Communicating Via a Processing Broadcast Satellite

Aaron D. Wyner [*], Jack K. Wolf [†], and Frans M.J. Willems [‡]

December 20, 2001

Abstract

Three dependent users are physically separated but communicate with each other via a satellite. Each user generates data which it stores locally. In addition, each user sends a message to the satellite. The satellite processes the messages received from the users and broadcasts one common message to all three users. Each user must be capable of reconstructing the data of the other two users based upon the broadcast message and its own stored data. Our problem is to determine the minimum amount of information which must be transmitted to and from the satellite. The solution to this problem is obtained for the case where subsequent data triples that are produced by the users are independent and identically distributed. The three symbols within each triple are assumed to be dependent. Crucial for the solution is an achievability proof that involves cascaded Slepian-Wolf [1973] source coding.

Keywords: correlated sources, data-compression, multi-user source coding, multi-user information theory, satellite communication, Slepian-Wolf coding.

1 Introduction

In this paper we consider a communication network in which each member of a set of users with dependent data must communicate its data to all other users. The users are physically separated and communicate with each other via a communications satellite which sends a broadcast message back to the users. Our problem is to characterize the set of achievable (in the usual Shannon Theory sense) uplink and downlink rates.

1.1 Definitions

We will state our results for the case of three users. The results easily generalize, however, to an arbitrary (finite) number of users. We start with a description of the model (see figure 1).

---

[*] A.T.& T. Bell Laboratories, Murray Hill, NJ 07974.

[†] Center for Magnetic Recording Research, University of California at San Diego, La Jolla, CA 92093.

[‡] Eindhoven University of Technology, Electrical Engineering Department, 5600 MB Eindhoven, The Netherlands.

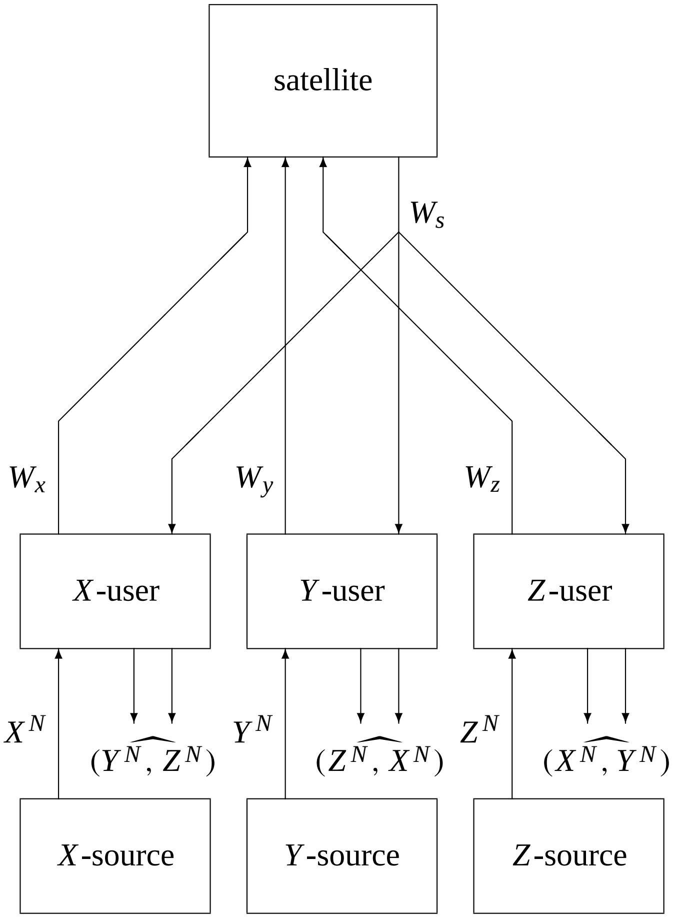

Figure 1: Three users (groundstations) communicating via a broadcasting satellite.

Let $\{X_n, Y_n, Z_n\}$, $n = 1, 2, \cdots$, be a sequence of independent drawings of the random triple (X, Y, Z) which takes values in the finite set $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Let

$$Q(x, y, z) = \Pr\{X = x, Y = y, Z = z\}, x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}, \tag{1}$$

thus $H(X, Y, Z) = \sum_{x,y,z} -Q(x, y, z) \log Q(x, y, z)$, etcetera.

There are three ”users” which have available as data sequences $\{X_n\}$, $\{Y_n\}$, and $\{Z_n\}$, $n = 1, 2, \cdots$, respectively. We will informally refer to the users as the ”X-user”, ”Y-user”, and ”Z-user”. Each user must communicate its data to the other users by sending a message to the satellite, which will in turn send a single ”broadcast” message to the three users. Based on its own data and the broadcast message, each user must be able to reconstruct the data corresponding to other users with high reliability.

For a sequence $\{u_n\}$, $n = 1, 2, \cdots$, and $N = 1, 2, \cdots$ let u^N denote the finite vector $(u_1, u_2, \cdots, u_N)$. Also for positive integers M, let $\mathcal{I}_M$ denote the set $\{1, 2, \cdots, M\}$. We now formally specify our problem.

A ”code” with parameters $(N, M_x, M_y, M_z, M_s, P_e)$ is defined by a set of seven mappings $\{e_x, e_y, e_z, s, d_x, d_y, d_z\}$, where the ”uplink encoder” mappings

$$\begin{aligned} e_x &: \mathcal{X}^N \to \mathcal{I}_{M_x}, \\ e_y &: \mathcal{Y}^N \to \mathcal{I}_{M_y}, \\ e_z &: \mathcal{Z}^N \to \mathcal{I}_{M_z}, \end{aligned} \tag{2}$$

the ”satellite” mapping

$$s : \mathcal{I}_{M_x} \times \mathcal{I}_{M_y} \times \mathcal{I}_{M_z} \to \mathcal{I}_{M_s}, \tag{3}$$

and the ”decoder” mappings

$$\begin{aligned} d_x &: \mathcal{X}^N \times \mathcal{I}_{M_s} \to \mathcal{Y}^N \times \mathcal{Z}^N, \\ d_y &: \mathcal{Y}^N \times \mathcal{I}_{M_s} \to \mathcal{Z}^N \times \mathcal{X}^N, \\ d_z &: \mathcal{Z}^N \times \mathcal{I}_{M_s} \to \mathcal{X}^N \times \mathcal{Y}^N. \end{aligned} \tag{4}$$

Let X^N, Y^N, and Z^N denote the random variables corresponding to the sequences that are produced by the sources. Assume that the uplink messages are

$$\begin{aligned} W_x &= e_x(X^N), \\ W_y &= e_y(Y^N), \\ W_z &= e_z(Z^N). \end{aligned} \tag{5}$$

Let the downlink broadcast message be

$$W_s = s(W_x, W_y, W_z). \tag{6}$$

Finally let

$$(\widehat{Y^N, Z^N}) = d_x(X^N, W_s) \tag{7}$$

be the X-user estimate of (Y^N, Z^N). Similarly let the Y-user estimate and Z-user estimate be

$$\begin{aligned}(\widehat{Z^N, X^N}) &= d_y(Y^N, W_s),\\ (\widehat{X^N, Y^N}) &= d_z(Z^N, W_s).\end{aligned} \tag{8}$$

The "error probability" is defined as

$$P_e \overset{\Delta}{=} \Pr\left\{(\widehat{Y^N, Z^N}) \neq (Y^N, Z^N) \text{ or } (\widehat{Z^N, X^N}) \neq (Z^N, X^N) \text{ or } (\widehat{X^N, Y^N}) \neq (X^N, Y^N)\right\}. \tag{9}$$

We say that a "rate" quadruple (R_x, R_y, R_z, R_s) is "achievable" if for all $\delta > 0$, and all N sufficiently large, there exists a code with parameters $(N, M_x, M_y, M_z, M_s, P_e)$ such that

$$\begin{aligned}\frac{1}{N}\log M_x &\leq R_x + \delta,\\ \frac{1}{N}\log M_y &\leq R_y + \delta,\\ \frac{1}{N}\log M_z &\leq R_z + \delta,\\ \frac{1}{N}\log M_s &\leq R_s + \delta,\end{aligned} \tag{10}$$

and

$$P_e \leq \delta. \tag{11}$$

Our problem is to find the set of achievable rate quadruples.

1.2 Statement of result

In this paper we will prove the following theorem:

Theorem 1 *The rate quadruple (R_x, R_y, R_z, R_s) is achievable if and only if*

$$\begin{aligned}R_x &\geq H(X|Y, Z),\\ R_y &\geq H(Y|Z, X),\\ R_z &\geq H(Z|X, Y),\\ R_y + R_z &\geq H(Y, Z|X),\\ R_z + R_x &\geq H(Z, X|Y),\\ R_x + R_y &\geq H(X, Y|Z),\\ R_s &\geq \max[H(Y, Z|X), H(Z, X|Y), H(X, Y|Z)].\end{aligned} \tag{12}$$

Before we will give the proof of the theorem we will make some remarks.

1.3 Remarks

1. The conditions given by the first six inequalities of (12) are not sufficient to communicate (X^N, Y^N, Z^N) to the satellite. From the Slepian-Wolf method [8], we know that in addition to these inequalities we would also need
$$R_x + R_y + R_z \geq H(X, Y, Z) \tag{13}$$
to accomplish this. But we do not require that the satellite knows (X^N, Y^N, Z^N) so there is no inconsistency.

2. Suppose that the satellite did in fact know (X^N, Y^N, Z^N). Then to communicate (Y^N, Z^N) to the X-user from the satellite a rate of $H(Y, Z|X)$ is necessary and sufficient. Similarly to communicate (Z^N, X^N) to the Y-user a rate not smaller than $H(Z, X|Y)$ is required. To communicate (X^N, Y^N) to the Z-user rate $H(X, Y|Z)$ is necessary. Theorem 1 asserts that all three tasks can be accomplished with a *single broadcast* message using a rate of $\max[H(X, Y|Z), H(Y, Z|X), H(Z, X|Y)]$, and in fact the satellite need not know (X^N, Y^N, Z^N) fully.

2 Proof of theorem 1

The proof consists of a converse result and an achievability proof. The converse is rather straightforward. To complete the achievability proof we cascaded two levels of Slepian-Wolf codes.

2.1 Converse

As usual this weak converse starts with a chain of (in)equalities that are centered around Fano's inequality (a):

$$\begin{aligned} H(Y^N, Z^N|X^N, W_s) &= H(Y^N, Z^N|X^N, W_s, \widehat{(Y^N, Z^N)}) \\ &\leq H(Y^N, Z^N|\widehat{(Y^N, Z^N)}) \\ &\overset{(a)}{\leq} h(P_{ex}) + P_{ex} \log(\|\mathcal{Y}\|^N \|\mathcal{Z}\|^N - 1) \\ &\leq \log 2 + N P_e \log \|\mathcal{X}\| \|\mathcal{Y}\| \|\mathcal{Z}\| = N\Delta(P_e, N), \end{aligned} \tag{14}$$

where
$$P_{ex} \triangleq \Pr\left\{\widehat{(Y^N, Z^N)} \neq (Y^N, Z^N)\right\}, \tag{15}$$

and
$$\Delta(p, N) \triangleq \frac{\log 2}{N} + p \cdot \log \|\mathcal{X}\| \|\mathcal{Y}\| \|\mathcal{Z}\|. \tag{16}$$

Note that $h(\cdot)$ is the binary entropy function. Similarly we obtain

$$\begin{aligned} H(Z^N, X^N|Y^N, W_s) &\leq N\Delta(P_e, N), \\ H(X^N, Y^N|Z^N, W_s) &\leq N\Delta(P_e, N). \end{aligned} \tag{17}$$

Now we continue with

$$\begin{aligned}\log M_x &\geq H(W_x)\\ &\geq I(W_x; X^N|Y^N, Z^N)\\ &= H(X^N|Y^N, Z^N) - H(X^N|Y^N, Z^N, W_x)\\ &\overset{(b)}{=} NH(X|Y, Z) - H(X^N|Y^N, W_y, Z^N, W_z, W_x, W_s)\\ &\geq NH(X|Y, Z) - H(X^N, Y^N|Z^N, W_s)\\ &\overset{(c)}{\geq} N(H(X|Y, Z) - \Delta(P_e, N)), \end{aligned} \tag{18}$$

where (b) follows from (5) and (6). Finally (c) follows from (17). In the same way we obtain

$$\begin{aligned}\log M_y &\geq N(H(Y|Z, X) - \Delta(P_e, N)),\\ \log M_z &\geq N(H(Z|X, Y) - \Delta(P_e, N)).\end{aligned} \tag{19}$$

Next we proceed with

$$\begin{aligned}\log M_y M_z &\geq H(W_y, W_z)\\ &\geq I(W_y, W_z; Y^N, Z^N|X^N)\\ &= H(Y^N, Z^N|X^N) - H(Y^N, Z^N|X^N, W_y, W_z)\\ &= NH(Y, Z|X) - H(Y^N, Z^N|X^N, W_x, W_y, W_z, W_s)\\ &\geq NH(Y, Z|X) - H(Y^N, Z^N|X^N, W_s)\\ &\geq N(H(Y, Z|X) - \Delta(P_e, N)).\end{aligned} \tag{20}$$

Analogously we find that

$$\begin{aligned}\log M_z M_x &\geq N(H(Z, X|Y) - \Delta(P_e, N)),\\ \log M_x M_y &\geq N(H(X, Y|Z) - \Delta(P_e, N)).\end{aligned} \tag{21}$$

Next

$$\begin{aligned}\log M_s &\geq H(W_s)\\ &\geq I(W_s; Y^N, Z^N|X^N)\\ &= H(Y^N, Z^N|X^N) - H(Y^N, Z^N|X^N, W_s)\\ &= N(H(Y, Z|X) - \Delta(P_e, N)),\end{aligned} \tag{22}$$

and in the same way

$$\begin{aligned}\log M_s &\geq N(H(Z, X|Y) - \Delta(P_e, N)),\\ \log M_s &\geq N(H(X, Y|Z) - \Delta(P_e, N)).\end{aligned} \tag{23}$$

Finally from (18) we obtain that an achievable quadruple (R_x, R_y, R_z, R_s) satisfies

$$R_x + \delta \geq \frac{\log M_x}{N} \geq H(X|Y, Z) - \Delta(P_e, N) \geq H(X|Y, Z) - \Delta(\delta, N) \tag{24}$$

for all $\delta > 0$ and all N large enough. Note that $\Delta(\delta, N) \to 0$ for $\delta \to 0$ and $N \to \infty$. This implies that

$$R_x \geq H(X|Y, Z), \tag{25}$$

for all achievable rate quadruples (R_x, R_y, R_z, R_s). In the same way we obtain the other inequalities in the statement of the theorem. This concludes the proof of converse.

2.2 Achievability proof

The achievability proof is based on Slepian-Wolf source coding techniques [8] (see also Cover [3]). It consists of two parts.

The first part deals with the ”uplink codes”. Note that the issue of constructing uplink codes is also considered by Cover, El Gamal, and Salehi [4]. These authors modeled the transmission links to the satellite as a multiple acces channel. Here we assume that these links are noiseless.

In the second part the construction of the ”dowlink code” is discussed. The important feature of this code is that the downlink signal carries three information streams simultaneously, while its rate is only determined by the rate of the ”largest” stream. A similar phenomenon was discovered by Sgarro [7] when he investigated source coding with side information at *several* decoders. It would be interesting to know whether and how Sgarro's result covers ours. The major problem in finding this out is that Sgarro's approach can only work if the satellite knows X^N, Y^N, and Z^N however.

2.2.1 The uplink codes

Fix $\epsilon > 0$. Let $\mathcal{A}_\epsilon^{(N)}(X, Y, Z)$ be the set of jointly ϵ-typical (x^N, y^N, z^N)-sequences (for definitions see e.g. Cover and Thomas [5], chapter 14).

Consider a random partitioning $\{A_1, A_2, \cdots, A_{M_x}\}$ with rate $R_x = (\log M_x)/N$ of the set $\mathcal{X}^N$. More precisely let

$$\Pr\{x^N \in A_m\} = \frac{1}{M_x} \text{ for } m = 1, 2, \cdots, M_x \text{ and } x^N \in \mathcal{X}^N. \tag{26}$$

In the same way create a random partitioning $\{B_1, B_2, \cdots, B_{M_y}\}$ with rate $R_y = (\log M_y)/N$ of the set $\mathcal{Y}^N$ and a random partitioning $\{C_1, C_2, \cdots, C_{M_z}\}$ with rate $R_z = (\log M_z)/N$ of the set $\mathcal{Z}^N$.

User X now observes the source sequence x^N, determines w_x such that $x^N \in A_{w_x}$, and sends this ”bin index” w_x to the satellite. Similarly user Y and user Z send the bin indices w_y and w_z respectively to the satellite.

If the X-user knew w_y and w_z it could try to determine y^N and z^N by choosing the pair $(\widehat{y^N, z^N})$ such that

$$(x^N, (\widehat{y^N, z^N})) \in \mathcal{A}_\epsilon^{(N)}(X, Y, Z) \text{ and } (\widehat{y^N, z^N}) \in A_{w_y} \times A_{w_z}. \tag{27}$$

If there is no unique pair $(\widehat{y^N, z^N})$ or no pair at all that satisfies (27), user X declares an error. We call this an *uplink error*. For the uplink error probability averaged[1] over the ensemble of random

---

[1] The ”overlines” in definition (28) denotes averaging over this ensemble.

partitionings $\{A_1, A_2, \cdots, A_{M_x}\}$, $\{B_1, B_2, \cdots, B_{M_y}\}$, and $\{C_1, C_2, \cdots, C_{M_s}\}$

$$\overline{P_{ex,u}} \overset{\Delta}{=} \overline{\Pr}\left\{(\widehat{Y^N, Z^N}) \neq (Y^N, Z^N)\right\} \tag{28}$$

we can show (see appendix A) that

$$\overline{P_{ex,u}} \leq 2\epsilon, \tag{29}$$

for sufficiently large N, provided that

$$\begin{aligned} R_y &\geq H(Y|Z, X) + 3\epsilon, \\ R_z &\geq H(Z|X, Y) + 3\epsilon, \\ R_y + R_z &\geq H(Y, Z|X) + 3\epsilon. \end{aligned} \tag{30}$$

Similarly for the two other uplink error probabilities (averaged over the ensemble of random partitionings) we obtain for large enough N that

$$\begin{aligned} \overline{P_{ey,u}} \overset{\Delta}{=} \overline{\Pr}\left\{(\widehat{Z^N, X^N}) \neq (Z^N, X^N)\right\} &\leq 2\epsilon, \\ \overline{P_{ez,u}} \overset{\Delta}{=} \overline{\Pr}\left\{(\widehat{X^N, Y^N}) \neq (X^N, Y^N)\right\} &\leq 2\epsilon, \end{aligned} \tag{31}$$

if in addition to (30)

$$\begin{aligned} R_x &\geq H(X|Y, Z) + 3\epsilon, \\ R_z + R_x &\geq H(Z, X|Y) + 3\epsilon, \\ R_x + R_y &\geq H(X, Y|Z) + 3\epsilon. \end{aligned} \tag{32}$$

Hence the total uplink error probability averaged over the ensemble of random partitionings $\{A_1, A_2, \cdots, A_{M_x}\}$, $\{B_1, B_2, \cdots, B_{M_y}\}$, and $\{C_1, C_2, \cdots, C_{M_s}\}$

$$\begin{aligned} \overline{P_{e,u}} &\overset{\Delta}{=} \overline{\Pr}\left\{(\widehat{Y^N, Z^N}) \neq (Y^N, Z^N) \text{ or } (\widehat{Z^N, X^N}) \neq (Z^N, X^N) \text{ or } (\widehat{X^N, Y^N}) \neq (X^N, Y^N)\right\} \\ &\leq 6\epsilon \end{aligned} \tag{33}$$

for N large enough, since it satisfies $\overline{P_{e,u}} \leq \overline{P_{ex,u}} + \overline{P_{ey,u}} + \overline{P_{ex,u}}$, if R_x, R_y, and R_z satisfy (30) and (32).

Finally we conclude that for all N large enough there must exist uplink codes (partitionings $\{A_1, A_2, \cdots, A_{M_x}\}$, $\{B_1, B_2, \cdots, B_{M_y}\}$, and $\{C_1, C_2, \cdots, C_{M_z}\}$) with rates satisfying (30) and (32), that achieve total uplink error probability $P_{e,u} \leq 6\epsilon$. Equivalently there exist uplink codes with rates satisfying (30) and (32), that achieve total probability $P_e \leq 6\epsilon$ if the downlink rate $R_s = R_x + R_y + R_z$. In what follows we will show that also rates R_s equal to the maximum of $H(Y, Z|X)$, $H(Z, X|Y)$, and $H(X, Y|Z)$ will suffice. Note that this is in general smaller than $R_x + R_y + R_z$.

2.2.2 The downlink code

For all sufficiently large N fix certain partitionings $\{A_1, A_2, \cdots, A_{M_x}\}$, $\{B_1, B_2, \cdots, B_{M_y}\}$, and $\{C_1, C_2, \cdots, C_{M_z}\}$, i.e. fix the uplink codes. Note that W_x, W_y, and W_z are the uplink messages that are produced from the data sequences X^N, Y^N, and Z^N, respectively.

Now we can define the sets (one for each value of N)

$$\begin{aligned}\mathcal{W}_{yz}(x^N) \;\; \overset{\Delta}{=} \;\; \Big\{(w_y, w_z) : \exists(y^N, z^N), y^N \in B_{w_y}, z^N \in C_{w_z}, \\ (x^N, y^N, z^N) \in \mathcal{A}_\epsilon^{(N)}(X, Y, Z)\Big\}. \qquad (34)\end{aligned}$$

Also consider similarly defined sets $\mathcal{W}_{zx}(y^N)$ and $\mathcal{W}_{xy}(z^N)$.

Now, again for each N, we create a random partitioning $(D_1, D_2, \cdots, D_{M_s})$ with rate $R_s = (\log M_s)/N$ of the set $\mathcal{I}_{M_x} \times \mathcal{I}_{M_y} \times \mathcal{I}_{M_z}$. More precisely let

$$\Pr\{(w_x, w_y, w_z) \in D_m\} = \frac{1}{M_s} \text{ for } m = 1, 2, \cdots, M_s \text{ and } (w_x, w_y, w_z) \in \mathcal{I}_{M_x} \times \mathcal{I}_{M_y} \times \mathcal{I}_{M_z}. \qquad (35)$$

The satellite, upon receiving the three bin indices w_x, w_y, and w_z, determines the super index w_s such that $(w_x, w_y, w_z) \in D_{w_s}$ and sends (broadcasts) this index w_s to all three users.

Consider the X-user. Since it knows x^N and the index w_x such that $x^N \in A_{w_x}$, it can determine the index pair $(\widehat{w_y, w_z})$ such that

$$(\widehat{w_y, w_z}) \in \mathcal{W}_{yz}(x^N) \text{ and } (w_x, (\widehat{w_y, w_z})) \in D_{w_s}, \qquad (36)$$

from w_s. If a pair satisfying (36) can not be found or if it is not unique a *downlink error* is declared by the X-user. For the corresponding downlink error probability averaged over the ensemble of random partitionings $\{D_1, D_2, \cdots, D_{M_s}\}$

$$\overline{P_{ex,d}} \overset{\Delta}{=} \overline{\Pr}\left\{(\widehat{W_y, W_z}) \neq (W_y, W_z)\right\} \qquad (37)$$

we can show (see appendix B) that

$$\overline{P_{ex,d}} \leq 2\epsilon, \qquad (38)$$

for N sufficient large, provided that

$$R_s \geq H(Y, Z|X) + 3\epsilon. \qquad (39)$$

Similarly for the other downlink error probabilities we obtain

$$\begin{aligned}\overline{P_{ey,d}} \overset{\Delta}{=} \overline{\Pr}\left\{(\widehat{W_z, W_x}) \neq (W_z, W_x)\right\} &\leq 2\epsilon, \\ \overline{P_{ez,d}} \overset{\Delta}{=} \overline{\Pr}\left\{(\widehat{W_x, W_y}) \neq (W_x, W_y)\right\} &\leq 2\epsilon, \qquad (40)\end{aligned}$$

for sufficiently large N, if

$$\begin{aligned}R_s &\geq H(Z, X|Y) + 3\epsilon, \\ R_s &\geq H(X, Y|Z) + 3\epsilon. \qquad (41)\end{aligned}$$

Thus for the total downlink error probability averaged over the ensemble of random partitionings $\{D_1, D_2, \cdots, D_{M_s}\}$ we get

$$\begin{aligned} \overline{P_{e,d}} &\stackrel{\Delta}{=} \overline{\Pr}\left\{(\widehat{W_y, W_z}) \neq (W_y, W_z) \text{ or } (\widehat{W_z, W_x}) \neq (W_z, W_x) \text{ or } (\widehat{W_x, W_y}) \neq (W_x, W_y)\right\} \\ &\leq 6\epsilon \end{aligned} \tag{42}$$

for N large enough, since $\overline{P_{e,d}} \leq \overline{P_{ex,d}} + \overline{P_{ey,d}} + \overline{P_{ex,d}}$, if R_s satisfies (39) and (41).

Therefore we conclude that there exists a downlink code (partitioning) $\{D_1, D_2, \cdots, D_{M_s}\}$ with rate satisfying (39) and (41), that achieves total downlink probability $P_{e,d} \leq 6\epsilon$ provided that

$$R_s \geq \max[H(Y, Z|X), H(Z, X|Y), H(X, Y|Z)] + 3\epsilon. \tag{43}$$

2.2.3 Total error probability

Assume that for the uplink codes we take the uplink rates R_x, R_y, and R_z satisfying (30), and (32). Then there exist uplink codes with total uplink error probability $P_{e,u} \leq 6\epsilon$.

Given these uplink codes, there exist downlink codes with rates satisfying (39) and (41) that result in a total downlink error probability $P_{e,d} \leq 6\epsilon$.

Furthermore note that if there occur no errors in the uplink part of the system and if there occur no errors in the downlink part then all users have correctly reconstructed the data sequences of the other two users. Hence there exists uplink and downlink codes that achieve total error probability

$$P_e \leq P_{e,u} + P_{e,d} \leq 12\epsilon, \tag{44}$$

for all N large enough. Letting $\epsilon \downarrow 0$ proves the achievability of rate quadruples satisfying the inequalities in the statement of the theorem.

3 Discussion

1. It is clear from the achievability proof that the uplink codes are based on standard Slepian-Wolf [8] techniques (random binning as proposed by Cover [3]). Interesting is that the downlink code is a code *cascaded* over the uplink code. The techniques that are used for the downlink code construction can again be regarded as Slepian-Wolf methods. Again random binning is involved and the set $\mathcal{W}_{yz}(x^N)$ is similar to a set of jointly typical sequences. The fact that the achievability proof involves *concatenation* of two Slepian-Wolf codes is maybe the most remarkable property of this paper.

2. So far we have not considered the *feedback* case. Instead of transmitting a single message to the satellite, each user can send a first message to the satellite, wait for a return message from the satellite, and depending on this return message send a second message to the satellite, receive a second return message, depending on this message send a third message, etcetera. We call this a scenario with feedback. More precisely assume that, for $b = 1, B$, the b-th uplink message $W_{x,b}$ is determined by the source sequence X^N and all messages

$W_{s,1}^{b-1} \overset{\Delta}{=} W_{s,1}, W_{s,2}, \cdots, W_{s,b-1}$ from the satellite that were received before, hence we have for the uplink messages

$$\begin{aligned} W_{x,b} &= e_{x,b}(X^N, W_{s,1}^{b-1}), \\ W_{y,b} &= e_{y,b}(Y^N, W_{s,1}^{b-1}), \\ W_{z,b} &= e_{z,b}(Z^N, W_{s,1}^{b-1}), \text{ for } b = 1, B. \end{aligned} \tag{45}$$

Similarly for the downlink messages we get

$$W_{s,b} = s_b(W_{x,1}^b, W_{y,1}^b, W_{z,1}^b), \text{ for } b = 1, B, \tag{46}$$

where $W_{x,1}^b \overset{\Delta}{=} W_{x,1}, W_{x,2}, \cdots, W_{x,b}$, etcetera. Finally let

$$\begin{aligned} (\widehat{Y^N, Z^N}) &= d_x(X^N, W_{s,1}^B), \\ (\widehat{Z^N, X^N}) &= d_y(Y^N, W_{s,1}^B), \\ (\widehat{X^N, Y^N}) &= d_z(Z^N, W_{s,1}^B). \end{aligned} \tag{47}$$

In appendix C it is shown that in the feedback scenario the set of achievable rate quadruples is the same as in the non-feedback case.

3. For two users X and Y, a rate triple (R_x, R_y, R_s) is achievable if and only if

$$\begin{aligned} R_x &\geq H(X|Y), \\ R_y &\geq H(Y|X), \\ R_s &\geq \max[H(Y|X), H(X|Y)]. \end{aligned} \tag{48}$$

As an example, suppose that (X, Y) is the "binary symmetric source", i.e. $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ and

$$Q(x,y) = \begin{cases} 1 - p_0 & \text{if } x = y, \\ p_0 & \text{if } x \neq y. \end{cases} \tag{49}$$

Then from the results of Körner and Marton [6], it is possible to communicate to the satellite $Z^N = X^N \oplus Y^N$ (where $\oplus$ denotes symbol-wise modulo two addition) using $R_x = R_y = h(p_0)$ (where $h(p_0) = -p_0 \log p_0 - (1 - p_0)\log(1 - p_0)$). Since $H(Z^N) = Nh(p_0)$, the satellite can simply broadcast an encoding of Z^N (using $R_s = h(p_0)$), and the X-user can recover Y^N $(= X^N \oplus Z^N)$, and the Y-user can recover X^N. Thus the rate triple $R_x = R_y = R_s = h(p_0)$ is achievable, consistent with theorem 1. But the Körner-Martin result does not generalize. Perhaps our results sheds some light on their problem.

4. The survey-paper of Berger [2] reported about the progress in the area of multi user source coding that was made in a very active period before 1977. More recently interest in this area is increasing again. E.g. in Yeung and Zhang [12] and Ahlswede et al. [1] source networks are considered in which a number of information sources are to be multicast

to certain sets of destinations. In multicasting it is usually assumed that all sources are independent. Very recently however, at the 2001 ISIT, Song and Yeung [9] presented results for multicast networks with dependent sources. This research is very much related to the problem that we have investigated here.

Epilog

Aaron Wyner and Jack Wolf formulated the problem that was addressed in this paper in the middle of the eighties. During a lecture at Philips Research Laboratories in April 1988 Jack Wolf mentioned the problem and presented results for the two-user case. After this lecture the three authors teamed up and were able to solve the general case. The result was presented first at the 1989 IEEE Information Theory Workshop at Cornell by Aaron Wyner [10] and later at the 1990 IEEE International Symposium on Information Theory in San Diego by Frans Willems [11]. Aaron Wyner died before this research could be written up for publication. The publication of a special issue of the Information Theory Transactions commemorating Aaron Wyner turned out to be a good opportunity to submit this result after all.

Acknowledgement

The authors thank the Associate Editor Shlomo Shamai and the anonymous reviewers for their constructive remarks and suggestions.

Appendix A: An upper bound for $\overline{P_{ex,u}}$

Define the events

$$\begin{aligned}
E_0 &= \left\{(X^N, Y^N, Z^N) \notin \mathcal{A}_\epsilon^{(N)}(X,Y,Z)\right\},\\
E_1 &= \left\{\exists \tilde{y}^N \neq Y^N : B(\tilde{y}^N) = B(Y^N) \text{ and } (X^N, \tilde{y}^N, Z^N) \in \mathcal{A}_\epsilon^{(N)}(X,Y,Z)\right\},\\
E_2 &= \left\{\exists \tilde{z}^N \neq Z^N : C(\tilde{z}^N) = C(Z^N) \text{ and } (X^N, Y^N, \tilde{z}^N) \in \mathcal{A}_\epsilon^{(N)}(X,Y,Z)\right\},\\
E_{12} &= \left\{\exists (\tilde{y}^N \neq Y^N, \tilde{z}^N \neq Z^N) : B(\tilde{y}^N) = B(Y^N), C(\tilde{z}^N) = C(Z^N),\right.\\
&\qquad\qquad\qquad \left.\text{and } (X^N, \tilde{y}^N, \tilde{z}^N) \in \mathcal{A}_\epsilon^{(N)}(X,Y,Z)\right\},
\end{aligned} \tag{50}$$

where $B(y^N)$ denotes the bin index of y^N and $C(z^N)$ the bin index of z^N. Note that the source outcomes X^N, Y^N, and Z^N are random variables but so are the ”mappings” $B(\cdot)$ and $C(\cdot)$.

We get an error if (X^N, Y^N, Z^N) is not in $\mathcal{A}_\epsilon^{(N)}(X,Y,Z)$ or if there is another typical triple $(X^N, \tilde{y}^N, \tilde{z}^N)$ with $\tilde{y}^N$ and $\tilde{z}^N$ in the same bins as Y^N and Z^N respectively.

By the union bound

$$\begin{aligned}
\overline{P_{ex,u}} &= \overline{\Pr}\{E_0 \cup E_1 \cup E_2 \cup E_{12}\} \\
&\leq \overline{\Pr}\{E_0\} + \overline{\Pr}\{E_1\} + \overline{\Pr}\{E_2\} + \overline{\Pr}\{E_{12}\}. \qquad (51)
\end{aligned}$$

We first consider E_0. The law of large numbers implies (see Cover and Thomas [5], Theorem 14.2.1) that for sufficiently large N

$$\overline{\Pr}\{E_0\} \leq \epsilon, \qquad (52)$$

for any $\epsilon > 0$. Now we can bound the other three terms. For the second term we get

$$\begin{aligned}
\overline{\Pr}\{E_1\} &= \overline{\Pr}\left\{\exists \tilde{y}^N \neq Y^N : B(\tilde{y}^N) = B(Y^N) \text{ and } (X^N, \tilde{y}^N, Z^N) \in \mathcal{A}_\epsilon^{(N)}(X, Y, Z)\right\} \\
&\leq \sum_{x^N, y^N, z^N} p(x^N, y^N, z^N) \sum_{\substack{\tilde{y}^N \neq y^N \\ (x^N, \tilde{y}^N, z^N) \in \mathcal{A}_\epsilon^{(N)}(X,Y,Z)}} \overline{\Pr}\{B(\tilde{y}^N) = B(y^N)\} \\
&\leq \sum_{x^N, y^N, z^N} p(x^N, y^N, z^N) \exp(-N R_y) \|\mathcal{A}_\epsilon^{(N)}(Y|z^N, x^N)\| \\
&\overset{(\alpha)}{\leq} \exp(-N R_y) \exp(N(H(Y|Z, X) + 2\epsilon) \\
&\leq \exp(-N\epsilon), \qquad (53)
\end{aligned}$$

if $R_y \geq H(Y|Z, X) + 3\epsilon$. Note that (α) follows from $\|\mathcal{A}_\epsilon^{(N)}(Y|z^N, x^N)\| \leq \exp(N(H(Y|Z, X) + 2\epsilon)$, see e.g. Theorem 14.2.2 in [5]. In the same way we can show that also

$$\begin{aligned}
\overline{\Pr}\{E_2\} &\leq \exp(-N\epsilon), \\
\overline{\Pr}\{E_{12}\} &\leq \exp(-N\epsilon), \qquad (54)
\end{aligned}$$

provided that $R_z \geq H(Z|X, Y) + 3\epsilon$ and $R_y + R_z \geq H(Y, Z|X) + 3\epsilon$. We finally conclude that

$$\overline{P_{ex,u}} \leq \epsilon + 3\exp(-N\epsilon) \leq 2\epsilon, \qquad (55)$$

for N sufficiently large, if only

$$\begin{aligned}
R_y &\geq H(Y|Z, X) + 3\epsilon, \\
R_z &\geq H(Z|X, Y) + 3\epsilon, \\
R_y + R_z &\geq H(Y, Z|X) + 3\epsilon. \qquad (56)
\end{aligned}$$

Appendix B: An upper bound for $\overline{P_{ex,d}}$

Define the events

$$\begin{aligned}
E_3 &= \left\{(W_y, W_z) \notin \mathcal{W}_{yz}(X^N)\right\}, \\
E_4 &= \left\{\exists (\tilde{w}_y, \tilde{w}_z) \neq (W_y, W_z) : D(W_x, \tilde{w}_y, \tilde{w}_z) = D(W_x, W_y, W_z)\right. \\
&\qquad \left. \text{and } (\tilde{w}_y, \tilde{w}_z) \in \mathcal{W}_{yz}(X^N)\right\}, \qquad (57)
\end{aligned}$$

where $D(w_x, w_y, w_z)$ denotes the bin index of (w_x, w_y, w_z). Note that the mapping $D(\cdot, \cdot, \cdot)$ is a random variable. Observe that the set $\mathcal{W}_{yz}(x^N)$ plays a similar role as the set of jointly typical sequences $\mathcal{A}_\epsilon^{(N)}(Y, Z|x^N)$. Later this will even become more clear. The method that we follow here is actually a Slepian-Wolf [8] (-Cover [3]) method.

We get an error if (W_y, W_z) is not in $\mathcal{W}_{yz}(X^N)$ or if there is another pair $(\tilde{w}_y, \tilde{w}_z) \in \mathcal{W}_{yz}(X^N)$ which, together with W_x, is in the same bin as (W_x, W_y, W_z).

By the union bound

$$\overline{P_{ex,d}} = \overline{\Pr}\{E_3 \cup E_4\} \leq \overline{\Pr}\{E_3\} + \overline{\Pr}\{E_4\}. \tag{58}$$

We first consider E_3. Observe that if a data sequence triple $(x^N, y^N, z^N) \in \mathcal{A}_\epsilon^{(N)}(X, Y, Z)$ occurs, the corresponding message pair $(w_y, w_z) \in \mathcal{W}_{yz}(x^N)$. Therefore we conclude (see Cover and Thomas [5], Theorem 14.2.1) that

$$\overline{\Pr}\{E_3\} \leq \overline{\Pr}\{(X^N, Y^N, Z^N) \notin \mathcal{A}_\epsilon^{(N)}(X, Y, Z)\} \leq \epsilon, \tag{59}$$

for sufficiently large N, for any $\epsilon > 0$. For the second term we get

$$\begin{aligned}
\overline{\Pr}\{E_4\} &= \overline{\Pr}\left\{\exists(\tilde{w}_y, \tilde{w}_z) \neq (W_y, W_z) : D(W_x, \tilde{w}_y, \tilde{w}_z) = D(W_x, W_y, W_z)\right. \\
&\qquad\qquad\qquad\qquad\qquad \left.\text{and } (\tilde{w}_y, \tilde{w}_z) \in \mathcal{W}_{yz}(X^N)\right\}, \\
&\leq \sum_{x^N, y^N, z^N} p(x^N, y^N, z^N) \sum_{\substack{(\tilde{w}_y, \tilde{w}_z) \neq (W_y, W_z) \\ (\tilde{w}_y, \tilde{w}_z) \in \mathcal{W}_{yz}(X^N)}} \overline{\Pr}\{D(W_x, \tilde{w}_y, \tilde{w}_z) = D(W_x, W_y, W_z)\} \\
&\leq \sum_{x^N, y^N, z^N} p(x^N, y^N, z^N) \exp(-NR_s) \|\mathcal{W}_{yz}(x^N)\| \\
&\overset{(\beta)}{\leq} \exp(-NR_s) \exp(N(H(Y, Z|X) + 2\epsilon) \\
&\leq \exp(-N\epsilon),
\end{aligned} \tag{60}$$

if $R_s \geq H(Y, Z|X) + 3\epsilon$. Note that (β) follows from $\|\mathcal{W}_{yz}(x^N)\| \leq \|\mathcal{A}_\epsilon^{(N)}(Y, Z|x^N)\| \leq \exp(N(H(Y, Z|X) + 2\epsilon)$, see again e.g. Theorem 14.2.2 in [5].

We finally conclude that

$$\overline{P_{ex,d}} \leq \epsilon + \exp(-N\epsilon) \leq 2\epsilon, \tag{61}$$

for N sufficiently large, if only

$$R_s \geq H(Y, Z|X) + 3\epsilon. \tag{62}$$

Appendix C: Achievable rate quadruples in the feedback case

In this appendix we will show that when a feedback scenario is allowed the set of achievable rate quadbruples is not larger than in the non-feedback case. To do so we only have to adjust the converse that was given in subsection 2.1. We briefly show the steps that are different.

We first show, using Fano's inequality, that also

$$H(Y^N, Z^N|X^N, W^B_{s,1}) \leq \log 2 + NP_e \log \|\mathcal{X}\| \|\mathcal{Y}\| \|\mathcal{Z}\| = N\Delta(P_e, N), \tag{63}$$

and similar inequalities hold for $H(Z^N, X^N|Y^N, W^B_{s,1})$ and $H(X^N, Y^N|Z^N, W^B_{s,1})$.

Assume that $W_{x,b} \in \{1, 2, \cdots, M_{x,b}\}$ for $b = 1, B$, etcetera. Now we continue with

$$\begin{aligned}
\sum_{b=1,B} \log M_{x,b} &\geq \sum_{b=1,B} H(W_{x,b}) \geq H(W^B_{x,1}) \\
&\geq I(W^B_{x,1}; X^N|Y^N, Z^N) \\
&= H(X^N|Y^N, Z^N) - H(X^N|Y^N, Z^N, W^B_{x,1}) \\
&\overset{(\gamma)}{=} NH(X|Y, Z) - H(X^N|Y^N, W^B_{y,1}, Z^N, W^B_{z,1}, W^B_{x,1}, W^B_{s,1}) \\
&\geq NH(X|Y, Z) - H(X^N, Y^N|Z^N, W^B_{s,1}) \\
&\geq N(H(X|Y, Z) - \Delta(P_e, N)),
\end{aligned} \tag{64}$$

where (γ) follows from (45) and (46). In the same way we obtain lower bounds for the sums $\sum_{b=1,B} \log M_{y,b}$ and $\sum_{b=1,B} \log M_{z,b}$. Next we proceed with the sum

$$\begin{aligned}
\sum_{b=1,B} \log M_{y,b}M_{z,b} &\geq \sum_{b=1,B} H(W_{y,b}, W_{z,b}) \geq H(W^B_{y,1}, W^B_{z,1}) \\
&\geq I(W^B_{y,1}, W^B_{z,1}; Y^N, Z^N|X^N) \\
&= H(Y^N, Z^N|X^N) - H(Y^N, Z^N|X^N, W^B_{y,1}, W^B_{z,1}) \\
&= NH(Y, Z|X) - H(Y^N, Z^N|X^N, W^B_{x,1}, W^B_{y,1}, W^B_{z,1}, W^B_{s,1}) \\
&\geq NH(Y, Z|X) - H(Y^N, Z^N|X^N, W^B_{s,1}) \\
&\geq N(H(Y, Z|X) - \Delta(P_e, N)),
\end{aligned} \tag{65}$$

and equivalently we get bounds for $\sum_{b=1,B} \log M_{z,b}M_{x,b}$ and $\sum_{b=1,B} \log M_{x,b}M_{y,b}$. Next

$$\begin{aligned}
\sum_{b=1,B} \log M_{s,b} &\geq \sum_{b=1,B} H(W_{s,b}) \geq H(W^B_{s,1}) \\
&\geq I(W^B_{s,1}; Y^N, Z^N|X^N) \\
&= H(Y^N, Z^N|X^N) - H(Y^N, Z^N|X^N, W^B_{s,1}) \\
&= N(H(Y, Z|X) - \Delta(P_e, N)),
\end{aligned} \tag{66}$$

and in the same way we obtain two more lower bounds for $\sum_{b=1,B} \log M_{s,b}$. Finally from (64) we obtain that an achievable quadruple (R_x, R_y, R_z, R_s) satisfies

$$R_x + \delta \geq \frac{\sum_{b=1,B} \log M_{x,b}}{N} \geq H(X|Y, Z) - \Delta(P_e, N) \geq H(X|Y, Z) - \Delta(\delta, N) \tag{67}$$

for all $\delta > 0$ and all N large enough. Note that in the definition of achievability for the feedback scenario (10) we should replace $\log M_x$ by $\sum_{b=1,B} M_{x,b}$, etcetera. Observe that $\Delta(\delta, N) \to 0$ for $\delta \to 0$ and $N \to \infty$. This implies that

$$R_x \geq H(X|Y, Z), \tag{68}$$

for all achievable rate quadruples (R_x, R_y, R_z, R_s). Etcetera.

References

[1] R. Ahlswede, N. Cai, S.-Y.R. Li, R.W. Yeung, ”Network information flow,” *IEEE Trans. Inform. Theory*, vol. 46, July 2000, pp. 1204-1216.

[2] T. Berger, ”Multiterminal source coding,” In G. Longo, editor, *The Information Theory Approach to Communications.* Springer-Verlag, New York, 1977.

[3] T.M. Cover, ”A proof of the data compression theorem of Slepian and Wolf for ergodic sources, *IEEE Trans. Inform. Theory*, vol. 22, March 1975, pp. 226-228.

[4] T.M. Cover, A. El Gamal, and M. Salehi, ”Multiple Access Channels with Arbitrarily Correlated Sources,” *IEEE Trans. Inform. Theory*, vol. 26, November 1980, pp. 648-657.

[5] T.M. Cover and J.A. Thomas, *Elements of Information Theory.* Wiley, New York 1991.

[6] J. Körner and K. Marton, ”How to encode the modulo-two sum of binary sources,” *IEEE Trans. Inform. Theory*, vol. 25, March 1979, pp. 219-221.

[7] A. Sgarro, ”Source Coding with Side Information at Several Decoders,” *IEEE Trans. Inform. Theory*, vol. 23, March 77, pp. 179-182.

[8] D. Slepian and J.K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. Inform. Theory*, vol. 19, July 1973, pp. 471-480.

[9] L. Song and R.W. Yeung, ”Network information flow - Multiple sources,” *Proceedings, 2001 IEEE International Symposium on Information Theory*, Washington DC, June 24-29, 2001, p. 102.

[10] F.M.J. Willems, J.K. Wolf, and A.D. Wyner, ”Communicating via a processing broadcast satellite,” *1989 IEEE/CAM Information Theory Workshop at Cornell*, Ithaca, June 25-29, 1989, p. 3-1.

[11] F.M.J. Willems, J.K. Wolf, and A.D. Wyner, ”Communicating via a processing broadcast satellite,” *Abstracts of Papers 1990 IEEE International Symposium on Information Theory*, San Diego, January 14-19, 1990, pp. 28-29.

[12] R.E. Yeung and Z. Zhang, ”Distributed source coding for satellite communications,” *IEEE Trans. Inform. Theory*, vol. 45, May 1999, pp. 1111-1120.

BIOGRAPHIES

Jack Keil Wolf is the Stephen O. Rice Professor of Electrical and Computer Engineering and a member of the Center for Magnetic Recording Research at the University of California-San Diego, La Jolla, CA. He received his B.S.E.E. degree from the University of Pennsylvania in 1956, and the M.S.E., M.A., and Ph.D. degrees from Princeton in 1957, 1958, and 1960, respectively.

He was a member of the Electrical Engineering Department at New York University from 1963 to 1965, and the Polytechnic Institute of Brooklyn from 1965 to 1973. He was chairman of the Department of Electrical and Computer Engineering at the University of Massachusetts from 1973 to 1975, and was Professor there from 1973 to 1984. Since 1984 he has been a Professor of Electrical and Computer Engineering and a member of the Center for Magnetic Recording Research at the University of California-San Diego. He also holds a part-time appointment at Qualcomm, Inc., San Diego, California. From 1971 to 1972 he was an NSF Senior Postdoctoral Fellow, and from 1979 to 1980 he held a Guggenheim Fellowship. His current interest is in signal processing for storage systems.

Dr. Wolf has been a Fellow of the IEEE since 1973. In 1993 he was elected to the National Academy of Engineering. He was recipient of the 1990 E. H. Armstrong Achievement Award of the IEEE Communications Society and was co-recipient of the 1975 IEEE Information Theory Group Paper Award for the paper ”Noiseless Coding for Correlated Information Sources” (co-authored with D. Slepian). He served on the Board of Governors of the IEEE Information Theory Group from 1970 to 1976 and from 1980 to 1986. Dr. Wolf was president of the IEEE Information Theory Group in 1974. He was International Chairman of Committee C of URSI from 1980 to 1983. Dr. Wolf was the recipient of the 1998 IEEE Koji Kobayashi Computers and Communications Award, ”for fundamental contributions to multi-user communications and applications of coding theory to magnetic data storage devices.” In May 2000, Dr. Wolf received a UCSD Distinguished Teaching Award. In June 2001 he received the 2001 Claude E. Shannon Award from the IEEE Information theory Society.

Frans M.J. Willems was born in Stein, The Netherlands, on June 26, 1954. He received the M.S. degree in electrical engineering from Eindhoven University of Technology, Eindhoven, The Netherlands, and the Ph.D. degree from the Catholic University of Louvain, Louvain, Belgium, in 1979 and 1982 respectively.

From 1979 to 1982 he was a research assistant at the Catholic University of Louvain. Since 1982, he is a staff member at the Electrical Engineering Department of Eindhoven University of Technology. His research contributions are in the areas of multi-user information theory and noiseless source coding.

Dr. Willems received the Marconi Young Scientist Award in 1982. From 1988 to 1990, he served as Associate Editor for Shannon Theory for the *IEEE Transactions on Information Theory*. He is co-recipient of the 1996 IEEE Information Theory Society Paper Award. From 1998 to 2000 he was a member of the Board of Governors of the IEEE Information Theory Society.