Chapter 12 Cooperative Strategies and LDPC Codes for Relay Networks

12.1 Brief Introduction to Cooperative Communication

12.1.1 Diversity from cooperation

The advantages of multiple-input multiple-output (MIMO) systems have been widely acknowledged, to the extent that certain transmit diversity methods (i.e., Alamouti signaling) have been incorporated into wireless standards. Although transmit diversity is clearly advantageous on a cellular base station, but it may not be practical for other scenarios. Specifically, due to size, cost or hardware limitations, a wireless user may not be able to support multiple transmit antennas. Well-known examples include most handsets due to the size or the nodes in a wireless sensor network because of the size and power.

The mobile wireless channel suffers from fading, meaning that the signal attenuation can vary significantly over the course of a given transmission. Transmitting independent copies of the signal generates diversity and can effectively combat the deleterious effects of fading. In particular, spatial diversity is generated by transmitting signals from different locations, thus allowing independently faded versions of the signals at the receiver.

Recently, a new class of techniques known as cooperative communication has been proposed, which allow single-antenna mobiles to reap some of the benefits of MIMO systems and generate diversity in a new and interesting way. The basic idea hidden behind cooperative communication is that single-antenna mobiles in a multi-user scenario can share their antennas in a manner that can effectively creates a virtual MIMO system.

For a preliminary explanation of the ideas behind cooperative communication, we use Fig. 12.1 that illustrates two mobile users (or agents) communicating with the same destination. Each mobile has one antenna and cannot individually generate spatial diversity. However, it may be possible for user 2 to receive the signal from user 1, in which case user 2 can forward some version of "overhead" information from user 1 along with its own transmitted data. Because the fading paths from two mobiles are statistically independent, this generates spatial diversity for the communication link between user 1 and the destination.

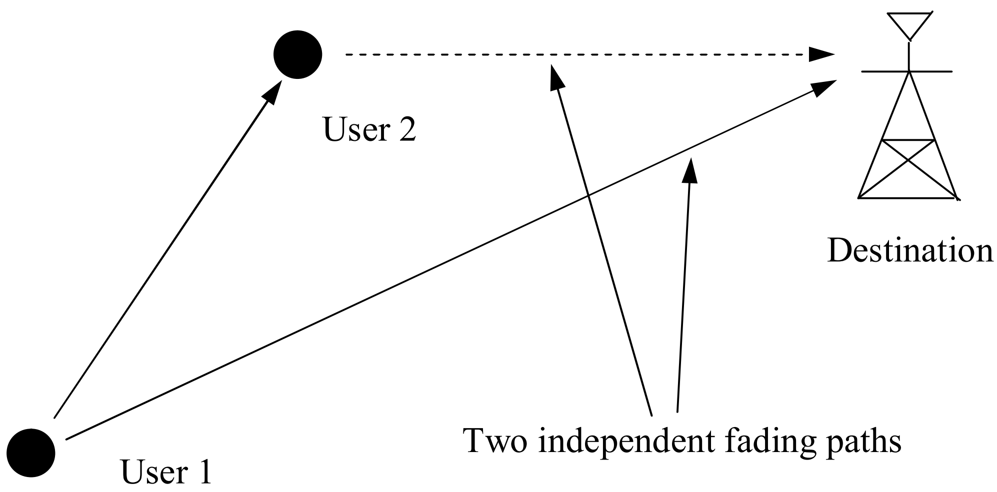


Fig. 12.1. An illustrative diagram of cooperative communication.

Generically, in a cooperative communication system, each wireless user is assumed to transmit data as well as act as a cooperative agent for another user. The illustrative diagram is shown in Fig. 12.2.

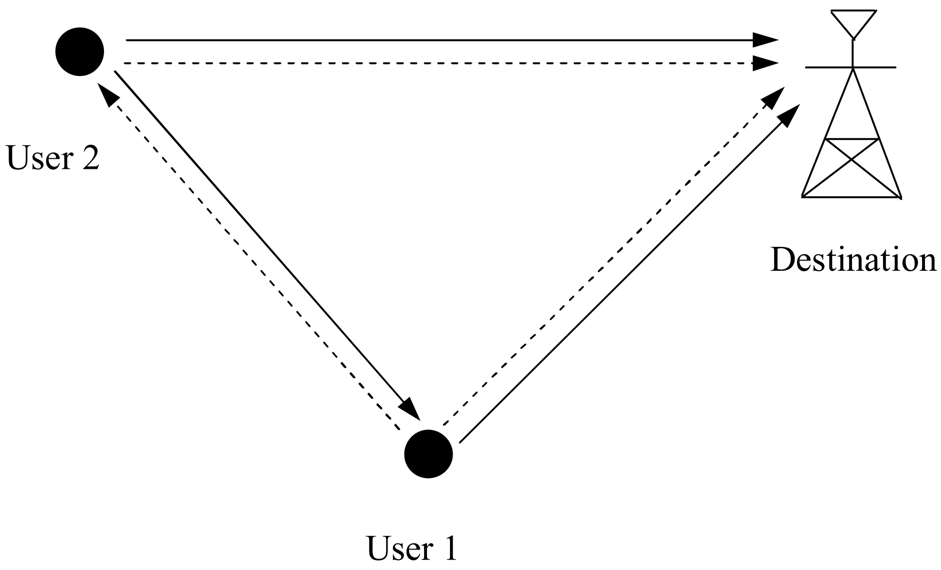


Fig. 12.2. Each mobile user is both a user and a relay in cooperative communication.

The users in the figures can be base stations or handsets, etc. The idea of cooperation is general, and perhaps even more suitable to ad hoc wireless networks and wireless sensor networks than cellular networks.

12.1.2 Some essential questions and trade-offs

Cooperation leads to interesting trade-offs in code rate and transmit power. In the case of power, one may argue on one hand that more power is needed because each user, when in cooperative mode, is transmitting for both users. On the other hand, the baseline transmit power for both users will be reduced because of diversity. In this trade-off, one hopes for a net reduction of transmit power, given everything else being constant.

Similar questions arise for the rate of the system. In cooperative communication each user transmits both its own bits as well as some information for its associated partner. One might think this may cause loss of rate in the overall system. The key question, what is the channel capacity of communication link by using cooperation compared with the conventional link without cooperation, is naturally proposed. Based on the obtained capacity, one can ideally employ a specific high efficient channel code, such as LDPC and turbo codes, etc., with sufficiently long block length and with code rate less than the capacity for cooperative system in order to achieve an error-free transmission between source and destination by cooperative strategies. Also, the spectrum efficiency of each user improve because, due to increasing capacity by cooperation diversity the channel code rate can be increased. Again a trade-off is observed. The premise of cooperation is that certain allocation strategies for the power and bandwidth of mobiles lead to significant gains in system performance [75][76].

10.1.3 Historical background for cooperative communication

The essential ideas behind cooperative communication can be traced back to the pioneer work of Cover and EI Gamal [74] on the information theoretic properties of the relay channel. This work analyzed the capacity of the simple three-node network only consisting of a source, a destination, and a relay, which were assumed to operate in the same band. Thus, the overall system can be decomposed into a broadcast channel from the viewpoint of the source and a multiple access channel from the viewpoint of the destination. An illustrative diagram of the relay channel is shown in Fig. 12.3, which consists of information source A, destination C and relay B.

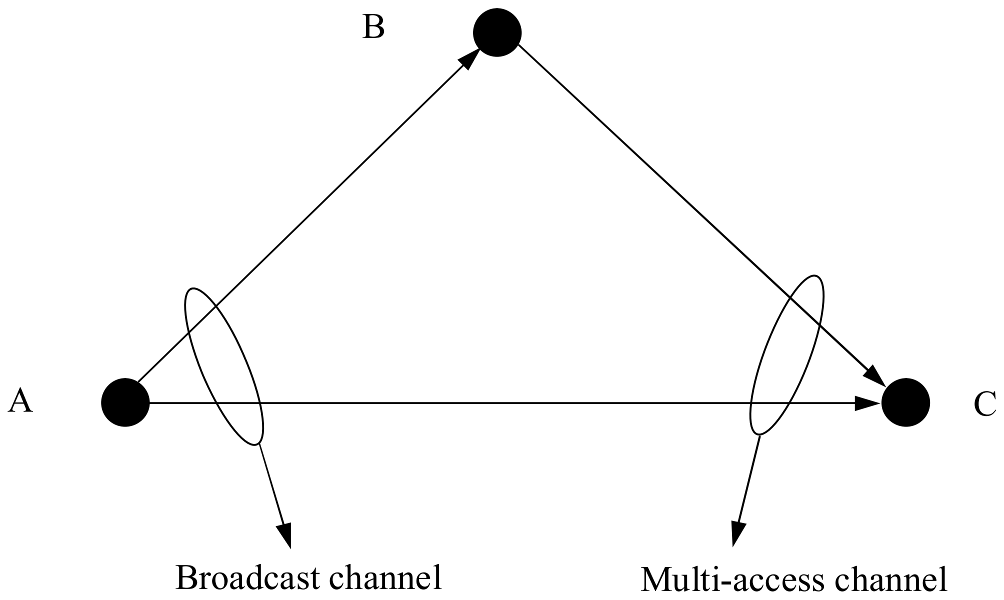


Fig. 12.3. A simple relay channel.

Many ideas that appeared later in the cooperative communications were first explored in [74]. However, in many respects the cooperative communication considered now is beyond and quite different from the original relay channel. First, recent developments are motivated by the concept of diversity on a fading channel [75][76][78][80], while the original work by Cover and EI Gamal [74] mainly is focused on capacity in an AWGN channel with deployment of relay. Second, in the relay channel, the sole purpose of relay is to help the main channel, while in cooperative communication the total system resources are fixed, and each user acts as an information source as well as a relay. Therefore, although the historical importance of [74] is indisputable, recent work in cooperation has taken a different emphasis, in particular for cooperation diversity and information theory for large scale wireless networks [75-80].

In this chapter we will first present a brief theoretical analysis on the capacity bound of cooperative communication over noisy channel, where the discrete memoryless multiple-level relay network is employed as a simple and typical example. Then we will apply low-density parity-check (LDPC) codes of the code rate less than the capacity bound to relay networks and properly distribute the information flow through the various transmissions in the networks. Finally, the performance of LDPC-coded system over relay channel is analyzed and the results of bit error rate (BER) are obtained by numerical simulations.

12.2 One relay channel

The relay channel was first introduced by van der Meulen [83]. The simplest case shown in Fig. 12.4 is the three-node system where node 1 purely functions as a relay to help the information transmission from node 0 to node 2. An immediate application of this framework, for example, is in satellite communications, where a relay satellite (node 1) is placed between the ground station (node 0) and destination satellite (node 2) in order to shorten the distance of a hop or prevent the scenario from line-out-of-sight between the ground station and the destination satellite.

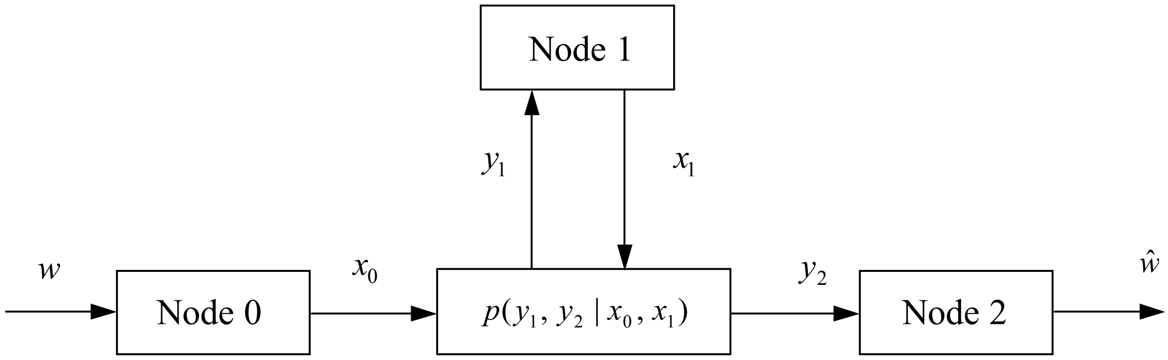


Fig. 12.4. The simplest one-relay channel.

The simplest discrete memoryless one-relay channel can be also depicted by Fig. 12.4, which can be denoted by ($\Omega_0 \times \Omega_1$, $p(y_1, y_2 | x_0, x_1)$, $\Psi_1 \times \Psi_2$), where Ω_0 and Ω_1 are the transmitter alphabets of nodes 0 and 1, respectively, Ψ_1 and Ψ_2 are the receiver alphabets of nodes 1 and 2, respectively, w and $\hat{w}$ are information message and estimated message, respectively, and a collection of probability distributions $p(y_1, y_2 \mid x_0, x_1)$ on (y_1, y_2)$\in \Psi_1 \times \Psi_2$ for each (x_0, x_1)$\in \Omega_0 \times \Omega_1$. The interpretation is that x_0 is the input to the channel from the source node 0, y_2 is the output of the channel to the destination node 2, and y_1 is the output received by the relay node 1. After processing y_1, the relay node 1 sends the input x_1 chosen as a function of its past parameters

$$x_1(t) = f_t(y_1(t-1), y_1(t-2), \ldots) \quad \text{for every } t \tag{12-1}$$

where $f_t(.)$ can be any causal function. Note that a one-step time delay is assumed in (12-1) to account for the signal processing time by the relay. Initially, information goes from source to the relay, and then from the relay to the destination. However, the destination needs to take into account both the inputs by the source and the relay in order to make a final decision. The destination node can make use of the signal coming from the source node, even though it may be relatively weaker than that from the relay node due to its greater distance from the source.

Definition 12.1: The discrete memoryless one-relay channel ($\Omega_0 \times \Omega_1$, $p(y_1, y_2 | x_0, x_1)$, $\Psi_1 \times \Psi_2$) shown in Fig. 12.4 is said to be degraded if

$$p(y_1, y_2 \mid x_0, x_1) = p(y_1 \mid x_0, x_1) p(y_2 \mid y_1, x_1) \tag{12-2a}$$

Remark on (12-2a):

$$p(y_1, y_2 \mid x_0, x_1) = p(y_2 \mid y_1, x_0, x_1) p(y_1 \mid x_0, x_1) \tag{12-2b}$$

Hence, by (12-2a) and (12-2b) can be equivalently expressed as

$$p(y_2 \mid y_1, x_0, x_1) = p(y_2 \mid y_1, x_1) \tag{12-2c}$$

which implies that $X_0 \to (X_1, Y_1) \to Y_2$ forms a Markov chain. This notion can be easily extended to multiple-level relay scenario.

12.3 Model of multiple-level relay channel

We begin with a definition of the discrete memoryless multiple-level relay channel, which is depicted by Fig. 12.5, by generalizing the basic ideal of the simplest discrete memoryless one-relay channel. Consider a channel with M+1 nodes. Let the source node be denoted by 0, the destination node by M, and let the other M−1 nodes be denoted sequentially as 1, 2, …., M−1 in arbitrary order. Assume each node $i \in \{0, 1, …., M\}$ sends $x_i(t) \in \Omega_i$ at time t, and each node $k \in \{1, 2, …, M\}$ receives $y_i(t) \in \Psi_k$ at time k, where the finite sets Ω_i and Ψ_k are the corresponding input and output alphabets for the corresponding nodes. The channel dynamics is described by the following probability function:

$$p(y_1, y_2, ..., y_M \big| x_0, x_1, ..., x_{M-1}) \tag{12-3}$$

for all

$$(x_0, x_1, ..., x_{M-1}) \in \Omega_0 \times \Omega_1 \times ..... \times \Omega_{M-1} \tag{12-4a}$$

and

$$(y_1, y_2, ..., y_M) \in \Psi_1 \times \Psi_2 \times ..... \times \Psi_M \tag{12-4b}$$

Analogously, we also assume a one-step time delay at every relay node to account for the signal processing time, so that for all $i \in \{1, 2, …., M-1\}$

$$x_i(t) = f_{i,t}(y_i(t-1), y_i(t-2), ...) \quad \text{for all } t \tag{12-5}$$

where $f_{i,t}$ can be any causal function.

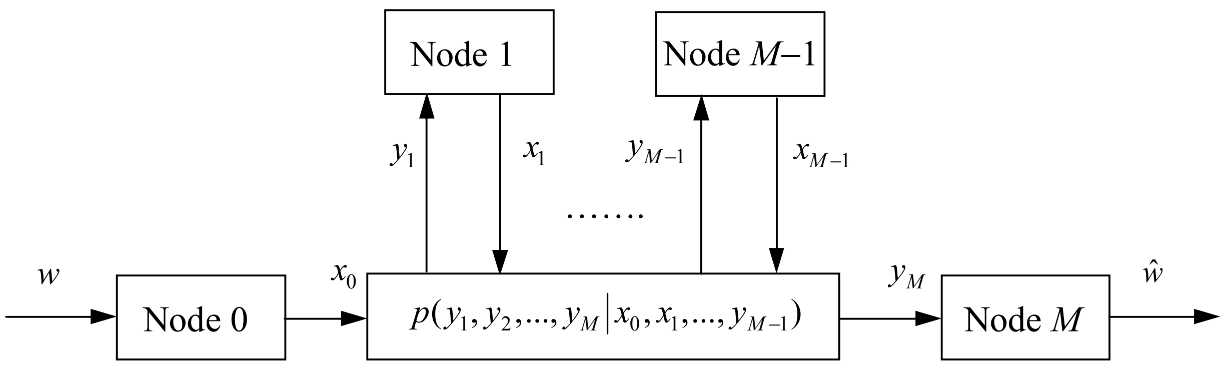


Fig. 12.5. The multiple-level relay channel.

We will consider the generalized degraded version of the multiple-level relay channel, which is given by the following definition

Definition 12.2: A discrete memoryless multiple-level relay channel is said to be degraded if

$$p(y_{k+1}, y_{k+2}, ..., y_M | y_k, x_0, ..., x_{k-1}, x_k, ...x_{M-1})$$
$$= p(y_{k+1}, y_{k+2}, ..., y_M | y_k, x_k, ...x_{M-1}) \quad \text{for } k = 1, 2, …, M-1 \tag{12-6}$$

Equivalently, (12-6) means that (X_0, X_1, …, X_{k-1})→(Y_k, X_k, …, X_{M-1})→(Y_{k+1}, Y_{k+2}, …, Y_M) forms a Markov chain for every k =1, 2, …., M–1. In the case of M=2, (12-6) reduces to (12-2a).

12.4 An achievable rate for the discrete memoryless multiple-level relay channel

In this section we will present the definitions of binary random coding and achievable rates for random codes over the discrete memoryless multiple-level relay channel.

Definition 12.3: The random block code $(2^{NR}, N, P_N)$ for a discrete memoryless multiple-level relay channel is made up of the following:

(**1**) A random variable W with probability $P(W=k)=1/2^{NR}$ for every $k \in \{1, 2, \ldots, 2^{NR}\}$.

(**2**) An encoding function F_0: $\{1, 2, \ldots, 2^{NR}\} \to \Omega_0^N$ for the source node 0, and relay function $f_{i,t}$: $\Psi_i^{t-1} \to \Omega_i$, t=2, 3, …., N for all the relay node $i \in \{1, 2, \ldots, M-1\}$, such that

$$\mathbf{x}_0 := [x_0(1), x_0(2), \ldots\ldots, x_0(N)] = F_0(W) \tag{12-7a}$$

$$x_i(t) = f_{i,t}(y_i(t-1), y_i(t-2), \ldots\ldots, y_i(1)) \tag{12-7b}$$

Note that $x_i(1)$ can be assigned any element in Ω_i.

(**3**) A decoding function G: $\Psi_M^N \to \{1, 2, \ldots, 2^{NR}\}$ for the destination node M with $\mathbf{y}_M \in \Psi_M^T$ as

$$\mathbf{y}_M = [y_M(1), y_M(2), \ldots\ldots, y_M(N)] \tag{12-8}$$

(**4**) The maximal probability of block error by the decoding strategy in (**3**) is

$$P_N := \max_{k \in \{1, 2, \ldots, 2^{NR}\}} \Pr\left(\mathrm{G}(\mathbf{y}_M) \neq k \mid W \neq k\right) \tag{12-9}$$

where N is the block length of the random code and R is the achievable rate for the transmission over discrete memoryless multiple-level channel. The exact definition of R is given as below

Definition 12.4: A rate R>0 is said to be achievable if there exists a *sequence* of $(2^{NR}, N, P_N)$ random block codes such that the maximal probability of block error P_N tends to zero as $N \to +\infty$.

Next, we will find the upper bound of achievable rate R that is guaranteed by the following theorem [82][77][79].

Theorem 12.1: For the discrete memoryless multiple-level relay channel defined above, the achievable rate R is bounded by

$$R < \max_{p(x_0, x_1, ..., x_{M-1})} \min_{1 \le k \le M} I(X_0, X_1, ..., X_{k-1}; Y_k \mid X_k, X_{k+1}, ..., X_{M-1}) \tag{12-10}$$

where $I(X_0, X_1, ..., X_{k-1}; Y_k | X_k, X_{k+1}, ..., X_{M-1})$ is the conditional mutual information between (X_0, X_1, …, X_{k-1}) and Y_k given (X_k, X_{k+1}, …, X_{M-1}). The proof of the theorem is given in Appendix 12-A.

For the degraded discrete memoryless multiple-level relay channel, the following theorem [82] shows that the right-hand side of (12-10) is actually the capacity.

Theorem 12.2: The capacity of the degraded discrete memoryless multiple-level relay channel is

$$C = \max_{p(x_0, x_1, ..., x_{M-1})} \min_{1 \le k \le M} I(X_0, X_1, ..., X_{k-1}; Y_k \mid X_k, X_{k+1}, ..., X_{M-1}) \tag{12-11}$$

More specifically, for a degraded discrete memoryless one-relay channel, i.e., M=2, we have

$$C = \max_{p(x_0, x_1)} \min\left\{ I(X_0; Y_1 | X_1), I(X_0, X_1; Y_2) \right\} \tag{12-12a}$$

For a degraded discrete memoryless two-relay channel, i.e., M=3, we have

$$C= \max_{p(x_0,x_1,x_2)} \min\{I(X_0;Y_1|X_1,X_2), I(X_0,X_1;Y_2|X_2,), I(X_0,X_1,X_2;Y_3)\} \quad (12\text{-}12b)$$

Remarks:

(A) We can intuitively image that there is an information flow from the source 0 to the destination node M along the path 0→1→2→……→M. Each node i gets the information one-step before the next node $i+1$. Hence, by the time the information reaches node k, all the upstream nodes (i.e., nodes with small index than k) have already obtained the same information and can thus cooperate with the node k. In other words, the transmitted signals from the upstream nodes can provide cooperation diversity to the received signal by the node k.

(B) Meanwhile, the conditioning with respect to the conditional mutual information is due to the reason that the downstream nodes (i.e., nodes with larger index than k) gets no more information than node k. In other words, the upstream nodes know messages input and output to the downstream nodes.

Example 12.1: For a discrete memoryless one-relay channel (M=2), we can provides more details in Fig. 12.6 according to the general principle given by Fig. 12.4.

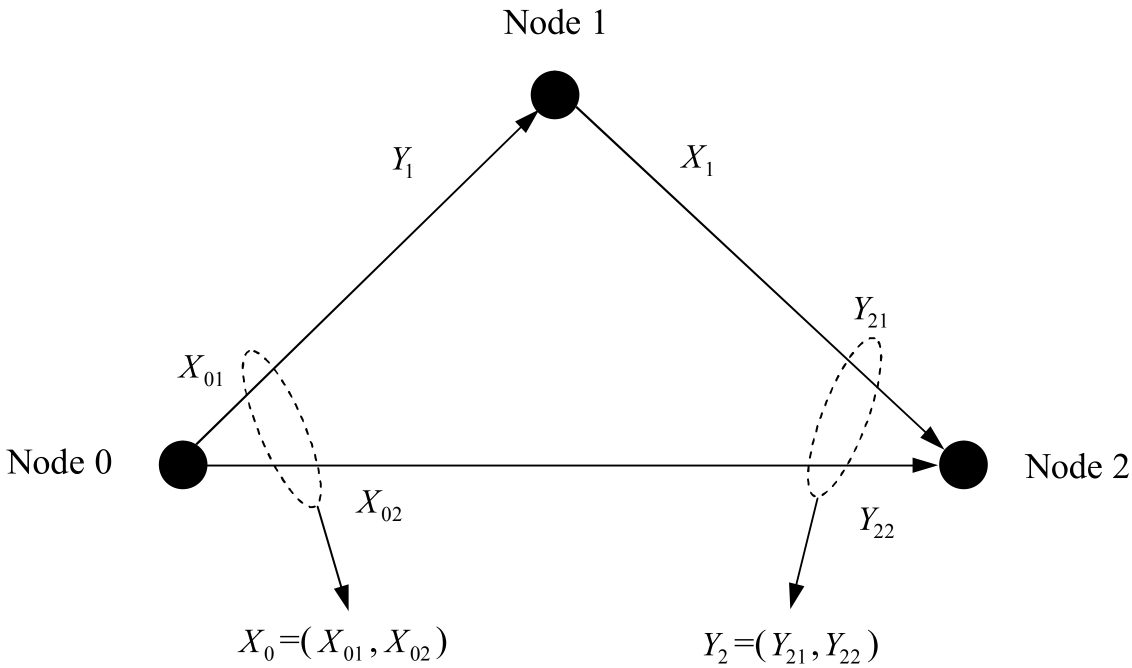


Fig. 12.6. An example of a one-relay channel.

If we make an assumption, which the outputs of the nodes are independent with each other and any output is only related to the inputs from the associated nodes, then the channel probability function can be greatly simplified as

$$p(y_1, y_2 | x_0, x_1) = p(y_1, (y_{21}, y_{22}) | (x_{01}, x_{02}))$$

$$= p(y_1|x_{01})\ p(y_{21}|x_1)\ p(y_{22}|x_{02}) \qquad (12\text{-}13)$$

The upper bound of the achievable rate becomes

$$R < \max_{p(x_0,x_1)} \min\left\{I(X_0;Y_1|X_1), I(X_0,X_1;Y_2)\right\}$$

$$= \max_{p(x_0,x_1)} \min\left\{I((X_{01},X_{02});Y_1|X_1), I((X_{01},X_{02}),X_1;(Y_{21},Y_{22}))\right\}$$

$$= \max_{p(x_0,x_1)} \min\left\{I(X_{01};Y_1),(I(X_{02};Y_{22})+I(X_1,Y_{21}))\right\} \qquad (12\text{-}14)$$

Note that we use the following identity for the second terms in (12-14).

$$I(X_0, X_1; Y_2) = I(X_1; Y_2) + I(X_0; Y_2 | X_1) \tag{12-15}$$

Example 12.2: For a discrete memoryless two-relay channel (*M*=3), which is shown in Fig. 12.7 based on the general principle depicted in Fig. 12.5.

Using the same assumption used by the last example, the channel probability function becomes

$$p(y_1, y_2, y_3 | x_0, x_1, x_2) = p(y_1, (y_{21}, y_{22}), (y_{31}, y_{32}, y_{33}) | (x_{01}, x_{02}, x_{03}), (x_{11}, x_{12}), x_2)$$

$$= p(y_1 | x_{01})\ p(y_{21} | x_{11})\ p(y_{22} | x_{02})\ p(y_{31} | x_2)\ p(y_{32} | x_{12})\ p(y_{33} | x_{03}) \tag{12-16}$$

The upper bound of the achievable rate can be expressed as

$$R < \max_{p(x_0, x_1, x_2)} \min\{I(X_0; Y_1 | X_1, X_2), I(X_0, X_1; Y_2 | X_2,), I(X_0, X_1, X_2; Y_3)\}$$

$$= \max_{p(x_0, x_1, x_2)} \min\{I((X_{01}, X_{02}, X_{03}); Y_1 | (X_{11}, X_{12}), X_2),$$

$$I((X_{01}, X_{02}, X_{03}), (X_{11}, X_{12}); (Y_{21}, Y_{22}) | X_2,),$$
$$I((X_{01}, X_{02}, X_{03}), (X_{11}, X_{12}), X_2; (Y_{31}, Y_{32}, Y_{33}))\}$$

$$= \max_{p(x_0, x_1, x_2)} \min\{I(X_{01}; Y_1), (I(X_1; Y_2) + I(X_0; Y_2 | X_1)), (I(X_1, X_2; Y_3) + I(X_0; Y_3 | X_1, X_2))\}$$

$$= \max_{p(x_0, x_1, x_2)} \min\{I(X_{01}; Y_1), (I(X_{11}; Y_{21}) + I(X_{02}; Y_{22})),$$

$$(I(X_2; Y_{31}) + I(X_{12}; Y_{32}) + I(X_{03}; Y_{33}))\} \tag{12-17}$$

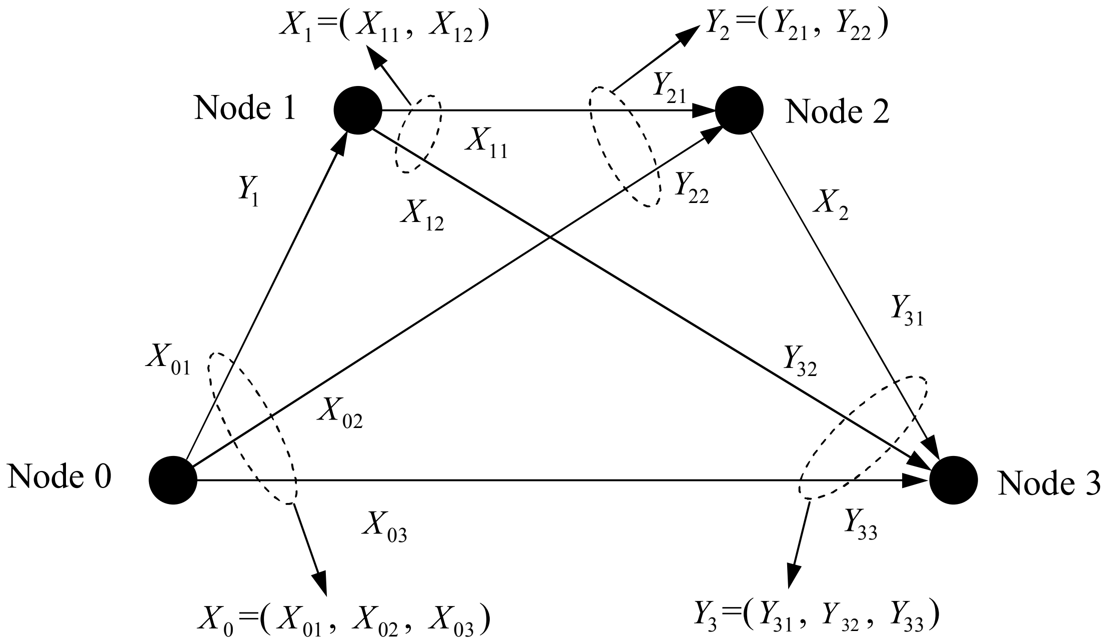


Fig. 12.7. An example of a two-relay channel.

Remark: If we allow an input $x_M \in \Omega_M$ to the relay channel from the destination node *M*, namely, the node k (k=1, 2, …., M) can both transmit and receive the signal to and from the relay channel, then it immediately follows from the Theorem 12.1 that the following rate [82] is achievable:

$$R < \max_{x_M \in \Omega_M} \max_{p(x_0, x_1, \ldots, x_M)} \min_{1 \le k \le M} I(X_0, X_1, \ldots, X_{k-1}; Y_k \mid X_k, X_{k+1}, \ldots, X_{M-1}, x_M) \quad (12\text{-}18)$$

However, we can achieve higher rate than (12-18) by the following theorem [82].

Theorem 12.3: For the discrete memoryless multiple-level relay channel defined above with an additional input $x_M \in \Omega_M$ from the destination node *M*, the following rate is achievable:

$$R < \max_{p(x_0, x_1, ..., x_M)} \min_{1 \le k \le M} I(X_0, X_1, ..., X_{k-1}; Y_k \mid X_k, X_{k+1}, ..., X_{M-1}, X_M) \quad (12\text{-}19)$$

The proof is straightforward since the conclusion above is an extension of the Theorem 12.1. Clearly, the upper bound of the achievable rate by (12-19) is larger than that by (10-18) that is regarded as a special case of the former one.

The following example will give the details for computing the upper bound of the achievable rate by (12-18) and (12-19).

Table 12.1 The channel dynamics $p(y_1, y_2 | x_0, x_1, x_2)$ for a one-relay channel (M=2).

(x_0, x_1, x_2) \ $(y_1\ y_2)$	(0, 0)	(0, 1)	(1, 0)	(1, 1)
(0, 0, 0)	1	0	0	0
(0, 1, 0)	0	1	0	0
(1, 0, 0)	1	0	0	0
(1, 1, 0)	0	1	0	0
(0, 0, 1)	1	0	0	0
(0, 1, 1)	1	0	0	0
(1, 0, 1)	0	0	1	0
(1, 1, 1)	0	0	1	0

Example 12.3: Consider a simple discrete memoryless one-relay channel (M=2) with channel probability function $p(y_1, y_2 | x_0, x_1, x_2)$ given by Table 12.1, and the alphabets for the input and output discrete symbols are

$$\Omega_0 = \Omega_1 = \Omega_2 = \Psi_1 = \Psi_2 = \{0,\ 1\} \tag{12-20}$$

(I) Calculate the upper bound of the achievable rate by (12-18)

$$R < \max_{x_2 \in \{0,\ 1\}} \max_{p(x_0,\ x_1)} \min\{I(X_0; Y_1 \mid X_1, x_2), I(X_0, X_1; Y_2 \mid x_2)\} \tag{12-21}$$

(1) The computation of $I(X_0; Y_1 | X_1, x_2 = 0)$:

The conditional mutual information $I(X_0; Y_1 | X_1, x_2)$ is evaluated as

$$I(X_0; Y_1 | X_1, x_2) = H(Y_1 | X_1, x_2) - H(Y_1 | X_0, X_1, x_2) \tag{12-22}$$

where

$$H(Y_1 | X_1, x_2) = -\sum_{y_1=0}^{1} \sum_{x_1=0}^{1} p(Y_1 = y_1 | X_1 = x_1, X_2 = x_2) p(X_1 = x_1, X_2 = x_2)$$

$$\times \log p(Y_1 = y_1 | X_1 = x_1, X_2 = x_2) \tag{12-23a}$$

$$H(Y_1 | X_0, X_1, x_2) = -\sum_{y_1=0}^{1} \sum_{x_0=0}^{1} \sum_{x_1=0}^{1} p(Y_1 = y_1 | X_0 = x_0, X_1 = x_1, X_2 = x_2)$$

$$\times p(X_0 = x_0, X_1 = x_1, X_2 = x_2)$$

$$\times \log p(Y_1 = y_1 | X_0 = x_0, X_1 = x_1, X_2 = x_2) \tag{12-23b}$$

(1-A) Evaluate $H(Y_1 | X_1, x_2 = 0)$

From Table 12.1 we have

$$p(Y_1 = 0, Y_2 = 0 | X_0 = 0, X_1 = 0, x_2 = 0) = 1 \tag{12-24a}$$

$$p(Y_1 = 0, Y_2 = 0 | X_0 = 1, X_1 = 0, x_2 = 0) = 1 \tag{12-24b}$$

By (12-24a) it yields

$$\begin{aligned} & p(Y_1 = 0, Y_2 = 0, X_0 = 0 | X_1 = 0, x_2 = 0) \\ = & \, p(Y_1 = 0, Y_2 = 0 | X_0 = 0, X_1 = 0, x_2 = 0) \; p(X_0 = 0 | X_1 = 0, x_2 = 0) \\ = & \, p(X_0 = 0 | X_1 = 0, x_2 = 0) \end{aligned} \tag{12-25a}$$

By (12-24b) we have

$$\begin{aligned} & p(Y_1 = 0, Y_2 = 0, X_0 = 1 | X_1 = 0, x_2 = 0) \\ = & \, p(Y_1 = 0, Y_2 = 0 | X_0 = 1, X_1 = 0, x_2 = 0) \; p(X_0 = 1 | X_1 = 0, x_2 = 0) \\ = & \, p(X_0 = 1 | X_1 = 0, x_2 = 0) \end{aligned} \tag{12-25b}$$

By (12-25a) and (12-25b),

$$\begin{aligned} & p(Y_1 = 0, Y_2 = 0 | X_1 = 0, x_2 = 0) \\ = & \, p(Y_1 = 0, Y_2 = 0, X_0 = 0 | X_1 = 0, x_2 = 0) \\ + & \, p(Y_1 = 0, Y_2 = 0, X_0 = 1 | X_1 = 0, x_2 = 0) \\ = & \, p(X_0 = 0 | X_1 = 0, x_2 = 0) + p(X_0 = 1 | X_1 = 0, x_2 = 0) = 1 \end{aligned} \tag{12-26a}$$

Hence,

$$p(Y_1 = 0, Y_2 = 1 | X_1 = 0, x_2 = 0) = 0 \tag{12-26b}$$

From (12-26a) and (12-26b)

$$p(Y_1 = 0 | X_1 = 0, x_2 = 0) = p(Y_1 = 0, Y_2 = 0 | X_1 = 0, x_2 = 0)$$

$$+ p(Y_1 = 0, Y_2 = 1 | X_1 = 0, x_2 = 0) = 1 \qquad \text{(12-27a)}$$

Therefore,

$$p(Y_1 = 1 | X_1 = 0, x_2 = 0) = 0 \qquad \text{(12-27b)}$$

From Table 12.1 we have

$$p(Y_1 = 0, Y_2 = 1 | X_0 = 0, X_1 = 1, x_2 = 0) = 1 \qquad \text{(12-28a)}$$

$$p(Y_1 = 0, Y_2 = 1 | X_0 = 1, X_1 = 1, x_2 = 0) = 1 \qquad \text{(12-28b)}$$

From (12-28a) we obtain

$$p(Y_1 = 0, Y_2 = 1, X_0 = 0 | X_1 = 1, x_2 = 0)$$

$$= p(Y_1 = 0, Y_2 = 1 | X_0 = 0, X_1 = 1, x_2 = 0)\ p(X_0 = 0 | X_1 = 1, x_2 = 0)$$

$$= p(X_0 = 0 | X_1 = 1, x_2 = 0) \qquad \text{(12-29a)}$$

Similarly, from (12-28b) it yields

$$p(Y_1 = 0, Y_2 = 1, X_0 = 1 | X_1 = 1, x_2 = 0)$$

$$= p(Y_1 = 0, Y_2 = 1 | X_0 = 1, X_1 = 1, x_2 = 0)\ p(X_0 = 1 | X_1 = 1, x_2 = 0)$$

$$= p(X_0 = 1 | X_1 = 1, x_2 = 0) \qquad \text{(10-29b)}$$

By (12-29a) and (12-29b),

$$p(Y_1 = 0, Y_2 = 1 | X_1 = 1, x_2 = 0)$$

$$= p(Y_1 = 0, Y_2 = 1, X_0 = 0 | X_1 = 1, x_2 = 0)$$

$$+ p(Y_1 = 0, Y_2 = 1, X_0 = 1 | X_1 = 1, x_2 = 0)$$

$$= p(X_0 = 0 | X_1 = 1, x_2 = 0) + p(X_0 = 1 | X_1 = 1, x_2 = 0) = 1 \qquad (12\text{-}30a)$$

Thus,

$$p(Y_1 = 0, Y_2 = 0 | X_1 = 1, x_2 = 0) = 0 \qquad (12\text{-}30b)$$

From (12-30a) and (12-30b)

$$p(Y_1 = 0 | X_1 = 1, x_2 = 0) = p(Y_1 = 0, Y_2 = 1 | X_1 = 1, x_2 = 0)$$

$$+ p(Y_1 = 0, Y_2 = 0 | X_1 = 1, x_2 = 0) = 1 \qquad (12\text{-}31a)$$

Hence

$$p(Y_1 = 1 | X_1 = 1, x_2 = 0) = 0 \qquad (12\text{-}31b)$$

Finally, by (12-27a), (12-27b), (12-31a), (12-31b) and (12-23a) we have

$$H(Y_1 | X_1, x_2 = 0) = 0 \text{ for any } p(x_0, x_1) \qquad (12\text{-}32)$$

(1-B) Find $H(Y_1 | X_0, X_1, x_2 = 0)$

From (12-24a) we have

$$p(Y_1 = 0, Y_2 = 1 | X_0 = 0, X_1 = 0, x_2 = 0) = 0 \qquad (12\text{-}33)$$

By (12-24a) and (12-33)

$$p(Y_1 = 0 | X_0 = 0, X_1 = 0, x_2 = 0) = 1 \qquad (12\text{-}34a)$$

Hence

$$p(Y_1 = 1 | X_0 = 0, X_1 = 0, x_2 = 0) = 0 \qquad (12\text{-}34b)$$

From (12-24b)

$$p(Y_1=0\,|\,X_0=1, X_1=0, x_2=0)=1 \tag{12-35a}$$

$$p(Y_1=1\,|\,X_0=1, X_1=0, x_2=0)=0 \tag{12-35b}$$

From (12-28a)

$$p(Y_1=0\,|\,X_0=0, X_1=1, x_2=0)=1 \tag{12-36a}$$

$$p(Y_1=1\,|\,X_0=0, X_1=1, x_2=0)=0 \tag{12-36b}$$

From (12-28b)

$$p(Y_1=0\,|\,X_0=1, X_1=1, x_2=0)=1 \tag{12-37a}$$

$$p(Y_1=1\,|\,X_0=1, X_1=1, x_2=0)=0 \tag{12-37b}$$

Finally, by (12-34), (12-35), (12-36), (12-37) and (12-23b) we have

$$H(Y_1\,|\,X_0, X_1, x_2=0)=0 \quad \text{for any } p(x_0,x_1) \tag{12-38}$$

(**1-C**) Find $I(X_0;Y_1\,|\,X_1, x_2=0)$

From (12-32), (12-38) and (12-22) we conclude that

$$I(X_0;Y_1\,|\,X_1, x_2=0)=0 \quad \text{for any } p(x_0,x_1) \tag{12-39}$$

(2) The computation of $I(X_0, X_1;Y_2\,|\,x_2=1)$

The conditional mutual information $I(X_0, X_1;Y_2\,|\,x_2)$ is evaluated as

$$I(X_0, X_1;Y_2\,|\,x_2)=H(Y_2\,|\,x_2)-H(Y_2\,|\,X_0, X_1, x_2) \tag{12-40}$$

where

$$H(Y_2\,|\,x_2)=-\sum_{y_2=0}^{1} p(Y_2=y_2\,|\,X_2=x_2)\,p(X_2=x_2)$$

$$\times \log p(Y_2 = y_2 | X_2 = x_2) \qquad (12\text{-}41a)$$

$$H(Y_2 | X_0, X_1, x_2) = -\sum_{y_2=0}^{1} \sum_{x_0=0}^{1} \sum_{x_1=0}^{1} p(Y_2 = y_2 | X_0 = x_0, X_1 = x_1, X_2 = x_2)$$

$$\times p(X_0 = x_0, X_1 = x_1, X_2 = x_2)$$

$$\times \log p(Y_2 = y_2 | X_0 = x_0, X_1 = x_1, X_2 = x_2) \qquad (12\text{-}41b)$$

(2-A) Find $H(Y_2 | x_2 = 1)$

From Table 12.1 we have

$$p(Y_1 = 0, Y_2 = 0 | X_0 = 0, X_1 = 0, x_2 = 1) = 1 \qquad (12\text{-}42a)$$

$$p(Y_1 = 0, Y_2 = 0 | X_0 = 0, X_1 = 1, x_2 = 1) = 1 \qquad (12\text{-}42b)$$

$$p(Y_1 = 1, Y_2 = 0 | X_0 = 1, X_1 = 0, x_2 = 1) = 1 \qquad (12\text{-}42c)$$

$$p(Y_1 = 1, Y_2 = 0 | X_0 = 1, X_1 = 1, x_2 = 1) = 1 \qquad (12\text{-}42d)$$

By (12-42a) and (12-42b)

$$p(Y_1 = 0, Y_2 = 0, X_0 = 0 | x_2 = 1)$$

$$= p(Y_1 = 0, Y_2 = 0, X_0 = 0, X_1 = 0 | x_2 = 1)$$

$$+ p(Y_1 = 0, Y_2 = 0, X_0 = 0, X_1 = 1 | x_2 = 1)$$

$$= p(Y_1 = 0, Y_2 = 0 | X_0 = 0, X_1 = 0, x_2 = 1)\ p(X_0 = 0, X_1 = 0 | x_2 = 1)$$

$$+ p(Y_1 = 0, Y_2 = 0 | X_0 = 0, X_1 = 1, x_2 = 1)\ p(X_0 = 0, X_1 = 1 | x_2 = 1)$$

$$= p(X_0 = 0, X_1 = 0 | x_2 = 1) + p(X_0 = 0, X_1 = 1 | x_2 = 1) \qquad (12\text{-}43a)$$

By (12-42c) and (12-42d)

$$p(Y_1=1, Y_2=0, X_0=1 \mid x_2=1)$$

$$= p(X_0=1, X_1=0 \mid x_2=1) + p(X_0=1, X_1=1 \mid x_2=1) \tag{12-43b}$$

By (12-43a) and (12-43b)

$$p(Y_1=0, Y_2=0, X_0=0 \mid x_2=1) + p(Y_1=1, Y_2=0, X_0=1 \mid x_2=1) = 1 \tag{12-44a}$$

Hence

$$p(Y_1=1, Y_2=0, X_0=0 \mid x_2=1) = 0 \tag{12-44b}$$

$$p(Y_1=0, Y_2=0, X_0=1 \mid x_2=1) = 0 \tag{12-44c}$$

By (12-44a), (12-44b) and (12-44c)

$$p(Y_2=0, X_0=0 \mid x_2=1) + p(Y_2=0, X_0=1 \mid x_2=1) = 1 \tag{12-45}$$

which is

$$p(Y_2=0 \mid x_2=1) = 1 \tag{12-46a}$$

and

$$p(Y_2=1 \mid x_2=1) = 0 \tag{12-46b}$$

Hence, by (12-46a), (12-46b) and (12-41a) we have

$$H(Y_2 \mid x_2=1) = 0 \quad \text{for any } p(x_0, x_1) \tag{12-47}$$

(2-B) Find $H(Y_2 \mid X_0, X_1, x_2)$

By (12-42a)

$$p(Y_1=1, Y_2=0 \mid X_0=0, X_1=0, x_2=1) = 0 \tag{12-48}$$

Using (12-42a) and (12-48) it yields

$$p(Y_2 = 0 | X_0 = 0, X_1 = 0, x_2 = 1) = 1 \quad \text{(12-49a)}$$

Equivalently,

$$p(Y_2 = 1 | X_0 = 0, X_1 = 0, x_2 = 1) = 0 \quad \text{(12-49b)}$$

By (12-42b)

$$p(Y_2 = 0 | X_0 = 0, X_1 = 1, x_2 = 1) = 1 \quad \text{(12-50a)}$$

$$p(Y_2 = 1 | X_0 = 0, X_1 = 1, x_2 = 1) = 0 \quad \text{(12-50b)}$$

By (12-42c)

$$p(Y_2 = 0 | X_0 = 1, X_1 = 0, x_2 = 1) = 1 \quad \text{(12-51a)}$$

$$p(Y_2 = 1 | X_0 = 1, X_1 = 0, x_2 = 1) = 0 \quad \text{(12-51b)}$$

By (12-42d)

$$p(Y_2 = 0 | X_0 = 1, X_1 = 1, x_2 = 1) = 1 \quad \text{(12-52a)}$$

$$p(Y_2 = 1 | X_0 = 1, X_1 = 1, x_2 = 1) = 0 \quad \text{(12-52b)}$$

Therefore, by (12-49), (12-50), (12-51), (12-52) and (12-41b) we conclude

$$H(Y_2 | X_0, X_1, x_2 = 1) = 0 \quad \text{for any } p(x_0, x_1) \quad \text{(12-53)}$$

(2-C) Find $I(X_0, X_1; Y_2 | x_2 = 1)$

From (12-47), (12-53) and (12-40) we conclude that

$$I(X_0, X_1; Y_2 | x_2 = 1) = 0 \quad \text{(12-54)}$$

Therefore, the upper bound of the achievable rate by (12-18) is

$$\max_{x_2\in\{0,1\}}\max_{p(x_0,x_1)}\min\{I(X_0;Y_1|X_1,x_2),I(X_0,X_1;Y_2|x_2)\}$$

$$=\max_{p(x_0,x_1)}\min\{I(X_0;Y_1|X_1,x_2=0),I(X_0,X_1;Y_2|x_2=1)\}=0 \tag{12-55}$$

(II) Calculate the upper bound of the achievable rate by (12-19)

$$R<\max_{p(x_0,x_1,x_2)}\min\{I(X_0;Y_1|X_1,X_2),I(X_0,X_1;Y_2|X_2)\} \tag{12-56}$$

where

$$I(X_0;Y_1|X_1,X_2)=I(X_0;Y_1|X_1,x_2=0)p(X_2=0)$$

$$+I(X_0;Y_1|X_1,x_2=1)p(X_2=1)=I(X_0;Y_1|X_1,x_2=1)p(X_2=1) \tag{12-57a}$$

$$I(X_0,X_1;Y_2|X_2)=I(X_0,X_1;Y_2|x_2=0)p(X_2=0)$$

$$+I(X_0,X_1;Y_2|x_2=1)p(X_2=1)=I(X_0,X_1;Y_2|x_2=0)p(X_2=0)$$

$$=\left[I(X_1;Y_2|x_2=0)+I(X_0;Y_2|X_1,x_2=0)\right]p(X_2=0) \tag{12-57b}$$

By (12-42a)~(12-42d) we can easily obtain

$$p(Y_1=0|X_0=0,X_1=0,x_2=1)=1 \tag{12-58a}$$

$$p(Y_1=0|X_0=0,X_1=1,x_2=1)=1 \tag{12-58b}$$

$$p(Y_1=1|X_0=1,X_1=0,x_2=1)=1 \tag{12-58c}$$

$$p(Y_1=1|X_0=1,X_1=1,x_2=1)=1 \tag{12-58d}$$

The probabilities (12-58a)~(12-58d) imply that

$$p(Y_1=X_0|X_1,x_2=1)=1 \tag{12-59}$$

Therefore, (12-57a) can be further simplified as

$$I(X_0;Y_1|X_1,X_2)=I(X_0;Y_1|X_1,x_2=1)p(X_2=1)$$

$$=H(X_0|X_1,x_2=1)p(X_2=1) \qquad (12\text{-}60)$$

By (12-24a), (12-24b), (12-28a) and (12-28b) we have

$$p(Y_2=0|X_0=0,X_1=0,x_2=0)=1 \qquad (12\text{-}61a)$$

$$p(Y_2=0|X_0=1,X_1=0,x_2=0)=1 \qquad (12\text{-}61b)$$

$$p(Y_2=1|X_0=0,X_1=1,x_2=0)=1 \qquad (12\text{-}61c)$$

$$p(Y_2=1|X_0=1,X_1=1,x_2=0)=1 \qquad (12\text{-}61d)$$

The probabilities (12-61a)~(12-61d) imply that

$$p(Y_2=X_1|X_0,x_2=0)=1 \qquad (12\text{-}62)$$

Thus, (12-57b) can be further simplified as

$$I(X_0,X_1;Y_2|X_2)=\left[I(X_1;Y_2|x_2=0)+I(X_0;Y_2|X_1,x_2=0)\right]p(X_2=0)$$

$$=H(X_1|x_2=0)p(X_2=0) \qquad (12\text{-}63)$$

Therefore, the upper bound of the achievable rate by (12-19) is

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

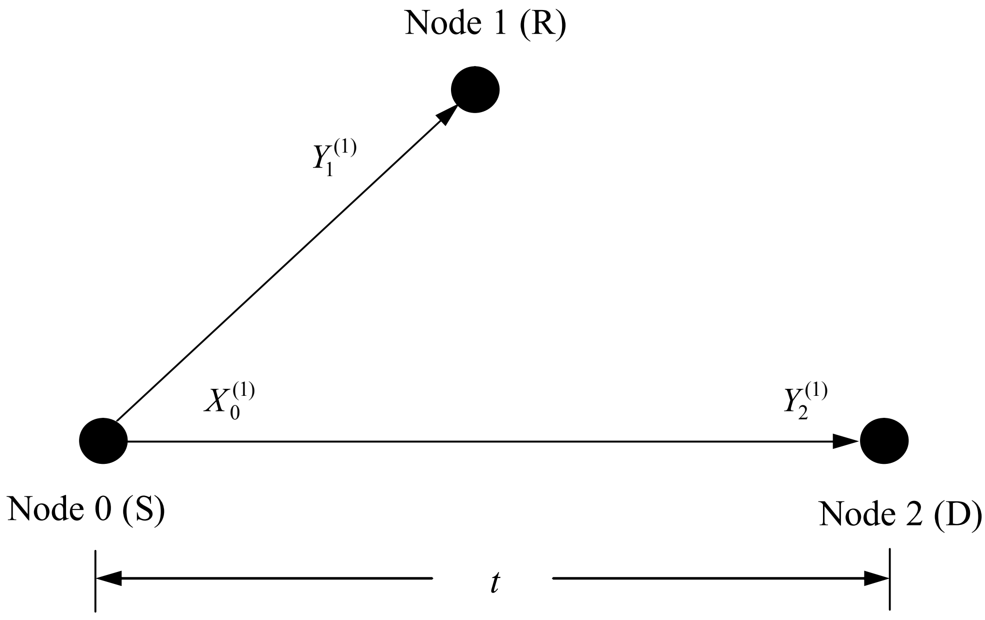
$$=\max_{p(x_0,x_1,x_2)} \min\{H(X_0|X_1,x_2=1)p(X_2=1),H(X_1|x_2=0)p(X_2=0)\}$$

$$\leq \max_{p(x_2)} \min\{p(X_2=1),p(X_2=0)\}=\frac{1}{2} \qquad (12\text{-}64)$$

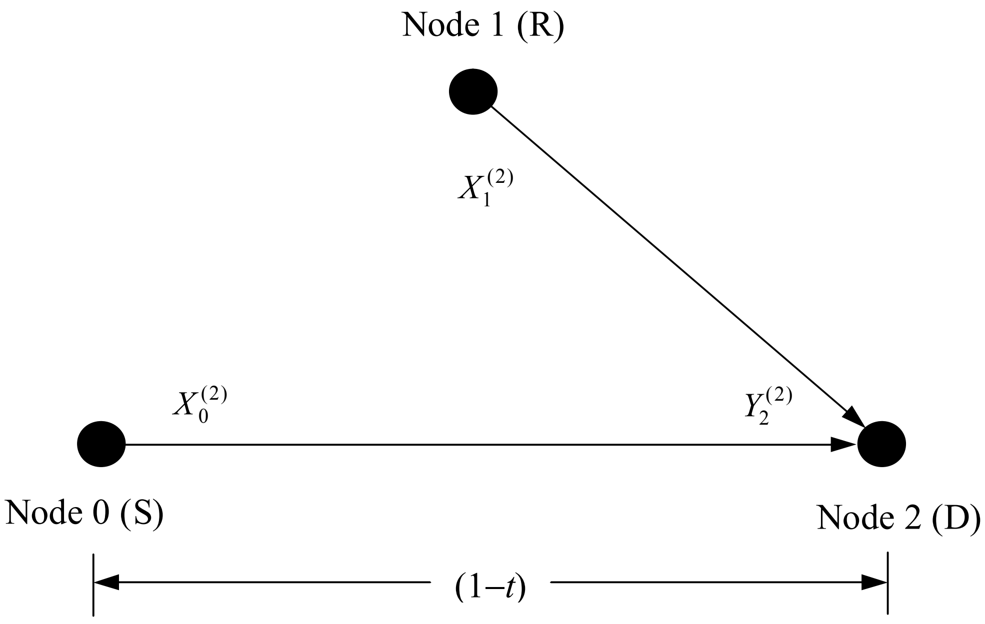
where the inequality in (12-64) is due to the binary entropy. From (12-55) and (12-64) we can see that the upper bound of achievable rate by (12-19) is larger than that by (12-18) in this example. □

12.5 Time-Division Half-Duplex One-Relay Channel

In this section we will present the applications of low density parity check (LDPC) codes to relay channel, which is particularly focused on one-relay with the time-division half-duplex relay mode. First, we will briefly revisit the fundamentals of half-duplex one-relay channel, which consists of broadcast (BC) mode and multiple-access (MAC) mode. Second, we will the structure of relay channel code with the achievable rate, including the encoding in BC and MAC modes, decoding at the end of MAC mode, LDPC codes design for relay channel with additive white Gaussian noise. Finally, the performance of LDPC codes over time-division half-duplex one-relay channel is presented.



(a) Normalized broadcast (BC) time fraction



(b) Normalized multiple-access (MAC) time fraction

Fig. 12.8. Half-duplex relay modes with (a) broadcast and (b) multiple-access time fractions.

(1) System description: In a half-duplex relay channel, the relay cannot transmit and receive simultaneously in the same band. We concentrate on time-division half-duplex relaying, where communication takes place over two time slots of normalized durations t and $t^{'}=1-t$. In the first time slot, S (node 0, source) transmit information message, which is received by both R (node 1, relay) and D (node 2, destination). We call this the broadcast (BC) mode of communication. In the second time slot, both S and R transmit information message to D. We refer to this as the multiple-access (MAC) mode. These two modes are depicted in Fig. 12.8. In this section whenever we mention the relay channel, we always refer to the time-division half-duplex relay channel. Meanwhile, we adopt the following conventions. Superscript 1 in the variables in Fig. 12.8 denotes the inputs/outputs during the broadcast time (BC) duration, while the superscript 2 denotes the inputs/outputs during the multiple-access (MAC) duration. We will use these notations in the sequel. The subscript 0, 1 and 2 denotes the source, relay and destination nodes, respectively, which have the same meaning as the previously introduced conventional one-relay channels. With the above conventions, we introduce the following channel model

$$Y_1^{(1)}=h_{SR}X_0^{(1)}+n_R^{(1)} \tag{12-65a}$$

$$Y_2^{(1)}=h_{SD}X_0^{(1)}+n_D^{(1)} \tag{12-65b}$$

$$Y_2^{(1)}=h_{SD}X_0^{(2)}+n_D^{(2)} \tag{12-65c}$$

where h_{SR} is the source-to-relay (SR) channel coefficient, and $n_R^{(1)}$ is additive Gaussian noise with zero mean and unit variance for the signal received by the relay node. Note that h_{SR} and $n_R^{(1)}$ are only considered in the broadcast duration. The SR channel gain is denoted by $\gamma_{SR} = |h_{SR}|^2$. The remaining expressions h_{SD}, $n_D^{(1)}$ and $n_D^{(2)}$ can be similarly interpreted. Note that the source-to-destination (SD) channel coefficient h_{SD} is used for both the broadcast and multiple-access durations. Also, perfect global channel knowledge is assumed at all nodes.

(2) Power constraints and component channel gains: An average global transmission power constraint is imposed on the nodes and expressed as

$$tP_S^{(1)} + t^{'}(P_S^{(2)} + P_R^{(2)}) \le P \tag{12-66}$$

where $P_S^{(1)} = E[(X_0^{(1)})^2]$ denotes the source transmission power in BC mode, $P_S^{(2)} = E[(X_0^{(2)})^2]$ and $P_R^{(2)} = E[(X_1^{(2)})^2]$ represent the source and relay transmission power in MAC mode, respectively, P represents the total system transmission power. Since additive Gaussian noise power is normalized to unit, thus P is also equivalent to the relay channel signal-to-noise ratio. A global power constraint is chosen instead of a per-node power constraint because it affords greater flexibility of power allocation and leads to higher achievable rates. Also, for fair comparison with direct communication, we ensure that the sum of the source and relay transmission powers in the relay channel equals the source transmission power for the direct link.

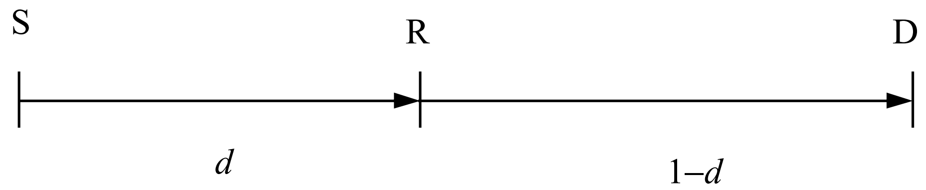


Fig. 12.9. Relay channel with source, relay and destination in a straight line.

For simplicity, the distance between source node S and destination D is normalized to unity, and relay node R is assumed to lie on the straight line joining S and D, which is shown in Fig. 12.9. The relay position, denoted by d, represents its distance from the source. In the above setting, the SD channel gain is γ_{SD}=1, the SR channel gain is γ_{SR}= $1/d^{\alpha}$, and the RD channel gain is $\gamma_{RD}=1/(1-d)^{\alpha}$, where α is the channel attenuation exponent. We use $\alpha=2$ in the following analysis.

12.6 Achievable Rate and Encoding/Decoding Structure

For the general time-division half-duplex one-relay channel, by using the Theorem 10.1 the decode-and-forward protocol can achieve the following rate

$$R=\sup_{0\le t\le 1}\min\{tI(X_0^{(1)};Y_1^{(1)})+t'I(X_0^{(2)};Y_2^{(2)}\mid X_1^{(2)}),$$

$$tI(X_0^{(1)};Y_2^{(1)})+t'I(X_0^{(2)},X_1^{(2)};Y_2^{(2)})\} \tag{12-67}$$

Note that R is absolutely less than the upper bound of the achievable rate that is obtained by maximizing over all the input probability distributions. The above mutual information expression can be evaluated for AWGN channel with binary inputs of equal probabilities, the interested reader can refer to ([85] Appendices I and II) for more details. In the following analysis we will investigate the rates of component codes in terms of mutual information, which can achieve the overall rate as expressed by (12-67).

A total of N symbols are transmitted, of which tN are transmitted in BC mode, and the rest are sent in MAC mode. The total information message (symbols or bits) from the source is divided into two independent pair (ω, v).

(1) Encoding in BC mode

In BC mode, the source encodes ω to generate a tN symbol-long codeword $c_{SR^{(1)}} \in C_{SR^{(1)}}$ with rate assigned by

$$R_{SR^{(1)}} = I(X_0^{(1)}; Y_1^{(1)}) \tag{12-68}$$

(2) Decoding the BC mode signals: The codeword $c_{SR^{(1)}}$ is corrupted by AWGN and received by both the relay R and the destination D. The relay can decode $c_{SR^{(1)}}$ reliably with error-free by using the infinite block length since the rate $R_{SR^{(1)}}$ is less than the capacity of the SR channel of BC mode. On the other hand, the direct distance between S and D is greater than the distance between S and R, thus the signal to noise ratio (SNR) at the destination end is lower than that at the intermediate relay if the source transmits the same power to both. As a consequence, the capacity of the SD link is less than that of the SR link, and the destination cannot decode the codeword $c_{SR^{(1)}}$ reliably with error-free even if the infinite block length is employed. Therefore, the destination only stores the received codeword for further decoding at the end of MAC mode with the cooperation from others.

(3) Encoding in MAC mode: The destination already has $tNI(X_0^{(1)};Y_2^{(1)})$ information bits in the form of the undecodable codeword $c_{SR^{(1)}}$ corrupted by noise. However, it still need an additional tN ($I(X_0^{(1)};Y_1^{(1)}) - I(X_0^{(1)};Y_2^{(1)})$) bits to reliably decode $c_{SR^{(1)}}$ [86]. These additional bits used to reliably decode $c_{SR^{(1)}}$ are jointly transmitted in MAC mode by the source S and the relay R in a codeword $c_{RD^{(2)}} \in C_{RD^{(2)}}$ of rate

$$R_{RD^{(2)}} = \frac{t}{t'}(I(X_0^{(1)};Y_1^{(1)}) - I(X_0^{(1)};Y_2^{(1)})) \tag{12-69}$$

The second part of the information message containing v bits is also transmitted in MAC mode using a codeword $c_{SD^{(2)}} \in C_{SD^{(2)}}$ to utilize the remaining capacity if the multiple access channel, which consists of the source and relay as the two transmitters and the destination S as the receiver. This information bits is transmitted by the source alone in MAC mode since the relay does not have access to new information. The amount of new information is bounded by the capacity region [86] of the multiple-access channel, and this information is therefore sent at a rate [85][86]

$$\begin{aligned} R_{SD^{(2)}} = \min\{ & I(X_0^{(2)}, X_1^{(2)};Y_2^{(2)}) - (t/t')[I(X_0^{(1)};Y_1^{(1)}) - I(X_0^{(1)};Y_2^{(1)})], \\ & I(X_0^{(2)};Y_2^{(2)} \mid X_1^{(2)}) \} \end{aligned} \tag{12-70}$$

(4) Decoding at the end of MAC mode:

The destination S first decodes the codewords $c_{RD^{(2)}}$ and $c_{SD^{(2)}}$ transmitted in MAC mode. These codes are transmitted at rates within the capacity of the MAC. It was argued in [85] that the source and relay rates in MAC mode correspond to a point on the multi-access capacity region that can be achieved by a pair of single-user codes.

After decoding $c_{RD^{(2)}}$ and $c_{SD^{(2)}}$, the destination S can decode the corrupted codeword $c_{SR^{(1)}}$ from BS mode using the information carried by $c_{RD^{(2)}}$ as side information. For decoding $c_{SR^{(1)}}$, the destination regards it as a codeword $c_{SD^{(1)}} \in C_{SD^{(1)}}$ of rate

$$R_{SD^{(1)}} = I(X_0^{(1)}; Y_2^{(1)}) \qquad (12\text{-}71)$$

which is less than $R_{SR^{(1)}}$ given by (12-68) due to the side information carried by $c_{RD^{(2)}}$. For example, if all codes were binary codes, then the information bits in $c_{RD^{(2)}}$ would act as additional parity information bits for $c_{SR^{(1)}}$.

12.7 LDPC Codes for the Time-Division Half-Duplex One-Relay Channel

(1) Simplifications for practical code design:

In [85] the following three observations are proposed based on the approximated simulations to simplify the practical code design.

First, the gain due to relaying is most prominent at low SNRs and negligible at high SNRs, which motivates the use of binary modulation.

Second, although there is an optimal value of correction between source and relay codewords in MAC mode that may maximize the achievable rate for Gaussian relay channels, in [85] the authors observe that the codebooks can be either completely correlated or completely independent without significant rate loss.

Last, the authors argue that the source and relay rates in MAC mode correspond to a point on the capacity region of the multiple-access channel that can be achieved by a successive interference canceling decoder at the destination.

(2) Motivation for binary coding:

At high SNR, the achievable rate of decode-and-forward relaying on Gaussian channel exceeds that of direct communication only by a constant independent of the SNR ([85] Appendix II). Thus, the ratio of the decode-and-forward rate to the single-user capacity approaches unity at high SNR. The relaying gain is maximum at low SNR, where binary modulation for each channel dimension is near optimum.

Fig. 12.10 compares the relay rate with that of direct communication for Gaussian as well as BPSK signaling when relay is in the middle between the node and destination, i.e., d =0.5. The results are in agreement with the claim [85] that relaying is most beneficial in the low SNR regime, where BPSK can achieve a significant fraction for the AWGN relay rate.

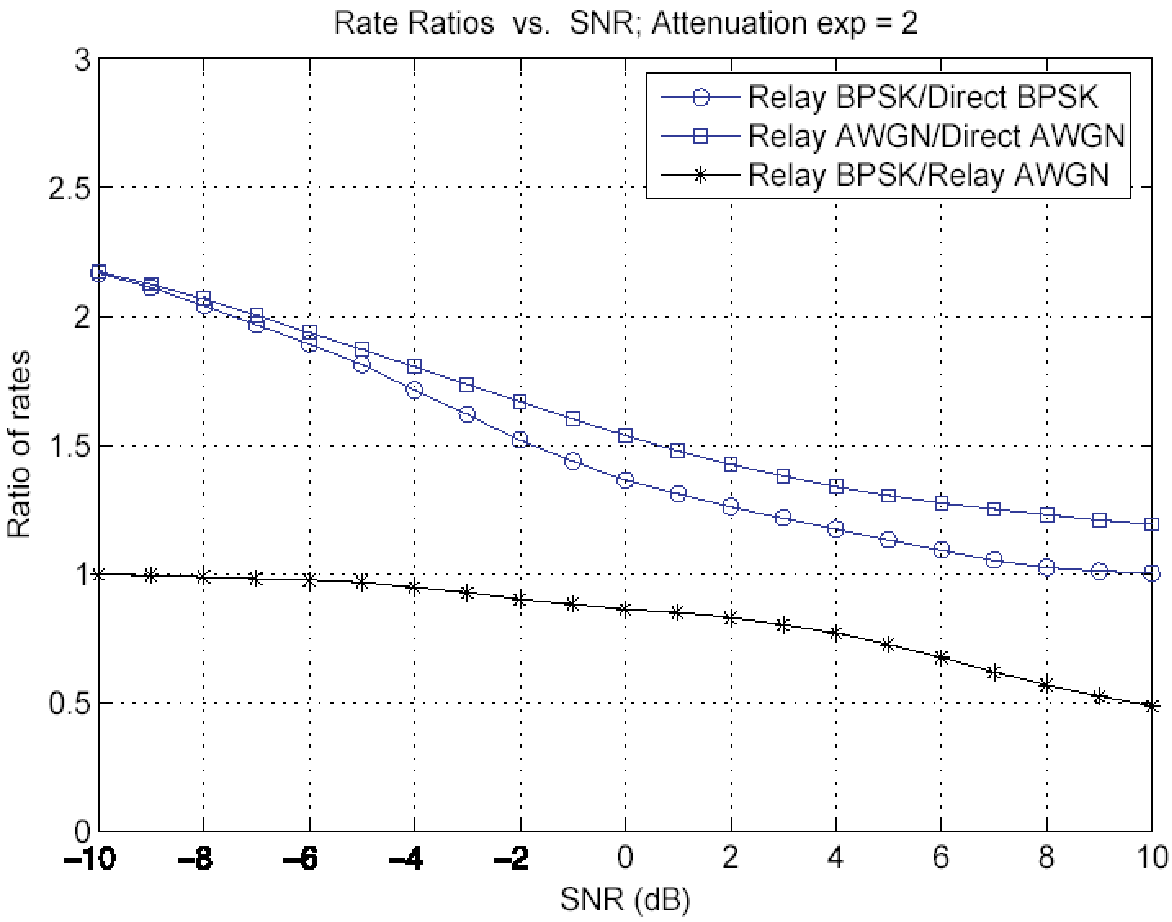


Fig. 12.10. Ratios of rates vs. SNR (dB) for direct and relay channels.

(2) MAC mode correlation

The achievable rate if decode-and-forward relay over AWGN links is maximized when the source and relay codebooks (codewords sets) are optimally correlated in MAC mode (85 Appendix). The correlation r reflects the fact that the source and the relay have common information since the relay has decoded the w portion of the total information in BC mode.

Any correlation between source and relay transmissions in MAC mode can be achieved using a pair of independent binary codewords. The details for construction of such codewords, the interested reader can refer to [82] regarding the finite length ε-typical semirandom sequences.

Let $c_{RD^{(2)}}$, $c_{SD^{(2)}}$ be a pair of binary codewords from independent codebooks $C_{RD^{(2)}}$ and $C_{SD^{(2)}}$, respectively. If the relay sends $C_{RD^{(2)}}$, and the source sends $rc_{RD_2}+(1-r)c_{SD_2}$, then the transmissions have correlation r. In the decoding the destination can proceed as if there were two transmitters sending their independent codewords $c_{RD^{(2)}}$ and $c_{SD^{(2)}}$ with transmission powers appropriately adjusted.

It is clear that r=0, 1 are two fundamental extremes, and all intermediate correlations can be achieved by superposing the two in the proper proportion. More specifically, when r=1, the source and relay send identical information in MAC mode, as a result the source send nothing new and there is no code $C_{SD^{(2)}}$. For r=0, the source and the relay send independent information in MAC mode, which implies that the source sends only new information through the codeword $c_{SD^{(2)}}$, whereas the relay alone sends $c_{RD^{(2)}}$ to assist the destination decode the BC mode codeword c_{SR_1}.

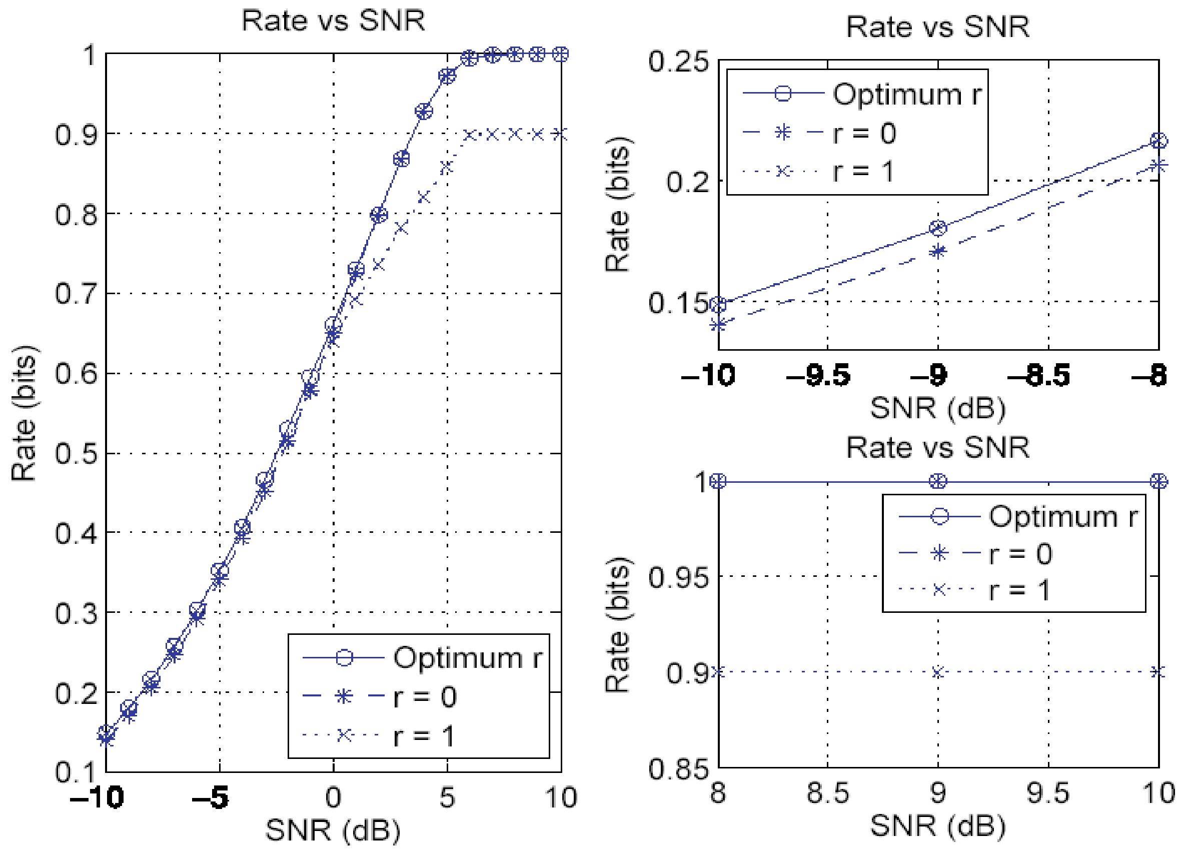


Fig. 12.11. Achievable rates for the relay channel with r=0, 1. and the optimum value of r (d=0.5). The subplots on the right show the behavior at low and high SNRs.

An interesting observation is carried out to observe the relationship between the achievable rate and r. In particular, we choose r=0, 1 and study the rate loss in comparison to a system that uses optimum correlation. The effect of correlation on the achievable rate is given in Fig. 12.11 by simulation, which essentially shows that there is no significant rate loss for two extremes and the optimum r over a wide SNR range, from −10dB to +10dB. Even for high SNR range the loss is still negligible. That is the crucial reason that leads to the design of the error-correction codes restricted in the case of r=0, 1.

(3) Optimality of successive decoding in MAC mode:

In the previous discussion of the information theoretic relay coding scheme, the source transmits additional information v at a rate, which is indicated by (12-70) within the capacity of the multiple-access channel. This rate is the smaller of $I(X_0^{(2)};Y_2^{(2)} \mid X_1^{(2)})$ and $I(X_0^{(2)},X_1^{(2)};Y_2^{(2)})-(t/t^{'})[I(X_0^{(1)};Y_1^{(1)})-I(X_0^{(1)};Y_2^{(1)})]$. For decode-and-forward relaying, it has been proved [85] that the best choice of t equates the two arguments of the min(.) function in (12-70). Therefore, from this point on, we have

$$\begin{aligned} R_{SD^{(2)}} &= I(X_0^{(2)},X_1^{(2)};Y_2^{(2)})-(t/t^{'})[I(X_0^{(1)};Y_1^{(1)})-I(X_0^{(1)};Y_2^{(1)})] \\ &= I(X_0^{(2)};Y_2^{(2)} \mid X_1^{(2)}) \end{aligned} \tag{12-72}$$

12.8 Encoding and Decoding Configurations for LDPC Codes in BS and MAC Modes

We are now ready to present the binary irregular LDPC coding schemes for r =0, 1. The two schemes differ only in MAC mode. We first describe the scenario for the full r=1, and then explain what is the difference for r=0. The detailed code design strategy [85] is described as follows:

(1). LDPC Code Design Strategy for r=1:

In BC mode, the source S uses an LDPC code $C_{SR^{(1)}}$ with an $Nt(1-R_{SR^{(1)}})\times Nt$ parity check matrix to transmit the codeword that conveys NtR_{SR_1} the information bits. The relay R decodes this codeword. Note that for the practical finite block length the relay cannot decode the codeword with error-free. The destination S stores it for further decoding. In MAC mode, S and R use the BC mode codeword as the basis for cooperation. Both nodes multiply the BS mode codeword with an $Nt\,(R_{SR^{(1)}}-R_{SD^{(1)}})\times Nt$ matrix to generate Nt $(R_{SR^{(1)}}-R_{SD^{(1)}})$ additional parity-check bits as side information to help the destination. These additional parity bits are then coded using an LDPC code $C_{RD^{(2)}}$ with an $Nt^{'}(1-R_{RD^{(2)}})\times Nt^{'}$ parity-check matrix and transmitted in a synchronized manner to the destination. These bits communicated by $C_{SD^{(1)}}$, in addition to the parity check bits of the original code $C_{SR^{(1)}}$, form a code $C_{SD^{(1)}}$ of lower rate $R_{SD^{(1)}}$ that can be decodable by the destination D. Fig. 12.12 shows the LDPC code structure [85] for $r=1$ with the absence of code $C_{SD^{(2)}}$.

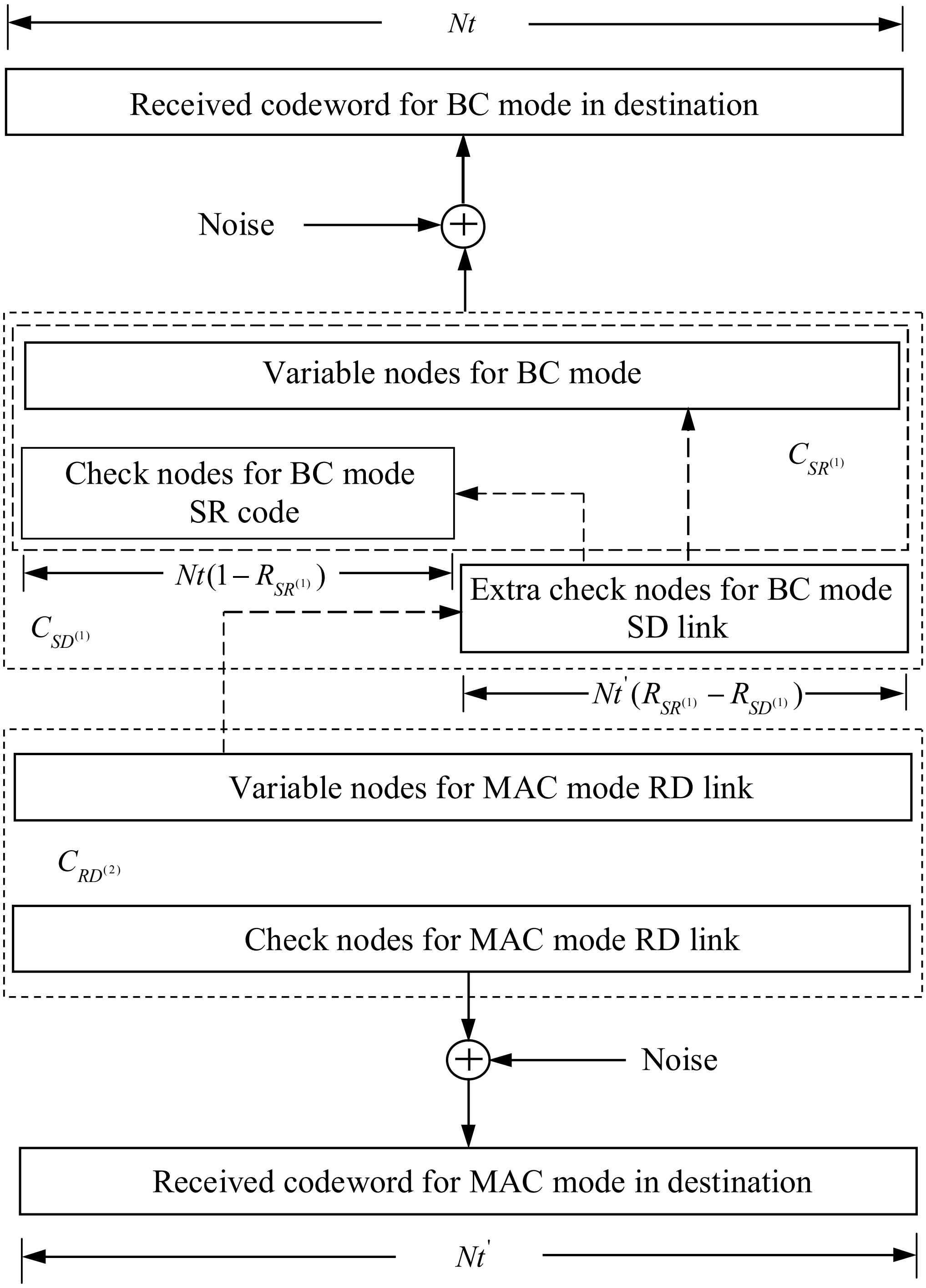


Fig. 12.12. LDPC code structure for r=1.

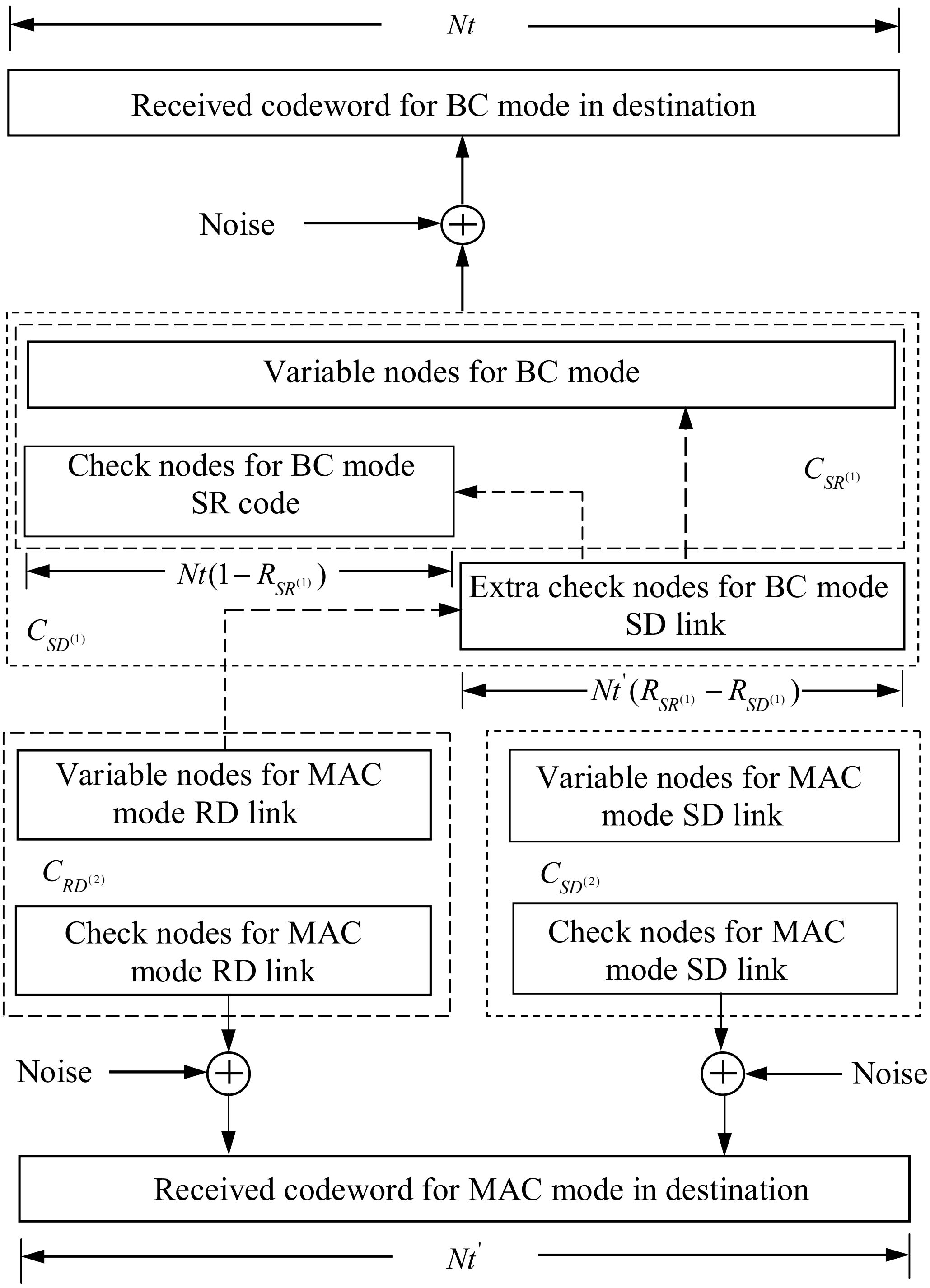


Fig. 12.13. LDPC code structure for r=0.

(2). LDPC Code Design Strategy for *r*=0:

The BC mode remains unchanged. In MAC mode, the source and relay send independent information. The relay first generates additional parity bits from the BC mode codeword by multiplying with a $Nt(R_{SR^{(1)}} - R_{SD^{(1)}}) \times Nt$ matrix, which is the same as for *r*=1. Then it uses an LDPC code $C_{RD^{(2)}}$ with an $Nt'(1-R_{RD^{(2)}}) \times Nt'$ parity check matrix to encode and transmit the additional parity information in MAC mode. The information carrier by this codeword enables decoding of the BC mode codeword at the end of MAC mode. The source, in MAC mode, uses an LDPC code $C_{SD^{(2)}}$ with an $Nt'(1-R_{SD^{(2)}}) \times Nt'$ parity-check matrix to send new information to the destination. At the end of MAC mode, D uses successive decoding to recover both the additional parity information and new source information. Finally, the additional parity bits received in MAC mode are employed to decode the received BC mode codeword at the destination D. Fig. 12.13 offers the LDPC code structure [85] for *r*=0 with independent parallel transmissions in MAC mode.

12.9 Performance of LDPC codes in Time-Division Half-Duplex One-Relay Channel with AWGN

After using the Gaussian approximation and linear optimization, introduced in the previous chapter, for the LDPC codes employed in BC and MAC modes, namely the codes $C_{SR^{(1)}}$, $C_{RD^{(2)}}$ and $C_{SD^{(2)}}$, we present the simulation results for LDPC codes over the time-division half-duplex one-relay channel, where the total transmission power P=−1dB and t=0.6 yields the best rate [85]. The BS codeword ($C_{SR^{(1)}}$) is chosen as 1.3×10^5 bits and the MAC mode codewords ($C_{RD^{(2)}}$ and $C_{SD^{(2)}}$) are selected as 0.7×10^5 bits long. The results of rate vs. SNRs are shown in Fig. 12.14 for 300 decoding iterations. We can clearly observe that the rates are basically the same for r=0, 1 for various SNRs and also comparable with the optimum r, which are all significantly large than the direct link without relay.

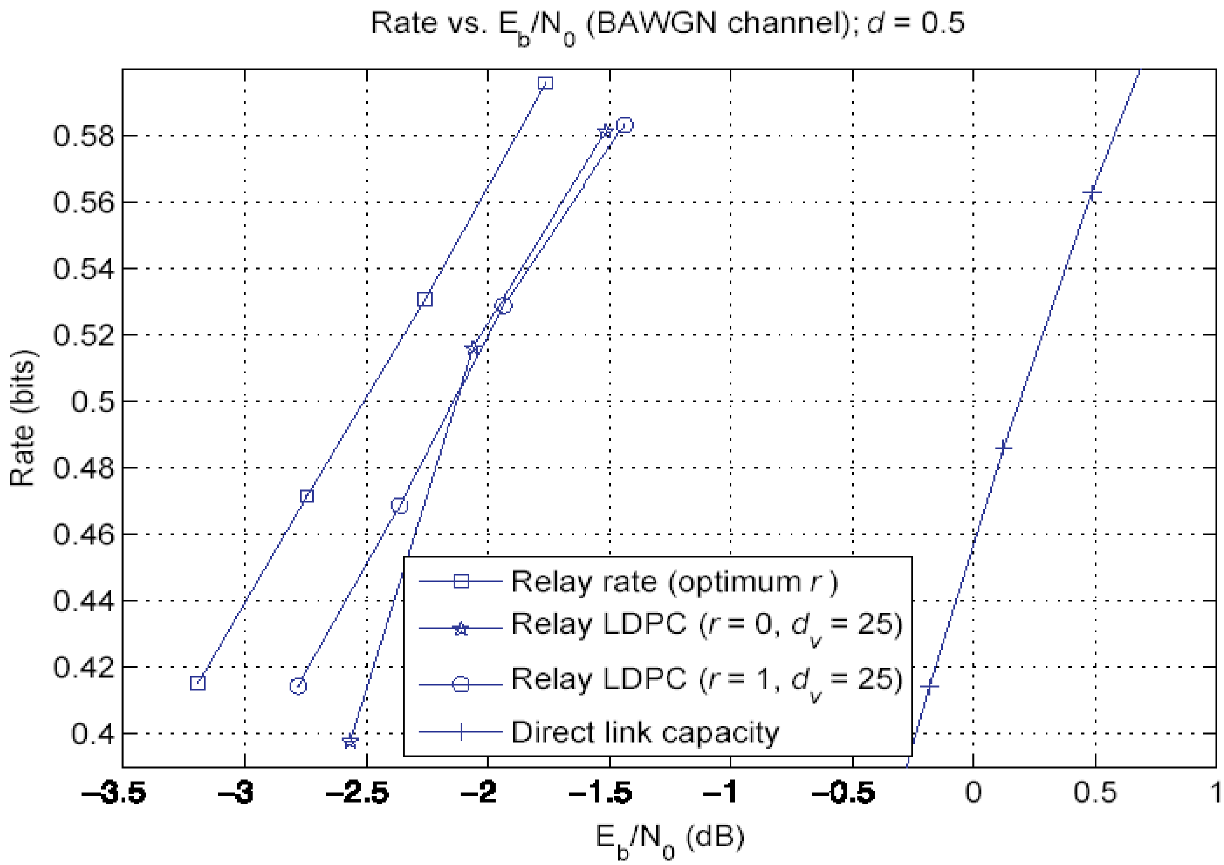


Fig. 12.14. Rate vs. E_b/N_0 (d=0.5, d_v=25) for r=0, 1 and the optimum value, the rate for direct link without relay is also plotted for comparison.

Fig. 12.15 offers the bit error rate (BER) vs. SNR (E_b / N_0) for each of the three component LDPC codes. The gap to the asymptotic threshold is nearly 1dB for the code $C_{SD^{(1)}}$ in MAC mode, while the codes $C_{RD^{(2)}}$ and $C_{SR^{(1)}}$ have significantly less gaps to their limits due to the optimization by the concentrated check degrees.

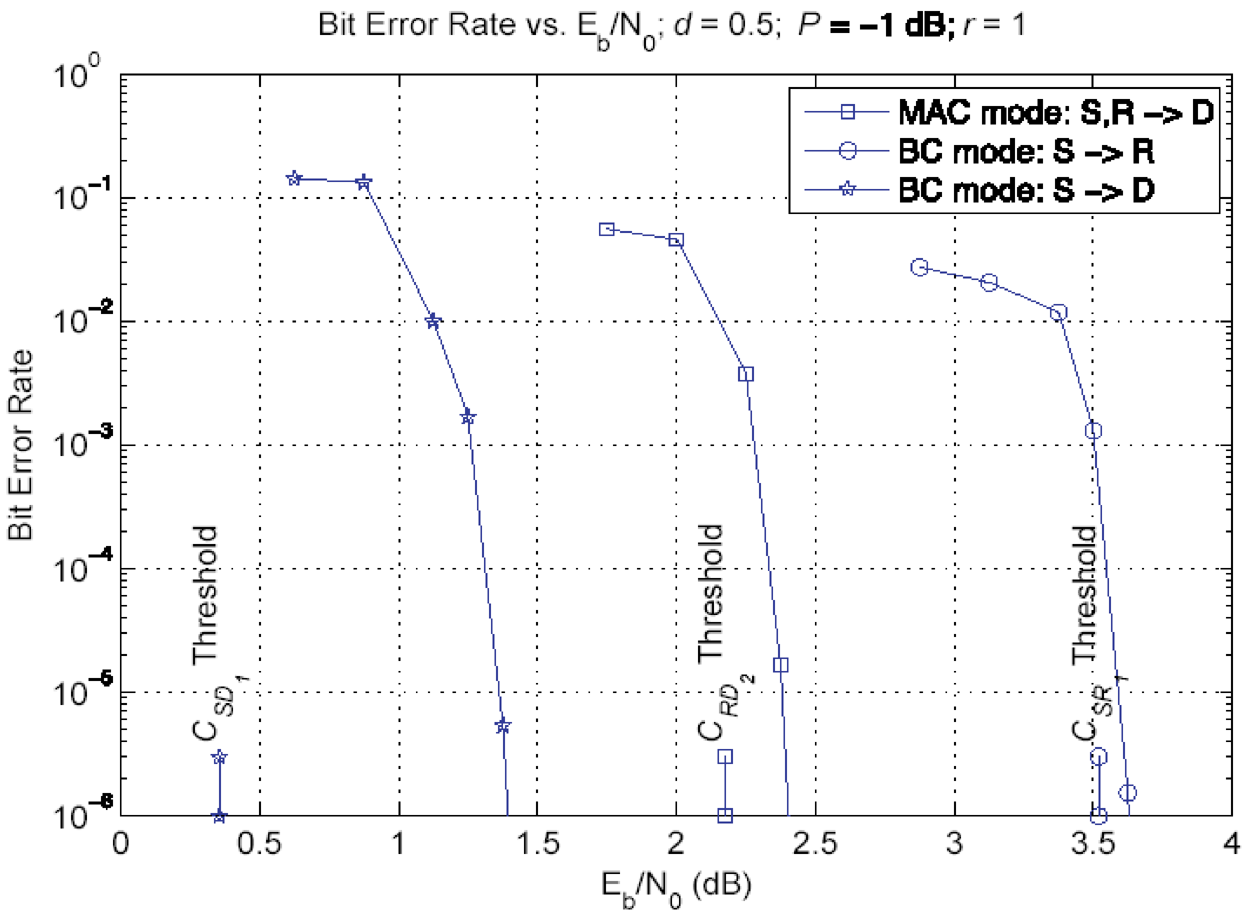


Fig. 12.15. Rate vs. E_b / N_0 for component LDPC codes ($P = -1$dB, r =1). Theoretical limits are indicated for comparison.

For the case of r=0, the BER performance of the component codes similar to that of the r=1 codes. The only difference is that there are two codes $C_{SD^{(2)}}$ and $C_{RD^{(2)}}$ in MAC mode, and the latter is decoded by treating the $C_{SD^{(2)}}$ as interference. Fig. 12.16 shows the BER performance of decoding $C_{RD^{(2)}}$ in the presence of interference that is modeled by a random bits sequence [85]. Once $C_{RD^{(2)}}$ has been decoded and subtracted out, the BER performance of $C_{SD^{(2)}}$ will be the same as that of a single-user code in Gaussian noise.

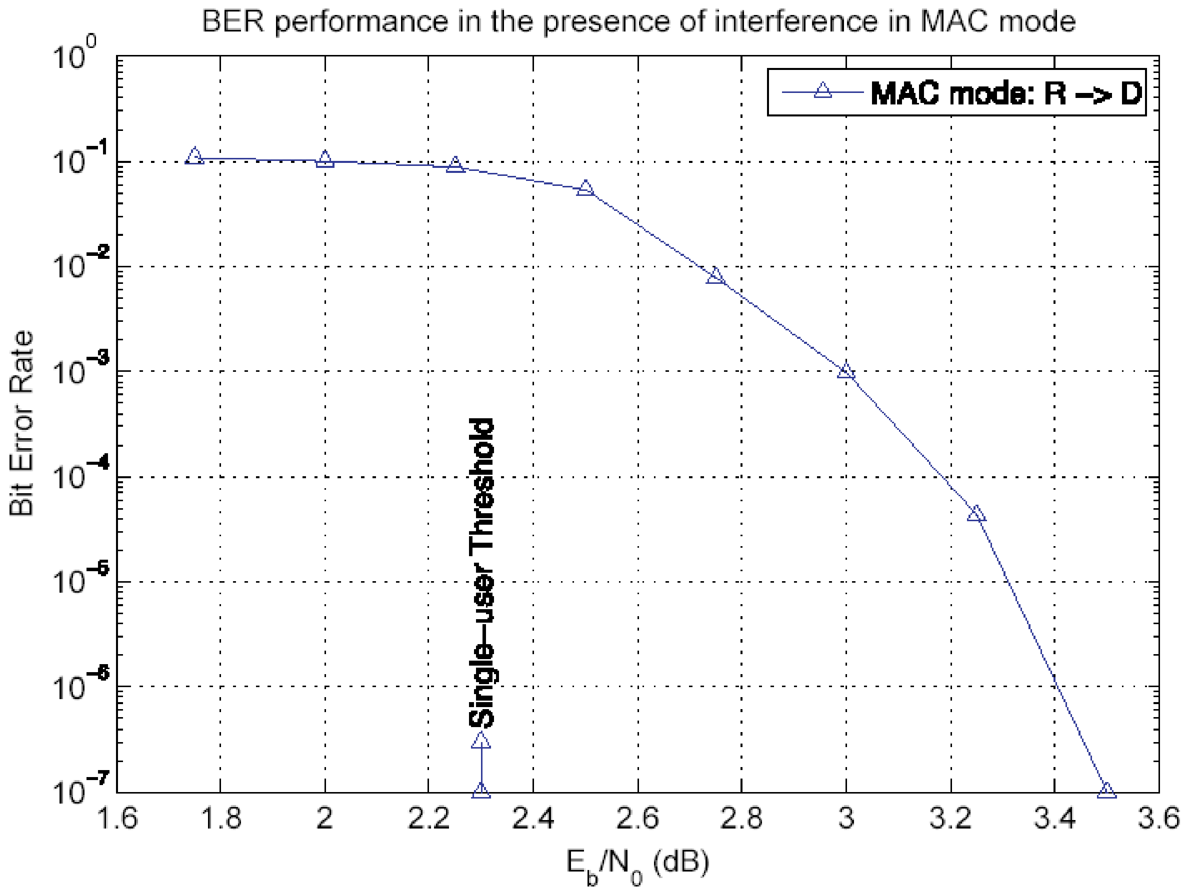


Fig. 12.16. Performance of the LDPC code $C_{RD^{(2)}}$ in MAC mode after 100 decoding iterations with interference from $C_{SD^{(2)}}$ ($P = -1$dB, $r = 1$). The single-user noise threshold is also indicated for comparison.

In summary, in the proposed relay coding scheme the relay can forward information only after successful decoding, which requires the entire codeword to be correct. Therefore, the component codes must have excellent error performance. This problem can be solved either by optimizing the parity check matrix to eliminate cycles in the graph, or by optimizing the variable/check node degree distributions. Decode-and-forward is only one relay protocol, code design for other efficient protocols, such as estimate-and-forward [81][85], are still required further investigations.

References

[1] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.

[2] R. M. Tanner, "A recursive approach to low complexity codes," IEEE Trans. Inform. Theory, vol. IT-27, pp. 533-547, Sept. 1981.

[3] V. Zablov and M. Pinsker, "Estimation of the error-correction complexity of Gallager low-density codes," *Probl. Pered.* Inform., vol. 11, pp. 23-26, jan. 1975.

[4] G. A. Margulis, "Explicit construction of graphs without short cycles and low density codes," Combinatorica, vol. 2, No. 1, pp. 71-78, 1982.

[5] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," Electric Lett., vol. 32, pp. 1645-1646, Aug. 1996.

[6] D. J. C. MacKay, "Good Error-correction Codes Based on Very Sparse Matrices," IEEE Trans. Inform. Theory, vol. 45, No. 2, pp. 399-431, Mar. 1999.

[7] T. J. Richardson and Rudiger L. Urbanke, "Efficient Encoding of Low-Density Parity-Check Codes," IEEE Trans. Inform. Theory, vol. 47, No. 2, pp. 638-656, Feb. 2001.

[8] T. J. Richardson, M. Amin Shkrollahi, and Rudiger L. Urbanke, "Design of Capacity-Approaching Irregular Low-Density Parity-Check Codes," IEEE Trans. Inform. Theory, vol. 47, No. 2, 619-637, Feb. 2001.

[9] T. J. Richardson and Rudiger L. Urbanke, "The Capacity of Low-Density Parity-Check Codes Under Message-Passing Decoding, "IEEE Trans. Inform. Theory, vol. 47, No. 2, pp. 599-618, Feb. 2001.

[10] C. Berrou, A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: Turbo Codes," IEEE Trans. Commun., vol. 44, No. 10, pp. 1261-1271, Oct. 1996.

[11] R. G. Gallager, "Low-Density Parity-Check Codes," IRE Transactions on Information Theory, pp. 21-28, Jan. 1962.

[12] S. Young Chung, G. D. Forney, T. J. Richardson and R. Urbanke, "On the Design of Low-Density Parity-Check Codes within 0.0045dB of the Shannon Limit," IEEE Commun., Lett. Vol. 5, No. 2, pp. 58-60, Feb. 2001.

[13] D. J. C. Mackay, S. T. Wilson and M. C. Davey, "Comparison of Constructions of Irregular Gallager Codes," IEEE Trans. Commun., vol. 47, No. 10, pp. 1449-1454, Oct. 1999.

[14] R. M. Tanner, "Minimum-Distance Bounds by Graph Analysis," IEEE Trans. Inform. Theory, vol. 47, No. 2, pp. 808-821, Feb. 2001.

[15] G. D. Forney, JR., "Codes on Graphs: Normal Realizations," IEEE Trans. Inform. Theory, vol. 47, No. 2, pp. 520-548, Feb. 2001.

[16] S. Y. Chung, T. J. Richardson and R. L. Urbanke, "Analysis of Sum-product Decoding of Low-Density Parity-Check Codes Using a Gaussian Approximation," IEEE Trans. Inform. Theory, vol. 47, No. 2, pp. 657-670, Feb. 2001.

[17] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi and D. A. Spielman, "Improved Low-Density Parity-Check Codes Using Irregular Graphs," IEEE Trans. Inform. Theory, vol. 47, No. 2, pp. 585-598, Feb. 2001.

[18] G. Miller, D. Burshtein, "Bounds on the Maximum-Likelihood Decoding Error Probability of Low-Density Parity-Check Codes," IEEE Trans. Inform. Theory, vol. 47, No. 7, pp. 2696-2710, Nov. 2001.

[19] B. M. Kurkoski, P. H. Siegel and J. K. Wolf, "Joint Message-Passing Decoding of LDPC Codes and Partial-Response Channel," IEEE Trans. Inform. Theory, vol. 48, No. 6, pp. 1410-1422, June 2002.

[20] D. Burshtein, M. Krivelevich, S. Litsyn and G. Miller, "Upper Bound on the Rate of LDPC Codes," IEEE Trans. Inform. Theory, vol. 48, No. 9, pp. 2437-2449, Sept. 2002.

[21] M. Luby, M. Mitzenmacher, A. Shokrollahi, D. Spielman and V. Stemann, "Practical loss-resilient codes," in Proc. 29th Anun., ACM Symp. Theory of Computing, pp. 150-159, 1997.

[22] R. J. McEliece, D. J. C. MacKay and J. F. Cheng, "Turbo Decoding as an Instance of Pearl's "Belief Propagation" Algorithm," IEEE Journal on Selected Areas in Communications. vol. 16, No. 2, pp. 140-152, Feb. 1998.

[23] F. R. Kschischang and B. J. Frey, "Iterative Decoding of Compound Codes by Probability Propagation in Graphical Models," IEEE Journal on Selected Areas in Communications, vol. 16, No. 2, pp. 219-230, Feb. 1998.

[24] B. J. Frey and F. R. Kschischang, “Early Detection and Trellis Splicing: Reduced-Complexity Iterative Decoding,” IEEE Journal on Selected Areas in Communications, vol. 16, No. 2, pp. 153-159, Feb. 1998.

[25] F. R. Kschischang, B. J. Frey and H. A. Loeliger, “Factor Graphs and the Sum-Product Algorithm,” IEEE Trans. on Inform. Theory, vol. 47, No. 2, pp. 498-519, Feb. 2001.

[26] Li Ping and K. Y. Wu, “Concatenated Tree Codes: A Low-Complexity, High-Performance Approach,” IEEE Trans. on Inform. Theory, vol. 47, No. 2, pp. 791-799, Feb. 2001.

[27] Lin Ping and W. K. Leung, “Decoding Low Density Parity Check Codes with Finite Quantization Bits,” IEEE Communications Letters, vol. 4, No. 2, pp. 62-64, Feb. 2000.

[28] D. Spielman, “Linear-Time Encodeable and Decodable Error-Correcting Codes,” IEEE Trans. Inform. Theory, vol. 42, pp. 1723-1731, Nov. 1996.

[29] J. Whittaker, Graphical Models in Applied Multivariable Statistics. Chichester, UK.: Wiley, 1990.

[30] F. V. Jensen, An Introduction to Bayesian Networks. New York: Spring-Verlag, 1996.

[31] G. Chartrand and L. Lesniak, Graphs & Digraphs, Third Edition, Boca Raton, London, New York and Washington, D.C, Chapman & Hall/CRC, 1996.

[32] Vera Pless, Introduction to the Theory of Error-Correcting Codes, New York, Chichester, Weinheim, Brisbane, Singapore, Toronto, A Wiley-Interscience Pulication JOHN WILEY & SONS, INC.

[33] R. M. Pyndiah, "Near-Optimal Decoding of Product Codes: Block Turbo Codes," IEEE Trans. on Commun., vol. 46, No. 8, pp. 1003-1010, Aug. 1998.

[34] S. Hirasawa, M. Kasahara, Y. Sugiyaha and T. Noamekawa, "Modified Product Codes," IEEE Trans. on Inform. Theory, vol.. IT-30, No. 2, pp. 299-306, Mar. 1984.

[35] H. Nickl, J. Hagenauer, and Burkert, "Approaching Shannon's capacity limit by 0.27dB using simple Hamming codes," IEEE Commun., Lett., vol. 1, pp. 130-132, Sept. 1997.

[36] D. Chase, "A Class of Algorithms for Decoding Block Codes with Channel Measurement Information," IEEE Trans. Inform. Theory, vol. IT-18, pp. 170-182, Jan. 1972.

[37] S. Benedetto D. Divsalar G. Montorsi and F. Pollara, "A Soft-Input Soft-Output APP Module for Iterative Decoding of Concatenated Codes," IEEE Commun., Lett., vol. 1, No. 1, pp. 22-24, Jan. 1997.

[38] J. Pear, "Fusion propagation and structuring in belief networks," Artif. Intell., vol. 29, pp. 241-288, 1986.

[39] L. Ping, S. Chen, and K. L. Yeung, "Iterative Decoding of Multidimensional Concatenated Signal Parity Check Codes," in Proc. IEEE Int. Communications, Conf. (ICC'98), pp. 131-135, June 1998.

[40] D. Divsalar and F. Pollara, "Multiple Turbo Codes," in Proc. IEEE MILCOM'95, vol. 1, pp. 279-285, 1995.

[41] M. Sipser and D. A. Spielman, "Expander Codes," IEEE Trans. on Inform. Theory, vol. 42, No. 6, pp. 1710-1722, Nov. 1996.

[42] M. P. C. Fossorier, M. Mihaljevic and H. Imai, "Reduced Complexity Iterative Decoding of Low-Density Parity-Check Codes Based on Belief Propagation," IEEE Trans. on Commun., vol. 47, no. 5, pp. 673-680, May 1999.

[43] R. Lucas, M. Fossorier, Y. Kou and S. Lin, "Iterative Decoding of One-Step Majority Logic Decodable Codes Based on Belief Propagation," IEEE Trans. Commun., vol. 48, no. 6, pp. 931-937, Jun. 2000.

[44] S. Litsyn and V. Shevelev, "On Ensembles of Low-Density Parity-Check Codes: Asymptotic Distance Distributions," IEEE Trans. on Inform. Theory, vol. 48, No. 4, pp. 887-908, April 2002.

[45] J. Hagenauer, E. Offer and L. Papke, "Iterative Decoding of Binary Block and Convolutional Codes," IEEE Trans. on Inform. Theory, vol. 42, No. 2, pp. 429-445, Mar. 1996.

[46] J. Chen, A. Dholakia, E. Eleftheriou, M. P. C. Fossorier and Xiao-Yu Hu, "Reduced -Complexity Decoding of LDPC Codes," IEEE Trans. on Commun., vol. 53, No. 8, pp. 1288−1299, August, 2005.

[47] A. Bag and G. David Forney, Jr., "Random Codes: Minimum Distances and Error Exponents," IEEE Trans. on Inform. Theory, vol. 48, No. 9, pp. 2568-2573, Sept. 2002.

[48] M. R. YazdaNi, S. Hemati and A. H. Banihashemi, "Improving Belief Propagation on Graphs with Cycles," IEEE Commun. Lett., vol. 8, no. 1, pp. 57-59, Jan. 2004.

[49] E. Eleftheriou, T. Mittelholzer and A. Dholakia, "Reduced-Complexity Decoding Algorithm for Low-Density Parity-Check Codes," IEE Electron Lett., vol. 37, pp. 102-104, Jan. 2001.

[50] A. Anastasopoulos, "A Comparison between the Sum-Product and the Min-Sum Iterative Detection Algorithms Based on Density Evolution," in Proc. IEEE Globecom, San Antonio, TX, Nov. 2001, pp. 1021−1025.

[51] J. Erfanian, S. Pasupathy and G. Gulak, "Reduced-Complexity Symbol Detectors with Parallel Structures for ISI Channels," IEEE Trans. Commun., vol. 42, No. 2-4. pp. 1661-1671, Feb.-Apr. 1994.

[52] J. Chen, and M. P. C. Fossorier, "Density Evolution for Two Improved BP-Basd Decoding Algorithms of LDPC Codes," IEEE Commun., Lett. Vol. 6, No. 5, pp. 208-210, May 2002.

[53] J. Chen, and M. P. C. Fossorier, “Near-Optimum Universal Belief-Propagation-Based Decoding of Low-Density Parity-Check Codes,” IEEE Trans. Commun., vol. 50, No. 3, pp. 406-414, Mar. 2002.

[54] Y. C. He, H. P. Li, S. H. Sun and L. Li, “Threshold-Based Design of Quantized Decoder for LDPC codes,” in Proc. IEEE *Int. Symp. Inf. Theory*, Yokohama, Japan, Jan.-Jul. 2003, p.149.

[55] H. Sankar and K. R. Narayanan, “Memory-Efficient Sum-Product Decoding of LDPC Codes,” IEEE Trans. on Commun., vol. 52, No. 8, pp. 1225-1230, Aug. 2004.

[56] M. M. Mansour and N. R. Shanbhag, “Memory-Efficient Turbo Decoder Architectures for LDPC Codes,” in Proc. IEEE Workshop Signal Processing Systems, pp. 159-164, Oct. 2002.

[57] M. Reza Soleymani, Y. Gao and U. Vilaipornsawai, Turbo Coding for Satellite and Wireless Communications, Kluwer Academic Publisher, 2002.

[58] K. Sripimanwat, Turbo Codes Applications: A Journey from a Paper to Realization, Springer 2005.

[59] M. Ardakani and Frank R. Kschischang, “Gear-Shift Decoding,” IEEE Trans. on Commun., vol. 54, No. 7, pp. 1235-1242, July 2006.

[60] L. Bazzj, T. J. Richardson and R. L. Urbanke, “Exact Threshold and Optimal Codes for the Binary-Symmetry Channel and Gallager’s Decoding Algorithm A,” IEEE Trans. on Inform. Theory, vol. 59, No.9, pp. 2010-2022, Sept. 2004.

[61] Irwin Miller and Marylees Miller, John E. Freund's Mathematical Statistics with Applications, Seventh Edition, Prentice Hall, 2004.

[62] H. Pishro-Nik and F. Fekri, "On Decoding of Low-Density Parity-Check Codes over the Binary Erasure Channel," IEEE Transactions on Inform. Theory, vol. 50, No. 3, pp. 439–453, Mar. 2004.

[63] P. Henrici, Applied and Computational Complex Analysis. New York: Wiley, 1974, vol. 1.

[64] David. G. Luenberger, "Linear and Nonlinear Programming, Second Edition, Ontario, Sydney, Addison-Wesley Publisher Company, 1984.

[65] Saul I. Gass, Linear Programming: Methods and Applications, Third Edition, McGraw-Hill Book Company, New York, Toronto, 1969.

[66] A. Schrijver, Theory of Linear and Integer Programming. New York; Wiley, 1986.

[67] J. M. Wozencraft and I. M. Jacobs, "Principles of Communication Engineering. New York: Wiley, 1965.

[68] I. S. Gradshtein and I. M. Ryzhik, Table of Integrals, Series, and Products. New York: Academic, 1965.

[69] S-Y. Chung, "On the Construction of some capacity-approaching coding schemes," Ph.D. dissertation, MIT, MA, 2000.

[70] P. E. O'Neil, "Asymptotics and random Matrices with row-sum and column-sum restrictions," Bull. Amer. Math. Soc., vol. 75, pp.1276-1282, 1969.

[71] I. J. Good and J. F. Crook, “The enumeration of arrays and a generalization related to contingency tables,” Discr. Math., vol. 19, no. 1, pp. 23-45, 1977.

[72] S. Litsyn and V. Shevelev, “On ensemble of low-density parity-check codes: Asymptotic distance distributions,” IEEE Trans. on Inform. Theory, vol. 48, No. 4, pp. 887-908, April, 2002.

[73] S. Litsyn and V. Shevelev, “Distance distributions in ensembles of irregular low-density parity-check codes,” IEEE Trans. on Inform. Theory, vol. 49, No. 12, pp. 3140-3159, Dec. 2003.

[74] T. M. Cover, A. A. El Gamal, “Capacity theorems for the relay channel,” IEEE Trans. On Inform. Theory, vol. 25, No. 5, pp. 572-584, Sept. 1979.

[75] A. Sendonaris, E. Erkip, B. Aazhang, “User cooperation diversity-Part I: System description,” IEEE Trans. on Commun., vol. 51, No. 11, pp. 1927-1938, Nov. 2003.

[76] A. Sendonaris, E. Erkip, B. Aazhang, “User cooperation diversity-Part II: Implementation aspects and performance analysis,” IEEE Trans. on Commun., vol. 51, No. 11, pp. 1939-1948, Nov. 2003.

[77] P. Gupta and P. R. Kumar, “Towards an information theory of large networks: An achievable rate region,” IEEE Trans. on Inform. Theory, vol. 49, No. 8, pp. 1877-1894, Aug. 2003.

[78] A. Host-Madsen and J. Zhang, "Capactity bound and power allocation for wireless relay channels," IEEE Trans. on Inform. Theory, vol. 51, No. 6, pp. pp. 1020-2040, June 2005.

[79] L. L. Xie and P. R. Kumar, "A network information theory for wireless communications: Scaling laws and optimal operation," IEEE Trans. on Inform. Theory, vol. 50, No. 5, pp. 748-767, May 2004.

[80] J. N. Laneman, David N. C. Tse and G. W. Wornell, 'Cooperative diversity in wireless networks: Efficient protocols and outage behaviour," IEEE Trans. on Inform. Theory, vol. 50, No. 12, pp. 3062-3080, Dec. 2004.

[81] G. Kramer, M. Gastpar and P. Gupta, "Cooperative strategies and capacity theorems for relay channel," IEEE Trans. on Inform. Theory, vol. 51, No. 9, pp. 3037-3063, Sept. 2005.

[82] L. L. Xie and P. R. Kumar, "An achievable rate for the multiple-level relay channel, IEEE Trans. on Inform. Theory, vol. 51, No. 4, pp. 1348-1358, April 2005.

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

[84] T. Cover and J. Thomas, Elements of Information Theory, New York: Wiley, 1991.

[85] A. Chakrabarti, A. de Baynast, A. Sabharwal, and B. Anzhang, "Low Density Parity Check Codes for the Rayleigh Channel," IEEE J-SAC, vol. 25, no. 2, pp. 280-291, Feb. 2007.

[86] D. Slepian and J. Wolf, "Noiseless coding of correlative information source," IEEE Trans. Inform. Theory, vol. 10, no. 4, pp. 471-480, July, 1973.

[87] S. M. Ali and Robert J. McEliece, "The generalized distributive law," IEEE Trans. on Inform. Theory, vol. 46, No. 2, pp. 325-343, March, 2000.

[88] A. R. Calderbank, G. David Forney, Jr., and Alexander Vardy, "Minimal tail-biting trellis: The Golay code and more," IEEE Trans. on Inform. Theory, vol. 45, No. 5, pp. 1435-1455, July, 1999.

[89] H. Imai et al., Essentials of Error-Control Coding Techniques, Academic Press, New York, 1990.

[90] G. Ungerboeck, "Channel coding with multilevel/phase signals," IEEE Trans. Inform. Theory, vol. 28, no. 1. pp. 55-67, Jan. 1982.

[91] G. Ungerboeck, J. Hagenauer, and T. Abdel-Nabi, "Coded 8PSK experimental modem for the INTELSAT SCPC system," Proceedings, ICDSC, 7th, pp. 299-304, 1986.

[92] M. Mouly and M. B. Pautet, The GSM System for Mobile Communications, ISBN 2-9507190-0-7, 1993.

[93] TIA/EIA/IS-95 interim standard, mobile station-base station compatibility standard for dual-mode wideband spread spectrum cellular systems, Telecommunications Industry Association, Washington, D.C., July 1993.

[94] C. E. Shannon, "A mathematical theory of communications," Bell Syst. Tech. J., vol. 27, pp. 379-423, July 1948.

[95] F. J. MacWilliams and N. J. A. Sloane, The Theory of Error Correcting Codes, North-Holland, New York, 1988.

[96] Christian B. Schlegel and Lance C. Perez, Trellis and Turbo Coding, IEEE Press, 2004.

[97] A. Roumy, S. Guemghar, G. Caire abd S. Verdu, “Design of Methods for Irregular Repeat-Accumulate Codes,” IEEE Trans. On Inform. Theory, vol. 50, no. 8, pp, 1711-1727, August 2004.

[98] D. Reynolds, Xiaodong Wang, “Turbo Multiuser Detection with Unknown Interferers,” IEEE Trans. on Commun., vol. 50, no. 4, pp. 616-622, April 2002.

[99] R. Koetter, A. C. Singer and M. Tuchler, “Turbo Equalization: An Iterative Equalization and Decoding Technique for Coded Data Transmission,” IEEE Signal Processing Magazine, vol. 21, no.1, pp. 67-80, Jan. 2004.

[100] S. Benedetto, D. Divsalar, G. Montorsi and F. Pollara, “Serial Concatenation of Interleaved Codes: Performance Analysis, Design, and Iterative Decoding,” IEEE Trans. on Inform. Theory, vol. 44, No. 3, pp. 909-926, May 1998.

[101] B. Vucetic and J. Yuan, Turbo Codes: Principles and Applications, Kluwer Academic Publisher, 2000.

[102] M. Reza Soleymani, Yingzi Gao and U. Vilaipornsawai, Turbo Coding for Satellite and Wireless Communications, Kluwer Academic Publishers, 2002.

[103] Ezio Biglieri, Coding for Wireless Channels, Springer, 2005.

[104] Stephen G. Wilson, Digital Modulation and Coding, Prentice Hall, 2996.

[105] Ron M. Roth, Introduction to Coding Theory, Cambridge university press, 2006.

[106] Berlekamp, E. R. Algebraic coding theory, McGraw-Hill, New York, 1968.

APPENDIX 12-A

The ε-Typical Sequence with Log-Likelihood Ratio

Definition 12.5: An independently and identically distributed (i.i.d) sequence $\mathbf{x}$=(x_1, x_2, …, x_N) is ε-typical with respect to $\lambda(x) = \log[p(x)/q(x)]$ if the following inequality holds for any ε>0 for sufficiently large N.

$$\left| \frac{1}{N}\log\left[\frac{p(\mathbf{x})}{q(\mathbf{x})}\right] - D(p(x)\,\|\,q(x)) \right| < \varepsilon \qquad \text{(A-1)}$$

where the random variable x represents an arbitrary element in $\mathbf{x}$, $p(x)$ and $q(x)$ are two probability distributions over an common alphabet Ω, and $D(p(x)\,\|\,q(x))$ is the relative entropy defined as

$$D(p(x)\,\|\,q(x)) = E_p\left[\log\frac{p(x)}{q(x)}\right] = \int_{-\infty}^{+\infty} p(x)\log\left[\frac{p(x)}{q(x)}\right]dx \qquad \text{(A-2)}$$

Remark: (A-1) can be further expressed as

$$\left| \frac{1}{N}\sum_{k=1}^{N}\log\left[\frac{p(x_i)}{q(x_i)}\right] - D(p(x)\,\|\,q(x)) \right| < \varepsilon \qquad \text{(A-3)}$$

Definition 12.6: The ε-typical region $T_\varepsilon^{(N)}(\mathbf{x})$ is defined as

$$T_\varepsilon^{(N)}(\mathbf{x}) = \{\, \mathbf{x} \mid N[D(p(x)\,|\,q(x)) - \varepsilon] \le \log[p(\mathbf{x})/q(\mathbf{x})] \le N[D(p(x)\,|\,q(x)) + \varepsilon] \,\} \qquad \text{(A-4)}$$

Theorem 12.2 (Fundamental theorem of large-deviation theory): Given two probability distributions over an common alphabet Ω, for any ε >0, the probability that an i.i.d sequence $\mathbf{x}$ distributed by $q(\mathbf{x})$ of length N is ε-typical for $p(x)\,|\,q(x)$ is bounded by

$$(1-\delta(N))\exp[-N(D(p(x)\,\|\,q(x))+\varepsilon)] \le \Pr\{\,\mathbf{x}\,|\,\mathbf{x}\ \ is\ \ \varepsilon-typical\ for\ p(x)\,|\,q(x)\,\}$$

$$\le \exp[-N(D(p(x)\,\|\,q(x))-\varepsilon)] \qquad \text{(A-5)}$$

where $\delta(N)\to 0$ as $N\to 0$.

The proof can be referred to [84] using Chernoff bound. The following results are the direct extensions from the above definitions and theorem, which are crucial for proving the theorem in Appendix 10-B.

Consider m i.i.d sequences where each $\mathbf{z}_i$=($z_{i,1}$, $z_{i,2}$, …, $z_{i,N}$) is a N-sequence for i=1, 2, …, m. A length-mN sequence $\mathbf{s}$ is formed as follows:

$$\mathbf{s}=(\mathbf{z}_1,\ \mathbf{z}_2,\ \ldots,\ \mathbf{z}_m) \qquad \text{(A-6)}$$

According to above definitions and theorem, the set $T_\varepsilon^{(N)}(\mathbf{s})$ of ε-typical sequences is given as

$$T_\varepsilon^{(N)}(\mathbf{s})=\left\{\mathbf{s}:\left|\frac{1}{N}\log\frac{1}{\Pr(\mathbf{s})}-H(S)\right|<\varepsilon\right\} \qquad \text{(A-7)}$$

where

$$\Pr(\mathbf{s})=\Pr(\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_m)=\prod_{n=1}^{N}\Pr(z_{1,n},z_{2,n},\ldots,z_{m,n}) \qquad \text{(A-8)}$$

and S is an arbitrary vector ($z_{1,n}$, $z_{2,n}$, …, $z_{m,n}$) for any $n\in\{1, 2, \ldots, N\}$.

(1) Let $\mathbf{s}$ be generated according to (A-8), then by law of large number and Chernoff bound

$$\Pr((\mathbf{z}_1,\mathbf{z}_2,\ldots,\mathbf{z}_m)\in T_\varepsilon^{(N)}(\mathbf{s}))\ge 1-\delta \qquad \text{(A-9)}$$

for any ε>0.

(2) Let $\mathbf{s}$ be generated according to

$$\Pr(\mathbf{s})=\prod_{n=1}^{N} p(z_{1,n} \mid z_{2,n}, z_{3,n}, ..., z_{m-1,n})\, p(z_{m,n} \mid z_{2,n}, z_{3,n}, ..., z_{m-1,n})$$

$$\times\, p(z_{2,n}, z_{3,n}, ..., z_{m-1,n}) \qquad \text{(A-10)}$$

Then based on Theorem 12.2, we have

$$\Pr((\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_m) \in T_{\varepsilon}^{(N)}(\mathbf{s})) < 2^{-N(I(Z_1; Z_m|Z_2, Z_3, ..., Z_m)-6\delta)} \qquad \text{(A-11)}$$

where Z_1, Z_2, …., Z_m are the corresponding component variables in S =($z_{1,n}$, $z_{2,n}$, …, $z_{m,n}$) for any $n \in \{1, 2, …, N\}$.

APPENDIX 12-B

The Proof of the Theorem 12.1 for Multiple-Level Relay Channel

Theorem 12.1: For the discrete memoryless multiple-level relay channel described in this chapter, the achievable rate R is bounded by

$$R< \max_{p(x_0,x_1,...,x_{M-1})} \min_{1<k<M} I(X_0,X_1,....,X_{k-1};Y_k \mid X_k,X_{k+1},....,X_{M-1}) \tag{B-1}$$

where $I(X_0,X_1,....,X_{k-1};Y_k \mid X_k,X_{k+1},....,X_{M-1})$ is the conditional mutual information between (X_0, X_1, …, X_{k-1}) and Y_k given (X_k, X_{k+1}, …, X_{M-1}).

Proof: In order to make the proof easier to follow and give a better understanding, we only present the one-relay case (M=2), which consists of all the essential ideas for the straightforward extension to the general case (M>2). In this proof we will use the so-called "ε-typical N-sequences", for more details the reader can refer to [84, Sections 8.6 and 14.2].

1. One-relay channel (M=2)

The block diagram of an one-relay channel is shown in Fig.12.4. We consider a binary block coding argument, where B blocks are considered for transmission from the source node and each of blocks contains N slots. A sequence of $B-1$ indices, $w(b) \in \{1, 2, ..., 2^{NR}\}$, b=1, 2,…, $B-1$, will be sent over in NB transmission slots. Note that as $B \to +\infty$, the rate $NR(B-1)/NB$ is arbitrarily close to R for any length N.

(1) Generate at random 2^{NR} independent and identically distributed (i.i.d) *N*-sequences Ω_1^N, each drawn according to

$$p(\mathbf{x}_1)=\prod_{n=1}^{N} p(x_{1,n}) \tag{B-2}$$

where the block codeword $\mathbf{x}_1=[x_{1,1}, x_{1,2}, ..., x_{1,N-1}, x_{1,N}]$. Note that $\mathbf{x}_1$ is a function of w_1 as $\mathbf{x}_1(w_1)$, $w_1 \in \{1, 2, ..., 2^{NR}\}$, to represent the selection of codeword from the codebook (codewords set) for the output of node 1.

(2) For each $\mathbf{x}_1(w_1)$, generate 2^{NR} conditionally independent *N*-sequences $\mathbf{x}_0(w_0|w_1)$, $w_0 \in \{1, 2, ..., 2^{NR}\}$, drawn independently according to

$$p(\mathbf{x}_0 \mid \mathbf{x}_1(w_1))=\prod_{n=1}^{N} p(x_{0,n} \mid x_{1,n}(w_1)) \tag{B-3}$$

This defines the *joint* codebook for node 0, 1 as

$$C_0 := \{\mathbf{x}_0(w_0|w_1),\ \mathbf{x}_1(w_1)\} \tag{B-4}$$

for each $\mathbf{x}_1(w_1)$.

Remark: The reason for this kind of *backward* codebook generation is based on the fundamental principle for relay channel that upstream nodes (with smaller index) know what the downstream nodes are going to transmit, and thus can adjust their own transmission accordingly.

Repeating the above process 1 and 2 independently once more, we can generate another random codebook C_1 similar to C_0 in (A-4). In each block *b*=1, 2, …., *B*, the codebook $C_{(b \bmod 2)}$ is used. Hence, in any two consecutive blocks, codewords from different blocks are independent and used alternatively. This property will be used in the analysis of the probability of block error.

Encoding: At the beginning of each block $b \in \{1, 2, \ldots., B\}$, node 1 has an estimate $\hat{w}_1(b-1)$ of $w(b-1)$, which is sent from the source (node 0), and sends the following N-sequence from the codebook $C_{(b \bmod 2)}$ in the block, which is denoted as

$$\vec{X}_1(b) := \mathbf{x}_1(\hat{w}_1(b-1)) \tag{B-5}$$

Also in the same block, node 0 sends the following N-sequence from the same block codebook $C_{(b \bmod 2)}$

$$\vec{X}_0(b) := \mathbf{x}_0(\hat{w}_0(b) \mid \hat{w}_1(b-1)) \tag{B-6}$$

Note that the estimate $\hat{w}_0(b) = w_0(b)$ for every $b \in \{1, 2, \ldots., B\}$ due to node 0 as the source. For the purpose of synchronization of all the nodes at the initial time, we set $\hat{w}_i(b_1) = w(b_1) = 1$ for every $b_1 \leq 0$, $i \in \{0, 1, 2\}$. The encoding process is depicted in Table 12.1.

Thus, every node $k \in \{1, 2\}$ receives an N-sequence as

$$\vec{Y}_k(b) := \vec{Y}_k(\vec{X}_0(b), \vec{X}_1(b)) \tag{B-7}$$

with probability

$$p(\vec{Y}_k(b) \mid \vec{X}_0(b), \vec{X}_1(b)) = \prod_{n=1}^{N} p(\vec{Y}_{k,n}(b) \mid \vec{X}_{0,n}(b), \vec{X}_{1,n}(b)) \tag{B-8}$$

where $\vec{Y}_{k,n}(b)$ is the nth element of the vector $\vec{Y}_k(b)$, and similar definitions hold for $\vec{X}_{i,n}(b)$ (i=0, 1).

Decoding: At the end of each block $b \in \{1, 2, \ldots, B\}$, decodings at node 1 and node 2 happen simultaneously, but independentl2y. The decoding criterion is

A: Node 1 declares that $\hat{w}_1(b) = w$ if w is the *unique* value in $\{1, 2, \ldots, 2^{NR}\}$ such that in the block b,

$$\{\mathbf{x}_0(w \mid \hat{w}_1(b-1)),\ \mathbf{x}_1(\hat{w}_1(b-1)),\ \vec{Y}_1(b)\} \in A_\varepsilon^{(N)}(X_0, X_1, Y_1) \tag{B-9}$$

where $A_\varepsilon^{(N)}(X_0, X_1, Y_1)$ denotes the set of ε-typical N-sequences [82][84] (X_0, X_1, Y_1). Otherwise, if no unique w as above exists, an error block is thus declared with $\hat{w}_1(b)$=0.

B: Node 2 declares that $\hat{w}_2(b-1) = w$ if w is the *unique* value in $\{1, 2, \ldots, 2^{NR}\}$ such that in *both* the blocks b and b–1

$$\{\mathbf{x}_1(w),\ \vec{Y}_2(b)\} \in A_\varepsilon^{(N)}(X_1, Y_2) \tag{B-10a}$$

and

$$\{\mathbf{x}_0(w \mid \hat{w}_2(b-2)),\ \mathbf{x}_1(\hat{w}_2(b-2)),\ \vec{Y}_2(b-1)\} \in A_\varepsilon^{(N)}(X_0, X_1, Y_2) \tag{B-10b}$$

Otherwise, if no unique w as above exists, an error block is declared with $\hat{w}_2(b-1)$=0.

Table 12.2: The encoding process for one-relay channel.

Block 1	Block 2	………	Block B
$\mathbf{x}_0(w(1) \mid 1)$	$\mathbf{x}_0(w(2) \mid w_1(1))$	………	$\mathbf{x}_0(w(B) \mid w_1(B-1))$
$\mathbf{x}_1(\hat{w}_1(0) = 1)$	$\mathbf{x}_1(\hat{w}_1(1))$	………	$\mathbf{x}_1(\hat{w}_1(B-1))$

Analysis of Probability of Block of Error: Denote the event that no decoding error is made in the first b blocks by

$$A_c(b) := \{ \hat{w}_k(b_1 - k + 1) = w(b_1 - k + 1), \text{ for all } b_1 \in \{1, \ldots, b\} \text{ and } k \in \{1, 2\} \} \quad \text{(B-11)}$$

and let its probability be

$$P_c(b) := \Pr(A_c(b)) \quad \text{(B-12)}$$

with $P_c(0) := 1$.

Then the probability that some decoding error is occurred at some nodes $k \in \{1, 2\}$ in some block $b \in \{1, 2, \ldots, B\}$ is

$$\begin{aligned} P_e &:= \Pr(\hat{w}_k(b - k + 1) \neq w(b - k + 1), \text{ for some } k \in \{1, 2\}, b \in \{1, 2, \ldots, B\}) \\ &= \sum_{b=1}^{B} \Pr(\hat{w}_k(b - k + 1)) \neq w(b - k + 1) \text{ for some } k \in \{1, 2\} \mid A_c(b-1)) \, p_c(b-1) \\ &\leq \sum_{b=1}^{B} \sum_{k=1}^{2} \Pr(\hat{w}_k(b - k + 1) \neq w(b - k + 1) \mid A_c(b-1)) \, p_c(b-1) \\ &= \sum_{b=1}^{B} \sum_{k=1}^{2} P_{e,k}(b) P_c(b-1) \end{aligned} \quad \text{(B-13)}$$

Where the last equation is due to union bound and $P_{e,k}(b)$ is defined as

$$P_{e,k}(b) = \Pr(\overline{w}_k(b - k + 1) \neq w(b - k + 1) \mid A_c(b-1)) \quad \text{(B-14)}$$

Thus, $P_{e,k}(b)$ is the probability that a decoding error happens at node k in block b, conditioned on the event without decoding error in the previous $b-1$ blocks.

Next, we now calculate $P_{e,k}(b)$, $k \in \{1, 2\}$. Since $A_c(b-1)$ is presumed to hold for every $k \in \{1, 2\}$ we have

$$\overline{w}_k(b_1-k+1)=w(b_1-k+1) \quad \text{for } 1\le b_1\le b-1 \tag{B-15}$$

Therefore, the decoding rule (B-9) with respect to node 1 is

$$\{\,\mathbf{x}_0(w\,|\,w(b-1)),\ \mathbf{x}_1(w(b-1)),\ \vec{Y}_1(\,\mathbf{x}_0(w(b)\,|\,w(b-1)),\mathbf{x}_1(w(b-1))\,)\}$$

$$\in T_\varepsilon^{(N)}(X_0, X_1, Y_1) \tag{B-16}$$

The decoding rule for node 2 is

$$\{\,\mathbf{x}_1(w),\ \vec{Y}_2(\,\mathbf{x}_0(w(b)\,|\,w(b-1)),\mathbf{x}_1(w(b-1))\,)\}\in T_\varepsilon^{(N)}(X_1, Y_2) \tag{B-17a}$$

for the block b with respect to (B-10a), and

$$\{\,\mathbf{x}_0(w\,|\,w(b-2)),\ \mathbf{x}_1(w(b-2)),\ \vec{Y}_2(\,\mathbf{x}_0(w(b-1)\,|\,w(b-2)),\mathbf{x}_1(w(b-2))\,)\}$$

$$\in T_\varepsilon^{(N)}(X_0, X_1, Y_2) \tag{B-17b}$$

for the block b–1 with respect to (B-10b). In order to simplify the expressions of analysis in the sequel, we let

$$W_1(b):=\{\,w\in\{1,\,2,\,\ldots,\,2^{NR}\}:\ w \text{ satisfies (B-16)}\} \tag{B-18a}$$

$$W_{2,0}(b):=\{\,w\in\{1,\,2,\,\ldots,\,2^{NR}\}:\ w \text{ satisfies (B-17a)}\} \tag{B-18b}$$

$$W_{2,1}(b):=\{\,w\in\{1,\,2,\ldots,\,2^{NR}\}:\ w \text{ satisfies (B-17b)}\} \tag{B-18c}$$

$$W_2(b):=W_{2,0}(b)\cap W_{2,1}(b) \tag{B-18d}$$

Then, $P_{e,k}(b)$ is the probability that $w(b-k+1)\notin W_k(b)$, or some $w^{'}\in W_k(b)$ but $w^{'}\neq w(b-w+1)$, conditioned on the event that no decoding error was made in the previous b–1 blocks. Thus, we get

$$P_{e,k}(b) = \Pr(\, w(b-k+1) \notin W_k(b), \textit{ or } \text{some } w' \in W_k(b) \text{ but } w' \neq w(b-k+1) \,|\, A_c(b-1)\,)$$

$$\leq \Pr(\, w(b-k+1) \notin W_k(b) \,|\, A_c(b-1)\,)$$

$$+\Pr(\text{some } w' \in W_k(b) \text{ but } w' \neq w(b-k+1) \,|\, A_c(b-1)\,)$$

$$= \frac{\Pr(w(b-k+1) \notin W_k(b))}{P_c(b-1)} + \frac{\Pr(\textit{some } \; w' \in W_k(b) \;\; \textit{but} \;\; w' \neq w(b-k+1))}{P_c(b-1)} \qquad \text{(B-19)}$$

Therefore, by (B-13)

$$P_e \leq \sum_{b=1}^{B} \sum_{k=1}^{2} [\, \Pr(\, w(b-k+1) \notin W_k(b)\,) + \Pr(\text{some } w' \in W_k(b) \text{ but } w' \neq w(b-k+1)\,)] \qquad \text{(B-20)}$$

For node 1 with (B-16) and sufficiently large *N*, we have following inequalities by (A-9) (A-11) for the *ε*-typical *N*-sequences under the condition $(Z_1, Z_2, Z_3)=(X_0,\ X_1,\ Y_1)$.

$$\Pr(w(b) \notin W_1(b)) < \delta \qquad \text{(B-21a)}$$

and for any $w' \neq w(b)$, we have

$$\Pr(w' \in W_1(b)) < 2^{-N(I(X_0;Y_1|X_1)-6\delta)} \qquad \text{(B-21b)}$$

Similarly, for node 2 with sufficiently large *N*, we also have

$$\Pr(w(b-1) \notin W_{2,0}(b)) < \delta \qquad \text{(B-22a)}$$

$$\Pr(w(b-1) \notin W_{2,1}(b)) < \delta \qquad \text{(A-22b)}$$

For any $w' \neq w(b-1)$ under the condition $(Z_1, Z_2, Z_3)=(X_0,\ X_1,\ Y_2)$, it yields

$$\Pr(w' \in W_{2,0}(b)) < 2^{-N(I(X_1;Y_2)-6\delta)} \qquad \text{(B-23a)}$$

$$\Pr(w' \in W_{2,1}(b)) < 2^{-N(I(X_0;Y_2|X_1)-6\delta)} \qquad \text{(B-23b)}$$

Hence, by (B-22a) and (B-22b)

$$\Pr(w(b-1) \notin W_2(b)) \le \sum_{j=0}^{1} \Pr(w(b-1)) \in W_{2,j}(b)) = 2\delta \tag{B-24}$$

and by (B-23a) and (B-23b)

$$\Pr(\text{some } w' \in W_2(b) \text{ but } w' \ne w(b-1))$$

$$\le \sum_{\substack{w' \in \{1,\ldots,2^{NR}\} \\ w' \ne w(b-1)}} \Pr(w' \in W_2(b)) = \sum_{\substack{w' \in \{1,\ldots,2^{NR}\} \\ w' \ne w(b-1)}} \prod_{j=0}^{1} \Pr(w' \in W_{2,j}(b))$$

$$\le (2^{NR}-1)\, 2^{-N(I(X_0,X_1;Y_2)-12\delta)}$$

$$= 2^{N(R-I(X_0,X_1;Y_2)+12\delta)} - 2^{-N(I(X_0,X_1;Y_2)-12\delta)} \tag{B-25}$$

where the first equality in (B-25) follows from the independence between the codebooks of any two consecutive blocks, and the second inequality is due to the identity

$$I(X_0, X_1; Y_2) = I(X_1; Y_2) + I(X_0; Y_2 | X_1) \tag{B-26}$$

Similarly, regarding to (B-21b) we have

$$\Pr(\text{some } w' \in W_1(b) \text{ but } w' \ne w(b))$$

$$\le (2^{NR}-1)\, 2^{-N(I(X_0;Y_1|X_1)-6\delta)}$$

$$= 2^{N(R-I(X_0;Y_1|X_1)+6\delta)} - 2^{-N(I(X_0;Y_1|X_1)-6\delta)} \tag{B-27}$$

For any R satisfying (B-1) for one-relay channel and also choosing ε small enough, we make N large enough such that

$$\Pr(\text{some } w' \in W_1(b) \text{ but } w' \ne w(b)) < \delta_1 \tag{B-28a}$$

$$\Pr(\text{some } w' \in W_2(b) \text{ but } w' \ne w(b-1)) < \delta_1 \tag{B-28b}$$

where ε_1 is an arbitrarily small positive number. Therefore, by (B-20), (B-21a), (B-24), (B-28a) and (B-28b),

$$P_e \le \sum_{b=1}^{B} (3\varepsilon + 2\varepsilon_1) = 3B\delta + 2B\delta_1 \tag{B-29}$$

which can be made arbitrarily small by letting $N \to +\infty$.

2. Multiple-relay channel (*M*>2)

Following the similar steps as above for one-relay channel, we can prove the scenario for multiple-relay channel (M>2). For more details, the interested reader can refer to bibliography [82]. □