A New Achievable Rate and the Capacity of a Class of Semi-Deterministic Relay Networks

Leila Ghabeli, Mohammad Reza Aref
Information Systems and Security Lab
Department of Electrical Engineering
Sharif University of Technology, Tehran, Iran
Email: ghabeli@ee.sharif.edu, aref@sharif.edu

Abstract—In this paper, we propose an information-theoretic constructive scheme based on generalized block Markov encoding strategy for obtaining an achievable rate for multirelay networks. The coding scheme is based on partial decoding scheme where the relay decodes only part of the transmitted message. The proposed achievable rate is then used to establish the capacity of a class of semi-deterministic relay networks composed of semi-deterministic relay channels. It is shown that the proposed rate will include those of previously proposed rates based on decode-and-forward strategy in some special cases.

I. INTRODUCTION

The relay channel, first introduced by Van der Meulen in [1], describes a single-user communication channel where a relay helps a sender-receiver pair in their communication. In [2], Cover and El Gamal proved a converse result for the relay channel, the so-called max-flow min-cut upper bound. Additionally they established, two coding approaches and three achievability results for the discrete-memoryless relay channel. They also presented the capacity of degraded, reversely degraded relay channel and the relay channel with full feedback. Moreover, the capacity of semi-deterministic relay channel was presented in [3].

The last few decades have seen tremendous growth in communication networks. The most popular examples are cellular voice, data networks and satellite communication systems. These and other similar applications have motivated researches to extend Shannon’s information theory to networks. In [4], Aref introduced deterministic relay networks with no interference and determined the unicast capacity of such networks by showing that a protocol that generalizes the usual cut-set bound of networking theory. In [5], the multicast capacity of Aref networks is characterized. They showed that the multicast capacity is given by the aforementioned information-theoretic cut-set bound.

Recent work on decode-and-forward for multiple relays appeared in [6]-[9]. In particular, Gupta and Kumar [6] applied irregular encoding/-successive decoding to multirelay networks in a manner similar to [4]. In their approach, in addition to the usual multihopping, i.e., the relays successively decode the source messages before these arrive at the destinations, the transmitters cooperate and each receiver uses several or all of its past channel output blocks to decode, and not only its most recent one. In [7], [8], Xie and Kumar developed regular encoding/sliding-window decoding for multiple relays, and showed that their scheme achieves better rates than those of [4], [6]. In [9], authors generalizes the compress-and-forward strategy and also gives an achievable rate when the relays use either decode-and-forward or compress-and-forward. Additionally they add partial decoding to the later method when there are two relays. In their scheme, first relay uses decode-and-forward, and second relay uses compress-and-forward. Second relay further partially decodes the signal from first relay before compressing its observation. They made the second relay output statistically independent of the first relay and the transmitter outputs.

In this work a new achievable rate for relay networks is developed based on generalized block Markov encoding. Generalized block Markov encoding is one of the coding scheme in that, the relay partially decode the transmitted message by the sender. It is in contrast with block Markov encoding scheme in which the relays completely decode the transmitted message by the sender. Generalized block Markov encoding was used to establish the capacity of two classes of relay channels, called, semi-deterministic relay channel and orthogonal relay channel [4], [10]. Here, we extend the concept of this encoding scheme to multirelay networks. In the proposed method, all the relays uses partial decoding and no independency is assumed between the outputs of the relays and the transmitter. It is shown that our achievable rate include those proposed in [6]-[9] in some special cases. In the proposed strategy, no direct communication between the sender and the second relay is considered. The scheme is shown to achieve capacity for a certain class of semi-deterministic relay networks composed of semi-deterministic relay channels. This concept can be applied to the networks that consist of both wire-line and wireless links.

The paper is organized as follows: Section II introduces modeling assumptions and notations. In section III, we first review some theorems and corollaries about generalized block Markov encoding scheme, then we drive a new achievable rate for relay networks. In section IV, we introduce a certain class of semi-deterministic relay networks with two relays.
and obtain its capacity. Section V generalizes the results of sections III, IV to multirelay networks. Finally, some concluding remarks are provided in section VI.

II. DEFINITIONS AND PRELIMINARIES

The discrete memoryless relay network is a model for the communication between a source $X_0$ and a sink $Y_0$ via $N$ intermediate nodes called relays. The relays receive signals from the source and other nodes and then transmit their information to help the sink to resolve its uncertainty about the message. To specify the network, we define $2N + 2$ finite sets: $\mathcal{X}_0 \times \mathcal{X}_1 \times \ldots \times \mathcal{X}_N \times \mathcal{Y}_0 \times \mathcal{Y}_1 \times \ldots \times \mathcal{Y}_N$ and a probability transition matrix $p(y_0, y_1, \ldots, y_N|x_0, x_1, \ldots, x_N)$ defined for all $(y_0, y_1, \ldots, y_N, x_0, x_1, \ldots, x_N) \in \mathcal{Y}_0 \times \mathcal{Y}_1 \times \ldots \times \mathcal{Y}_N \times \mathcal{X}_0 \times \mathcal{X}_1 \times \ldots \times \mathcal{X}_N$. In this model $X_0$ is the input to the network, $Y_0$ is the ultimate output, $Y_i$ is the $i$th relay output and $X_i$ is the $i$th relay input.

An $(M, n)$ code for the network consists of a set of integers $\mathcal{W} = \{1, 2, \ldots, M\}$, an encoding function $x_m^n : \mathcal{W} \rightarrow \mathcal{X}_0^n$, a set of relay function $\{f_{ij}\}$ such that

$$x_{ij} = f_{ij}(y_{i1}, y_{i2}, \ldots, y_{ij-1}), \quad 1 \leq j \leq n, \quad 1 \leq i \leq N,$$

i.e., $x_{ij}$ is the $j$th component of $x_m^n = (x_{i1}, \ldots, x_{in})$, and a decoding function $g : x_m^n \rightarrow \mathcal{W}$. For generality, all functions are allowed to be stochastic functions.

Let $y_i^{N-1} = (y_{i1}, y_{i2}, \ldots, y_{ij-1})$. The input $x_{ij}$ is allowed to depend only on the past received signals at the $i$th node, i.e., $(y_{i1}, \ldots, y_{ij-1})$. The network is memoryless in the sense that $(y_0, y_1, \ldots, y_N)$ depends on the past $(x_0^n, x_1^n, \ldots, x_N^n)$ only through the present transmitted symbols $(x_0^n, x_1^n, \ldots, x_N^n)$. Therefore, the joint probability mass function on $\mathcal{W} \times \mathcal{X}_0 \times \mathcal{X}_1 \times \ldots \times \mathcal{X}_N \times \mathcal{Y}_0 \times \mathcal{Y}_1 \times \ldots \times \mathcal{Y}_N$ is given by

$$p(w, x_0^n, x_1^n, \ldots, x_N^n, y_0^n, y_1^n, \ldots, y_N^n) = \prod_{i=1}^n p(x_0|w)p(x_1|y_i^{N-1})p(y_i|y_i^{N-1})p(y_0|w)$$

where $p(w)$ is the probability distribution on the message $w \in \mathcal{W}$. If the message $w \in \mathcal{W}$ is sent, let $\lambda(w) = \Pr\{g(y_0^n) = w|w = w\}$, denote the conditional probability of error. Define the average probability of error of the code, assuming a uniform distribution over the set of all messages $w \in \mathcal{W}$, as $P_e^n = \frac{1}{M} \sum w \lambda(w)$. Let $\lambda_n = \max \lambda(w)$ be the maximal probability of error for the $(M, n)$ code. The rate $R$ of an $(M, n)$ code is defined to be $R = \frac{1}{n} \log M$ bits/transmission. The rate $R$ is said to be achievable by the network if, for any $\epsilon > 0$, and for all $n$ sufficiently large, there exists an $(M, n)$ code with $M \geq 2^{nR}$ such that $P_e^n < \epsilon$. The capacity $C$ of the network is the supremum of the set of achievable rates.

III. GENERALIZED BLOCK MARKOV ENCODING

In [3] generalized block Markov encoding is defined as a special case of Theorem 7 in [2]. In this coding scheme, the relay does not completely decode the transmitted message by the sender. Instead the relay only decodes part of the message transmitted by the sender. A block Markov encoding timeframe is again used in this scheme such that the relay decodes part of the message transmitted in the previous block and cooperates with the sender to transmit the decoded part of the message to the receiver in current block. Recall the following version of [2, Theorem 7].

**Theorem 1** [2, Theorem 7]: For any relay network $(X_0 \times X_1, p(y_0, y_1|x_0, x_1), Y_0 \times Y_1)$ the capacity $C$ is lower bounded by

$$C \geq \sup \min \{I(U; Y_1|X_1V) + I(X_0; Y_0Y_1|X_1U), I(Y_0; Y_1|X_1V) + I(X_0; Y_0Y_1|X_1U)\}$$

(1)

where the supremum is taken over all joint probability mass functions of the form

$$p(u, v, x_0, x_1, y_0, y_1) = p(v) p(u|v) p(x_0|u) p(x_1|v) p(y_0|y_1, x_0) p(y_1|y_1, x_1, u)$$

(2)

subject to the constraint

$$I(\hat{Y}_1; Y_0X_1|U) \leq I(X_0; Y_0|V).$$

(3)

By substituting $\hat{Y}_1 \equiv \emptyset$, $V \equiv X_1$, and $U \equiv (X_1, U)$ in (1)-(3), the following theorem for generalized block Markov encoding is obtained.

**Theorem 2** [3, Theorem]: For any relay network $(X_0 \times X_1, p(y_0, y_1|x_0, x_1), Y_0 \times Y_1)$ the capacity $C$ is lower bounded by

$$C \geq \max \min \{I(X_0X_1; Y_0), I(U; Y_1|X_1) + I(X_0; Y_0|X_1)\}$$

(4)

where the maximum is taken over all joint probability mass functions of the form

$$p(u, x_0, x_1, y_0, y_1) = p(u, x_0, x_1) p(y_0, y_1|x_0, x_1)$$

(5)

such that $U \rightarrow (X_0, X_1) \rightarrow (Y_0, Y_1)$ form a Markov chain.

If we choose the random variable $U = X_0$, it satisfies the Markovity criterion and the result of block Markov coding directly follows as,

$$C \geq \max \min \{I(X_0X_1; Y_0), I(X_0; Y_1|X_1)\}$$

(6)

The above expression introduces the capacity of degraded relay channel as shown in [2].

Moreover, by substituting $U = Y_1$ in (4), the following corollary is obtained.

**Corollary** [3, Corollary]: If $y_1$ is a deterministic function of $x_0$ and $x_1$, then

$$C = \max p(x_0|x_1) \min \{I(X_0X_1; Y_0), H(Y_1|X_1) + I(X_0; Y_0|X_1Y_1)\}$$

(7)

Next, we generalize the relation of Theorem 2 to the two relay networks and prove the main theorem of this paper as follows,

**Theorem 3**: For any relay networks $(X_0 \times X_1 \times X_2, p(y_0, y_1, y_2|x_0, x_1, x_2), Y_0 \times Y_1 \times Y_2)$ the capacity $C$ is
lower bounded by
\[
C \geq \sup_{p(x_0,x_1,x_2,u_1,u_2)} \min\{I(X_0X_1X_2;Y_0),
I(U_2;Y_2|X_2) + I(X_0X_1;Y_0|U_2X_2),
I(U_1;Y_1|X_1X_2U_2) + I(X_0;Y_0|X_1X_2U_2)\}
\]
where the supremum is over all joint probability mass functions
\[p(x_0,x_1,x_2,u_1,u_2)\] on \(U_1 \times U_2 \times X_0 \times X_1 \times X_2\) such that
\[\left(U_1,U_2\right) \rightarrow (X_0,X_1,X_2) \rightarrow (Y_0,Y_1,Y_2)\]
form a Markov chain.

**Proof:** In this encoding scheme, the message of the transmitter is divided into two parts. The first part is decoded by the first relay and the receiver can only make an estimate of it, while the second part is directly decoded by the receiver. Again the message of the first relay is divided into two parts. The first part is decoded by the next relay and the receiver can only make an estimate of it, while the second part is directly decoded by the receiver. The sender and the relays cooperate in next transmission blocks to remove the receiver’s uncertainty about the previous parts of the message.

The encoding and decoding methods differ from those of Cover and El Gamal. For encoding, we make use of regular block Markov superposition encoding and for decoding we make use of backward decoding [11]. We consider \(B\) blocks of transmission, each of \(n\) symbols. A sequence of \(B - 2\) messages, \(w_{0i} \times w_{1i} \times w_{2i} \in [1,2^{nR_0}] \times [1,2^{nR_1}] \times [1,2^{nR_2}],\) \(i = 1,2,\ldots,B - 2,\) will be sent over the channel in \(nB\) transmissions. In each \(n\)-block \(b = 1,2,\ldots,B,\) we shall use the same set of codewords. We consider only the probability of error in each block as the total average probability of error can be upper bounded by the sum of the decoding error probabilities at each step, under the assumption that no error propagation from the previous steps has occurred [12].

**Random Coding:** The random codewords to be used in each block are generated as follows:

a) Choose \(2^{nR_2}\) i.i.d. \(x_{2i}\) each with probability \(p(x_{2i}) = \prod_{i=1}^{n} p(x_{2i})\). Label these as \(x_{2i}(w_{22}), w_{22} \in [1,2^{nR_2}]\).

b) For every \(x_{2i}(w_{22}), \) generate \(2^{nR_2}\) i.i.d. \(u_{2i}\) with probability
\[
p(u_{2i}|x_{2i}(w_{22})) = \prod_{i=1}^{n} p(u_{2i}|x_{2i}(w_{22}))
\]
Label these \(u_{2i}(w_{22}, w_{21}), w_{21} \in [1,2^{nR_2}]\).

c) For each \(u_{2i}(w_{22}, w_{21}), x_{2i}(w_{22})\), generate \(2^{nR_1}\) i.i.d. \(x_{1i}\) each with probability
\[
p(x_{1i}|u_{2i}(w_{22}, w_{21}), x_{2i}(w_{22})) = \prod_{i=1}^{n} p(x_{1i}|u_{2i}(w_{22}, w_{21}), x_{2i}(w_{22}))
\]
Label these \(x_{1i}(w_{22}, w_{21}, w_{11}), w_{11} \in [1,2^{nR_1}]\).

d) For each \(x_{1i}(w_{22}, w_{21}, w_{11}), \) generate \(2^{n(R_1 + R_2)}\) i.i.d. \(u_{1i}\) each with probability
\[
p(u_{1i}|x_{1i}(w_{22}, w_{21}, w_{11}), x_{2i}(w_{22})) = \prod_{i=1}^{n} p(u_{1i}|x_{1i}(w_{22}, w_{21}, w_{11}), x_{2i}(w_{22}))
\]
Label these as \(u_{1i}(w_{22}, w_{21}, w_{11}, w_{12}), (w_{11}, w_{12}) \in [1,2^{nR_1}] \times [1,2^{nR_2}]\).

e) For each \(u_{1i}(w_{22}, w_{21}, w_{11}, w_{12}), \) generate \(2^{nR_0}\) i.i.d. \(x_{0i}\) each with probability
\[
p(x_{0i}|u_{1i}(w_{22}, w_{21}, w_{11}, w_{12}), x_{1i}(w_{22}, w_{21}, w_{11}),
\]
\[u_{2i}(w_{22}, w_{21}), x_{2i}(w_{22})) = \prod_{i=1}^{n} p(x_{0i}|u_{1i}(w_{22}, w_{21}, w_{11}, w_{12}), x_{1i}(w_{22}, w_{21}, w_{11}),
\]
\[u_{2i}(w_{22}, w_{21}), x_{2i}(w_{22}))
\]
Label these as \(x_{0i}(w_{22}, w_{21}, w_{11}, w_{12}, w_{0}), w_{0} \in [1,2^{nR_0}]\).

The indices \(w_{11}\) and \(w_{21}\) represent the indices \(w_{1}\) and \(w_{2}\) of the previous block respectively, while \(w_{22}\) represents the index \(w_{21}\) of the previous block. The transmitter and relay encoders send the following codewords,
\[
x_{0i}(1,1,1,1), x_{1i}(1,1,1), x_{2i}(1,1,1)
\]
in block \(i = 1,\) the following codewords
\[
x_{0i}(1,1,1,1,1,1,1,1,1), x_{1i}(1,1,1,1,1,1), x_{2i}(1,1,1,1,1,1)
\]
in each block \(i = 2,\) the following codewords
\[
x_{0i}(1,1,1,1,1,1,1,1,1,1,1,1,1), x_{1i}(1,1,1,1,1,1,1,1,1), x_{2i}(1,1,1,1,1,1,1,1,1)
\]
in each block \(i = 2,\ldots,B - 2,\) the following codewords
\[
x_{0i}(1,1,1,1,1,1,1,1,1,1,1,1,1), x_{1i}(1,1,1,1,1,1,1,1,1,1,1,1), x_{2i}(1,1,1,1,1,1,1,1,1,1,1,1)
\]
in block \(i = B - 1,\) and the following codewords
\[
x_{0i}(1,1,1,1,1,1,1,1,1,1,1,1,1), x_{1i}(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1), x_{2i}(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)
\]
in block \(i = B.\)

**Decoding at the relays:** Suppose the second relay has decoded accurately \(w_{2,i-1}\) from the previous block. It determines \(\hat{w}_{2,i}\) such that \((\hat{w}_{2,i-1}, \hat{w}_{2,i}), x_{0i}(\hat{w}_{2,i-1}, \hat{w}_{2,i}), y_{0i}(\hat{w}_{2,i-1}, \hat{w}_{2,i}, i))\) are jointly \(\epsilon\)-typical. \(\hat{w}_{2,i} = w_{2,i}\) with high probability if
\[
R_2 < I(U_2;Y_2|X_2)
\]
and \(n\) is sufficiently large.

Suppose the first relay has decoded accurately \((w_{1,i-1}, w_{2,i-1})\) from the previous block. It determines \((\hat{w}_{1,i}, \hat{w}_{2,i})\) such that
\[
\left(x_{0i}(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i}, \hat{w}_{2,i}), x_{1i}(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i}, \hat{w}_{2,i}),
\right.
\]
\[
x_{2i}(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i}, \hat{w}_{2,i})) = \prod_{i=1}^{n} p(x_{0i}(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i}, \hat{w}_{2,i}), x_{1i}(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i}, \hat{w}_{2,i}))
\]
are jointly \(\epsilon\)-typical. \((\hat{w}_{1,i}, \hat{w}_{2,i}) = (w_{1,i}, w_{2,i})\) with high probability if
\[
R_1 + R_2 < I(U_1;Y_1|X_1X_2U_2)
\]

283
and $n$ is sufficiently large.

Decoding at the Receiver:
Assume that $w_{2,i}$ has been decoded accurately. The receiver determines the unique $\hat{w}_{2,i-1}$ such that
\[
(u^n_2(\hat{w}_{2,i-1}, \hat{w}_{2,i}), x^n_2(\hat{w}_{2,i-1}), y^n(i))
\]
are jointly $\epsilon$-typical. $\hat{w}_{2,i-1} = w_{2,i-1}$ with high probability if
\[
R_2 < I(X_2 U_2; Y_0)
\]
and $n$ is sufficiently large.

Assume that $w_{1,i}$ has been decoded accurately. The receiver determines the unique $\hat{w}_{1,i-1}$ such that
\[
(x^n_1(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i-1}, \hat{w}_{2,i}, \hat{w}_{0,i}),
\hat{u}_1(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i-1}, \hat{w}_{2,i}),
\hat{x}_1^n(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i-1}),
\hat{u}_0^n(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i}, \hat{w}_{0,i}),
\hat{x}_0^n(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i}, \hat{w}_{0,i}), y^n(i))
\]
are jointly $\epsilon$-typical. $\hat{w}_{1,i-1} = w_{1,i-1}$ with high probability if
\[
R_1 < I(X_1 U_1; Y_0 X_2 U_2)
\]
and $n$ is sufficiently large.

The receiver determines the unique $\hat{w}_{0,i}$ such that
\[
(x^n_0(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i-1}, \hat{w}_{1,i}, \hat{w}_{2,i}, \hat{w}_{0,i}),
\hat{u}_1(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i-1}, \hat{w}_{2,i}),
\hat{x}_1^n(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i-1}),
\hat{u}_0^n(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i}, \hat{w}_{0,i}),
\hat{x}_0^n(\hat{w}_{2,i-2}, \hat{w}_{2,i-1}, \hat{w}_{1,i}, \hat{w}_{0,i}), y^n(i))
\]
are jointly $\epsilon$-typical. $\hat{w}_{0,i} = w_{0,i}$ with high probability if
\[
R_0 < I(X_0; Y_0 X_1 X_2 U_2)
\]
and $n$ is sufficiently large.

Now, (12), (13), (14) and (9) results the first term of (8). And (10), (13), (14) and (9) results the second term of (8). And (11) and (14) results the third term of (8). □

Remarks:
1) Theorem 2 can be considered as a special case of Theorem 3, by replacing $X_2 = U_2 = \emptyset$ in (8).

2) By putting $U_1 = X_0$ and the assumption that the second relay output is statistically independent of the first relay and the transmitter outputs, the rate of (8) is reduced to
\[
C \geq \sup_{p(x_0,x_1,u_2)p(x_2)} \min \{I(X_0 X_1 X_2; Y_0), I(X_0; Y_0 X_1 X_2 U_2), I(U_2; Y_2 | X_2) + I(X_0 X_1; Y_0 U_2 X_2), H(Y_0 Y_2 | X_2) - H(Y_0 X_0 X_1 X_2), H(Y_0 Y_1 | X_1 X_2) - H(Y_0 X_0 X_1 X_2)\}
\]
which includes the proposed rate [9, Theorem 5], by omitting the compress-and-forward variable in that equation. Recall that in the scheme of [9, Theorem 5], first relay uses decode-and-forward, and second relay uses compress-and-forward. Second relay further partially decodes the signal from first relay before compressing its observation. The second relay output is assumed to be statistically independent of the outputs of the first relay and the transmitter.

3) If we choose the random variable $U_1 = X_0$ and $U_2 = X_1$, they satisfy the Markovity criterion and result the following,
\[
C \geq \sup_{p(x_0,x_1,x_2)} \min \{I(X_0 X_1 X_2; Y_0), I(X_0; Y_0 X_1 X_2) + I(X_1; Y_2 | X_2) + I(X_0; Y_0 | X_1 X_2)\}
\]
Now we compare the above equation with the result of block Markov coding for two relay networks [4], [6]-[8] as follows,
\[
C \geq \sup_{p(x_0,x_1,x_2)} \min \{I(X_0 X_1 X_2; Y_0), I(X_0; Y_0 X_1 X_2) + I(X_0 X_1; Y_2 | X_2)\}
\]
The first two terms of (16) and (17) are the same. About the third terms, depending on the network conditions each of them can be greater or less than the other.

4) In the proposed strategy, no direct communication between the sender and the second relay is considered. In future work we will consider all communication links between each pair of the nodes in the network and present a coding strategy that include the proposed rates in [4], [6]-[9].

IV. A CLASS OF SEMI-DETERMINISTIC RELAY NETWORKS

In this section we introduce a class of semi-deterministic relay network and show that the capacity of these networks is obtained by using the proposed method of previous section, i.e., it coincides with the max flow min cut upper bound.

Fig. 1 shows a class of semi deterministic relay networks with two relays in which $y_1 = h_1(x_0, x_1, x_2)$ and $y_2 = h_2(x_1, x_2)$. In this figure deterministic and nondeterministic links are shown by solid and dash lines respectively. In the following theorem we prove the capacity of the network of fig. 1.

Theorem 4: The capacity of the relay network of fig. 1 having $y_1 = h_1(x_0, x_1, x_2)$ and $y_2 = h_2(x_1, x_2)$ is given by
\[
C = \sup_{p(x_0,x_1,x_2)} \{I(X_0 X_1 X_2; Y_0), H(Y_0 Y_2 | X_2) - H(Y_0 X_0 X_1 X_2), H(Y_0 Y_1 | X_1 X_2) - H(Y_0 X_0 X_1 X_2)\}
\]
Proof: for achievability result, we note that if the source knows the initial symbol $x_{11}$ and $x_{21}$ in a block of $n$ transmission and if the relay encoding functions are deterministic, then at the $i$th transmission, the source $x_0$ can compute $y_{11},...,y_{i-1}, y_{21},...,y_{2i-1}, x_{1i}, x_{2i}$ and $y_{1i} = h_1(x_{0i}, x_{1i}, x_{2i})$. This means that source will know in advance which sequence will be received by the
relays. By replacing $U_k = Y_k$ for $k = 1, 2$ in (8), the relation in (18) is resulted.

The converse part of this theorem follows easily from max-flow min-cut theorem. For the network with two relay this bound reduces to

$$C \leq \sup_{p(x_0, x_1, x_2)} \min \{ I(X_0; X_1 X_2), I(X_0; Y_0 Y_1 Y_2 | X_1 X_2),
I(X_0 X_1; Y_0 Y_2 | X_2), I(X_0 X_2; Y_0 Y_1 | X_1) \}$$

(19)

and the fact that for the above network with $y_1 = h_1(x_0, x_1, x_2)$ and $y_2 = h_2(x_1, x_2)$, deterministic functions, the following are true,

$$I(X_0; Y_0 Y_1 | X_1 X_2) = H(Y_0 Y_1 | X_1 X_2) - H(Y_0 | X_0 X_1 X_2)$$

(20)

$$I(X_0 X_1; Y_0 Y_2 | X_2) = H(Y_0 Y_2 | X_2) - H(Y_0 Y_2 | X_0 X_1 X_2)$$

(21)

$$I(X_0 X_2; Y_0 Y_1 | X_1) = H(Y_0 Y_1 | X_1 X_2) + I(X_2; Y_0 Y_1 | X_1)$$

(22)

From the above equation, it is observed that (22) is greater than (20). Thus (19) reduces to (18). □

V. GENERALIZATION TO MULTI RELAY NETWORK

In this section we generalize Theorems 3,4 to the networks with $N$ relays in Theorems 5,6 respectively.

Theorem 5: For any relay network $(X_0 \times X_1 \times \ldots \times X_N, p(y_0, y_1, \ldots, y_N | x_0, x_1, \ldots, x_N), Y_0 \times Y_1 \times \ldots \times Y_N)$ the capacity $C$ is lower bounded by

$$C \geq \sup_{p(x_0, x_1, \ldots, x_N)} \min \{ I(\{X_k\}_{k=0}^N; Y_0), \min_{1 \leq i \leq N} \{ I(\{U_i; Y_i | X_i X_{k=0}^i \}_{k=i+1}^N) + I(\{X_k\}_{k=0}^i; Y_0 | \{X_k U_k\}_{k=i+1}^N) \} \}$$

(23)

where the supremum is over all joint probability mass functions $p(x_0, x_1, \ldots, x_N, u_1, \ldots, u_N)$ on $X_0 \times \ldots \times X_N \times U_1 \times \ldots \times U_N$ such that $(U_1, \ldots, U_N) \rightarrow (X_0, \ldots, X_N) \rightarrow (Y_0, \ldots, Y_N)$ form a Markov chain.

The proof of the theorem can be obtained via induction, the familiar method that is usually used in the case of generalization of the theorems to the multi relay network, as was done in [13] for degraded Gaussian multirelay channel.

Now, consider the semi-deterministic relay networks with $N$ relays as shown in Fig. 2. In this figure deterministic and nondeterministic links are shown by solid and dash lines respectively. By replacing $U_k = Y_k$ for $k = 1, \ldots, N$, in (23), the capacity of the network of Fig. 2 is resulted. It is expressed in the following theorem. It can be easily proved that this rate coincides with max-flow min-cut bound.

Theorem 6: For a class of semi deterministic relay network $(A_0 \times A_1 \times \ldots \times A_N, p(y_0, y_1, \ldots, y_N | x_0, x_1, \ldots, x_N), Y_0 \times Y_1 \times \ldots \times Y_N)$ having $y_k = h_k(x_{k-1}, \ldots, x_N)$ for $k = 1, \ldots, N$, the capacity $C$ is given by,

$$C = \sup_{p(x_0, x_1, \ldots, x_N)} \min_{1 \leq i \leq N} \{ H(Y_0 | \{X_k\}_{k=0}^i) - H(Y_0 | \{X_k\}_{k=0}^i) \},$$

(24)

VI. CONCLUSION

In this paper, we have derived a new achievable rate for multirelay network. The strategies make use of regular block Markov superposition encoding and Willems backward decoding. We show that capacity-defining achievable rate result for a class of semi deterministic relay networks is a special case of the proposed scheme. We also show that the proposed rate will include that of presented in [4], [6],[9] in some special cases. In the proposed strategy, no direct communication between the sender and the second relay is considered. In future work we will present a coding strategy that considers direct communication between each pair of the nodes and thus includes the previously proposed rates.

REFERENCES