variance," Ph.D. dissertation in engineering, University of California, Los Angeles, 1975.
[5] ——, "Consistent estimates of parameters in continuous time systems," presented at the IMA Conf. on Analysis and Optimization of Stochastic Systems, Oxford, England, Sept. 1978. (Also, Memo. NR 225, Dept. of Applied Math., Twente Univ. of Tech., Enschede, The Netherlands, Aug. 1978.)
[6] K. W. Iliff and R. E. Maine, "Observations on maximum likelihood estimation of aerodynamic characteristics from flight data," *J. Guidance and Control*, vol. 2, pp. 228–234, May–June 1979.
[7] E. Tse and J. J. Anton, "On identifiability of parameters," *IEEE Trans. Automat. Cont.*, vol. AC-17, pp. 637–646, Oct. 1972.
[8] K. J. Åström and P. Eykoff, "System identification—A survey," *Automatica*, vol. 7, pp. 123–162, 1971.
[9] R. K. Mehra and D. G. Lainiotis, Eds., *System Identification: Advances and Case Studies*. New York: Academic, 1976.
[10] G. C. Goodwin and R. L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*. New York: Academic, 1977.
[11] Special Issue on System Identification and Time-Series Analysis, *IEEE Trans. Automat. Contr.*, vol. AC-19, Dec. 1974.
[12] R. B. Ash, *Real Analysis and Probability*. New York: Academic, 1972.
[13] R. Bellman, *Introduction to Matrix Analysis*. New York: McGraw-Hill, 1960.
[14] J. L. Doob, *Stochastic Processes*. New York: Wiley, 1953.
[15] E. Wong, *Stochastic Processes in Information and Dynamical Systems*. New York: McGraw-Hill, 1971.
[16] S. Kullback, *Information Theory and Statistics*. New York: Wiley, 1959.
[17] H. Cramér, *Mathematical Methods of Statistics*. Princeton, NJ: Princeton Univ., 1946.
[18] A. Wald, "Note on the consistency of the maximum likelihood estimate," *Ann. Math. Statist.*, vol. 20, pp. 595–601, 1949.
[19] J. Wolfowitz, "On Wald's proof of the consistency of the maximum likelihood estimates," *Ann. Math. Statist.*, vol. 20, pp. 601–602, 1949.
[20] Y. Baram and N. R. Sandell, Jr., "An information theoretic approach to dynamical systems modeling and identification," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 61–66, Feb. 1978.
[21] Y. Baram, "A sufficient condition for consistent discrimination between stationary Gaussian models," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 958–960, Oct. 1978.
[22] R. M. Hawkes and J. B. Moore, "Performance bounds for adaptive estimation," *Proc. IEEE*, vol. 64, pp. 1143–1150, Aug. 1976.
[23] B. D. O. Anderson, J. B. Moore, and R. M. Hawkes, "Model approximations via prediction error identification," *Automatica*, vol. 14, pp. 615–622, Nov. 1978.
[24] L. Ljung, "On consistency of prediction error identification methods," in *System Identification: Advances and Case Studies*, R. K. Mehra and D. G. Lainiotis, Eds. New York: Academic, 1976.
[25] P. E. Caines, "Prediction error identification methods for stationary stochastic processes," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 500–505, Aug. 1976.
[26] J. Rissanen and P. E. Caines, "The strong consistency of maximum likelihood estimators for ARMA processes," *Ann. Math. Statist.*, vol. 7, pp. 297–315, March 1979.

The Capacity of the Semideterministic Relay Channel

ABBAS EL GAMAL AND MOHAMMAD AREF

***Abstract*—The capacity of the class of relay channels with sender x_1, a relay sender x_2, a relay receiver $y_1 = f(x_1, x_2)$, and ultimate receiver y is proved to be**

$$C = \max_{p(x_1, x_2)} \min \{ I(X_1, X_2; Y), H(Y_1 | X_2) + I(X_1; Y | X_2, Y_1) \}.$$

The relay channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1 | x_1, x_2), \mathcal{Y} \times \mathcal{Y}_1)$ is a model for the communication between a sender x_1 and a receiver y through two paths; a direct path from x_1 to y, and a path from x_1 to y through the relay (y_1, x_2). The relationship among the received symbols y and y_1 and the transmitted ones x_1 and x_2 is given by the probability transition matrix $p(y, y_1 | x_1, x_2)$.

The problem of finding the capacity of the relay channel (i.e., maximum rate of transmission from x_1 to y) was first studied by van der Muelen [1]. In [2] Cover and El Gamal established the capacity when the relay channel is *degraded*, *reversely degraded*, and when *feedback* is added from the receivers y and y_1 to both senders x_1 and x_2. A general upper bound to capacity and an achievable rate (lower bound) were also established in [2].

In this note, we show that a special case of the lower bound to the capacity given in [2, theorem 7] is in fact the capacity when the relay receiver y_1 is a deterministic function of x_1 and x_2. First, recall the following version of [2, theorem 7].

Theorem 7 [2]: For any relay channel, the following rate R^* is achievable:

$$R^* = \sup \{ \min \{ I(U; Y_1 | X_2, V) + I(X_1; Y, \hat{Y}_1 | X_2, U), I(V; Y) + I(U; Y | X_2, V) + I(X_1; Y, \hat{Y}_1 | X_2, U) \} \}, \quad (1)$$

where the supremum is taken over all joint probability mass functions of the form

$$p(u, v, x_1, x_2, y, y_1, \hat{y}_1) = p(v)(u | v) p(x_1 | u) p(x_2 | v) \cdot p(y, y_1 | x_1, x_2) p(\hat{y}_1 | x_2, y_1, u), \quad (2)$$

subject to the constraint

$$I(\hat{Y}_1; Y_1 | Y, X_2, U) \le I(X_2; Y | V). \quad (3)$$

Now, substituting $\hat{Y}_1 \equiv \varnothing$, $V \equiv X_2$, and $U \equiv (X_2, Q)$ in (1)–(3), we obtain the following theorem.

Theorem: For any relay channel, the following rate R_0 is achievable:

$$R_0 = \sup \min \{ I(X_1, X_2; Y), I(Q; Y_1 | X_2) + I(X_1; Y | X_2, Q) \}, \quad (4)$$

where the supremum is over all probability mass functions of the form

$$p(q, x_1, x_2, y, y_1) = p(q, x_1, x_2) p(y, y_1 | x_1, x_2), \quad (5)$$

and Q is an arbitrary random variable.

Corollary: If y_1 is a deterministic function of x_1 and x_2, then

$$C = R_0 = \max_{p(x_1, x_2)} \min \{ I(X_1, X_2; Y), H(Y_1 | X_2) + I(X_1; Y | X_2, Y_1) \} \quad (6)$$

Proof: That (6) is a special case of R_0 can be easily seen by setting $Q \equiv Y_1$. The converse part of the corollary follows easily from the outer bound of [2, theorem 4], and the fact that if $y_1 = f(x_1, x_2)$, then $I(X_1; Y, Y_1 | X_2) = H(Y_1 | X_2) + I(X_1; Y | X_2, Y_1)$, and the corollary is proved.

Remarks: If both y and y_1 are deterministic functions of (x_1, x_2), i.e., if the relay channel is deterministic, then it follows from the corollary that the capacity is given by

$$C = \max_{p(x_1, x_2)} \min \{ H(Y), H(Y, Y_1 | X_2) \}.$$

Manuscript received September 5, 1979; revised June 16, 1981. This work was supported in part by the Joint Services Electronics Program through the Air Force Office of the Scientific Research (AFSC) under Contract F44620-76-C-0061, and in part by the National Science Foundation under Grant ENG 78-23334. This paper was partially presented at the International Symposium on Information Theory, Grignano, Italy, 1979.

Abbas El Gamal is with Stanford University, Stanford, CA 94305.
Mohammad Aref was with Stanford University, Stanford, CA 94305.

Acknowledgment

The author is indebted to the referees for their invaluable remarks during the preparation of this paper.

References

[1] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probl.*, vol. 3, pp. 120–154, 1971.
[2] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 572–584, Sept. 1979.

0018-9448/82/0500-0536$00.75 ©1982 IEEE