A Bound on the Financial Value of Information

ANDREW R. BARRON, MEMBER, IEEE, AND
THOMAS M. COVER, FELLOW, IEEE

Abstract —**It will be shown that each bit of information at most doubles the resulting wealth in the general stock market setup. This information bound on the growth of wealth is actually attained for certain probability distributions on the market investigated by Kelly. The bound will be shown to be a special case of the result that the increase in exponential growth of wealth achieved with true knowledge of the stock market distribution F over that achieved with incorrect knowledge G is bounded above by $D(F\|G)$, the entropy of F relative to G.**

I. Introduction

Let $X \geq 0$, $X \in R^m$ denote a random stock market vector, with the interpretation that X_i is the ratio of the price of the ith stock at the end of an investment period to the price at the beginning. Let $B = \{b \in R^m\colon\ b_i \geq 0, \sum_{i=1}^m b_i = 1\}$, be the set of all portfolios b, where b_i is the proportion of wealth invested in the ith stock. The resulting wealth is

$$S = \sum_{i=1}^{m} b_i X_i = b^t X. \tag{1}$$

This is the wealth resulting from a unit investment allocated to the m stocks according to the portfolio b.

II. Doubling Rate

Now let $F(x)$ be the probability distribution function of the stock vector X. We define the doubling rate $W(X)$ for the market by

$$W(X) = \max_{b \in B} \int \log b^t x \, dF(x). \tag{2}$$

The units for W are "doubles per investment." All logarithms in this correspondence are to the base 2. Let $b^* = b^*(F)$ denote a portfolio achieving $W(X)$. Note that $W(X)$ is a real number, a functional of F; the apparent dependence of W on X is for notational convenience.

Necessary and sufficient conditions for b to maximize $E \log b^t X$ are

$$E\frac{X_i}{b^t X} = 1, \qquad \text{for } b_i > 0$$

$$E\frac{X_i}{b^t X} \leq 1, \qquad \text{for } b_i = 0. \tag{3}$$

These are the Kuhn–Tucker conditions characterizing $b^*(F)$ (see Bell and Cover [3], Cover [4], and Finkelstein and Whitley [5]).

If current wealth is reallocated according to b^* in repeated independent investments against stock vectors $X_1, X_2, \cdots$ independent identically distributed (i.i.d.) according to $F(x)$, then the wealth S_n^* at time n is given by

$$S_n^* = \prod_{i=1}^{n} b^{*t} X_i. \tag{4}$$

The strong law of large numbers for products yields

$$(S_n^*)^{1/n} = 2^{(1/n)\sum_{i=1}^n \log b^{*t} X_i} \to 2^W, \tag{5}$$

with probability one. Moreover, no other portfolio achieves a higher exponent (Breiman [1]; Algoet and Cover [9]).

Now suppose side information Y is available. Here Y could be world events, the behavior of a correlated market, or past information on previous outcomes X. Again we define the maximum expected logarithm of the wealth, but this time we allow the portfolio b to depend on Y. Let the doubling rate for side information be

$$W(X|Y) = \max_{b(y)} \iint \log b^t(y) x \, dF(x, y) \tag{6}$$

and let $b^*(y) = b^*(F_{X|y})$ be the portfolio achieving $W(X|Y)$. It can be shown that $b^*(y)$ maximizes the conditional expected logarithm of the wealth $E\{\log b^t X | Y = y\}$.

In repeated investments against $X_1, X_2, \cdots, X_n$ where (X_i, Y_i) are i.i.d. $\sim F(x, y)$, and $b^*(Y_i)$ is the portfolio used at investment time i given side information Y_i, we have resulting wealth

$$S_n^{**} = \prod_{i=1}^{n} b^{*t}(Y_i) X_i \tag{7}$$

with asymptotic behavior

$$(S_n^{**})^{1/n} \to 2^{W(X|Y)} \tag{8}$$

with probability one. It follows that the ratio of wealth with side information to that without side information has limit

$$\left(\frac{S_n^{**}}{S_n^*}\right)^{1/n} \to 2^{W(X|Y) - W(X)} \tag{9}$$

with probability one.

Let the difference between the maximum expected logarithm of wealth with Y and without Y be

$$\Delta = W(X|Y) - W(X). \tag{10}$$

Thus Δ is the increment in doubling rate due to the side information Y. It is this difference that we wish to bound. As an example, if $\Delta = 1$ then the information Y yields an additional doubling of the capital in each investment period. Finally, we observe from (6) and (2) that $\Delta \geq 0$. Information never hurts. Kelly [6] identified Δ with the mutual information for a "horse-race" stock market, a result we will generalize here.

III. Mutual Information and Relative Entropy

The relative entropy (or Kullback Leibler information number) of probability distributions F and G is

$$D(F\|G) = \int \log(f/g) \, dF \tag{11}$$

where f and g are the respective densities with respect to any dominating measure. (*Note:* D is infinite if $g(x)$ is zero on a set of positive probability with respect to F.)

The relative entropy may be interpreted as the error exponent for the hypothesis test F versus G (Stein's lemma; see Chernoff [2]). Another interpretation of the relative entropy for a discrete random vector $X \sim P$ is that $D(P\|Q)$ is the expected increase in description length of the Shannon–Fano code based on the incorrect distribution Q.

Let X, Y be two random variables with joint distribution P_{XY}. The relative entropy between the conditional distribution $P_{X|Y}$

Manuscript received January 8, 1988; revised February 15, 1988. This work was supported in part by the National Science Foundation under a research fellowship and under Contract NCR-85-20136 A1 and in part by the Office of Naval Research under Contract N00014-86-K-06. This work was partially presented at the 6th Annual Symposium on Information Theory, Tashkent, USSR, September 1984.

A. R. Barron is with the Department of Statistics, University of Illinois, 725 South Wright Street, Champaign, IL 61282.

T. M. Cover is with the Departments of Electrical Engineering and Statistics, Stanford University, Durand, Room 121, Stanford, CA 94305.

IEEE Log Number 8824063.

and the marginal distribution P_X is the mutual information

$$I(X;Y) = \int D(P_{X|Y=y} \| P_X) P_Y(dy) = D(P_{XY} \| P_X P_Y). \quad (12)$$

Of the many alternative expressions for I, the most evocative is the identity

$$I(X;Y) = H(X) - H(X|Y) \quad (13)$$

where $H(X)$ is the entropy of X and $H(X|Y)$ is the conditional entropy. Thus I is the amount the entropy of X is decreased by knowledge of Y. One can compare (13) with (10) to see why a relationship between Δ and I might be expected.

The mutual information I can also be interpreted as the information rate achievable in communication over the communication channel $P(x, y)$. There is also an interpretation of $I(X;Y)$ in terms of efficient descriptions. Since $H(X)$ bits are required to describe the value of the random variable X (if X is discrete), and since $H(X|Y)$ bits are required to describe X given knowledge of Y, the decrement in the expected description length of X is given by $H(X) - H(X|Y) = I(X;Y)$.

In summary, the mutual information $I(X;Y)$ is 1) the decrease in the entropy of X when Y is made available, 2) the number of bits by which the expected description length of X is reduced by knowledge of Y, 3) The rate in bits at which Y can communicate with X by appropriate choice of Y, 4) the error exponent for the hypothesis test (X,Y) independent versus (X,Y) dependent.

IV. Portfolios Based on Incorrect Distributions

Suppose that it is believed that $\boldsymbol{X} \sim G(\boldsymbol{x})$ when in fact $\boldsymbol{X} \sim F(\boldsymbol{x})$. Thus the incorrect portfolio $\boldsymbol{b}^*(G)$ is used instead of $\boldsymbol{b}^*(F)$. The doubling rate associated with portfolio $\boldsymbol{b}$ and distribution F can be written

$$W(\boldsymbol{b}, F) = \int \log \boldsymbol{b}^t \boldsymbol{x} \, dF(\boldsymbol{x}) \quad (14)$$

with resulting growth of wealth

$$S_n \doteq 2^{nW(\boldsymbol{b}, F)}. \quad (15)$$

The decrement in exponent from using $\boldsymbol{b}^*(G)$ is

$$\Delta W(F,G) = W(\boldsymbol{b}^*(F), F) - W(\boldsymbol{b}^*(G), F). \quad (16)$$

The following theorem is central to our results.

Theorem 1:

$$0 \le \Delta W(F,G) \le D(F\|G). \quad (17)$$

Proof: The first inequality $0 \le \Delta$ follows by the optimality of $\boldsymbol{b}^*(F)$ for the distribution F. The second inequality $\Delta \le D$ is shown to be a consequence of the optimality of $\boldsymbol{b}^*(G)$ for the distribution G. Let F and G have densities f and g with respect to some dominating measure. The result $\Delta \le D$ is trivially true if $D(F\|G) = \infty$, so it is henceforth assumed that D is finite (whence $F \ll G$).

Let

$$S_1^* = \boldsymbol{b}^{*t}(F)\boldsymbol{X} \qquad S_2 = \boldsymbol{b}^{*t}(G)\boldsymbol{X} \quad (18)$$

be the wealth factors corresponding to the optimal portfolios with respect to F and G. From the Kuhn–Tucker conditions the wealth factor S_2 is strictly positive with probability one with respect to G (and with respect to F since $F \ll G$). It follows (again since $F \ll G$) that the set $A = \{x: \; S_2 > 0, \; f(\boldsymbol{x}) > 0, \; g(\boldsymbol{x}) > 0\}$ has probability one with respect to F. Then

$$\begin{aligned} \Delta W(F,G) &= \int_A \left(\log \frac{S_1^*}{S_2} \right) dF \\ &= \int_A \log \left(\frac{S_1^*}{S_2} \frac{g}{f} \frac{f}{g} \right) dF \\ &= \int_A \log \left(\frac{S_1^*}{S_2} \frac{g}{f} \right) dF + D(F\|G) \\ &\le \log \int_A \frac{S_1^*}{S_2} \, dG + D(F\|G) \\ &\le D(F\|G) \end{aligned} \quad (19)$$

where the first inequality follows from the concavity of the logarithm and the second from the Kuhn–Tucker conditions for the optimality of $\boldsymbol{b}^*(G)$ for the distribution G.

We can improve Theorem 1 by normalizing $\boldsymbol{X}$. Let $\tilde{F}$ denote the distribution of $\boldsymbol{X}/\sum X_i$. We note that $E(\log \boldsymbol{b}_1^t \boldsymbol{X}/\boldsymbol{b}_2^t \boldsymbol{X})$ depends on the distribution $F(\boldsymbol{x})$ only through the distribution of $\boldsymbol{X}/\sum_{i=1}^m X_i \sim \tilde{F}$.

Corollary:

$$\Delta W(F,G) \le D(\tilde{F}\|\tilde{G}).$$

Remark: Another relationship between W and D is shown by Móri [13]. The doubling rate $W = W(\boldsymbol{b}^*(F), F)$ is equal to the minimum of $D(F\|G)$ over all distributions G for which $E_G X_i \le 1$, for $i = 1, 2, \cdots, m$.

V. The Information Bound for Side Information

We now ask how Δ and I are related for the stock market. We have

$$\Delta = E \log \frac{\boldsymbol{b}^{*t}(Y)\boldsymbol{X}}{\boldsymbol{b}^{*t}\boldsymbol{X}} \quad (20)$$

and

$$I = E \log \frac{f(\boldsymbol{X}, Y)}{f(\boldsymbol{X}) f(Y)} \quad (21)$$

where $(\boldsymbol{X}, Y) \sim F(\boldsymbol{x}, y)$. The first involves wealth and depends on the values $\boldsymbol{X}$ takes on. The second involves information and depends on $\boldsymbol{X}$ and Y only through the density $f(\boldsymbol{x}, y)$. The following theorem establishes that the increment Δ in the doubling rate resulting from side information Y is less than or equal to the mutual information I.

Theorem 2:

$$0 \le \Delta \le I(\boldsymbol{X}; Y). \quad (22)$$

Proof: For any y, let $P_{X|y}$ be the conditional distribution for $\boldsymbol{X}$ given that $Y = y$ and let P_X be the marginal distribution for $\boldsymbol{X}$. Also let $\boldsymbol{b}^{**} = \boldsymbol{b}^*(P_{X|y})$. Apply Theorem 1, with $P_{X|y}$ and P_X in place of F and G, respectively, to obtain

$$0 \le E\left[\log \frac{\boldsymbol{b}^{**t}\boldsymbol{X}}{\boldsymbol{b}^{*t}\boldsymbol{X}} \middle| Y = y \right] \le D(P_{X|y} \| P_X). \quad (23)$$

Averaging with respect to the distribution of Y yields

$$0 \le \Delta \le I. \quad (24)$$

Remark: An alternative proof of this theorem, based on money ratio tests and Stein's lemma, appears in [7].

VI. Sequential Portfolio Estimation

Here we show that a good sequence of estimates of the true market distribution leads to asymptotically optimal growth rate of wealth. First we generalize Theorem 1 to handle the sequential setting.

Let $X_1, X_2, \cdots, X_n$ be a sequence of random stock vectors with joint probability distribution P^n. The log-optimal sequential strategy uses the portfolio $b_i^* = b^*(P_{X_i|x_1, x_2, \cdots, x_{i-1}})$, which maximizes the conditional expected value of $\log b^t X_i$ given that $X_1 = x_1, \cdots, X_{i-1} = x_{i-1}$. Suppose that instead of b^*, we use portfolios $\hat{b}_i = b^*(Q_{X_i|x_1, \cdots, x_{i-1}})$ which are optimal for an incorrect distribution Q^n for the sequence $X_1, \cdots, X_n$.

Let $P_{X_i|X_1, \cdots, X_{i-1}}$ and $Q_{X_i|X_1, \cdots, X_{i-1}}$ be the regular conditional distributions associated with P^n and Q^n, respectively. We compare the resulting wealth

$$\hat{S}_n = \prod_{i=1}^{n} \hat{b}_i^T X_i \tag{25}$$

with the wealth

$$S_n^* = \prod_{i=1}^{n} b_i^{*T} X_i. \tag{26}$$

Theorem 3:

$$0 \le E \log \frac{S_n^*}{\hat{S}_n} \le D(P^n \| Q^n). \tag{27}$$

Proof: Application of Theorem 1 shows that

$$0 \le E\left[\left(\log \frac{b_i^{*t} X_i}{\hat{b}_i^t X_i}\right) \middle| X_1, \cdots, X_{i-1}\right] \le D\left(P_{X_i|X^{i-1}} \| Q_{X_i|X^{i-1}}\right). \tag{28}$$

Averaging with respect to the distribution of $X^{i-1} = (X_1, \cdots, X_{i-1})$ and then summing for $i = 1, 2, \cdots, n$ yields

$$\begin{aligned} 0 &\le \sum_{i=1}^{n} E\left(\log \frac{b_i^{*t} X_i}{\hat{b}_i^t X_i}\right) \\ &\le \sum_{i=1}^{n} ED\left(P_{X_i|X^{i-1}} \| Q_{X_i|X^{i-1}}\right) \\ &= D(P^n \| Q^n) \end{aligned} \tag{29}$$

by the chain rule, completing the proof.

Suppose $X_1, X_2, \cdots$ are independent with unknown density $p(x)$. Clearly, the optimal portfolio b^* does not depend on the time i or on the past. However, if $p(x)$ is unknown, a series of estimators of the distribution $\hat{P}_i(\cdot)$ corresponding to density estimators $\hat{p}_i(x)$ based on the past X^{i-1} may be used to obtain asymptotically optimal portfolios $\hat{b}_i = b^*(\hat{P}_i)$. It is often the case (see Barron [11], [12]) that there exists a sequence of estimators $\hat{P}_n$ converging to P in the sense that $ED(P \| \hat{P}_n) \to 0$, at least in the Cesaro sense, i.e.,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} ED(P \| \hat{P}_i) = 0. \tag{30}$$

In this case Theorem 3 applies with $Q_{X_i|X^{i-1}}$ given by the estimator $\hat{P}_i(\cdot)$ to yield

$$\lim E \frac{1}{n} \log \frac{S_n^*}{\hat{S}_n} = 0. \tag{31}$$

It follows that the actual wealth $\hat{S}_n$ is close to the log-optimal wealth S_n^* as shown in the following theorem.

Theorem 4: Let $X_1, X_2, \cdots$ be i.i.d. $\sim P$. Let $\hat{P}_n$ be a sequence of estimators of the true distribution P such that

$$\frac{1}{n} \sum_{i=1}^{n} ED(P \| \hat{P}_i) \to 0 \tag{32}$$

and let

$$\hat{S}_n = \prod_{i=1}^{n} \hat{b}_i^t X_i \tag{33}$$

where

$$\hat{b}_i = b^*(\hat{P}_i). \tag{34}$$

Let

$$S_n^* = \prod_{i=1}^{n} b^{*t}(P) X_i \tag{35}$$

be the optimal wealth sequence. Then

$$\hat{S}_n = S_n^* 2^{n\, o(1)}, \tag{36}$$

where $o(1) \to 0$ in probability.

Consequently, if S_n^* has an exponential growth rate W^*, then $\hat{S}_n$ has the same asymptotic exponent.

Proof: To see that $\hat{S}_n / S_n^* = 2^{n\, o(1)}$ in probability, first observe that by Markov's inequality

$$P\left\{\frac{\hat{S}_n}{S_n^*} > 2^{n\epsilon}\right\} \le 2^{-n\epsilon} E \frac{\hat{S}_n}{S_n^*} \le 2^{-n\epsilon} \tag{37}$$

where the inequality $E(\hat{S}_n / S_n^*) \le 1$ follows from the Kuhn–Tucker conditions for the optimality of b^* (see Bell and Cover [8]).

On the other hand, using the notation $y^+ = \max\{0, y\}$, $y^- = \max\{0, -y\}$,

$$\begin{aligned} P\left\{\frac{S_n^*}{\hat{S}_n} > 2^{n\epsilon}\right\} &= P\left\{\log \frac{S_n^*}{\hat{S}_n} > n\epsilon\right\} \\ &\le \frac{1}{n\epsilon} E\left(\log \frac{S_n^*}{\hat{S}_n}\right)^+ \\ &\le \frac{1}{\epsilon}\left[\frac{1}{n} E \log \frac{S_n^*}{\hat{S}_n} + \frac{1}{n}\right] \end{aligned} \tag{38}$$

where the first inequality follows from Markov's inequality and the second from

$$\begin{aligned} E\left(\log S_n^* / \hat{S}_n\right)^- &= E \log \max\left\{\hat{S}_n / S_n^*, 1\right\} \\ &\le E \log\left(1 + \hat{S}_n / S_n^*\right) \\ &\le \log\left(1 + E\left(\hat{S}_n / S_n^*\right)\right) \le \log 2 = 1 \end{aligned} \tag{39}$$

by the concavity of the logarithm and the Kuhn–Tucker conditions. Combining (31), (37) and (38), we have $\hat{S}_n / S_n^* = 2^{n\, o(1)}$ in probability, as claimed.

VII. Examples

We first give an example due to Kelly [6] in which $\Delta = I$. Here the stock market is a horse race, which, in the setup of (1), consists of a probability mass function $P\{X = O_i e_i\} = p_i$, $i = 1, 2, \cdots, m$, where e_i is a unit vector with a 1 in the ith place and 0's elsewhere, O_i equals the win odds (O_i for 1), and p_i is the probability that the ith horse wins the race.

Then

$$\begin{aligned} W(X) &= \max_{b} E \log b^t X \\ &= \max_{b} \sum_{i=1}^{m} p_i \log b_i O_i \\ &= \sum p_i \log O_i - H(X) \end{aligned} \tag{40}$$

where $H(X) = -\sum_{i=1}^{m} p_i \log p_i$. Also $b^* = p$, i.e., the optimal portfolio is to bet in proportion to the win probabilities, regardless of the odds.

For side information Y, where (X, Y) has a given distribution, a similar calculation yields

$$W(X|Y) = \sum p_i \log O_i - H(X|Y) \tag{41}$$

and

$$b_i^* = P(X = O_i e_i | y), \qquad i = 1, 2, \cdots, m.$$

Here the optimal portfolio is to bet in proportion to the conditional probabilities, given Y. Subtracting (40) from (41), we have

$$\Delta = W(X|Y) - W(X) = H(X) - H(X|Y) = I(X;Y). \tag{42}$$

Consequently, the information bound on Δ is tight.

Of course, it sometimes happens that the information Y about the market is useless for investment purposes. The next example has $\Delta = 0$, $I = 1$. Let $X = (1, 1/2)$ with probability $1/2$, and $X = (1, 3/4)$ with probability $1/2$. Let $Y = X$. An investment in the first stock always returns the investment, but an investment in the second stock may cut the investment capital to either $1/2$ or $3/4$ depending on the outcome X. It would be foolish to invest in the second stock, since the first stock dominates its performance. Thus $b^* = b^*(y) = (1, 0)$ for all y, and $\Delta = 0$. On the other hand, since the outcomes of X are equally likely, and $Y = X$, we see

$$\begin{aligned} I(X;Y) &= I(X;X) = H(X) - H(X|X) \\ &= H(X) = 1 \quad \text{bit.} \end{aligned} \tag{43}$$

Thus a bit of information is available, but $\Delta = 0$ and the growth rate is not improved.

VIII. Conclusion

We offer one final interpretation. Recall that $H(X) - H(X|Y) = I(X;Y)$ is the decrement in the expected description length of X due to the side information Y. Hence the inequality $\Delta \le I$ has the interpretation that the increment in the doubling rate of the market X is less than the decrement in the description rate of X.

IX. Acknowledgment

We would like to thank R. O. Duda for speculations that led to the statement of Theorem 2.

References

[1] L. Breiman, "Optimal gambling systems for favorable games," in *Proc. 4th Berkeley Symp.*, vol. 1, pp. 65–78, 1961.
[2] H. Chernoff, "Large-sample theory: Parametric case," *Ann. Math. Stat.*, pp. 1–22, 1956.
[3] R. Bell and T. Cover, "Competitive optimality of logarithmic investment," *Math. Operations Res.*, vol. 5, no. 2, pp. 161–166, 1980.
[4] T. Cover, "An algorithm for maximizing expected log investment return," *IEEE Trans. Inform. Theory*, vol. IT-30, no. 2, pp. 369–373, 1984.
[5] M. Finkelstein and R. Whitley, "Optimal strategies for repeated games," *Adv. Appl. Prob.*, vol. 13, pp. 415–428, 1981.
[6] J. Kelly, "New interpretation of information rate," *Bell Syst. Tech. J.*, vol. 35, pp. 917–926, 1956.
[7] T. Cover, "A bound on the monetary value of information," Stanford Statist. Dept. Tech. Rep. 52, 1984.
[8] R. Bell and T. Cover, "Game theoretic optimal portfolios," *Management Science*, vol. 34, no. 6, pp. 724–733, 1988.
[9] P. Algoet and T. Cover, "Asymptotic optimality and asymptotic equipartition properties of log-optimum investment," *Ann. Prob.*, vol. 16, no. 2, pp. 876–898, 1988.
[10] T. Cover and D. Gluss, "Empirical Bayes stock market portfolios," *Adv. Appl. Math.*, vol. 7, pp. 170–181, 1986.
[11] A. Barron, "Are Bayes rules consistent in information?" in *Open Problems in Communication and Computation*, T. Cover and B. Gopinath, Eds. New York: Springer-Verlag, 1987, pp. 85–91.
[12] A. Barron, "The exponential convergence of posterior probabilities with implications for Bayes estimators of density functions," submitted to *Ann. Statist.*, 1987.
[13] T. Móri, "*I*-divergence geometry of distributions and stochastic gains," in *Proc. 3rd Pannonian Symp. on Math. Stat.*, J. Mogyoródi, I. Vincze, W. Wertz, Eds., Visegrád, Hungary, 1982, pp. 231–238.

Code Construction for the Noiseless Binary Switching Multiple-Access Channel

PETER VANROOSE

Abstract —**The noiseless coding problem is considered for a recently introduced discrete memoryless multiple-access channel that is a counterpart to the well-known binary adder channel. Upper and lower bounds on the number of codewords in a uniquely decodable code pair are given, from which the zero-error capacity region of this channel is derived. This region coincides with the classical capacity region of this channel. The proof uses the new notion of second-order distance of a code. For several values of n and k, good code pairs of block length n are constructed with the first code being $[n,k]$-linear. Some of these are found to be optimal. Furthermore, some convolutional codes are investigated that yield additional good rate pairs.**

I. The Binary Switching Multiple-Access Channel

Consider the classical two-access communication situation of Fig. 1, where two separate senders attempt to communicate to a third user, the receiver. The channel accepts two binary input streams, transmitted at the same symbol rate and divided into blocks of the same length n (assuming bit and block synchronism), and outputs a ternary stream according to the bitwise *deterministic* transitions depicted in Fig. 2: $y = x_1/x_2$, where division by 0 gives ∞. We call this channel the *binary switching multiple-access channel* (BS-MAC). Note that this model is completely noiseless. We will consider the problem where the decoder has to reconstruct the two messages *error-free*. This of course restricts the information rate of the input streams.

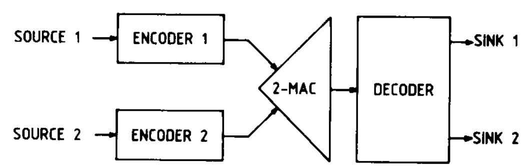

Fig. 1. Two-access communication system.

The only other deterministic binary two-input multiple-access channel (2-MAC) with ternary output is the *binary adder channel* (BAC), where the channel operation is $y = x_1 + x_2$. Noiseless coding for this channel has been studied by various authors (e.g. [1]–[3]). Surprisingly, the channels differ in many ways which will become clear in this correspondence.

Its relationship to the BAC was the first reason for studying the BS-MAC. Because both channels have the same input and

Manuscript received August 19, 1987; revised January 25, 1988. This work was supported in part by the Belgian Program for the Reinforcement of the Scientific Potential in the New Technologies and in part by the Ministry of Science Policy, Belgium, under Project GOA 83/88-53. This material was partially presented at the IEEE International Symposium on Information Theory, Ann Arbor, MI, October 1986, at the 7th Symposium on Information Theory in the Benelux, Deventer, The Netherlands, May 1987, and at the 3rd Joint Soviet-Swedish International Workshop on Information Theory, Sochi, USSR, May 1987.

The author is with the Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200 B, B-3030 Heverlee, Belgium.

IEEE Log Number 8824064.

0018-9448/88/0900-1100$01.00 ©1988 IEEE