion is captured by the following lemma, which shows that every realizable virtual broadcast channel $p_{\hat{S}_1, \hat{S}_2|S}$ is dominated by the physical broadcast channel $p_{Y_1, Y_2|X}$ in the sense that Marton's inner bound of $p_{\hat{S}_1, \hat{S}_2|S}$ with input distribution p_S is contained in the capacity region of $p_{Y_1, Y_2|X}$. The proof of this lemma is relegated to Appendix D.

Lemma 2: If a virtual broadcast channel $p_{\hat{S}_1, \hat{S}_2|S}$ is realizable through the physical broadcast channel $p_{Y_1, Y_2|X}$ with bandwidth expansion ratio ρ, then

$$\mathcal{C}_{\mathsf{in}}(p_S, p_{\hat{S}_1, \hat{S}_2|S}) \subseteq \rho \mathcal{C}(p_{Y_1, Y_2|X}).$$

The following necessary condition for the lossy source broadcast problem is a simple consequence of Lemma 2.

Theorem 5: If distortion pair (d_1, d_2) is achievable under distortion measures w_1 and w_2 subject to bandwidth expansion constraint κ, then there exists a virtual broadcast channel $p_{\hat{S}_1, \hat{S}_2|S}$ with $\mathbb{E}[w_i(S, \hat{S}_i)] \leq d_i$, $i = 1, 2$, such that

$$\mathcal{C}_{\mathsf{in}}(p_S, p_{\hat{S}_1, \hat{S}_2|S}) \subseteq \kappa \mathcal{C}(p_{Y_1, Y_2|X}).$$

Proof: Let (d_1, d_2) be a distortion pair that is achievable under distortion measures w_1 and w_2 subject to bandwidth expansion constraint κ. In view of Definition 1 and the discussion in Section II, for every $\epsilon > 0$, there exists a virtual broadcast channel $p_{\hat{S}_1^{(\epsilon)}, \hat{S}_2^{(\epsilon)}|S}$ realizable through the physical broadcast channel $p_{Y_1, Y_2|X}$ with bandwidth expansion ratio $\rho \leq \kappa + \epsilon$ such that $\mathbb{E}[w_i(S, \hat{S}_i^{(\epsilon)})] \leq d_i + \epsilon$, $i = 1, 2$. It follows from Lemma 2 that, for such $p_{\hat{S}_1^{(\epsilon)}, \hat{S}_2^{(\epsilon)}|S}$, we have

$$\mathcal{C}_{\mathsf{in}}(p_S, p_{\hat{S}_1^{(\epsilon)}, \hat{S}_2^{(\epsilon)}|S}) \subseteq (\kappa + \epsilon) \mathcal{C}(p_{Y_1, Y_2|X}).$$

Since $\{p_{\hat{S}_1^{(\epsilon)}, \hat{S}_2^{(\epsilon)}|S} : \epsilon > 0\}$ can be viewed as a subset of $\{\pi \in \mathbb{R}_+^{|\mathcal{S}| \times |\hat{\mathcal{S}}_1| \times |\hat{\mathcal{S}}_2|} : \sum_{\hat{s}_1 \in \hat{\mathcal{S}}_1, \hat{s}_2 \in \hat{\mathcal{S}}_2} \pi(s, \hat{s}_1, \hat{s}_2) = 1, s \in \mathcal{S}\}$, which is compact under the Euclidean distance, one can find a sequence $\epsilon_1, \epsilon_2, \cdots$ converging to zero such that

$$\lim_{k \to \infty} p_{\hat{S}_1^{(\epsilon_k)}, \hat{S}_2^{(\epsilon_k)}|S} = p_{\hat{S}_1, \hat{S}_2|S}$$

for some $p_{\hat{S}_1, \hat{S}_2|S}$ with $\mathbb{E}[w_i(S, \hat{S}_i)] \leq d_i$, $i = 1, 2$. Now the proof can be completed via a simple limiting argument. ∎

B. Application to the Problem of Broadcasting Correlated Sources

In view of (6), one can readily deduce from Theorem 5 the following outer bound on the admissible source region.

Theorem 6: If $p_{(S_1, S_2)}$ is admissible for broadcast channel $p_{Y_1, Y_2|X}$ subject to bandwidth expansion constraint κ, then

$$\mathcal{C}(p_{(S_1, S_2)}) \subseteq \kappa \mathcal{C}(p_{Y_1, Y_2|X}). \tag{10}$$

The following result provides a complete characterization of the source admissible region and a rigorous justification of the source-message correspondence in the case where $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ form a Markov chain.

Corollary 3: A source distribution $p_{(S_1, S_2)}$ with $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ forming a Markov chain is admissible for broadcast channel $p_{Y_1, Y_2|X}$ subject to bandwidth expansion constraint κ if and only if

$$(H(S_0), H(S_1|S_0), H(S_2|S_0)) \in \kappa \mathcal{C}(p_{Y_1, Y_2|X}). \tag{11}$$

Proof: The proof of the "if" part is based on a simple separation-based scheme. The transmitter first compresses S_0^m via entropy coding and maps the resulting bits to the common message M_0; given S_0^m, the transmitter further compresses S_i^m via conditional entropy coding and maps the resulting bits to the private message M_i, $i = 1, 2$. Note that (11) ensures the existence of a good broadcast channel code such that receiver i can recover (M_0, M_i) and consequently S_i^m with high probability, $i = 1, 2$.

The "only if" part follows from Theorem 6 as well as the fact that $(H(S_0), H(S_1|S_0), H(S_2|S_0)) \in \mathcal{C}(p_{(S_1, S_2)})$ when $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ form a Markov chain. ∎

In view of Theorem 4 and Corollary 3, the necessary conditions in Theorem 3 and Theorem 6 are equivalent to (8) and (11), respectively, when $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ form a Markov chain. It is known [22] that in general $\mathcal{C}_{\mathsf{out}}(p_{Y_1, Y_2|X})$ can be strictly larger than $\mathcal{C}(p_{Y_1, Y_2|X})$. So it is possible to find an example for which

$$(H(S_0), H(S_1|S_0), H(S_2|S_0)) \notin \kappa \mathcal{C}(p_{Y_1, Y_2|X}),$$
$$(H(S_0), H(S_1|S_0), H(S_2|S_0)) \in \kappa \mathcal{C}_{\mathsf{out}}(p_{Y_1, Y_2|X}).$$

This means[9] that Theorem 6 cannot be deduced from Theorem 3.

Note that both (7) and (10) imply

$$\mathcal{C}(p_{(S_1, S_2)}) \subseteq \kappa \mathcal{C}_{\mathsf{out}}(p_{Y_1, Y_2|X}). \tag{12}$$

We shall show that (12) suffices to recover several existing results. Let $\mathcal{C}_{\mathsf{D}}(p_{Y_1, Y_2|X})$ denote the capacity region of broadcast channel $p_{Y_1, Y_2|X}$ with degraded message sets, i.e.,

$$\mathcal{C}_{\mathsf{D}}(p_{Y_1, Y_2|X}) = \{(R_0, R_2) : (R_0, 0, R_2) \in \mathcal{C}(p_{Y_1, Y_2|X})\}.$$

---

[9]We believe that Theorem 3 also cannot be deduced from Theorem 6.

It is known [23] that $\mathcal{C}_{\mathsf{D}}(p_{Y_1,Y_2|X})$ is given by the set of $(R_0, R_2) \in \mathbb{R}_+^2$ satisfying

$$R_0 \leq I(V; Y_1),$$
$$R_2 \leq I(X; Y_2|V),$$
$$R_0 + R_2 \leq I(X; Y_2)$$

for some $p_{V,X,Y_1,Y_2} = p_{V|X} p_X p_{Y_1,Y_2|X}$ with $|\mathcal{V}| \leq |\mathcal{X}| + 1$. Moreover, it can be verified that

$$\{(R_0, R_2) : (R_0, 0, R_2) \in \mathcal{C}_{\mathsf{out}}(p_{Y_1,Y_2|X})\} = \mathcal{C}_{\mathsf{D}}(p_{Y_1,Y_2|X}). \tag{13}$$

The following result is a special case of [24, Th. 2 and 3].

Corollary 4: A source distribution $p_{(S_1,S_2)}$ with S_1 being a deterministic function of S_2 is admissible for broadcast channel $p_{Y_1,Y_2|X}$ subject to bandwidth expansion constraint κ if and only if

$$(H(S_1), H(S_2|S_1)) \in \kappa \mathcal{C}_{\mathsf{D}}(p_{Y_1,Y_2|X}). \tag{14}$$

Proof: The proof of the "if" part is based on a simple separation-based scheme. The transmitter first compresses S_1^m via entropy coding and maps the resulting bits to the common message M_0; given S_1^m, the transmitter further compresses S_2^m via conditional entropy coding and maps the resulting bits to the private message M_2. Note that (14) ensures the existence of a good broadcast channel code such that receiver 1 can recover M_0 and consequently S_1^m with high probability while receiver 2 can recover (M_0, M_2) and consequently S_2^m with high probability.

The "only if" part follows by (12) and (13) as well as the fact that $(H(S_1), 0, H(S_2|S_1)) \in \mathcal{C}(p_{(S_1,S_2)})$ when S_1 is a deterministic function of S_2. ∎

We say $p_{Y_2|X}$ is more capable than $p_{Y_1|X}$ if $I(X; Y_2) \geq I(X; Y_1)$ for all p_X [25] [26, p. 121]. It can be verified that

$$\{(0, R_1, R_2) \in \mathcal{C}_{\mathsf{out}}(p_{Y_1,Y_2|X})\} = \mathcal{C}_{\mathsf{D}}(p_{Y_1,Y_2|X}) \tag{15}$$

when $p_{Y_2|X}$ is more capable than $p_{Y_1|X}$.

The following result [6, Th. 4] is a dual version of Corollary 4.

Corollary 5: A source distribution $p_{(S_1,S_2)}$ is admissible for broadcast channel $p_{Y_1,Y_2|X}$ (with $p_{Y_2|X}$ more capable than $p_{Y_1|X}$) subject to bandwidth expansion constraint κ if and only if (14) holds.

Proof: The proof of the "if" part is the same as that for Corollary 4. The "only if" part follows by (12) and (15) as well as the fact that $(0, H(S_1), H(S_2|S_1)) \in \mathcal{C}(p_{(S_1,S_2)})$. ∎

VI. CONCLUSION

We have established two necessary conditions for the lossy source broadcast problem (Theorem 1 and Theorem 5), from which new outer bounds on the admissible source region (Theorem 3 and Theorem 6) are deduced. It is expected that the idea of deriving converse results via suitable comparisons between the virtual channel (induced by the source(s) and the reconstruction(s)) and the physical channel has potential applications beyond the lossy source broadcast problem considered in the present paper.

APPENDIX A
PROOF OF LEMMA 1

Let $(X^{\rho m}, Y_1^{\rho m}, Y_2^{\rho m}, \hat{S}_1^m, \hat{S}_2^m)$ be jointly distributed with S^m according to

$$p_{S^m}(s^m) p_{X^{\rho m}|S^m}(x^{\rho m}|s^m) p_{Y_1^{\rho m}, Y_2^{\rho m}|X^{\rho m}}(y_1^{\rho m}, y_2^{\rho m}|x^{\rho m})$$

$$\times \prod_{i=1}^{2} p_{\hat{S}_i^m|Y_i^{\rho m}}(\hat{s}_i^m|y_i^{\rho m}), \tag{16}$$

where

$$p_{S^m}(s^m) = \prod_{t=1}^{m} p_S(s(t)),$$

$$p_{Y_1^{\rho m}, Y_2^{\rho m}|X^{\rho m}}(y_1^{\rho m}, y_2^{\rho m}|x^{\rho m}) = \prod_{q=1}^{\rho m}(y_1(q), y_2(q)|x(q)).$$

Note that the induced conditional distribution[10] $p_{\hat{S}_1^m, \hat{S}_2^m|S^m}$ is degraded with respect to $p_{Y_1^{\rho m}, Y_2^{\rho m}|X^{\rho m}}$; in fact, every $p_{\hat{S}_1^m, \hat{S}_2^m|S^m}$ degraded with respect to $p_{Y_1^{\rho m}, Y_2^{\rho m}|X^{\rho m}}$ can be obtained in this way.

Let $(U_0^m, \cdots, U_L^m)$ be jointly distributed with $(S^m, X^{\rho m}, Y_1^{\rho m}, Y_2^{\rho m}, \hat{S}_1^m, \hat{S}_2^m)$ such that $(U_0^m, \cdots, U_L^m) \leftrightarrow S^m \leftrightarrow (X^{\rho m}, Y_1^{\rho m}, Y_2^{\rho m}, \hat{S}_1^m, \hat{S}_2^m)$ form a Markov chain, and $(U_0(t), \cdots, U_L(t), S(t))$, $t = 1, \cdots, m$, are independent and identically distributed. Let T be a random variable independent of $(U_0^m, \cdots, U_L^m, S^m, \hat{S}_1^m, \hat{S}_2^m)$ and uniformly distributed over $\{1, \cdots, m\}$. Define

$$U_i = U_i(T), \quad i = 0, \cdots, L,$$
$$S = S(T),$$
$$\hat{S}_i = \hat{S}_i(T), \quad i = 1, 2.$$

The following properties of $(U_0, \cdots, U_L, S, \hat{S}_1, \hat{S}_2)$ can be easily verified:

1) the distribution of $(U_0, \cdots, U_L, S)$ is identical with that of $(U_0(t), U_1(t), U_2(t), S(t))$ for every t;
2) $(U_0, \cdots, U_L) \leftrightarrow S \leftrightarrow (\hat{S}_1, \hat{S}_2)$ form a Markov chain;
3) $p_{\hat{S}_1, \hat{S}_2|S}$ is the single-letterized version[11] of $p_{\hat{S}_1^m, \hat{S}_2^m|S^m}$.

Note that

$$\sum_{i=1}^{k} I(U_{\mathcal{A}_i}^m; \hat{S}_{a(i)}^m | U_{\cup_{j=1}^{i-1} \mathcal{A}_j}^m)$$

$$= \sum_{i=1}^{k} \sum_{t=1}^{m} I(U_{\mathcal{A}_i}(t); \hat{S}_{a(i)}^m | U_{\cup_{j=1}^{i-1} \mathcal{A}_j}^m, U_{\mathcal{A}_i}^{t-1})$$

$$= \sum_{i=1}^{k} \sum_{t=1}^{m} I(U_{\mathcal{A}_i}(t); \hat{S}_{a(i)}^m, U_{\cup_{j=1}^{i-1} \mathcal{A}_j}^{t-1}, U_{\cup_{j=1}^{i-1} \mathcal{A}_j, t+1}^m,$$

$$U_{\mathcal{A}_i}^{t-1} | U_{\cup_{j=1}^{i-1} \mathcal{A}_j}(t))$$

----

[10]If $p_S(s) > 0$ for all $s \in \mathcal{S}$, then $p_{\hat{S}_1^m, \hat{S}_2^m|S^m}$ is uniquely given by (4). If $p_S(s) = 0$ for some $s \in \mathcal{S}$, then the conditional distribution in (4) is not the only one that is compatible with the joint distribution in (16); in this case we simply use (4) as the definition of $p_{\hat{S}_1^m, \hat{S}_2^m|S^m}$.

[11]Strictly speaking, $p_{\hat{S}_1, \hat{S}_2|S}(\cdot, \cdot|s)$ is uniquely specified only for those $s \in \{s' \in \mathcal{S} : p_S(s') > 0\}$. However, this suffices for our purpose since the results in the present paper depend on $p_{\hat{S}_1, \hat{S}_2|S}$ only through $p_S p_{\hat{S}_1, \hat{S}_2|S}$.

$$\geq \sum_{i=1}^{k}\sum_{t=1}^{m} I(U_{\mathcal{A}_i}(t); \hat{S}_{a(i)}(t)|U_{\cup_{j=1}^{i-1}\mathcal{A}_j}(t))$$

$$= m\sum_{i=1}^{k} I(U_{\mathcal{A}_i}(T); \hat{S}_{a(i)}(T)|U_{\cup_{j=1}^{i-1}\mathcal{A}_j}(T), T)$$

$$= m\sum_{i=1}^{k} I(U_{\mathcal{A}_i}(T); \hat{S}_{a(i)}(T), T|U_{\cup_{j=1}^{i-1}\mathcal{A}_j}(T))$$

$$\geq m\sum_{i=1}^{k} I(U_{\mathcal{A}_i}(T); \hat{S}_{a(i)}(T)|U_{\cup_{j=1}^{i-1}\mathcal{A}_j}(T))$$

$$= m\sum_{i=1}^{k} I(U_{\mathcal{A}_i}; \hat{S}_{a(i)}|U_{\cup_{j=1}^{i-1}\mathcal{A}_j}).$$

On the other hand, we have

$$\sum_{i=1}^{k} I(U^m_{\mathcal{A}_i}; \hat{S}^m_{a(i)}|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j}) \leq \sum_{i=1}^{k} I(U^m_{\mathcal{A}_i}; Y^{\rho m}_{a(i)}|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j}). \tag{17}$$

We shall show that, for $l = 1, \cdots, k$,

$$\sum_{i=l}^{k} I(U^m_{\mathcal{A}_i}; Y^{\rho m}_{a(i)}|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j})$$

$$\leq \sum_{q=1}^{\rho m} I(Y^{\rho m}_{2,q+1}; Y_1(q)|U^m_{\cup_{j=1}^{l-1}\mathcal{A}_j}, Y_1^{q-1})$$

$$+ \sum_{i=l}^{k}\sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_{a(i)}(q)|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j}, Y_1^{q-1}, Y^{\rho m}_{2,q+1}), \tag{18}$$

which, in light of the Csiszár sum identity [26, p. 25], is equivalent to

$$\sum_{i=l}^{k} I(U^m_{\mathcal{A}_i}; Y^{\rho m}_{a(i)}|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j})$$

$$\leq \sum_{q=1}^{\rho m} I(Y_1^{q-1}; Y_2(q)|U^m_{\cup_{j=1}^{l-1}\mathcal{A}_j}, Y^{\rho m}_{2,q+1})$$

$$+ \sum_{i=l}^{k}\sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_{a(i)}(q)|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j}, Y_1^{q-1}, Y^{\rho m}_{2,q+1}). \tag{19}$$

First consider the case $l = k$. If $a(k) = 1$, we have

$$I(U^m_{\mathcal{A}_i}; Y^{\rho m}_{a(k)}|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j})$$

$$= \sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_1(q)|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j}, Y_1^{q-1})$$

$$\leq \sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}, Y^{\rho m}_{2,q+1}; Y_1(q)|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j}, Y_1^{q-1})$$

$$= \sum_{q=1}^{\rho m} I(Y^{\rho m}_{2,q+1}; Y_1(q)|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j}, Y_1^{q-1})$$

$$+ \sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_1(q)|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j}, Y_1^{q-1}, Y^{\rho m}_{2,q+1});$$

if $a(k) = 2$, we have

$$I(U^m_{\mathcal{A}_i}; Y^{\rho m}_{a(k)}|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j})$$

$$= \sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_2(q)|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j}, Y^{\rho m}_{2,q+1})$$

$$\leq \sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}, Y_1^{q-1}; Y_2(q)|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j}, Y^{\rho m}_{2,q+1})$$

$$= \sum_{q=1}^{\rho m} I(Y_1^{q-1}; Y_2(q)|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j}, Y^{\rho m}_{2,q+1})$$

$$+ \sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_2(q)|U^m_{\cup_{j=1}^{k-1}\mathcal{A}_j}, Y_1^{q-1}, Y^{\rho m}_{2,q+1}).$$

Therefore, (18) and (19) hold when $l = k$. Now we proceed by induction on l. If $a(l) = 1$, we have

$$\sum_{i=l}^{k} I(U^m_{\mathcal{A}_i}; Y^{\rho m}_{a(i)}|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j})$$

$$= I(U^m_{\mathcal{A}_l}; Y^{\rho m}_1|U^m_{\cup_{j=1}^{l-1}\mathcal{A}_j}) + \sum_{i=l+1}^{k} I(U^m_{\mathcal{A}_i}; Y^{\rho m}_{a(i)}|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j})$$

$$\leq I(U^m_{\mathcal{A}_l}; Y^{\rho m}_1|U^m_{\cup_{j=1}^{l-1}\mathcal{A}_j})$$

$$+ \sum_{q=1}^{\rho m} I(Y^{\rho m}_{2,q+1}; Y_1(q)|U^m_{\cup_{j=1}^{l}\mathcal{A}_j}, Y_1^{q-1})$$

$$+ \sum_{i=l+1}^{k}\sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_{a(i)}(q)|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j}, Y_1^{q-1}, Y^{\rho m}_{2,q+1}) \tag{20}$$

$$= \sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_l}; Y_1(q)|U^m_{\cup_{j=1}^{l-1}\mathcal{A}_j}, Y_1^{q-1})$$

$$+ \sum_{q=1}^{\rho m} I(Y^{\rho m}_{2,q+1}; Y_1(q)|U^m_{\cup_{j=1}^{l}\mathcal{A}_j}, Y_1^{q-1})$$

$$+ \sum_{i=l+1}^{k}\sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_{a(i)}(q)|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j}, Y_1^{q-1}, Y^{\rho m}_{2,q+1})$$

$$= \sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_l}, Y^{\rho m}_{2,q+1}; Y_1(q)|U^m_{\cup_{j=1}^{l-1}\mathcal{A}_j}, Y_1^{q-1})$$

$$+ \sum_{i=l+1}^{k}\sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_{a(i)}(q)|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j}, Y_1^{q-1}, Y^{\rho m}_{2,q+1})$$

$$= \sum_{q=1}^{\rho m} I(Y^{\rho m}_{2,q+1}; Y_1(q)|U^m_{\cup_{j=1}^{l-1}\mathcal{A}_j}, Y_1^{q-1})$$

$$+ \sum_{i=l}^{k}\sum_{q=1}^{\rho m} I(U^m_{\mathcal{A}_i}; Y_{a(i)}(q)|U^m_{\cup_{j=1}^{i-1}\mathcal{A}_j}, Y_1^{q-1}, Y^{\rho m}_{2,q+1}),$$

where (20) follows by the induction hypothesis. Therefore, (18) holds when $a(l) = 1$. Similarly, it can be shown that (19) holds when $a(l) = 2$. This finishes the induction argument.

Let Q be a random variable independent of $(U_0^m, \cdots, U_L^m, X^{\rho m}, Y_1^{\rho m}, Y_2^{\rho m})$ and uniformly distributed

over $\{1, \cdots, \rho m\}$. Define

$$V_i = (U_i^m, Y_1^{Q-1}, Y_{2,Q+1}^{\rho m}, Q), \quad i = 0, \cdots, L,$$
$$X = X(Q),$$
$$Y_i = Y_i(Q), \quad i = 1, 2.$$

It is clear that $(V_0, \cdots, V_L) \leftrightarrow X \leftrightarrow (Y_1, Y_2)$ form a Markov chain; moreover, p_X does not depend on the choice of $p_{U_0,\cdots,U_L|S}$. Continuing from (17),

$$\sum_{i=1}^{k} I(U_{\mathcal{A}_i}^m; \hat{S}_{a(i)}^m | U_{\cup_{j=1}^{i-1} \mathcal{A}_j}^m)$$

$$\leq \sum_{q=1}^{\rho m} I(U_{\mathcal{A}_1}^m, Y_1^{q-1}, Y_{2,q+1}^{\rho m}; Y_{a(1)}(q))$$

$$+ \sum_{i=2}^{k} \sum_{q=1}^{\rho m} I(U_{\mathcal{A}_i}^m; Y_{a(i)}(q) | U_{\cup_{j=1}^{i-1} \mathcal{A}_j}^m, Y_1^{q-1}, Y_{2,q+1}^{\rho m}) \quad (21)$$

$$= \rho m I(U_{\mathcal{A}_1}^m, Y_1^{Q-1}, Y_{2,Q+1}^{\rho m}; Y_{a(1)}(Q) | Q)$$

$$+ \rho m \sum_{i=2}^{k} I(U_{\mathcal{A}_i}^m; Y_{a(i)}(Q) | U_{\cup_{j=1}^{i-1} \mathcal{A}_j}^m, Y_1^{Q-1}, Y_{2,Q+1}^{\rho m}, Q)$$

$$\leq \rho m I(U_{\mathcal{A}_1}^m, Y_1^{Q-1}, Y_{2,Q+1}^{\rho m}, Q; Y_{a(1)}(Q))$$

$$+ \rho m \sum_{i=2}^{k} I(U_{\mathcal{A}_i}^m; Y_{a(i)}(Q) | U_{\cup_{j=1}^{i-1} \mathcal{A}_j}^m, Y_1^{Q-1}, Y_{2,Q+1}^{\rho m}, Q)$$

$$= \rho m \sum_{i=1}^{k} I(V_{\mathcal{A}_i}; Y_{a(i)} | V_{\cup_{j=1}^{i-1} \mathcal{A}_j}),$$

where (21) is due to (18) and (19) as well as the fact that

$$\sum_{q=1}^{\rho m} I(Y_{2,q+1}^{\rho m}; Y_1(q) | Y_1^{q-1})$$

$$+ \sum_{q=1}^{\rho m} I(U_{\mathcal{A}_1}^m; Y_{a(1)}(q) | Y_1^{q-1}, Y_{2,q+1}^{\rho m})$$

$$= \sum_{q=1}^{\rho m} I(Y_1^{q-1}; Y_2(q) | Y_{2,q+1}^{\rho m})$$

$$+ \sum_{q=1}^{\rho m} I(U_{\mathcal{A}_1}^m; Y_{a(1)}(q) | Y_1^{q-1}, Y_{2,q+1}^{\rho m})$$

$$\leq \sum_{q=1}^{\rho m} I(U_{\mathcal{A}_1}^m, Y_1^{q-1}, Y_{2,q+1}^{\rho m}; Y_{a(1)}(q)).$$

This completes the proof of Lemma 1.

APPENDIX B
PROOF OF THEOREM 1

According to Lemma 1, for every virtual broadcast channel $p_{\hat{S}_1,\hat{S}_2|S}$ realizable through the physical broadcast channel $p_{Y_1,Y_2|X}$ with bandwidth expansion ratio ρ, there exists an input distribution p_X such that, for any $p_{U_0,\cdots,U_L,S,\hat{S}_1,\hat{S}_2} =$

$p_{U_0,\cdots,U_L|S} p_S p_{\hat{S}_1,\hat{S}_2|S}$, one can find $p_{V_0,\cdots,V_L,X,Y_1,Y_2} = p_{V_0,\cdots,V_L|X} p_X p_{Y_1,Y_2|X}$ satisfying

$$\sum_{i=1}^{k} I(U_{\mathcal{A}_i}; \hat{S}_{a(i)} | U_{\cup_{j=1}^{i-1} \mathcal{A}_j}) \leq \rho \sum_{i=1}^{k} I(V_{\mathcal{A}_i}; Y_{a(i)} | V_{\cup_{j=1}^{i-1} \mathcal{A}_j}) \quad (22)$$

for any $\mathcal{A}_i \subseteq \{0, \cdots, L\}$ and $a(i) \in \{1, 2\}$, $i = 1, \cdots, k$. Now choose $L = 2$. Setting $k = 1$, $\mathcal{A}_1 = \{0\}$, and $a(1) = 1$ in (22) gives

$$I(U_0; \hat{S}_1) \leq \rho I(V_0; Y_1). \quad (23)$$

Setting $k = 1$, $\mathcal{A}_1 = \{0\}$, and $a(1) = 2$ in (22) gives

$$I(U_0; \hat{S}_2) \leq \rho I(V_0; Y_2). \quad (24)$$

Setting $k = 1$, $\mathcal{A}_1 = \{0, 1\}$, and $a(1) = 1$ in (22) gives

$$I(U_0, U_1; \hat{S}_1) \leq \rho I(V_0, V_1; Y_1). \quad (25)$$

Setting $k = 1$, $\mathcal{A}_1 = \{0, 2\}$, and $a(1) = 2$ in (22) gives

$$I(U_0, U_2; \hat{S}_2) \leq \rho I(V_0, V_2; Y_2). \quad (26)$$

Setting $k = 2$, $\mathcal{A}_1 = \{0\}$, $\mathcal{A}_2 = \{2\}$, $a(1) = 1$, and $a(2) = 2$ in (22) gives

$$I(U_0; \hat{S}_1) + I(U_2; \hat{S}_2|U_0) \leq \rho[I(V_0; Y_1) + I(V_2; Y_2|V_0)]. \quad (27)$$

Setting $k = 2$, $\mathcal{A}_1 = \{0\}$, $\mathcal{A}_2 = \{1\}$, $a(1) = 2$, and $a(2) = 1$ in (22) gives

$$I(U_0; \hat{S}_2) + I(U_1; \hat{S}_1|U_0) \leq \rho[I(V_0; Y_2) + I(V_1; Y_1|V_0)]. \quad (28)$$

Setting $k = 2$, $\mathcal{A}_1 = \{0, 1\}$, $\mathcal{A}_2 = \{2\}$, $a(1) = 1$, and $a(2) = 2$ in (22) gives

$$I(U_0, U_1; \hat{S}_1) + I(U_2; \hat{S}_2|U_0, U_1)$$
$$\leq \rho[I(V_0, V_1; Y_1) + I(V_2; Y_2|V_0, V_1)]. \quad (29)$$

Setting $k = 2$, $\mathcal{A}_1 = \{0, 2\}$, $\mathcal{A}_2 = \{1\}$, $a(1) = 2$, and $a(2) = 1$ in (22) gives

$$I(U_0, U_2; \hat{S}_2) + I(U_1; \hat{S}_1|U_0, U_2)$$
$$\leq \rho[I(V_0, V_2; Y_2) + I(V_1; Y_1|V_0, V_2)]. \quad (30)$$

Setting $k = 3$, $\mathcal{A}_1 = \{0\}$, $\mathcal{A}_2 = \{2\}$, $\mathcal{A}_3 = \{1\}$, $a(1) = a(3) = 1$, and $a(2) = 2$ in (22) gives

$$I(U_0; \hat{S}_1) + I(U_2; \hat{S}_2|U_0) + I(U_1; \hat{S}_1|U_0, U_2)$$
$$\leq \rho[I(V_0; Y_1) + I(V_2; Y_2|V_0) + I(V_1; Y_1|V_0, V_2)]. \quad (31)$$

Setting $k = 3$, $\mathcal{A}_1 = \{0\}$, $\mathcal{A}_2 = \{1\}$, $\mathcal{A}_3 = \{2\}$, $a(1) = a(3) = 2$, and $a(2) = 1$ in (22) gives

$$I(U_0; \hat{S}_2) + I(U_1; \hat{S}_1|U_0) + I(U_2; \hat{S}_2|U_0, U_1)$$
$$\leq \rho[I(V_0; Y_2) + I(V_1; Y_1|V_0) + I(V_2; Y_2|V_0, V_1)]. \quad (32)$$

Let $\mathcal{R}(p_S, p_{\hat{S}_1,\hat{S}_2|S})$ denote the set of $(r_1, \cdots, r_{10}) \in \mathbb{R}_+^{10}$ satisfying

$$r_1 \leq I(U_0; \hat{S}_1),$$
$$r_2 \leq I(U_0; \hat{S}_2),$$

$$r_3 \leq I(U_0, U_1; \hat{S}_1),$$
$$r_4 \leq I(U_0, U_2; \hat{S}_2),$$
$$r_5 \leq I(U_0; \hat{S}_1) + I(U_2; \hat{S}_2|U_0),$$
$$r_6 \leq I(U_0; \hat{S}_2) + I(U_1; \hat{S}_1|U_0),$$
$$r_7 \leq I(U_0, U_1; \hat{S}_1) + I(S; \hat{S}_2|U_0, U_1),$$
$$r_8 \leq I(U_0, U_2; \hat{S}_2) + I(S; \hat{S}_1|U_0, U_2),$$
$$r_9 \leq I(U_0; \hat{S}_1) + I(U_2; \hat{S}_2|U_0) + I(S; \hat{S}_1|U_0, U_2),$$
$$r_{10} \leq I(U_0; \hat{S}_2) + I(U_1; \hat{S}_1|U_0) + I(S; \hat{S}_2|U_0, U_1)$$

for some $p_{U_0,U_1,U_2,S,\hat{S}_1,\hat{S}_2} = p_{U_0,U_1,U_2|S} p_S p_{\hat{S}_1,\hat{S}_2|S}$; analogously, let $\mathcal{R}(p_X, p_{Y_1,Y_2|X})$ denote the set of $(r_1, \cdots, r_{10}) \in \mathbb{R}_+^{10}$ satisfying

$$r_1 \leq I(V_0; Y_1),$$
$$r_2 \leq I(V_0; Y_2),$$
$$r_3 \leq I(V_0, V_1; Y_1),$$
$$r_4 \leq I(V_0, V_2; Y_2),$$
$$r_5 \leq I(V_0; Y_1) + I(V_2; Y_2|V_0),$$
$$r_6 \leq I(V_0; Y_2) + I(V_1; Y_1|V_0),$$
$$r_7 \leq I(V_0, V_1; Y_1) + I(X; Y_2|V_0, V_1),$$
$$r_8 \leq I(V_0, V_2; Y_2) + I(X; Y_1|V_0, V_2),$$
$$r_9 \leq I(V_0; Y_1) + I(V_2; Y_2|V_0) + I(X; Y_1|V_0, V_2),$$
$$r_{10} \leq I(V_0; Y_2) + I(V_1; Y_1|V_0) + I(X; Y_2|V_0, V_1)$$

for some $p_{V_0,V_1,V_2,X,Y_1,Y_2} = p_{V_0,V_1,V_2|X} p_X p_{Y_1,Y_2|X}$. It can be shown (see [27, Remark 3.6]) that (23)-(32) can be stated equivalently as

$$\mathcal{R}(p_S, p_{\hat{S}_1,\hat{S}_2|S}) \subseteq \rho \mathcal{R}(p_X, p_{Y_1,Y_2|X}). \tag{33}$$

Moreover, the following argument by Nair [18] indicates that, to compute $\mathcal{R}(p_S, p_{\hat{S}_1,\hat{S}_2|S})$ and $\mathcal{R}(p_X, p_{Y_1,Y_2|X})$, it suffices to consider $|\mathcal{U}_0| \leq |\mathcal{S}| + 5$, $|\mathcal{V}_0| \leq |\mathcal{X}| + 5$, $|\mathcal{U}_i| \leq |\mathcal{S}|$, and $|\mathcal{V}_i| \leq |\mathcal{X}|$, $i = 1, 2$. We shall only give the proof for $\mathcal{R}(p_S, p_{\hat{S}_1,\hat{S}_2|S})$ since $\mathcal{R}(p_X, p_{Y_1,Y_2|X})$ can be treated in the same way. The main idea is that it suffices for U_1 and U_2 to preserve the extreme points of $\mathcal{R}(p_S, p_{\hat{S}_1,\hat{S}_2|S})$, and then U_0 can be used to convexify the region. Note that every convex combination of the constraints in the definition of $\mathcal{R}(p_S, p_{\hat{S}_1,\hat{S}_2|S})$ can be written in the form

$$\lambda_1 H(\hat{S}_1) + \lambda_2 H(\hat{S}_2) + \lambda_3 I(S; \hat{S}_1) + \lambda_4 I(S; \hat{S}_2)$$
$$+ \lambda_5 H(\hat{S}_1|U_0) + \lambda_6 H(\hat{S}_2|U_0) + \lambda_7 H(\hat{S}_1|U_0, U_1)$$
$$+ \lambda_8 H(\hat{S}_2|U_0, U_2) + \lambda_9 H(\hat{S}_1|U_0, U_2) + \lambda_{10} H(\hat{S}_2|U_0, U_1),$$

which depends on $p_{U_0,U_1,U_2,S}$ only through $p_{U_0,U_1,S}$ and $p_{U_0,U_2,S}$. First fix p_{U_0}. For every $U_0 = u_0$, one can find $p_{S,U_1|U_0}(\cdot, \cdot|u_0)$ with $|\mathcal{U}_1| \leq |\mathcal{S}|$ that preserves $p_{S|U_0}(\cdot|u_0)$ and $\lambda_7 H(\hat{S}_1|U_0 = u_0, U_1) + \lambda_{10} H(\hat{S}_2|U_0 = u_0, U_1)$; similarly, one can find $p_{S,U_2|U_0}(\cdot, \cdot|u_0)$ with $|\mathcal{U}_2| \leq |\mathcal{S}|$ that preserves $p_{S|U_0}(\cdot|u_0)$ and $\lambda_8 H(\hat{S}_2|U_0 = u_0, U_2) + \lambda_9 H(\hat{S}_1|U_0 = u_0, U_2)$. We can get a consistent joint distribution $p_{U_0,U_1,U_2,S}$ by setting $p_{U_0,U_1,U_2,S} = p_{U_0,S} p_{U_1|U_0,S} p_{U_2|U_0,S}$. Finally, it suffices to have $|\mathcal{U}_0| \leq |\mathcal{S}| + 5$ for preserving p_S, $H(\hat{S}_i|U_0)$, $H(\hat{S}_i|U_0, U_1)$, and $H(\hat{S}_i|U_0, U_2)$, $i = 1, 2$.

Let (d_1, d_2) be a distortion pair that is achievable under distortion measures w_1 and w_2 subject to bandwidth expansion constraint κ. In view of Definition 1 and the discussion in Section II, for every $\epsilon > 0$, there exists a virtual broadcast channel $p_{\hat{S}_1^{(\epsilon)}, \hat{S}_2^{(\epsilon)}|S}$ realizable through the physical broadcast channel $p_{Y_1,Y_2|X}$ with bandwidth expansion ratio $\rho \leq \kappa + \epsilon$ such that $\mathbb{E}[w_i(S, \hat{S}_i^{(\epsilon)})] \leq d_i + \epsilon$, $i = 1, 2$. It follows by (33) that, for such $p_{\hat{S}_1^{(\epsilon)}, \hat{S}_2^{(\epsilon)}|S}$, we have

$$\mathcal{R}(p_S, p_{\hat{S}_1^{(\epsilon)}, \hat{S}_2^{(\epsilon)}|S}) \subseteq (\kappa + \epsilon)\mathcal{R}(p_{X^{(\epsilon)}}, p_{Y_1,Y_2|X})$$

for some $p_{X^{(\epsilon)}}$. Since $\{(p_{\hat{S}_1^{(\epsilon)}, \hat{S}_2^{(\epsilon)}|S}, p_{X^{(\epsilon)}}) : \epsilon > 0\}$ can be viewed as a subset of $\{(\pi, \pi') \in \mathbb{R}_+^{|\mathcal{S}| \times |\hat{S}_1| \times |\hat{S}_2|} \times \mathbb{R}_+^{|\mathcal{X}|} : \sum_{\hat{s}_1 \in \hat{S}_1, \hat{s}_2 \in \hat{S}_2} \pi(s, \hat{s}_1, \hat{s}_2) = 1, s \in \mathcal{S},$ and $\sum_{x \in \mathcal{X}} \pi'(x) = 1\}$, which is compact under the Euclidean distance, one can find a sequence $\epsilon_1, \epsilon_2, \cdots$ converging to zero such that

$$\lim_{k \to \infty} p_{\hat{S}_1^{(\epsilon_k)}, \hat{S}_2^{(\epsilon_k)}|S} = p_{\hat{S}_1,\hat{S}_2|S},$$
$$\lim_{k \to \infty} p_{X^{(\epsilon_k)}} = p_X$$

for some $p_{\hat{S}_1,\hat{S}_2|S}$ with $\mathbb{E}[w_i(S, \hat{S}_i)] \leq d_i$, $i = 1, 2$, and p_X. Now a simple limiting argument yields

$$\mathcal{R}(p_S, p_{\hat{S}_1,\hat{S}_2|S}) \subseteq \kappa \mathcal{R}(p_X, p_{Y_1,Y_2|X}). \tag{34}$$

Note that $\mathcal{R}(p_X, p_{Y_1,Y_2|X})$ is a convex set. As a consequence, (34) holds if and only if $\kappa \mathcal{R}(p_X, p_{Y_1,Y_2|X})$ contains all extreme points of $\mathcal{R}(p_S, p_{\hat{S}_1,\hat{S}_2|S})$. To realize all such extreme points, it suffices to consider $|\mathcal{U}_0| \leq |\mathcal{S}|$. This completes the proof of Theorem 1.

APPENDIX C
PROOF OF THEOREM 4

We shall only prove that (8) implies the necessary condition in Theorem 3 when $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ form a Markov chain since the other direction is straightforward.

Note that the necessary condition in Theorem 3 can be written equivalently as

$$\mathcal{R}(p_{(S_1,S_2)}) \subseteq \kappa \mathcal{R}(p_X, p_{Y_1,Y_2|X}) \tag{35}$$

for some p_X, where $\mathcal{R}(p_{(S_1,S_2)})$ is the set of $(r_1, \cdots, r_{10}) \in \mathbb{R}_+^{10}$ satisfying

$$r_1 \leq I(U; S_1),$$
$$r_2 \leq I(U; S_2),$$
$$r_3 \leq H(S_1),$$
$$r_4 \leq H(S_2),$$
$$r_5 \leq I(U; S_1) + H(S_2|U),$$
$$r_6 \leq I(U; S_2) + H(S_1|U),$$
$$r_7 \leq I(U; S_1) + H(S_1, S_2|U),$$
$$r_8 \leq I(U; S_2) + H(S_1, S_2|U),$$
$$r_9 \leq I(U; S_1) + H(S_1, S_2|U),$$
$$r_{10} \leq I(U; S_2) + H(S_1, S_2|U)$$

for some $p_{U,(S_1,S_2)} = p_{U|(S_1,S_2)} p_{(S_1,S_2)}$ with $|\mathcal{U}| \leq |\mathcal{S}| + 2$, and $\mathcal{R}(p_X, p_{Y_1,Y_2|X})$ is defined in Appendix B. On the

other hand, (8) can be written equivalently as

$$(H(S_0), H(S_1|S_0), H(S_2|S_0)) \in \kappa \mathcal{C}_{\text{out}}(p_X, p_{Y_1,Y_2|X}) \quad (36)$$

for some p_X. Therefore, it suffices to show that (36) implies (35) when $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ form a Markov chain. Throughout the proof we assume p_X is fixed.

It is clear that both $\mathcal{R}(p_{(S_1,S_2)})$ and $\mathcal{R}(p_X, p_{Y_1,Y_2|X})$ are closed convex sets. Let $\lambda_1, \cdots, \lambda_{10}$ be arbitrary non-negative numbers. We have

$$\max_{(r_1,\cdots,r_{10})\in\mathcal{R}(p_{(S_1,S_2)})} \sum_{i=1}^{10} \lambda_i r_i$$

$$= \max_{p_{U|(S_1,S_2)}} \lambda_1 I(U; S_1) + \lambda_2 I(U; S_2) + \lambda_3 H(S_1)$$
$$+ \lambda_4 H(S_2) + \lambda_5[I(U; S_1) + H(S_2|U)]$$
$$+ \lambda_6[I(U; S_2) + H(S_1|U)]$$
$$+ \lambda_7[I(U; S_1) + H(S_1, S_2|U)]$$
$$+ \lambda_8[I(U; S_2) + H(S_1, S_2|U)]$$
$$+ \lambda_9[I(U; S_1) + H(S_1, S_2|U)]$$
$$+ \lambda_{10}[I(U; S_2) + H(S_1, S_2|U)] \quad (37)$$

$$= \max_{p_{U|(S_1,S_2)}} (\lambda_1 + \lambda_3 + \lambda_5 + \lambda_7 + \lambda_9)H(S_1)$$
$$+ (\lambda_2 + \lambda_4 + \lambda_6 + \lambda_8 + \lambda_{10})H(S_2)$$
$$- (\lambda_1 + \lambda_5 - \lambda_6 + \lambda_7 + \lambda_9)H(S_1|U)$$
$$- (\lambda_2 - \lambda_5 + \lambda_6 + \lambda_8 + \lambda_{10})H(S_2|U)$$
$$+ (\lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})H(S_1, S_2|U)$$

$$= \max_{p_{U|(S_1,S_2)}} (\lambda_1 + \lambda_3 + \lambda_5 + \lambda_7 + \lambda_9)H(S_1)$$
$$+ (\lambda_2 + \lambda_4 + \lambda_6 + \lambda_8 + \lambda_{10})H(S_2)$$
$$- (\lambda_1 + \lambda_2)H(S_0|U)$$
$$- (\lambda_1 + \lambda_5 - \lambda_6 + \lambda_7 + \lambda_9)H(S_1|S_0, U)$$
$$- (\lambda_2 - \lambda_5 + \lambda_6 + \lambda_8 + \lambda_{10})H(S_2|S_0, U)$$
$$+ (\lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})H(S_1, S_2|S_0, U)$$

$$\leq \max_{p_{U|(S_1,S_2)}} (\lambda_1 + \lambda_3 + \lambda_5 + \lambda_7 + \lambda_9)H(S_1)$$
$$+ (\lambda_2 + \lambda_4 + \lambda_6 + \lambda_8 + \lambda_{10})H(S_2)$$
$$- (\lambda_1 + \lambda_2)H(S_0|U)$$
$$- (\lambda_1 + \lambda_5 - \lambda_6 - \lambda_8 - \lambda_{10})H(S_1|S_0, U)$$
$$- (\lambda_2 - \lambda_5 + \lambda_6 - \lambda_7 - \lambda_9)H(S_2|S_0, U), \quad (38)$$

where the last inequality follows from the fact that

$$H(S_1, S_2|S_0, U) \leq H(S_1|S_0, U) + H(S_2|S_0, U). \quad (39)$$

Let $a = \lambda_1 + \lambda_5 - \lambda_6 - \lambda_8 - \lambda_{10}$ and $b = \lambda_2 - \lambda_5 + \lambda_6 - \lambda_7 - \lambda_9$. Consider the following four possible cases.

1) $a \leq 0$ and $b \leq 0$: The maximum value of (38) is attained when $U = S_0$.
2) $a \geq 0$ and $b \leq 0$: The maximum value of (38) is attained when $U = S_1$.
3) $a \leq 0$ and $b \geq 0$: The maximum value of (38) is attained when $U = S_2$.
4) $a \geq 0$ and $b \geq 0$: The maximum value of (38) is attained when $U = (S_1, S_2)$.

It is clear that the equality holds in (39) for the following four choices of U:
1) $U = S_0$,
2) $U = S_1$,
3) $U = S_2$,
4) $U = (S_1, S_2)$.

Therefore, the maximum value of (37) is also attained by one of these four choices of U; as a consequence, for the necessary condition in Theorem 3, there is no loss of generality in restricting U to such choices. Note that (36) can be expressed alternatively as

$$H(S_0) \leq \kappa \min\{I(V_0^*; Y_1), I(V_0^*; Y_2)\}, \quad (40)$$
$$H(S_1) \leq \kappa[\min\{I(V_0^*; Y_1), I(V_0^*; Y_2)\} + I(V_1^*; Y_1|V_0^*)], \quad (41)$$
$$H(S_2) \leq \kappa[\min\{I(V_0^*; Y_1), I(V_0^*; Y_2)\} + I(V_2^*; Y_2|V_0^*)], \quad (42)$$
$$H(S_1, S_2) \leq \kappa[\min\{I(V_0^*; Y_1), I(V_0^*; Y_2)\} + I(V_1^*; Y_1|V_0^*) + I(X; Y_2|V_0^*, V_1^*)], \quad (43)$$
$$H(S_1, S_2) \leq \kappa[\min\{I(V_0^*; Y_1), I(V_0^*; Y_2)\} + I(V_2^*; Y_2|V_0^*) + I(X; Y_1|V_0^*, V_2^*)] \quad (44)$$

for some $p_{V_0^*,V_1^*,V_2^*,X,Y_1,Y_2} = p_{V_0^*,V_1^*,V_2^*|X} p_X p_{Y_1,Y_2|X}$. Setting $U = S_0$ in Theorem 3 yields the same set of constraints. On the other hand, when $U = S_1$, the necessary condition in Theorem 3 can be written as

$$H(S_0) \leq \kappa I(V_0; Y_2), \quad (45)$$
$$H(S_1) \leq \kappa I(V_0; Y_1), \quad (46)$$
$$H(S_2) \leq \kappa I(V_0, V_2; Y_2), \quad (47)$$
$$H(S_2) \leq \kappa[I(V_0; Y_2) + I(V_1; Y_1|V_0) + I(X; Y_2|V_0, V_1)], \quad (48)$$
$$H(S_1, S_2) \leq \kappa[I(V_0; Y_1) + I(V_2; Y_2|V_0)], \quad (49)$$
$$H(S_1, S_2) \leq \kappa[I(V_0, V_1; Y_1) + I(X; Y_2|V_0, V_1)] \quad (50)$$

for some $p_{V_0,V_1,V_2,X,Y_1,Y_2} = p_{V_0,V_1,V_2|X} p_X p_{Y_1,Y_2|X}$. By choosing $V_0 = V_1 = (V_0^*, V_1^*)$ and $V_2 = X$, we can see that (40)$\Rightarrow$(45), (41)$\Rightarrow$(46), (42)$\Rightarrow$(47), (42)$\Rightarrow$(48), (43)$\Rightarrow$(49), and (43)$\Rightarrow$(50). The case $U = S_2$ follows by symmetry. When $U = (S_1, S_2)$, the necessary condition in Theorem 3 can be written as

$$H(S_1) \leq \kappa I(V_0; Y_1), \quad (51)$$
$$H(S_2) \leq \kappa I(V_0; Y_2), \quad (52)$$

for some $p_{V_0,X,Y_1,Y_2} = p_{V_0|X} p_X p_{Y_1,Y_2|X}$. By choosing $V_0 = X$, we can see that (41)$\Rightarrow$(51) and (42)$\Rightarrow$(52). Hence, (35) is indeed implied by (36) when $S_1 \leftrightarrow S_0 \leftrightarrow S_2$ form a Markov chain. This completes the proof of Theorem 4.

APPENDIX D
PROOF OF LEMMA 2

Let $(\hat{S}_1^m, \hat{S}_2^m)$ be jointly distributed with S^m according to

$$p_{S^m}(s^m) p_{\hat{S}_1^m, \hat{S}_2^m|S^m}(\hat{s}_1^m, \hat{s}_2^m|s^m),$$

where $p_{S^m}(s^m) = \prod_{t=1}^{m} p_S(s(t))$. We assume that $p_{\hat{S}_1^m, \hat{S}_2^m|S^m}$ is degraded with respect to $p_{Y_1^{\rho m}, Y_2^{\rho m}|X^m}$,

where $p_{Y_1^{\rho m}, Y_2^{\rho m}|X^m}(y_1^{\rho m}, y_2^{\rho m}|x^m) = \prod_{q=1}^{\rho m} p_{Y_1, Y_2|X}$ $(y_1(q), y_2(q)|x(q))$. As a consequence,

$$\rho \mathcal{C}(p_{\hat{S}_1^m, \hat{S}_2^m|S^m}) \subseteq \mathcal{C}(p_{Y_1^{\rho m}, Y_2^{\rho m}|X^m}) = \rho m \mathcal{C}(p_{Y_1, Y_2|X}). \quad (53)$$

Let $\tilde{\mathcal{C}}_{\mathsf{in}}(p_{S^m}, p_{\hat{S}_1^m, \hat{S}_2^m|S^m})$ denote the set of $(R_0, R_1, R_2) \in \mathbb{R}_+^3$ satisfying

$$R_0 \le \min\{I(U_0^m; \hat{S}_1^m), I(U_0^m; \hat{S}_2^m)\},$$
$$R_0 + R_i \le I(U_0^m, U_i^m; \hat{S}_i^m), \quad i = 1, 2,$$
$$R_0 + R_1 + R_2 \le \min\{I(U_0^m; \hat{S}_1^m), I(U_0^m; \hat{S}_2^m)\}$$
$$+ I(U_1^m; \hat{S}_1^m|U_0^m) + I(U_2^m; \hat{S}_2^m|U_0^m)$$
$$- I(U_1^m; U_2^m|U_0^m)$$

for some (U_0^m, U_1^m, U_2^m) jointly distributed with $(S^m, \hat{S}_1^m, \hat{S}_2^m)$ such that $(U_0^m, U_1^m, U_2^m) \leftrightarrow S^m \leftrightarrow (\hat{S}_1^m, \hat{S}_2^m)$ form a Markov chain, and $(U_0(t), U_1(t), U_2(t), S(t))$, $t = 1, \cdots, m$, are independent and identically distributed. It is clear that

$$\tilde{\mathcal{C}}_{\mathsf{in}}(p_{S^m}, p_{\hat{S}_1^m, \hat{S}_2^m|S^m}) \subseteq \mathcal{C}_{\mathsf{in}}(p_{S^m}, p_{\hat{S}_1^m, \hat{S}_2^m|S^m})$$
$$\subseteq \mathcal{C}(p_{\hat{S}_1^m, \hat{S}_2^m|S^m}). \quad (54)$$

Let T be a random variable independent of $(U_0^m, U_1^m, U_2^m, S^m, \hat{S}_1^m, \hat{S}_2^m)$ and uniformly distributed over $\{1, \cdots, m\}$. Define

$$U_i = U_i(T), \quad i = 0, 1, 2,$$
$$S = S(T),$$
$$\hat{S}_i = \hat{S}_i(T), \quad i = 1, 2.$$

Note that

$$I(U_0^m; \hat{S}_i^m) = \sum_{t=1}^m I(U_0(t); \hat{S}_i^m|U_0^{t-1})$$
$$= \sum_{t=1}^m I(U_0(t); \hat{S}_i^m, U_0^{t-1})$$
$$\ge \sum_{t=1}^m I(U_0(t); \hat{S}_i(t))$$
$$= m I(U_0(T); \hat{S}_i(T)|T)$$
$$= m I(U_0(T); \hat{S}_i(T), T)$$
$$\ge m I(U_0(T); \hat{S}_i(T))$$
$$= m I(U_0; \hat{S}_i), \quad i = 1, 2;$$

moreover,

$$I(U_i^m; \hat{S}_i^m|U_0^m)$$
$$= \sum_{t=1}^m I(U_i(t); \hat{S}_i^m|U_0^m, U_i^{t-1}),$$
$$= \sum_{t=1}^m I(U_i(t); \hat{S}_i^m, U_0^{t-1}, U_{0,t+1}^m, U_i^{t-1}|U_0(t))$$
$$\ge \sum_{t=1}^m I(U_i(t); \hat{S}_i(t)|U_0(t))$$
$$= m I(U_i(T); \hat{S}_i(T)|U_0(T), T)$$
$$= m I(U_i(T); \hat{S}_i(T), T|U_0(T))$$
$$\ge m I(U_i(T); \hat{S}_i(T)|U_0(T))$$
$$= m I(U_i; \hat{S}_i|U_0), \quad i = 1, 2,$$

and

$$I(U_1^m; U_2^m|U_0^m) = m I(U_1; U_2|U_0).$$

Therefore, we have

$$m \mathcal{C}_{\mathsf{in}}(p_S, p_{\hat{S}_1, \hat{S}_2|S}) \subseteq \tilde{\mathcal{C}}_{\mathsf{in}}(p_{S^m}, p_{\hat{S}_1^m, \hat{S}_2^m|S^m}). \quad (55)$$

Combining (53), (54), and (55) completes the proof of Lemma 2.

ACKNOWLEDGMENT

REFERENCES

[1] T. S. Han and M. H. M. Costa, "Broadcast channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 5, pp. 641–650, Sep. 1987.

[2] G. Kramer and C. Nair, "Comments on 'broadcast channels with arbitrarily correlated sources,'" in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jun./Jul. 2009, pp. 2777–2779.

[3] P. Minero and Y.-H. Kim, "Correlated sources over broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jun./Jul. 2009, pp. 2780–2784.

[4] A. A. Gohari and V. Anantharam, "An outer bound to the admissible source region of broadcast channels with arbitrarily correlated sources and channel variations," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Monticello, IL, USA, Sep. 2008, pp. 301–308.

[5] A. A. Gohari and V. Anantharam. (2010). "Converses for discrete memoryless multiterminal networks." [Online]. Available: http://www.eecs.berkeley.edu/~ananth/2008+/ConversesForDMMN.pdf

[6] G. Kramer, Y. Liang, and S. Shamai (Shitz), "Outer bounds on the admissible source region for broadcast channels with dependent sources," in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, USA, Feb. 2009, pp. 169–172.

[7] C. Nair. (2011). "A note on outer bounds for broadcast channel." [Online]. Available: http://arxiv.org/abs/1101.0640

[8] A. Gohari and V. Anantharam, "Infeasibility proof and information state in network information theory," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5992–6004, Oct. 2014.

[9] K. Khezeli and J. Chen, "An improved outer bound on the admissible source region for broadcast channels with correlated sources," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 466–470.

[10] C. Tian, S. Diggavi, and S. Shamai, "Approximate characterizations for the Gaussian source broadcast distortion region," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 124–136, Jan. 2011.

[11] C. Tian, J. Chen, S. N. Diggavi, and S. Shamai (Shitz), "Optimality and approximate optimality of source-channel separation in networks," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 904–918, Feb. 2014.

[12] A. A. Gohari and V. Anantharam, "Evaluation of Marton's inner bound for the general broadcast channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 608–619, Feb. 2012.

[13] A. Gohari, private communication, Oct. 2013.

[14] 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 (ISIT)*, Istanbul, Turkey, Jul. 2013, pp. 1272–1276.

[15] S. I. Gel'fand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Inf. Transmiss.*, vol. 16, no. 1, pp. 17–25, Jan./Mar. 1980.

[16] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest, Hungary: Akadémiai Kiadó, 1981.

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

[18] C. Nair, private communication, Oct./Nov. 2013.

[19] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems Control Inf. Theory*, vol. 2, no. 2, pp. 149–162, 1973.

[20] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM J. Appl. Math.*, vol. 28, no. 1, pp. 100–113, Jan. 1975.

[21] T. S. Han, "The capacity region for the deterministic broadcast channel with a common message (Coresp.)," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 122–125, Jan. 1981.

[22] Y. Geng, A. Gohari, C. Nair, and Y. Yu, "The capacity region for two classes of product broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Saint Petersburg, Russia, Jul./Aug. 2011, pp. 1544–1548.

[23] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 60–64, Jan. 1977.

[24] W. Kang and G. Kramer, "Broadcast channel with degraded source random variables and receiver side information," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Toronto, ON, Canada, Jul. 2008, pp. 1711–1715.

[25] J. Körner and K. Marton, "Comparison of two noisy channels," in *Topics in Information Theory* (Colloquia Mathematica Societatis), I. Csiszár and P. Elias, Eds. Amsterdam, The Netherlands: North Holland, 1977, pp. 411–423.

[26] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[27] 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.

Kia Khezeli received the B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2012 and the M.A.Sc. degree in electrical and computer engineering from McMaster University, Hamilton, ON, Canada in 2014. He has been pursuing a Ph.D. degree in electrical and computer engineering at Cornell University, Ithaca, NY, since August 2014. His research interests include information theory and stochastic optimization. He received the Outstanding Thesis Award from McMaster University and the Jacobs Fellowship from Cornell University in 2014.

Jun Chen (S'03–M'06) received the B.E. degree with honors in communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2001 and the M.S. and Ph.D. degrees in electrical and computer engineering from Cornell University, Ithaca, NY, in 2004 and 2006, respectively.

He was a Postdoctoral Research Associate in the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign, Urbana, IL, from September 2005 to July 2006, and a Postdoctoral Fellow at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, from July 2006 to August 2007. Since September 2007 he has been with the Department of Electrical and Computer Engineering at McMaster University, Hamilton, ON, Canada, where he is currently an Associate Professor. His research interests include information theory, wireless communications, and signal processing.

He received several awards for his research, including the Josef Raviv Memorial Postdoctoral Fellowship in 2006, the Early Researcher Award from the Province of Ontario in 2010, and the IBM Faculty Award in 2010. He is currently serving as an Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.