- [14] E. Kurtas and M. Salehi, "Source-channel matching for a simple network," in *Proc. 27th Annual Conf. on Information Sciences and Systems* (Baltimore, MD), Mar. 1993.
- [15] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems. New York: Academic Press, 1981.
- [16] A. ElGamal and T. M. Cover, "Multiple user information theory," *Proc. IEEE*, vol. 68, pp. 1466–1483, Dec. 1980.
- [17] M. Salehi, "Matching of correlated sources to random multiple access channels," in Abstracts of Papers, Int. Symp. on Information Theory and its Applications (Waikiki, HI), Nov. 1990.

# Feedback Does Not Increase the Capacity of Discrete Channels with Additive Noise

Fady Alajaji

Abstract— We consider discrete-time finite alphabet channels with additive random noise. We show that output feedback does not increase the capacity of such channels. This result holds in the most general case; i.e., for arbitrary additive noise processes.

Index Terms— Shannon theory, output feedback, capacity, discrete channels with memory, additive noise.

#### I. INTRODUCTION

We consider discrete (discrete-time finite alphabet) channels with additive random noise. Note that such channels need not be memoryless; in general, they have memory. The Gilbert burst-noise channel [5], as well as the Polya-contagion channel [1], belong to the class of such channels. We assume that these channels are each accompanied by a noiseless, delayless feedback channel with large capacity. Intuitively, it is plausible that if we use feedback on channels with memory, then we can use some encoding techniques at the transmitter end in order to combat the channel noise and hence increase the channel capacity. However, we reach the seemingly surprising result that the capacity of the additive channels with feedback does not exceed their respective capacity without feedback. This is demonstrated for *arbitrary* (nonstationary, nonergodic in general) additive noise processes, using recent results by Verdú and Han on a general channel capacity formula [10].

For these channels, the capacities with and without feedback are equal because additive noise channels are *symmetric* channels. By this we mean that the *inf-information* rate between input and output processes is maximized by an equally likely independent and identically distributed (iid) input process. Furthermore, this maximizing equally likely iid input process yields an equally likely iid output process.

In earlier related work, Shannon [9] showed that feedback does not increase the capacity of discrete memoryless channels. The same result was proven to be true for continuous alphabet channels

Manuscript received March 18, 1993; revised September 2, 1994. This work was supported in part by the National Science Foundation under Grant NCR-8957623 and by the NSF Engineering Research Centers Program CDR-8803012. Parts of the material in this correspondence were presented at the 1993 Conference on Information Sciences and Systems, The Johns Hopkins University, Baltimore, MD, March 24–26, 1993.

The author is with the Electrical Engineering Department, Institute for Systems Research, University of Maryland, College Park, MD 20742 USA. IEEE Log Number 9408066.

with additive white Gaussian noise. Later, Cover and Pombra [4] and others considered continuous alphabet channels with additive nonwhite Gaussian noise and showed that feedback increases their capacity by at most half a bit; similarly, it has been shown [4] that feedback can at most double the capacity of a nonwhite Gaussian channel.

#### II. CAPACITY WITH NO FEEDBACK

Consider a discrete channel with common input, noise, and output q-ary alphabet A where  $A = \{0, 1, \dots, q-1\}$ , described by the following equation:  $Y_n = X_n \oplus Z_n$ , for  $n = 1, 2, 3, \dots$ , where

- $\oplus$  represents the addition operation modulo q.
- The random variables X<sub>n</sub>, Z<sub>n</sub>, and Y<sub>n</sub> are, respectively, the input, noise, and output of the channel.
- {X<sub>n</sub>} \perp {Z<sub>n</sub>}, i.e., the input and noise sequences are independent from each other.
- The noise process  $\{Z_n\}_{n=1}^{n=\infty}$  is an *arbitrary* random process (nonstationary, nonergodic in general).

Note that additive channels defined above, are "nonanticipatory" channels; where by "nonanticipatory" we mean channels with no input memory (i.e., historyless) and no anticipation (i.e., causal) [6]. A channel is said to have no anticipation if for a given input and a given input—output history, its current output is independent of future inputs. Furthermore, a channel is said to have no input memory if its current output is independent of previous inputs. Refer to [6] for more rigorous definitions of causal and historyless channels.

We furthermore note that discrete additive noise channels have a *symmetry* property. By this we mean that their input—output infinformation rate is maximized by an equally likely iid input process, which also yields an equally likely iid output process. This is due to the facts that the input and noise processes of these channels are independent from each other, the addition operation (modulo q) is invertible, and the input and output alphabets are *finite* and have the *same cardinality*.

A channel code with blocklength n and rate R consists of an encoder<sup>1</sup>

$$f: \{1, 2, \cdots, 2^{nR}\} \to A^n$$

and a decoder

$$g: A^n \to \{1, 2, \cdots, 2^{nR}\}.$$

The encoder represents the message  $V \in \{1, 2, \dots, 2^{nR}\}$  with the codeword  $f(V) = X^n = [X_1, X_2, \dots, X_n]$  which is then transmitted over the channel; at the receiver, the decoder observes the channel output  $Y^n = [Y_1, Y_2, \dots, Y_n]$ , and chooses as its estimate of the message  $\hat{V} = g(Y^n)$ . A decoding error occurs if  $\hat{V} \neq V$ .

For additive channels,  $Y_i = X_i \oplus Z_i$  for all i. We assume that V is uniformly distributed over  $\{1, 2, \dots, 2^{nR}\}$ . The probability of decoding error<sup>2</sup> is thus given by

$$P_{\epsilon}^{(n)} = \frac{1}{2^{nR}} \sum_{k=1}^{2^{nR}} \Pr \left\{ g(\boldsymbol{Y}^n) \neq \boldsymbol{V} \mid \boldsymbol{V} = k \right\} = \Pr \left\{ g(\boldsymbol{Y}^n) \neq \boldsymbol{V} \right\}.$$

<sup>1</sup>The number of messages is  $2^{nR}$ . If  $2^{nR}$  is not an integer, then we replace it with  $\lceil 2^{nR} \rceil$ . We will however write it as  $2^{nR}$  for notational simplicity.

<sup>2</sup>We consider the average error probability. The analysis remains unchanged if we work with the maximal error probability since the capacity of a single-user channel with known statistics is the same under both error probability criteria.

We say that a rate R is achievable (admissible) if there exists a sequence of codes with blocklength n and rate R such that

$$\lim_{n\to\infty} P_e^{(n)} = 0.$$

The objective is to find an admissible sequence of codes with as high a rate as possible. The capacity of the channel is defined as the supremum of the rate over all admissible sequences of codes. We denote it by  $C_{NFB}$ , to stand for capacity with no feedback.

In [10], Verdú and Han derived a formula for the operational capacity of arbitrary single-users channels (not necessarily stationary, ergodic, information-stable, etc.). The (nonfeedback) capacity was shown to equal the supremum, over all input processes, of the input-output *inf-information rate* defined as the liminf in probability of the normalized information density

$$C_{NFB} = \sup_{X^n} \underline{I}(X^n; Y^n) \tag{1}$$

where  $X^n=(X_1,X_2,\cdots,X_n)$ , for  $n=1,2,\cdots$ , is the block input vector and  $Y^n=(Y_1,Y_2,\cdots,Y_n)$  is the corresponding output sequence induced by  $X^n$  via the channel  $W^{(n)}=P_{Y^n|X^n}\colon A^n\to B^n$ ;  $n=1,2,\cdots$ , which is an arbitrary sequence of n-dimensional conditional output distributions from  $A^n$  to  $B^n$ , where A and B are the input and output alphabets, respectively.

The symbol  $\underline{I}(X^n; Y^n)$  appearing in (1) is the *inf-information* rate between  $X^n$  and  $Y^n$  and is defined as the *liminf* in probability of the sequence of normalized information densities  $(1/n) i_{X^nY^n}(X^n; Y^n)$ , where

$$i_{X^nY^n}(a^n; b^n) = \log_2 \frac{P_{Y^n|X^n}(b^n \mid a^n)}{P_{Y^n}(b^n)}.$$
 (2)

The *liminf in probability* of a sequence of random variables is defined as follows: if  $A_n$  is a sequence of random variables, then its *liminf in probability* is the supremum of all reals  $\alpha$  for which  $P(A_n \leq \alpha) \to 0$  as  $n \to \infty$ . Similarly, its *limsup in probability* is the infimum of all reals  $\beta$  for which  $P(A_n \geq \beta) \to 0$  as  $n \to \infty$ . Note that these two quantities are always defined; if they are equal, then the sequence of random variables converges in probability to a constant (which is  $\alpha$ ).

Using (1) as well as the properties of the inf-information rate derived in [10], we obtain that the inf-information rate in (1) is maximized for equiprobable iid  $X^n$  (symmetry property), yielding the following expression for the nonfeedback capacity of our discrete channel with arbitrary additive noise:

$$C_{NFB} = \log_2(q) - \overline{H}(Z^n)$$
 (3)

where  $Z^n = (Z_1, Z_2, \dots, Z_n)$  and  $\overline{H}(Z^n)$  is the sup-entropy rate of the additive noise process  $\{Z_n\}$ , which is defined as the limsup in probability of the normalized noise entropy density

$$\frac{1}{n}\log_2\frac{1}{P_{Z^n}(Z^n)}.$$

### III. CAPACITY WITH FEEDBACK

We now consider the corresponding problem for the discrete additive channel with complete output feedback. By this we mean that there exists a "return channel" from the receiver to the transmitter; we assume this return channel is noiseless, delayless, and has large capacity. The receiver uses the return channel to inform the transmitter what letters were actually received; these letters are received at the transmitter before the next letter is transmitted, and therefore can be used in choosing the next transmitted letter.

A feedback code with blocklength n and rate R consists of a sequence of encoders

$$f_i: \{1, 2, \cdots, 2^{nR}\} \times A^{i-1} \to A$$

for  $i = 1, 2, \dots, n$ , along with a decoding function

$$g: A^n \to \{1, 2, \cdots, 2^{nR}\}.$$

The interpretation is simple: If the user wishes to convey message  $V \in \{1, 2, \cdots, 2^{nR}\}$  then the first code symbol transmitted is  $X_1 = f_1(V)$ ; the second code symbol transmitted is  $X_2 = f_2(V, Y_1)$ , where  $Y_1$  is the channel's output due to  $X_1$ . The third code symbol transmitted is  $X_3 = f_3(V, Y_1, Y_2)$ , where  $Y_2$  is the channel's output due to  $X_2$ . This process is continued until the encoder transmits  $X_n = f_n(V, Y_1, Y_2, \cdots, Y_{n-1})$ . At this point the decoder estimates the message to be  $g(Y^n)$ , where  $Y^n = [Y_1, Y_2, \cdots, Y_n]$ .

Assuming our additive channel  $Y_i = X_i \oplus Z_i$  where  $\{Z_i\}$  is an arbitrary noise process. Again, we assume that V is uniformly distributed over  $\{1, 2, \cdots, 2^{nR}\}$ , and we define the probability of error and achievability as in Section II.

Note, however, that because of the feedback,  $X^n$  and  $Z^n$  are no longer independent;  $X_i$  may depend on  $Z^{i-1}$ .

We denote the capacity of the channel with feedback by  $C_{FB}$ . As before,  $C_{FB}$  is the supremum of all admissible feedback code rates.

We now state the key result [10, Theorem 4] which is a new converse approach based on a simple new lower bound on the error probability of an arbitrary channel code as a function of its size.

*Lemma:* Let  $(n, M, \epsilon)$  represent a channel block code with blocklength n, M codewords, and (average) error probability  $\epsilon$ . Then every  $(n, M, \epsilon)$  code satisfies

$$\epsilon \ge P\left[\frac{1}{n}i_{X^nY^n}(X^n;Y^n) \le \frac{1}{n}\log_2 M - \gamma\right] - \exp(-\gamma n)$$
 (4)

for every  $\gamma > 0$ , where  $X^n$  places probability mass 1/M on each codeword.

We now obtain our main result:

Theorem: Feedback does not increase the capacity of discrete channels with arbitrary additive noise

$$C_{FB} = C_{NFB} = \log_2(q) - \overline{H}(Z^n). \tag{5}$$

*Proof:* We start by noting that the result given in the above lemma still holds if we replace the input vector  $X^n$  by the message random variable V where V is uniform over the set of messages  $\{1, 2, \dots, M\}$ . That is, every  $(n, M, \epsilon)$  feedback code satisfies

$$\epsilon \ge P\left[\frac{1}{n}i_{VY^n}(V;Y^n) \le \frac{1}{n}\log_2 M - \gamma\right] - \exp\left(-\gamma n\right)$$
 (6)

for every  $\gamma > 0$ , where V is uniform over  $\{1, 2, \cdots, M\}$ .

We refer to the sequence  $(n, M, \epsilon_n)$  of feedback codes with vanishingly small error probability (i.e.,  $\epsilon_n \to 0$  as  $n \to \infty$ ) as a reliable feedback code sequence.

Using (6), we first show that

$$C_{FB} \le \sup_{X^n} \underline{I}(V; Y^n) \tag{7}$$

where the supremum is taken over all possible feedback encoding schemes.<sup>3</sup>

We prove (7) by contradiction. Assume that for some  $\rho > 0$ 

$$C_{FB} = \sup_{V^n} \underline{I}(V; Y^n) + 3\rho. \tag{8}$$

$$\begin{array}{ll} ^3 \sup \underline{I}(V;Y^n) &= \sup_{X^n = (f_1(V),\, f_2(V,Y_1), \cdots, f_n(V,Y^{n-1}))} \underline{I}(V;Y^n) = \\ \sup_{(f_1,f_2, \cdots, f_n)} \underline{I}(V;Y^n). \end{array}$$

By definition of capacity, there exists a reliable feedback code sequence with rate

$$R = \frac{1}{n} \log_2 M > C_{FB} - \rho. \tag{9}$$

Now using (6) (with  $\gamma = \rho$ ) along with (8) and (9), we obtain that the error probability of the sequence  $(n, M, \epsilon_n)$  of feedback codes must be lower bounded by

$$\epsilon_n \ge P\left[\frac{1}{n}i_{VY^n}(V;Y^n) \le \sup_{X^n} \underline{I}(V;Y^n) + \rho\right] - \exp(-\rho n).$$
 (10)

However, by definition of  $\underline{I}(V; Y^n)$  the probability in the right side of (10) cannot vanish asymptotically; therefore, contradicting the fact that  $\epsilon_n \to 0$  as  $n \to \infty$ . Thus (7) is proved.

Now using the properties of the inf-information rate in [10], we can write

$$\underline{I}(V; Y^n) \le \overline{H}(Y^n) - \overline{H}(Y^n \mid V) \le \log_2(q) - \overline{H}(Y^n \mid V). \tag{11}$$

The conditional sup-entropy rate  $\overline{H}(Y^n\mid V)$  is the limsup in probability (according to  $P_{VY^n}$ ) of

$$\frac{1}{n}\log_2\frac{1}{P_{Y^n\mid V}(Y^n\mid V)}.$$

That is,  $\overline{H}(Y^n \mid V)$  is the infimum of all reals  $\beta$  such that

$$\Pr\left\{\frac{1}{n}\log_2\frac{1}{P_{Y^n|V}(Y^n\mid V)}\geq \beta\right\}\to 0, \qquad \text{as } n\to\infty.$$

But we can write

$$\begin{split} \Pr\left\{\frac{1}{n}\log_2\frac{1}{P_{Y^n\mid V}(Y^n\mid V)} \geq \beta\right\} &= \sum_v P(V=v)\\ &\cdot \sum_{y^n: P(Y^n=y^n\mid V=v) \leq 2^{-n\beta}} P(Y^n=y^n\mid V=v). \end{split}$$

Now, letting

$$f_i \stackrel{\Delta}{=} f_i(v, y^{i-1})$$

and

$$f^{i} \stackrel{\Delta}{=} [f_{1}(v), f_{2}(v, y_{1}), \cdots, f_{i}(v, y^{i-1})] = [f_{1}, f_{2}, \cdots, f_{i}]$$

we have

$$P(Y^{n} = y^{n} | V = v)$$

$$= \prod_{i=1}^{n} P(Y_{i} = y_{i} | Y^{i-1} = y^{i-1}, V = v)$$

$$= \prod_{i=1}^{n} P(X_{i} \oplus Z_{i} = y_{i} | Y^{i-1} = y^{i-1}, V = v, X_{i} = f_{i})$$

$$= \prod_{i=1}^{n} P(Z_{i} = y_{i} - f_{i} | Y^{i-1} = y^{i-1}, V = v, X_{i} = f_{i})$$

$$= \prod_{i=1}^{n} P(Z_{i} = y_{i} - f_{i} | Y^{i-1} = y^{i-1}, V = v, X_{i} = f_{i})$$
(13)

$$= \prod_{i=1}^{n} P(Z_i = y_i - f_i \mid Y^{i-1} = y^{i-1}, V = v,$$

$$X^{i} = f^{i}, Z^{i-1} = y^{i-1} - f^{i-1})^{4}$$
 (12)

$$= \prod_{i=1}^{n} P(Z_i = y_i - f_i \mid Z^{i-1} = y^{i-1} - f^{i-1})$$
 (15)

$$=P(Z^n=y^n-f^n). (16)$$

Here

· Equation (12) follows from the fact that

$$X_i = f_i(V, Y_1, \cdots, Y_{i-1})$$

due to feedback.

- Equation (13) holds since  $P(Z \oplus X = y \mid X = x) = P(Z = y x \mid X = x)$ .
- Equation (14) follows from the fact that given V and Y<sup>i-1</sup>, we know all the previous transmitted letters X<sub>1</sub>, X<sub>2</sub>,..., X<sub>i-1</sub> and thus we can recover all the previous noise letters Z<sub>j</sub> = Y<sub>j</sub> X<sub>j</sub> (mod q) for j = 1, 2,..., i 1.
- Equation (15) follows from the fact that  $Z_i$  and  $(V, Y^{i-1}, X^i)$  are conditionally independent given  $Z^{i-1}$ .

Hence

$$\begin{split} & \Pr\left\{\frac{1}{n}\log_{2}\frac{1}{P_{Y^{n}\mid V}(Y^{n}\mid V)} \geq \beta\right\} \\ & = \sum_{v} P(V=v) \sum_{y^{n}: P(Z^{n}=y^{n}-f^{n}) \leq 2^{-n\beta}} P(Z^{n}=y^{n}-f^{n}) \\ & = \sum_{v} P(V=v) \sum_{z^{n}: P(Z^{n}=z^{n}) \leq 2^{-n\beta}} P(Z^{n}=z^{n}) \\ & = \sum_{z^{n}: P(Z^{n}=z^{n}) \leq 2^{-n\beta}} P(Z^{n}=z^{n}). \end{split}$$

Therefore, we obtain that

$$\overline{H}(Y^n \mid V) = \overline{H}(Z^n). \tag{17}$$

Thus from (7), (11), and (17) we conclude that

$$C_{FB} \le \log_2(q) - \overline{H}(Z^n) = C_{NFB}. \tag{18}$$

But by definition of a feedback code,  $C_{FB} \ge C_{NFB}$  since a nonfeedback code is a special case of a feedback code. Thus we get

$$C_{FB} = C_{NFB} = \log_2(q) - \overline{H}(Z^n). \tag{19}$$

Corollary 1: If the noise process is stationary, then its sup-entropy rate is equal to the supremum over the entropies of almost every ergodic component of the stationary noise [10]. Thus (19) reduces to the formula derived by Parthasarathy [8] and Kieffer [7]

$$C_{FB} = C_{NFB} = \log_2(q) - \operatorname{ess}_{\Theta} \sup h(Z_{\theta})$$

where

•  $h(Z_{\theta})$  is the entropy rate of the  $\theta$  th ergodic component of the stationary noise process

$$h(Z_{\theta}) \stackrel{\Delta}{=} \lim_{n \to \infty} \frac{1}{n} H_n(Z_{\theta}^n)$$

with

$$H_n(Z_{\theta}^n) \stackrel{\Delta}{=} - \sum_{z^n \in A^n} P_{Z_{\theta}}^{(n)}(z^n) \log_2 P_{Z_{\theta}}^{(n)}(z^n)$$

• and the essential supremum is defined by

$$\operatorname{ess}_{\Theta} \sup f(\theta) \stackrel{\Delta}{=} \inf [r] : dG(f(\theta) \leq r) = 1]$$

where G is a probability measure defined on the event space of  $\Theta$ , the set of the ergodic components of the noise process.

<sup>4</sup>where the modulo q difference between two vectors  $a^n = (a_1, \dots, a_n)$  and  $b^n = (b_1, \dots, b_n)$  is defined as  $a^n - b^n \stackrel{\triangle}{=} (a_1 - b_1, \dots, a_n - b_n)$ .

Corollary 2: If the noise process is stationary ergodic, then its sup-entropy rate is equal to the entropy rate of the noise (by the Shannon-McMillan theorem). Thus (19) reduces to the familiar expression of Shannon

$$C_{FB} = C_{NFB} = \log_2\left(q\right) - \lim_{n \to \infty} \frac{1}{n} H(Z^n)$$

where  $H(Z^n)$  is the entropy of the noise vector  $Z^n$ .

Corollary 2 can be directly proven (in a similar way as for the general case of arbitrary noise) using Fano's inequality [3]. Corollary 1 can also be proven using the Ergodic Decomposition Theorem for stationary processes and the properties of averaged channels [3].

Observation: The reason why output feedback potentially increases the capacity of additive nonwhite Gaussian channels [4] is because for continuous alphabet channels we have power constraints on the input, which upon optimization may increase  $\lim_{n \to \infty} \frac{1}{n} H(Y^n)$ (assuming, for example, that the noise is stationary-ergodic) when feedback is used; while for discrete channels this quantity is upperbounded by  $log_2(q)$  and cannot be increased with feedback. In particular for discrete additive channels, the output entropy rate is equal to  $\log_2(q)$  without feedback (symmetry property). It is therefore suspected that feedback might increase the capacity of discrete additive channels if we impose power constraints on the input.5

# IV. CONCLUSIONS

In this work, we considered a discrete additive noise channel with output feedback. We showed that the capacity of the channel without feedback equals its capacity with feedback. This was shown for arbitrary additive noise processes.

In [2], [3], we introduce the notion of symmetric channels with memory. These channels are obtained by combining an input process with an arbitrary noise process that is independent of the input, and possess the symmetry property defined earlier. We show that feedback does not also increase the capacity of these channels. Additive noise channels belong to the class of symmetric channels. The effect of feedback on the capacity of additive noise channels that are subject to average cost constraints on their input sequences is also addressed in [2], [3]. In this case, it is shown that the capacity-cost function can be increased by feedback.

Future studies may involve the investigation of the capacity of nonsymmetric channels with feedback, like the "AND" channel (the multiplicative channel with alphabet  $\{0, 1\}$ ) or the real adder channel. It is conjectured that feedback does cause an increase in capacity for such channels.

# ACKNOWLEDGMENT

The author wishes to thank Prof. I. Csiszàr and Prof. T. Fuja for their very valuable advice and constructive criticism. The author wishes also to thank Prof. S. Verdú for introducing [10] to him.

## REFERENCES

- [1] F. Alajaji and T. Fuja, "A communication channel modeled by the spread of disease," in Proc. IEEE Int. Symp. on Information Theory (San Antonio, TX, Jan. 1993).
- \_, "Effect of feedback on the capacity of discrete additive channels with memory," in Proc. IEEE Int. Symp. on Information Theory (Trondheim, Norway, June 1994).
- <sup>5</sup>This is indeed demonstrated in [3].

- [3] F. Alajaji, "New results on the analysis of discrete communication channels with memory," Ph.D. dissertation, Dept. Elec. Eng., Univ. of Maryland, College Park, Aug. 1994.
- T. M. Cover and S. Pombra, "Gaussian feedback capacity," IEEE Trans.
- Inform. Theory, vol. 35, pp. 37-43, 1989. E. N. Gilbert, "Capacity of burst-noise channels," Bell Syst. Tech. J., vol. 39, pp. 1253-1265, 1960.
- R. M. Gray and D. S. Ornstein, "Block coding for discrete stationary  $\overline{d}$ -continuous noisy channels," IEEE Trans. Inform. Theory, vol. IT-25, pp. 292-306, 1979.
- J. C. Kieffer, "A general formula for the capacity of stationary nonanticipatory channels," *Inform. Contr.*, vol. 26, pp. 381–391, 1974.
- K. R. Parthasarathy, "Effective entropy rate and transmission of information through channels with additive random noise," Sankhya, vol. A(25), pp. 75-84, 1963.
- C. E. Shannon, "The zero-error capacity of a noisy channel," IRE Trans. Inform. Theory, vol. IT-2, pp. 8-19, 1956.
- S. Verdú and T. S. Han, "A general formula for channel capacity," IEEE Trans. Inform. Theory, vol. 40, pp. 1147-1157, July 1994.

# New Bounds on the Information Rate of Secret Sharing Schemes

Carlo Blundo, Alfredo De Santis, Member, IEEE, Antonio Giorgio Gaggia, and Ugo Vaccaro

Abstract-A secret sharing scheme permits a secret to be shared among participants in such a way that only qualified subsets of participants can recover the secret, but any nonqualified subset has absolutely no information on the secret. In this correspondence we derive new limitations on the information rate of secret sharing schemes, that measures how much information is being distributed as shares as compared to the size of the secret key, and the average information rate, that is the ratio between the secret size and the arithmetic mean of the size of the shares. By applying the substitution technique, we are able to construct many new examples of access structures where the information rate is bounded away from 1. The substitution technique is a method to obtain a new access structure by replacing a participant in a previous structure with a new access structure.

Index Terms-Data security, cryptography, secret sharing, information

### I. INTRODUCTION

A secret sharing scheme is a method to distribute a secret s among a set of participants  $\mathcal{P}$  in such a way that only qualified subsets of  $\mathcal{P}$  can reconstruct the value of s whereas any other (nonqualified) subset of  ${\mathcal P}$  cannot determine anything about the value of the secret.

Blakley [3] and Shamir [20] initiated the study of secret sharing schemes, giving algorithm to realize  $(k,\,n)$  threshold schemes. A (k, n) threshold scheme allows a secret to be shared among nparticipants in such a way that any k of them can recover the

Manuscript received April 6, 1993; revised October 10, 1994. This work was partially supported by the Italian Ministry of University and Scientific Research in the framework of the project "Algoritmi, Modelli di Calcolo e Structure Informative" and by the National Council of Research. Part of the work was performed while one of the authors (C. Blundo) was visiting the Department of Computer Science and Engineering, University of Nebraska-Lincoln.

The authors are with Dipartimento di Informatica ed Applicazioni, Università di Salerno, 84081 Baronissi (SA), Italy.

IEEE Log Number 9408642.