Optimal Transmission Strategy and Explicit Capacity Region for Broadcast Z Channels

Bike Xie, *Student Member, IEEE*, Miguel Griot,*Student Member, IEEE*, Andres I. Vila Casado, *Student Member, IEEE* and Richard D. Wesel, *Senior Member, IEEE*

Abstract

This paper provides an explicit expression for the capacity region of the two-user broadcast Z channel and proves that the optimal boundary can be achieved by independent encoding of each user. Specifically, the information messages corresponding to each user are encoded independently and the OR of these two encoded streams is transmitted. Nonlinear turbo codes that provide a controlled distribution of ones and zeros are used to demonstrate a low-complexity scheme that operates close to the optimal boundary.

Keywords

broadcast channel, broadcast Z channel, capacity region, turbo codes, nonlinear turbo codes.

I. Introduction

Degraded broadcast channels were first studied by Cover in [1] and a formulation of the capacity region was established in [2], [3] and [4]. Superposition encoding is the key idea to achieve the optimal boundary of the capacity region for degraded broadcast channels [5]. With superposition encoding for degraded broadcast channels, the data sent to the user with the most degraded channel is encoded first. Given the encoded bits for that user, an appropriate codebook for the second most degraded channel user is selected, and so forth. Hence superposition encoding is, in general, a joint encoding scheme. However, combining independently encoded streams, one for each user, is an optimal scheme for some broadcast channels including broadcast Gaussian channels [1] and broadcast binary-symmetric channels [1] [2].

Successive decoding is a natural decoding scheme for superposition encoding [1] [2] [5]. With successive decoding for degraded broadcast channels, each receiver first decodes the

This work was supported by the Defence Advanced Research Project Agency SPAWAR Systems Center, San Diego, California under Grant N66001-02-1-8938. This paper was presented in part at the Information Theory Workshop 2007.
The authors are with the Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA (e-mail:xbk@ee.ucla.edu; mgriot@ee.ucla.edu; avila@ee.ucla.edu; wesel@ee.ucla.edu).

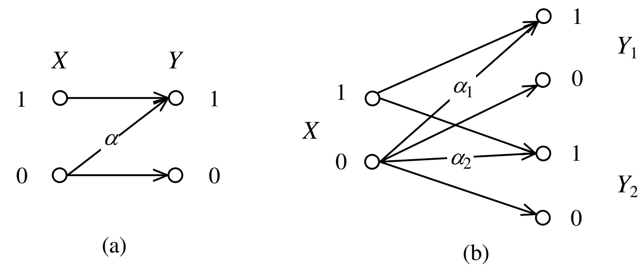

Fig. 1. (a) Z channel. (b) Broadcast Z channel.

data sent to the user with the most degraded channel. Conditioning on the decoded data for that user, each receiver determines the codebook for the user with the second most degraded channel and decodes that data, and so forth until the desired user's data is decoded. The performance of successive decoding for degraded broadcast channels is very close to optimal decoding under normal operating conditions.

Turbo codes [6] and Low-Density Parity-Check (LDPC) codes [7] perform close to the Shannon limit. Coding approaches for broadcast channels were studied in [8] and [9]. In [8], LDPC codes provided reliable transmission over two-user broadcast channels with additive white Gaussian noise (AWGN) and fading known at the receiver only. In [9], a superposition turbo coding scheme performs within 1dB of the capacity region boundary for broadcast Gaussian channels. Both of these approaches are designed specifically for broadcast Gaussian channels and used linear codes. For multi-user binary adder channels, nonlinear trellis codes were studied and designed in [10].

The Z channel is the binary-asymmetric channel shown in Fig. 1(a). The capacity of the Z channel was studied in [11]. Nonlinear trellis codes were designed to maintain a low ones density for the Z channel in [12] and parallel concatenated nonlinear turbo codes were designed for the Z channel in [13]. This paper focuses on the study of the two-user broadcast Z channel shown in Fig. 1(b). This paper provides an explicit expression of the capacity region for the two-user broadcast Z channel and shows that independent encoding with successive decoding can achieve the boundary of this capacity region.

This paper is organized as follows. Section II introduces definitions and notation for broadcast channels. Section III provides the explicit expression of the capacity region for the two-user broadcast Z channel and proves that independent encoding can achieve the optimal boundary of the capacity region. Section IV presents nonlinear-turbo codes designed to achieve the optimal boundary, and Section V provides the simulation results. Section VI delivers the conclusions.

II. Definitions and Preliminaries

A. Degraded broadcast channels

The general representation of a discrete memoryless broadcast channel is given in Fig. 2. A single signal X is broadcast to M users through M different channels $A_1, \cdots, A_M$. Channel A_{i+1} is a physically degraded version of channel A_i (and thus the broadcast channel $X \rightarrow Y_i, Y_{i+1}$ is physically degraded) if $p(y_i, y_{i+1}|x) = p(y_i|x)p(y_{i+1}|y_i)$ [5]. A physically degraded broadcast channel with M users is shown in Fig. 3. Since each user decodes its received signal without collaboration, only the marginal transition probabilities $p(y_1|x), p(y_2|x), \cdots, p(y_M|x)$ of the component channels $A_1, A_2, \cdots, A_M$ affect receiver performance. Hence, the *stochastically* degraded broadcast channel is defined in [2] and [5].

Let A_i be a channel with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}_\mathrm{i}$, and transition probability $p_i(y_i|x)$. Let A_{i+1} be another channel with the same input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}_{\mathrm{i}+1}$, and transition probability $p_{i+1}(y_{i+1}|x)$. A_{i+1} is a stochastically degraded version of A_i if there exists a transition probability $q(y_{i+1}|y_i)$ such that

$$p_{i+1}(y_{i+1}|x) = \sum_{y_i \in \mathcal{Y}_\mathrm{i}} q(y_{i+1}|y_i) p_i(y_i|x). \tag{1}$$

A broadcast channel with receivers $Y_1, Y_2 \cdots, Y_M$ is a stochastically degraded broadcast channel if every component channel A_i is a stochastically degraded version of A_{i-1} for all $i = 2, \cdots, M$ [2]. Since the marginal transition probabilities $p(y_1|x), p(y_2|x), \cdots, p(y_M|x)$ completely determine a stochastically degraded broadcast channel, we can model any stochas-

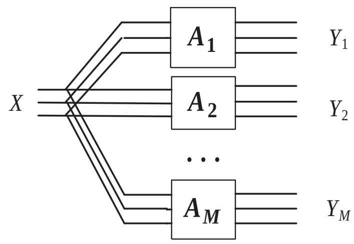

Fig. 2. Broadcast channel.

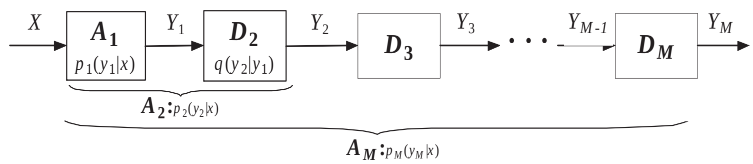

Fig. 3. Physically degraded broadcast channel.

tically degraded broadcast channel as a physically degraded broadcast channel with the same marginal transition probabilities.

Theorem 1 ([2] [4]) The capacity region for the two-user stochastically degraded broadcast channel $X \to Y_1 \to Y_2$ is the convex hull of the closure of all (R_1, R_2) satisfying

$$R_2 \leq I(X_2; Y_2) \qquad R_1 \leq I(X; Y_1|X_2), \tag{2}$$

for some joint distribution $p(x_2)p(x|x_2)p(y_1, y_2|x)$, where the auxiliary random variable X_2 has cardinality bounded by $|\mathcal{X}_2| \leq \min\{|\mathcal{X}|, |\mathcal{Y}_1|, |\mathcal{Y}_2|\}$. ($U$ was used as the auxiliary random variable in [2] [4]. In this paper, we use X_2 instead of U because the auxiliary random variable X_2 corresponds to the second user's encoded stream in the independent encoding scheme to be proposed.)

B. The broadcast Z channel

The Z channel, shown in Fig. 1(a), is a binary-asymmetric channel with the transition probability matrix

$$T = \begin{bmatrix} 1 & \alpha \\ 0 & 1-\alpha \end{bmatrix},$$

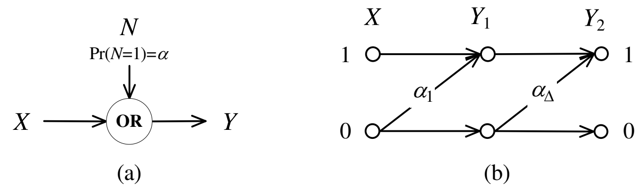

Fig. 4. (a) OR operation view of Z channel. (b) Physically degraded broadcast Z channel.

where $0 \leq \alpha \leq 1$. If symbol 1 is transmitted, symbol 1 is received with probability 1. If symbol 0 is transmitted, symbol 1 is received with probability α and symbol 0 is received with probability $1-\alpha$. We can model the Z channel as the OR operation of the channel input X and Bernoulli noise N with parameter α as shown in Fig. 4(a)). In an OR Multiple Access Channel, each user appears to transmit over a Z channel when the other users are treated as noise [13]. Thus, in an OR network with multiple transmitters and multiple receivers, each transmitter transmitting to more than one receiver sees a broadcast Z channel if other transmitters transmitting to those receivers are treated as noise. The two-user broadcast Z channel with the marginal transition probability matrices

$$T_1 = \begin{bmatrix} 1 & \alpha_1 \\ 0 & 1-\alpha_1 \end{bmatrix} \qquad T_2 = \begin{bmatrix} 1 & \alpha_2 \\ 0 & 1-\alpha_2 \end{bmatrix}$$

is shown in Fig. 1(b), where $0 \leq \alpha_1 \leq \alpha_2 \leq 1$. Because broadcast Z channels are stochastically degraded, we can model any broadcast Z channel as a physically degraded broadcast Z channel as shown in Fig. 4(b), where

$$\alpha_\Delta = \frac{\alpha_2 - \alpha_1}{1 - \alpha_1}. \tag{3}$$

III. Optimal Transmission Strategy for the Two-User Broadcast Z Channel

Since the broadcast Z channel is stochastically degraded, its capacity region can be obtained directly from Theorem 1. The capacity region for the broadcast Z channel $X \to Y_1 \to Y_2$ as shown in Fig. 5 is the convex hull of the closure of all (R_1, R_2) satisfying

$$\begin{aligned} R_2 &\le I_2 = I(X_2; Y_2) \\ &= H\big((\bar{\mu}_2\gamma + \mu_2\mu_1)(1-\alpha_2)\big) - \bar{\mu}_2 H\big(\gamma(1-\alpha_2)\big) - \mu_2 H\big(\mu_1(1-\alpha_2)\big), \qquad (4) \\ R_1 &\le I_1 = I(X; Y_1 | X_2) \\ &= \bar{\mu}_2\big(H(\gamma(1-\alpha_1)) - \gamma H(1-\alpha_1)\big) + \mu_2\big(H(\mu_1(1-\alpha_1)) - \mu_1 H(1-\alpha_1)\big), \qquad (5) \end{aligned}$$

for some probabilities μ_1, μ_2, γ, where $\mu_1 = \Pr(x = 0 | x_2 = 0)$, $\mu_2 = \Pr(x_2 = 0)$, $\gamma = \Pr(x = 0 | x_2 = 1)$, $H(\cdot)$ is the binary entropy function, $\bar{\mu}_1 = 1 - \mu_1$, $\bar{\mu}_2 = 1 - \mu_2$ and

$$\alpha_2 = \Pr\{y_2 = 1 | x = 0\} = 1 - (1-\alpha_1)(1-\alpha_\Delta). \qquad (6)$$

Each particular choice of (μ_1, μ_2, γ) in Fig. 5 specifies a particular transmission strategy and a rate pair (I_1, I_2). The optimal boundary of a capacity region is the set of all Pareto optimal points (I_1, I_2), for which it is impossible to increase rate I_1 without decreasing rate I_2 or vice versa. A transmission strategy is optimal if and only if it achieves a rate pair point on the optimal boundary. We call a set of transmission strategies *sufficient* if all rate pairs on the optimal boundary can be achieved by using these strategies and time sharing. Furthermore, a set of transmission strategies is *strongly sufficient* if these strategies can achieve all rate pairs on the optimal boundary without using time sharing. Equations (4) and (5) give a set of pentagons that yield the capacity region through their convex hull, but do *not* explicitly show the optimal transmission strategies or derive the boundary of the

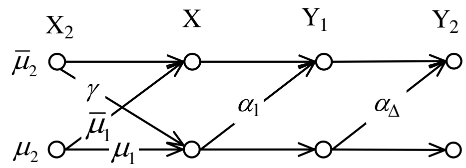

Fig. 5. Information theoretic diagram of the system.

capacity region.

A. *Optimal transmission strategies*

The following theorem identifies a set of optimal transmission strategies and provides an explicit expression of the boundary of the capacity region.

Theorem 2: For a broadcast Z channel with $0 < \alpha_1 < \alpha_2 < 1$, the set of the optimal transmission strategies (μ_1, μ_2, γ), which satisfy

$$\gamma = 0, \tag{7}$$

$$\frac{1}{(1-\alpha_1)(e^{H(1-\alpha_1)/(1-\alpha_1)}+1)} \le \mu_1 \le 1, \tag{8}$$

and

$$\begin{aligned} &\ln(1-\mu_1(1-\alpha_1))\big(H(\mu_1(1-\alpha_2)) - \mu_1(1-\alpha_2)\ln\frac{1-\mu_2\mu_1(1-\alpha_2)}{\mu_2\mu_1(1-\alpha_2)}\big) \\ &= \ln(1-\mu_1(1-\alpha_2))\big(H(\mu_1(1-\alpha_1)) - \mu_1 H(1-\alpha_1)\big), \end{aligned} \tag{9}$$

are strongly sufficient. In other words, all rate pairs on the optimal boundary of the capacity region can be achieved by using exactly the transmission strategies described in (7-9) without the need of time sharing. Furthermore, applying (7-9) to (4) and (5) yields an explicit expression of the optimal boundary of the capacity region.

Before proving Theorem 2, we present and prove some preliminary results. From (4) and (5), we can see that the transmission strategies (μ_1, μ_2, γ) and $(\gamma, 1-\mu_2, \mu_1)$ have the same transmission rate pairs. Therefore, we assume $\gamma \leq \mu_1$ in the rest of the section without loss of generality.

Theorem 3: For a broadcast Z channel with $0 < \alpha_1 < \alpha_2 < 1$, any transmission strategy (μ_1, μ_2, γ) with $0 < \mu_2 < 1, 0 < \gamma < \mu_1$ is not optimal.

The proof is given in Appendix A.

Corollary 1: The set of all the transmission strategies with $\gamma = 0$ is sufficient for any broadcast Z channel with $0 < \alpha_1 < \alpha_2 < 1$.

Proof: From Theorem 3, we know that the transmission strategy (μ_1, μ_2, γ) is optimal only if at least one of these four equations $\mu_2 = 0$, $\mu_2 = 1$, $\gamma = \mu_1$, $\gamma = 0$ is true. Hence the set of all the transmission strategies with $\mu_2 = 0$, $\mu_2 = 1$, $\gamma = \mu_1$ or $\gamma = 0$ is sufficient. When $\mu_2 = 0$, $\mu_2 = 1$ or $\gamma = \mu_1$, the transmission rate for the second user, I_2 in equation (4), is zero. (This optimal rate pair is the point B in Fig. 6(a), which will see later.) Since this point can also be achieved by the transmission strategy with $\gamma = 0$, $\mu_2 = 1$ and $\mu_1 = \arg\max(H(x(1-\alpha_1)) - xH(1-\alpha_1))$, all the optimal rate pairs on the optimal boundary of the capacity region can be achieved by using the transmission strategies with $\gamma = 0$ and time sharing. Thus, the set of all the transmission strategies with $\gamma = 0$ is sufficient. □

From Corollary 1, we can set $\gamma = 0$ in Fig. 5 without losing any part of the capacity region and so the designed virtual channel $X_2 \to X$ is a Z channel. Since we can consider the output of a Z channel as the OR operation of two Bernoulli random variables, an independent encoding scheme that works well for the broadcast Z channel will be introduced later in this paper.

Applying $\gamma = 0$ to (4) and (5) yields

$$R_2 \leq I_2 = H(\mu_2\mu_1(1-\alpha_2)) - \mu_2 H(\mu_1(1-\alpha_2)), \tag{10}$$

$$R_1 \leq I_1 = \mu_2 H(\mu_1(1-\alpha_1)) - \mu_2\mu_1 H(1-\alpha_1). \tag{11}$$

By Corollary 1, the capacity region is the convex hull of the closure of all rate pairs (R_1, R_2) satisfying (10) and (11) for some probability μ_1, μ_2. However, *not* all transmission strategies of $(\mu_1, \mu_2, \gamma = 0)$ achieve the optimal boundary of the capacity region. Since any optimal transmission strategy maximizes $I_1 + \lambda I_2$ for some nonnegative λ, we solve the optimization problem of maximizing $I_1 + \lambda I_2$ for any fixed $\lambda \geq 0$ in order to find the constraints on μ_1 and μ_2 for optimal transmission strategies. Theorem 4 provides the solution to this maximization problem.

Theorem 4: The optimal solution to the maximization problem

$$\begin{aligned} \text{maximize} \quad & I_1 + \lambda I_2 \\ \text{subject to} \quad & I_2 = H(\mu_2\mu_1(1-\alpha_2)) - \mu_2 H(\mu_1(1-\alpha_2)) \\ & I_1 = \mu_2 H(\mu_1(1-\alpha_1)) - \mu_2\mu_1 H(1-\alpha_1) \\ & 0 \leq \mu_2 \leq 1, 0 \leq \mu_1 \leq 1, \end{aligned} \tag{12}$$

is unique and it is given below for any fixed $\lambda \geq 0$.

Define

$$\varphi(x) = \frac{\ln(1-(1-\alpha_1)x)}{\ln(1-(1-\alpha_2)x)} \tag{13}$$

and

$$\psi(x) = \frac{1}{xe^{H(x)/x} + x}. \tag{14}$$

Case 1: if $0 \leq \lambda \leq \varphi(\psi(1-\alpha_1))$, then the optimal solution is $\mu_2^* = 1, \mu_1^* = \psi(1-\alpha_1)$, which satisfies (8) and (9), and the corresponding rate pair is $I_1^* = H(\mu_1^*(1-\alpha_1)) - \mu_1^* H(1-\alpha_1)$, $I_2^* = 0$.

Case 2: if $\lambda \geq \varphi(1)$, then the optimal solution is $\mu_2^* = \psi(1-\alpha_2), \mu_1^* = 1$, which also satisfies (8) and (9), and the corresponding rate pair is $I_1^* = 0, I_2^* = H(\mu_2^*(1-\alpha_2)) - \mu_2^* H(1-\alpha_2)$.

Case 3: if $\varphi(\psi(1-\alpha_1)) < \lambda < \varphi(1)$, then the optimal solution given below also satisfies (8) and (9):

$$\mu_1^* = \varphi^{-1}(\lambda) = \frac{e^\lambda - 1}{e^\lambda(1-\alpha_2) - (1-\alpha_1)} \tag{15}$$

and

$$\begin{aligned}&\ln(1-\mu_1^*(1-\alpha_1))\big(H(\mu_1^*(1-\alpha_2)) - \mu_1^*(1-\alpha_2)\ln\frac{1-\mu_2^*\mu_1^*(1-\alpha_2)}{\mu_2^*\mu_1^*(1-\alpha_2)}\big)\\ &= \ln(1-\mu_1^*(1-\alpha_2))\big(H(\mu_1^*(1-\alpha_1)) - \mu_1^* H(1-\alpha_1)\big).\end{aligned} \tag{16}$$

The proof is given in Appendix B. Combining Case 1,2 and 3, we conclude that (μ_1, μ_2) is a maximizer of (12) if and only if the pair (μ_1, μ_2) satisfies (8) and (9). In other words, if (μ_1, μ_2) doesn't satisfy (8) or (9), (μ_1, μ_2) can not be a maximizer of (12), and thus the transmission strategy $(\mu_1, \mu_2, \gamma = 0)$ is not optimal. Since the set of the transmission strategies with $\gamma = 0$ is sufficient by Corollary 1, the set of all the transmission strategies satisfying (7-9) is also sufficient. Therefore the capacity region is the convex hull of the closure of all rate pairs (R_1, R_2) satisfying (10) and (11) for some μ_1, μ_2 which satisfy (8) and (9).

A sketch of the capacity region is shown with two upper bounds in Fig. 6(a). From Case 1 in Theorem 4, the point B corresponds to the largest transmission rate for the first user. The first upper bound is the tangent of the capacity region at the point B, and its slope is $-1/\varphi(\psi(1-\alpha_1))$. From Case 2, the point A provides the largest transmission rate for the second user. The second upper bound is the tangent of the capacity region at the point A, and its slope is $-1/\varphi(1)$. Case 3 gives us the optimal boundary of the capacity region except the points A and B.

Given α_1 and α_2, which completely describe a two-user degraded broadcast Z channel, the optimal boundary of the capacity region can be explicitly described by (8-11). For any μ_1 in the range of (8), the value of the unique associated μ_2 follows from (9). The curve of the optimal boundary of the capacity region is then the set of (I_1, I_2) pairs satisfying (10) and (11) for these μ_1 and associated μ_2. For example, for $\alpha_1 = 0.15$ and $\alpha_2 = 0.6$, the range of μ_1

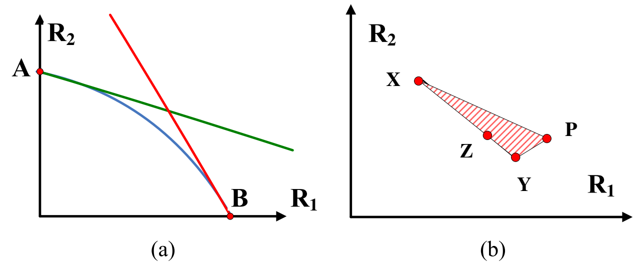

Fig. 6. (a) The capacity region and two upper bounds. (b) Point Z can not be on the boundary of the capacity region.

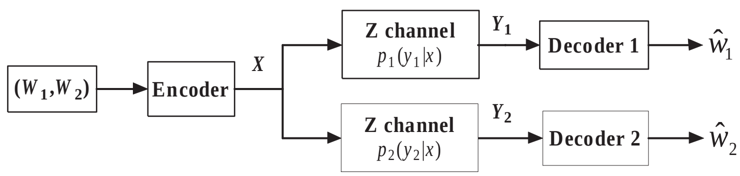

Fig. 7. Communication system for 2-user broadcast Z channels.

values is $0.445 \leq \mu_1 \leq 1$ and the associated capacity region boundary is plotted in Fig. 12.

Now we prove Theorem 2. Since we have proved that the set of all the transmission strategies satisfying (7-9) is sufficient, we only need to show that any rate pair on the optimal boundary of the capacity region can be achieved without using time sharing.

Proof by contradiction: Suppose the point Z in Fig. 6(b) is on the optimal boundary of the capacity region for the broadcast Z channel and this point can only be achieved by time sharing of the points X and Y, which can be directly achieved by using transmission strategies satisfying (7-9). Clearly, the slope of the line segment XY is neither zero nor minus infinity. Denote $-k, 0 < k < \infty$ as the slope of XY. The points X and Y provide the same value of $I_1 + \frac{1}{k}I_2$. By Theorem 4, the optimal solution to the maximization problem of $\max(I_1 + \lambda I_2)$ is unique, and so neither X nor Y maximizes $(I_1 + \frac{1}{k}I_2)$. Thus, there exists an achievable point P such that this point is on the right upper side of the line XY. Since and the triangle $\triangle XYP$ is in the capacity region, the point Z must not be on the optimal boundary of the capacity region (contradiction).$\square$

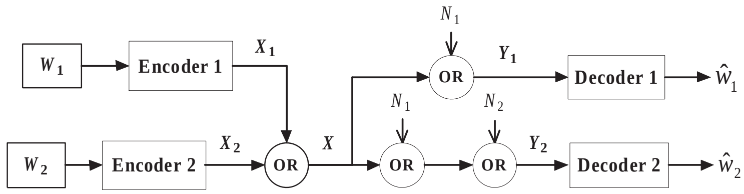

Fig. 8. Optimal transmission strategy for broadcast Z channels.

B. Independent encoding scheme

The communication system for the two-user broadcast Z channel is shown in Fig. 7. In a general scheme, the transmitter jointly encodes the independent messages W_1 and W_2, which is potentially too complex to implement. Theorem 2 demonstrates that there exists an independent encoding scheme which achieves the optimal boundary of the capacity region. Since $\gamma = 0$ is strongly sufficient, the designed channel $X_2 \rightarrow X$ is a Z channel. Thus, the broadcast signal X can be constructed as the OR of two Bernoulli random variables X_1 and X_2. This construction of X is an independent encoding scheme. The system diagram of the independent encoding scheme is shown in Fig. 8. First the messages W_1 and W_2 are encoded separately and independently. X_1 and X_2 are two binary random variables with $\Pr\{X_j = 1\} = \bar{\mu}_j$ and $\Pr\{X_j = 0\} = \mu_j$, where $\bar{\mu}_j + \mu_j = 1$ for $j = 1, 2$. The transmitter broadcasts X, which is the OR of X_1 and X_2. From Theorem 2, this independent encoding scheme with any choice of (μ_1, μ_2) satisfying (8) and (9) achieves a rate pair (I_1, I_2) arbitrarily close to the optimal boundary of the capacity region if the codes for X_1 and X_2 are properly chosen and have sufficiently large block lengths.

IV. Nonlinear-Turbo Codes for the Two-User Broadcast Z Channel

In this section we show a practical implementation of the transmission strategy for the two-user broadcast Z channel. As proved in Section III, the optimal boundary is achieved by transmitting the OR of the encoded data of each user, provided that the density of ones of each of these encoded streams is chosen properly. Hence, a family of codes that provides a controlled density of ones is required. We use the nonlinear turbo codes, introduced in

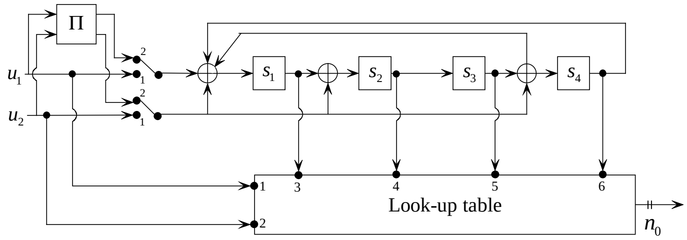

Fig. 9. 16-state nonlinear turbo code structure, with $k_0 = 2$ input bits per trellis section.

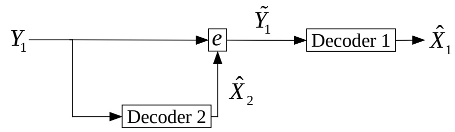

Fig. 10. Decoder structure for user 1.

[13], to provide the needed controlled density of ones. Nonlinear turbo codes are parallel concatenated trellis codes with k_0 input bits and n_0 output bits per trellis section. A look-up table assigns the output label for each branch of the trellis so that the required ones density is achieved. Each constituent encoder for the turbo code in this paper is a 16-state trellis code with $k_0 = 2$ and the trellis structure shown in Fig. 9. The output labels are assigned via a constrained search that provides the required ones density for each case, using the tools presented in [13] for the Z Channel. The output labels for the codes with rate pair $(R_1 = 1/6, R_2 = 1/6)$, which is simulated on a broadcast Z channel with $\alpha_1 = 0.15, \alpha_2 = 0.6$, are listed in Table I.

Receiver 1 uses successive decoding as shown in Fig. 10. Denote as $\hat{X}_2$ the decoded stream corresponding to user 2. Since the transmitted data is $x = x_1(\mathrm{OR})x_2$, whenever a bit $x_2 = 1$, there is no information about x_1, and x_1 can be considered an erasure. Hence, the input stream to Decoder 1 is

$$\hat{y}_1 = e(y_1, \hat{x}_2) = \begin{cases} y_1 & \text{if} \quad \hat{x}_2 = 0, \\ e & \text{if} \quad \hat{x}_2 = 1. \end{cases} \tag{17}$$

Therefore, Decoder 2 sees a Z Channel with erasures as shown in Fig. 11. The tools

TABLE I

LABELING FOR CONSTITUENT TRELLIS CODES. RATES $R_1 = 1/6, R_2 = 1/6$. ROWS REPRESENT THE STATE $s_1s_2s_3s_4$, COLUMNS REPRESENT THE INPUT u_1u_2. LABELING IN OCTAL NOTATION.

User 1				
state	input			
	00	01	10	11
0000	40	20	10	04
0001	20	40	04	10
0010	10	04	02	01
0011	04	10	01	02
0100	02	01	40	20
0101	01	02	20	40
0110	42	21	14	05
0111	21	42	05	14
1000	01	02	04	10
1001	02	01	10	04
1010	04	10	20	40
1011	10	04	40	20
1100	05	14	21	42
1101	14	05	42	21
1110	20	40	01	02
1111	40	20	02	01

User 2				
state	input			
	00	01	10	11
0000	07	34	62	51
0001	34	07	51	62
0010	25	16	43	70
0011	16	25	70	43
0100	61	13	54	26
0101	13	61	26	54
0110	23	15	52	64
0111	15	23	64	52
1000	70	43	16	25
1001	43	70	25	16
1010	51	62	34	07
1011	62	51	07	34
1100	64	52	15	23
1101	52	64	23	15
1110	26	54	13	61
1111	54	26	61	13

presented in [13] were general enough to be applied to the Z Channel with erasures. Note that if α_1 is much smaller than α_2 we can use hard decoding in Decoder 2 instead of soft decoding without any loss in performance. Since the code for user 2 is designed for a Z Channel with 0-to-1 crossover probability $1-(1-\alpha_2)\mu_1$, and the channel perceived by Decoder 2 in user 1 is a Z-Channel with crossover probability $1-(1-\alpha_1)\mu_1 < 1-(1-\alpha_2)\mu_1$, the bit error rate of $\hat{x}_2$ is negligible compared to the bit error rate of Decoder 1. In fact, in all the simulations shown in Section V, which include 100 frame errors of user 1, none of the errors were produced by Decoder 2.

V. RESULTS

We simulate the transmission strategy for the two-user broadcast Z channel with crossover probabilities $\alpha_1 = 0.15$ and $\alpha_2 = 0.6$, using nonlinear turbo codes, with the structure shown

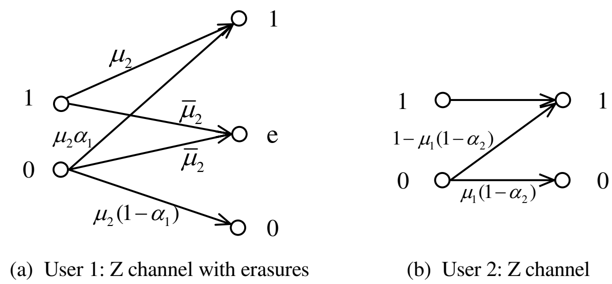

Fig. 11. Perceived channel by each decoder.

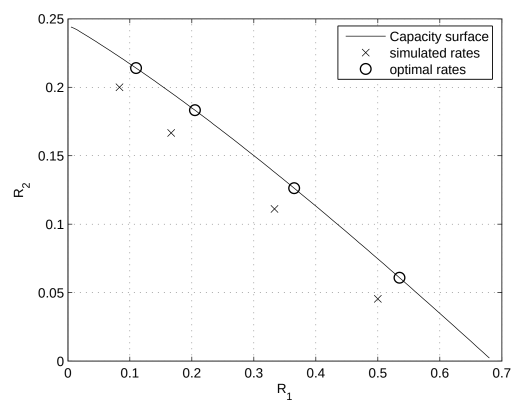

Fig. 12. Broadcast Z channel with crossover probabilities $\alpha_1 = 0.15$ and $\alpha_2 = 0.6$ for receiver 1 and 2 respectively: achievable capacity region, simulated rate pairs (R_1, R_2) and their corresponding optimal rates.

in Fig. 9. Fig. 12 shows the capacity region for the broadcast Z channel and identifies the simulated rate pairs. It also shows the optimal rate pairs which is used to compute the ones densities of each code. The output labels for the codes with each simulated rate pair are listed at [14]. For each of these four simulated rate pairs, the loss in mutual information from the associated optimal rate is only 0.04 bits or less in R_1 and only 0.02 bits or less in R_2. Table II shows bit error rates for each rate pair, the ones densities $\bar{\mu}_1$ and $\bar{\mu}_2$, and the interleaver lengths K_1 and K_2 used for each code. For simplicity, we chose K_1 and K_2 so that the codeword length n would be the same for user 1 and user 2, except for rate pairs $R_1 = 1/2$ and $R_2 = 1/22$, where one codeword length of user 2 is twice the length of user 1.

TABLE II

BER FOR TWO-USER BROADCAST Z CHANNEL WITH CROSSOVER PROBABILITIES $\alpha_1 = 0.15$ AND $\alpha_2 = 0.6$.

R_1	R_2	$\bar{\mu}_1$	$\bar{\mu}_2$	K_1	K_2	BER$_1$	BER$_2$
1/12	1/5	0.106	0.56	4800	1700	2.54×10^{-5}	1.24×10^{-5}
1/6	1/6	0.196	0.5	2048	2048	7.01×10^{-6}	5.33×10^{-6}
1/3	1/9	0.336	0.3739	4608	1536	7.13×10^{-6}	6.70×10^{-6}
1/2	1/22	0.463	0.1979	5632	1024	9.27×10^{-7}	3.27×10^{-6}

VI. Conclusions

This paper presented an optimal transmission strategy for the broadcast Z channel with independent encoding and successive decoding. We proved that any point on the optimal boundary of the capacity region can be achieved by independently encoding the messages corresponding to different users and transmitting the OR of the encoded signals. Also, the distributions of the outputs of each encoder that achieve the optimal boundary were provided. Nonlinear-turbo codes that provide a controlled distribution of ones and zeros in their codewords were used to demonstrate a low-complexity scheme that works close to the optimal boundary.

Appendices

Appendix A

Here we prove Theorem 3. In (4) and (5), denote

$$I_1(\mu_1, \mu_2, \gamma) = I(X; Y_1|X_2)\big|_{\mu_1,\mu_2,\gamma} \tag{18}$$

$$I_2(\mu_1, \mu_2, \gamma) = I(X_2; Y_2)\big|_{\mu_1,\mu_2,\gamma} \tag{19}$$

$$I_{1,2}(\mu_1, \mu_2, \gamma) = (I_1, I_2)\big|_{\mu_1,\mu_2,\gamma}. \tag{20}$$

The transmission strategy (μ_1, μ_2, γ) achieves the rate pair $I_{1,2}(\mu_1, \mu_2, \gamma)$. The theorem is true if we can increase both I_1 and I_2 when $0 < \mu_2 < 1, 0 < \gamma < \mu_1$.

Firstly we compare the strategies (μ_1, μ_2, γ) and $(\mu_1 + \bar{\mu}_2\delta_1, \mu_2, \gamma - \mu_2\delta_1)$ for a small positive number $\delta_1 > 0$.

$$\begin{aligned}
\Delta_1 I_1 &= I_1(\mu_1 + \bar{\mu}_2\delta_1, \mu_2, \gamma - \mu_2\delta_1) - I_1(\mu_1, \mu_2, \gamma) \\
&\simeq \frac{\partial I_1(\mu_1 + \bar{\mu}_2\delta_1, \mu_2, \gamma - \mu_2\delta_1)}{\partial \delta_1}\Big|_{\delta_1=0} \delta_1 \\
&= -\mu_2\bar{\mu}_2(1-\alpha_1)\left(\ln\frac{1-\gamma(1-\alpha_1)}{\gamma(1-\alpha_1)} + \ln\frac{\mu_1(1-\alpha_1)}{1-\mu_1(1-\alpha_1)}\right)\delta_1 \\
&< 0
\end{aligned} \tag{21}$$

and

$$\begin{aligned}
\Delta_1 I_2 &= I_2(\mu_1 + \bar{\mu}_2\delta_1, \mu_2, \gamma - \mu_2\delta_1) - I_2(\mu_1, \mu_2, \gamma) \\
&\simeq \frac{\partial I_2(\mu_1 + \bar{\mu}_2\delta_1, \mu_2, \gamma - \mu_2\delta_1)}{\partial \delta_1}\Big|_{\delta_1=0} \delta_1 \\
&= \mu_2\bar{\mu}_2(1-\alpha_2)\left(\ln\frac{1-\gamma(1-\alpha_2)}{\gamma(1-\alpha_2)} + \ln\frac{\mu_1(1-\alpha_2)}{1-\mu_1(1-\alpha_2)}\right)\delta_1 \\
&> 0.
\end{aligned} \tag{22}$$

The small change of the rate pair $(\Delta_1 I_1, \Delta_1 I_2)$ is shown Fig. 13. Point A is the rate pair of the transmission strategy (μ_1, μ_2, γ), the arrow Δ_1 shows the small movement of the rate pair $(\Delta_1 I_1, \Delta_1 I_2)$.

Secondly we compare the strategies (μ_1, μ_2, γ) and $(\mu_1 + (\gamma - \mu_1)\delta_2, \mu_2 + \mu_2\delta_2, \gamma)$ for a small positive number $\delta_2 > 0$.

$$\begin{aligned}\Delta_2 I_1 &= I_1(\mu_1 + (\gamma - \mu_1)\delta_2, \mu_2 + \mu_2\delta_2, \gamma) - I_1(\mu_1, \mu_2, \gamma) \\ &\simeq \frac{\partial I_1(\mu_1 + (\gamma - \mu_1)\delta_2, \mu_2 + \mu_2\delta_2, \gamma)}{\partial \delta_2}\Big|_{\delta_2 = 0} \delta_2 \\ &= -\mu_2\delta_2 \left\{ \gamma(1-\alpha_1) \ln \frac{\mu_1}{\gamma} + (1 - \gamma(1-\alpha_1)) \ln \frac{1 - \mu_1(1-\alpha_1)}{1 - \gamma(1-\alpha_1)} \right\} \\ &= \mu_2\delta_2 D(\gamma(1-\alpha_1) \parallel \mu_1(1-\alpha_1)) \\ &> 0 \end{aligned} \tag{23}$$

and

$$\begin{aligned}\Delta_2 I_2 &= I_2(\mu_1 + (\gamma - \mu_1)\delta_2, \mu_2 + \mu_2\delta_2, \gamma) - I_2(\mu_1, \mu_2, \gamma) \\ &\simeq \frac{\partial I_2(\mu_1 + (\gamma - \mu_1)\delta_2, \mu_2 + \mu_2\delta_2, \gamma)}{\partial \delta_2}\Big|_{\delta_2 = 0} \delta_2 \\ &= \mu_2\delta_2 \left\{ \gamma(1-\alpha_2) \ln \frac{\mu_1}{\gamma} + (1 - \gamma(1-\alpha_2)) \ln \frac{1 - \mu_1(1-\alpha_2)}{1 - \gamma(1-\alpha_2)} \right\} \\ &= -\mu_2\delta_2 D(\gamma(1-\alpha_2) \parallel \mu_1(1-\alpha_2)) \\ &< 0, \end{aligned} \tag{24}$$

where $D(p \parallel q)$ is the relative entropy between the distributions p and q. The arrow Δ_2 in Fig. 13 shows the small movement of the rate pair $(\Delta_2 I_1, \Delta_2 I_2)$.

Now we show that

$$\frac{\Delta_1 I_2}{\Delta_1 I_1} < \frac{\Delta_2 I_2}{\Delta_2 I_1} < 0. \tag{25}$$

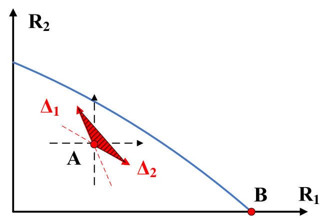

Fig. 13. Capacity region and the changes of rate pairs.

$$
\begin{aligned}
&\frac{\Delta_1 I_2}{\Delta_1 I_1} < \frac{\Delta_2 I_2}{\Delta_2 I_1} \\
&\Leftrightarrow \frac{D(\gamma(1-\alpha_2) \parallel \mu_1(1-\alpha_2)) + \ln\frac{1-\gamma(1-\alpha_2)}{1-\mu_1(1-\alpha_2)}}{D(\gamma(1-\alpha_1) \parallel \mu_1(1-\alpha_1)) + \ln\frac{1-\gamma(1-\alpha_1)}{1-\mu_1(1-\alpha_1)}} > \frac{D(\gamma(1-\alpha_2) \parallel \mu_1(1-\alpha_2))}{D(\gamma(1-\alpha_1) \parallel \mu_1(1-\alpha_1))} \\
&\Leftrightarrow \frac{D(\gamma(1-\alpha_1) \parallel \mu_1(1-\alpha_1))}{\ln\frac{1-\gamma(1-\alpha_1)}{1-\mu_1(1-\alpha_1)}} > \frac{D(\gamma(1-\alpha_2) \parallel \mu_1(1-\alpha_2))}{\ln\frac{1-\gamma(1-\alpha_2)}{1-\mu_1(1-\alpha_2)}} \\
&\Leftrightarrow f(x) = \frac{D(\gamma x \parallel \mu_1 x)}{\ln\frac{1-\gamma x}{1-\mu_1 x}} \text{is monotonically increasing in the domain of } 0 < x < 1 \\
&\Leftrightarrow f'(x) = \left\{ \ln\frac{\gamma x}{\mu_1 x} \ln\frac{1-\gamma x}{1-\mu_1 x} - \left(\ln\frac{1-\gamma x}{1-\mu_1 x}\right)^2 + \ln\frac{\gamma x}{\mu_1 x}\left(\frac{1}{1-\gamma x} - \frac{1}{1-\mu_1 x}\right) \right\} \gamma \left(\ln\frac{1-\gamma x}{1-\mu_1 x}\right)^{-2} \\
&\qquad > 0. \qquad (26)
\end{aligned}
$$

Let $a = 1 - \gamma x$ and $b = 1 - \mu_1 x$. We have $0 < b < a < 1$ and want to show that

$$g(a, b) = \ln\frac{a}{b} \ln\frac{1-a}{1-b} - \left(\ln\frac{a}{b}\right)^2 + \ln\frac{1-a}{1-b}\left(\frac{1}{a} - \frac{1}{b}\right) > 0. \qquad (27)$$

Since

$$\frac{\partial^2 g(a, b)}{\partial a \partial b} = -\frac{(a-b)^2}{a^2 b^2 (1-a)(1-b)} < 0, \qquad (28)$$

and

$$\left.\frac{\partial g(a, b)}{\partial a}\right|_{b=a} = 0 \quad \forall 0 < a < 1, \qquad (29)$$

it is true that

$$\frac{\partial g(a,b)}{\partial a} > 0 \quad \forall 0 < b < a < 1. \tag{30}$$

It implies from (30) and the fact $g(b,b) = 0, \forall 0 < b < 1$ that $g(a,b) > 0, \forall 0 < b < a < 1$. Thus, the inequality (25) is true, which means that the slope of Δ_1 is smaller than that of Δ_2 in Fig. 13. Hence, the achievable shaded region is on the upper right side of the point A. Therefore, we can increase the rate pair $I_{1,2}(\mu_1, \mu_2, \gamma)$ together and the strategy (μ_1, μ_2, γ) is not optimal when $0 < \mu_2 < 1$ and $0 < \gamma < \mu_1$.$\square$

Appendix B

Here we prove Theorem 4. In problem (12), the objective function $I_1 + \lambda I_2$ is bounded and the domain $0 \leq \mu_1, \mu_2 \leq 1$ is closed, so the maximum exists and can be attained. First we discuss some possible optimal solutions and then we show that only one of them is the optimum for any fixed $\lambda \geq 0$.

Case 0: If $\mu_1 = 0$ or $\mu_2 = 0$ or $\mu_1 = \mu_2 = 1$, then $I_1 = I_2 = 0$ and so it can not be the optimum.

Case 1: If $\mu_2 = 1$ and $0 < \mu_1 < 1$, then $I_2 = 0$ and

$$\frac{\partial I_1}{\partial \mu_1} = (1-\alpha_1)\ln\frac{1-\mu_1(1-\alpha_1)}{\mu_1(1-\alpha_1)} - H(1-\alpha_1) = 0 \tag{31}$$

$$\Rightarrow \mu_1^* = \frac{1}{(1-\alpha_1)(e^{H(1-\alpha_1)/(1-\alpha_1)}+1)}. \tag{32}$$

Case 2: If $\mu_1 = 1$ and $0 < \mu_2 < 1$, then $I_1 = 0$ and

$$\frac{\partial I_2}{\partial \mu_2} = (1-\alpha_2)\ln\frac{1-\mu_2(1-\alpha_2)}{\mu_2(1-\alpha_2)} - H(1-\alpha_2) = 0 \tag{33}$$

$$\Rightarrow \mu_2^* = \frac{1}{(1-\alpha_2)(e^{H(1-\alpha_2)/(1-\alpha_2)}+1)}. \tag{34}$$

Case 3: If $0 < \mu_1, \mu_2 < 1$, then the optimum is attained when

$$\mu_2 \frac{\partial(I_1 + \lambda I_2)}{\partial \mu_2} - \mu_1 \frac{\partial(I_1 + \lambda I_2)}{\partial \mu_1} = 0$$

$$\Rightarrow \ln(1 - \mu_1^*(1 - \alpha_1)) = \lambda \ln(1 - \mu_1^*(1 - \alpha_2)), \tag{35}$$

and

$$\begin{aligned}
&\frac{\partial(I_1 + \lambda I_2)}{\partial \mu_2} = 0 \\
\Rightarrow &\lambda \left\{ H(\mu_1^*(1 - \alpha_2)) - \mu_1^*(1 - \alpha_2) \ln \frac{1 - \mu_2^* \mu_1^*(1 - \alpha_2)}{\mu_2^* \mu_1^*(1 - \alpha_2)} \right\} = \big(H(\mu_1^*(1 - \alpha_1)) - \mu_1^* H(1 - \alpha_1)\big) \\
\Rightarrow &\ln(1 - \mu_1^*(1 - \alpha_1)) \left\{ H(\mu_1^*(1 - \alpha_2)) - \mu_1^*(1 - \alpha_2) \ln \frac{1 - \mu_2^* \mu_1^*(1 - \alpha_2)}{\mu_2^* \mu_1^*(1 - \alpha_2)} \right\} \\
&= \ln(1 - \mu_1^*(1 - \alpha_2)) \big(H(\mu_1^*(1 - \alpha_1)) - \mu_1^* H(1 - \alpha_1)\big).
\end{aligned} \tag{36}$$

For any fixed $\lambda \geq 0$, the optimal solution is in Case 1,2 or 3.

Lemma 1: Function $\varphi(x) = \frac{\ln(1-(1-\alpha_1)x)}{\ln(1-(1-\alpha_2)x)}$ is monotonically increasing in the domain of $0 \leq x \leq 1$ when $\alpha_1 < \alpha_2$.

Lemma 2: The solution in Case 1 can not be the optimum when $\lambda > \varphi(\psi(1 - \alpha_1))$.

Proof: When $\mu_2 = 1$ and $\mu_1 = \psi(1 - \alpha_1)$, $\frac{\partial I_2}{\partial \mu_1} = 0$ and $\frac{\partial I_1}{\partial \mu_1} = 0$. Therefore, for any fixed λ, $\frac{\partial(I_1 + \lambda I_2)}{\partial \mu_1} = 0$. When $\lambda = \varphi(\mu_1) = \varphi(\psi(1 - \alpha_1))$, (35) holds, and so

$$\frac{\partial(I_1 + \lambda I_2)}{\partial \mu_2}\Big|_{\mu_2=1, \mu_1=\psi(1-\alpha_1)} = \frac{\partial(I_1 + \varphi(\psi(1 - \alpha_1)) I_2)}{\partial \mu_2}\Big|_{\mu_2=1, \mu_1=\psi(1-\alpha_1)} = 0.$$

Since $\frac{\partial I_2}{\partial \mu_2}\Big|_{\mu_2=1,\mu_1=\psi(1-\alpha_1)} = \ln(1-\psi(1-\alpha_1)\cdot(1-\alpha_2)) < 0$, when $\lambda > \varphi(\psi(1-\alpha_1))$,

$$\begin{aligned}
&\frac{\partial(I_1+\lambda I_2)}{\partial \mu_2}\Big|_{\mu_2=1,\mu_1=\psi(1-\alpha_1)} \\
&= \frac{\partial I_1}{\partial \mu_2}\Big|_{\mu_2=1,\mu_1=\psi(1-\alpha_1)} + \lambda\frac{\partial I_2}{\partial \mu_2}\Big|_{\mu_2=1,\mu_1=\psi(1-\alpha_1)} \\
&< \frac{\partial I_1}{\partial \mu_2}\Big|_{\mu_2=1,\mu_1=\psi(1-\alpha_1)} + \varphi(\psi(1-\alpha_1))\frac{\partial I_2}{\partial \mu_2}\Big|_{\mu_2=1,\mu_1=\psi(1-\alpha_1)} \\
&= \frac{\partial(I_1+\varphi(\psi(1-\alpha_1))I_2)}{\partial \mu_2}\Big|_{\mu_2=1,\mu_1=\psi(1-\alpha_1)} \\
&= 0.
\end{aligned} \tag{37}$$

Therefore, Case 1 can not be the optimum when $\lambda > \varphi(\psi(1-\alpha_1))$. □

Lemma 3: The solution in Case 2 can not be the optimum when $\lambda < \varphi(1)$.

Proof: When $\mu_2 = \psi(1-\alpha_2)$ and $\mu_1 = 1$, $\frac{\partial I_2}{\partial \mu_2} = 0$ and $\frac{\partial I_1}{\partial \mu_2} = 0$. Therefore, for any fixed λ, $\frac{\partial(I_1+\lambda I_2)}{\partial \mu_2} = 0$. When $\lambda = \varphi(\mu_1) = \varphi(1)$, (35) holds, and so

$$\frac{\partial(I_1+\lambda I_2)}{\partial \mu_1}\Big|_{\mu_2=\psi(1-\alpha_2),\mu_1=1} = \frac{\partial(I_1+\varphi(1) I_2)}{\partial \mu_1}\Big|_{\mu_2=\psi(1-\alpha_2),\mu_1=1} = 0. \tag{38}$$

Since $\frac{\partial I_2}{\partial \mu_1}\Big|_{\mu_2=\psi(1-\alpha_2),\mu_1=1} = -\psi(1-\alpha_2)\ln\alpha_2 > 0$, when $\lambda < \varphi(1)$,

$$\begin{aligned}
&\frac{\partial(I_1+\lambda I_2)}{\partial \mu_1}\Big|_{\mu_2=\psi(1-\alpha_2),\mu_1=1} \\
&= \frac{\partial I_1}{\partial \mu_2}\Big|_{\mu_2=\psi(1-\alpha_2),\mu_1=1} + \lambda\frac{\partial I_2}{\partial \mu_2}\Big|_{\mu_2=\psi(1-\alpha_2),\mu_1=1} \\
&< \frac{\partial I_1}{\partial \mu_2}\Big|_{\mu_2=\psi(1-\alpha_2),\mu_1=1} + \varphi(1)\frac{\partial I_2}{\partial \mu_2}\Big|_{\mu_2=\psi(1-\alpha_2),\mu_1=1} \\
&= \frac{\partial(I_1+\varphi(1) I_2)}{\partial \mu_1}\Big|_{\mu_2=\psi(1-\alpha_2),\mu_1=1} \\
&= 0.
\end{aligned} \tag{39}$$

Therefore, Case 2 can not be the optimum when $\lambda < \varphi(1)$. □

Lemma 4: The solution to (35) exists in $(0,1)$ and is unique for any λ in the range of $\varphi(0) < \lambda < \varphi(1)$.

Proof: (35) is equivalent to $\varphi(\mu_1^*) = \lambda$. From Lemma 1, $\varphi(\mu_1)$ is monotonically increasing. Therefore, when $\varphi(0) < \lambda < \varphi(1)$, the solution $\mu_1^*$is unique and $\mu_1^* \in (0,1)$.□

Lemma 5: The unique solution (μ_1^*, μ_2^*) to (35) and (36) in Case 3 is the optimum if $\varphi(\psi(1-\alpha_1)) < \lambda < \varphi(1)$.

Proof: From Lemma 4, the solution μ_1^* to (35) is unique if $\varphi(\psi(1-\alpha_1)) < \lambda < \varphi(1)$. From (36),

$$\begin{aligned} m(\mu_2) = &\big(H(\mu_1^*(1-\alpha_2)) - \mu_1^*(1-\alpha_2)\ln\frac{1-\mu_2\mu_1^*(1-\alpha_2)}{\mu_2\mu_1^*(1-\alpha_2)}\big)\ln(1-\mu_1^*(1-\alpha_1)) \\ &- \big(H(\mu_1^*(1-\alpha_1)) - \mu_1^* H(1-\alpha_1)\big)\ln(1-\mu_1^*(1-\alpha_2)) \\ &= 0. \end{aligned} \tag{40}$$

Clearly, $m(\mu_2)$ is monotonically increasing,

$$\lim_{\mu_2 \to 0} m(\mu_2) = -\infty < 0, \tag{41}$$

and

$$\begin{aligned} &\varphi(\psi(1-\alpha_1)) < \lambda < \varphi(1) \\ \Rightarrow &\mu_1^* > \psi(1-\alpha_1) \\ \Rightarrow &m(1) > 0. \end{aligned} \tag{42}$$

That means the unique solution μ_2^* to (36) is in the domain of $0 \le \mu_2 \le 1$. Furthermore, when $\varphi(\psi(1-\alpha_1)) < \lambda < \varphi(1)$, by Lemma 2 and Lemma 3, Case 1 or Case 2 can not be the optimum because

$$\frac{\partial(I_1+\lambda I_2)}{\partial\mu_2}\Big|_{\mu_2=1,\mu_1=\psi(1-\alpha_1)} < 0, \tag{43}$$

$$\frac{\partial(I_1+\lambda I_2)}{\partial\mu_1}\Big|_{\mu_1=1,\mu_2=\psi(1-\alpha_2)} < 0. \tag{44}$$

Therefore, Case 3 is the optimum.□

Lemma 6: The unique solution $(\mu_2^*=1, \mu_1^*=\psi(1-\alpha_1))$ in Case 1 is the optimum if $0 \le \lambda \le \varphi(\psi(1-\alpha_1))$.

Proof: When $0 \le \lambda \le \varphi(\psi(1-\alpha_1))$, Case 3 is not optimal because there is no solution $\mu_1 \in (0,1)$ to (35). Case 2 is not optimal by Lemma 3. Hence, Case 1 is the optimum.□

Lemma 7: The unique solution $(\mu_2^*=\psi(1-\alpha_2), \mu_1^*=1)$ in Case 2 is the optimum if $\lambda \ge \varphi(1)$.

Proof: When $\lambda \ge \varphi(1)$, Case 3 is not optimal because there is no solution $\mu_2 \in (0,1)$ to (36). Case 1 is not optimal by Lemma 2. Hence, Case 2 is the optimum.□

From Lemma 5,6 and 7, Theorem 4 is immediately proved.□

References

[1] T. M. Cover. Broadcast channels. *IEEE Trans. Inform. Theory*, IT-18:2–14, January 1972.

[2] P. P. Bergmans. Random coding theorem for broadcast channels with degraded components. *IEEE Trans. Inform. Theory*, IT-19:197–207, March 1973.

[3] P. P. Bergmans. A simple converse for broadcast channels with additive white Gaussian noise. *IEEE Trans. Inform. Theory*, IT-20:279–280, March 1974.

[4] R. G. Gallager. Capacity and coding for degraded broadcast channels. *Probl. Pered. Inform.*, 10:3–14, July–Sept. 1974.

[5] T. M. Cover. Comments on broadcast channels. *IEEE Trans. Inform. Theory*, 44:2524–2530, October 1998.

[6] C. Berrou, A. Glavieux and P. Thitimajshima. Near shannon limit error-correcting coding and decoding: turbo-codes. *Proc. ICC'93*, pages 873–890, May 1993.

[7] R. G. Gallager. *Low-Density Parity-Check Codes.* PhD thesis, MIT, Cambridge, MA, 1963.

[8] P. Berlin and D. Tuninetti. LDPC codes for Gaussian broadcast channels. *Signal Proc. Advances in Wireless Commun., 2004 IEEE 5th Workshop on*, pages 444–448, 2004.

[9] T. W. Sun, R. D. Wesel, M. R. Shane and K. Jarett. Superposition turbo-TCM for multi-rate broadcast. *IEEE Trans. on Commun.*, 52:368–371, 2004.

[10] P. R. Chevillat. N-user trellis coding for a class of multiple-access channels. *IEEE Trans. on Info. Theo.*, IT-27:114–120, 1981.

[11] S. W. Golomb. The limiting behavior of the Z-channel. *IEEE Trans. Inform. Theory*, IT-26:372, May 1980.

[12] M. Griot, A.I. Vila Casado, W-Y. Weng, H. Chan, J. Basak, E. Yablanovitch, I. Verbauwhede, B. Jalali and R. D. Wesel. Trellis codes with low ones density for the OR multiple access channel. In *IEEE ISIT 2006*, July 2006.

[13] M. Griot, A.I. Vila Casado, and R.D. Wesel. Non-linear turbo codes for interleaver-division multiple access on the OR channel. In *GLOBECOM '06. IEEE Global Telecomm. Conf.*, 27 Nov. - 1 Dec. 2006.

[14] Nonlinear turbo codes for broadcast Z channels. *[Online]. Available: http://www.ee.ucla.edu/˜csl/files/codes/bzc.html.*