Deterministic Relay Networks with State Information

Sung Hoon Lim  
School of EECS  
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  
School of EECS  
KAIST  
Daejeon, Korea  
Email: sychung@ee.kaist.ac.kr

Abstract—Motivated by fading channels and erasure channels, the problem of reliable communication over deterministic relay networks is studied, in which relay nodes receive a function of the incoming signals and a random network state. An achievable rate is characterized for the case in which destination nodes have full knowledge of the state information. If the relay nodes receive a linear function of the incoming signals and the state in a finite field, then the achievable rate is shown to be optimal, meeting the cut-set upper bound on the capacity. This result generalizes the work of Avestimehr, Diggavi, and Tse on deterministic networks with state dependency, the work of Dana, Gowalikar, Palanki, Hassibi, and Effros on linear erasure networks with interference, and the work of Smith and Vishwanath on linear erasure networks with broadcast.

I. INTRODUCTION

In their celebrated paper [1] that opened the field of network coding, Ahlswede et al. found the multicast capacity of wireline networks. For wireless networks, however, there are some new challenges for reliable communication compared to the wireline network. Among them are broadcast and interference, and there has been some work that deals with these two features. In [2], the multicast capacity was shown for networks that have deterministic channels with broadcast, but without interference at the receivers. Deterministic networks were further studied in [3] to incorporate interference at the receiving nodes, where the capacity for linear finite field networks was found. These rather simple models were shown to give good insights in solving real-world network problems. For example in [4], Avestimehr et al. were able to approximately characterize the capacity of Gaussian relay networks within some constant gap using a similar approach used for deterministic networks. Although previous models consider broadcast and interference, they did not explicitly consider another important feature in wireless communications. The wireless medium in real-world communications suffer fading, which in turn cause severe degradation of the transmitted signal. Although the deterministic model can be a good abstraction in understanding broadcast and interference, it does not fully capture the effect of fading in wireless networks. In this sense, the erasure network in which transmitted symbols get erased at random provides a simple model that captures the fading characteristics. In [5], Dana et al. considered the erasure networks with broadcast and no interference, where the erasures are at the traversing edges. Smith and Vishwanath [6] considered an erasure network without broadcast, where the interference is modeled as a linear finite field sum of incoming signals that are not erased. In both [5] and [6], if the destination node has perfect knowledge of the state information, they showed that the capacity is given by the cut-set bound.

In this paper, we consider a deterministic network in which the observation at each node is a function of the incoming signals and a random state. The channel state affecting the relay and destination nodes is assumed to be perfectly known at the destinations. We give an achievable rate for this class of networks, and show that the associated coding scheme achieves the capacity for the case in which the relay and destination nodes receive a linear function of the incoming signals and the state over a finite field. This result generalizes the work of Dana et al. and the work of Smith and Vishwanath on linear erasure networks to handle both interference and broadcast. As for deterministic networks, our result generalizes the work of Avestimehr, Diggavi, and Tse on the deterministic networks to deterministic state-dependent networks.

II. PROBLEM STATEMENT AND PRELIMINARIES

In the following we will give useful definitions for later use. Upper case letters denote random variables (e.g., X, Y, S) and lower case letters represent scalars (e.g., x, y, s). Calligraphic letters (e.g., A) denote sets and the cardinality of the set is denoted by |A|. Subscripts are used to specify node and time indices. For example, X_u and X_u,t denotes the signal sent at node u and the signal sent at node u at time t, respectively. To represent a sequence of random variables we use the notation X^n_u = X_u,1,...,X_u,n. We will frequently use random variables subscripted by sets to denote the set of random variables indexed with elements in the set. For example, X_A = {X_a : a ∈ A} and X^n_A = {X^n_a : a ∈ A}.

We consider a network G = (V,E) where V and E are the set of nodes and directed edges, respectively. Without loss of generality, we let V = {1,...,|V|} and index the source node with 1. We use D and R = V−(1∪D) to denote the set of destination nodes and relay nodes respectively. The network has one channel input X_u ∈ X_u associated with each node u ∈ V, where X_u is the alphabet of X_u. This incorporates the
broadcast nature of the network. Each node \( v \in \mathcal{V} \) observes

\[
Y_v = f_v (X_{\mathcal{N}_v}, S),
\]

where the input neighbors \( \mathcal{N}_v \) of \( v \) is defined as \( \mathcal{N}_v = \{ u : (u, v) \in \mathcal{E} \} \). The random variable \( S \) is a random state affecting nodes, which is independent of the source message. The state sequence is memoryless and stationary with \( p(s^T) = \prod_{t=1}^T p(s_t) \). We assume that each destination \( d \in \mathcal{D} \) has side information of the state sequence. The source node wishes to send a common message \( m \in [2^nR] \triangleq \{ 1, \ldots, 2^nR \} \) to all destination nodes.

A \((2^nR, n)\) code consists of a source encoding function \( \phi_1 \), relay encoding functions \( \phi_{u,i}, v \in \mathcal{V} - \{ \{ 1 \} \cup \mathcal{D} \}, i \in \{ 1, \ldots, n \} \), and decoding functions \( \psi_d, d \in \mathcal{D} \), where

\[
\begin{align*}
\phi_1 & : [2^nR] \rightarrow \mathcal{X}_1^n, \\
\phi_{u,i} & : \mathcal{Y}_u^{i-1} \rightarrow \mathcal{X}_v, i \in \{ 1, \ldots, n \}, v \in \mathcal{R}, \\
\psi_d & : \mathcal{Y}_d^n \times S^n \rightarrow [2^nR], d \in \mathcal{D}
\end{align*}
\]

where \( M \) is uniformly distributed over \([2^nR]\). The probability of error is defined by

\[
P^e = \Pr\{ \psi_d(Y^d, S^n) \neq M \text{ for some } d \in \mathcal{D} \}.
\]

A rate \( R \) is said to be achievable if there exist a sequence of \((2^nR, n)\) codes with \( P^e \rightarrow 0 \) as \( n \rightarrow \infty \).

For each \( d \in \mathcal{D} \), a cut \( \mathcal{U}_d \subset \mathcal{V} \) is a subset of nodes such that \( 1 \in \mathcal{U}_d \) and \( d \in \mathcal{U}_d \). We will omit the destination index when it is clear from the context. We define a boundary of a cut as \( \partial(\mathcal{U}) = \{ u : (u, v) \in \mathcal{E}, u \in \mathcal{U}, v \in \mathcal{V} \} \) and the boundary of a complement of a cut as \( \partial(\mathcal{U}^c) = \{ v : (u, v) \in \mathcal{E}, u \in \mathcal{U}, v \in \mathcal{V} \} \).

We say that a node \( v \) is in layer \( l \) if all directed paths from the source to \( v \) has \( l \) hops. Let \( L \) be the longest distance from the source node to any node. We say that a network is layered with \( L \) layers if every node in \( \mathcal{V} \) belong to some layer \( l \in \{ 0, \ldots, L \} \). The set of nodes in layer \( l \) is denoted by \( \mathcal{V}_l \).

Without loss of generality we will assume that \( \mathcal{V}_0 = \{ 1 \} \).

For a random variable \( X \sim p(x) \), the set \( T_{e}^{(a)} \) of \( e \)-typical \( n \)-sequences \( x^n \) is defined as

\[
T_{e}^{(a)} \triangleq \{ x^n : |\pi(a|x^n) - p(a)| \leq \delta \cdot p(a), \forall a \in \mathcal{X} \}
\]

where \( \pi(a|x^n) \) is the relative frequency of the symbol \( a \) in the sequence \( x^n \).

### III. MAIN RESULT

#### A. General state dependent networks

Given a class of relay networks as defined in [1], the multicast capacity \( C \) is upper bounded by

\[
C \leq \max_{p(x)} \min_{d \in \mathcal{D}} \min_{\mathcal{U}_d} H(Y_{d|L} | X_{d|L}, S).
\]

The upper bound is from the cut-set bound [8] Theorem 15.10.1 by treating the state information as additional outputs to the destinations, and using the fact that the state sequences are independent of the message, the memoryless property of the channel, and the deterministic nature of the channel given \( S \).

**Remark 1:** The cut-set bound is given by [2] whether we assume that the relay nodes have state information or not, as long as the state information at the relays are causal (i.e., \( x_{v,i} = \phi_{v,i}(y_{v,i}^{i-1}, s^i) \)) and destination nodes have the state information.

As our main result we state the following theorem.

**Theorem 1:** For the multicast relay network \( G = (\mathcal{V}, \mathcal{E}) \) in [1], if all destination nodes in \( \mathcal{D} \) have side information of the state, then the capacity \( C \) of the network is lower bounded by

\[
C \geq \max_{\mathcal{U}_d} \min_{d \in \mathcal{D}} \min_{\mathcal{U}_d} H(Y_{d|L} | X_{d|L}, S).
\]

The proof of this theorem will be given in Sections [IV] and [V].

**Remark 2:** Theorem [1] includes the special case of unicast networks if \( |\mathcal{D}| = 1 \).

**Example 1 ([5, Theorem 1]):** Consider a network with output symbols \( Y_v = \{ Y_{u,v} : u \in \mathcal{N}_v \} \), where \( Y_{u,v} \) is the observation at node \( v \) through the edge \((u, v)\). Thus, the receiving nodes receives a separate output for each link connected to the node, i.e., has no interference. Let the output random variables take values from \( \mathcal{Y} = \mathcal{X} \cup \{ e \} \), where the symbol \( e \) is the erasure symbol. Each channel output \( Y_{u,v} \) is given by the transmitted signal \( X_u \) with probability \( 1 - \epsilon_{u,v} \) or an erasure symbol \( e \) with probability \( \epsilon_{u,v} \). Let \( S_{u,v,i} \) be a random variable indicating erasure occurrence across channel \((u, v) \) at time \( i \). If an erasure occurs on link \((u, v) \) at time \( i \), the value of \( S_{u,v,i} \) will be one, otherwise zero. Let \( S^n = \{ S^n_{u,v} : u \in \mathcal{N}_v \} \). If the destination nodes have the \( S^n \) sequence as side information, this channel falls into the channel model described in Section [III] since the output at each relay is a function of the incoming signals and \( S^n \). It can be shown that the cut-set bound is achieved by the uniform product distribution. Hence, the capacity of this channel is given by [3] with equality.

#### B. Linear finite field fading networks

Consider a finite field \((GF(q))\) network in which each node \( v \in \mathcal{V} \) observes

\[
Y_v = \sum_{u \in \mathcal{N}_v} S_{u,v} X_u
\]
where $Y_v, X_u, u \in \mathcal{N}_o, S_{u,v}, u \in \mathcal{N}_o$, are in $GF(q)$. If we assume that $S_{u,v}, \forall (u,v) \in \mathcal{E}$ is known at the destination nodes, this channel falls into the class of channels in Section 11.

Let $Y_{U^c}$ and $X_{U^c}$ be vectors of observations in $\bar{\partial}(\mathcal{U}^c)$ and input signals in $\bar{\partial}(\mathcal{U})$, respectively. These are of observations and input signals of nodes that have an edge passing through the cut. We define a transfer matrix of an arbitrary cut $U$ as $G_{U^c}$ such that it satisfies $Y_{U^c} = G_{U^c}X_{U^c}$.

Thus, the random matrix $G_{U^c}$ consists of zeros when there is no connection between the nodes and $S_{u,v}$ if $u \in \bar{\partial}(\mathcal{U})$ and $v \in \bar{\partial}(\mathcal{U}^c)$. The column index represents the sending node index in $\bar{\partial}(\mathcal{U})$ and row index represents the receiving node index in $\bar{\partial}(\mathcal{U}^c)$. For the example in Figure 1 we have the expression

$$
\begin{bmatrix}
Y_3 \\
Y_2
\end{bmatrix}
= 
\begin{bmatrix}
S_{1,3} & S_{2,3} & 0 \\
S_{2,3} & S_{2,4} & 0
\end{bmatrix}
\begin{bmatrix}
X_1 \\
X_2
\end{bmatrix}
$$

for the cut $\mathcal{U} = \{1, 2\}$.

**Theorem 2:** The multicast capacity of the linear finite field fading network (6) is

$$
C = \min_{d \in D} \min_{\mathcal{U}^c} E[\text{rank}(G_{U^c})] \log q.
$$

**Proof:** Proof is omitted due to space limitations.  

**Remark 3:** For the special case of $S \in \{0, 1\}$, Theorem 2 includes the capacity result for linear finite field erasure networks with broadcast and interference.  

IV. PROOF OF THEOREM 1 FOR LAYERED NETWORKS

We begin by showing the achievability of Theorem 1 for layered networks with $D = \{d\}$. The multicast network is a simple extension of the single destination network and will be treated later.

We use a block Markov encoding scheme in which we divide the message $m$ into $K$ parts $m_k, k \in \{1, \ldots, K\}$. We code in $K + L - 1$ blocks of length $n$. Message $m_k$ takes values from $[2^{nR}]$ for all $k$ and the overall rate is given by $R_K = \log(K+L-1)$. We will use two types of indexing for the inputs, outputs, and state. We will use $s^n(j)$ to denote the state sequence when message $m_j$ is being sent at the source node. For the set of observations and input sequences at layer $l$ carrying message $m_j$, we will use the notation

$$
y^n_v(m_j) \triangleq \{y^n_v(x^n_{X_e}(m_j), s^n(j+l)) : v \in \mathcal{V}_l\}
$$

and

$$
x^n_v(m_j) \triangleq \{x^n_v(y^n_{X_e}(m_j)) : v \in \mathcal{V}_l\},
$$

respectively. For example, (5) denotes the set of observation sequences of the nodes in layer $l$ when $m_j$ is received. Due to the layered structure of the network and the coding strategy, which will be explained in the following, the observation sequences corresponding to the $j$th message at layer $l$ are functions of $s^n(j+l)$. This will be explained in more detail in the following.

**Codebook generation:** Fix $p(x_u)$ for all $u \in \mathcal{V} - \{d\}$. Randomly and independently generate $2^{nR}$ sequences $x^n_v(m)$, $m \in [2^{nR}]$, each according to $\prod_{i=1}^{L+1} p(x_{1,i})$. For each $u \in \mathcal{V} - \{1\}$, randomly and independently generate $x^n_u(y^n_v)$ sequences for each $y^n_v \in Y^n_v$, according to $\prod_{i=1}^{L+1} p(x_{1,i}).$

**Encoding:** To send message $m_j$, $j \in \{1, \ldots, K\}$, the encoder sends $x^n_v(m_j)$, while at each layer $l$, node $v \in \mathcal{V}_l$ sends $x^n_v(y^n_v(m_j-l))$.

**Decoding:** When the destination receives $y^n_v(m_j)$, it also has $\{s^n(1), \ldots, s^n(j+L)\}$ from previous observations. Assuming the previous blocks were decoded with arbitrarily small error, the receiver declares that a message was sent if it is a unique index $m_j \in [2^{nR}]$ such that

$$
\bigcap_{l=0}^{L-1} \left\{ (x^n_{\mathcal{V}_l}(m_j), y^n_{\mathcal{V}_{l+1}}(m_j), s^n(j+l)) \in T_{E}^{(n)} \right\};
$$

otherwise an error is declared.

From the encoding we can see that there is a block delay at layer $l$, $l \in \{1, \ldots, L\}$. When the source sends message $m_j$, the relays in layer 1 send $x^n_{\mathcal{V}_2}(m_j)$, where the relays in layer 2 send $x^n_{\mathcal{V}_3}(m_j-2)$ and so on. Accordingly, when the source sends the $j$th block, received observation sequence of node $v \in \mathcal{V}_l$ is a function of $x^n_{\mathcal{V}_l}(m_j-l-1)$ and $s^n(j)$, which gives (6). Table 1 shows the coding strategy for a simple diamond network given in Figure 2.

The decoding is a typicality check over an intersection of disjoint sets. Recall that from (5) and (6), as message $m_j$ traverses through the network, the message is being affected by a different state at each layer. Therefore, we require that all inputs and outputs of that layer and a state (corresponding to the specific block time) are uniquely jointly typical.

Before dealing with arbitrarily large networks, we will first give a proof for a simple diamond network. Consider a diamond network depicted in Fig. 2 at the top of the next page. The relay nodes $\{a, b\}$ in layer 1 receives $Y_a, Y_b$ which are deterministic functions of $X_1$ and $S$. The destination node in layer 2 observes $Y_d$, which is a deterministic function of $X_a, X_b$, and $S$. Without loss of generality, we will assume that $m_j = 1$ was sent, and show the decoding and probability of error analysis for the $j$th block. We will omit the message index for simplicity. There are two types of error events:

$$
E_0 \triangleq (A_1^m \cap A_2^m)^c \text{ and } E_1 \triangleq \bigcup_{m \neq 1} (A_1^m \cap A_2^m)
$$

where

$$
A_1^m \triangleq \left\{ (X^n_1(m), Y^n_a(m), Y^n_d(m), S^n(j)) \in T_{E}^{(n)} \right\},
$$

and

$$
A_2^m \triangleq \left\{ (X^n_1(m), X^n_b(m), Y^n_d(1), S^n(j+1)) \in T_{E}^{(n)} \right\}.
$$

For the first error event, we have $P(E_0) \rightarrow 0$ as $n \rightarrow \infty$ by the law of large numbers. We will decompose $E_1$ into four
disjoint events. Let
\[
B_Q^m \triangleq \{ Y_Q^n(m) \neq Y_Q^n(1), Y_{Q^c}^n(m) = Y_{Q^c}^n(1) \}
\]
where \( Q \subseteq \{a, b\} \) and \( Q^c = \{a, b\} - Q \). We have four such events since \( \{a, b\} \) has four subsets. Then the probability of \( E_1 \) is given by
\[
P(E_1) = P \left( \bigcup_{m \neq 1} (A_1^m \cap A_2^m) \right) \leq \sum_{m \neq 1} P \{ A_1^m \cap A_2^m \} \tag{7}
\]
and
\[
P(E_1) = \sum_{m \neq 1} \sum_{Q \subseteq \{a, b\}} P \{ A_1^m \cap A_2^m \cap B_Q^m \} \tag{8}
\]
where in (7) we have used the union bound and (8) is from the fact that \( B_Q^m \) are partitions that cover the whole set. Thus, we have decomposed \( E_1 \) into four disjoint events. The event \( A_1^m \cap B_Q^m \) implies
\[
\left\{ (X_1^m(m), Y_1^m(m), Y_1^m(1), S^m(j)) \in T_e^{(n)} \right\} \tag{9}
\]
and \( A_2^m \cap B_Q^m \) implies
\[
\left\{ (X_0^m(m), Y_0^m(1), S^m(j + 1)) \in T_e^{(n)} \right\} \tag{10}
\]
since \( X_0^m(m) = X_0^m(Y_0^m(m)) \). Since (9) and (10) are independent events, we have
\[
P \{ A_1^m \cap A_2^m \cap B_Q^m \} \leq 2^{n(I(X_1, Y_1, Y_2|S) - 3\epsilon) + n(I(X_2, Y_2|X_1, S) - 6\epsilon)}
\]
where in the last step we have used the Markov structure of the network. Similar to the previous steps, we can bound the other events by
\[
P \{ A_1^m \cap A_2^m \cap B_Q^m \} \leq 2^{-n(H(Y_2|X_2, S) - 6\epsilon)} \tag{12}
\]
and
\[
P \{ A_1^m \cap A_2^m \cap B_Q^m \} \leq 2^{-n(H(Y_2|S, X_2) - 6\epsilon)} \tag{13}
\]
and
\[
P \{ A_1^m \cap A_2^m \cap B_Q^m \} \leq 2^{-n(H(Y_2|S) - 3\epsilon)} \tag{14}
\]
Combining (8), (11), (12), (13), and (14), we get \( P(E_1) \to 0 \) as \( n \to \infty \) if
\[
R < \min \left\{ H(Y_a, Y_b, Y_d|S, X_a, X_b) - 6\epsilon, \right. \\
\left. H(Y_d|S) - 3\epsilon, \right. \\
\left. H(Y_b, Y_d|S, X_b) - 6\epsilon, \right. \\
\left. H(Y_a, Y_d|S, X_a) - 6\epsilon \right\} \tag{15}
\]
where \( Q_l = V_l \cap Q \) and \( Q_l^c = V_l - Q_l \). Then,
\[
P \left( \bigcap_{l=0}^{L-1} A_n^l \cap B_n^\ell \right) \leq \prod_{l=0}^{L-1} 2^{-n(H(Y_{Q_l^c}^{Q_l} | X_{Q_l^c}^c, S) - 3\epsilon)}
\]
\[
= \prod_{l=0}^{L-1} 2^{-n(H(Y_{Q_l^c}^{Q_l} | X_{Q_l^c}^c, S) - 3\epsilon)}.
\] (18)

From (17) and (18) we get
\[
P(E_1) \leq \sum_{m \neq 1} \sum_{Q \subseteq R} \prod_{l=0}^{L-1} 2^{-n(H(Y_{Q_l^c}^{Q_l} | X_{Q_l^c}^c, S) - 3\epsilon)}
\]
\[
= \sum_{m \neq Q} \sum_{Q \subseteq R} 2^{-n \sum_{l=0}^{L-1} (H(Y_{Q_l^c}^{Q_l} | X_{Q_l^c}^c, S) - 3\epsilon)}
\]
\[
\leq \sum_{Q \subseteq R} 2^{nR_2} 2^{-n \sum_{l=0}^{L-1} (H(Y_{Q_l^c}^{Q_l} | X_{Q_l^c}^c, S) - 3\epsilon)}
\]
\[
= \sum_{Q \subseteq R} 2^{nR_2} 2^{-n(H(Y_{Q_l^c}^{Q_l} | X_{Q_l^c}^c, S) - \epsilon')}.
\]
where \( \epsilon' = 3L \epsilon \) and \( \mathcal{U}^c = \{ Q^c, d \} \) which gives a cut in the network. Thus, \( P(E_1) \rightarrow 0 \) as \( n \rightarrow \infty \) if
\[
R < \min_{U} H(Y_{\mathcal{U}^c}^{\mathcal{U}^c}, X_{\mathcal{U}^c}, S) - \epsilon',
\]
which proves Theorem 1 for layered networks with a single destination.

**Remark 4:** Consider a semi-deterministic layered network \( \mathcal{G} = (\mathcal{V}, \mathcal{E}) \) where each node \( v \in \mathcal{V} - \{ d \} \) observes \( Y_v = f_v(X_{N_v}, Y_d) \) and the final destination gets \( Y_d \sim p(y_d | x_{N_d}) \), i.e., a stochastic output. Using the coding scheme above we can show that all rates \( \bar{R} \) that satisfies
\[
R < \min_{U} I(X_{\mathcal{U}^c}^{\mathcal{U}^c} ; Y_{\mathcal{U}^c}^{\mathcal{U}^c} | X_{\mathcal{U}^c}^{\mathcal{U}^c})
\]
are achievable for unicast.

For the multicast scenario we declare an error if any of the nodes in \( \mathcal{D} \) makes an error. Using the union bound and the same line of proof as in Section IV for each \( d \in \mathcal{D} \), we can show that the probability of error is arbitrarily small for sufficiently large \( n \) if
\[
R < \max_{I \in \mathcal{V}} \min_{U_d} \min_{p(x_d)} H(Y_{\mathcal{U}_d^c}^{\mathcal{U}_d^c} | X_{\mathcal{U}_d^c}, S).
\]

**V. ARBITRARY NETWORKS**

For extending the layered network result to arbitrary networks we use the same line of proof as done in [3] that unfolds \( \mathcal{G} \) into a time-extended network. We will just give an outline of the proof. For more details on unfolding \( \mathcal{G} \), we refer to [8] due to space limitations. Given an arbitrary network \( \mathcal{G} \), we unfold the original network over \( T \) stages to get a layered network \( \mathcal{G} \). Using the coding scheme for the unfolded layered network, we can achieve
\[
R < \frac{1}{T} \max_{I \in \mathcal{V}} \min_{U} H(Y_{\mathcal{U}^c}^{\mathcal{U}^c}, X_{\mathcal{U}^c}^{\mathcal{U}^c}, S)
\] (19)
duplicate paths of the original network. Using Lemma 6.2 in [3] (by including a state random variable in the conditional entropies) we have the relation
\[
(T + N - 1) \min_{U} H(Y_{\mathcal{U}^c}^{\mathcal{U}^c}, X_{\mathcal{U}^c}^{\mathcal{U}^c}, S) \leq H(Y_{\mathcal{U}^c}^{\mathcal{U}^c}, X_{\mathcal{U}^c}^{\mathcal{U}^c}, S)
\] (20)
where \( N = 2^{\lvert V \rvert - 2} \). We also have for any distribution,
\[
\min_{U} H(Y_{\mathcal{U}^c}^{\mathcal{U}^c}, X_{\mathcal{U}^c}^{\mathcal{U}^c}, S) \leq T \min_{U} H(Y_{\mathcal{U}^c}^{\mathcal{U}^c}, X_{\mathcal{U}^c}^{\mathcal{U}^c}, S),
\] (21)
since the right hand side corresponds to taking the minimum over only steady cuts (subset of all possible cuts). Combining (20) with (21) we have
\[
\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{U} \max_{p(x)} \min_{U} H(Y_{\mathcal{U}^c}^{\mathcal{U}^c}, X_{\mathcal{U}^c}^{\mathcal{U}^c}, S) \leq \max_{U} \sum_{U} \min_{p(x)} H(Y_{\mathcal{U}^c}^{\mathcal{U}^c}, X_{\mathcal{U}^c}^{\mathcal{U}^c}, S).
\]
Finally, with the relations (19) and (21), we can show that rates arbitrary close to the right hand side of (3) are achievable for sufficiently large \( T \).

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

[2] 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, June 2006.

[3] A. Avestimehr, S. N. Diggavi, and D. Tse, “Wireless network information flow,” in Proc. Forty-Fifth Annual Allerton Conf. Commun., Contr. Comput., Monticello, IL, Sept. 2007.

[4] ——, “Approximate capacity of Gaussian relay networks,” in Proc. IEEE Int. Symp. Information Theory, Toronto, Canada, 2008, pp. 474–478.

[5] A. 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, March 2006.

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

[7] A. Orlitsky and J. Roche, “Coding for computing,” IEEE Trans. Inf. Theory, vol. 47, no. 4, pp. 1242–1250, June 2001.

[8] T. M. Cover and J. A. Thomas, Elements of Information Theory, 2nd ed. New York: Wiley, 2006.