Published in IET Communications
Received on 10th March 2008
Revised on 10th July 2008
doi: 10.1049/iet-com.2008.0141

ISSN 1751-8628

Comprehensive partial decoding approach for two-level relay networks

L. Ghabeli   M.R. Aref

Information Systems and Security Lab (ISSL), Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran
E-mail: ghabeli@ee.sharif.edu

Abstract: Partial decoding scheme is a scheme in which each relay decodes only part of the transmitted message. Obviously, the achievable rates proposed by the partial decoding scheme subsume the achievable rates proposed by the full decoding scheme. The other motivation of using partial decoding scheme is that there are some special classes of relay networks such as semi-deterministic and orthogonal relay networks such that their capacities are obtained via this scheme. The authors propose a comprehensive partial decoding scheme based on regular encoding/sliding window decoding analysis to propose a new achievable rate for two-level relay networks. In contrast with the previously proposed methods, here the authors consider all possible partial decoding states that can occur between the different parts of the messages of the source and the relays in a two-level relay network. In this way, the common and private parts of the message transmitted by the source are defined to be decoded by the appropriate relays. Moreover, in the proof, the authors take advantage of regular encoding/sliding window decoding scheme that has superiorities to regular encoding/backward decoding and irregular encoding/random partitioning, in having less delay and yielding higher rates, respectively.

1 Introduction

The relay channel, first introduced by van der Meulen [1], describes a single-user communication channel where a relay helps a sender–receiver pair in their communication. Cover and El Gamal [2] 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. Partial decoding scheme or generalised block Markov encoding was defined in [3] as a special case of the proposed coding scheme by Cover and El Gamal [2, Theorem 7]; in this scheme, the relay does not completely decode the transmitted message by the sender, instead the relay decodes only part of the message transmitted by the sender. Partial decoding scheme was used to establish the capacity of two classes of relay channels called semi-deterministic relay channel [3, 4] and orthogonal relay channel [5].

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 researchers to extend Shannon's information theory to networks. One of the remarkable work in this area has been applying the proposed encoding schemes by Cover and El Gamal for the relay channel to the multi-relay networks [6–14]. In particular, Gupta and Kumar [6] applied irregular encoding/successive decoding to multi-relay networks in a manner similar to [4]. Xie and Kumar [7, 8] developed regular encoding/sliding-window decoding for multiple relays, and showed that their scheme achieves better rates than that of [4, 6]. Regular encoding/backward decoding was similarly generalised in [9]. The achievable rates of the two regular encoding strategies turn out to be the same. However, the delay of sliding-window decoding is much less than that of backward decoding. Regular encoding/sliding-window decoding is therefore currently the preferred variant of multi-hopping in the sense that it achieves the best rates in the simplest way. In [10], the authors generalise the compress-and-forward strategy and also give 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 decodes the signal partially from first relay before compressing its observation. They made the second relay output statistically independent of the first relay and the transmitter outputs. In [11, 12], parity-forwarding protocol is introduced, and a structured generalisation of decode-and-forward strategies for multiple-relay networks with feed-forward structure based on such protocol is proposed. In their method, each relay chooses a selective set of previous nodes in the network and decodes all messages of those nodes. Parity forwarding was shown to improve previous decode-and-forward strategies, and it achieves the capacity of new forms of degraded multi-relay networks.

There are also some new works done for generalising partial decoding scheme to multi-relay networks [13–15]. The motivation of using partial decoding scheme is that there are some special classes of relay networks such as semi-deterministic and orthogonal relay networks that their capacities are obtained via this scheme. In [13], a generalisation of partial decoding scheme was applied to two-relay networks; in this method, all the relays in the network successively decode only part of the messages transmitted by the previous relay before they arrive at the destination. In this way, using auxiliary random variables that indicate the message parts results in the flexibility in defining some special classes of relay networks that the proposed rate obtain their exact capacities, e.g. it was shown that the capacity of feed-forward semi-deterministic can be obtained by the proposed method. In [14], the results of [13] were generalised to N-relay networks and the capacity of feed-forward semi-deterministic and orthogonal relay networks with N relays was obtained. In [15], a mixed strategy for multi-relay networks was derived based on using a combined approach of partial decoding and the ideas of successive refinement with different side information at the receivers.

In this work, we propose a comprehensive generalisation of partial decoding scheme for two-level relay network and obtain a better achievable rate. In the proof, we use the method of regular encoding/sliding window decoding that is presented for full decoding strategy in [7, 8]. In contrast with the previously proposed partial decoding approach [13–15], in the proposed method in this paper, we consider all possible partial decoding states that can occur between different parts of the transmitted message by the source and the relays in a two-relay network, e.g. (1) second relay can decode part of the transmitted message by the source in addition to the part of the transmitted message by the first relay, (2) there are individual parts of the transmitted message by the source that is only allowed to be decoded by the related relay and (3) there is common part of the transmitted message by the source that can be decoded by both relays. Another preference of our method is taking advantage of regular encoding/sliding window decoding strategy [7, 8], which potentially yields higher rates for relay networks in comparison with random partitioning, and it has less delay in comparison with regular encoding/backward decoding scheme, which are used in the previously proposed partial decoding scheme [13–15]. These properties of regular encoding/backward decoding scheme have been proved in [8]. It is also shown that the proposed rate in this paper subsumes the well-known previously presented rate by Xie and Kumar [8], which is based on decode-and-forward strategy.

The rest of this paper is organised as follows: Section 2 introduces modelling assumptions and notations. In Section 3, we review some theorems and concepts about sequential relaying based on partial decoding scheme. In Section 4, we drive a new achievable rate for relay networks based on comprehensive partial decoding approach. Finally, some concluding remarks are provided in Section 5.

2 Definitions and preliminaries

The discrete memoryless relay network shown in Fig. 1 [4, Fig. 2.1] 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: $\mathcal{X}_0 \times \mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times \mathcal{Y}_0 \times \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_N$ and a probability transition matrix $p(y_0, y_1, \ldots, y_N \mid x_0, x_1, \ldots, x_N)$ defined for all $(y_0, y_1, \ldots, y_N, x_0, x_1, \ldots, x_N) \in \mathcal{Y}_0 \times \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_N \times \mathcal{X}_0 \times \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$. In this model, X_0 is the input to the network, Y_0 is the ultimate output, Y_i is the ith relay output and X_i is the ith relay input.

An (M, n) code for the network consists of a set of integers $\mathcal{W} = \{1, 2, \ldots, M\}$, an encoding function $x_0^n: \mathcal{W} \to \mathcal{X}_0^n$, a set of relay function $\{f_{ij}\}$ such that

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

i.e. $x_{ij} \triangleq j$th component of $x_i^n \triangleq (x_{i1}, \ldots, x_{in})$, and a decoding function $g: \mathcal{Y}_0^n \to \mathcal{W}$. For generality, all functions are allowed to be stochastic functions.

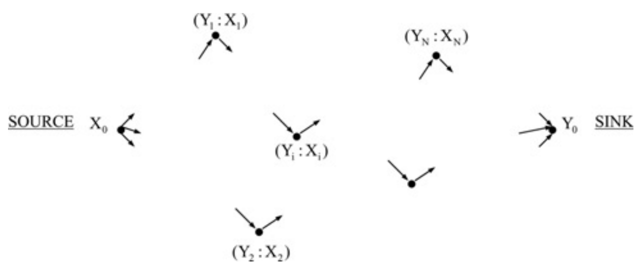

Figure 1 *General discrete memoryless relay network*

doi: 10.1049/iet-com.2008.0141

Let $y_i^{j-1} = (y_{i1}, y_{i2}, \ldots, y_{i,j-1})$. The input x_{ij} is allowed to depend only on the past received signals at the ith node, i.e. $(y_{i1}, \ldots, y_{i,j-1})$. The network is memoryless in the sense that $(y_{0i}, y_{1i}, \ldots, y_{Ni})$ depends on the past $(x_0^i, x_1^i, \ldots, x_N^i)$ only through the present transmitted symbols $(x_{0i}, x_{1i}, \ldots, x_{Ni})$. Therefore the joint probability mass function on $\mathcal{W} \times \mathcal{X}_0 \times \mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times \mathcal{Y}_0 \times \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_N$ is given by

$$p(w, x_0^n, x_1^n, \ldots, x_N^n, y_0^n, y_1^n, \ldots, y_N^n)$$
$$= p(w)\prod_{i=1}^{n} p(x_{0i} \mid w)p(x_{1i} \mid y_1^{i-1})\ldots p(x_{Ni} \mid y_N^{i-1})$$
$$\times p(y_{0i}, \ldots, y_{Ni} \mid x_{0i}, \ldots, x_{Ni})$$

where $p(w)$ is the probability distribution on the message $w \in \mathcal{W}$. If the message $w \in \mathcal{W}$ is sent, let $\lambda(w) \triangleq Pr\{g(Y_0^n) \neq W | W = w\}$ 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 $\bar{P}_e^n = (1/M)\sum_w \lambda(w)$. Let $\lambda_n \triangleq \max_{w\in\mathcal{W}} \lambda(w)$ be the maximal probability of error for the (M, n) code. The rate R of an (M, n) code is defined to be $R = (1/n)\log M$ bits/transmission. The rate R is said to be achievable by the network if, for any $\varepsilon > 0$, and for all n sufficiently large, there exists an (M, n) code with $M \geq 2^{nR}$ such that $\bar{P}_e^n < \epsilon$. The capacity $\mathcal{C}$ of the network is the supremum of the set of achievable rates.

Next, we review some basic properties of typical sequences that will be used later. For more details, see [16, Secs. 15.2] and [8, Sec. IV].

Let $(Z_1, Z_2, \ldots, Z_m)$ denote a finite collection of discrete random variables with some joint distribution $p(z_1, z_2, \ldots, z_m)$ for

$$(z_1, z_2, \ldots, z_m) \in \mathcal{Z}_1 \times \mathcal{Z}_2 \times \cdots \times \mathcal{Z}_m$$

Definition: The set $A_\epsilon^{(n)}$ of ϵ-typical T-sequences $(z_1^n, z_2^n, \ldots, z_m^n)$ is defined by

$$A_\epsilon^{(n)}(Z_1, Z_2, \ldots, Z_m)$$
$$:= \left\{(z_1^n, z_2^n, \ldots, z_m^n) : \left| -\frac{1}{n}\Pr(s^n) - H(S)\right|\right.$$
$$\left. < \epsilon, \forall S \subseteq \{Z_1, Z_2, \ldots, Z_m\}\right\}$$

where each $z_i^n = (z_{i1}, z_{i2}, \ldots, z_{in})$ is an n-vector, $i = 1, 2, \ldots, m$ and s^n is defined as follows: If $S = (Z_{i_1}, Z_{i_2}, \ldots, Z_{i_l})$ then $s^n = (z_{i_1}^n, z_{i_2}^n, \ldots, z_{i_l}^n)$ and

$$Pr(s^n) = Pr(z_{i_1}^n, z_{i_2}^n, \ldots, z_{i_l}^n) = \prod_{t=1}^{n} P(z_{i_1,t}, z_{i_2,t}, \ldots, z_{i_l,t})$$

Lemma 1: For any $\epsilon > 0$, let an n-sequence $(z_1^n, z_2^n, \ldots, z_m^n)$ be generated according to

$$\prod_{t=1}^{n} p(z_{1t}, z_{2t}, \ldots, z_{mt})$$

Then

$$Pr\Big((z_1^n, z_2^n, \ldots, z_m^n) \in A_\epsilon^{(n)}(Z_1, Z_2, \ldots, Z_m)\Big) \geq 1 - \epsilon$$

Lemma 2: For any $\epsilon > 0$, let an n-sequence $(z_1^n, z_2^n, \ldots, z_m^n)$ be generated according to

$$\prod_{t=1}^{n} p(z_{1t}|z_{2t}, \ldots, z_{m-1,t})$$
$$\times p(z_{mt}|z_{2t}, \ldots, z_{m-1,t})p(z_{2t}, \ldots, z_{m-1,t})$$

Then

$$Pr\Big((z_1^n, z_2^n, \ldots, z_m^n) \in A_\epsilon^{(n)}(Z_1, Z_2, \ldots, Z_m)\Big)$$
$$< 2^{-n\left(I(Z_1;Z_m|Z_2,\ldots,Z_{m-1})-6\epsilon\right)}$$

Lemma 3 [15, Lemma 1]: For any $\epsilon > 0$, let an n-sequence (x^n, y^n, u^n, w^n) be generated according to

$$\prod_{t=1}^{n} p(x \mid u)p(y \mid u, w)p(w, u)$$

Then

$$Pr\Big((x^n, y^n, u^n, w^n) \in A_\epsilon^{(n)}(X, Y, U, W)\Big) < 2^{-n(I(X;YW|U)-6\epsilon)}$$

Proof:

$$P_e = \sum_{(x^n,y^n,u^n,w^n)\in A_\epsilon^{(n)}(X,Y,U,W)} p(x \mid u)p(y \mid u, w)p(w, u) \quad (1)$$
$$= \|A_\epsilon^{(n)}\| 2^{-n(H(X|U)-2\epsilon)} 2^{-n(H(Y|WU)-2\epsilon)} 2^{-n(H(WU)-\epsilon)} \quad (2)$$
$$= 2^{n(H(XYWU)+\epsilon)} 2^{-n(H(X|U)-2\epsilon)} 2^{-n(H(Y|WU)-2\epsilon)} 2^{-n(H(WU)-\epsilon)} \quad (3)$$
$$= 2^{-n(I(X;YW|U)-6\epsilon)} \quad (4)$$

if the Markov condition $X \rightarrow U \rightarrow W$ holds, then $I(X; YW|U) = I(X; Y|WU)$. □

3 Sequential relaying scheme based on partial decoding

In [3], partial decoding scheme is defined as a special case of Theorem 7 in [2]. In this scheme, the relay does not completely decode the transmitted message by the sender. Instead, the relay decodes only part of the message transmitted by the sender. A block Markov encoding time-frame 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 sink in the current block. The following theorem expresses the obtained rate via partial decoding scheme.

Theorem 1 [3, Theorem]: For any relay network $(\mathcal{X}_0 \times \mathcal{X}_1, p(y_0, y_1|x_0, x_1), \mathcal{Y}_0 \times \mathcal{Y}_1)$, the capacity $\mathcal{C}$ is lower bounded by

$$\mathcal{C} \geq \max_{p(x_0,x_1)} \min\{I(X_0X_1; Y_0), I(U; Y_1|X_1) + I(X_0; Y_0|X_1U)\} \quad (5)$$

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

$$p(u, x_0, x_1, y_0, y_1) = p(u, x_0, x_1).p(y_0, y_1|x_0, x_1) \quad (6)$$

such that $U \to (X_0, X_1) \to (Y_0, Y_1)$ 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

$$\mathcal{C} \geq \max_{p(x_0,x_1)} \min\{I(X_0X_1; Y_0), I(X_0; Y_1|X_1)\} \quad (7)$$

The above expression introduces the capacity of degraded relay channel as shown in [2]. Moreover, by substituting $U = Y_1$ in (5), the capacity of semi-deterministic relay channel in which y_1 is a deterministic function of x_0 and x_1 is obtained [3, Corollary]

$$\mathcal{C} = \max_{p(x_0,x_1)} \min\{I(X_0X_1; Y_0), H(Y_1|X_1) + I(X_0; Y_0|X_1Y_1)\} \quad (8)$$

In [13], an achievable rate based on partial decoding is proposed, in which the second relay decode only part of the message of the first relay and the first relay decode only part of the message of the source. In Fig. 2, individual parts of the messages are shown. In this figure, U_1 denotes part of the message source that is decoded by the first relay. U_2 denotes part of the message of the first relay that is decoded by the second relay. The following theorem expresses this rate.

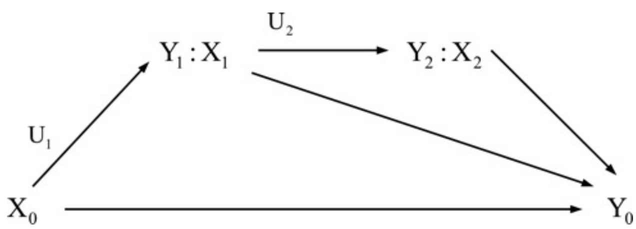

Figure 2 *Two-relay network with the individual parts of the transmitted message by the source and the relays shown in it, as expressed in Theorem 2*

Theorem 2 [13, Theorem 3]: For any relay network $(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2, p(y_0, y_1, y_2 \mid x_0, x_1, x_2), \mathcal{Y}_0 \times \mathcal{Y}_1 \times \mathcal{Y}_2)$, the capacity $\mathcal{C}$ is lower bounded by

$$\begin{aligned}\mathcal{C} \geq \sup_{p(u_1,u_2,x_0,x_1,x_2)} \min\{ & I(X_0, X_1, X_2; Y_0), \\ & I(U_2; Y_2|X_2) + I(X_0, X_1; Y_0|U_2, X_2), \\ & I(U_1; Y_1|X_1, X_2, U_2) + I(X_0; Y_0|X_1, X_2, U_1, U_2)\}\end{aligned} \quad (9)$$

where the supremum is over all joint probability mass functions $p(u_1, u_2, x_0, x_1, x_2)$ on $\mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$ such that

$$(U_1, U_2) \to (X_0, X_1, X_2) \to (Y_0, Y_1, Y_2) \quad (10)$$

form a Markov chain.

Remark: If we choose $U_1 = X_0$ and $U_2 = X_1$, they satisfy the Markovity criterion and result in the following rate

$$\begin{aligned}\mathcal{C} \geq \sup_{p(x_0,x_1,x_2)} \min\{ & I(X_0X_1X_2; Y_0), I(X_0; Y_1 \mid X_1X_2), \\ & I(X_1; Y_2 \mid X_2) + I(X_0; Y_0 \mid X_1X_2)\}\end{aligned} \quad (11)$$

Now we compare the above equation with the result of full decoding strategy [8, Theorem 3.2], in which each relay can decode all the source messages and all messages of the previous relay. The rate is as follows

$$\begin{aligned}\mathcal{C} \geq \sup_{p(x_0,x_1,x_2)} \min\{ & I(X_0X_1X_2; Y_0), \; I(X_0; Y_1 \mid X_1X_2), \\ & I(X_0X_1; Y_2 \mid X_2)\}\end{aligned} \quad (12)$$

By comparison of (11) and (12), it is seen that the first two terms of (11) and (12) are the same. About the third term, depending on the network conditions, each of them can be greater or less than the other.

In [15], a mixed strategy is derived for multiple relay networks using a combined approach of partial decoding and compress-and-forward based on random partitioning. In the partial decoding section of their theorem, they consider that only the source message is separated into parts and not the relay messages. Moreover, they did not consider any part of the message source that can be decoded only by the first relay. Parts of the message source that are decoded by the relays in this scheme are shown in Fig. 3. In this figure U_s^1 denotes part of the message source that is decoded by both relays and U_s^2 denotes part of the message source that is decoded by the second relay.

doi: 10.1049/iet-com.2008.0141

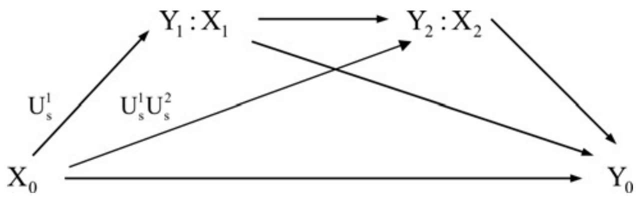

Figure 3 *Two-relay network with the individual parts of the transmitted message by the source shown in it*

4 Comprehensive partial decoding approach for two-level relay networks

In this section, we propose a comprehensive partial decoding scheme based on regular encoding/sliding window decoding. In this method, we consider all possible states of partial decoding, which can occur in the two-level relay network. In Fig. 4, individual parts of the messages are shown. In this figure, U_{01}^1 denotes common part of the source message that is decoded by both relays. U_{01}^2 denotes private part of the source message that is decoded by the first relay. U_{02} denotes private part of the source message that is decoded by the second relay. U_{12} denotes part of the transmitted message by the first relay that is decoded by the second relay. U_{20} denotes the way of the transmission of U_{12} to the destination. The rates of U_{01}^1, U_{01}^2 and U_{12} are shown by R_{01}^1, R_{01}^2 and R_{02}, respectively. The rate of the part of the message source that is directly decoded by the receiver is shown by R_{00}. As it is shown next, we do not need to define any auxilliary random variable for this part of the source message. The proposed rate is shown by the next theorem.

Theorem 3: For any relay networks $(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2, p(y_0, y_1, y_2 \mid x_0, x_1, x_2), \mathcal{Y}_0 \times \mathcal{Y}_1 \times \mathcal{Y}_2)$, the capacity $\mathcal{C}$ is lower bounded by

$$\mathcal{C} \geq \sup_{p(u_{01}^1, u_{01}^2, u_{02}, u_{12}, u_{20}, x_0, x_1, x_2)} R \tag{13}$$

$$R = R_{00} + R_{01}^1 + R_{01}^2 + R_{02} \tag{14}$$

$$R_{00} < I(X_0; Y_0 \mid U_{01}^1 U_{01}^2 U_{02} X_1 U_{12} X_2 U_{20}) \tag{15}$$

$$R_{02} < \min\{I(U_{02}; Y_2 \mid U_{01}^1 U_{12} X_2 U_{20}), I(U_{02}; Y_0 \mid U_{01}^1 X_1 U_{12} X_2 U_{20}) + I(X_2; Y_0 \mid U_{20})\} \tag{16}$$

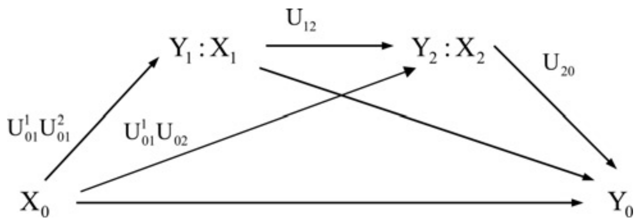

Figure 4 *Two-relay network with the individual parts of the transmitted message by the source and the relays shown in it, as expressed in Theorem 3*

$$R_{01}^1 < \min\{I(U_{01}^1; Y_1 \mid X_1 U_{12} U_{20}), I(U_{01}^1 U_{12}; Y_2 \mid X_2 U_{20}), I(U_{01}^1; Y_0 \mid X_1 U_{12} X_2 U_{20}) + I(U_{12}; Y_0 \mid X_2 U_{20}) + I(U_{20}; Y_0)\} \tag{17}$$

$$R_{01}^2 < \min\{I(U_{01}^2; Y_1 \mid U_{01}^1 X_1 U_{12} U_{20}), I(U_{01}^2; Y_0 \mid U_{02} U_{01}^1 X_1 U_{12} X_2 U_{20}) + I(X_1; Y_0 \mid U_{12} X_2 U_{20})\} \tag{18}$$

where the supremum is over all joint probability mass functions $p(x_0, x_1, x_2, u_{01}^1, u_{01}^2, u_{02}, u_{12}, u_{20})$ on $\mathcal{U}_{01}^1 \times \mathcal{U}_{01}^2 \times \mathcal{U}_{02} \times \mathcal{U}_{12} \times \mathcal{U}_{20} \times \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$ such that

$$(U_{01}^1, U_{01}^2, U_{02}, U_{12}, U_{20}) \to (X_0, X_1, X_2) \to (Y_0, Y_1, Y_2) \tag{19}$$

form a Markov chain.

Proof: In this encoding scheme, the message of the transmitter is divided into four parts. The first part is decoded only by two relays, the second part is decoded by the first relay, the third part is decoded by the second relay and the fourth part is directly decoded by the receiver. The message of first relay is also divided into two parts. The first part is decoded by second relay, and the receiver can only make an estimate of it, whereas 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.

We make use of regular block Markov superposition for encoding, and for decoding, we make use of a sliding window decoding scheme [8]. We consider B blocks of transmission, each of n symbols. A sequence of $B-2$ messages, $w_{00,i} \times w_{01,i}^1 \times w_{01,i}^2 \times w_{02,i} \in [1, 2^{nR_{00}}] \times [1, 2^{nR_{01}^1}] \times [1, 2^{nR_{01}^2}] \times [1, 2^{nR_{02}}]$, $i = 1, 2, \ldots, B-2$, will be sent over the channel in nB transmissions. $w_{01,i}^1$ is decoded only by two relays, $w_{01,i}^2$ is decoded by the first relay, $w_{02,i}$ is decoded by the second relay and $w_{00,i}$ is directly decoded by the receiver. 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.

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

Random coding: For any joint probability mass function

$$\begin{aligned} &p(x_0, x_1, x_2, u_{01}^1, u_{01}^2, u_{02}, u_{12}, u_{20}) \\ &= p(x_0 \mid x_1, x_2, u_{01}^1, u_{01}^2, u_{02}, u_{12}, u_{20}) \\ &\quad p(u_{01}^2 \mid u_{01}^1, x_1, u_{12}, u_{20}) \quad p(u_{02} \mid u_{01}^1, u_{12}, x_2, u_{20}) \\ &\quad p(u_{01}^1 \mid u_{12}, u_{20}) \ p(x_1 \mid u_{12}, u_{20}) p(u_{12} \mid u_{20}) \ p(x_2 \mid u_{20}) p(u_{20}) \end{aligned} \tag{20}$$

doi: 10.1049/iet-com.2008.0141

on the product set

$$\mathcal{U}_{20} \times \mathcal{U}_{12} \times \mathcal{U}_{02} \times \mathcal{U}_{01}^1 \times \mathcal{U}_{01}^2 \times \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$$

(a) Choose $2^{nR_{01}^1}$ i.i.d. u_{20}^n each with probability

$$p(u_{20}^n) = \prod_{i=1}^{n} p(u_{20,i})$$

Label these as $u_{20}^n(m_{20}^1)$, $m_{20}^1 \in [1, 2^{nR_{01}^1}]$.

(b) For every $u_{20}^n(m_{20}^1)$, generate $2^{nR_{01}^1}$ i.i.d u_{12}^n with probability

$$p(u_{12}^n \mid u_{20}^n) = \prod_{i=1}^{n} p(u_{12,i} \mid u_{20,i})$$

Label these $u_{12}^n(m_{12} \mid m_{20}^1)$, $m_{12} \in [1, 2^{nR_{01}^1}]$.

(c) For each $u_{20}^n(m_{20}^1)$ and $u_{12}^n(m_{12} \mid m_{20}^1)$ generate $2^{nR_{01}^1}$ i.i.d u_{01}^{1n} each with probability

$$p(u_{01}^{1n} \mid u_{12}^n, u_{20}^n) = \prod_{i=1}^{n} p(u_{01,i}^1 \mid u_{12,i}, u_{20,i})$$

Label these $u_{01}^{1n}(w_{01}^1 \mid m_{12}, m_{20}^1)$, $w_{01}^1 \in [1, 2^{nR_{01}^1}]$.

(d) For each $u_{20}^n(m_{20}^1)$, generate $2^{nR_{02}}$ i.i.d. x_2^n each with probability

$$p(x_2^n \mid u_{20}^n) = \prod_{i=1}^{n} p(x_{2i} \mid u_{20,i})$$

Label these as $x_2^n(m_{20}^2 \mid m_{20}^1)$, $m_{20}^2 \in [1, 2^{nR_{02}}]$.

(e) For each $u_{01}^{1n}(w_{01}^1 \mid m_{12}, m_{20}^1)$ and $u_{12}^n(m_{12} \mid m_{20}^1)$ $x_2^n(m_{20}^2 \mid m_{20}^1)$, generate $2^{nR_{02}}$ i.i.d u_{02}^n each with probability

$$p(u_{02}^n \mid u_{01}^{1n}, u_{12}^n, x_2^n) = \prod_{i=1}^{n} p(u_{02,i} \mid u_{01,i}^1, u_{12,i}^n, x_{2i}^n)$$

Label these $u_{02}^n(w_{02} \mid w_{01}^1, m_{12}, m_{20}^1, m_{20}^2)$,$w_{02} \in [1, 2^{nR_{02}}]$.

(f) For each $u_{20}^n(m_{20}^1)$ and $u_{12}^n(m_{12} \mid m_{20}^1)$, generate $2^{nR_{01}^2}$ i.i.d x_1^n each with probability

$$p(x_1^n \mid u_{12}^n, u_{20}^n) = \prod_{i=1}^{n} p(x_{1i} \mid u_{12,i}, u_{20,i})$$

Label these $x_1^n(m_{10} \mid m_{12}, m_{20}^1)$, $m_{10} \in [1, 2^{nR_{01}^2}]$.

(g) For each $u_{20}^n(m_{20}^1)$, $u_{12}^n(m_{12} \mid m_{20}^1)$, $x_1^n(m_{10} \mid m_{12}, m_{20}^1)$ and $u_{01}^{1n}(w_{01}^1 \mid m_{12}, m_{20}^1)$ generate $2^{nR_{01}^2}$ i.i.d u_{01}^{2n} each with probability

$$p(u_{01}^{2n} \mid u_{01}^{1n}, x_1^n, u_{12}^n, u_{20}^n) = \prod_{i=1}^{n} p(u_{01,i}^2 \mid u_{01,i}^1, x_{1i}, u_{12,i}, u_{20,i})$$

Label these as $u_{01}^{2n}(w_{01}^2 \mid w_{01}^1, m_{10}, m_{12}, m_{20}^1)$, $w_{01}^2 \in [1, 2^{nR_{01}^2}]$.

(h) For each $u_{20}^n(m_{20}^1)$, $x_2^n(m_{20}^2 \mid m_{20}^1)$, $u_{12}^n(m_{12} \mid m_{20}^1, m_{20}^2)$, $x_1^n(m_{10} \mid m_{12}, m_{20}^1)$, $u_{01}^{1n}(w_{01}^1 \mid m_{12}, m_{20}^1)$, $u_{02}^n(w_{02} \mid w_{01}^1, m_{12}, m_{20}^1, m_{20}^2)$ and $u_{01}^{2n}(w_{01} \mid m_{10}, m_{12}, m_{20}^1, m_{20}^2)$, generate $2^{nR_{00}}$ i.i.d x_0^n each with probability

$$p(x_0^n | u_{01}^{2n}, u_{02}^n, u_{01}^{1n}, x_1^n, u_{12}^n, x_2^n, u_{20}^n) = \prod_{i=1}^{n} p(x_{0i} | u_{01,i}^2, u_{02,i}, u_{01,i}^1, x_{1i}, u_{12,i}, x_{2i}, u_{20,i})$$

Label these as $x_0^n(w_{00} \mid w_{01}^2, w_{01}^1, w_{02}, m_{10}, m_{12}, m_{20}^1, m_{20}^2)$, $w_{00} \in [1, 2^{nR_{00}}]$.

In the defined codewords, the index m_{20}^1 represents the index m_{12} of the previous block. The index m_{12} represents the index w_{01}^1 of the previous block. The index m_{20}^2 represents the index w_{02} of two previous blocks. The index m_{10} represents the index w_{01}^2 of the previous block.

Based on this encoding scheme, the total rate transmitted by the sender is expressed as

$$R = R_{00} + R_{01}^1 + R_{01}^2 + R_{02} \tag{21}$$

This defines the joint codebook $\mathcal{C}_0$ for all the transmitter nodes.

Repeating the above process (a)–(h) independently six times, we generate another five random codebooks $\{\mathcal{C}_i\}_{i=1}^5$ similar to $\mathcal{C}_0$. We will use these six codebooks in a sequential way as follows: In block $b = 1, \ldots, B$, the codebook $\mathcal{C}_{b \bmod 6}$ is used. Hence, in any six consecutive blocks, codewords from different blocks are independent. This is a property we will use in the analysis of the probability of error.

Before the transmission, the codebooks $\{\mathcal{C}_i\}_{i=0}^5$ are revealed to all nodes.

The transmitter and relay encoders send the following codewords:

$$x_0^n(w_{00,i} \mid w_{01,i}^1, w_{01,i}^2, w_{02,i}, 1, 1, 1, 1), \quad x_1^n(1 \mid 1, 1), \; x_2^n(1 \mid 1)$$

in block $i = 1$, the following codewords

$$x_0^n(w_{00,i} \mid w_{01,i}^1, w_{01,i}^2, w_{02,i}, w_{01,i-1}^1, w_{01,i-1}^2, 1, 1), \quad x_1^n(\hat{w}_{01,i-1}^2 \mid \hat{w}_{01,i-1}^1, 1), \; x_2^n(1 \mid 1)$$

IET Commun., 2009, Vol. 3, Iss. 4, pp. 585–596
doi: 10.1049/iet-com.2008.0141

in each block $i = 2$, the following codewords

$$x_0^n(w_{00,i} \mid w_{01,i}^1, w_{01,i}^2, w_{02,i}, w_{01,i-1}^1, w_{01,i-1}^2, w_{02,i-2}, w_{01,i-2}^1)$$
$$x_1^n(\hat{w}_{01,i-1}^2 \mid \hat{w}_{01,i-1}^1, \hat{w}_{01,i-2}^1), x_2^n(\hat{w}_{02,i-2} \mid \hat{w}_{01,i-2}^1)$$

in each block $i = 3, \ldots, B-2$, the following codewords

$$x_0^n(1 \mid 1, 1, 1, w_{01,B-2}^1, w_{01,B-2}^2, w_{02,B-3}, w_{01,B-3}^1)$$
$$x_1^n(\hat{w}_{01,B-2}^2 \mid \hat{w}_{01,B-2}^1, \hat{w}_{01,B-3}^1), x_2^n(\hat{w}_{02,B-3} \mid \hat{w}_{01,B-3}^1)$$

in block $i = B-1$, and the following codewords

$$x_0^n(1 \mid 1, 1, 1, 1, 1, w_{02,B-2}, w_{01,B-2}^1), x_1^n(1 \mid 1, \hat{w}_{01,B-2}^1),$$
$$x_2^n(\hat{w}_{02,B-2} \mid \hat{w}_{01,B-2}^1)$$

Decoding: Assume that at the end of block $(i-1)$, the first relay knows $\{\hat{w}_{01,1}^1, \hat{w}_{01,2}^1, \ldots, \hat{w}_{01,i-1}^1\}$, $\{\hat{m}_{12,1}, \hat{m}_{12,2}, \ldots, \hat{m}_{12,i}\}$, $\{\hat{m}_{10,1}, \hat{m}_{10,2}, \ldots, \hat{m}_{10,i}\}$ and $\{\hat{m}_{20,1}^1, \hat{m}_{20,2}^1, \ldots, \hat{m}_{20,i}^1\}$. The second relay knows $\{\hat{w}_{02,1}, \hat{w}_{02,2}, \ldots, \hat{w}_{02,i-2}\}$, $\{\hat{w}_{01,1}^1, \hat{w}_{01,2}^1, \ldots, \hat{w}_{01,i-2}^1\}$, $\{\hat{m}_{12,1}, \hat{m}_{12,2}, \ldots, \hat{m}_{12,i-1}\}$, $\{\hat{m}_{20,1}^1, \hat{m}_{20,2}^1, \ldots, \hat{m}_{20,i}^1\}$ and $\{\hat{m}_{20,1}^2, \hat{m}_{20,2}^2, \ldots, \hat{m}_{20,i}^2\}$. The receiver knows $\{\hat{w}_{01,1}^1, \hat{w}_{01,2}^1, \ldots, \hat{w}_{01,i-3}^1\}$, $\{\hat{w}_{02,1}^1, \hat{w}_{02,2}^1, \ldots, \hat{w}_{02,i-3}^1\}$, $\{\hat{w}_{02,1}, \hat{w}_{02,2}, \ldots, \hat{w}_{02,i-3}\}$, $\{m_{12,1}, m_{12,2}, \ldots, m_{12,i-2}\}$, $\{\hat{m}_{10,1}, \hat{m}_{10,2}, \ldots, \hat{m}_{10,i-2}\}$, $\{\hat{m}_{20,1}^2, \hat{m}_{20,2}^2, \ldots, \hat{m}_{20,i-1}^2\}$ and $\{\hat{m}_{20,1}^1, \hat{m}_{20,2}^1, \ldots, \hat{m}_{20,i-1}^1\}$. At the end of block $i = B$, decoding is performed in the following manner.

(1) (*At the first relay*) Suppose the first relay has decoded accurately $w_{01,i-1}^1$ and by knowing $\hat{m}_{10,i}$, $\hat{m}_{12,i}$, $\hat{m}_{20,i}^1$ from the previous block. It determines $\hat{w}_{01,i}^1$ such that

$$\{u_{01}^{1n}(\hat{w}_{01,i}^1 \mid \hat{m}_{12,i}, \hat{m}_{20,i}^1), x_1^n(\hat{m}_{10,i} \mid \hat{m}_{12,i}, \hat{m}_{20,i}^1),$$
$$u_{12}^n(\hat{m}_{12,i} \mid \hat{m}_{20,i}^1), u_{20}^n(\hat{m}_{20,i}^1), y_1^n(i)\} \in A_\epsilon^n$$

$\hat{w}_{01,i}^1 = w_{01,i}^1$ with high probability if

$$R_{01}^1 < I(U_{01}^1; Y_1 | X_1 U_{12} U_{20}) \tag{22}$$

and n is sufficiently large.

The proof is based on Lemma 3, and based on the fact that according to (20), U_{01}^1 and X_1 are conditionally independent given (U_{12}, U_{20}), or $(U_{01}^1 \to (U_{12}, U_{20}) \to X_1)$.

(2) (*At the first relay*) Suppose the first relay has decoded accurately $w_{01,i-1}^2$ and by knowing $\hat{w}_{01,i}^1$, $\hat{m}_{10,i}$, $\hat{m}_{12,i}$, $\hat{m}_{20,i}^1$ from the previous block. It determines $\hat{w}_{01,i}^2$ such that

$$\left\{\begin{array}{c} u_{01}^{2n}(\hat{w}_{01,i}^2 \mid \hat{w}_{01,i}^1, \hat{m}_{10,i}, \hat{m}_{12,i}, \hat{m}_{20,i}^1), u_{01}^{1n}(\hat{w}_{01,i}^1 \mid \hat{m}_{12,i}, \hat{m}_{20,i}^1), \\ x_1^n(\hat{m}_{10,i} \mid \hat{m}_{12,i}, \hat{m}_{20,i}^1), u_{12}^n(\hat{m}_{12,i} \mid \hat{m}_{20,i}^1), u_{20}^n(\hat{m}_{20,i}^1), y_1^n(i) \end{array}\right\}$$
$$\in A_\epsilon^n$$

according to Lemma 2, $\hat{w}_{01,i}^2 = w_{01,i}^2$ with high probability if

$$R_{01}^2 < I(U_{01}^2; Y_1 \mid U_{01}^1 X_1 U_{12} U_{20}) \tag{23}$$

and n is sufficiently large.

(3) (*At the second relay*) Suppose the second relay has decoded accurately $w_{01,i-2}^1$. By knowing $m_{20,i}^1 = w_{01,i-2}^1$, $m_{20,i-1}^1 = w_{01,i-3}^1$, $m_{20,i}^2 = w_{02,i-2}^1$ and $m_{20,i-1}^2 = w_{02,i-3}^1$, it determines $\hat{w}_{01,i-1}^1$ or equivalently $\hat{m}_{12,i}$, if $\hat{w}_{01}^1$ is the unique value in $\{1, \ldots, 2^{nR_{01}^1}\}$ such that in the blocks $i-1$ and i,

$$\{u_{12}^n(\hat{m}_{12,i}|\hat{m}_{20,i}^1), x_2^n(\hat{m}_{20,i}^2|\hat{m}_{20,i}^1), u_{20}^n(\hat{m}_{20,i}^1), y_2^n(i)\} \in A_\epsilon^n \tag{24}$$

$$\left\{\begin{array}{c} u_{01}^{1n}(\hat{w}_{01,i-1}^1 \mid \hat{m}_{12,i-1}, \hat{m}_{20,i-1}^1), x_2^n(\hat{m}_{20,i-1}^2 \mid \hat{m}_{20,i-1}^1), \\ u_{12}^n(\hat{m}_{12,i-1} \mid \hat{m}_{20,i-1}^1), u_{20}^n(\hat{m}_{20,i-1}^1), y_2^n(i-1) \end{array}\right\} \in A_\epsilon^n \tag{25}$$

$\hat{w}_{01,i-1}^1 = w_{01,i-1}^1$ with high probability if

$$R_{01}^1 < I(U_{01}^1 U_{12}; Y_2 \mid X_2 U_{20}) \tag{26}$$

and n is sufficiently large. □

Proof: See Section 7.1

(4) (*At the second relay*) Suppose the second relay has decoded accurately $w_{02,i-2}$. It determines $\hat{w}_{02,i-1}$, if w_{02} is the unique value in $\{1, \ldots, 2^{nR_{02}}\}$ such that

$$\left\{\begin{array}{c} u_{02}^n(w_{02,i-1} \mid w_{01,i-1}^1, m_{12,i-1}, m_{20,i-1}^1, m_{20,i-1}^2), \\ u_{01}^{1n}(\hat{w}_{01,i-1}^1 \mid \hat{m}_{12,i-1}, \hat{m}_{20,i-1}^1), u_{12}^n(\hat{m}_{12,i-1} \mid \hat{m}_{20,i-1}^1), \\ x_2^n(\hat{m}_{20,i-1}^2 \mid \hat{m}_{20,i-1}^1), u_{20}^n(\hat{m}_{20,i-1}^1), y_2^n(i-1) \end{array}\right\} \in A_\epsilon^n$$

according to Lemma 2, $\hat{w}_{02,i} = w_{02,i}$ with high probability if

$$R_{02} < I(U_{02}; Y_2 | U_{01}^1 U_{12} X_2 U_{20}) \tag{27}$$

and n is sufficiently large.

(5) (*At the receiver*) Suppose the receiver has decoded accurately $\hat{w}_{01,i-3}^1$ and $\hat{w}_{01,i-3}^2$, thus by knowing $\hat{m}_{20,i-1}^1 = \hat{w}_{01,i-3}^1$, $\hat{m}_{20,i-2}^1 = \hat{w}_{01,i-4}^1$, $\hat{m}_{20,i-1}^2 = \hat{w}_{01,i-3}^2$ and $\hat{m}_{20,i-2}^2 = \hat{w}_{01,i-4}^2$ from the previous blocks, it determines $\hat{w}_{01,i-2}^1$ or equivalently $\hat{m}_{12,i-1}$or $\hat{m}_{20,i}^1$, if w_{01}^1 is the unique value in $\{1, \ldots, 2^{nR_{01}^1}\}$ such that in the blocks $i-2$, $i-1$ and i,

$$\{u_{20}^n(\hat{m}_{20,i}^1), y_0^n(i)\} \in A_\epsilon^n \tag{28}$$

$$\{u_{12}^n(\hat{m}_{12,i-1} \mid \hat{m}_{20,i-1}^1), x_2^n(\hat{m}_{20,i-1}^2 \mid \hat{m}_{20,i-1}^1),$$
$$u_{20}^n(\hat{m}_{20,i-1}^1), y_0^n(i-1)\} \in A_\epsilon^n \tag{29}$$

$$\left\{\begin{array}{c} u_{01}^{1n}(\hat{w}_{01,i-2}^1 \mid \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1), \\ x_1^n(\hat{m}_{10,i-2} \mid \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1), u_{12}^n(\hat{m}_{12,i-2} \mid \hat{m}_{20,i-2}^1), \\ x_2^n(\hat{m}_{20,i-2}^2 \mid \hat{m}_{20,i-2}^1), u_{20}^n(\hat{m}_{20,i-2}^1), y_0^n(i-2) \end{array}\right\} \in A_\epsilon^n \tag{30}$$

doi: 10.1049/iet-com.2008.0141

Otherwise, if no unique index w_{01}^1 as above exists, the error is declared. This can be successfully done with high probability if

$$R_{01}^1 < I(U_{01}^1; Y_0 \mid X_1 U_{12} X_2 U_{20}) + I(U_{12}; Y_0 \mid X_2 U_{20}) + I(U_{20}; Y_0) \quad (31)$$

and n is sufficiently large. □

Proof: See Section 7.2

(6) (*At the receiver*) Suppose the receiver has decoded accurately $\hat{w}_{01,i-2}^1$, thus by knowing $\hat{m}_{20,i}^1 = \hat{w}_{01,i-2}^1$, $\hat{m}_{12,i-2} = \hat{w}_{01,i-3}^1$ and $\hat{m}_{20,i-2}^2 = \hat{w}_{02,i-4}$ from the previous blocks, it determines $\hat{w}_{02,i-2}$ or equivalently $\hat{m}_{20,i}^2$, if w_{02} is the unique value in $\{1, \ldots, 2^{nR_{02}}\}$ such that in the blocks $i-2$ and i

$$\{x_2^n(\hat{m}_{20,i}^2 \mid \hat{m}_{20,i}^1), u_{20}^n(\hat{m}_{20,i}^1), y_0^n(i)\} \in A_\epsilon^n \quad (32)$$

$$\left\{\begin{array}{c} u_{02}^n(\hat{w}_{02,i-2} \mid \hat{w}_{01,i-2}^1, \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1, \hat{m}_{20,i-2}^2), \\ u_{01}^{1n}(\hat{w}_{01,i-2}^1 \mid \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1), \\ x_1^n(\hat{m}_{10,i-2} \mid \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1), \\ u_{12}^n(\hat{m}_{12,i-2} \mid \hat{m}_{20,i-2}^1), x_2^n(\hat{m}_{20,i-2}^2 \mid \hat{m}_{20,i-2}^1), \\ u_{20}^n(\hat{m}_{20,i-2}^1), y_0^n(i-2) \end{array}\right\} \in A_\epsilon^n \quad (33)$$

Otherwise, if no unique index w_{02} as above exists, the error is declared. This can be successfully done with high probability if

$$R_{02} < I(U_{02}; Y_0 \mid U_{01}^1 X_1 U_{12} X_2 U_{20}) + I(X_2; Y_0 \mid U_{20}) \quad (34)$$

and n is sufficiently large. □

Proof: See Section 7.3

(7) (*At the receiver*) Suppose the receiver has decoded accurately $\hat{w}_{01,i-2}^1$, thus by knowing $\hat{m}_{20,i-1}^1 = \hat{w}_{01,i-3}^1$, $\hat{m}_{20,i-2}^1 = \hat{w}_{01,i-4}^1$, $\hat{m}_{12,i-1} = \hat{w}_{01,i-2}^1$, $\hat{m}_{12,i-2} = \hat{w}_{01,i-3}^1$, $\hat{m}_{20,i-2}^2 = \hat{w}_{02,i-4}$ and $\hat{m}_{20,i-1}^2 = \hat{w}_{02,i-3}$ from the previous blocks, it determines $\hat{w}_{01,i-2}^2$ or equivalently $\hat{m}_{10,i-1}$, if w_{01}^2 is the unique value in $\{1, \ldots, 2^{nR_{01}^2}\}$, respectively, such that in the blocks $i-2$ and $i-1$,

$$\left\{\begin{array}{c} x_1^n(\hat{m}_{10,i-1} \mid \hat{m}_{12,i-1}, \hat{m}_{20,i-1}^1), u_{12}^n(\hat{m}_{12,i-1} \mid \hat{m}_{20,i-1}^1), \\ x_2^n(\hat{m}_{20,i-1}^2 \mid \hat{m}_{20,i-1}^1), u_{20}^n(\hat{m}_{20,i-1}^1), y_0^n(i-1) \end{array}\right\} \in A_\epsilon^n \quad (35)$$

$$\left\{\begin{array}{c} u_{01}^{2n}(\hat{w}_{01,i-2}^2 \mid \hat{w}_{01,i-2}^1, \hat{m}_{10,i-2}, \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1), \\ u_{02}^n(\hat{w}_{02,i-2} \mid \hat{w}_{01,i-2}^1, \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1, \hat{m}_{20,i-2}^2), \\ u_{01}^{1n}(\hat{w}_{01,i-2}^1 \mid \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1), \\ x_1^n(\hat{m}_{10,i-2} \mid \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1), \\ u_{12}^n(\hat{m}_{12,i-2} \mid \hat{m}_{20,i-2}^1), x_2^n(\hat{m}_{20,i-2}^2 \mid \hat{m}_{20,i-2}^1), \\ u_{20}^n(\hat{m}_{20,i-2}^1), y_0^n(i-2) \end{array}\right\} \in A_\epsilon^n \quad (36)$$

Otherwise, if no unique index w_{01}^2 as above exists, the error is declared. This can be successfully done with high probability if

$$R_{01}^2 < I(U_{01}^2; Y_0 \mid U_{02} U_{01}^1 X_1 U_{12} X_2 U_{20}) + I(X_1; Y_0 \mid U_{12} X_2 U_{20}) \quad (37)$$

and n is sufficiently large. □

Proof: See Section 7.4

(8) (*At the receiver:*) Suppose the receiver has decoded accurately $w_{01,i-2}^1$, $w_{01,i-2}^2$, $w_{02,i-2}$, the receiver declares that $\hat{w}_{00,i-2} = w_{00}$ if w_{00} is the unique value in $\{1, \ldots, 2^{nR_{00}}\}$ such that in the blocks $i-2$,

$$\left\{\begin{array}{c} x_0^n(\hat{w}_{00,i-2} \mid \hat{w}_{01,i-2}^2, \hat{w}_{01,i-2}^1, \hat{w}_{02,i-2}, \hat{m}_{10,i-2}, \hat{m}_{12,i-2}, \\ \hat{m}_{20,i-2}^1, \hat{m}_{20,i-2}^2), u_{01}^{2n}(\hat{w}_{01,i-2}^2 \mid \hat{w}_{01,i-2}^1, \hat{m}_{10,i-2}, \\ \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1), u_{02}^n(\hat{w}_{02,i-2} \mid \hat{w}_{01,i-2}^1, \hat{m}_{12,i-2}, \\ \hat{m}_{20,i-2}^1, \hat{m}_{20,i-2}^2), \\ u_{01}^{1n}(\hat{w}_{01,i-2}^1 \mid \hat{m}_{12,i-2}, \hat{m}_{20,i-2}^1), x_1^n(\hat{m}_{10,i-2} \mid \hat{m}_{12,i-2}, \hat{m}_{20,i-1}^1), \\ u_{12}^n(\hat{m}_{12,i-2} \mid \hat{m}_{20,i-2}^1), x_2^n(\hat{m}_{20,i}^2 \mid \hat{m}_{20,i}^1), \\ u_{20}^n(\hat{m}_{20,i-2}^1), y_0^n(i-2) \end{array}\right\} \in A_\epsilon^n$$

according to Lemma 2, $\hat{w}_{00,i-2} = w_{00,i-2}$ with high probability if

$$R_{00} < I(X_0; Y_0 \mid U_{01}^1 U_{01}^2 U_{02} X_1 U_{12} X_2 U_{20}) \quad (38)$$

and n is sufficiently large.

Now the relations of Theorem 3 can be obtained by the following argument: (14) and (15) are the same as (21) and (38), respectively. (16) results from (27) and (34). (17) results from (22) and (26). (18) is obtained from (23) and (37). □

Remarks:

(1) Theorem 1 can be considered as a special case of Theorem 3 by replacing $U_{02} = X_0$, $Y_2 = Y_0$, $U_{01}^1 = U$, $U_{12} = X_1$ and $U_{01}^2 = U_{20} = X_2 = Y_0 = 0$ in (13)–(18).

(2) If we choose $U_{01}^1 = X_0$, $U_{12} = X_1$, $U_{20} = X_2$ and $U_{01}^2 = U_{02} = 0$ in (13)–(18), they satisfy the Markovity

doi: 10.1049/iet-com.2008.0141

criterion and the result of block Markov encoding for two relay networks, (12), [8, Theorem 3.2] is obtained. It shows the priority of Theorem 3 to Theorem 2 [13, Theorem 3], which does not contain the rate of [8, Theorem 3.2] in general. Moreover, in Theorem 3, we take advantage of sliding window decoding scheme that has less delay than backward decoding, which was used in [13]. In the backward decoding scheme, we should delay decoding operation until the last block is received.

(3) By comparing our strategy of partial decoding, by the method presented in [15], it is seen that in our method there is some part of the decoded message by the first relay that is not decoded by the second relay. This case considers more general condition than the method of [15], in which all parts of the message decoded by the second relay are also decoded by the first relay. Another preference of our method is using regular encoding-sliding window decoding strategy, which in general yields higher rates than random partitioning that was used in [15].

5 Conclusion

This paper presents a new achievable rate based on a partial decoding scheme for the two-level relay network. An application of regular encoding/sliding window decoding is presented to implement the proposed rate. The proposed partial decoding scheme is called comprehensive partial decoding scheme, because in this strategy, we consider all possible states of partial decoding that can occur in a two-level relay network, as shown in Fig. 4. In this way, the common and private parts of the message transmitted by the source are specified to be decoded by the appropriate relays; moreover, the second relay can decode part of the transmitted message by the first relay. In the proof of Theorem 3, we take advantage of regular encoding/sliding window decoding scheme that has superiorities to regular encoding/backward decoding and irregular encoding/random partitioning, in having less delay and yielding higher rates, respectively. It is shown that the proposed rate includes previously presented rates based on decode-and-forward and partial decode-and-forward strategy.

6 Acknowledgments

This work was supported by Iranian National Science Foundation (INSF) under contract no. 84,5193-2006 and International Telecommunication Research Center (ITRC) under contract no. 500/1892.

7 References

[1] VAN DER MEULEN E.C.: 'Three terminal communication channels', *Adv. Appl. Probab.*, 1971, **3**, pp. 120–154

[2] COVER T.M., EL GAMAL A.: 'Capacity theorems for the relay channel', *IEEE Trans. Inf. Theory*, 1979, **IT-25**, (5), pp. 572–584

[3] EL GAMAL A., AREF M.: 'The capacity of the semideterministic relay channel', *IEEE Trans. Inf. Theory*, 1982, **IT-28**, (3), p. 536

[4] AREF M.R.: 'Information flow in relay networks', PhD Thesis, Stanford University, Stanford, CA, 1980

[5] EL GAMAL A., ZAHEDI S.: 'Capacity of relay channels with orthogonal components', *IEEE Trans. Inf. Theory*, 2005, **51**, (5), pp. 1815–1817

[6] GUPTA P., KUMAR P.R.: 'Towards an information theory of large networks: an achievable rate region', *IEEE Trans. Inf. Theory*, 2003, **49**, (8), pp. 1877–1894

[7] XIE L.L., KUMAR P.R.: 'A network information theory for wireless communication: scaling laws and optimal operation', *IEEE Trans. Inf. Theory*, 2004, **50**, (5), pp. 748–767

[8] XIE L.L., KUMAR P.R.: 'An achievable rate for the multiple-level relay channel', *IEEE Trans. Inf. Theory*, 2005, **51**, (4), pp. 1348–1358

[9] KRAMER G., GASTPAR M., GUPTA P.: 'Capacity theorems for wireless relay channels'. Proc. 41st Annu. Allerton Conf. Commun., Control, Computing, Monticello, IL, October 2003, pp. 1074–1083

[10] KRAMER G., GASTPAR M., GUPTA P.: 'Cooperative strategies and capacity theorems for relay networks', *IEEE Trans. Inf. Theory*, 2005, **51**, (9), pp. 3037–3063

[11] RAZAGHI P., YU W.: 'Parity forwarding for multiple-relay networks'. Proc. IEEE Int. Symp. Inf. Theory, Seattle, WA, 9–14 July 2006, pp. 1678–1682

[12] RAZAGHI P., YU W.: 'A structured generalization of decode-and-forward for the multiple relay network'. Proc. IEEE Int. Symp. Inf. Theory, Nice, France, 24–29 June 2007, pp. 271–275

[13] GHABELI L., AREF M.R.: 'A new achievable rate and the capacity of a class of semi-deterministic relay networks'. Proc. IEEE Int. Symp. Inf. Theory, Nice, France, 24–29 June 2007, pp. 281–285

[14] GHABELI L., AREF M.R.: 'A new achievable rate and the capacity of some classes of multilevel relay network', *EURASIP J. Wirel. Commun. Netw.*, 2008, **8**, (3),

[15] ROST P., FETTWEIS G.: 'Analysis of a mixed strategy for multiple relay networks', *IEEE Trans. Inf. Theory.*, 2009, **55**, (1), pp. 174–189

[16] COVER T.M., THOMAS J.A.: 'Elements of information theory' (Wiley, 2006, 2nd edn.)

8 Appendixes

8.1 Proof of (26)

Let

$$\mathcal{W}_{2,0}(i) := \{w_{01}^1 \in \{1, \ldots, 2^{nR_{01}^1}\}\colon w_{01}^1 \text{ satisfies (24)}\}$$

$$\mathcal{W}_{2,1}(i) := \{w_{01}^1 \in \{1, \ldots, 2^{nR_{01}^1}\}\colon w_{01}^1 \text{ satisfies (25)}\}$$

$$\mathcal{W}_2(i) := \mathcal{W}_{2,0}(i) \cap \mathcal{W}_{2,1}(i)$$

Then, the probability that some decoding error is made at this step is

$$\begin{aligned} P_{e_2} &:= Pr(w_{01,i-1}^1 \notin \mathcal{W}_2(i) \quad \text{or some } w' \in \mathcal{W}_2(i), \\ &\quad \text{but } w' \neq w_{01,i-1}^1) \\ &< Pr(w_{01,i-1}^1 \notin \mathcal{W}_2(i)) + Pr(\text{some } w' \in \mathcal{W}_2(i), \quad (39) \\ &\quad \text{but } w' \neq w_{01,i-1}^1) \quad (40) \end{aligned}$$

According to Lemma 1, it is concluded that

$$Pr\left(w_{01,i-1}^1 \notin \mathcal{W}_2(i) \le \sum_{j=0}^{1} Pr(w_{01,i-1}^1 \notin \mathcal{W}_{2,j}(i)\right) \le 2\epsilon \quad (41)$$

and by considering (24), according to Lemma 3, based on the fact that U_{12} and X_2 are conditionally independent given U_{20}, it is concluded that

$$Pr(w' \in \mathcal{W}_{2,0}(i)) < 2^{-n(I(U_{12};Y_2|X_2U_{20})-6\epsilon)} \quad (42)$$

By considering (25), according to Lemma 3 and based on the fact that U_{01}^1 and X_2 are conditionally independent given (U_12, U_20), it is concluded that

$$Pr(w' \in \mathcal{W}_{2,1}(i)) < 2^{-n(I(U_{01}^1;Y_2|U_{12}X_2U_{20})-6\epsilon)} \quad (43)$$

Hence

$$\begin{aligned} &Pr[\text{some } w' \in \mathcal{W}_2(i) \text{ but } w' \neq w_{01,i-1}^1] \\ &\quad \le \sum_{\substack{w' \in [1,2^{nR_{01}^1}] \\ w' \neq w_{01,i-1}}} Pr(w' \in \mathcal{W}_2(i)) \\ &\quad = \sum_{\substack{w' \in [1,2^{nR_{01}^1}] \\ w' \neq w_{01,i-2}}} \prod_{j=0}^{1} Pr(w' \in \mathcal{W}_{2,j}(i)) \quad (44) \\ &\le (2^{nR_{01}^1} - 1)2^{-n(I(U_{01}^1;Y_2|U_{12}X_2U_{20})+I(U_{12};Y_2|X_2U_{20})-12\epsilon)} \\ &= (2^{nR_{01}^1} - 1)2^{-n(I(U_{01}^1U_{12};Y_2|X_2U_{20})-12\epsilon)} \quad (45) \end{aligned}$$

where (44) follows from the independence between the codewords of any two consecutive blocks, which is obtained from the independence between the codewords of any six consecutive blocks as considered in code construction. Equation (45) follows from (42) and (43).

For any R_{01}^1 satisfying (26), by choosing ϵ small enough, we can make n large enough such that for any $\epsilon_1 > 0$, we get (45) to be less than ϵ_1. Hence, by (39) and (41), $P_{e_2} \le (2\epsilon + \epsilon_1)$, which can be made arbitrarily small by letting $n \to \infty$.

8.2 Proof of (31)

Let

$$\mathcal{W}_{r1,0}(i) := \{w_{01}^1 \in \{1, \ldots, 2^{nR_{01}^1}\}\colon w_{01}^1 \text{ satisfies (28)}\}$$

$$\mathcal{W}_{r1,1}(i) := \{w_{01}^1 \in \{1, \ldots, 2^{nR_{01}^1}\}\colon w_{01}^1 \text{ satisfies (29)}\}$$

$$\mathcal{W}_{r1,2}(i) := \{w_{01}^1 \in \{1, \ldots, 2^{nR_{01}^1}\}\colon w_{01}^1 \text{ satisfies (30)}\}$$

$$\mathcal{W}_{r1}(i) := \mathcal{W}_{r1,0}(i) \cap \mathcal{W}_{r1,1}(i) \cap \mathcal{W}_{r1,2}(i)$$

Then, the probability that some decoding error is made at this step is

$$\begin{aligned} P_{e_{01}} &:= Pr(w_{01,i-2}^1 \notin \mathcal{W}_{r1}(i) \\ &\quad \text{or some } w' \in \mathcal{W}_{r1}(i), \text{ but } w' \neq w_{01,i-2}^1) \quad (46) \\ &< Pr(w_{01,i-2}^1 \notin \mathcal{W}_{r1}(i)) \\ &\quad + Pr(\text{some } w' \in \mathcal{W}_{r1}(i), \text{ but } w' \neq w_{01,i-2}^1) \quad (47) \end{aligned}$$

According to Lemma 1, it is concluded that

$$Pr\left(w_{01,i-2}^1 \notin \mathcal{W}_{r1}(i) \le \sum_{j=0}^{2} Pr(w_{01,i-2}^1 \notin \mathcal{W}_{r1,j}(i)\right) \le 3\epsilon \quad (48)$$

By considering (29), according to Lemma 2, it is concluded that

$$Pr(w' \in \mathcal{W}_{r1,0}(i)) < 2^{-n(I(U_{20};Y_0)-6\epsilon)} \quad (49)$$

By considering (29), according to Lemma 3 and based on the fact that U_{12} and X_2 are conditionally independent given U_{20}, it is concluded that

$$Pr(w' \in \mathcal{W}_{r1,1}(i)) < 2^{-n(I(U_{12};Y_0|X_2U_{20})-6\epsilon)} \quad (50)$$

By considering (30), according to Lemma 3 and based on the fact that U_{01}^1 and (X_1, X_2) are conditionally independent given (U_{12}, U_{20}), it is concluded that

$$Pr(w' \in \mathcal{W}_{r1,2}(i)) < 2^{-n(I(U_{01}^1;Y_0|X_1U_{12}X_2U_{20})-6\epsilon)} \quad (51)$$

IET Commun., 2009, Vol. 3, Iss. 4, pp. 585–596

doi: 10.1049/iet-com.2008.0141

Hence,

$$Pr[\text{some } w' \in \mathcal{W}_{r1}(i) \text{ but } w' \neq w^1_{01,i-2}]$$
$$\leq \sum_{\substack{w' \in [1, 2^{nR^1_{01}}] \\ w' \neq w_{01,i-2}}} Pr(w' \in \mathcal{W}_{r1}(i))$$
$$= \sum_{\substack{w' \in [1, 2^{nR^1_{01}}] \\ w' \neq w_{01,i-2}}} \prod_{j=0}^{2} Pr(w' \in \mathcal{W}_{r1,j}(i)) \tag{52}$$
$$\leq (2^{nR^1_{01}} - 1) \times 2^{-n(I(U^1_{01};Y_0|X_1U_{12}X_2U_{20})+I(U_{12};Y_0|X_2U_{20})+I(U_{20};Y_0)-18\epsilon)} \tag{53}$$

where (52) follows from the independence between the codewords of any three consecutive blocks, which is obtained from the independence between the codewords of any six consecutive blocks as considered in code construction. (53) follows from (49), (50) and (51).

For any R^1_{01} satisfying (31), by choosing ϵ small enough, we can make n large enough such that for any $\epsilon_1 > 0$, we get (53) to be less than ϵ_1. Hence, by (46) and (48), $P_{e_{01}} \leq (3\epsilon + \epsilon_1)$, which can be made arbitrarily small by letting $n \to \infty$.

8.3 Proof of (34)

Let

$$\mathcal{W}_{r2,0}(i) := \{w_{02} \in \{1, \ldots, 2^{nR_{02}}\}: w_{02} \text{ satisfies } (32)\}$$
$$\mathcal{W}_{r2,1}(i) := \{w_{02} \in \{1, \ldots, 2^{nR_{02}}\}: w_{02} \text{ satisfies } (33)\}$$
$$\mathcal{W}_{r2}(i) := \mathcal{W}_{r2,0}(i) \cap \mathcal{W}_{r2,1}(i)$$

Then, the probability that some decoding error is made at the receiver in some block $i \in \{1, \ldots, B\}$ is

$$P_{e_{02}} := Pr(w_{02,i-2} \notin \mathcal{W}_{r2}(i) \text{ or some } w' \in \mathcal{W}_{r2}(i), \text{ but } w' \neq w_{02,i-2}) \tag{54}$$
$$< Pr(w_{02,i-2} \notin \mathcal{W}_{r2}(i)) + Pr(\text{some } w' \in \mathcal{W}_{r2}(i), \text{ but } w' \neq w_{02,i-2}) \tag{55}$$

According to Lemma 1, it is concluded that

$$Pr(w_{02,i-2} \notin \mathcal{W}_{r2}(i) \leq \sum_{j=0}^{1} Pr(w_{02,i-2} \notin \mathcal{W}_{r2,j}(i)) \leq 2\epsilon \tag{56}$$

By considering (32), according to Lemma 2, it is concluded that

$$Pr(w' \in \mathcal{W}_{r2,0}(i)) < 2^{-n(I(X_2;Y_0|U_{20})-6\epsilon)} \tag{57}$$

By considering (33), according to Lemma 3 and based on the fact U_{02} and X_1 are conditionally independent given (U^1_{01}, X_2, U_{20}), it is concluded that

$$Pr(w' \in \mathcal{W}_{r2,1}(i)) < 2^{-n(I(U_{02};Y_0|U^1_{01}X_1U_{12}X_2U_{20})-6\epsilon)} \tag{58}$$

Hence,

$$Pr[\text{some } w' \in \mathcal{W}_{r2}(i) \text{ but } w' \neq w_{02,i-2}]$$
$$\leq \sum_{\substack{w' \in [1, 2^{nR_{02}}] \\ w' \neq w_{02,i-2}}} Pr(w' \in \mathcal{W}_{r2}(i))$$
$$= \sum_{\substack{w' \in [1, 2^{nR_{02}}] \\ w' \neq w_{02,i-2}}} \prod_{j=0}^{1} Pr(w' \in \mathcal{W}_{r2,j}(i)) \tag{59}$$
$$\leq (2^{nR_{02}} - 1) \times 2^{-n(I(U_{02};Y_0|U_{01}X_1X_2U_{20})+I(X_2;Y_0|U_{20})-12\epsilon)} \tag{60}$$

where (59) follows from the independence between the codewords of any three consecutive blocks, which is obtained from the independence between the codewords of any six consecutive blocks as considered in code construction. Equation (60) follows from (57) and (58).

For any R_{02} satisfying (34), by choosing ϵ small enough, we can make n large enough such that for any $\epsilon_1 > 0$, we get (60) to be less than ϵ_1. Hence, by (54) and (56), $P_{e_{02}} \leq (2\epsilon + \epsilon_1)$, which can be made arbitrarily small by letting $n \to \infty$.

8.4 Proof of (37)

Let

$$\mathcal{W}_{r3,0}(i) := \{w^2_{01} \in \{1, \ldots, 2^{nR_{02}}\}: w^2_{01} \text{ satisfies } (35)\}$$
$$\mathcal{W}_{r3,1}(i) := \{w^2_{01} \in \{1, \ldots, 2^{nR_{02}}\}: w^2_{01} \text{ satisfies } (36)\}$$
$$\mathcal{W}_{r3}(i) := \mathcal{W}_{r3,0}(i) \cap \mathcal{W}_{r2,1}(i)$$

Then, the probability that some decoding error is made at the receiver in some block $i \in \{1, \ldots, B\}$ is

$$P_{e_{03}} := Pr(w^2_{01,i-2} \notin \mathcal{W}_{r3}(i) \text{ or some } w' \in \mathcal{W}_{r3}(i), \text{ but } w' \neq w^2_{01,i-2})$$
$$< Pr(w^2_{01,i-2} \notin \mathcal{W}_{r3}(i)) + Pr(\text{some } w' \in \mathcal{W}_{r3}(i), \text{ but } w' \neq w^2_{01,i-2}) \tag{61}$$

According to Lemma 1, it is concluded that

$$Pr\left(w^2_{01,i-2} \notin \mathcal{W}_{r3}(i) \leq \sum_{j=0}^{1} Pr(w^2_{01,i-2} \notin \mathcal{W}_{r3,j}(i)\right) \leq 2\epsilon \tag{62}$$

By considering (35), according to Lemma 3 and based on the fact X_1 and X_2 are conditionally independent given (U_{12}, U_{20}), it is concluded that

$$Pr(w' \in \mathcal{W}_{r3,0}(i)) < 2^{-n(I(X_1;Y_0|U_{12}X_2U_{20})-6\epsilon)} \quad (63)$$

By considering (36), according to Lemma 3 and based on the fact U_{01}^2 and (U_{02}, X_2) are conditionally independent given $(U_{01}^1, X_1, U_{12}, U_{20})$, it is concluded that

$$Pr(w' \in \mathcal{W}_{r3,1}(i)) < 2^{-n(I(U_{01}^2;Y_0|U_{02}U_{01}^1X_1U_{12}X_2U_{20})-6\epsilon)} \quad (64)$$

Hence,

$$Pr[\text{some } w' \in \mathcal{W}_{r3}(i) \text{ but } w' \neq w_{01,i-2}^2]$$
$$\leq \sum_{\substack{w' \in [1, 2^{nR_{01}^2}] \\ w' \neq w_{01,i-2}^2}} Pr(w' \in \mathcal{W}_{r3}(i))$$
$$= \sum_{\substack{w' \in [1, 2^{nR_{01}^2}] \\ w' \neq w_{01,i-2}^2}} \prod_{j=0}^{1} Pr(w' \in \mathcal{W}_{r3,j}(i)) \quad (65)$$
$$\leq (2^{nR_{01}^2} - 1) \times 2^{-n(I(U_{01}^2;Y_0|U_{02}U_{01}^1X_1U_{12}X_2U_{20})+I(X_1;Y_0|U_{12}X_2U_{20})-12\epsilon)} \quad (66)$$

where (65) follows from the independence between the codewords of any three consecutive blocks, which is obtained from the independence between the codewords of any six consecutive blocks as considered in code construction. (66) follows from (63) and (64).

For any R_{01}^2 satisfying (37), by choosing ϵ small enough, we can make n large enough such that for any $\epsilon_1 > 0$, we get (66) to be less than ϵ_1. Hence, by (61) and (62), $P_{e_{03}} \leq (2\epsilon + \epsilon_1)$ which can be made arbitrarily small by letting $n \to \infty$.

IET Commun., 2009, Vol. 3, Iss. 4, pp. 585–596
doi: 10.1049/iet-com.2008.0141