State-Dependent Broadcast Channels with Reversible Input Constraints

Viswanathan Ramachandran

***Abstract*—A joint message communication and input reconstruction problem over a two-user state-dependent degraded discrete-memoryless broadcast channel is considered. The state process is assumed to be independent and identically distributed (i.i.d.), and known non-causally at the transmitter as well as the non-degraded receiver. The two receivers have to decode the messages from the transmitter, while the degraded receiver also needs to estimate the channel input codeword to meet a prescribed distortion limit. A complete characterization of the optimal rates versus distortion performance is provided. The tight characterization is also illustrated by means of an example of an additive binary broadcast channel with a Hamming distortion constraint for the input reconstruction at the degraded receiver.**

***Index Terms*—Broadcast channel, codeword reconstruction, constrained communication, Gelfand-Pinsker coding, Hamming distortion, rate-distortion, state-dependent channel.**

I. Introduction

COMMUNICATION over channels whose transition probability depends upon an external state process have long been studied [1]–[5]. Much attention has been dedicated in the literature to the case of independent and identically distributed (i.i.d.) states that are available at either the encoder or the decoder, or both. State-dependent channels have also been investigated in multi-terminal settings, see for instance the broadcast model in [6]. The reader is referred to the survey work [7] for a comprehensive overview of channel coding problems in the presence of state information.

In a state-dependent communication system, the decoder is often interested in reconstructing certain information embedded in the transmission, in addition to reliably recovering the source message. Sutivong *et al.* [8], for example, considered a simultaneous message transmission and channel state amplification problem. They provided a complete characterisation of the optimal tradeoff between message transmission rate and state information reconstruction distortion over the Gaussian dirty paper channel [9] as assessed by mean-squared error measure. Kim *et al.* [10] then extended this problem from the Gaussian to the general discrete-memoryless scenario, with state information reconstruction fidelity evaluated by a list-size uncertainty reduction metric. Choudhuri *et al.* [11] subsequently analyzed the casual discrete-memoryless case with a general distortion metric. A multi-user extension of state estimation in a Gaussian multiple-access scenario was addressed in [12], wherein a complete characterization of the optimal rate-distortion trade-off region was obtained. This state information reconstruction problem sparked extensive research on other information reconstruction problems such as state masking [13], [14], partial/remote state estimation [15], [16], common reconstruction [17], [18], and information embedding motivated by Witsenhausen's counterexample [19], [20].

Communication systems with constrained inputs have been studied in various contexts. For instance, Sumszyk and Steinberg [21] focused on lossless input reconstruction over a state-dependent single-user channel in an information embedding context. Bandemer and El Gamal [22] discussed a setting where the amount of disturbance caused to other users by a given transmission must be minimized. Zhang *et al.* [23] investigated channel coding subject to signal estimation constraints in the absence of state information at the encoder. However, we note that multi-terminal settings with communication and input reconstruction constraints have received scant attention in the literature, which is of interest in several applications, as described below.

In this study, we address a communication problem over a discrete-memoryless state-dependent degraded broadcast channel with an input reconstruction constraint at the degraded receiver, with noncausal state information at the transmitter and the non-degraded receiver. This is of interest, for instance in watermarking systems [24], where the encoder must encode messages to two receivers, one of which has access to the covertext (which corresponds to the state process), while the other receiver must be able to reproduce the stegotext (which corresponds to the channel input codeword) for retransmission to another destination. For this problem, we explore and characterize the optimal tradeoff between reliable communication rates and reconstruction distortion using a general distortion measure. Our bounds also yield a tight result for a binary broadcast channel with Hamming distortion. The achievability proof relies on adapting the approach for state-dependent broadcast channels from [6], while the outer bound relies on appropriate auxiliary identifications to manipulate the non-i.i.d. channel input sequence. Compared to a joint source-channel coding problem (see for instance [25, Chapter 3]) which usually involves transmission of an i.i.d. source over a given channel in a lossy/lossless manner, the problem addressed in this paper instead involves an estimation of the message-bearing non-i.i.d. channel input codeword at the receiver. The term 'reversible input constraints' is used in this context to refer to such recovery of the input from an observation of the channel output sequence

The rest of the paper is organized as follows. The sys-

Manuscript received May 3, 2023; revised August 4, 2023; approved for publication by Kim, Sang-Hyo, Division 1 Editor, September. 9, 2023.

The author is with the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden, email: visra@kth.se.

Digital Object Identifier: 10.23919/JCN.2023.000045

1229-2370/23/$10.00 © 2023 KICS

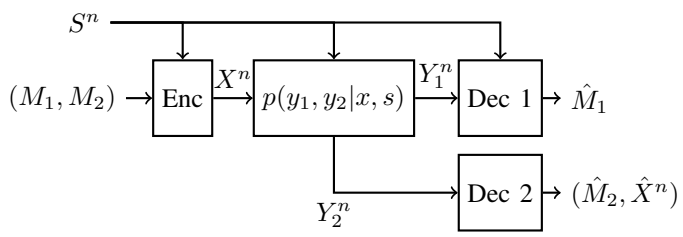

Fig. 1. Channel model, where $p(y_1, y_2|x, s) = p(y_1|x, s)p(y_2|y_1)$.

tem model and main result (Theorem. 1) are introduced in Section II. Section III contains the proof of achievability for Theorem. 1 while its converse is proved in Section IV. A tight characterization is worked out for the binary additive broadcast channel in Section V. Concluding remarks are given in Section VI.

Notations: Random variables/vectors are denoted by uppercase letters, while their realizations are denoted by the corresponding lower case letters. A sequence $(X_1, X_2, \cdots, X_n)$ is denoted by X^n for convenience.

II. System Model and Results

Consider the system shown in Fig. 1. Let $\mathcal{X}, \mathcal{S}, \mathcal{Y}_1, \mathcal{Y}_2, \hat{\mathcal{X}}$ be finite sets, and $p_S(s)$ be a probability mass function on $\mathcal{S}$. The given system is that of a discrete-memoryless broadcast channel (DM-BC) with discrete memoryless states, denoted by the tuple $(\mathcal{S}, p_S(s), \mathcal{X}, p_{Y_1,Y_2|X,S}(y_1, y_2|x, s), \mathcal{Y}_1, \mathcal{Y}_2)$. The channel is characterized by an input alphabet $\mathcal{X}$, state alphabet $\mathcal{S}$, output alphabet $\mathcal{Y}_1 \times \mathcal{Y}_2$, and conditional probability mass function $p_{Y_1,Y_2|X,S}(y_1, y_2|x, s)$. The channel and the states are assumed to be memoryless, i.e.,

$$p(y_1^n, y_2^n|x^n, s^n) = \prod_{i=1}^{n} p(y_{1i}, y_{2i}|x_i, s_i), \;\; p(s^n) = \prod_{i=1}^{n} p(s_i).$$

We assume that the channel is degraded, i.e.,

$$p_{Y_1,Y_2|X,S}(y_1, y_2|x, s) = p_{Y_1|X,S}(y_1|x, s)p_{Y_2|Y_1}(y_2|y_1). \quad (1)$$

The channel state is assumed to be known non-causally to the transmitter as well as the non-degraded receiver Y_1 (also referred to as the strong receiver), while the degraded receiver Y_2 (also known as the weak receiver) is uninformed. We wish to communicate messages (M_1, M_2) over the channel, and at the same time also reconstruct the input codeword at the weak receiver (depicted $\hat{X}^n$ in Fig. 1) to meet a distortion constraint. The distortion measure is defined as

$$d : \mathcal{X} \times \hat{\mathcal{X}} \to [0, \infty), \quad (2)$$

where $\hat{\mathcal{X}}$ is the reconstruction alphabet.

Definition 1. *An $(n, 2^{nR_1}, 2^{nR_2})$ code consists of two message sets $\{1, 2, \cdots, 2^{nR_j}\}, j \in \{1, 2\}$ on which M_j are assumed to be uniformly distributed, an encoder map that assigns a codeword $x^n \in \mathcal{X}^n$ to each $m_j \in \{1, 2, \cdots, 2^{nR_j}\}, j \in \{1, 2\}$ and $s^n \in \mathcal{S}^n$, a decoder map at the strong receiver that assigns an estimate $\hat{m}_1 \in \{1, 2, \cdots, 2^{nR_1}\}$ to each received sequence $(y_1^n, s^n) \in \mathcal{Y}_1^n \times \mathcal{S}^n$, and a decoder map at the weak receiver that assigns a pair of estimates $(\hat{m}_2, \hat{x}^n) \in \{1, 2, \cdots, 2^{nR_2}\} \times \hat{\mathcal{X}}^n$ to each received sequence $y_2^n \in \mathcal{Y}_2^n$. The average probability of error is given by*

$$P_e^{(n)} = \Pr\Big((\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)\Big). \quad (3)$$

A rate-distortion triple (R_1, R_2, D) is said to be achievable if there exists a sequence of $(n, 2^{nR_1}, 2^{nR_2})$ codes such that

$$\lim_{n\to\infty} P_e^{(n)} = 0, \quad (4)$$

and

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\Big[d(X_i, \hat{X}_i)\Big] \leq D. \quad (5)$$

The rate-distortion trade-off region $\mathcal{C}$ is the closure of the set of all achievable (R_1, R_2, D) triples. The main result of this paper is stated next.

Theorem 1. *The rate-distortion trade-off region $\mathcal{C}$ for the degraded DM-BC with non-causal state information known to the encoder as well as the strong receiver, and input reconstruction constraints at the weak receiver, is the closure of the set that contains all triples (R_1, R_2, D) that satisfy*

$$R_2 \leq I(U; Y_2) - I(U; S), \quad (6)$$

$$R_1 \leq I(X; Y_1|U, S), \quad (7)$$

$$D \geq \mathbb{E}[d(X, \phi(U, Y_2))], \quad (8)$$

for some joint probability distribution of the form

$$P_{S,U,X,Y_1,Y_2} = P_S P_{U|S} P_{X|U,S} P_{Y_1|X,S} P_{Y_2|Y_1}, \quad (9)$$

and some function $\phi : \mathcal{U} \times \mathcal{Y}_2 \to \hat{\mathcal{X}}$, with the auxiliary random variable cardinality bounded as $|\mathcal{U}| \leq |\mathcal{X}| \cdot |\mathcal{S}| + 2$.

Proof. The achievability is proved in Section III, while the converse proof is given in Section IV. ■

We note that expression (6) also appears in the capacity of a state-dependent single-user channel with non-causal state information at the transmitter [3]. Expression (7) also appears as the rate constraint for the strong receiver in a degraded state-dependent broadcast channel [6]. However, the key difference in this paper compared to known works is the addition of constraint (8) in our region that takes care of the codeword distortion tolerance requirement.

Remark 1. *The definition in* (1) *corresponds to a* physically degraded *broadcast channel, i.e. $(X, S) \to Y_1 \to Y_2$ is a Markov chain. On the other hand, the broadcast channel is said to be* stochastically degraded *if there exists $\tilde{Y}_1$ such that $\tilde{Y}_1|\{X = x, S = s\} \sim p_{Y_1|X,S}(\tilde{y}_1|x, s)$, i.e. $\tilde{Y}_1$ has the same conditional pmf as Y_1 given (X, S), and $(X, S) \to \tilde{Y}_1 \to Y_2$ is a Markov chain (see also [25, Chapter 5]).*

As far as the probability of error requirement (4) *is concerned, since the capacity region of the broadcast channel $p_{Y_1,Y_2|X,S}(y_1, y_2|x, s)$ depends upon the marginal distributions $p_{Y_1|X,S}(y_1|x, s)$ and $p_{Y_2|X,S}(y_2|x, s)$, no distinction needs to be made between physical and stochastic degradation. Furthermore, notice that the distortion requirement* (5) *depends only upon the marginal distribution of Y_2, i.e. $p_{Y_2|X,S}(y_2|x, s)$, and not on the marginal distribution of Y_1. This is because the input reconstruction $\hat{X}^n$ and the corresponding expected distortion $\mathbb{E}[d(X^n, \hat{X}^n)]$ depend only upon*

the degraded output Y_2^n, and not on the non-degraded output Y_1^n. Hence, even with respect to the distortion requirement, no distinction needs to be made between physical and stochastic degradation. Therefore, the proofs in the sequel go through even if the broadcast channel is stochastically degraded, and henceforth we shall simply refer to the broadcast channel as being degraded.

III. Achievability Proof of Theorem 1

The achievability is proven using a combination of Gelfand-Pinsker coding and superposition coding. Here, a U codebook is built first which shall be a compression codebook for the state sequence and is binned according to the weak user's message. Then for each u^n sequence, a conditional codebook shall be generated according to the strong user's message. We denote the set of jointly $\epsilon-$typical $n-$length sequences with respect to a joint distribution $p_{A,B}(a,b)$ by $\mathcal{T}_\epsilon^n(A,B)$, as in [25]. Fix the pmf $p(u|s)p(x|u,s)$ and function $\phi(u,y_2)$.

Codebook generation:
Randomly and independently generate $2^{n(R_2+R_2')}$ sequences $u^n(m_2,j_2)$, $m_2 \in [1:2^{nR_2}]$ and $j_2 \in [1:2^{nR_2'}]$ i.i.d. according to $\prod_{i=1}^n p_U(u_i)$ and divide them into 2^{nR_2} bins $\mathcal{B}_U(m_2)$ for $m_2 \in [1:2^{nR_2}]$. For each $u^n(m_2,j_2)$ sequence, randomly and conditionally independently generate 2^{nR_1} sequences $x^n(m_2,j_2,m_1)$ for $m_1 \in [1:2^{nR_1}]$ i.i.d. according to $\prod_{i=1}^n p_{X|U}(x_i|u_i)$.

Encoding:
Firstly, given $m_2 \in [1:2^{nR_2}]$ and s^n, find a sequence $u^n(m_2,j_2)$ with $j_2 \in \mathcal{B}_U(m_2)$ such that $(u^n(m_2,j_2),s^n) \in \mathcal{T}_{\epsilon'}^n(U,S)$. Declare error if no such index is found. Given $m_1 \in [1:2^{nR_1}]$, pick the sequence $x^n(m_2,j_2,m_1)$ in the conditional codebook corresponding to the chosen $u^n(m_2,j_2)$ sequence, which is then transmitted over the channel.

Decoding:
Let $\epsilon > \epsilon'$. The weak receiver declares that $\hat{m}_2$ is sent if it is the unique message such that $(u^n(\hat{m}_2,\hat{j}_2),y_2^n) \in \mathcal{T}_\epsilon^n(U,Y_2)$ for some $\hat{j}_2 \in \mathcal{B}_U(\hat{m}_2)$. Declare error if no such message is found or if more than one are found. At the strong user, we use simultaneous decoding [25, eq. (6.3)]. The strong user declares that $\hat{m}_1$ is sent if it is the unique message such that $(u^n(\hat{m}_2,\hat{j}_2),s^n,x^n(\hat{m}_2,\hat{j}_2,\hat{m}_1),y_1^n) \in \mathcal{T}_\epsilon^n(U,X,S,Y_1)$ for some $\hat{m}_2$ and $\hat{j}_2 \in \mathcal{B}_U(\hat{m}_2)$. Declare error if no such message is found or if more than one are found.

Analysis of probability of error:
Assume without loss of generality that the messages $M_1=1$ and $M_2=1$ were sent, and the index of the chosen U^n sequence for $M_1=1$, $M_2=1$ and S^n is J_2. The encoding error events are as follows:

$$\mathcal{E}_1 = \{(U^n(1,j_2),S^n) \notin \mathcal{T}_{\epsilon'}^n(U,S)\ \forall\ j_2 \in \mathcal{B}_U(1)\}. \quad (10)$$

The decoding error events at the receivers are as follows:

$$\begin{aligned}\mathcal{E}_2 &= \{(U^n(m_2,j_2),Y_2^n) \in \mathcal{T}_\epsilon^n(U,Y_2) \\ &\qquad \text{for some } m_2 \neq 1, j_2 \in \mathcal{B}_U(m_2)\},\\ \mathcal{E}_3 &= \{(U^n(1,J_2),X^n(1,J_2,m_1),S^n,Y_1^n) \\ &\qquad \in \mathcal{T}_\epsilon^n(U,X,S,Y_1) \text{ for some } m_1 \neq 1\}, \quad (11)\\ \mathcal{E}_4 &= \{(U^n(m_2,j_2),X^n(m_2,j_2,1),S^n,Y_1^n) \\ &\qquad \in \mathcal{T}_\epsilon^n(U,X,S,Y_1) \text{ for some } m_2 \neq 1, j_2 \in \mathcal{B}_U(m_2)\}, \quad (12)\\ \mathcal{E}_5 &= \{(U^n(m_2,j_2),X^n(m_2,j_2,m_1),S^n,Y_1^n) \\ &\qquad \in \mathcal{T}_\epsilon^n(U,X,S,Y_1) \text{ for some } m_2 \neq 1, \\ &\qquad\quad j_2 \in \mathcal{B}_U(m_2), m_1 \neq 1\}. \quad (13)\end{aligned}$$

The overall error event $\mathcal{E}$ is the union of the five error events listed above. By the union bound, we have:

$$\Pr(\mathcal{E}) \leq \sum_{i=1}^5 \Pr(\mathcal{E}_i). \quad (14)$$

By the covering lemma [25], the probability of encoding errors goes to zero as $n \to \infty$ as follows:

$$\Pr(\mathcal{E}_1) \to 0 \text{ as } n \to \infty \text{ if } R_2' \geq I(U;S) + \delta'(\epsilon'), \quad (15)$$

where $\delta'(\epsilon') \to 0$ as $\epsilon' \to 0$. By the packing lemma [25], the probability of decoding errors goes to zero as $n \to \infty$ as follows:

$$\Pr(\mathcal{E}_2) \to 0 \text{ as } n \to \infty \text{ if } R_2 + R_2' \leq I(U;Y_2) - \delta(\epsilon), \quad (16)$$

$$\Pr(\mathcal{E}_3) \to 0 \text{ as } n \to \infty \text{ if } R_1 \leq I(X;Y_1|U,S) - \delta(\epsilon), \quad (17)$$

$$\Pr(\mathcal{E}_4), \Pr(\mathcal{E}_5) \to 0 \text{ as } n \to \infty \text{ if}$$
$$R_1 + R_2 + R_2' \leq I(U,X;Y_1,S) - 2\delta(\epsilon), \quad (18)$$

where $\delta(\epsilon) \to 0$ as $\epsilon \to 0$. However, we note that the constraint (18) follows from (16) and (17), and thus rendered inoperative due to the degraded nature of the channel. This can be seen as follows, by adding together (16) and (17):

$$\begin{aligned}R_1 + R_2 + R_2' &\leq I(U;Y_2) + I(X;Y_1|U,S) - 2\delta(\epsilon)\\ &\overset{(a)}{\leq} I(U;Y_1) + I(X;Y_1|U,S) - 2\delta(\epsilon)\\ &\leq I(U;Y_1,S) + I(X;Y_1,S|U) - 2\delta(\epsilon)\\ &= I(U,X;Y_1,S) - 2\delta(\epsilon), \quad (19)\end{aligned}$$

where (a) follows from the degraded nature of the channel. Putting together equations (15) through (17), we arrive at the rate constraints:

$$\begin{aligned}R_1 &\leq I(X;Y_1|U,S) - \delta(\epsilon),\\ R_2 &\leq I(U;Y_2) - I(U;S) - \delta(\epsilon) - \delta'(\epsilon'). \quad (20)\end{aligned}$$

Distortion analysis:
The mentioned distortion can be obtained by making the input estimate at the weak receiver on a per-letter basis. Since the function $\phi(U,Y_2)$ satisfies the distortion constraint, it follows from the random codebook construction that as $n \to \infty$, we have

$$\begin{aligned}\mathbb{E}[d(X^n,\phi(U^n,Y_2^n))] &= \frac{1}{n}\sum_{i=1}^n \mathbb{E}[d(X_i,\phi(U_i,Y_{2i}))]\\ &\xrightarrow{n\to\infty} \mathbb{E}[d(X,\phi(U,Y_2))] \leq D. \quad (21)\end{aligned}$$

Thus the achievable region is proved.

IV. Converse Proof of Theorem 1

Given an achievable triple (R_1, R_2, D), we need to prove that there exists a joint distribution of the form $P_S P_{U|S} P_{X|U,S} P_{Y_1|X,S} P_{Y_2|Y_1}$ and a function $\phi(\cdot)$, such that the rate-distortion constraints in Theorem 1 hold. To bound the weak user's rate, consider the following chain of inequalities:

$$\begin{aligned}
nR_2 = H(M_2) &\overset{(a)}{\leq} I(M_2; Y_2^n) - I(M_2; S^n) + n\epsilon_n \\
&= \sum_{i=1}^{n} \{I(M_2; Y_{2i}|Y_2^{i-1}) - I(M_2; S_i|S_{i+1}^n)\} + n\epsilon_n \\
&\overset{(b)}{\leq} \sum_{i=1}^{n} \{I(M_2, Y_2^{i-1}; Y_{2i}) - I(M_2, S_{i+1}^n; S_i)\} + n\epsilon_n \\
&= \sum_{i=1}^{n} \{I(M_2, Y_2^{i-1}, S_{i+1}^n; Y_{2i}) - I(M_2, Y_2^{i-1}, S_{i+1}^n; S_i)\} \\
&\quad + \sum_{i=1}^{n} \{-I(S_{i+1}^n; Y_{2i}|M_2, Y_2^{i-1}) \\
&\qquad + I(Y_2^{i-1}; S_i|M_2, S_{i+1}^n)\} + n\epsilon_n \\
&\overset{(c)}{=} \sum_{i=1}^{n} \{I(M_2, Y_2^{i-1}, S_{i+1}^n; Y_{2i}) - I(M_2, Y_2^{i-1}, S_{i+1}^n; S_i)\} \\
&\quad + n\epsilon_n \\
&\overset{(d)}{=} \sum_{i=1}^{n} \{I(U_i; Y_{2i}) - I(U_i; S_i)\} + n\epsilon_n, \qquad (22)
\end{aligned}$$

where (a) follows from Fano's inequality with $\epsilon_n \to 0$ as $n \to \infty$ and since M_2 is independent of S^n, (b) follows since the state process is i.i.d., (c) follows from the Csiszar-sum identity, while (d) follows with an auxiliary identification of $U_i = (M_2, Y_2^{i-1}, S_{i+1}^n)$. Thus, we obtain

$$\begin{aligned}
R_2 &\leq \frac{1}{n} \sum_{i=1}^{n} \{I(U_i; Y_{2i}) - I(U_i; S_i)\} + \epsilon_n \\
&\overset{(a)}{=} I(U_Q; Y_{2,Q}|Q) - I(U_Q; S_Q|Q) + \epsilon_n \\
&\overset{(b)}{\leq} I(Q, U_Q; Y_{2,Q}) - I(Q, U_Q; S_Q) + \epsilon_n \\
&\overset{(c)}{=} I(U; Y_2) - I(U; S) + \epsilon_n, \qquad (23)
\end{aligned}$$

where (a) follows by introducing a time-sharing random variable Q uniform on $[1, 2, \ldots, n]$ that is independent of everything else, (b) follows since S_Q is independent of Q and (c) follows by defining $U = (Q, U_Q)$, $X = X_Q$, $S = S_Q$, $Y_1 = Y_{1,Q}$ and $Y_2 = Y_{2,Q}$. Taking limits as $n \to \infty$ which makes $\epsilon_n \to 0$ completes the bound. Next, to bound the strong user's rate, consider the following chain of inequalities:

$$\begin{aligned}
nR_1 = H(M_1) &\overset{(a)}{=} H(M_1|M_2, S^n) \\
&\overset{(b)}{\leq} I(M_1; Y_1^n|M_2, S^n) + n\epsilon_n \\
&\overset{(c)}{=} I(M_1, X^n; Y_1^n|M_2, S^n) + n\epsilon_n \\
&= \sum_{i=1}^{n} I(M_1, X^n; Y_{1i}|M_2, S^n, Y_1^{i-1}) + n\epsilon_n \\
&\overset{(d)}{=} \sum_{i=1}^{n} I(M_1, X^n; Y_{1i}|M_2, S^n, Y_1^{i-1}, Y_2^{i-1}) + n\epsilon_n \\
&= \sum_{i=1}^{n} \big\{H(Y_{1i}|M_2, S^n, Y_1^{i-1}, Y_2^{i-1}) \\
&\qquad - H(Y_{1i}|M_1, M_2, X^n, S^n, Y_1^{i-1}, Y_2^{i-1})\big\} + n\epsilon_n \\
&\overset{(e)}{\leq} \sum_{i=1}^{n} \big\{H(Y_{1i}|M_2, Y_2^{i-1}, S_{i+1}^n, S_i) \\
&\qquad - H(Y_{1i}|M_2, Y_2^{i-1}, S_{i+1}^n, X_i, S_i)\big\} + n\epsilon_n \\
&\overset{(f)}{=} \sum_{i=1}^{n} I(X_i; Y_{1i}|U_i, S_i) + n\epsilon_n, \qquad (24)
\end{aligned}$$

where (a) follows since M_1 is independent of (M_2, S^n), (b) follows from Fano's inequality, (c) follows since X^n is completely determined by (M_1, M_2, S^n), (d) follows from the degraded nature of the broadcast channel, (e) follows from the memorylessness of the channel, while (f) follows from the earlier identification of $U_i = (M_2, Y_2^{i-1}, S_{i+1}^n)$. Thus,

$$\begin{aligned}
R_1 &\leq \frac{1}{n} \sum_{i=1}^{n} I(X_i; Y_{1i}|U_i, S_i) + \epsilon_n \\
&= I(X_Q; Y_{1,Q}|Q, U_Q, S_Q) + \epsilon_n \\
&= I(X; Y_1|U, S) + \epsilon_n. \qquad (25)
\end{aligned}$$

We next verify the expected distortion. By the distortion constraint in assumption, we have

$$\begin{aligned}
D &\geq \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[d(X_i, \hat{X}_i)] \\
&\overset{(a)}{=} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[d(X_i, \phi_i(U_i, Y_{2i}))] \\
&= \mathbb{E}_Q[\mathbb{E}[d(X_Q, \phi_Q(U_Q, Y_{2,Q}))]|Q] \\
&= \mathbb{E}[d(X_Q, \phi_Q(U_Q, Y_{2,Q}))] \\
&\overset{(b)}{=} \mathbb{E}[d(X_Q, \phi(Q, U_Q, Y_{2,Q}))] \\
&\overset{(c)}{=} \mathbb{E}[d(X, \phi(U, Y_2))], \qquad (26)
\end{aligned}$$

where (a) follows by taking $\phi_i(U_i, Y_{2i}) = \hat{X}_i$, (b) follows by defining $\phi : (Q, U_Q) \mapsto \phi_Q(U_Q)$ and (c) follows since $(U_Q, Q) = U, X_Q = X$, and $Y_{2,Q} = Y_2$. Finally, the bound on the auxiliary random variable cardinality $|\mathcal{U}|$ follows using the support lemma [25, Appendix C], as detailed in Appendix A. The converse proof is complete.

V. Binary Additive Broadcast Channel with Reversible Inputs

Suppose the input, state and reconstruction alphabets are restricted to be binary, i.e., $\mathcal{X} = \mathcal{S} = \hat{\mathcal{X}} = \{0, 1\}$. For the second receiver's link, we consider an additive binary channel given by

$$Y_2 = X \oplus S \oplus N_2, \qquad (27)$$

where $\oplus$ denotes modulo-two addition, while the state S is a Bernoulli random variable independent of everything else, i.e., $S \sim \text{Ber}(p), p \in [0, 1/2]$. Here, N_2 is another Bernoulli random variable specified by $N_2 \sim \text{Ber}(q_2), q_2 \in [0, 1/2]$. On the other hand, the first receiver's link is another binary

additive channel specified by

$$Y_1 = X \oplus S \oplus N_1, \tag{28}$$

where N_1 is independent of everything else and is Bernoulli distributed, i.e., $N_1 \sim \mathrm{Ber}(q_1), q_1 \in [0, 1/2]$ with $q_1 < q_2$. Notice that we can express

$$Y_2 = Y_1 \oplus \tilde{N}, \tag{29}$$

where $\tilde{N} \sim \mathrm{Ber}(\tilde{q})$ and $q_2 = q_1 * \tilde{q}$. At the weak receiver, the reconstruction distortion is defined in terms of the Hamming measure

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\Big[X_i \oplus \hat{X}_i(Y_2^n)\Big] \leq D. \tag{30}$$

We have the following theorem.

Theorem 2. *For the input reconstruction problem over the given binary additive broadcast channel, the rate-distortion trade-off region is the closure of the set that contains all triples (R_1, R_2, D) that satisfy*

$$R_2 \leq 1 - H_2(\alpha * q_2) + H_2(D * q_2) - H_2(p), \tag{31}$$

$$R_1 \leq H_2(\alpha * q_1) - H_2(q_1), \tag{32}$$

for some $\alpha \in [0, 1/2]$, with $H_2(a) = -a\log_2 a - (1-a)\log_2(1-a)$ being the binary entropy function and $a * b = a(1-b) + b(1-a)$ denoting binary convolution.

Proof. The proof is based on an evaluation of the bounds in Theorem 1 for the given channel model. This is given in the following two subsections. The achievability is based on splitting the information transmission into two parts, one of which is for reliable communication while the other is intended for lossy input compression. ■

A. Proof of achievability for the binary broadcast channel

Choose independent random variables $(U', \tilde{S})$ such that $U' \sim \mathrm{Ber}(1/2)$ and $\tilde{S} \sim \mathrm{Ber}((p-(D*q_2))/(1-2(D*q_2)))$. The joint distribution of $(S, \tilde{S})$ is such that $\tilde{S}$ is an input to a binary symmetric channel with crossover probability $D * q_2$, and S is the corresponding output distributed as $S \sim \mathrm{Ber}(p)$. In other words, we have

$$S = \tilde{S} \oplus N' \oplus N_2, \tag{33}$$

where $N' \sim \mathrm{Ber}(D)$ is independent of everything else.

We now construct the auxiliary random variable U, the channel input X and the input codeword estimate at the weak receiver as

$$U = (U', U'') \triangleq (U', U' \oplus \tilde{S}), \tag{34}$$

$$X = U' \oplus S \oplus N_2', \tag{35}$$

$$\hat{X} = U' \oplus \tilde{S} \oplus N_2' \oplus N_2, \tag{36}$$

where $N_2' \sim \mathrm{Ber}(\alpha)$ is independent of (U', S, N'). We now evaluate the mutual information terms in Theorem 1. The achievable rate for the weak user is

$$\begin{aligned} R_2 &= I(U;Y_2) - I(U;S) \\ &= H(X \oplus S \oplus N_2) - H(X \oplus S \oplus N_2 | U', U' \oplus \tilde{S}) \\ &\quad - H(S) + H(S|U', U' \oplus \tilde{S}) \\ &= H(U' \oplus N_2 \oplus N_2') - H(U' \oplus N_2 \oplus N_2' | U', U' \oplus \tilde{S}) \\ &\quad - H_2(p) + H(\tilde{S} \oplus N' \oplus N_2 | U', U' \oplus \tilde{S}) \\ &= H(U' \oplus N_2 \oplus N_2') - H(N_2 \oplus N_2') \\ &\quad - H_2(p) + H(N' \oplus N_2) \\ &= 1 - H_2(\alpha * q_2) - H_2(p) + H_2(D * q_2). \end{aligned} \tag{37}$$

The achievable rate for the strong user is

$$\begin{aligned} R_1 &= I(X;Y_1|U,S) \\ &= H(X \oplus S \oplus N_1 | U', U' \oplus \tilde{S}, S) \\ &\quad - H(X \oplus S \oplus N_1 | U', U' \oplus \tilde{S}, X, S) \\ &= H(U' \oplus N_2' \oplus N_1 | U', S, N') - H(N_1) \\ &= H(N_2' \oplus N_1) - H(N_1) \\ &= H_2(\alpha * q_1) - H_2(q_1). \end{aligned} \tag{38}$$

The distortion constraint can be verified as follows:

$$\mathbb{E}[d(X, \hat{X})] = \mathbb{E}[X \oplus \hat{X}] = \mathbb{E}[S \oplus \tilde{S} \oplus N_2] = \mathbb{E}[N'] = D. \tag{39}$$

This completes the proof of achievability.

B. Proof of converse for the binary broadcast channel

The converse proof involves invoking the single-letter expressions in Theorem 1. For the weak user's rate, we have

$$\begin{aligned} R_2 &\leq I(U;Y_2) - I(U;S) \\ &= H(Y_2) - H(Y_2|U) - H(S) + H(S|U) \\ &= H(Y_2) - H(Y_2|U,S) - H(S) + H(S|U,Y_2) \\ &\overset{(a)}{=} H(Y_2) - H(Y_2|U,S) - H(S) + H(S|U,Y_2,\hat{X}) \\ &\overset{(b)}{\leq} 1 - H(Y_2|U,S) - H_2(p) + H(S \oplus Y_2|U,Y_2,\hat{X}) \\ &= 1 - H(Y_2|U,S) - H_2(p) + H(X \oplus N_2|U,Y_2,\hat{X}) \\ &\leq 1 - H(Y_2|U,S) - H_2(p) + H(X \oplus \hat{X} \oplus N_2) \\ &\overset{(c)}{\leq} 1 - H(Y_2|U,S) - H_2(p) + H_2(D * q_2), \end{aligned} \tag{40}$$

where (a) follows since the estimate $\hat{X}$ is determined by (U, Y_2), (b) follows since Y_2 is binary, while (c) follows since $\Pr(X \neq \hat{X}) \leq D$ and by invoking the monotonicity of the binary entropy function $H_2(\cdot)$ on $[0, p]$. For the strong user's rate, we have

$$\begin{aligned} R_1 &\leq I(X;Y_1|U,S) \\ &= H(Y_1|U,S) - H(Y_1|U,S,X) \\ &= H(Y_1|U,S) - H(N_1) = H(Y_1|U,S) - H_2(q_1). \end{aligned} \tag{41}$$

Since the following holds

$$1 \geq H(Y_2|U,S) \geq H(Y_2|U,S,X) = H_2(q_2), \tag{42}$$

there exists $\alpha \in [0, 1/2]$ such that

$$H(Y_2|U,S) = H_2(\alpha * q_2). \tag{43}$$

Substituting (43) in (40), we obtain

$$R_2 \leq 1 - H_2(\alpha * q_2) - H_2(p) + H_2(D * q_2). \tag{44}$$

Now let $0 \leq H^{-1}(\nu) \leq 1/2$ be the inverse of the binary entropy function. By the degraded nature of the channel and applying Mrs. Gerber's lemma [25]

$$H_2(\alpha * q_2) = H(Y_2|U,S)$$

$$= H(Y_1 \oplus \tilde{N}|U,S) \geq H\big(H^{-1}(H(Y_1|U,S)) * \tilde{q}\big), \tag{45}$$

where $\tilde{N}$ and $\tilde{q}$ are as in (29). This implies that

$$H(Y_1|U,S) \leq H_2(\alpha * q_1). \tag{46}$$

Substituting (46) in (41), we obtain

$$R_1 \leq H_2(\alpha * q_1) - H_2(q_1). \tag{47}$$

This completes the proof of converse.

C. Numerical Illustration

We now plot the trade=off region in Theorem 2 for an example system with the parameters $p = 0.4, q_1 = 2/9$, and $q_2 = 4/9$ in Fig. 2. In particular, we plot the trade-off between the communication rates R_1 and R_2 for three fixed values of the input reconstruction distortion D. It is seen that demanding smaller distortion values (which amounts to a better quality input reconstruction) results in smaller achievable communication rates, and vice versa. For instance, the set of achievable rate pairs R_1 and R_2 for $D = 0.1$ (which corresponds to a better quality input reconstruction) are much smaller compared to the corresponding achievable rates for $D = 0.3$ (which corresponds to a poorer quality input reconstruction).

We have also plotted the trade-off between R_2 and D for the same parameters in Fig. 3, wherein it is observed that a higher D in fact leads to a higher R_2 (and vice-versa). This is unlike a traditional rate-distortion trade-off, where a higher description rate leads to a lower distortion. In contrast, the demand for a higher rate in our setting leads to a poorer quality input codeword estimate at the second receiver, i.e. a higher D (and vice-versa). This is because of the intrinsic tension between the dual requirements at the second receiver.

We also compare the region in Theorem 2 to the corresponding setting without any input reconstruction constraints, as depicted in Fig. 4. Naturally, it is observed that the introduction of an input reconstruction constraint restricts the region of achievable rates. In other words, larger rates are seen to be achievable in the absence of the input reconstruction constraint.

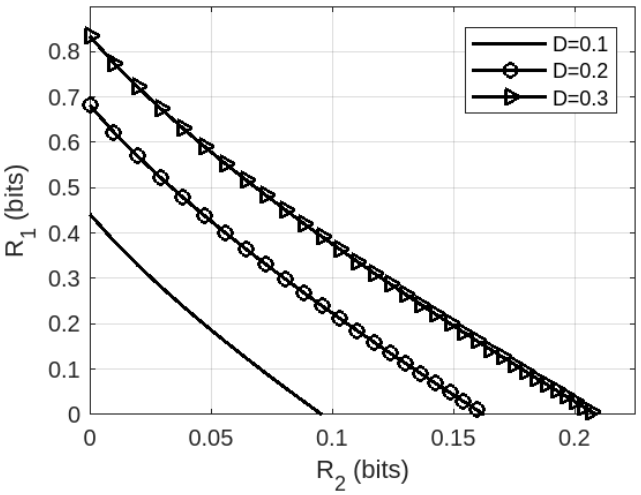

Fig. 2. Plot of the achievable communication rate trade-offs (R_1, R_2) for different distortion values D incurred in input reconstruction, for a broadcast channel with parameters specified as $p = 0.4, q_1 = 2/9$, and $q_2 = 4/9$.

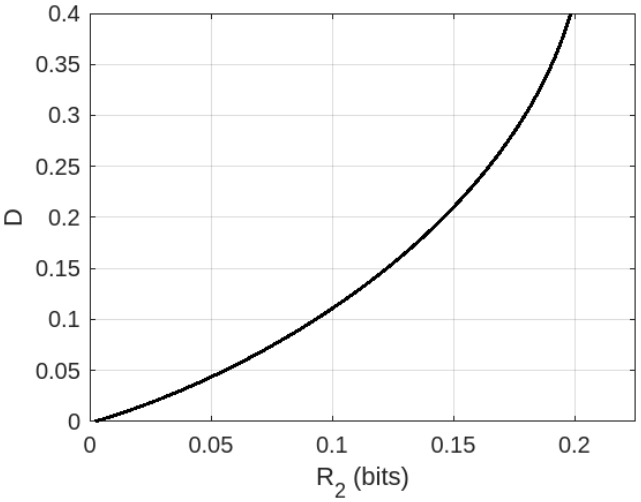

Fig. 3. Plot of the achievable distortion D versus the rate R_2, for the same broadcast channel with parameters $p = 0.4, q_1 = 2/9$, and $q_2 = 4/9$.

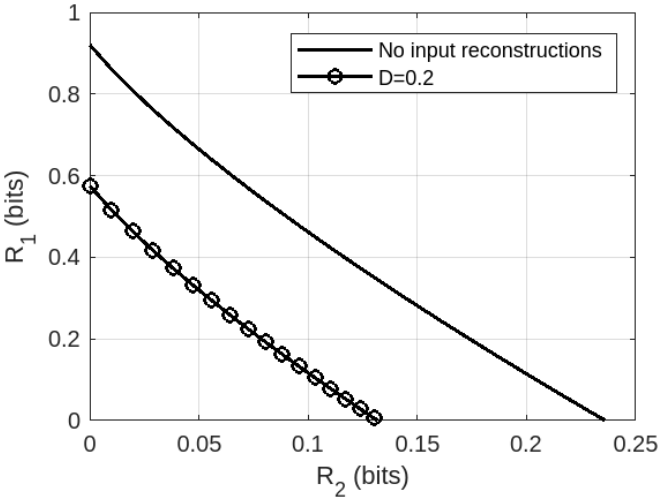

Fig. 4. Comparison of achievable rate regions for a given D versus the case of no input reconstructions.

VI. Conclusion

A state-dependent degraded broadcast channel with input reconstruction requirements was investigated, and the optimal trade-off between the message communication rates and the input reconstruction distortion was characterized. The setting where the strong receiver is also uninformed of the state process appears to be an interesting and challenging problem for further investigations.

VII. Acknowledgements

The author would like to thank the Editor and two anonymous reviewers for their insightful comments, which has greatly helped the content and presentation of this paper.

APPENDIX A
PROOF OF CARDINALITY BOUND ON $|\mathcal{U}|$

In this section, we shall prove the cardinality bound on the auxiliary random variable U in Theorem 1, thereby showing that our characterization is computable. We use standard cardinality bounding techniques in the information theory literature, see [26]. We shall prove that $|\mathcal{U}| \leq |\mathcal{S}|\cdot|\mathcal{X}|+2$. Let $U \sim F(u)$ and $(S,X)|U=u \sim p(s,x|u)$, where U takes values in $\mathcal{U}$. Given (U,S,X), consider the following $|\mathcal{S}|\cdot|\mathcal{X}|+2$ continuous real-valued functions of $\pi = p(s,x|u)$

$$g_j(p(s,x|u) = \pi(j), j = 1,\cdots,|\mathcal{S}|\cdot|\mathcal{X}|-1, \tag{48}$$

$$g_j(p(s,x|u) = H(Y_1|S,U=u), j = |\mathcal{S}|\cdot|\mathcal{X}|, \tag{49}$$

$$g_j(p(s,x|u) = H(S|U=u) - H(Y_2|U=u), j = |\mathcal{S}|\cdot|\mathcal{X}|+1, \tag{50}$$

$$g_j(p(s,x|u) = \mathbb{E}[d(X,\hat{X})|U=u], j = |\mathcal{S}|\cdot|\mathcal{X}|+2. \tag{51}$$

By the Fenchel-Eggleston strengthening of Caratheodory's theorem [26], there exists a random variable U' with $|\mathcal{U}'| \leq |\mathcal{S}|\cdot|\mathcal{X}|+2$ such that $p(s,x)$, $\mathbb{E}[d(X,\hat{X})|U]$, $H(Y_1|U,S)$ and $H(S|U) - H(Y_2|U)$ are preserved:

$$\begin{aligned}\int_{\mathcal{U}} p(s,x|u)dF(u) &= p(s,x)\\ &= \sum_{u'\in\mathcal{U}'} p(s,x|u')p(u'),\end{aligned} \tag{52}$$

$$\begin{aligned}H(S|U) - H(Y_2|U)&\\ &= \int_{\mathcal{U}} (H(S|U=u) - H(Y_2|U=u))dF(u)\\ &= \sum_{u'\in\mathcal{U}'} (H(S|U'=u') - H(Y_2|U'=u'))p(u')\\ &= H(S|U') - H(Y_2|U'),\end{aligned} \tag{53}$$

$$\begin{aligned}\mathbb{E}[d(X,\hat{X})|U]&\\ &= \int_{\mathcal{U}} \mathbb{E}[d(X,\hat{X})|U=u]dF(u)\\ &= \sum_{u'\in\mathcal{U}'} \mathbb{E}[d(X,\hat{X})|U'=u']p(u')\\ &= \mathbb{E}[d(X,\hat{X})|U'],\end{aligned} \tag{54}$$

$$\begin{aligned}H(Y_1|S,U)&\\ &= \int_{\mathcal{U}} H(Y_1|S,U=u)dF(u)\\ &= \sum_{u'\in\mathcal{U}'} H(Y_1|S,U'=u')p(u')\\ &= H(Y_1|S,U').\end{aligned} \tag{55}$$

We can thus write the following:

$$\begin{aligned}I(U;Y_2) - I(U;S)&\\ &= H(Y_2) - H(S) - H(Y_2|U) + H(S|U)\\ &= H(Y_2) - H(S) - H(Y_2|U') + H(S|U')\\ &= I(U';Y_2) - I(U';S).\end{aligned} \tag{56}$$

Moreover, it follows that:

$$\begin{aligned}I(X;Y_1|U,S)&\\ &= H(Y_1|U,S) - H(Y_1|X,S)\\ &= H(Y_1|U',S) - H(Y_1|X,S)\\ &= I(X;Y_1|U',S).\end{aligned} \tag{57}$$

Thus it suffices to consider auxiliary random variables U such that:

$$|\mathcal{U}| \leq |\mathcal{S}|\cdot|\mathcal{X}|+2. \tag{58}$$

This completes the proof of the cardinality of the auxiliary random variable U in Theorem 1.

REFERENCES

[1] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Research Development*, vol. 2, no. 4, pp. 289–293, 1958.

[2] A. V. Kuznetsov and B. S. Tsybakov, "Coding in a memory with defective cells," *Prob. peredachi informatsii*, vol. 10, no. 2, pp. 52–60, 1974.

[3] S. Gelfand and M. Pinsker, "Coding for channels with random parameters," *Probl. Contr. Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[4] C. Heegard and A. Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inform. Theory*, vol. 29, no. 5, pp. 731–739, 1983.

[5] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2007–2019, 1999.

[6] Y. Steinberg, "Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *IEEE Trans. Inform. Theory*, vol. 51, no. 8, pp. 2867–2877, 2005.

[7] G. Keshet, Y. Steinberg, and N. Merhav, "Channel coding in the presence of side information," *Foundations Trends Commun. Inform. Theory*, vol. 4, no. 6, pp. 445–586, 2007.

[8] A. Sutivong, M. Chiang, T. M. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1486–1495, 2005.

[9] M. H. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, 1983.

[10] Y.-H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 1850–1859, 2008.

[11] C. Choudhuri, Y.-H. Kim, and U. Mitra, "Causal state communication," *IEEE Trans. Inform. Theory*, vol. 59, no. 6, pp. 3709–3719, 2013.

[12] V. Ramachandran, S. R. B. Pillai, and V. M. Prabhakaran, "Joint state estimation and communication over a state-dependent Gaussian multiple access channel," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6743–6752, 2019.

[13] N. Merhav and S. Shamai, "Information rates subject to state masking," *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 2254–2261, 2007.

[14] O. O. Koyluoglu, R. Soundararajan, and S. Vishwanath, "State amplification subject to masking constraints," *IEEE Trans. Inform. Theory*, vol. 62, no. 11, pp. 6233–6250, 2016.

[15] Y.-K. Chia, R. Soundararajan, and T. Weissman, "Estimation with a helper who knows the interference," *IEEE Trans. Inform. Theory*, vol. 59, no. 11, pp. 7097–7117, 2013.

[16] C. Tian, B. Bandemer, and S. Shamai, "Gaussian state amplification with noisy observations," *IEEE Trans. Inform. Theory*, vol. 61, no. 9, pp. 4587–4597, 2015.

[17] Y. Steinberg, "Coding and common reconstruction," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 4995–5010, 2009.

[18] K. Kittichokechai, T. J. Oechtering, and M. Skoglund, "Coding with action-dependent side information and additional reconstruction requirements," *IEEE Trans. Inform. Theory*, vol. 61, no. 11, pp. 6355–6367, 2015.

[19] P. Grover, A. B. Wagner, and A. Sahai, "Information embedding and the triple role of control," *IEEE Trans. Inform. Theory*, vol. 61, no. 4, pp. 1539–1549, 2015.

[20] C. Choudhuri and U. Mitra, "On Witsenhausen's counterexample: The asymptotic vector case," in *Proc. IEEE ITW*, 2012.

[21] O. Sumszyk and Y. Steinberg, "Information embedding with reversible stegotext," in *Proc. IEEE ISIT*, 2009.

[22] B. Bandemer and A. El Gamal, "Communication with disturbance constraints," *IEEE Trans. Inform. Theory*, vol. 60, no. 8, pp. 4488–4502, 2014.

[23] W. Zhang, S. Vedantam, and U. Mitra, "Joint transmission and state estimation: A constrained channel coding approach," *IEEE Trans. Inform. Theory*, vol. 57, no. 10, pp. 7084–7095, 2011.

[24] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Inform. Theory*, vol. 49, no. 3, pp. 563–593, 2003.

[25] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge University Press, 2011.
[26] M. Salehi, "Cardinality bounds on auxiliary variables in multiple-user theory via the method of ahlswede and körner," Dept. Stat., Stanford Univ., Stanford, CA, Tech. Rep, vol. 33, 1978.

Viswanathan Ramachandran was born in Kerala, India. He received the Ph.D. degree in Electrical Engineering from the Indian Institute of Technology Bombay in 2020. He worked as a Visiting Researcher at the Tata Institute of Fundamental Research during 2020. During 2021-2022, he was a Postdoctoral Fellow with the Department of Electrical Engineering, Technical University of Eindhoven, the Netherlands. Since 2023, he is a Postdoctoral Fellow with the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Sweden. His research interests are in information theory, particularly in multi-terminal settings, with applications to wireless and optical channels. He was a recipient of the Naik and Rastogi Award for Excellence in Ph.D. research from the Indian Institute of Technology Bombay in 2021, and also the Marie-Curie individual postdoctoral fellowship 2022 from the European Commission.