On State-Dependent Broadcast Channels, with Side-Information at the Transmitter

K. G. Nagananda†, Chandra R. Murthy‡ and Shalinee Kishore†

†Dept. of ECE  ‡Dept. of ECE
Lehigh University  Indian Institute of Science
Bethlehem, PA 18105, USA  Bangalore 560012, India
{kgn209, skishore}@lehigh.edu  cmurthy@ece.iisc.ernet.in

Abstract

In this paper, we derive information-theoretic performance limits for three classes of two-user state-dependent discrete memoryless broadcast channels, with noncausal side-information at the encoder. The first class of channels comprises a sender broadcasting two independent messages to two non-cooperating receivers; for channels of the second class, each receiver is given the message it need not decode; and the third class comprises channels where the sender is constrained to keep each message confidential from the unintended receiver. To derive inner bounds, we employ an extension of Marton’s achievability scheme for the classical two-user broadcast channel, results from the second moment method, an extension of the technique proposed by Kramer and Shamai for broadcast channels with receiver side-information, and stochastic encoders to satisfy confidentiality requirements. Outer bounds are derived by following the procedure used to prove the converse theorem for Gel’fand-Pinsker’s channels with random parameters; and confidentiality constraints are utilized for deriving outer bounds for channels belonging to the third class. For channels of the second class, the bounds are tight, thereby yielding the capacity region.

Index Terms

State-dependent broadcast channels, side-information, rate regions, outer bounds.

The authors would like to thank Parthajit Mohapatra for helpful comments and discussion. The authors would also like to thank Tobias Oechtering, for pointing out an error in an earlier draft of this paper, in the proof of Theorem 3.2.
I. INTRODUCTION

An information-theoretic study of broadcast channels (BC) was initiated first by Cover in [1]. In the classical setting, the BC comprises a sender who wishes to transmit $k$ independent messages to $k$ noncooperative receivers. The largest known inner bound on the capacity region when $k = 2$ was derived by Marton [2]; capacity outer bounds for BC have appeared in [3] - [7]. Several variants of this classical setting have also received considerable attention. One of the most prominent variants is the state-dependent BC with side-information, where the probability distribution characterizing the channel depends on a state process, and with the channel state made available as side information at the transmitter, or at the receiver, or at both ends. Capacity inner bounds for the two-user BC with noncausal side-information at the transmitter were derived in [8], where Marton’s achievability scheme was extended to state dependent channels. In [9], inner and outer bounds were derived for the degraded BC with noncausal side-information at the transmitter; the capacity region was derived when side-information was provided to the encoder in a causal manner. The capacity region for BC with receiver side-information was derived in [10], where a genie provides each receiver with the message it need not decode. A slightly different model was considered in [11], where a sender broadcasts blocks of data to multiple receivers, with each receiver having prior side-information of some subset of the other blocks.

Yet another issue in wireless communications, owing to the broadcast nature of the wireless medium, is related to information security. That is, the broadcast nature of wireless networks facilitates malicious or unauthorized access to confidential data, denial of service attacks, corruption of sensitive data, etc. An information-theoretic approach to address problems related to security has gained rapid momentum, and is commonly referred to as information-theoretic confidentiality or wireless physical-layer security [12]. An information-theoretic approach to secure broadcasting was inspired by the pioneering work of Csiszár and Körner [13], who derived capacity bounds for the two-user BC, when the sender transmits a private message to receiver 1 and a common message to both receivers, while keeping the private message confidential from receiver 2. In [14], capacity bounds were derived for BC where a sender broadcasts two independent messages to two receivers, while keeping each message confidential from the unintended receiver. Capacity results and bounds for Gaussian BC with confidential messages were reported in [15] - [17]. The reader is referred to [18] for a comprehensive review of physical-layer security in BC.
A. Our contribution

Fig. 1. State-dependent broadcast channels with side-information at the transmitter: (a) Class I; (b) Class II; and (c) Class III.

In this paper, our main goal is to analyze the impact of side information and confidentiality constraints on the information theoretic performance limits of BC. To this end, we derive capacity bounds on the following three classes of two-user discrete memoryless BC, with noncausal side-information, for e.g., fading in the wireless medium, interference caused by neighboring nodes in the network, etc. at the encoder:

1) Class I: A sender broadcasts two independent messages to two non-cooperating receivers (see Fig. 1(a)). An inner bound for this class of channels was derived by Steinberg and Shamai in [8], by extending Marton’s achievability scheme [2] to include noncausal side-information at the encoder. However, in this paper, we extend Marton’s achievability scheme and use results from the second moment method [19] to derive an inner bound. Our proof is simpler and generalizes well to derive an inner bound for channels of Class III (described below). An outer bound is derived employing the procedure used to prove the converse theorem for Gel’fand-Pinsker’s (GP) channels with random parameters [20]. The bounds are shown to be tight for individual rate constraints, but can be improved upon for the sum-rate. An example for Class I channels is a base station transmitting to two mobile receivers, in the presence of a priori known interference from a transmitter located in the vicinity of the base station.

2) Class II: A sender broadcasts two independent messages to two receivers, with each receiver having a priori knowledge of the message it need not decode (see Fig. 1(b)). An example of this scenario is full-duplex communications between two nodes, aided
by a relay. The relay node broadcasts the messages to the terminals, with each terminal knowing its own message. Class II channels are also addressed in [21], where an inner bound matching our result was derived; however, there was only an outline for deriving an outer bound. In this paper, an inner bound is derived by extending the method proposed by Kramer and Shamai in [10], to include transmitter side-information for BC where each receiver has knowledge of the other’s message; our proof is much simpler than the one presented in [21]. Furthermore, our outer bounds are derived using arguments from the proof of converse for GP’s channel and are shown to be tight, thereby yielding the capacity region for this class of channels.

3) Class III: A sender broadcasts two independent messages to two receivers, such that each message is kept confidential from the unintended receiver (see Fig. [c]). The achievability theorem is proved by employing the technique used to derive an inner bound for Class I channels, in conjunction with a stochastic encoder to satisfy confidentiality constraints. The technique to derive outer bounds hinges on the confidentiality requirements. We also derive a genie-aided outer bound, where a genie gives a receiver the message it need not decode, while the other receiver computes the equivocation treating this message as side-information. We also suggest a tighter outer bound for the sum rate of this class of channels. As an example for this class of channels, we can extend the example considered for Class I channels, with the additional constraint of keeping each message ignorant from the unintended receiver.

For all the three classes of channels, Csiszár’s sum identity [22] plays a central role in establishing the capacity outer bounds. The results in this paper demonstrate that, owing to rate-penalties for dealing with side-information and satisfying confidentiality constraints, the rate region for channels of Class III is smaller than that for Class I, which is further smaller compared to the classical two-user BC. Note that, the comparisons are presented primarily to illustrate the role of side information and secrecy constraints on the achievable rates. However, since each model is characterized by its own probability distribution, these comparisons should be made with caution. The initial results of this work have been submitted to a conference [23], [24].

The remainder of the paper is organized as follows. In Section II, we introduce the notation used and provide a mathematical model for the discrete memoryless version of the channels
considered in this paper. In Section III we summarize the main results of this paper, by describing inner and outer bounds for all the channel models. In Section IV we prove the achievability theorem and provide related discussion. The proof of the converse theorem is provided in Section V. Finally, we conclude the paper in Section VI. The analysis of the probability of error at the encoder for channels of Class I and Class III is relegated to the appendix.

II. SYSTEM MODEL & NOTATION

The channels belonging to Class I, Class II and Class III are denoted $C_1, C_2$ and $C_3$, respectively. Calligraphic letters are used to denote finite sets, with a probability function defined on them. $N$ is the number of channel uses, and $n = 1, \ldots, N$ denotes the channel index. Uppercase letters denote random variables (RV), while boldface uppercase letters denote a sequence of RVs. The following notation for a sequence of RVs is useful: $Y_1^N \triangleq (Y_{1,1}, \ldots, Y_{1,N})$; $Y_{1,n}^{n-1} \triangleq (Y_{1,1}, \ldots, Y_{1,n-1})$; and $Y_{1,n+1}^N \triangleq (Y_{1,n+1}, \ldots, Y_{1,N})$. Lowercase letters are used to denote particular realizations of RVs, and boldface lowercase letters denote $N$-length vectors. The sender is denoted $S$ and the receivers are denoted $D_t$, where $t = 1, 2$ denotes the receiver index. Discrete random variables (RV) $X \in \mathcal{X}$ and $Y_t \in \mathcal{Y}_t$ denote the channel input and outputs, respectively. The encoder of $S$ is supplied with side-information $W \in \mathcal{W}^N$, in a noncausal manner. The channel is assumed to be memoryless and is characterized by the conditional distribution $p(Y_1, Y_2|X, W) = \prod_{n=1}^{N} p(Y_{n+1}, Y_{n+2} | X_n, W_n)$. For sake of brevity, in the remainder of this paper, we use $p(x)$ to denote $p(X = x)$. Unless otherwise stated, $p(x) = \prod_{n=1}^{N} p(x_n)$.

To transmit its messages, $S$ generates two RVs $M_t \in \mathcal{M}_t$, where $\mathcal{M}_t = \{1, \ldots, 2^{N R_t}\}$ denotes a set of message indices. Without loss of generality, $2^{N R_t}$ is assumed to be an integer, with $R_t$ being the transmission rate intended to $D_t$. $M_t$ denotes the message $S$ intends to transmit to $D_t$, and is assumed to be independently generated and uniformly distributed over the finite set $\mathcal{M}_t$. Integer $m_t \in \mathcal{M}_t$ is a particular realization of $M_t$ and denotes the message-index.

Given the conditional distribution characterizing the channel, a $((2^{NR_1}, 2^{NR_2}), N, P_e^{(N)})$ code for the channels $C_1$ and $C_2$ comprises $N$ encoding functions $f$, such that $X = f(m_1, m_2, W)$; for the channel $C_3$, it comprises a stochastic encoder, which is defined by the matrix of conditional probabilities $\phi(X|m_1, m_2, W)$, such that $\sum_x \phi(X|m_1, m_2, W) = 1$. Here, $\phi(X|m_1, m_2, W)$ denotes the probability that a pair of message-indices $(m_1, m_2)$ is encoded as $X \in \mathcal{X}^N$ to be
transmitted by $S$, in the presence of noncausal side-information $W$. For all channel models, there are two decoders $g_t : \mathcal{Y}_t^N \to \mathcal{M}_t$.

The average probability of decoding error for the code, averaged over all codes, is $P_e^{(N)} = \max\{P_{e,1}^{(N)}, P_{e,2}^{(N)}\}$, where $P_{e,t}^{(N)} = \sum_{m_1,m_2} \sum_{W \in \mathcal{W}^N} \frac{1}{2^{N(r_1^t + r_2^t)}} \Pr [g_t(Y_t^N) \neq m_t | m_1, m_2, W \text{ sent}]$. A rate pair $(R_1, R_2)$ is said to be achievable for the channel $C_c$; $c = 1, 2, 3$, if there exists a sequence of $((2^{NR_1}, 2^{NR_2}), N, P_e^{(N)})$ codes $\forall \epsilon > 0$ and sufficiently small, such that $P_e^{(N)} \leq \epsilon$ as $N \to \infty$. Furthermore, for the channel $C_3$, the following constraints [25] on the conditional entropy must be satisfied for $(R_1, R_2)$ to be considered achievable:

\begin{align}
NR_1 - H(M_1|Y_2) &\leq N\epsilon, \\
NR_2 - H(M_2|Y_1) &\leq N\epsilon.
\end{align}

The capacity region is defined as the closure of the set of all achievable rate pairs $(R_1, R_2)$.

### III. Statement of Results

In this section, we state first the achievability theorem and then present outer bounds for all the channel models described in the previous section. Let $C_c$ denote the capacity region of the channel $C_c$. We use the following auxiliary RVs defined on finite sets: $U \in \mathcal{U}$, $V_1 \in \mathcal{V}_1$ and $V_2 \in \mathcal{V}_2$. For the channel $C_1$, $V_1$ and $V_2$ are constrained to satisfy the Markov chain $(V_1, V_2) \to X \to (Y_1, Y_2)$; for the channel $C_2$, we have $U \to X \to (Y_1, Y_2)$; and for the channel $C_3$, $U$, $V_1$ and $V_2$ satisfy the Markov chain $U \to (V_1, V_2) \to X \to (Y_1, Y_2)$. For the channel $C_1$, we consider the set $\mathcal{P}_1$ of all joint probability distributions $p_1(.)$ that can be factored as $p(w)p(v_1, v_2|w)p(x|w, v_1, v_2)p(y_1, y_2|x)$. For the channel $C_2$, we consider the set $\mathcal{P}_2$ of all joint probability distributions $p_2(.)$ of the form $p(w)p(u|w)p(x|w, u)p(y_1, y_2|x)$. For the channel $C_3$, we consider the set $\mathcal{P}_3$ of all joint probability distributions $p_3(.)$ that can be written as $p(w)p(u)p(v_1, v_2|w, u)p(x|w, v_1, v_2)p(y_1, y_2|x)$.

#### A. Achievable rate regions

1) For a given $p_1(.) \in \mathcal{P}_1$, a lower bound on the capacity region for $C_1$ is described by the set $\mathcal{R}_{1, in}(p_1)$, which is defined as the union over all distributions $p_1(.)$ of the convex hull
of the set of all rate pairs \((R_1, R_2)\) that simultaneously satisfy (3) - (5).

\[
R_1 \leq I(V_1; Y_1) - I(V_1; W), \tag{3}
\]

\[
R_2 \leq I(V_2; Y_2) - I(V_2; W), \tag{4}
\]

\[
R_1 + R_2 \leq I(V_1; Y_1) + I(V_2; Y_2) - I(V_1; V_2) - I(V_1; V_2; W). \tag{5}
\]

2) For a given \(p_2(\cdot) \in \mathcal{P}_2\), a lower bound on the capacity region for \(C_2\) is described by the set \(\mathcal{R}_{2,\text{in}}(p_2)\), which is defined as the union over all distributions \(p_2(\cdot)\) of the convex-hull of the set of all rate pairs \((R_1, R_2)\) that simultaneously satisfy (6) - (7).

\[
R_1 \leq I(U; Y_1) - I(U; W), \tag{6}
\]

\[
R_2 \leq I(U; Y_2) - I(U; W). \tag{7}
\]

3) For a given \(p_3(\cdot) \in \mathcal{P}_3\), an inner bound on the capacity region for \(C_3\) is described by the set \(\mathcal{R}_{3,\text{in}}(p_3)\), which is defined as the union over all distributions \(p_3(\cdot)\) of the convex-hull of the set of all rate pairs \((R_1, R_2)\) that simultaneously satisfy (8) - (10).

\[
R_1 \leq I(V_1; Y_1|U) - \max[I(V_1; Y_2|U, V_2), I(V_1; W|U)], \tag{8}
\]

\[
R_2 \leq I(V_2; Y_2|U) - \max[I(V_2; Y_1|U, V_1), I(V_2; W|U)], \tag{9}
\]

\[
R_1 + R_2 \leq I(V_1; Y_1|U) + I(V_2; Y_2|U) - I(V_1; Y_2|U, V_2) - I(V_2; Y_1|U, V_1) - I(V_1; V_2|U) - I(V_1; V_2; W|U). \tag{10}
\]

**Theorem 3.1:** Let \(\mathcal{R}_{c,\text{in}} = \bigcup_{p_c(\cdot) \in \mathcal{P}_c} \mathcal{R}_{c,\text{in}}(p_c)\). The region \(\mathcal{R}_{c,\text{in}}\) is an achievable rate region for \(C_c\), i.e., \(\mathcal{R}_{c,\text{in}} \subseteq C_c\), where \(c = 1, 2, 3\).

**B. Outer bounds**

1) For a given \(p_1(\cdot) \in \mathcal{P}_1\), an outer bound for \(C_1\) is described by the set \(\mathcal{R}_{1,\text{out}}(p_1)\), which is defined as the union of all rate pairs \((R_1, R_2)\) that simultaneously satisfy (11) - (13).

\[
R_1 \leq I(V_1; Y_1) - I(V_1; W), \tag{11}
\]

\[
R_2 \leq I(V_2; Y_2) - I(V_2; W), \tag{12}
\]

\[
R_1 + R_2 \leq I(V_1; Y_1) + I(V_2; Y_2) - I(V_1; W) - I(V_2; W). \tag{13}
\]
2) For a given $p_2(.) \in P_2$, an outer bound for $C_2$ is described by the set $R_{2,\text{out}}(p_2)$, which is defined as the union of all rate pairs $(R_1, R_2)$ that simultaneously satisfy (14) - (15).

\[
R_1 \leq I(U; Y_1) - I(U; W), \\
R_2 \leq I(U; Y_2) - I(U; W).
\]

(14) \hspace{1cm} (15)

3) For a given $p_3(.) \in P_3$, an outer bound for $C_3$ is described by the set $R_{3,\text{out}}(p_3)$, which is defined as the union of all rate pairs $(R_1, R_2)$ that simultaneously satisfy (16) - (18).

\[
R_1 \leq \min[I_1, I_1^*], \\
R_2 \leq \min[I_2, I_2^*], \\
R_1 + R_2 \leq \min[I_{12}, I_{12}^*],
\]

(16) \hspace{1cm} (17) \hspace{1cm} (18)

where $I_1, \ldots, I_{12}^*$ are given by (19) - (24), respectively.

\[
I_1 \triangleq I(V_1; Y_1|U) - I(V_1; Y_2|U) + H(W|U, V_1), \\
I_2 \triangleq I(V_2; Y_2|U) - I(V_2; Y_1|U) + H(W|U, V_2), \\
I_{12} \triangleq I(V_1; Y_1|U) + I(V_2; Y_2|U) - I(V_1; Y_2|U) - I(V_2; Y_1|U)
\]

\[+ H(W|U, V_1) + H(W|U, V_2). \]

(19) \hspace{1cm} (20) \hspace{1cm} (21)

\[
I_1^* \triangleq I(V_1; Y_1|U, V_2) - I(V_1; Y_2|U, V_2) + H(W|U, V_1, V_2), \\
I_2^* \triangleq I(V_2; Y_2|U, V_1) - I(V_2; Y_1|U, V_1) + H(W|U, V_1, V_2), \\
I_{12}^* \triangleq I(V_1; Y_1|U, V_2) + I(V_2; Y_2|U, V_1) - I(V_1; Y_2|U, V_2)
\]

\[-I(V_2; Y_1|U, V_1) + 2H(W|U, V_1, V_2). \]

(22) \hspace{1cm} (23) \hspace{1cm} (24)

*Theorem 3.2:* Let $R_{c,\text{out}} = \bigcup_{p_c(.) \in P_c} R_{c,\text{out}}(p_c)$. The region $R_{c,\text{out}}$ is an outer bound for $C_c$, i.e., $C_c \subseteq R_{c,\text{out}}$, where $c = 1, 2, 3$.

Some remarks:

1) For the channels $C_1$ and $C_3$, a bound on the sum rate $R_1 + R_2$ is obtained by using the fact that if $R_1 \leq I_A$ and $R_2 \leq I_B$, then $\forall \alpha, \beta \in \mathbb{Z}^+$, $\alpha R_1 + \beta R_2 \leq \alpha I_A + \beta I_B$, where $I_A$ and $I_B$ are mutual information expressions.

2) For the channel $C_1$, the bounds on the individual rates are tight, whereas the sum-rate bound can be improved upon. For the channel $C_2$, the bounds are tight, thereby yielding...
the capacity region for the channel; it is also a genie-aided outer bound for the channel $C_1$.

3) For the channel $C_3$, the expressions (22) - (24) are obtained by letting a hypothetical genie give $D_1$ message $M_2$, while $D_2$ computes the equivocation using $M_2$ as side-information.

IV. PROOF OF THEOREM 3.1

In this section, we prove Theorem 3.1 for the three channel models presented in the previous section. For any $\epsilon > 0$, we denote by $A^{(N)}_\epsilon (P_X)$ an $\epsilon$-typical set comprising sequences picked from the distribution $p(x)$. For all the channel models, the encoder at $S$ is given an $\epsilon$-typical sequence $W \in A^{(N)}_\epsilon (P_W)$ in a noncausal manner.

A. Proof of Theorem 3.1 for the channel $C_1$

For the channel $C_1$, generate $2^{N[R_t+R'_t]}$ independent typical sequences $V_t(i_t, j_t) \in A^{(N)}_\epsilon (P_{V_t}); t = 1, 2$. Here, $i_t \in \{1, \ldots, 2^{NR_t}\}; j_t \in \{1, \ldots, 2^{N'R_t}\}$. Uniformly distribute $2^{N[R_t+R'_t]}$ sequences into $2^{NR_t}$ bins, so that each bin, indexed by $i_t$, comprises $2^{N'R_t}$ sequences. To send the message pair $(m_1 = i_1, m_2 = i_2)$, the encoder at $S$ looks for a pair $(j_1, j_2)$ that satisfies the following joint typicality condition: $E_S \triangleq \{(W, V_1(i_1, j_1), V_2(i_2, j_2)) \in A^{(N)}_\epsilon (P_{W, V_1, V_2})\}$. An error is declared at the encoder of $S$, if it is not possible to find the $(j_1, j_2)$-pair to satisfy the condition $E_S$. The encoder error analysis can be found in Appendix A. The channel input sequence is $X \in A^{(N)}_\epsilon (P_{X|W, V_1, V_2})$.

At the destination $D_t$, the decoder looks for $(\hat{i}_t, \hat{j}_t)$ that satisfies the following joint typicality condition: $E_{D_t} \triangleq \{(V_t(\hat{i}_t, \hat{j}_t), Y_t) \in A^{(N)}_\epsilon (P_{V_t, Y_t})\}$. An error is declared at decoder of $D_t$, if it not possible to find a unique integer $\hat{i}_t$ to satisfy the condition $E_{D_t}$. From the union of events bound, the probability of decoder error at $D_t$ can be upper bounded as follows: $P^{(N)}_{\epsilon, D_t} \leq \Pr(E_{D_t}^c, E_S) + \sum_{\hat{i}_t \neq i_t} \Pr(E_{D_t}|E_S)$. From the asymptotic equipartition property (AEP) [26], $\forall \epsilon > 0$ and sufficiently small; and for large $N$, $\Pr(E_{D_t}^c | E_S) \leq \epsilon$. Further, for $\hat{i}_t \neq i_t$, $\Pr(E_{D_t}|E_S) \leq 2^{-N[I(V_t; Y_t) - \epsilon]}$. Therefore, we have $P^{(N)}_{\epsilon, D_t} \leq \epsilon + 2^{N[R_t+R'_t]}2^{-N[I(V_t; Y_t) - \epsilon]}$, leading us to conclude that, for any $\epsilon_0 > 0$ and sufficiently small; and for large $N$, $P^{(N)}_{\epsilon, D_t} \leq \epsilon_0$ if

$$R_t + R'_t < I(V_t; Y_t).$$  

(25)
For the channel $C_1$, the rate inequalities (25) and the bounds on the binning rates (52) - (54) (see Appendix A) are combined to obtain an achievable rate region given by (3) - (5). This completes the proof of Theorem 3.1 for the channel $C_1$. We now employ the results of GP’s channel with random parameters [20] to obtain a pictorial representation of the rate region (see Fig. 2(a)). When $R_2 = 0$, the channel resembles a single-user channel $(S, D_1)$ with side-information and $S$ can transmit at the maximum achievable $R_1$ given by (3), denoted by point $H$. At the point $H$, the maximum achievable $R_2$ is given by the point $E_1 \equiv I(V_2; Y_2) - I(V_1; V_2) - I(W; V_2)$; this is obtained by treating the channel $(S, D_2)$ as a single-user channel with side-information. Therefore, the rectangle $OHGE_1$ is achievable. By exchanging $R_1$ and $R_2$ and following similar arguments the points $E_1$, given by (4), and $F_1 \equiv I(V_1; Y_1) - I(V_1; V_2 | U) - I(W; V_1)$ are achievable. Hence, the rectangle $OEFF_1$ is also achievable. Since the points $F$ and $G$ are shown to be achievable, any point which lies on the line $FG$ can also be achieved by deriving a bound on the binning rates (see (52) - (54), Appendix A). This leads to a sum rate bound given by (5). Finally, owing to convexity of the rate region, any point in the interior of the line $FG$ is also achievable. Therefore, an achievable rate region for $C_1$ is described by the pentagon $OEFGH$.

In the absence of side-information, i.e., $\mathcal{W} = \{\phi\}$, the channel reduces to the classical two-user BC whose rate region is described by the convex-hull of the set of all rate pairs $(R_1, R_2)$.
that satisfy the following inequalities (see the pentagon OIJKL in Fig. 2(b)):

\[
\begin{align*}
R_1 &\leq I(V_1; Y_1), \\
R_2 &\leq I(V_2; Y_2), \\
R_1 + R_2 &\leq I(V_1; Y_1) + I(V_2; Y_2) - I(V_1; V_2).
\end{align*}
\]

Due to the rate-penalty for dealing with the side information that is unavailable at the receivers, the achievable region for the channel \( C_1 \) is strictly smaller than that for the classical BC.

**B. Proof of Theorem 3.1 for the channel \( C_2 \)**

For the channel \( C_2 \), we consider the following two cases.

1) When \( R_1 \leq R_2 \): Generate \( 2^{N(R_2 + R^*)} \) typical sequences \( \mathbf{U}(i, j) \in A_i^{(N)}(P_U); i \in \{1, \ldots, 2^{NR_2}\}; j \in \{1, \ldots, 2^{NR^*}\} \). Uniformly distribute these sequences into \( 2^{NR_2} \) bins, so that each bin comprises \( 2^{NR^*} \) sequences. The bins are indexed by \( i \). Define now the following mappings:

\[
 m_t \in \{1, \ldots, 2^{NR_t}\} \implies \text{Int}(m_t) \in \{0, \ldots, 2^{NR_t} - 1\}; t = 1, 2,
\]

where \( \text{Int}(\alpha) \) denotes an integer to represent \( \alpha \). To transmit the message pair \( (m_1, m_2) \), compute \( (\text{Int}(m_1) + \text{Int}(m_2) \mod 2^{NR_2}) \). By construction, the bin index

\[
i \triangleq \text{Int}^{-1}(\text{Int}(m_1) + \text{Int}(m_2) \mod 2^{NR_2}).
\]

Given the sequence \( \mathbf{W} \), the encoder looks for an integer \( j \) to satisfy the following joint typicality condition: \( (\mathbf{U}(i, j), \mathbf{W}) \in A_i^{(N)}(P_{W,U}) \). Finally, \( \mathbf{X} \triangleq \mathbf{f}(\mathbf{U}(i, j), \mathbf{W}) \) is transmitted in \( N \) channel uses.

At receiver \( D_1 \), given \( m_2 \), the decoder looks for the pair \( (i, \hat{m}_1, j) \) such that the following joint typicality condition is satisfied: \( E_{D_1} \triangleq \{(\mathbf{U}(\text{Int}^{-1}(\text{Int}(\hat{m}_1) + \text{Int}(m_2) \mod 2^{NR_2})) , j), \mathbf{Y}_1 \} \in A_i^{(N)}(P_{U,Y_1}) \}. \) From AEP, it can be shown that \( \Pr(E_{D_1}^c) \leq \delta_1; \forall \delta_1 > 0 \) and sufficiently small; and for large \( N \), if \( R_1 + R^* \leq I(U; Y_1) \). Similarly, it can be shown that \( \Pr(E_{D_2}^c) \leq \delta_2; \forall \delta_2 > 0 \) and sufficiently small; and for large \( N \), if \( R_2 + R^* \leq I(U; Y_2) \). Additionally, by following a procedure similar to the one presented in Appendix \( A \), we bound the binning rate as follows:

\[
 R^* > I(U; W).
\]

Therefore, \( m_1 \) (resp. \( m_2 \)) can be reliably decoded at \( D_1 \) (resp. \( D_2 \)) if

\[
\begin{align*}
R_1 &\leq I(U; Y_1) - I(U; W), \\
R_2 &\leq I(U; Y_2) - I(U; W).
\end{align*}
\]
2) When $R_2 \leq R_1$: By symmetry, we get the same rate bounds as in (29) and (30).

This completes the proof of Theorem 3.1 for the channel $C$. Note that, each bound in (29) - (30) is the capacity of GP’s single-user channel with noncausal side-information. In the absence of side-information, i.e., $\mathcal{W} = \{\phi\}$, we get $R_t \leq I(U; Y_t) = I(X; Y_t)$, which represents the capacity region of BC when each receiver is given the message it need not decode 10. Furthermore, we show in Section V-B that, the bounds given by (29) - (30) are tight, thereby establishing the capacity region for the channel $C$.

C. Proof of Theorem 3.1 for the channel $C_3$

For the channel $C_3$, generate a typical sequence $\mathbf{U} \in A_{t}^{(N)}(P_U)$, known to all nodes in the network. Generate $2^{N[R_t + R' + R^*_t]}$ independent typical sequences $\mathbf{V}_t(i_t, j_t, k_t) \in A_{t}^{(N)}(P_{V_t})$; $i_t \in \{1, \ldots, 2^{NR_t}\}$; $j_t \in \{1, \ldots, 2^{NR'_t}\}$; $k_t \in \{1, \ldots, 2^{NR^*_t}\}$. Uniformly distribute $2^{N[R_t + R' + R^*_t]}$ sequences into $2^{NR_t}$ bins, so that each bin, indexed by $i_t$, comprises $2^{N[R'_t + R^*_t]}$ sequences. Uniformly distribute $2^{N[R'_t + R^*_t]}$ sequences into $2^{NR'_t}$ sub-bins indexed by $(i_t, j_t)$, so that each sub-bin comprises $2^{NR^*_t}$ sequences.

To send the message pair $(m_1, m_2)$, $S$ employs a stochastic encoder. In the bin indexed by $i_t$, randomly pick a sub-bin indexed $(i_t, j_t)$. The encoder then looks for a pair $(k_1, k_2)$ that satisfies the following joint typicality condition: $(\mathcal{W}, \mathbf{V}_1(i_1, j_1, k_1), \mathbf{V}_2(i_2, j_2, k_2)) \in A_{t}^{(N)}(P_{W,V_1,V_2|U})$. The channel input sequence $\mathbf{X} \in A_{t}^{(N)}(P_{X|W,V_1,V_2})$ is transmitted in $N$ uses of the channel.

At the destination $D_t$, given $\mathbf{U}$, the decoder picks $k_t$ that satisfies the following joint typicality condition: $E_{D_t} \triangleq \{(\mathbf{V}_1(i_t, j_t, k_t), \mathbf{Y}_t) \in A_{t}^{(N)}(P_{V_t,Y_t|U})\}$. An error is declared at the decoder of $D_t$ if it not possible to find an integer $\hat{i}_t$ satisfying $E_{D_t}$. From union of events bound, the probability of decoder error at $D_t$ can be upper bounded as follows: $P_{e,D_t}^{(N)} \leq \Pr(E_{D_t}|E_S) + \sum_{i_t \neq \hat{i}_t} \sum_{j_t,k_t} \Pr(E_{D_t}|E_S)$. From AEP [26], $\forall \epsilon > 0$ and sufficiently small; and for large $N$, $\Pr(E_{D_t}|E_S) \leq \epsilon$ and for $\hat{i}_t \neq i_t$, we have $\Pr(E_{D_t}|E_S) \leq 2^{-N[I(V_t; Y_t|U) - \epsilon]}$. Therefore, $P_{e,D_t}^{(N)} \leq \epsilon + 2^{N[R_t + R'_t + R^*_t]} 2^{-N[I(V_t; Y_t|U) - \epsilon]}$. For any $\epsilon_0 > 0$ and sufficiently small; and for large $N$, $P_{e,D_t}^{(N)} \leq \epsilon_0$ if

$$R_t + R'_t + R^*_t < I(V_t; Y_t|U).$$

(31)

The equivocation at the decoder of $D_2$ is calculated by first considering the following lower bound: $H(M_1|Y_2^N) \geq H(M_1|Y_2^N, U^N, V_2^N)$. Following the procedure in [14, Section V-B] and
using the fact that $M_1 \to (U^N, V_1^N, V_2^N) \to Y_2^N$ forms a Markov chain, we get

$$H(M_1 | Y_2^N) \geq H(V_1^N | U^N) - I(V_1^N; V_2^N | U^N) - H(V_1^N | M_1, U^N, V_2^N, Y_2^N) - I(V_1^N; Y_2^N | U^N, V_2^N). \tag{32}$$

Now, $\forall \epsilon_l > 0; l = 4, \ldots, 10$ and sufficiently small; and for large $N$, the terms in (32) can be written as

$$H(V_1^N | U^N) \overset{(a)}{=} N[R_1 + R'_1 + R''_1]; I(V_1^N; V_2^N | U^N) \overset{(b)}{=} N I(V_1; V_2 | U) + N \epsilon_4;$$

$$H(V_1^N | M_1, U^N, V_2^N, Y_2^N) \overset{(c)}{\leq} N \epsilon_5; I(V_1^N; Y_2^N | U^N, V_2^N) \overset{(d)}{=} N I(V_1; Y_2 | U, V_2) + N \epsilon_6,$$

where (a) follows from the codebook construction; (b) and (d) follow from standard techniques (for e.g., see [14, Lemma 3]); and (c) is proved in [14, Lemma 2]). A similar procedure is followed to calculate the equivocation at the decoder at $D_1$. Finally, the security constraints (1) and (2) are satisfied by letting

$$R'_1 = I(V_1; Y_2 | U, V_2) - \epsilon_7; R''_1 = I(V_1; V_2 | U) - \epsilon_8; \tag{33}$$

$$R'_2 = I(V_2; Y_1 | W, U, V_1) - \epsilon_9; R''_2 = I(V_1; V_2 | W, U) - \epsilon_{10}. \tag{34}$$

For the channel $C_3$, rate inequalities (31), constraints (33) - (34) and bounds on the binning rates (55) - (57) (see Appendix A) are combined to obtain the rate region described by (8) - (10). This completes the proof of Theorem 3.1 for the channel $C_3$. Using a combination of results from GP’s channel with random parameters [20] and wiretap channels with side-information [27], we obtain a pictorial representation of the rate region for the channel $C_3$ as shown in Fig. 3(a). The arguments used to obtain this schematic are similar to those used for the channel $C_1$; therefore, we briefly explain the construction of Fig. 3(a) below.

The point $A_1$ corresponds to the maximum achievable $R_1$ (when $R_2 = 0$) and is given by (8). Exchanging $R_1$ and $R_2$ we get the point $C_1$ given by (9). The points $B_1 \equiv I(V_2; Y_2 | U) - I(V_2; Y_1 | U, V_1) - \max[I(V_1; V_2 | U), I(W; V_2 | U)]$ and $D_1 \equiv I(V_1; Y_1 | U) - I(V_1; Y_2 | U, V_2) - \max[I(V_1; V_2 | U), I(W; V_1 | U)]$ are achievable by treating channels $(S, D_2)$ and $(S, D_1)$, respectively, as wiretap channels with side-information. The line $E_1F_1$ corresponds to the sum rate bound given by (10). Finally, owing to convexity of the rate region, any point in the interior of the line $E_1F_1$ is also achievable. Therefore, an achievable rate region for $C_3$ is described by the pentagon $OA_1F_1E_1C_1$. 

DRAFT November 23, 2011
If the confidentiality constraints (1) - (2) are relaxed, the channel $C_3$ reduces to the channel $C_1$, whose rate region is described by (3) - (5) (see the pentagon OEGFH in Fig. 3(b)). Further, in the absence of side-information, i.e., $\mathcal{W} = \{\phi\}$, the channel reduces to the classical two-user BC whose rate region is described by (26) - (28) (see the pentagon OIJKL in Fig. 3(b)). Lastly, if the encoder satisfies confidentiality constraints in the absence of side-information, the channel $C_3$ reduces to BC with two independent and confidential messages whose rate region was first characterized by Liu et. al [14]. It is described by the convex-hull of the set of all rate pairs $(R_1, R_2)$ that satisfy the following inequalities:

\begin{align}
R_1 &\leq I(V_1; Y_1|U) - I(V_1; Y_2|U) - I(V_1; V_2|U), \\
R_2 &\leq I(V_2; Y_2|U) - I(V_2; Y_1|U) - I(V_1; V_2|U).
\end{align}

The rate region for BC with side-information (3) - (5) is smaller than that of the classical BC (26) - (28), due to the rate-penalty for side-information. And, the rate-penalties for side-information and confidentiality constraints make the achievable region of channel $C_3$ smaller than that for $C_1$. This provides the necessary intuition for the size (though, they are not to-scale in Fig. 3(b)) of the pentagon OIJKL, which subsumes OEGFH which further subsumes OABCD.

Fig. 3. (a) Pictorial representation of the rate region for Class III channels; (b) the rate regions of the classical BC, Class I and Class III channels.
V. PROOF OF THEOREM 3.2

A. Proof of Theorem 3.2 for the channel $C_1$

For the channel $C_1$, $\forall \epsilon > 0$ and sufficiently small; and for large $N$, $R_1$ can be bounded as follows:

$$NR_1 = H(M_1) = I(M_1; Y_1^N) + H(M_1|Y_1^N)$$

$$(a) \leq I(M_1; Y_1^N) + N\epsilon \overset{(b)}{=} \sum_{n=1}^{N}[H(Y_{1,n}|Y_1^{n-1}) - H(Y_{1,n}|Y_1^{n-1}, M_1)] + N\epsilon$$

$$\leq \sum_{n=1}^{N}[H(Y_{1,n}) - H(Y_{1,n}|Y_1^{n-1}, M_1)] + N\epsilon = \sum_{n=1}^{N} I(M_1, Y_1^{n-1}; Y_{1,n}) + N\epsilon$$

$$= \sum_{n=1}^{N} [I(M_1, Y_1^{n-1}, W_{n+1}^{N}; Y_{1,n}) - I(W_{n+1}^{N}; Y_{1,n}|M_1, Y_1^{n-1})] + N\epsilon$$

$$\leq \sum_{n=1}^{N} [I(M_1, Y_1^{n-1}, W_{n+1}^{N}; Y_{1,n}) - I(Y_1^{n-1}; W_{n+1}|M_1, W_{n+1}^{N})] + N\epsilon$$

$$(c) \leq \sum_{n=1}^{N} [I(M_1, Y_1^{n-1}, W_{n+1}^{N}; Y_{1,n}) - I(Y_1^{n-1}; W_{n+1}|M_1, W_{n+1}^{N})] + N\epsilon$$

where (a) follows from Fano’s inequality [26], (b) follows from the chain rule, (c) follows from the fact that conditioning reduces entropy, (d) follows from Csiszár’s sum identity and (e) is due to the fact that $(M_i, W_{n+1}^{N})$ is independent of $W_n$. We let $V_{1,n} = (M_1, W_{n+1}^{N}, Y_1^{n-1})$ and note that this choice satisfies the Markov chain requirement $V_1 \rightarrow X \rightarrow (Y_1, Y_2)$, specified in Section III for the channel $C_1$. Thus, we get

$$NR_1 \leq \sum_{n=1}^{N} I(V_{1,n}; Y_{1,n}) - I(V_{1,n}; W_n) + N\epsilon. \quad (37)$$

Proceeding in a similar manner and letting $V_{2,n} = (M_2, W_{n+1}^{N}, Y_2^{n-1})$, we get

$$NR_2 \leq \sum_{n=1}^{N} I(V_{2,n}; Y_{2,n}) - I(V_{2,n}; W_n) + N\epsilon. \quad (38)$$

The following is a bound on the sum-rate:

$$N(R_1 + R_2) \leq \sum_{n=1}^{N} [I(V_{1,n}; Y_{1,n}) + I(V_{2,n}; Y_{2,n}) - I(V_{1,n}; W_n) - I(V_{2,n}; W_n)] + 2N\epsilon. \quad (39)$$
B. Proof of Theorem 3.2 for the channel $C_2$

For the channel $C_2$, $\forall \epsilon > 0$ and sufficiently small; and for large $N$, $R_1$ can be bounded as follows:

$$NR_1 = H(M_1) = I(M_1; Y_1^N) + H(M_1|Y_1)$$

\[\overset{(a)}{\leq} I(M_1; Y_1^N) + N\epsilon \overset{(b)}{\leq} I(M_1; Y_1^N, M_2) + N\epsilon = I(M_1; Y_1^N|M_2) + N\epsilon\]

\[\overset{(c)}{=} \sum_{n=1}^{N} [H(Y_{1,n}|Y_1^{n-1}, M_2) - H(Y_{1,n}|Y_1^{n-1}, M_1, M_2)] + N\epsilon\]

\[\overset{(d)}{\leq} \sum_{n=1}^{N} [H(Y_{1,n}) - H(Y_{1,n}|Y_1^{n-1}, M_1, M_2)] + N\epsilon\]

\[= \sum_{n=1}^{N} I(M_1, M_2; Y_1^{n-1}, Y_{1,n}) + N\epsilon \leq \sum_{n=1}^{N} I(M_1, M_2, Y_1^{n-1}, Y_2^{n-1}, Y_{1,n}) + N\epsilon\]

\[= \sum_{n=1}^{N} [I(M_1, M_2, Y_1^{n-1}, Y_2^{n-1}, W_{n+1}^N; Y_{1,n}) - I(W_{n+1}^N, Y_{1,n}|M_1, M_2, Y_1^{n-1}, Y_2^{n-1})] + N\epsilon\]

\[\overset{(e)}{=} \sum_{n=1}^{N} [I(M_1, M_2, Y_1^{n-1}, Y_2^{n-1}, W_{n+1}^N; Y_{1,n}) - I(Y_1^{n-1}, M_2, Y_2^{n-1}, W_{n+1}^N|Y_{1,n})] + N\epsilon\]

\[\overset{(f)}{=} \sum_{n=1}^{N} [I(M_1, M_2, Y_1^{n-1}, Y_2^{n-1}, W_{n+1}^N; Y_{1,n}) - I(M_1, M_2, Y_1^{n-1}, Y_2^{n-1}, W_{n+1}^N|W_n)] + N\epsilon,\]

where (a) follows from Fano’s inequality, (b) follows from the data-processing inequality, (c) follows from chain rule, (d) follows from the fact that conditioning reduces entropy, (e) follows from Csiszár’s sum identity and (f) is due to the fact that $(M_1, M_2, Y_1^{n-1}, Y_2^{n-1}, W_{n+1}^N)$ is independent of $W_n$. We let $U_n \triangleq (M_1, M_2, Y_1^{n-1}, Y_2^{n-1}, W_{n+1}^N)$ and note that this choice satisfies the Markov chain requirement $U \rightarrow X \rightarrow (Y_1, Y_2)$ specified in Section III for the channel $C_2$, we get

$$NR_1 \leq \sum_{n=1}^{N} I(U_n; Y_{1,n}) - I(U_n; W_n) + N\epsilon. \quad (40)$$

By symmetry, we get the following bound on $R_2$:

$$NR_2 \leq \sum_{n=1}^{N} I(U_n; Y_{2,n}) - I(U_n; W_n) + N\epsilon. \quad (41)$$
C. Proof of Theorem 3.2 for the channel $C_3$

For the channel $C_3$, $\forall \epsilon > 0$ and sufficiently small; and for large $N$, $R_1$ can be bounded as follows:

\[
NR_1 = H(M_1) = I(M_1; Y_1^N) + H(M_1| Y_1^N) \leq I(M_1; Y_1^N) + N\epsilon \leq I(M_1; Y_1^N) - I(M_1; Y_2^N) + 2N\epsilon = \sum_{n=1}^{N} [I(M_1; Y_{1,n}| Y_{1,1,n+1}^N) - I(M_1; Y_{1,n}^N, Y_{2,n}| Y_{2,1,n+1}^N)] + 2N\epsilon \leq \sum_{n=1}^{N} [I(M_1, W_{n}; Y_{1,n}| Y_{1,1,n+1}^N, Y_{2,n}^N) - I(M_1; Y_{1,n}^N, Y_{2,n}| Y_{2,1,n+1}^N) + 2N\epsilon] \leq \sum_{n=1}^{N} [I(M_1; Y_{1,n}| Y_{1,1,n+1}^N, Y_{2,n}^N) + I(W_n; Y_{1,n}| M_1, Y_{1,1,n+1}^N, Y_{2,n}^N) - I(M_1; Y_{2,n}| Y_{1,1,n+1}^N, Y_{2,n}^N)] + 2N\epsilon = \sum_{n=1}^{N} [I(M_1; Y_{1,n}| Y_{1,1,n+1}^N, Y_{2,n}^N) + H(W_n|M_1, Y_{1,1,n+1}^N, Y_{2,n}^N) - H(W_n|M_1, Y_{1,1,n+1}^N, Y_{2,n}^N) - I(M_1; Y_{2,n}| Y_{1,1,n+1}^N, Y_{2,n}^N)] + 2N\epsilon \leq \sum_{n=1}^{N} [I(M_1; Y_{1,n}| Y_{1,1,n+1}^N, Y_{2,n}^N) + H(W_n|M_1, Y_{1,1,n+1}^N, Y_{2,n}^N) - I(M_1; Y_{2,n}| Y_{1,1,n+1}^N, Y_{2,n}^N)] + 2N\epsilon,
\]

where (a) is from Fano’s inequality, (b) is from confidentiality constraints, (c) and (d) follow from Csiszár’s sum identity and (e) is the chain rule for mutual information. Letting $U_n \triangleq (Y_{1,1,n+1}^N, Y_{2,n}^N)$; and $V_{1,1} = \cdots = V_{1,N} \triangleq M_1$, where $U$ and $V$ satisfy the Markov chain $U \rightarrow V \rightarrow X$ specified in Section III for the channel $C_3$, we get

\[
NR_1 \leq \sum_{n=1}^{N} [I(V_{1,n}; Y_{1,n}| U_n) + H(W_n| U_n, V_{1,n}) - I(V_{1,n}; Y_{2,n}| U_n)] + 2N\epsilon. \tag{42}
\]
Proceeding in a similar fashion and letting \( V_{2,1} = \cdots = V_{2,N} \triangleq M_2 \),
\[
NR_2 \leq \sum_{n=1}^{N} [I(V_{2,n}; Y_{2,n}|U_n) + H(W_n|U_n, V_{2,n}) - I(V_{2,n}; Y_{1,n}|U_n)] + 2N\epsilon. \tag{43}
\]

The sum-rate bound is given by
\[
N(R_1 + R_2) \leq \sum_{n=1}^{N} [I(V_{1,n}; Y_{1,n}|U_n) + I(V_{2,n}; Y_{2,n}|U_n) - I(V_{1,n}; Y_{2,n}|U_n) - I(V_{2,n}; Y_{1,n}|U_n)]
\]
\[
+ H(W_n|U_n, V_{1,n}) + H(W_n|U_n, V_{2,n}) + 4N\epsilon. \tag{44}
\]

For the channel \( C_3 \), we also derive a genie-aided outer bound by letting a hypothetical genie
give \( D_1 \) message \( M_2 \), while \( D_2 \) computes the equivocation using \( M_2 \) as side-information. \( \forall \epsilon > 0 \)
and sufficiently small; and for large \( N \), \( R_1 \) can be upper bounded as follows:
\[
NR_1 = H(M_1) \leq H(M_1|Y_2^N) + N\epsilon \leq H(M_1, M_2|Y_2^N) + N\epsilon
\]
\[
= H(M_1|Y_2^N, M_2) + H(M_2|Y_2^N) + N\epsilon \leq H(M_1|Y_2^N, M_2) + N\epsilon
\]
\[
\leq H(M_1|Y_2^N, M_2) - H(M_1|Y_1^N) + N\epsilon \leq H(M_1|Y_2^N, M_2) - H(M_1|Y_1^N, M_2) + N\epsilon
\]
\[
\leq I(M_1; Y_1^N|M_2) - I(M_1; Y_2^N|M_2) + 2N\epsilon
\]
\[
= \sum_{n=1}^{N} [I(M_1; Y_{1,n}|Y_{1,n+1}^N, M_2) - I(M_1; Y_{2,n}|Y_{2,n-1}^N, M_2)] + 2N\epsilon
\]
\[
\leq \sum_{n=1}^{N} [I(M_1; W_n; Y_{1,n}|Y_{1,n+1}^N, Y_{2,n-1}^N, M_2) - I(M_1; Y_{2,n}|Y_{1,n+1}^N, Y_{2,n-1}^N, M_2)] + 2N\epsilon
\]
\[
= \sum_{n=1}^{N} [I(M_1; Y_{1,n}|Y_{1,n+1}^N, Y_{2,n-1}^N, M_2) + I(W_n; Y_{1,n}|M_1, Y_{1,n+1}^N, Y_{2,n-1}^N, M_2)
\]
\[
- I(M_1; Y_{2,n}|Y_{1,n+1}^N, Y_{2,n-1}^N, M_2)] + 2N\epsilon
\]
\[
= \sum_{n=1}^{N} [I(M_1; Y_{1,n}|Y_{1,n+1}^N, Y_{2,n-1}^N, M_2) + H(W_n|M_1, Y_{1,n+1}^N, Y_{2,n-1}^N, M_2)
\]
\[
- H(W_n|M_1, Y_{1,n}, Y_{1,n+1}^N, Y_{2,n-1}^N, M_2) + I(M_1; Y_{2,n}|Y_{1,n+1}^N, Y_{2,n-1}^N, M_2)] + 2N\epsilon
\]
where (a) follows since the genie gives $D_1$ message $M_2$, (b) and (c) follow from Csiszár’s sum identity. Letting $U_n \triangleq (Y_{1,n}^{N_n}, Y_{2,n}^{-1})$, $V_{1,1} = \cdots = V_{1,N} \triangleq M_1$ and $V_{2,1} = \cdots = V_{2,N} \triangleq M_2$, where $U$, $V_1$ and $V_2$ satisfy the Markov chains $U \rightarrow V_1 \rightarrow X$ and $U \rightarrow V_2 \rightarrow X$ specified in Section III for the channel $C_3$, $R_1$ can be bounded as

$$NR_1 \leq \sum_{n=1}^{N} [I(V_{1,n}; Y_{1,n} | U_n, V_{2,n}) + H(W_n | U_n, V_{1,n}, V_{2,n}) - I(V_{1,n}; Y_{2,n} | U_n, V_{2,n})] + 2N\varepsilon. \quad (45)$$

Similarly,

$$NR_1 \leq \sum_{n=1}^{N} [I(V_{2,n}; Y_{2,n} | U_n, V_{1,n}) + H(W_n | U_n, V_{1,n}, V_{2,n}) - I(V_{2,n}; Y_{1,n} | U_n, V_{1,n})] + 2N\varepsilon. \quad (46)$$

A bound on the sum rate is given by

$$N(R_1 + R_2) \leq \sum_{n=1}^{N} [I(V_{1,n}; Y_{1,n} | U_n, V_{2,n}) + I(V_{2,n}; Y_{2,n} | U_n, V_{1,n}) - I(V_{1,n}; Y_{2,n} | U_n, V_{2,n}) - I(V_{2,n}; Y_{1,n} | U_n, V_{1,n}) + 2H(W_n | U_n, V_{1,n}, V_{2,n}) + 4N\varepsilon. \quad (47)$$

Finally, a time sharing RV $Q$, which is uniformly distributed over $N$ symbols and independent of the RVs $M_1, M_2, W, U, V_1, V_2, X, Y_1$ and $Y_2$ is introduced for the single letter characterization of the above derived outer bounds. Applying the procedure similar to the one presented in [26, Chapter 15.3.4] on (37) - (39); (40) - (41); (42) - (44); and (45) - (47), we get the outer bounds (11) - (13), (14) - (15) and (16) - (18). This completes the proof of Theorem 3.2 for the channel $C_3; c = 1, 2, 3$.

For the channel $C_3$, the outer bound on $R_1 + R_2$ can be made tighter by the following procedure. From (16) - (24), we see that

$$R_1 + R_2 \leq I_1 + I_2, \quad (48)$$

$$R_1 + R_2 \leq I_1^* + I_2^*, \quad (49)$$

Therefore,

$$R_1 + R_2 \leq \min[I_1 + I_2^*, I_2 + I_1^*]. \quad (50)$$
We show now that the bound (50) is a tighter bound than (48) and (49). It is easy to see that

\[ I_1 + I_2 = I_1^* + I_2^* + I(W; V_1|U, V_2) + I(W; V_2|U, V_1). \]

Consider \(2(I_1 + I_2) = 2[I_1^* + I_2^* + I(W; V_1|U, V_2) + I(W; V_2|U, V_1)]\), which implies the following:

\[
\begin{align*}
\min[I_1 + I_2^*, I_2 + I_1^*] & \leq I_1 + I_2, \\
\min[I_1 + I_2^*, I_2 + I_1^*] & \leq I_1^* + I_2^*.
\end{align*}
\]

Therefore, the sum rate bound given by (50) is tighter than (48) and (49).

VI. CONCLUSIONS

We presented inner and outer bounds on the capacity region of three classes of two-user discrete memoryless broadcast channels, with noncausal side-information at the encoder. To prove the achievability theorem, we used an extension of Marton’s coding scheme, results from the second moment method, an extension of a method proposed by Kramer and Shamai for broadcast channels with receiver side-information, and stochastic encoders to satisfy confidentiality requirements. Outer bounds were derived using results from Gel’fand-Pinsker’s channel and utilizing confidentiality constraints. For channels where each receiver has \textit{a priori} knowledge of the message of the other receiver, we showed that the bounds are tight, thereby yielding the capacity region for that class of channels. Future work could involve, among other things, considering causal side-information at the encoder/decoder; Gaussian channel models; and deriving tighter bounds on the sum-rates.

APPENDIX A

Here, we upper bound the probability of encoder error for the channel \(C_1\), by using results from the second moment method [19]. This method was also employed in [28] and [29, Chap. 7, pp. 354] to provide an alternative proof of Marton’s achievability scheme. An error is declared at the encoder of \(S\) if it is not possible to find a pair \((i_1, i_2)\) to satisfy the condition \(E_S \triangleq \{(W, V_1(i_1, j_1), V_2(i_2, j_2)) \in A^{(N)}(P_{W, V_1, V_2})\}\). Let \(P_{e, E_S}\) denote the probability of error at the encoder, \textit{i.e.}, \(P_{e, E_S} \triangleq \Pr(E_S)\). Let \(I\) be an indicator RV that the event \(E_S\) has occurred. Let
\[ Q = \sum_{j_1, j_2} I, \quad \tilde{Q} = \mathbb{E}[Q]; \text{ and } \text{Var}[Q] = \mathbb{E}[(Q - \tilde{Q})^2], \text{ where } \mathbb{E}(\cdot) \text{ denotes the expectation operator.} \]

\[ P_{e,E_8} \text{ can be upper bounded as follows:} \]

\[ P_{e,E_8} = \Pr(Q = 0) \leq \text{Var}[Q]/\tilde{Q}^2, \quad (51) \]

where \((i)\) follows from Markov’s inequality for non-negative RVs. Consider now

\[ \tilde{Q} = \sum_{j_1, j_2} \mathbb{E}(I) \geq \sum_{j_1, j_2} (1 - \delta^{(N)}) 2^{-N[I(V_1;V_2|U)+I(V_1,V_2;W|U)+4\epsilon]} \]

\[ = \left(1 - \delta^{(N)}\right) 2^{-N[R_1^* + R_2^* - I(V_1;V_2|U) - I(V_1,V_2;W|U) - 4\epsilon]} \]

Next, consider \(\text{Var}[Q] = \sum_{j_1,j_2} \sum_{j_1',j_2'} \{ \mathbb{E}[I(j_1,j_2)I(j_1',j_2')] - \mathbb{E}[I(j_1,j_2)]\mathbb{E}[I(j_1',j_2')] \}. \) We have the following four cases:

1) If \(j_1' \neq j_1 \) and \(j_2' \neq j_2\), then \((I(j_1,j_2) \text{ and } I(j_1',j_2'))\) are independent and \(\text{Var}[Q] = 0\).
2) If \(j_1' = j_1 \) and \(j_2' = j_2\), then \(\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] = \mathbb{E}[I(j_1,j_2)] \leq 2^{-N[I(V_1;V_2)+I(V_1,V_2;W) - 4\epsilon]} \).
3) If \(j_1' \neq j_1 \) and \(j_2' = j_2\), then \(\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] \leq 2^{-N[I(V_1;V_2|U)+I(V_1,V_2;W)+I(V_1;V_2,W) - 4\epsilon]} \).
4) If \(j_1' = j_1 \) and \(j_2' \neq j_2\), then \(\mathbb{E}[I(j_1,j_2)I(j_1',j_2')] \leq 2^{-N[I(V_1;V_2|U)+I(V_1,V_2;W)+I(V_1;V_2,W) - 6\epsilon]} \).

Substituting for \(\tilde{Q}\) and \(\text{Var}[Q]\) in \((51)\), we can show that \(P(E_8) \leq \delta_C^{(N)}, \forall \delta_C^{(N)} > 0 \) and sufficiently small; and for \(N\) large, if the following conditions are simultaneously satisfied:

\[ R_1' > I(W; V_1) - \epsilon_1, \quad (52) \]

\[ R_2' > I(W; V_2) - \epsilon_2, \quad (53) \]

\[ R_1' + R_2' > I(V_1; V_2) + I(V_1, V_2; W) - \epsilon_3. \quad (54) \]

Similar analysis is done to bound the binning rates for the channel \(C_3\). The probability of encoder error \(P(E_8) \leq \delta_C^{(N)}, \forall \delta_C^{(N)} > 0 \) and sufficiently small; and for \(N\) large, if the following conditions are simultaneously satisfied:

\[ R_1^* > I(W; V_1|U) - \epsilon_{11}, \quad (55) \]

\[ R_2^* > I(W; V_2|U) - \epsilon_{12}, \quad (56) \]

\[ R_1^* + R_2^* > I(V_1; V_2|U) + I(V_1, V_2; W|U) - \epsilon_{13}. \quad (57) \]
REFERENCES


