Coding Theorem for Secret Sharing Communication Systems with Two Noisy Channels

HIROSUKE YAMAMOTO, MEMBER, IEEE

Abstract—**The coding theorem is proved for the secret sharing communication system (SSCS) with two noisy channels, each of which is a broadcast channel characterized by $P(y_j z_j|x_j)$, $j=1,2$; it is assumed that the legitimate channel $(X_j \to Y_j)$ is less noisy than the wiretapped channel $(X_j \to Z_j)$. The code (f,ϕ) for the SSCS is defined by two mappings: $(X_1^{N_1}, X_2^{N_2}) = f(S^K, T)$ and $\hat{S} = \phi(Y_1^{N_1}, Y_2^{N_2})$ where T is an arbitrarily chosen random number and S is an independent identically distributed source output that must be transmitted to the legitimate receiver with an arbitrarily small error. The rate of each channel is given by N_j/K while the security level for each wiretapper can be evaluated by $(1/K)H(S^K|Z_j^{N_j})$. The admissible region of rates and security levels is given by a "single-letter characterization."**

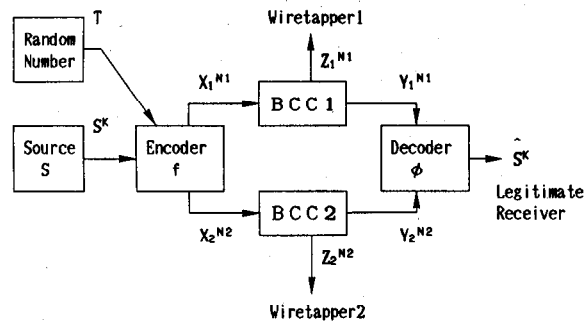

Fig. 1. SSCS with two noisy channels.

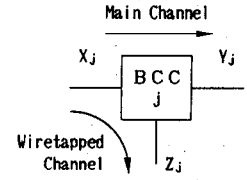

Fig. 2. Broadcast channel.

I. Introduction

TO ATTAIN safe information transmission via several parallel channels, the secret sharing communication system (SSCS) [1] has been studied as an extension of both Shannon's cipher system [2] and the secret sharing system [3]. The coding theorem was proved for the SSCS with two or three channels in [1], but the channels were assumed to be noiseless.

In this paper the SSCS with two noisy channels (shown in Fig. 1) is considered. The specifications of the system are as follows. The source S is a finite memoryless source. The information S must be transmitted to the legitimate receiver without errors via two noisy channels (BCC1 and BCC2). Because unauthorized persons may eavesdrop on the information S via the noisy channel, the information S must be kept as secret from them as possible. Each noisy channel can be regarded as a discrete memoryless broadcast channel (BCC) $P(y_j z_j|x_j)$, $j=1,2$ (shown in Fig. 2), which consists of a main channel $(X_j \to Y_j)$ and a wiretapped channel $(X_j \to Z_j)$. We assume for simplicity that each main channel is less noisy than each wiretapped channel, respectively. The relation "Channel $(X_j \to Y_j)$ is less noisy than channel $(X_j \to Z_j)$" means that for every random variable forming a Markov chain $V_j \to X_j \to Y_j Z_j$, the following inequality holds:

$$I(V_j; Y_j) \geq I(V_j; Z_j) \tag{1}$$

(see [4] for details on the notion of "less noisy").

To realize a secure transmission in this system, we use a block code. Since the two channels do not have the same characteristics in general, a different codeword length is used for each BCC (say N_1 and N_2 for BCC1 and BCC2, respectively) per K source output symbols. The encoder f can utilize an arbitrary random number T besides the source output S^K to randomize the codewords $X_1^{N_1} = (X_{11}, X_{12}, \cdots, X_{1N_1})$ and $X_2^{N_2} = (X_{21}, X_{22}, \cdots, X_{2N_2})$. Since T can be chosen arbitrarily, the encoder f can be restricted to deterministic functions without loss of generality. Hence $X_1^{N_1}$ and $X_2^{N_2}$ can be assumed to be uniquely determined from both S^K and T by the encoder f. The decoder reproduces $\hat{S}^K$ from both $Y_1^{N_1}$ and $Y_2^{N_2}$. The security level of this system is measured by $((1/K)H(S^K|Z_1^{N_1}), (1/K)H(S^K|Z_2^{N_2}))$.

In this paper the coding theorem for the foregoing system is proved, and the admissible region of rates and security levels is completely obtained. The precise statement of the problem and the main results are given in

Manuscript received February 15, 1988; revised July 18, 1988. This paper was presented in part at the IEEE International Symposium on Information Theory, Kobe, Japan, June 1988.

The author is with the Department of Communications and Systems, University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, 182 Japan.

IEEE Log Number 8927894.

Section II. The theorem is proved in Section III. A binary example is given in Section IV.

The SSCS with two noisy channels was discussed earlier as a generalization of the SSCS with two noiseless channels. However (see Fig. 1), if only one channel is considered, the system reduces to the wiretap channel system, studied by Wyner [5] and extended by Csiszár and Körner [4]. Hence the SSCS with two noisy channels can be regarded as an extension of the concept of the wiretap channel. The relation of our results to previously known ones [1], [4], [5] is examined in Section II.

II. Formal Statement of the Problem and Main Results

Let the source output $\{S_k\}_{k=1}^{\infty}$ be a sequence of independent identically distributed (i.i.d.) random variables (RV's), S taking values in a finite discrete set $\mathcal{S}$. Each channel in Fig. 1 is a broadcast channel (shown in Fig. 2) characterized by $P(y_j z_j|x_j)$, $x_j \in \mathcal{X}_j$, $y_j \in \mathcal{Y}_j$, $z_j \in \mathcal{Z}_j$ $(j=1,2)$ where $\mathcal{X}_j, \mathcal{Y}_j, \mathcal{Z}_j$ are finite discrete sets. The BCC has the property that the main channel $(X_j \to Y_j)$ is less noisy than the wiretapped channel $(X_j \to Z_j)$. The code (f,ϕ) is defined by two mappings:

$$f\colon \mathcal{S}^K \times \mathcal{T} \to \mathcal{X}_1^{N_1} \times \mathcal{X}_2^{N_2} \tag{2}$$

$$\phi\colon \mathcal{Y}_1^{N_1} \times \mathcal{Y}_2^{N_2} \to \mathcal{S}^K \tag{3}$$

$$\left(\boldsymbol{X}_1^{N_1}, \boldsymbol{X}_2^{N_2}\right) = f(\boldsymbol{S}^K, T) \tag{4}$$

$$\hat{\boldsymbol{S}}^K = \phi\left(\boldsymbol{Y}_1^{N_1}, \boldsymbol{Y}_2^{N_2}\right) \tag{5}$$

where T is a random number and takes values in a discrete finite set $\mathcal{T}$. T and $\mathcal{T}$ can be chosen arbitrarily.

The rate of each channel is given by N_j/K. The decoding error of the legitimate receiver can be evaluated by $E(1/K)D_H(\boldsymbol{S}^K, \hat{\boldsymbol{S}}^K)$ where D_H is the Hamming distance function while the security level of $\boldsymbol{S}^K$ for each eavesdropper can be evaluated by $(1/K)H(\boldsymbol{S}^K|\boldsymbol{Z}_j^{N_j})$.

Definition 1: (R_1, R_2, h_1, h_2) is admissible for the SSCS shown in Fig. 1 if a code (f,ϕ) and a random number T exist such that for any given $\epsilon > 0$ and K sufficiently large,

$$\frac{N_j}{K} \le R_j + \epsilon \tag{6}$$

$$\frac{1}{K} H\left(\boldsymbol{S}^K|\boldsymbol{Z}_j^{N_j}\right) \ge h_j - \epsilon, \left(0 \le h_j \le H(S)\right) \tag{7}$$

$$E\frac{1}{K} D_H\left(\boldsymbol{S}^K, \phi\left(\boldsymbol{Y}_1^{N_1}, \boldsymbol{Y}_2^{N_2}\right)\right) \le \epsilon. \tag{8}$$

Definition 2: The admissible region $\mathcal{R}_{\text{sscs}}$ is defined by

$$\mathcal{R}_{\text{sscs}} \triangleq \left\{(R_1, R_2, h_1, h_2)\colon (R_1, R_2, h_1, h_2) \text{ is admissible}\right\}. \tag{9}$$

Our main results are as follows:

Theorem 1: Assume that the main channel is less noisy than the wiretapped channel in each BCC. Then,

$$(R_1, R_2, h_1, h_2) \in \mathcal{R}_{\text{sscs}}, 0 \le h_1, h_2 \le H(S) \tag{10}$$

if and only if (iff) there exist constants r_j and RV's X_j, Y_j, Z_j $(j=1,2)$ such that

$$0 \le \frac{r_j}{R_j} \le I(X_j; Y_j) \tag{11}$$

$$\frac{h_j - r_i}{R_j} \le I(X_j; Y_j) - I(X_j; Z_j) \tag{12}$$

$$r_1 + r_2 \ge H(S), \qquad (i,j) = (1,2),(2,1). \tag{13}$$

Proof: See Section III.

Corollary 1: Under the same condition as in Theorem 1,

$$(R_1, R_2, h_1, h_2) \in \mathcal{R}_{\text{sscs}}, 0 \le h_1, h_2 \le H(S)$$

if and only if there exist RV's X_j, Y_j, Z_j $(j=1,2)$ such that

$$h_1 \le \left[I(X_1; Y_1) - I(X_1; Z_1)\right] R_1 + I(X_2; Y_2) R_2 \tag{14}$$

$$h_2 \le \left[I(X_2; Y_2) - I(X_2; Z_2)\right] R_2 + I(X_1; Y_1) R_1 \tag{15}$$

$$H(S) \le I(X_1; Y_1) R_1 + I(X_2; Y_2) R_2. \tag{16}$$

Proof: Equations (11)–(13) can be rewritten as follows:

$$h_j \le \left[I(X_j; Y_j) - I(X_j; Z_j)\right] R_j + r_i \tag{17}$$

$$0 \le r_j \le I(X_j; Y_j) R_j \tag{18}$$

$$H(S) \le r_1 + r_2. \tag{19}$$

Hence if (R_1, R_2, h_1, h_2) satisfies (11)–(13), it also satisfies (14)–(16). On the other hand, if (R_1, R_2, h_1, h_2) satisfies (14)–(16), it also satisfies (11)–(13) with $r_j = I(X_j; Y_j) R_j$.

The secrecy capacity or secrecy capacity region can be defined when the information can be sent to the legitimate receiver in perfect secrecy (cf. [4], [5]). In the SSCS with two BCC's, the information $\boldsymbol{S}^K$ can be kept secret from each eavesdropper perfectly if $h_1 = h_2 = H(S)$. Hence we define the secrecy capacity region $\mathcal{R}_{\text{sscs}}^S$ as follows:

$$\mathcal{R}_{\text{sscs}}^S \triangleq \left\{(R_1, R_2)\colon (R_1, R_2, H(S), H(S)) \in \mathcal{R}_{\text{sscs}}\right\}. \tag{20}$$

We can then easily obtain the following corollary.

Corollary 2: Under the same condition as in Theorem 1, $(R_1, R_2) \in \mathcal{R}_{\text{sscs}}^S$ if and only if there exist RV's X_j, Y_j, Z_j $(j=1,2)$ such that

$$H(S) \le \left[I(X_1; Y_1) - I(X_1; Z_1)\right] R_1 + I(X_2; Y_2) R_2 \tag{21}$$

$$H(S) \le \left[I(X_2; Y_2) - I(X_2; Z_2)\right] R_2 + I(X_1; Y_1) R_1. \tag{22}$$

We now investigate the special cases of Fig. 1. In the case that the eavesdropper can obtain the same information as the legitimate receiver, the admissible region reduces to

$$h_1 \le I(X_2; Y_2) R_2 \tag{23}$$

$$h_2 \le I(X_1; Y_1) R_1 \tag{24}$$

$$H(S) \le I(X_1; Y_1) R_1 + I(X_2; Y_2) R_2 \tag{25}$$

by substituting $Z_j = Y_j$ $(j=1,2)$ in (14)–(16). In the case in which the two channels are noiseless, we can obtain the

admissible region by letting $X_j = Y_j = Z_j$ $(j=1,2)$ as follows:

$$h_1 \le \log|\mathcal{X}_1| R_1 \tag{26}$$

$$h_2 \le \log|\mathcal{X}_2| R_2 \tag{27}$$

$$H(S) \le \log|\mathcal{X}_1| R_1 + \log|\mathcal{X}_2| R_2. \tag{28}$$

The region (26)–(28) coincides with the admissible region of the SSCS with two noiseless channels [1]. Furthermore, we note from (23)–(25) that the admissible region in the case $Y_j = Z_j$ can be achieved by concatenating the usual error correcting code and the code for the SSCS with two noiseless channels.

In the case that only one channel is available, say BCC1, we get the admissible region by substituting $R_2 = 0$ and $h_2 = H(S)$ in (14)–(16) as follows:

$$\frac{h_1}{R_1} \le \frac{H(S)}{R_1} \le I(X_1; Y_1) \tag{29}$$

$$0 \le \frac{h_1}{R_1} \le I(X_1; Y_1) - I(X_1; Z_1). \tag{30}$$

This region coincides with the result derived by Csiszár–Körner [4, cor. 4][1], which is a generalized version of the wiretap channel coding theorem [5].

III. Proof of Theorem 1

A. Proof of the Converse Part of Theorem 1[2]

If $(R_1, R_2, h_1, h_2) \in \mathcal{R}_{\text{sscs}}$, then the code (f, ϕ) and a random number T exist that satisfy (6)–(8). Let $\boldsymbol{X}_j, \boldsymbol{Y}_j, \boldsymbol{Z}_j$ $(j=1,2)$ be the RV's characterized by the code (f, ϕ). Since $\boldsymbol{X}_1, \boldsymbol{X}_2$ are uniquely determined from $(\boldsymbol{S}, T)$ by the encoder f, a Markov chain relationship $\boldsymbol{Y}_1 \to \boldsymbol{X}_1 \to \boldsymbol{S}T \to \boldsymbol{X}_2 \to \boldsymbol{Y}_2$ holds. Hence we have

$$\begin{aligned} I(\boldsymbol{S}T; \boldsymbol{Y}_1) &= I(\boldsymbol{S}T; \boldsymbol{Y}_1|\boldsymbol{Y}_2) + I(\boldsymbol{Y}_1; \boldsymbol{Y}_2) \\ &\ge I(\boldsymbol{S}T; \boldsymbol{Y}_1|\boldsymbol{Y}_2) \\ &= I(\boldsymbol{S}T; \boldsymbol{Y}_1\boldsymbol{Y}_2) - I(\boldsymbol{S}T; \boldsymbol{Y}_2) \end{aligned} \tag{31}$$

where the first equality follows from [4, lemma 1] or [6, (2.3.18)]. From (8) and Fano's inequality we obtain

$$\begin{aligned} H(\boldsymbol{S}|\boldsymbol{Y}_1\boldsymbol{Y}_2) &\le \Pr\{\boldsymbol{S} \ne \phi(\boldsymbol{Y}_1, \boldsymbol{Y}_2)\} \log(|\mathcal{S}^K| - 1) \\ &\quad + h(\Pr\{\boldsymbol{S} \ne \phi(\boldsymbol{Y}_1, \boldsymbol{Y}_2)\}) \\ &\le K\epsilon \log|\mathcal{S}| + h(\epsilon) \triangleq K\epsilon_0 \end{aligned} \tag{32}$$

where $\epsilon_0 \to 0$ as $\epsilon \to 0$. Combining (31) and (32), the following inequality holds:

$$\begin{aligned} &I(\boldsymbol{S}T; \boldsymbol{Y}_1) + I(\boldsymbol{S}T; \boldsymbol{Y}_2) \\ &\quad \ge I(\boldsymbol{S}T; \boldsymbol{Y}_1\boldsymbol{Y}_2) \\ &\quad = H(\boldsymbol{S}T) - H(\boldsymbol{S}T|\boldsymbol{Y}_1\boldsymbol{Y}_2) \\ &\quad = H(\boldsymbol{S}) + H(T|\boldsymbol{S}) - H(T|\boldsymbol{S}\boldsymbol{Y}_1\boldsymbol{Y}_2) - H(\boldsymbol{S}|\boldsymbol{Y}_1\boldsymbol{Y}_2) \\ &\quad \ge H(\boldsymbol{S}) + I(T; \boldsymbol{Y}_1\boldsymbol{Y}_2|\boldsymbol{S}) - K\epsilon_0. \end{aligned} \tag{33}$$

[1] In [4], the rate is defined as $R_1 = K/N_1$ as opposed to $R_1 = N_1/K$ in this paper.

[2] In this section superscripts on vectors are omitted for simplicity.

We now define r_j $(j=1,2)$ by

$$Kr_j \triangleq I(\boldsymbol{S}T; \boldsymbol{Y}_j). \tag{34}$$

Then from (33) we get

$$\begin{aligned} r_1 + r_2 &\ge \frac{1}{K}[H(\boldsymbol{S}) + I(T; \boldsymbol{Y}_1\boldsymbol{Y}_2|\boldsymbol{S}) - K\epsilon_0] \\ &\ge H(S) - \epsilon_0. \end{aligned} \tag{35}$$

Furthermore, the following inequality can be obtained from (33):

$$\begin{aligned} Kr_1 &= I(\boldsymbol{S}T; \boldsymbol{Y}_1) \\ &\ge H(\boldsymbol{S}) + I(T; \boldsymbol{Y}_1\boldsymbol{Y}_2|\boldsymbol{S}) - K\epsilon_0 - I(\boldsymbol{S}T; \boldsymbol{Y}_2) \\ &= H(\boldsymbol{S}|\boldsymbol{Z}_2) + I(\boldsymbol{S}; \boldsymbol{Z}_2) - I(\boldsymbol{S}; \boldsymbol{Y}_2) \\ &\quad - I(T; \boldsymbol{Y}_2|\boldsymbol{S}) + I(T; \boldsymbol{Y}_1\boldsymbol{Y}_2|\boldsymbol{S}) - K\epsilon_0 \\ &\ge H(\boldsymbol{S}|\boldsymbol{Z}_2) + I(\boldsymbol{S}; \boldsymbol{Z}_2) - I(\boldsymbol{S}; \boldsymbol{Y}_2) - K\epsilon_0 \\ &= H(\boldsymbol{S}|\boldsymbol{Z}_2) + \sum_{t=1}^{N_2} [I(\boldsymbol{S}; Z_{2t}|\boldsymbol{Y}_2^{t-1}\tilde{\boldsymbol{Z}}_2^{t+1}) \\ &\quad - I(\boldsymbol{S}; Y_{2t}|\boldsymbol{Y}_2^{t-1}\tilde{\boldsymbol{Z}}_2^{t+1})] - K\epsilon_0 \end{aligned} \tag{36}$$

where

$$\boldsymbol{Y}_2^{t-1} = (Y_{21}, Y_{22}, \cdots, Y_{2t-1}),$$

$$\tilde{\boldsymbol{Z}}_2^{t+1} = (Z_{2t+1}, Z_{2t+2}, \cdots, Z_{2N_2}),$$

and the last equality follows from [4, lemma 7].

To simplify (36), we follow the technique shown in [4]. Let us introduce an RV J independent of S, T, Y_2, and Z_2 and distributed over $\{1,2,\cdots,N_2\}$.

Set

$$\begin{aligned} &U_2 \triangleq \boldsymbol{Y}_2^{J-1}\tilde{\boldsymbol{Z}}_2^{J+1}J, \qquad V_2 \triangleq U_2\boldsymbol{S} \\ &X_2 \triangleq X_{2J}, \qquad Y_2 \triangleq Y_{2J}, \qquad Z_2 \triangleq Z_{2J}. \end{aligned} \tag{37}$$

Then (36) becomes

$$Kr_1 \ge H(\boldsymbol{S}|\boldsymbol{Z}_2) + N_2[I(V_2; Z_2|U_2) - I(V_2; Y_2|U_2)] - K\epsilon_0. \tag{38}$$

Since $U_2 \to V_2 \to X_2 \to Y_2Z_2$ is a Markov chain, applying [4, lemma 1] we get

$$\begin{aligned} Kr_1 &\ge H(\boldsymbol{S}|\boldsymbol{Z}_2) + N_2[I(V_2; Z_2) - I(U_2; Z_2) \\ &\quad - I(V_2; Y_2) + I(U_2; Y_2)] - K\epsilon_0 \\ &= H(\boldsymbol{S}|\boldsymbol{Z}_2) + N_2[I(X_2; Z_2) \\ &\quad - I(X_2; Z_2|V_2) - I(U_2; Z_2) \\ &\quad - I(X_2; Y_2) + I(X_2; Y_2|V_2) + I(U_2; Y_2)] - K\epsilon_0. \end{aligned} \tag{39}$$

Since the main channel $(X_j \to Y_j)$ is less noisy than the wiretapped channel $(X_j \to Z_j)$,

$$I(U_2; Y_2) \ge I(U_2; Z_2) \tag{40}$$

$$I(X_2; Y_2|V_2) \ge I(X_2; Z_2|V_2). \tag{41}$$

Hence

$$Kr_1 \ge H(\boldsymbol{S}|\boldsymbol{Z}_2) + N_2[I(X_2; Z_2) - I(X_2; Y_2)] - K\epsilon_0. \tag{42}$$

On the other hand, Kr_1 is upper bounded as follows:

$$\begin{aligned} 0 \le Kr_1 &\le I(\boldsymbol{ST}; \boldsymbol{Y}_1) \\ &\le {}^1 I(\boldsymbol{X}_1; \boldsymbol{Y}_1) \\ &= H(\boldsymbol{Y}_1) - H(\boldsymbol{Y}_1|\boldsymbol{X}_1) \\ &= \sum_{t=1}^{N_1} \left[H(Y_{1t}|\boldsymbol{Y}_1^{t-1}) - H(Y_{1t}|\boldsymbol{Y}_1^{t-1}\boldsymbol{X}_1) \right] \\ &\le {}^2 \sum_{t=1}^{N_1} \left[H(Y_{1t}) - H(Y_{1t}|X_{1t}) \right] \\ &= N_1 I(X_1; Y_1|J) \\ &\le {}^3 N_1 I(X_1; Y_1) \end{aligned} \tag{43}$$

where inequalities 1–3 hold for the following reasons:

1) $ST \to X_1 \to Y_1$ is a Markov chain.
2) the BCC is memoryless;
3) $J \to X_1 \to Y_1$ is a Markov chain.

Since $\boldsymbol{S}, \boldsymbol{Z}_j, N_1, K$ satisfy (6) and (7), the following inequalities are established from (42) and (43), respectively:

$$\begin{aligned} h_2 - r_1 &\le \frac{1}{K} H(\boldsymbol{S}|\boldsymbol{Z}_2) + \epsilon - r_1 \\ &\le \frac{N_2}{K} \left[I(X_2; Y_2) - I(X_2; Z_2) \right] + (\epsilon + \epsilon_0) \\ &\le (R_2 + \epsilon) \left[I(X_2; Y_2) - I(X_2; Z_2) \right] \\ &\quad + (\epsilon + \epsilon_0) \end{aligned} \tag{44}$$

$$\begin{aligned} 0 \le r_1 &\le \frac{N_1}{K} I(X_1; Y_1) \\ &\le (R_1 + \epsilon) I(X_1; Y_1). \end{aligned} \tag{45}$$

Similarly, we can prove

$$h_1 - r_2 \le (R_1 + \epsilon) \left[I(X_1; Y_1) - I(X_1; Z_1) \right] + (\epsilon + \epsilon_0) \tag{46}$$

$$0 \le r_2 \le (R_2 + \epsilon) I(X_2; Y_2). \tag{47}$$

Since (35) and (44)–(47) can be obtained for any $\epsilon > 0$, (11)–(13) hold.

B. Lemma on Broadcast Channels

Before going to the details of the proof of the direct part of Theorem 1, we shall establish a lemma concerning broadcast channels. Let us consider a system with a memoryless BCC $P(yz|X)$ in Fig. 3, $x \in \mathcal{X}$, $Y \in \mathcal{Y}$, $z \in \mathcal{Z}$. The encoder and decoder are defined by

$$g: I_{M_1} \times I_{M_2} \to \mathcal{X}^n \tag{48}$$

$$\psi: \mathcal{Y}^n \to I_{M_1} \times I_{M_2} \tag{49}$$

where $I_{M_j} \triangleq \{0, 1, 2, \cdots, M_j - 1\}$. Let W_1, W_2 be messages

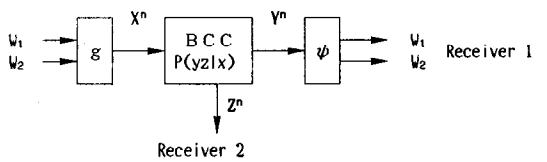

Fig. 3. Broadcast communication system.

from the sender to the receiver 1. W_1 must be kept secret from the receiver 2, who is an eavesdropper, while W_2 may not be kept secret. W_1 and W_2 are independent of each other and taking values over I_{M_1} and I_{M_2}, respectively.

If for any given $\epsilon > 0$ and sufficiently large n there exists a code (g, ψ) such that

$$\frac{1}{n} \log M_j \ge R_j^* - \epsilon, \qquad j = 1, 2 \tag{50}$$

$$\frac{1}{n} H(W_1|\boldsymbol{Z}^n) \ge R_1^* - \epsilon \tag{51}$$

$$P_{Y|Z}^n(\psi(\boldsymbol{y}^n) \ne (W_1, W_2)|W_1, W_2) \le \epsilon, \tag{52}$$

(R_1^*, R_2^*) is said to be admissible for the system shown in Fig. 3. Then the following lemma holds.

Lemma 1: Let $P(x)$ be an arbitrary probability distribution over $\mathcal{X}$. If

$$R_1^* \le I(X; Y) - I(X; Z) \tag{53}$$

$$R_2^* \le I(X; Z), \tag{54}$$

then (R_1^*, R_2^*) is admissible for the system shown in Fig. 3.

Lemma 1 is proved in the Appendix. This lemma means that W_1 and W_2 can be transmitted to the receiver 1 at rate $I(X; Y) - I(X; Z)$ and $I(X; Z)$, respectively, with an arbitrarily small error. Furthermore, W_1 can be kept entirely secret from the receiver 2.

C. Proof of the Direct Part of Theorem 1

Let h_1, h_2, R_1, and R_2 satisfy (11)–(13) for some RV's X_j, Y_j, Z_j and constants r_j $(j = 1, 2)$. Then we must show that a code (f, ϕ) exists that satisfies (6)–(8). Let r_j' be the constant that satisfies the right inequality in (11) with equality. Since $r_j' \ge r_j$, the following inequalities hold:

$$\frac{h_j - r_i'}{R_j} \le I(X_j; Y_j) - I(X_j; Z_j) \tag{55}$$

$$r_1' + r_2' \ge H(S). \tag{56}$$

Furthermore, let h_j' be the constant that satisfies (55) with equality. Then we have

$$h_j' \ge h_j. \tag{57}$$

Letting

$$R_j = \frac{N_j}{K}, \tag{58}$$

r_j' and h_j' satisfy

$$0 \le \frac{K\left[h_j' - r_i'\right]}{N_j} = I(X_j; Y_j) - I(X_j; Z_j) \tag{59}$$

$$\frac{Kr_j'}{N_j} = I(X_j; Y_j). \tag{60}$$

We now construct a code by applying a typical sequence technique for the RV's X_j, Y_j, Z_j satisfying (59) and (60). Let $\mathcal{T}_S$ be the set of typical sequences of S. Then it is

well-known that

$$\Pr\left\{\boldsymbol{S}^K \in \mathcal{T}_S\right\} \geq 1-\epsilon_K,\ \epsilon_K \to 0\ (K\to\infty) \tag{61}$$

$$2^{-K[H(S)+\epsilon_K]} \leq \Pr\left\{\boldsymbol{S}^K = \boldsymbol{s}^K | \boldsymbol{S}^K \in \mathcal{T}_s\right\} \leq 2^{-K[H(S)-\epsilon_K]} \tag{62}$$

$$2^{K[H(S)-\epsilon_K]} \leq |\mathcal{T}_S| \leq 2^{K[H(S)+\epsilon_K]}. \tag{63}$$

Therefore, if $K[H(S)+\epsilon_K]$ bits can be transmitted to the legitimate receiver with an arbitrarily small error for sufficiently large K, (8) is satisfied. Let $K[H(S)+\epsilon_K]$ bits divide into $a_1 \sim a_5$ as shown in Fig. 4. If $h_j' > H(S)+\epsilon_K$, then $h_j' - [H(S)+\epsilon_K]$ dummy bits are added. (See Fig. 5.) Each a_t has the following bits:

a_1: $K[H(S)+\epsilon_K - h_1']$ bits

a_2: $K[h_1' - r_2']$ bits

a_3: $K[r_1' + r_2' - H(S) - \epsilon_K]$ bits

a_4: $K[h_2' - r_1']$ bits

a_5: $K[H(S)+\epsilon_K - h_2']$ bits.

(For simplicity, we consider these bits as integers because they can be approximated by integers with any desired accuracy for sufficiently large K.)

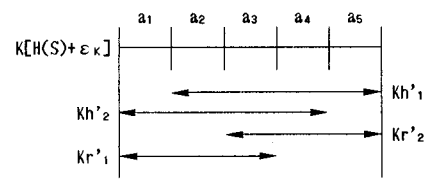

Fig. 4. Partition of $K[H(S)+\epsilon_K]$ bits.

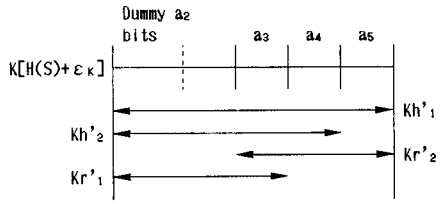

Fig. 5. Modified partition in case of $h_1' > K[H(S)+\epsilon_K]$.

Let T be a uniform random number having $K[r_1' + r_2' - H(S) - \epsilon_K]$ bits and independent of S. These a_t and T are divided into $(a_1, a_2, a_3 \oplus T)$ and (T, a_4, a_5) to be transmitted via BCC1 and BCC2, respectively, where $\oplus$ represents bitwise modulo-two sum.

Since $(a_1, a_2, a_3 \oplus T)$ is coded with codeword length N_1, the rates of a_2 and $(a_1, a_3 \oplus T)$ are given by

$$\frac{K}{N_1}(h_1' - r_2') = I(X_1; Y_1) - I(X_1; Z_1) \tag{64}$$

$$\frac{K}{N_1}(r_1' - (h_1' - r_2')) = I(X_1; Z_1), \tag{65}$$

respectively (because of (59) and (60)). Therefore, from Lemma 1, a_2 and $(a_1, a_3 \oplus T)$ can be transmitted to the legitimate receiver with an arbitrarily small error, and a_2 can be kept entirely secret from the eavesdropper 1.

Similarly, since (T, a_4, a_5) is coded with codeword length N_2, the rates of a_4 and (T, a_5) are given by

$$\frac{K}{N_2}(h_2' - r_1') = I(X_2; Y_2) - I(X_2; Z_2) \tag{66}$$

$$\frac{K}{N_2}(r_2' - (h_2' - r_1')) = I(X_2; Z_2), \tag{67}$$

respectively. Hence a_4 and (T, a_5) can be transmitted to the legitimate receiver with an arbitrarily small error, and a_4 can be kept entirely secret from the eavesdropper 2.

Since the legitimate receiver can obtain $a_1, a_2, a_3 \oplus T, T, a_4, a_5$, it can then reproduce $\hat{\boldsymbol{S}} \in \mathcal{T}_S$ with an arbitrarily small error. This means that (8) holds.

Eavesdropper 1 may obtain a_1 and $a_3 \oplus T$. However, since T is an independent uniform random number, the eavesdropper can know only a_1 having $K[H(S)+\epsilon_K - h_1']$ bits. Therefore, from the equiprobability of the typical sequences (see (62)), the following inequality holds:

$$\begin{aligned}\frac{1}{K}H\left(\boldsymbol{S}|\boldsymbol{Z}_1^{N_1}\right) &\geq \frac{1}{K}\left\{KH(S) - K\left(H(S)+\epsilon_K - h_1'\right)\right\} - \epsilon_K' \\ &= h_1' - \epsilon_K'' \\ &\geq h_1 - \epsilon_K'' \end{aligned} \tag{68}$$

where $\epsilon_K', \epsilon_K'' \to 0$ $(K\to\infty)$. Similarly, for eavesdropper 2 we have

$$\begin{aligned}\frac{1}{K}H\left(\boldsymbol{S}|\boldsymbol{Z}_1^{N_1}\right) &\geq h_2' - \epsilon_K'' \\ &\geq h_2 - \epsilon_K''. \end{aligned} \tag{69}$$

From (58), (68), and (69) we can obtain (6) and (7). Hence, $(R_1, R_2, h_1, h_2) \in \mathcal{R}_{\text{sscs}}$.

IV. An Example of the SSCS with Binary Broadcast Channels

Let us consider a binary example. Let each BCC be constructed by three binary symmetric channels (BSC's) as shown in Fig. 6, where the bit error probability P_{jt} $(t = 1,2,3$ and $j = 1,2)$ is restricted to

$$0 \leq P_{jt} \leq 0.5 \qquad P_{j2} \leq P_{j3} \tag{70}$$

because the main channel is assumed to be less noisy than the wiretapped channel.

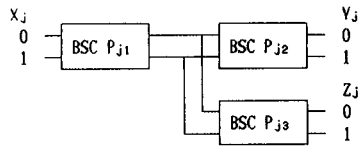

Fig. 6. Broadcast channel consisting of three BSC's.

Let $\Pr\{X=0\}=q_j$ $(j=1,2)$; then we have

$$I(X_j;Y_j)=h(q_j*P_{j1}*P_{j2})-h(P_{j1}*P_{j2}) \tag{71}$$

$$\begin{aligned} I(X_j;Y_j)-I(X_j;Z_j) &= \left[h(q_j*P_{j1}*P_{j2})-h(q_j*P_{j1}*P_{j3})\right] \\ &\quad -\left[h(P_{j1}*P_{j2})-h(P_{j1}*P_{j3})\right] \end{aligned} \tag{72}$$

where $a*b=a(1-b)+(1-a)b$ and $h(x)=-x\log x-(1-x)\log(1-x)$. To maximize the region given by (14)–(16), we have to obtain the optimum q_j such that both (71) and (72) are maximized. Equation (71) is clearly maximized at $q_j=0.5$. Equation (72) is also maximized at $q_j=0.5$ because

$$h(q_j*P_{j1}*P_{j2})\le h(q_j*P_{j1}*P_{j3}) \tag{73}$$

holds for any $q_j, P_{j1}, P_{j2}, P_{j3}$ from (70). Hence $\mathscr{R}_{\text{sscs}}$ for this example is given by

$$\begin{aligned} \mathscr{R}_{\text{sscs}}=\{(R_1,R_2,h_1,h_2):& \\ h_1 &\ge b_1R_1+a_2R_2 \\ h_2 &\ge b_2R_2+a_1R_1 \\ H(S) &\ge a_1R_1+a_2R_2\} \end{aligned} \tag{74}$$

where

$$a_j=1-h(P_{j1}*P_{j2}) \tag{75}$$

$$b_j=h(P_{j1}*P_{j3})-h(P_{j1}*P_{j2}). \tag{76}$$

From Corollary 2 we can also obtain the secrecy capacity region $\mathscr{R}^s_{\text{sscs}}$ for this example as follows:

$$\begin{aligned} \mathscr{R}^s_{\text{sscs}}=\{(R_1,R_2):& \\ H(S) &\ge b_1R_1+a_2R_2 \\ H(S) &\ge b_2R_2+a_1R_1\}. \end{aligned} \tag{77}$$

Note that the degraded BCC of Fig. 7 has equivalent characteristics to the BCC of Fig. 6 with

$$P_{j4}=P_{j1}*P_{j2} \tag{78}$$

$$P_{j5}=\frac{P_{j1}*P_{j3}-P_{j1}*P_{j2}}{1-2(P_{j1}*P_{j2})}. \tag{79}$$

Therefore, we can obtain $\mathscr{R}_{\text{sscs}}$ and $\mathscr{R}^s_{\text{sscs}}$ of the degraded BCC of Fig. 7 from (74) and (77), respectively, by letting

$$a_j=1-h(P_{j4}) \tag{80}$$

$$b_j=h(P_{j4}*P_{j5})-h(P_{j4}). \tag{81}$$

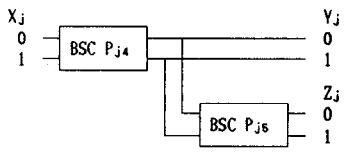

Fig. 7. Broadcast channel equivalent to Fig. 6.

V. Conclusion

The coding theorem of the SSCS with two noisy channels is proved. Furthermore, it is shown that the results of [1], [4], and [5] can be derived from the main theorem of this paper.

The BCC's in this paper are restricted to be "less noisy" for simplicity. We could consider the similar coding problems for the SSCS with two "not less noisy" BCC's or three or more noisy channels; however, these are still open problems.

Appendix
Proof of Lemma 1

Lemma 1 can be proved by using the following lemma.

Lemma A: If $I(X;Y)\ge I(X;Z)$, then for every n there exists a set $\{x^n_{jk}\}\subset\mathscr{T}_X$ where $j\in I_{M_J}\triangleq\{0,1,\cdots,M_J-1\}$ and $k\in I_{M_K}\triangleq\{0,1,\cdots,M_K-1\}$ with the following property. There exist pairwise disjoint subsets $\mathscr{B}_{jk}\subset\mathscr{T}_{Y|X}(x^n_{jk})$ and subsets $\mathscr{C}_{jk}\subset\mathscr{T}_{Z|X}(x^n_{jk})$, of which those with the same index k are pairwise disjoint such that

$$P^n_{Y|X}(\mathscr{B}_{jk}|x^n_{jk})\ge 1-\epsilon_n \tag{A1}$$

$$P^n_{Z|X}(\mathscr{C}_{jk}|x^n_{jk})\ge 1-\epsilon_n \tag{A2}$$

and as $n\to\infty$,

$$\epsilon_n\to 0 \tag{A3}$$

$$(1/n)\log M_J\to I(X;Z) \tag{A4}$$

$$(1/n)\log M_K\to I(X;Y)-I(X;Z). \tag{A5}$$

In the foregoing lemma, $\mathscr{T}_X$ is the set of X-typical sequences such that

$$|N(a|x^n)-nP_X(a)|\le r_n,\qquad \text{for all } a\in\mathscr{X} \tag{A6}$$

where $N(a|x^n)$ is the number of occurrences of a in x^n, and $\{r_n\}$ is a fixed sequence of positive numbers such that $r_nn^{-0.5}\to\infty$, $r_nn^{-1}\to 0$. $\mathscr{T}_{Z|X}$ is the set of all $z^n\in\mathscr{Z}^n$, $Z|X$-generated by X-typical sequences $x^n\in\mathscr{X}^n$ such that

$$|N(ac|x^nz^n)-N(a|x^n)P_{Z|X}(c|a)|\le r_n \tag{A7}$$

for all $a\in\mathscr{X}$, $c\in\mathscr{Z}$. $\mathscr{T}_{Y|X}$ is also defined similarly. If $z^n\in\mathscr{T}_{Z|X}$, we have from (A6) and (A7)

$$|N(c|z^n)-nP_Z(c)|\le 2r_n|\mathscr{X}|\triangleq r'_n \tag{A8}$$

for all $c\in\mathscr{Z}$. Hence $z^n\in\mathscr{T}_{Z|X}$ means $z^n\in\mathscr{T}_Z$. Furthermore it is well-known that

$$P^n_Z(z^n\in\mathscr{T}_Z)\ge 1-\epsilon_n,\qquad \epsilon_n\to 0\,(n\to\infty) \tag{A9}$$

$$\frac{1}{n}H(z^n|z^n\in\mathscr{T}_Z)\le H(Z)+\epsilon_n. \tag{A10}$$

Proof of Lemma A: Lemma A can be obtained by letting $U=$ constant in [4, lemma 2].

Proof of Lemma 1: We can prove Lemma 1 in a way similar to the proof of [4, lemma 3]. Suppose that (53) and (54) hold. Let $M_j=2^{nR^*_j}$, $j=1,2$; then from (A4) and (A5), we can obtain sets $I_{M_{V_j}}\triangleq\{0,1,\cdots,M_{V_j}-1\}$ for (I_{M_1},I_{M_2}) defined in Section III-B and (I_{M_J},I_{M_K}) defined in Lemma A such that M_J, M_{V_j}, M_K, M_J satisfy the following conditions with arbitrary accuracy for sufficiently large n:

$$M_1\times M_{V_1}=M_K \tag{A11}$$

$$M_2\times M_{V_2}=M_J. \tag{A12}$$

Therefore, by using the code $\{x_{jk}^n\}$ and the decoding region $\mathscr{B}_{jk}$ in Lemma A, we can transmit $w_1 \in I_{M_1}$ and $w_2 \in I_{M_2}$ with sufficiently small error. This means that (52) holds.

To complete the proof, it remains to show (51). Let $V_1 \in I_{M_{V_1}}$ and $V_2 \in I_{M_{V_2}}$ be uniform random numbers independent of W_1 and W_2. Let $\boldsymbol{X}^n$ be the channel input RV taking values in the codeword $\{x_{jk}^n\}$. Then

$$\begin{aligned} H(W_1|\boldsymbol{Z}) &= H(W_1\boldsymbol{Z}) - H(\boldsymbol{Z}) \\ &\geq H(W_1\boldsymbol{Z}|V_1) - H(\boldsymbol{Z}) \\ &= H(W_1\boldsymbol{X}\boldsymbol{Z}|V_1) - H(\boldsymbol{X}|W_1\boldsymbol{Z}V_1) - H(\boldsymbol{Z}) \\ &= H(W_1\boldsymbol{X}|V_1) + H(\boldsymbol{Z}|W_1\boldsymbol{X}V_1) \\ &\quad - H(\boldsymbol{X}|W_1\boldsymbol{Z}V_1) - H(\boldsymbol{Z}) \\ &= H(\boldsymbol{X}|V_1) + H(\boldsymbol{Z}|\boldsymbol{X}) - H(\boldsymbol{X}|W_1\boldsymbol{Z}V_1) - H(\boldsymbol{Z}) \end{aligned} \tag{A13}$$

where the last equality follows from the fact that W_1 is uniquely determined by $\boldsymbol{X}$ and $(W_1, V_1) \to \boldsymbol{X} \to \boldsymbol{Z}$ forms a Markov chain.

Since W_1, W_2, V_1, V_2 are mutually independent, we have

$$\begin{aligned} \frac{1}{n}H(\boldsymbol{X}|V_1) &= \frac{1}{n}\log M_1 + \frac{1}{n}\log M_J \\ &\geq R_1^* + I(X;Z) - \epsilon_n. \end{aligned} \tag{A14}$$

Furthermore, since the BCC is memoryless, $H(\boldsymbol{Z}|\boldsymbol{X})$ can be represented by

$$\begin{aligned} H(\boldsymbol{Z}|\boldsymbol{X}) = -\sum_{x \in \{x_{jk}^n\}} \Pr\{\boldsymbol{X} = \boldsymbol{x}\} \sum_{a \in \mathscr{X}} N(a|\boldsymbol{x}) \\ \cdot \sum_{c \in \mathscr{Z}} P_{Z|X}(c|a) \log P_{Z|X}(c|a). \end{aligned} \tag{A15}$$

Since $\boldsymbol{x}$ is a typical sequence, (A6) holds. Hence we have

$$(1/n)H(\boldsymbol{Z}|\boldsymbol{X}) \geq H(Z|X) - \epsilon_n. \tag{A16}$$

If $W_1 \in I_{M_1}$ and $V_1 \in I_{M_{V_1}}$ are given, $k \in I_{M_K}$ is also given. Then we can decode $\boldsymbol{X}$ from $\boldsymbol{Z}$ by using the decoding region $\{\mathscr{C}_{jk}\}$. Therefore, we obtain from Fano's lemma,

$$(1/n)H(\boldsymbol{X}|W_1\boldsymbol{Z}V_1) \leq \epsilon_n. \tag{A17}$$

To obtain a bound of $H(\boldsymbol{Z})$, define RV U by

$$U \triangleq \begin{cases} 1, & \text{if } \boldsymbol{z} \in \mathscr{T}_Z \\ 0, & \text{if } \boldsymbol{z} \notin \mathscr{T}_Z \end{cases}. \tag{A18}$$

Then

$$\begin{aligned} H(\boldsymbol{Z}) = H(\boldsymbol{Z}U) &\leq H(U) + H(\boldsymbol{Z}|U=1) \\ &\quad + H(\boldsymbol{Z}|U=0)\Pr\{U=0\}. \end{aligned} \tag{A19}$$

From (A9) and (A10), we have

$$\frac{1}{n}H(\boldsymbol{Z}) \leq h(\epsilon_n) + H(Z) + \epsilon_n + |\mathscr{Z}|\epsilon_n. \tag{A20}$$

Substituting (A14), (A16), (A17), (A20) into (A13), we obtain

$$\begin{aligned} \frac{1}{n}H(W_1|\boldsymbol{Z}) &\geq R_1^* + I(X;Z) + H(Z|X) - H(Z) - \epsilon_n' \\ &= R_1^* - \epsilon_n', \qquad \epsilon_n' \to 0 \text{ as } n \to \infty. \end{aligned} \tag{A21}$$

REFERENCES

[1] H. Yamamoto, "On secret sharing communication systems with two or three channels," *IEEE Trans. Inform. Theory*, vol. IT-32, no. 3, pp. 387–393, May 1986.

[2] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 565–715, Oct. 1949.

[3] E. D. Karnin, J. W. Greene, and M. E. Hellman, "On secret sharing systems," *IEEE Trans. Inform. Theory*, vol. IT-29, no. 1, pp. 35–41, Jan. 1983.

[4] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[5] A. D. Wyner, "The wire-tap channels," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[6] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[7] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.