Semideterministic Relay Networks

Sung Hoon Lim
Department of EE
KAIST
Daejeon, Korea
Email: sunghlim@kaist.ac.kr

Young-Han Kim
Department of ECE
UCSD
La Jolla, CA 92093, USA
Email: yhk@ucsd.edu

Sae-Young Chung
Department of EE
KAIST
Daejeon, Korea
Email: sychung@ee.kaist.ac.kr

***Abstract*—An alternative coding scheme is presented for the class of semideterminsitic relay networks, which recovers the *noisy network coding inner bound* by Lim, Kim, El Gamal, and Chung for this special case.**

I. Introduction

Recently, Lim, Kim, El Gamal, and Chung [1] presented the *noisy network coding inner bound* on the capacity region for a general multi-source multicast network (see the current Proceedings) via a new network compress–forward coding scheme. This paper studies a special case of the single-source multicast semideterministic network with source input X_1, $N-2$ input–output pairs $(X_2, Y_2), \ldots, (X_{N-1}, Y_{N-1})$, and a set of D destination nodes $\mathcal{D} = \{N, \ldots, N+D-1\}$, where output symbols $Y_N, \ldots, Y_{N+D-1}$ can be expressed as deterministic functions of $(X_1, \ldots, X_{N-1}, Y_d)$ for each $d \in \mathcal{D}$, and presents an alternative coding scheme to achieve the same capacity lower bound:

$$C \geq \max_{\prod_{k=1}^{N-1} p(x_k)} \min_{d \in \mathcal{D}} \min_{\substack{\mathcal{S} \subseteq \{1,\ldots,N\} \\ 1 \in \mathcal{S}, d \in \mathcal{S}^c}} I(X(\mathcal{S}); Y(\mathcal{S}^c) | X(\mathcal{S}^c))$$

where $X(\mathcal{S}) = \{X_k : k \in \mathcal{S}\}$ and $Y(\mathcal{S}^c) = \{Y_k : k \in \mathcal{S}^c\}$. This result extends several previous results including multicast network coding over noiseless networks by Ahlswede, Cai, Li, and Yeung [2], coding for deterministic networks by Aref [3], by Ratnakar and Kramer [4], and by Avestimehr, Diggavi, and Tse [5], coding for wireless erasure networks by Dana, Gowaikar, Palanki, Hassibi, and Effros [6] and by Smith and Vishwanath [7].

The main contribution of the paper is a block based coding scheme, in which the sender transmits a single message index over multiple blocks using independent codes, the relays map *all* previously received blocks to respective codewords (unlike the network compress–forward coding scheme [8] which relays the immediate past block), and the receiver decodes the message based on all received blocks. As will be shown, the determinism of the channel and the relay operation based on the entire past history allows a simpler decoding that the general network compress–forward coding scheme.

II. Problem Statement and the Main Result

A discrete memoryless multicast relay network consists of a source input $\mathcal{X}_1$, $N-2$ relay input-output alphabet pairs $(\mathcal{X}_2, \mathcal{Y}_2), \ldots, (\mathcal{X}_{N-1}, \mathcal{Y}_{N-1})$, D desitnation outputs alphabets $\mathcal{Y}_N, \ldots, \mathcal{Y}_{N+D-1}$, and a collection of conditional pmfs $p(y_2, \ldots, y_{N+D-1} | x_1, \ldots, x_{N-1})$, one for each $(x_1, \ldots, x_{N-1}) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_{N-1}$. Source node 1 wishes to send a message $M \in [1 : 2^{nR}] := \{1, \ldots, 2^{nR}\}$ to a set of destination nodes, $\mathcal{D} = [N : N+D-1]$, with the help of a set of $N-2$ relay nodes $\mathcal{R} := [2 : N-1]$. Note that there is no loss of generality in assuming that a destination node does not relay the received information; we can always split a single node into a relay and destination node with identical channel observations.

For each $d \in \mathcal{D}$, a cut $\mathcal{S}$ is a subset of nodes such that $1 \in \mathcal{S}$ and $d \in \mathcal{S}^c$. Random vectors $X^n = (X_1, \ldots, X_n)$ and information theoretic quantities will be denoted following the notation in [9].

A $(2^{nR}, n)$ code consists of a source encoding function

$$x_1^n : [1 : 2^{nR}] \to \mathcal{X}_1^n,$$

a set of relay encoding functions

$$x_{ki} : \mathcal{Y}_k^{i-1} \to \mathcal{X}_k, \quad i \in [1:n], k \in \mathcal{R},$$

and a set of decoding functions

$$\hat{m}_d : Y_d^n \to [1 : 2^{nR}], \quad d \in \mathcal{D}.$$

Assume that the message M is uniformly distributed over $[1 : 2^{nR}]$. The probability of error is defined by

$$P_e^{(n)} = \mathsf{P}\{\hat{m}_d(Y_d^n) \neq M \text{ for some } d \in \mathcal{D}\}.$$

A rate R is said to be achievable for the multicast relay network if there exists a sequence of $(2^{nR}, n)$ codes with $\mathsf{P}_e^{(n)} \to 0$ as $n \to \infty$. The capacity C is the supremum of the achievable rates.

A discrete memoryless multicast relay network is said to be *semideterministic* if relay output symbols

$$Y_k = f_{kd}(X(\mathcal{T}), Y_d) \tag{1}$$

are deterministic functions of $X(\mathcal{T}) = (X_1, \ldots, X_{N-1})$ and Y_d for each $d \in \mathcal{D}$ and each $k \in \mathcal{R}$. In comparison to the semideterministic relay channel model (source node 1, destination node 3, and relay node 2) considered by El Gamal and Aref [10], i.e.,

$$Y_2 = f_2(X_1, X_2) \tag{2}$$

there is an additional variable Y_d in the functional relationship.

This slight difference in appearance enables to model larger classes of networks. For example, consider a multicast network in which source node 1 sends a common message to two destination nodes $\mathcal{D} = \{3,4\}$ with the help of one relay $\mathcal{R} = \{2\}$. Assume that the destination outputs are given by $Y_3 = X_1 + X_2 + Z$ and $Y_4 = X_1 + X_2 - Z$, and the relay output is given by $Y_2 = Z$ where Z is an additive noise independent of (X_1, X_2). Since the relay observation Y_2 is a deterministic function of both (X_1, X_2, Y_3) and (X_1, X_2, Y_4), this channel is a special case of (2) but cannot be represented by the deterministic models in [10], [11], [5] directly.

We are ready to state the main result of the paper, the proof of which will be given in Section III.

Theorem 1: The capacity C of a semideterministic multicast relay network is lower bounded by

$$C \geq \max_{\prod_{k\in\mathcal{T}} p(x_k)} \min_{d\in\mathcal{D}} \min_{\substack{\mathcal{S}\subseteq\mathcal{T}\cup\mathcal{D}: \\ 1\in\mathcal{S}, d\in\mathcal{S}^c}} I(X(\mathcal{S}); Y(\mathcal{S}^c)|X(\mathcal{S}^c)). \quad (3)$$

Recall the cutset upper bound on the capacity [12]:

$$C \leq \max_{p(x_\mathcal{T})} \min_{d\in\mathcal{D}} \min_{\substack{\mathcal{S}\subseteq\mathcal{T}\cup\mathcal{D}: \\ 1\in\mathcal{S}, d\in\mathcal{S}^c}} I(X(\mathcal{S}); Y(\mathcal{S}^c)|X(\mathcal{S}^c)). \quad (4)$$

The lower bound (3) is of the same form as the cutset upper bound except that the maximization is over the set of product input distributions. Hence, if the maximum in the cutset upper bound (4) for a given network is achieved by a product distribution, then both bounds are tight and the capacity is given by (3) with equality.

III. Proof of Theorem 1

The achievability of the multicast semideterministic relay network involves a new coding technique which differs from the usual block Markov coding scheme [13], [14] in two aspects. Instead of splitting the message into multiple indices and sending each of the indices per block, the source node sends a single message index over b blocks using independently generated codebooks. Furthermore, each relay node transmits a sequence that is mapped from all previous block observations instead of transmitting a sequence mapped from only the most recent block observation as in block Markov coding schemes. Therefore, the encoding operations at the source and relays

$$\begin{gathered}(\mathbf{x}_{11}(m), \ldots, \mathbf{x}_{1b}(m)), \\ (\mathbf{x}_{k1}, \mathbf{x}_{k2}(\mathbf{y}_{k1}), \ldots, \mathbf{x}_{kb}(\mathbf{y}_{k1}, \ldots, \mathbf{y}_{k,b-1})), \quad k \in \mathcal{R}, \\ \mathbf{x}_{kj} := x_{k,(j-1)n+1}, \ldots, x_{k,jn},\end{gathered}$$

closely resemble the most general relay operation

$$\begin{gathered}(x_{1,1}(m), \ldots, x_{1n}(m)), \\ (x_{k1}, x_{k2}(y_{k1}), \ldots, x_{kn}(y_{k1}, \ldots, y_{k,n-1})), \quad k \in \mathcal{R},\end{gathered}$$

except that here a block of n transmissions is the basic time unit.

We first consider the single-relay single-destination case (standard relay channel [15], [13]), i.e., $N = 3, \mathcal{R} = \{2\}, \mathcal{D} = \{3\}$, to illustrate the coding scheme.

A message $m \in [1 : 2^{nbR}]$ will be sent over b blocks of n transmissions. Thus, the overall rate of the code is R. We use separate codebooks for each block $j \in [1 : b]$. The codebooks are generated as follows:

Codebook generation: Fix $p(x_1)p(x_2)$. For each $j \in [1 : b]$, randomly and independently generate 2^{nbR} sequences $\mathbf{x}_{1j}(m)$, $m \in [1 : 2^{nbR}]$, each according to $\prod_{i=1}^n p(x_{1,(j-1)n+i})$. For $j = 1$, randomly and independently generate a sequence $\mathbf{x}_{2j}$ according to $\prod_{i=1}^n p(x_{2,(j-1)n+i})$. For $j \in [2 : b]$, randomly and independently generate $\mathbf{x}_{2j}(\mathbf{y}_{21}, \ldots, \mathbf{y}_{2,j-1})$ sequences for each $(\mathbf{y}_{2,1}, \ldots, \mathbf{y}_{2,j-1}) \in \mathcal{Y}_2^{(j-1)n}$ according to $\prod_{i=1}^n p(x_{2,(j-1)n+i})$. This defines the codebook $\mathcal{C}_j = \{\mathbf{x}_{1j}(m), \mathbf{x}_{2j}(\mathbf{y}_2^{j-1}) : m \in [1 : 2^{nbR}], \mathbf{y}_2^{j-1} \in \mathcal{Y}_2^{(j-1)n}\}$ for $j \in [1 : b]$. The codebooks are revealed to both encoders and the decoder before any transmission takes place.

Encoding: The use of codebooks is synchronized, i.e., at any given block index j, all nodes use the codebook indexed with j. Let $\mathbf{y}_{2j}$ and $\mathbf{y}_{3j}$ be the corresponding output sequences at the relay and destination when the network uses the j-th codebook, respectively. To send a message $m \in [1 : 2^{nbR}]$, the encoder transmits $\mathbf{x}_{1j}(m)$ for each block $j \in [1 : b]$. For $j = 1$, the relay encoder transmits $\mathbf{x}_{2,1}$. For $j \in [2 : b]$, the relay node sends $\mathbf{x}_{2j}(\mathbf{y}_2^{j-1})$.

Decoding: Define recursively

$$\begin{aligned}\mathbf{x}_{2j}(m) &:= \mathbf{x}_{2j}(\mathbf{y}_2^{j-1}(m)), \\ \mathbf{y}_{2j}(m) &:= \mathbf{y}_{2j}(\mathbf{x}_{1j}(m), \mathbf{x}_{2j}(m), \mathbf{y}_{3j}).\end{aligned}$$

The decoder declares that a message $m \in [1 : 2^{nbR}]$ is sent if it is the unique index such that

$$(\mathbf{x}_{1j}(m), \mathbf{x}_{2j}(m), \mathbf{y}_{2j}(m), \mathbf{y}_{3j}) \in \mathcal{T}_\epsilon^{(n)}(X_1, X_2, Y_2, Y_3)$$

for all $j \in [1 : b]$; otherwise an error is declared. Here $\mathcal{T}_\epsilon^{(n)}(X_1, X_2, Y_2, Y_3)$ denotes the ϵ-typical set [16], [9] for (X_1, X_2, Y_2, Y_3).

Analysis of the probability of error: Let

$$\mathcal{A}_j(m) := \{(\mathbf{X}_{1j}(m), \mathbf{X}_{2j}(m), \mathbf{Y}_{2j}(m), \mathbf{Y}_{3j}) \in \mathcal{T}_\epsilon^{(n)}\}.$$

Assuming $m = 1$ was sent, we have two sources of error:

$$\mathcal{E}_0 := \cup_{j=1}^b \mathcal{A}_j^c(1) \text{ and } \mathcal{E}_1 := \cup_{m\neq 1} \cap_{j=1}^b \mathcal{A}_j(m).$$

The probability $\mathsf{P}(\mathcal{E}_0) \to 0$ as $n \to \infty$ by the law of large numbers. Let

$$\mathcal{B}(j_2) := \{ j_2 \text{ is the smallest index } j \text{ such that } \mathbf{Y}_{2j}(m) \neq \mathbf{Y}_{2j}(1)\} \quad (5)$$

where $j_2 \in [1 : b+1]$. We define the event $\mathcal{B}(b+1)$ as the event that there is no index j such that $\mathbf{Y}_{2j}(m) \neq \mathbf{Y}_{2j}(1)$. The probability of $\mathcal{E}_1$ is given by

$$\mathsf{P}(\mathcal{E}_1) = \mathsf{P}\left(\bigcup_{m\neq 1} \bigcap_{j=1}^b \mathcal{A}_j(m)\right) \leq \sum_{m\neq 1} \mathsf{P}\left(\bigcap_{j=1}^b \mathcal{A}_j(m)\right).$$

Then,

$$\mathsf{P}\left(\bigcap_{j=1}^b \mathcal{A}_j(m)\right) = \sum_{j_2} \mathsf{P}\left(\bigcap_{j=1}^b \mathcal{A}_j(m) \cap \mathcal{B}(j_2)\right)$$

$$
\begin{aligned}
&\overset{(a)}{\leq} \sum_{j_2} \mathsf{P}\Bigg(\bigcap_{j=1}^{j_2-1}\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(1),\mathbf{Y}_{2j}(1),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\} \\
&\qquad \cap \bigcap_{j=j_2+1}^{b}\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(m),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\}\cap\mathcal{B}(j_2)\Bigg) \\
&\overset{(b)}{\leq} \sum_{j_2} \mathsf{P}\Bigg(\bigcap_{j=1}^{j_2-1}\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(1),\mathbf{Y}_{2j}(1),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\} \\
&\qquad \cap \bigcap_{j=j_2+1}^{b}\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(m),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\} \\
&\qquad \cap\{\mathbf{Y}_{2j_2}(m)\neq\mathbf{Y}_{2j_2}(1)\}\Bigg) \\
&\overset{(c)}{\leq} \sum_{j_2}\prod_{j=1}^{j_2-1}\mathsf{P}((\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(1),\mathbf{Y}_{2j}(1),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}) \\
&\qquad \cdot\mathsf{P}\Bigg(\bigcap_{j=j_2+1}^{b}\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(m),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\} \\
&\qquad \cap\{\mathbf{Y}_{2j_2}(m)\neq\mathbf{Y}_{2j_2}(1)\}\Bigg) \\
&\leq \sum_{j_2}\prod_{j=1}^{j_2-1}\mathsf{P}((\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(1),\mathbf{Y}_{2j}(1),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}) \\
&\qquad \cdot\mathsf{P}\Bigg(\bigcap_{j=j_2+1}^{b}\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(m),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\} \\
&\qquad \Big|\mathbf{Y}_{2j_2}(m)\neq\mathbf{Y}_{2j_2}(1)\Bigg) \\
&\overset{(d)}{\leq} \sum_{j_2}\prod_{j=1}^{j_2-1}\mathsf{P}((\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(1),\mathbf{Y}_{2j}(1),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}) \\
&\qquad \cdot\prod_{j=j_2+1}^{b}\mathsf{P}((\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(m),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)} \\
&\qquad |\mathbf{Y}_{2j_2}(m)\neq\mathbf{Y}_{2j_2}(1)),
\end{aligned}
$$

where

(a) follows by omitting $\mathcal{A}_{j_2}(m)$ in the intersection, noting that $\mathcal{B}(j_2)\subseteq\{\mathbf{Y}_2^{j_2-1}(m)=\mathbf{Y}_2^{j_2-1}(1)\}$, and using the following property [16, Lemma 21] of the ϵ-typical set:

$$
\begin{aligned}
\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(m),\mathbf{Y}_{2j}(m),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\} \\
\subseteq\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(m),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\},
\end{aligned}
$$

(b) follows since $\mathcal{B}(j_2)\subseteq\{\mathbf{Y}_{2j_2}(m)\neq\mathbf{Y}_{2j_2}(1)\}$,

(c) follows since the events

$$\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(1),\mathbf{Y}_{2j}(1),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\}, j<j_2,$$

are mutually independent, and

(d) follows since the events

$$\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(m),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\}, j>j_2,$$

are mutually conditionally independent given $\{\mathbf{Y}_{2j_2}(m)\neq\mathbf{Y}_{2j_2}(1)\}$.

Furthermore,

$$
\begin{aligned}
&\mathsf{P}\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(1),\mathbf{Y}_{2j}(1),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}\} \\
&= \sum_{(x_1^n,x_2^n,y_2^n,y_3^n)\in\mathcal{T}_\epsilon^{(n)}} p(x_1)\cdot p(x_2,y_2,y_3) \\
&\leq 2^{-n(I(X_1;Y_2,Y_3|X_2)-\delta(\epsilon))}
\end{aligned}
$$

by the joint typicality lemma [9]. Similarly for $j>j_2$

$$
\begin{aligned}
&\mathsf{P}\{(\mathbf{X}_{1j}(m),\mathbf{X}_{2j}(m),\mathbf{Y}_{3j})\in\mathcal{T}_\epsilon^{(n)}|\mathbf{Y}_{2j_2}(m)\neq\mathbf{Y}_{2j_2}(1)\} \\
&\quad\leq 2^{-n(I(X_1,X_2;Y_3)-\delta_1(\epsilon))}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\mathsf{P}(\mathcal{E}_1) &\leq \sum_{m\neq1}\sum_{j_2}\prod_{j=1}^{j_2-1}2^{-n(I(X_1;Y_2,Y_3|X_2)-\delta(\epsilon))} \\
&\quad\cdot\prod_{j=j_2+1}^{b}2^{-n(I(X_1,X_2;Y_3)-\delta_1(\epsilon))} \\
&\leq 2^{nbR}2^{\log(b+1)} \\
&\quad\cdot 2^{-n(b-1)\min\{I(X_1;Y_2,Y_3|X_2),I(X_1,X_2;Y_3\}-\delta_2(\epsilon))}.
\end{aligned}
$$

Therefore, $\mathsf{P}(\mathcal{E}_1)\to0$ as $n\to\infty$ if

$$R<\frac{b-1}{b}(\min\{I(X_1;Y_2,Y_3|X_2),I(X_1,X_2;Y_3)\}-\delta_3(\epsilon)).$$

Finally, letting $b\to\infty$ and $\epsilon\to0$, we have the achievability of any rate

$$R<\min\{I(X_1;Y_2,Y_3|X_2),I(X_1,X_2;Y_3)\}.$$

We now generalize the coding scheme to the general semideterministic network with multiple destination nodes.

Codebook generation: Fix $\prod_{k\in\mathcal{T}}p(x_k)$. For each $j\in[1:b]$, randomly and independently generate 2^{nbR} sequences $\mathbf{x}_{1j}(m)$, $m\in[1:2^{nbR}]$, each according to $\prod_{i=1}^n p(x_{1,(j-1)n+i})$. For $j=1$, $k\in\mathcal{R}$, randomly and independently generate a sequence $\mathbf{x}_{kj}$ according to $\prod_{i=1}^n p(x_{k,(j-1)n+i})$. For each $j\in[2:b]$, $k\in\mathcal{R}$, randomly and independently generate $\mathbf{x}_{kj}(\mathbf{y}_{k1},\ldots,\mathbf{y}_{k,j-1})$ sequences for each $(\mathbf{y}_{k1},\ldots,\mathbf{y}_{k,j-1})\in\mathcal{Y}_k^{(j-1)n}$ according to $\prod_{i=1}^n p(x_{k,(j-1)n+i})$. This defines the codebook $\mathcal{C}_j=\{\mathbf{x}_{1j}(m),\mathbf{x}_{kj}(\mathbf{y}_k^{j-1}):m\in[1:2^{nbR}],\mathbf{y}_k^{j-1}\in\mathcal{Y}_k^{(j-1)n},k\in\mathcal{R}\cup\mathcal{D}\}$ for $j\in[1:b]$.

The codebooks are revealed to both encoders and the decoder before any transmission takes place.

Encoding: The use of codebooks is synchronized, i.e., at any given block index j, all nodes use the codebook indexed with j. Let $\mathbf{y}_{kj}$ be the corresponding output sequence at node $k\in\mathcal{R}\cup\mathcal{D}$ when the network uses the j-th codebook. To send a message $m\in[1:2^{nbR}]$, the source encoder transmits $\mathbf{x}_{1j}(m)$ at each block $j\in[1:b]$. For $j=1$, each relay node $k\in\mathcal{R}$ the relay encoder transmits $\mathbf{x}_{k1}$. For $j\in[2:b]$, each relay node $k\in\mathcal{R}$ the relay encoder transmits $\mathbf{x}_{kj}(\mathbf{y}_k^{j-1})$.

Decoding: For $\mathcal{S}\subseteq\mathcal{R}$, define recursively

$$
\begin{aligned}
&\mathbf{x}_j(\mathcal{S}(m)):=\{\mathbf{x}_{kj}(\mathbf{y}_k^{j-1}(m)):k\in\mathcal{S}\}, \\
&\mathbf{y}_j(\mathcal{S}(m)):=\{\mathbf{y}_{kj}(\mathbf{x}_{1j}(m),\ldots,\mathbf{x}_{N-1,j}(m),\mathbf{y}_{dj}):k\in\mathcal{S}\}.
\end{aligned}
$$

Decoder d finds a unique message $m \in [1:2^{nbR}]$ such that

$$\{(\mathbf{x}_{1j}(m), \mathbf{x}_j(\mathcal{R}(m)), \mathbf{y}_j(\mathcal{R}(m)), \mathbf{y}_{dj}) \in \mathcal{T}_\epsilon^{(n)}\} \qquad (6)$$

for all $j \in [1:b]$.

Analysis of the probability of error: Let

$$\mathcal{A}_{jd}(m) := \{(\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{R}(m)), \mathbf{Y}_j(\mathcal{R}(m)), \mathbf{Y}_{dj}) \in \mathcal{T}_\epsilon^{(n)}\}.$$

Assuming $m = 1$ was sent, we have two sources of error for decoder $d \in \mathcal{D}$:

$$\mathcal{E}_{0d} := \cup_{j=1}^b \mathcal{A}_{jd}^c(1) \text{ and } \mathcal{E}_{1d} := \cup_{m\neq 1} \cap_{j=1}^b \mathcal{A}_{jd}(m).$$

The probability $\mathsf{P}(\mathcal{E}_{0d}) \to 0$ as $n \to \infty$ by the law of large numbers. Let $\mathbf{j}_\mathcal{R} := (j_2, \ldots, j_{N-1})$ where $j_k \in [1:b+1]$, $k \in \mathcal{R}$. Define the event

$$\mathcal{B}(\mathbf{j}_\mathcal{R}) := \{j_k \text{ is the smallest index } j \text{ such that } \mathbf{Y}_{kj}(m) \neq \mathbf{Y}_{kj}(1), k \in \mathcal{R}\}.$$

Thus, when the event $\mathcal{B}(\mathbf{j}_\mathcal{R})$ occurs, an element j_k in $\mathbf{j}_\mathcal{R}$ represents the smallest block index in which relay node k observes a different output under message index m and 1. If $\mathbf{Y}_{kj}(m) = \mathbf{Y}_{kj}(1)$ for all $j \in [1:b]$, $j_k = b+1$.

The probability of $\mathcal{E}_{1d}$ is given by

$$\mathsf{P}(\mathcal{E}_{1d}) = \mathsf{P}\left(\bigcup_{m\neq 1}\bigcap_{j=1}^b \mathcal{A}_{jd}(m)\right) \leq \sum_{m\neq 1} \mathsf{P}\left(\bigcap_{j=1}^b \mathcal{A}_{jd}(m)\right).$$

Then,

$$\begin{aligned}
\mathsf{P}\left(\bigcap_{j=1}^b \mathcal{A}_{jd}(m)\right) &= \sum_{\mathbf{j}_\mathcal{R}} \mathsf{P}\left(\bigcap_{j=1}^b \mathcal{A}_{jd}(m) \cap \mathcal{B}(\mathbf{j}_\mathcal{R})\right) \\
&\stackrel{(e)}{=} \sum_{\mathbf{j}_\mathcal{R}} \mathsf{P}\Bigg(\bigcap_{j=1}^b \{(\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{S}_j(m)), \mathbf{X}_j(\mathcal{S}_j^c(1)), \\
&\qquad \mathbf{Y}_j(\mathcal{S}_j(m)), \mathbf{Y}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_{dj}) \in \mathcal{T}_\epsilon^{(n)}\} \cap \mathcal{B}(\mathbf{j}_\mathcal{R})\Bigg) \\
&\stackrel{(f)}{\leq} \sum_{\mathbf{j}_\mathcal{R}} \mathsf{P}\Bigg(\bigcap_{j=1}^b \{(\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{S}_j(m)), \\
&\qquad \mathbf{X}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_{dj}\} \cap \mathcal{B}(\mathbf{j}_\mathcal{R})\Bigg) \\
&\stackrel{(g)}{\leq} \sum_{\mathbf{j}_\mathcal{R}} \mathsf{P}\Bigg(\bigcap_{j=1}^b \{(\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{S}_j(m)), \\
&\qquad \mathbf{X}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_{dj}) \in \mathcal{T}_\epsilon^{(n)}\} \\
&\qquad \cap \{\mathbf{Y}_{k,j_k}(m) \neq \mathbf{Y}_{k,j_k}(1), k \in \mathcal{R}\}\Bigg) \\
&\leq \sum_{\mathbf{j}_\mathcal{R}} \mathsf{P}\Bigg(\bigcap_{j=1}^b \{(\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{S}_j(m)), \\
&\qquad \mathbf{X}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_{dj}) \in \mathcal{T}_\epsilon^{(n)}\} \\
&\qquad \Big| \mathbf{Y}_{k,j_k}(m) \neq \mathbf{Y}_{k,j_k}(1), k \in \mathcal{R}\Bigg) \\
&\stackrel{(h)}{=} \sum_{\mathbf{j}_\mathcal{R}} \prod_{j=1}^b \mathsf{P}((\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{S}_j(m)), \\
&\qquad \mathbf{X}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_{dj}) \in \mathcal{T}_\epsilon^{(n)} \\
&\qquad |\mathbf{Y}_{k,j_k}(m) \neq \mathbf{Y}_{k,j_k}(1), k \in \mathcal{R}) \\
&\leq \sum_{\mathbf{j}_\mathcal{R}} \prod_{j\notin\mathbf{j}_\mathcal{R}} \mathsf{P}((\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{S}_j(m)), \\
&\qquad \mathbf{X}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_{dj}) \in \mathcal{T}_\epsilon^{(n)} \\
&\qquad |\mathbf{Y}_{k,j_k}(m) \neq \mathbf{Y}_{k,j_k}(1), k \in \mathcal{R})
\end{aligned}$$

where $\mathcal{S}_j := \{k \in \mathcal{R} : j \geq j_k\}$, $\mathcal{S}_j^c := \{k \in \mathcal{R} : j < j_k\}$,

(e) follows since

$$\mathcal{B}(\mathbf{j}_\mathcal{R}) \subseteq \{\mathbf{X}_j(\mathcal{S}_j^c(m)) = \mathbf{X}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_j(\mathcal{S}_j^c(m)) = \mathbf{Y}_j(\mathcal{S}_j^c(1)), j \in [1:b]\},$$

(f) follows by [16, Lemma 21], i.e.,

$$\begin{aligned}
&\{(\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{S}_j(m)), \mathbf{X}_j(\mathcal{S}_j^c(1)), \\
&\qquad \mathbf{Y}_j(\mathcal{S}_j(m)), \mathbf{Y}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_{dj}) \in \mathcal{T}_\epsilon^{(n)}\} \\
&\subseteq \{(\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{S}_j(m)), \mathbf{X}_j(\mathcal{S}_j^c(1)), \\
&\qquad \mathbf{Y}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_{dj}) \in \mathcal{T}_\epsilon^{(n)}\},
\end{aligned}$$

(g) follows since $\mathcal{B}(\mathbf{j}_\mathcal{R}) \subseteq \{\mathbf{Y}_{k,j_k}(m) \neq \mathbf{Y}_{k,j_k}(1), k \in \mathcal{R}\}$, and

(h) follows since the events are mutually conditionally independent given $\{\mathbf{Y}_{k,j_k}(m) \neq \mathbf{Y}_{k,j_k}(1), k \in \mathcal{R}\}$.

Furthermore, for $j \notin \mathbf{j}_\mathcal{R}$

$$\begin{aligned}
&\mathsf{P}\{(\mathbf{X}_{1j}(m), \mathbf{X}_j(\mathcal{S}_j(m)), \mathbf{X}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_j(\mathcal{S}_j^c(1)), \mathbf{Y}_{dj}) \in \mathcal{T}_\epsilon^{(n)} \\
&\qquad |\mathbf{Y}_{k,j_k}(m) \neq \mathbf{Y}_{k,j_k}(1), k \in \mathcal{R}\} \\
&\leq 2^{-n(I(X_{1j}, X(\mathcal{S}_j); Y(\mathcal{S}_j^c), Y_d | X(\mathcal{S}_j^c)) - \delta_4(\epsilon))} \\
&\leq 2^{-n(\min_{\mathcal{S}:1\in\mathcal{S}, d\in\mathcal{S}^c} I(X(\mathcal{S}); Y(\mathcal{S}^c)|X(\mathcal{S}^c)) - \delta_4(\epsilon))}
\end{aligned}$$

by the joint typicality lemma [9] where, $\mathcal{S} := \{1\} \cup \mathcal{S}_j$ is a cut between 1 and $d \in \mathcal{D}$. Thus,

$$\begin{aligned}
\mathsf{P}(\mathcal{E}_{1d}) &\leq \sum_{m\neq 1} \sum_{\mathbf{j}_\mathcal{R}} \prod_{j\notin\mathbf{j}_\mathcal{R}} 2^{-n(\min_\mathcal{S} I(X(\mathcal{S}); Y(\mathcal{S}^c)|X(\mathcal{S}^c)) - \delta_4(\epsilon))} \\
&\leq 2^{nbR} 2^{(N-2)\log(b+1)} \\
&\qquad \cdot 2^{-n(b-N+2)(\min_\mathcal{S} I(X(\mathcal{S}); Y(\mathcal{S}^c)|X(\mathcal{S}^c)) - \delta_4(\epsilon))}.
\end{aligned}$$

Therefore, $\mathsf{P}(\mathcal{E}_{1d}) \to 0$ as $n \to \infty$ if

$$R < \frac{b-N+2}{b} \cdot \left(\min_{\substack{\mathcal{S}\subseteq\mathcal{T}\cup\mathcal{D} \\ 1\in\mathcal{S}, d\in\mathcal{S}^c}} I(X(\mathcal{S}); Y(\mathcal{S}^c)|X(\mathcal{S}^c)) - \delta_5(\epsilon)\right). \qquad (7)$$

The probability of decoding error goes to zero for each destination node $d \in \mathcal{D}$ as $n \to \infty$, provided that the rate condition satisfies (7). By the union of events bound, the probability of error goes to zero as $n \to 0$. Finally, by letting $b \to \infty$ and $\epsilon \to 0$, we have the achievability of any rate

$$R < \min_{\substack{\mathcal{S}\subseteq\mathcal{T}\cup\mathcal{D} \\ :1\in\mathcal{S}, d\in\mathcal{S}^c}} I(X(\mathcal{S}); Y(\mathcal{S}^c)|X(\mathcal{S}^c)) \qquad (8)$$

which completes the proof of Theorem 1.

Acknowledgment

The work of Sung Hoon Lim and Sae-Young Chung is partially supported by the MKE, Korea, under the ITRC support program supervised by the IITA (IITA-2009-C1090-0902-0005), and the work of Young-Han Kim is partially supported by the National Science Foundation CAREER award CCF-0747111.

References

[1] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," in *Proc. Information Theory Workshop*, Cairo, Egypt, Jan. 2010.

[2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[3] M. R. Aref, "Information flow in relay networks," Ph.D. Thesis, Stanford University, Stanford, CA, Oct. 1980.

[4] N. Ratnakar and G. Kramer, "The multicast capacity of deterministic relay networks with no interference," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2425–2432, 2006.

[5] S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow," 2007, submitted to *IEEE Trans. Inf. Theory,* 2007. [Online]. Available: http://arxiv.org/abs/0710.3781/

[6] A. F. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 789–804, 2006.

[7] B. Smith and S. Vishwanath, "Unicast transmission over multiple access erasure networks: Capacity and duality," in *Proc. Information Theory Workshop*, Tahoe City, California, May 2007, pp. 331–336.

[8] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," 2009.

[9] A. El Gamal and Y.-H. Kim, *Lecture Notes on Network Information Theory*, Stanford University and UCSD, 2009.

[10] A. El Gamal and M. R. Aref, "The capacity of the semideterministic relay channel," *IEEE Trans. Inf. Theory*, vol. 28, no. 3, p. 536, 1982.

[11] T. M. Cover and Y.-H. Kim, "Capacity of a class of deterministic relay channels," in *Proc. IEEE International Symposium on Information Theory*, Nice, France, June 2007, pp. 591–595.

[12] A. El Gamal, "On information flow in relay networks," in *Proc. IEEE National Telecom Conference*, Nov. 1981, vol. 2, pp. D4.1.1–D4.1.4.

[13] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[14] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[15] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.

[16] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, 2001.