Decode-Forward Relaying with State Available Noncausally at the Relay

Jie SHEN1*, Da-jin WANG2, Ou WANG3 and Geng ZHANG1

1China Electric Power Research Institute Beijing, China
2Spreadtrum Communications, Inc. Shanghai, China
3School of Electronic Engineering, Beijing University of Posts and Telecommunications. Beijing, China

*Corresponding author

Keywords: Relay channel, Channel state information, Decode-forward, Discrete memoryless multicast network.

Abstract. The problem of relay channel (RC) with noncausal channel state information (CSI) available at the relay is considered. With the CSI, the relay can help communication in two ways: 1) by relaying message information; 2) by conveying the CSI to help the destination decode. In previous work, Zaidi et al. established a lower bound by letting the relay send message information by performing Gelfand-Pinsker (GP) coding. While in our schemes, we combine the two ways by letting the relay send message information as well as the compressed CSI to the receivers. We investigate three decode-forward (DF) lower bounds. The first two bounds are obtained by transmitting the message information and the CSI directly. We find this scheme is surprisingly equivalent to Zaidi’s. The third bound is established by exploiting a new coding scheme. We show that the third bound is tighter than others. We also generalize RC to discrete memoryless multicast network (DM-MN). To our knowledge, we are the first to characterize the DF lower bound and cutset upper bound for DM-MN with noncausal CSI only at the relay nodes. We also give the condition under which DF lower bound is tight.

Introduction

The problem of state dependent (SD) channel with channel state information (CSI) causally or noncausally available at the communication nodes has been extensively studied [1-6]. Causal CSI at the transmitter was first studied by Shannon [1]. For noncausal CSI, Gelfand and Pinsker [2] introduced the Gelfand-Pinsker (GP) coding and characterized the capacity of the SD single-user channel with noncausal CSI only at the transmitter. Costa [4] applied GP coding to the Gaussian channel and showed that known interference does not reduce capacity by dirty paper coding (DPC).

The capacity of relay channel has been exhaustively studied in [7], and some (by now standard) techniques were proposed, such as decode-forward (DF) coding and compress-forward (CF) coding. Recently, the state dependent relay channel (SDRC) with discrete memoryless (DM) channel state has been studied excessively. For SD-RC with noncausal CSI, the DF lower bound is established in [8,9]. And the CF lower bound is characterized in [10,11]. Zaidi et al. studied the case where noncausal CSI at only the relay and gave the DF lower bound and an upper bound [9]. In [12], the DF lower bound for the noncausal CSI only at the source was derived.

The key contribution of our paper (and the key difference from all previous work) is that the relay not only performs GP coding, but also sends compressed CSI to the destination. So the destination can use the received CSI to help it decode. In all the previous work on SD-RC with CSI [8,9,10,13], the CSI is only used for precoding. That is using Shannon strategy for causal CSI and GP coding for noncausal CSI. For the utility of the noncausal CSI, we introduce a new coding scheme to send both the message information and the CSI. Thus we get a tighter bound based on this new scheme.

Generally, the transmitter should not send the CSI, because the transmission of CSI will occupy the channel and lower the rate of message information. Moreover the merit of CSI can’t outperform
the message information [6]. However for the relay in a SD-RC, we show that transmitting CSI is reasonable, that is which send on multicast or broadcast channels by helpers without occupy information channel. We get three DF lower bounds and compare them with Zaidi’s in [9]. Our first two DF lower bounds are obtained by letting the relay send message information and CSI directly. Although this scheme is different to Zaidi’s in [9], the bounds corresponding to the two different schemes are surprisingly equivalent. Finally we characterize a tighter bound by exploiting a new coding scheme.

We also generalize SD-RC to discrete memoryless multicast network (DM-MN) with state. To our knowledge, we are the first to establish the DF lower bound and cutset upper bound for DM-MN with state. Then we give the condition under which the DF lower bound is tight.

The notations used in this work are as follows. Upper case letters denote random variables, e.g. $X$; lower case letters denote realizations of random variables, e.g. $x$; and calligraphic letters designate alphabets, e.g. $\mathcal{X}$. The shorthand notation $X^i$ indicates a sequence of random variables $(X_i, X_{i+1}, \ldots, X_j)$ and $x^i$ denotes a particular realization of a random sequence $X^i$. Specially, $X^i$ denotes $X^i_1$ and $x^i$ denotes $x^i_1$.

The rest of the paper is organized as follows. In section II, channel models are given. In section III, we focus on SD-RC, and characterize several DF lower bounds. In this section, we introduce a new coding scheme, by which we can make use of the noncausal CSI to send both the message information and the compressed CSI. In section IV, we generalize SD-RC to DM-MN with state. In section V, we establish the cutset upper bound for DM-MN with noncausal CSI only at the relay nodes and give the condition under which the DF strategy is optimal. In section VI, we give the conclusion.

Preliminaries

In this paper, We consider DM channel and DM CSI only. A SD-RC with CSI consists of five finite sets $X_1, X_2, Y_1, Y_2, S$ and a collection of conditional pmfs $p(y_2, y_3 | x_1, x_2, s)$ on $Y_2 \times Y_3$. Source $X_1$ wishes to send a message $M$ to destination $Y_3$ with the help of the relay $(X_2, Y_2)$. The relay is informed of the noncausal CSI $S$. The CSI is noncausal in the sense that the entire sequence of the channel state is known in advance and can be used for encoding from time $i = 1$.

A $(2^a, n)$ code for a SD-RC consists of:

1. A message set $[1: 2^a]$;
2. A source encoder that assigns a codeword $x^a_i(m)$ to each message $m \in [1: 2^a]$;
3. With noncausal CSI $s^a_2$, the relay encoder that assigns at time $i \in [1:n]$ a symbol $x^i_2(y^{i-1}_2, s^a_2)$ to each past received sequence $y^{i-1}_2 \in Y^{i-1}_2$ and CSI $s^a_2 \in S^a_2$;
4. A decoder that assigns a message $\hat{m}$ or an error message $e$ to each received sequence $y^n_3 \in Y^n_3$.

The channel is memoryless in the sense that current received symbols $(Y_2, Y_3)$ and the past symbols $(X^{i-1}_1, X^{i-1}_2, Y^{i-1}_2, Y^{i-1}_3)$ are conditionally independent given the current transmitted symbols $(X_i, X_{i+1})$. We assume that the message $M$ is uniformly distributed over $[1: 2^a]$ and the average probability of error is $P_e(a) = P\{M \neq \hat{M}\}$. The rate $R$ is said to be achievable for the SD-RC if there
exists a sequence of \((2^{2^n}, n)\) codes with \(P_e^{(n)} \to 0\) and the capacity \(C\) of the SD-RC is the supremum of all achievable rates.

An \(N\)-node discrete memoryless multicast network (DM-MN) with state \((X_1 \times \cdots \times X_N, p(y_2, \cdots, y_N | x_1, \cdots, x_N, s_1, \cdots, s_N), Y_2, \cdots, Y_N)\) consists of a source alphabet \(X_1, N-1\) sender-receiver alphabet pairs \((X_i, Y_i), j \in [2:N]\), and a collection of conditional pmfs \(p(y_2, \cdots, y_N | x_1, \cdots, x_N, s_1, \cdots, s_N)\). Here \(s_1, \ldots, s_N\) is the noncausal partial CSI available at each node. In this paper, we consider the case where only relays know the partial noncausal CSI, so we just set \(s_i = s_N = \emptyset\). Source node 1 wishes to send a message \(M\) to a set \(\mathcal{D} \subseteq [2:N]\) of destination nodes.

Figure 2. A DM-MN with different noncausal CSI at its nodes. The whole network can be divided into two sets:

The one denoted by \(S\) contains the source and the other denoted by \(S^C\) contains the destination.

In a large network, different nodes subject to different channel state, which can be modeled as partial CSI with regard to the whole CSI of the network.

A \((2^{2^n}, n)\) code for a DM-MN with state is similar to the SD-RC, which contains:
1. A message set \([1:2^{2^n}]\);
2. A source encoder that assigns at time \(i \in [1:n]\) a codeword \(x_i^m(m)\) to each message \(m \in [1:2^{2^n}]\);
3. A set of \(N-1\) relay encoders: Encoder \(k\) assigns a symbol \(x_k(y_{s_k}^{i-1}, s_i^k)\) to every received sequence \(y_{s_k}^{i-1}\) and CSI \(s_i^k\) for \(i \in [1:n]\);
4. A set of decoders: Decoder \(k, k \in \mathcal{D}\), assigns a message \(\tilde{m}\) or an error message \(\varepsilon\) to each received sequence \(y_i^m\) and CSI \(s_i^m\).

For DM-MN, we assume that each node only knows CSI partially, i.e. the CSI at each relay nodes is different and correlated. This assumption is reasonable for a large network, different nodes in a distance may suffer from different channel state. So each node only cares about its own channel state regardless of others, which is modeled as partial CSI instead of the overall CSI of the whole network.

**DF Lower Bound for RC**

In this section, we establish three DF lower bounds for SD-RC with noncausal CSI available at the relay. The first two bounds are obtained by directly transmitting message information and the CSI. Unlike the scheme in [8,9], the relay doesn’t perform GP coding but sends compressed CSI to the destination. The decoder then uses the received CSI to help it decode the message both from the source and the relay. The third bound is based on a new coding scheme, and it is tighter than the first two bounds.

**Scheme I**

In scheme I: the relay sends message information and CSI compression index. For point to point channel with noncausal CSI available only at the transmitter, GP coding achieves the capacity. Although the CSI is useful at the receiver side, the transmitter would better not send CSI to the receiver. Because the transmission of CSI will lower the rate of message; and the merit of CSI can’t outperform the message information [6]. So typically the transmitter sends the message only. But in the scenario of RC, the CSI received at the destination can be used twice: help the destination
decode the message 1) from the source; 2) from the relay. So the sending of CSI exactly compensates the loss of message information. By letting the relay send message information and CSI directly, we get the first lower bound as follows.

**Theorem 1:** The capacity of the SD-RC with noncausal CSI $S_2$ available at the relay is lower bounded by

$$C \geq \max_{p(x_1)p(y_2)p(z_2|x_1,y_2)} \min \{ I(X_1;Y_2 | S_2, U_1, X_2), I(X_1;U_1,X_2;Y_2) - I(Y_2;U_1) \}$$

s.t. $0 \leq R \leq I(X_2;S_2|U_1)$

The auxiliary random variable $U_1$ contains the message information for relaying, and $X_2$ contains the combination of message information and the compressed CSI.

The lower bound established by using DF strategy and GP coding [9] is shown as below.

$$C \geq \max_{p(x_1)p(y_2)p(z_2|x_1,y_2)} \min \{ I(X_1;Y_2 | S_2, U_1, X_2), I(X_1;U_1,U_2;Y_2) - I(U_2;S_2 | U_1) \}$$

The bound (2) in [9] is based upon a technique called codeword splitting, combining DF relaying. In that technique the relay uses two codebooks. The first one is joint with the source to transmit information cooperatively, and the second one is the GP coding codebook for the utility of CSI at the relay. The bound (1) in Theorem 1 is obtained by simply letting the relay directly transmit message information with compressed CSI. Although the schemes corresponding to (1) and (2) are totally different, the two bounds are surprisingly equivalent as revealed in the following theorem.

**Theorem 2:** The lower bound (1) and (2) are equivalent.

To compare the bound (1) and (2), replace $X_2$ in (1) with $U_2$. Now every pair of identical symbols in the two bounds follow the same form of probability distribution. The second terms of the two bounds are identical. For the first terms, note that $Y_2 \rightarrow (U_2, S_2) \rightarrow (X_2, S_2)$ forms a Markov chain. So $I(X_1;Y_2 | S_2, U_1, X_2) \leq I(X_1;Y_2 | S_2, U_1, U_2)$. The equality holds when $X_2$ is a deterministic mapping of $(U_1, U_2)$. In fact, this term represents the constraint between the source and the relay. As the coding scheme for the source corresponding to (1) and (2) are the same, the two constraints should be identical.

That means for the relay, transmitting compressed CSI with the message information has the same effect with relaying message by GP coding.

**Scheme II**

In scheme 2: Similar coding of scheme 1 but with inde-pendent codewords for message information and CSI compression index. The relay sends both the message information and the CSI. As the two kinds of information are independent, they can be sent separately. Denote the message information by $U_1$ and the compressed CSI by $U_2$. The relay sends $X_2$ as a deterministic function of $(U_1, U_2)$.

**Corollary 1:** The capacity of the SD-RC with noncausal CSI $S_2$ available at the relay is lower bounded by

$$C \geq \max_{p(x_1)p(y_2)p(z_2|x_1,y_2)} \min \{ I(X_1;Y_2 | S_2, X_2), I(X_1,X_2;Y_2) - I(U_2;S_2) \}$$

where $X_2 = (U_1, U_2)$.

In the scheme with respect to the bound (1), the relay generates compressed CSI $U_2$ according to the message information $U_1$ and CSI $S_2$. Since the message information and the CSI are independent, there is no benefit in generating $U_2$ according to $U_1$. So the bound (3) and (1) are equivalent. According to Theorem 2, the bound (3) is equivalent to (2) too.
With the noncausal CSI, the relay can perform GP coding to send the \((u_1^*, u_2^*)\) pairs: For each \((u_1^*, u_2^*)\) pair, randomly and independently generate a subcodebook to match the CSI \(s_2^*\). However, this scheme fails to make use of the relationship between the compressed CSI \(U_2\) and the CSI \(S_2\). Here we exploit a new coding scheme and get the following result.

**Scheme III**

In scheme 3: The paper combines Gelfand-Pinsker (GP) binning scheme and scheme 2. Although, GP code binning is to remove to effect of CSI at the destination, in multicast channel the relay has no knowledge of whether the CSI is non-causally or not, under such circumstance, the helper can send compression index and not consumption of extern bandwidth.

**Theorem 3**: The capacity of the SD-RC with noncausal CSI \(S_2\) available at the relay is lower bounded by

\[
C \geq \max_{u_1, u_2} \min\{I(X_1; Y_2 | S_2, U_1, X_2), I(X_1; Y_3 | U_1, U_2) + I(W; Y_1) - I(W; S_2 | U_1) - I(U_1; S_2)\}
\]

where the maximum takes over the pmf of the form

\[
p(u_1) p(x_1 | u_1) p(u_2 | x_2) p(w_2 | u_1, u_2, s_2)
\]

For the proof, we only give the details of the new coding scheme. The rest part is similar to the proof of Theorem 1 in the appendix.

1. At the relay, the CSI \(s_2^*(j)\) is compressed as \(u_2^*(j)\) at rate \(R' = I(U_2; S_2)\). Here \(j\) is the index of the coding block.
2. For each sequence \(u_2^*(l_j)\), \(l_j \in [1: 2^{R'}]\), randomly and independently generate \(2^{N_R}\) sequences \(w^R\).
3. Then partition these \(2^{N_R}\) sequences \(w^R\) into \(2^{N_R}\) bins indexed \(m_j\).

![Figure 3. The new coding scheme.](image)

In this way, there are \(2^{N_R} \times 2^{N_S}\) sequences for each \(u_2^*(m_j)\), and \(2^{N_R} \times 2^{N_S}\) sequences in a subcodebook corresponding to each \((u_2^*(l_j), u_1^*(m_j))\) pair. Given the CSI \(S_2^*\), assume that the relay wants to send \(U_1^*(M_j)\) and \(U_2^*\). It finds the sequences \(W^R\) satisfies that \((U_2^*(l_j), W^R, S_2^*) \in T^{N_R}_{w} \) and \(W^R \in B^{N_S}(M_j)\).

Note that \((S_2^*, U_2^*)\) is already a joint typicality pair and \(w^R\) is generated according to \(U_2^*\). By the covering lemma, it is enough for only \(2^{N_R} \times 2^{N_S}\) sequences in each subcodebook rather than \(2^{N_R} \times 2^{N_S}\). The relay then transmits a deterministic mapping of \((W, S_2^*)\), that is \(x_3 = x_3(w, s_2^*)\).

The new coding scheme can be taken as an improved GP coding for sending \((u_1^*, u_2^*)\) pairs. So the term \(I(W; Y_1) - I(W; S_2 | U_1)\) is the capacity of GP channel. The term \(I(X_1; Y_3)\) in (1) is the capacity of common point to point channel without noncausal CSI. So \(I(W; Y_1) - I(W; S_2 | U_1)\) is greater than \(I(X_1; Y_3)\) and the bound of (4) is tighter than (1).
DF Lower Bound for DM-MN

In this section, we generalize RC to DM-MN and establish the DF lower bound for DM-MN with noncausal CSI at the relay nodes. With noncausal CSI, the relay nodes directly transmit compressed CSI and message information to the next nodes. To the best of our knowledge, we are the first to get the DF lower bound for DM-MN with noncausal CSI available at the relay nodes.

**Theorem 4:** The capacity of the DM-MN $p(y_1, \cdots, y_N, x^N, s^N)$ with noncausal CSI at the relay nodes and with any set $D \subseteq [2:n]$ of destination nodes is lower bounded by

$$C \geq \max_{p(x^N, s^N, k) \in [1:N-1]} \min \{I(X^k; Y_k | X^N, S_k), I(X^N; S^k)\}$$  \hspace{1cm} (5)

**Remark 1:** The DM-MN reduce to RC by setting $N = 3$, $S_i = S_j = \emptyset$ and $Y_i = X_j = \emptyset$. And it is easy to verify that Theorem 4 reduce to Corollary 1 in this case.

In the proof, we use block Markov encoding and sliding window decoding [15,16].

**Codebook generation:** Fix the pmf $p(x^N, s^N)$ that achieves the lower bound.

1. Randomly generate a sequence $u^x_k(j)$ according to $\prod_{l=1}^{n} p(u_{l})$.
2. For each relay node $k = N - 1, N - 2, \cdots, 1$, and for each $(u^x_k(m_{j-k}) | m_{j-k+1}), \cdots, u^x_{N-1}(m_{j-N+2}), u^x_{N}(j))$, generate $2^N$ conditionally independent sequences $w^x_k(m_{j-k+1}, m_{j-N+2})$, $m_{j-k+1} \in [1:2^{2n}]$, each according to

$$P_{w^x_k(m_{j-k+1}, m_{j-N+2})}(w_k | u_{j-k+1}(m_{j-k}), \cdots, u_{N-1}(m_{j-N+2}), u^x_{N}(j))$$

This defines the codebooks

$$C_j = \{w^x_k(m_{j-k}, m_{j-N+2}), \cdots, w^x_{N-1}(m_{j-N+2}), w^x_{N}(j)\}$$

where $m_{j-N+2}, \cdots, m_{j} \in [1:2^{2n}]$, $l_{j} \in [1:2^{m}]$.

3. For each relay node $k = N - 1, N - 2, \cdots, 1$ and for each $(u^x_k(m_{j-k}) | m_{j-k+1}), \cdots, u^x_{N-1}(m_{j-N+2}), u^x_{N}(j))$, generate $2^N$ conditionally independent sequences $v^x_k(l_{j-k+1}, m_{j-N+2})$, $l_{j-k+1} \in [1:2^{2n}]$, each according to

$$P_{v^x_k(l_{j-k+1}, m_{j-N+2})}(v_k | u_{j-k+1}(l_{j-k}), \cdots, u_{N-1}(l_{j-N+2}), u^x_{N}(j))$$

**Encoding:** Let $m_{j} \in [1:2^{n}]$ be the new message to be sent in block $j$, the encoder transmits $x^x_{j}(m_{j} | m_{j-k+1})$ from the codebook $C_j$. At the end of block $j \in [2:N]$, each according to $(\hat{m}_j, \hat{l}_j)$ of the message $m_j$ and a compressed version of its CSI $(\hat{l}_j)$. In block $j = 1$, it finds $w^x_{k}(\hat{m}_j, \hat{l}_j)$ from the codebook $C_{j=1}$ and $v^x_{k}(\hat{l}_j)$, then sends $x^x_{j}(\hat{m}_j, \hat{l}_j)$. Node $N$ sends $x^x_{N}(j)$ in block $j \in [1:N]$.

**Decoding:** At the end of $(j \in [N-2])$ block, upon receiving $y^x_{j}(j \in [N-2])$, the relay node $k$ finds the unique $\hat{m}_j$ that satisfies the following conditions simultaneously:

$$(X^x_{j}(\hat{m}_j, \hat{l}_j), X^x_{j}, \cdots, X^x_{N}(j), Y^x_{j}(j)) \in T'^{x}_{j}$$

$$(X^x_{j}(\hat{m}_j, \hat{l}_j), X^x_{j}, \cdots, X^x_{N}(j+1), Y^x_{j}(j + 1)) \in T'^{x}_{j}$$

$$(X^x_{N}(j + 1), X^x_{N}, \cdots, X^x_{j}(j + k - 1), Y^x_{j}(j + k - 1)) \in T'^{x}_{j}$$

here the dependence of code words on previous message indices $m_{j-k+1}$ is suppressed for brevity.

Following similar steps to the proof of the sliding-window decoding for the relay channel [14], by the independence of the codebooks, the LLN and the joint typicality lemma, it can be shown that the probability of error tends to 0 as $n \to \infty$, provided that

263
\[
R < \sum_{k=1}^{K} I(X_k;Y_k \mid X_{k-1},S_k) - I(X_k;S_k) - \delta(e)
\]
\[
= I(X^{k-1};Y_k \mid X^N_k, S_k) - I(X^{k-1};S^{k-1})
\]

So we get (5), that completes the proof.

Remark 2: Theorem 4 also holds for causal CSI. In that case we don’t use GP coding and the relays just convey compressed CSI to the destination. To decode the message of current slot \( j \), the decoder of the destination can wait for the corresponding compressed CSI \( u^N_2(j) \) coming in the next time slot \( j+1 \).

**Cutset Upper Bound**

In this section, we derive the cutset upper bound for DM-MN with the noncausal CSI at the relay nodes and give a condition under which the DF strategy is optimal. We also specialize our results to SD-RC and compare with those in [9].

**Theorem 5**: The capacity of the DM-MN \( p(y_2, \ldots, y_N \mid x^N, s^N) \) with noncausal CSI available only at the relay nodes is upper bounded by

\[
C \leq \max_{p(x|y^N)} \min_{S \in \mathcal{S}} \{ I(X(S);Y(N) \mid X, S) - I(X(S);S) \} \tag{6}
\]

The proof can be found in the appendix. In order to investigate the condition under which the lower bound coincides with the upper bound, we give another form as below. It specifies the cutset \( S \) as the form \( S = [1 : k] \), \( k \in [1 : N - 1] \).

**Corollary 2**: The cutset upper bound with cutsets of the form \( S = [1 : k] \), \( k \in [1 : N - 1] \) can be written as

\[
C \leq \max_{p(x|y^N)} \min_{S \in \mathcal{S}} \{ I(X(N);Y(N) \mid X, S) - I(X(N);Y(N), S) \} \tag{7}
\]

For classic DM-MN (without channel state), the DF lower bound is tight when the DM-MN is physically degraded [14,17], i.e.,

\[
p(y^N_2 \mid x^N, y^{k-1}) = p(y^N_2 \mid x^N, y_{k-1}) \quad \text{for } k \in [2 : N] \tag{8}
\]

where \( y_1 = \emptyset \) by convention.

For DM-MN with noncausal CSI available at the relay nodes, we have the lower bound (5) and the upper bound (7). Consider that the CSI at node \( k \in [2 : N] \) satisfies

\[
p(y^N_k \mid x^N, s^N, y^{k-1}) = p(y^N_k \mid x^N, s_{k-1}, y_{k-1}) \tag{9}
\]

That means for each node \( k \in [2 : N] \), \( (S_2, \ldots, S_{k-1}, S_{k+1}, \ldots, S_N) \rightarrow (X^N, S_{k+1}) \rightarrow Y^N_{k+1} \) forms a Markov chain. So we have

\[
\max_{p(x|y^N)} \min_{S \in \mathcal{S}} \{ I(X^N;Y^N_{k+1} \mid X^N_{k+1}, S^N) \}
= \max_{p(x|y^N)} \min_{S \in \mathcal{S}} \{ I(X^N;Y^N \mid X^N_{k+1}, S) \} \tag{10}
\]

Since \( S^k \rightarrow X^k \rightarrow Y^N_{k+1} \rightarrow X^N_{k+1} \) forms a Markov chain, the follow equality holds

\[
\max_{p(x|y^N)} \min_{S \in \mathcal{S}} \{ I(X^N;S \mid X^N_{k+1}) \} = \max_{p(x|y^N)} \min_{S \in \mathcal{S}} \{ I(X^N;S) \} \tag{11}
\]

It is easy to see the DF lower bound (5) and the cutset upper bound (7) are coincide when the equality (11) holds for all the nodes \( k \in [2 : N] \). So we get a condition under which DF lower bound is tight.
**Theorem 6:** The lower bound (5) is tight whenever (9) holds for all the nodes $k \in [2:N]$.

By setting $N = 3$, DM-MN reduces to RC. The bound (7) reduces to (13) and we get the upper bound for RC with noncausal CSI at the relay.

**Corollary 3:** The capacity of the RC $p(y_2, y_3 | x_1, x_2, s)$ with noncausal CSI $S$ available at the relay node is upper bounded by

$$C \leq \max_{p(x_1;x_2)} \min \{I(X_1; Y_1 | S, X_2), I(X_1, X_2; Y_1 | S) - I(X_2; S)\}$$

The upper bound in [9] is shown as follows.

$$C \leq \max_{p(x_1;x_2)} \min \{I(X_1; Y_1 | S, X_2), I(X_1, X_2; Y_1 | S) - I(X_2; S | Y_1)\}$$

The only difference is that the term $I(X_1; S | Y_1)$ in [9] is replaced by $I(X_2; S)$ in (13).

According to Theorem 6, the lower bound (3) is tight if the following equation holds.

$$p(y_2, y_3 | x_1, x_2, s) = p(y_2, y_3 | x_1, x_2)$$

That means $X_2$ contains all the information of $S$. This needs the channel between the relay to the destination is better enough to transmit both the message and the CSI.

**Simulation**

The available rate region for the three schemes under Gaussian channel with noncausal CSI at source, as the assumption of [20] is plotted for $P_1 = Q = N_3 = 10dB$, and $P_2 = 16dB, N_2 = 0dB$.

![Figure 4. GP based achievable rate region for general Gaussian DM-MN with non-causal CSI only at the source, in different scheme.](image)

where setting the power and noise constraint for source as $P_1, N_1$, and $Q_1$ represents the bin of compressed signal’s index.

This scenery is parameterized by the fraction of the relay’s power dedicated to relay’s private message, and we can see the trade-off between $R_{13}$ and $R_{12}$. As illustrated in Fig. 4, when the fraction is increased, the maximum achievable $R_{13}$ is decreased. For example, the case of scheme 3 where the relay sends no private message to the destination, and Gaussian SDRC reduces to Gaussian DM-MN. Therefore, for this case we can compare our CF based achievable region, with the DF based achievable region for DM-MN with informed source only.

We illustrate this outer bound, DF and CF based rate regions in Fig. 5 for examples of noise configuration: $N_2 = N_3 = 10dB, P_1 = 21dB$. 

$$\begin{align*}
(N_2 = N_3 = 10dB, P_1 = 21dB).
\end{align*}$$
We can also get situation of $N_3 > N_2$ and $N_4 > N_2$, all these situations $Q = 10 dB, P_2 = 26 dB$. For these power sets, CF outperforms DF in SD-RC (state-independent). So, in the state-dependent version of this channel it can outperform DF which is shown in Fig. 5. In these examples, if $P_2$ is also increased, CF based bound becomes closer to its respective outer bound. We remark that these power sets are chosen to have cases that can be presented in one figure 5. It is also interesting to note that for all these schemes, maximum achievable $R_{12}$ (when $R_{13} = 0$) coincides with that of derived for DM-MN (state-independent), and this confirms that a complete state cancellation is performed for sending $U_5$ on the source-relay link for $R_{13} = 0$. Simulation shows that the scheme 3’s rate outer bound is tighter than the other twos.

Conclusion

In this paper, three DF lower bounds for SD-RC are established. The first two bounds, which are equivalent to Zaidi’s, are obtained by letting the relay send message information and compressed CSI directly. The third bound, which is tighter, is obtained by exploiting a new coding scheme. Then SD-RC is generalized to DM-MN with state. The DF lower bound and cutset upper bound are obtained for DM-MN with state. The condition, under which the DF lower bound is tight, is also given.

Appendix A

Proof of Theorem 1

In this scheme, the relay directly sends a description of message information by Wyner-Ziv coding [18,19] and compressed CSI obtained by lossy source coding [14].

Codebook generation:
1. Randomly and independently generate $2^{nR}$ sequences $u^*_i(m_{j-1})$, $m_{j-1} \in [1:2^n]$, each according to $\prod_{i=1}^nP_{t_i}(u_i)$; 
2. For each $u^*_i(m_{j-1})$, randomly and conditionally independently generate $2^{nR}$ sequences $x^*_j(l_j | m_{j-1})$, each according to $\prod_{i=1}^nP_{X|U}(x_{ji} | u_i)$; 
3. For each $u^*_i(m_{j-1})$, randomly and conditionally independently generate $2^{nR}$ sequences $x^*_j(m_j | m_{j-1})$, each according to $\prod_{i=1}^nP_{X|U}(x_{ji} | u_i)$.

Encoding:
Let $m_j \in [1:2^n]$ be the message to be sent in block $j$ and $m_{j-1}$ be the message sent in the previous block. The encoder transmits $x^*_j(m_j | m_{j-1})$.

Decoding and analysis of probability of error:
At the end of block $j$, upon receiving $x^*_j(j)$, the relay finds the unique $\hat{m}_j$ such that
(A-14) for such $m_j$, the relay finds $l_j$ such that

\[(x_1^*(m_j), u_1^*(m_{j-1}), y_1^*(j), x_1^*(j)) \in T_{e}^{(n)}\]

The relay sends $x_1^*(l_j | m_j)$ at block $j+1$. Here the index $m_j$ is the Wyner-Ziv compressed message information to cooperate with the source. $l_{j-1}$ is the index of compressed CSI of block $j$ obtained by lossy source coding [14].

Correct decoding and encoding need $R < I(X_1, Y, U_1; S_2 | U_1)$ and $R > I(X_2; S_2 | U_1)$.

The decoder uses sliding window decoding. Note that $x_1^*(l_{j-1}, m_j)$ contains the CSI $s^*(j)$, which is corresponding to $x_1^*(m_j, l_{j-1})$. Upon receiving $y_1^*(j)$ at the end of block $j$, it finds $\hat{m}_j$ such that:

\[(x_1^*(\hat{m}_j, l_{j-1}), u_1^*(m_{j-1}), y_1^*(j), x_1^*(l_{j-1}, m_j)) \in T_{e}^{(n)}\]

\[(u_1^*(m_j), x_1^*(l_j, m_j), y_1^*(j + 1)) \in T_{e}^{(n)}\]

Now let’s bound the probability of error. Suppose that the sending indices $M_j = 1$, $L_{j-1} = L_j = 1$. Here we want the decoder to decode both the message and the compressed CSI correctly. Since the codebooks are generated independently for each block, these two events are independent. Hence, by the LLN and the joint typicality lemma, the probability of error $\to 0$ as $n \to \infty$ if

\[R < I(X_1; Y | U_1, X_2) + I(U_1, X_2; Y_3) - I(X_2; S_2 | U_1)\]

\[= I(X_1, U_1, X_2; Y_3) - I(X_2; S_2 | U_1)\]

**Appendix B**

**Proof of Theorem 5**

Note that only the relay nodes know the CSI, so the destination $k$ estimates $\hat{m}_j$ according to received $Y_1^*$.

Let $k \in \mathcal{D}$ and $S \subseteq \{1 : N\}$ such that $1 \in S$ and $k \in S'$ . Then by Fano’s inequality,

\[H(M | Y^*(S')) \leq H(M | Y_1^*) \leq n \epsilon_n \text{ with } \epsilon_n \to 0 \text{ as } n \to \infty\]

Consider

\[nR = H(M) \leq I(M; Y^*(S')) = \sum_{i=1}^{n} I(M; Y_i(S') | Y^{i-1}(S')) = \sum_{i=1}^{n} I(M; Y_i(S') | Y^{i-1}(S'), X_i(S'))\]

\[\leq \sum_{i=1}^{n} I(M; Y_i^{i-1}(S'), Y_i(S') | X_i(S')) = \sum_{i=1}^{n} I(M; Y_i^{i-1}(S'), Y_i(S') | S_i, X_i(S'))\]

\[= \sum_{i=1}^{n} [I(M, Y_i^{i-1}(S'), S_i^{n_i}(S); Y_i(S') | S_i, X_i(S')) - I(Y_i(S') | S_i^{n_i}(S) | M, Y_i^{i-1}(S'), S_i, X_i(S'))]\]

\[= \sum_{i=1}^{n} [I(M, Y_i^{i-1}(S'), S_i^{n_i}(S); Y_i(S') | S_i, X_i(S')) - I(Y_i(S') | S_i^{n_i}(S) | M, Y_i^{i-1}(S'), S_i, X_i(S'))]\]

\[= \sum_{i=1}^{n} [I(M, Y_i^{i-1}(S'), S_i^{n_i}(S); Y_i(S') | S_i, X_i(S')) - I(Y_i^{i-1}(S') | S_i^{n_i}(S) | M, S_i^{n_i}(S), X_i(S'))]\]

\[= \sum_{i=1}^{n} [I(M, S_i^{n_i}(S); Y_i(S') | S_i, X_i(S')) - I(M, S_i^{n_i}(S) | S_i, X_i(S'))]]\]

The term $n \epsilon_n$ is suppressed for brevity. The equality (a) follows from the fact that $S_i(S')$ is independent of $(M, Y_i^{i-1}(S'))$; (b) follows from $S_i^{n_i}(S)$ is independent of $S_i$; and (c) follows from the Csiszár sum identity. Define the encoder’s sending signal $X_i(S)$ from $(M, S_i^{n_i}(S))$, then $(M, S_i^{n_i}(S)) \to X_i(S) \to Y_i(S')$ forms a Markov chain. Thus
\[ nR \leq \sum_{i=1}^{n} [I(X_i(S);Y_i(S')|S_i,X_j(S')) - I(X_i(S);S_j(S)|X_j(S'))] + n\epsilon_n \]

By introducing a time-sharing random variable \( Q \sim \text{Unif}[1:n] \) and independent of all other random sequences, we have

\[ nR \leq nI(X_0(S);Y_0(S')|S_0,X_0(S')) - nI(X_0(S);S_0(S)|X_0(S')) + n\epsilon_n \]

that completes the proof.

References


