The Capacity Region of a Product Gaussian Broadcast Channel with Degraded Message Sets

Jae Won Yoo and Tie Liu
Dept. of Electrical & Computer Engineering
Texas A&M University
College Station, TX 77843
E-mail: {yoojw78,tieliu}@tamu.edu

Yingbin Liang
Dept. of Electrical Engineering & Computer Science
Syracuse University
Syracuse, NY 13244
E-mail: yliang06@syr.edu

Abstract—**A product Gaussian broadcast channel with three receivers and degraded message sets is considered. A single-letter characterization of the capacity region was obtained in a recent work of Nair and El Gamal based on superposition coding and indirect decoding. Through Lagrangian theory and an extremal entropy inequality, the single-letter expression of Nair and El Gamal is evaluated for the specific product Gaussian model, which leads to an explicit characterization of the capacity region.**

I. INTRODUCTION

Broadcast is a fundamental nature of wireless communication: any receiver within the transmission range can listen to the source and potentially decode some of the messages. With appropriate coding architecture, the broadcast nature of wireless communication can be used to the advantage of simultaneously transmitting to several receivers at high rates. Understanding the limits and the appropriate coding architectures that can harness the broadcast advantage of wireless communication is an important subject of network information theory [1].

Most of the previous work focused on one of the following two scenarios:

1) to deliver the *same* messages to each of the receivers, usually known as the *multicast* problem; and
2) to deliver completely *distinct* messages to different receivers, namely the *private message* problem.

Formally, the distinction between these two broadcast scenarios can be identified by the configurations of the message sets associated with each of the receivers. For the multicast problem, the intended message sets for each of the receivers are *identical*. For the private message problem, the intended message sets for each of the receivers are *mutually exclusive*. Clearly, the appropriate coding architecture depends on the configurations of the message sets.

Between these two "extreme" broadcast scenarios, the multicast and the private message problems, there is a rich collection of "intermediate" problems with message sets of interesting configurations and significant engineering appeal. A good example is the *degraded message set* problems first considered in [2], which can be used to model broadcast scenarios with a progressively encoded source and receivers of different quality-of-service requirement.
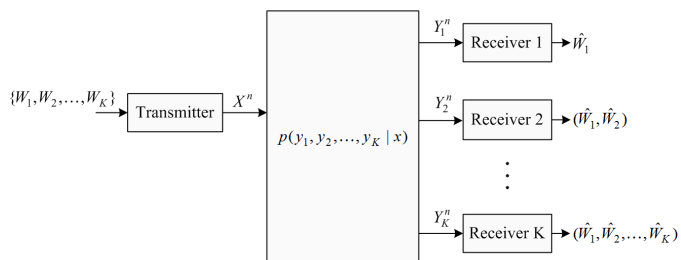

Fig. 1.   Broadcast channel with degraded message sets.

Fig. 1 illustrates a general discrete memoryless broadcast channel with degraded message sets. The transmitter has a total of K independent messages $(M_1, M_2, \ldots, M_K)$. Each of the K receivers demands a subset of messages from the transmitter. The message set $\mathcal{S}_k$ intended for receiver k is given by

$$\mathcal{S}_k = \{M_1, M_2, \ldots, M_k\}, \quad k = 1, 2, \ldots, K.$$

Clearly, we have

$$\mathcal{S}_1 \subseteq \mathcal{S}_2 \subseteq \cdots \subseteq \mathcal{S}_K$$

and hence the name "degraded message sets".

For the degraded message set problem, there is a natural communication strategy based on superposition coding [3] and direct decoding. With K independent messages at the transmitter and K receivers, an K-layer superposition code can be built with the kth layer from the bottom representing message M_k. Receiver k decodes messages $(W_1, W_2, \ldots, W_k)$ by *directly* decoding all the bottom layers up to the kth. For $K = 2$, it was shown in [2] that this natural strategy is also optimal in achieving the capacity region of the channel. For $K \geq 3$, however, finding the capacity region of the discrete memoryless broadcast channel with degraded message sets remains an open problem in network information theory.

In an excellent contribution [4], Nair and El Gamal considered a special three-receiver discrete memoryless broadcast channel with degraded message sets and presented a precise single-letter characterization of the capacity region. Specifically, in [4], it was assumed that:

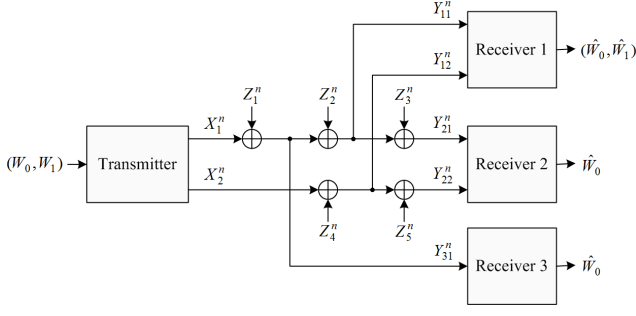1) receiver 2 is degraded with respect to receiver 1, i.e., $X - Y_1 - Y_2$ forms a Markov for any input distribution

Fig. 2. Product Gaussian broadcast channel with degraded message sets.

    $p(x)$; and

2) the rate of message M_2 is set to be zero so in defacto, there are only two independent messages M_1 and M_3 at the transmitter.

Under these two assumptions, Nair and El Gamal [4] proved a surprising result that the natural scheme that uses direct decoding is, in general, *suboptimal*. Instead, a coding scheme that uses *indirect decoding* [4] can always achieve the capacity region of the channel.

    Building on the result of [4], in this paper, we consider a specific product Gaussian broadcast channel with degraded message sets and provide an *explicit* characterization of the capacity region. The main tools used in this characterization are Lagrangian theory [5] and an extremal entropy inequality of Liu and Viswanath [6]. It is worth mentioning that the exact same product Gaussian model was also considered in the original work of Nair and El Gamal [4], and characterizing the capacity region was posted as an open problem.

II. CHANNEL MODEL

    As shown in Fig. 2, consider a discrete-time memoryless product Gaussian broadcast channel with three receivers. At each time sample, the received signals at receivers 1, 2 and 3 are given by $Y_1 = (Y_{11}, Y_{12})$, $Y_2 = (Y_{21}, Y_{22})$ and $Y_3 = Y_{31}$, respectively, where

$$Y_{31} = X_1 + Z_1, \quad Y_{11} = Y_{31} + Z_2, \quad Y_{21} = Y_{11} + Z_3 \atop Y_{12} = X_2 + Z_4, \quad Y_{22} = Y_{12} + Z_5. \qquad (1)$$

Here, $X = (X_1, X_2)$ is the channel input, and Z_i, $i = 1, 2, 3, 4, 5$, are Gaussian noise with zero means with co-variance N_i, respectively, and are assumed to be mutually independent of each other. We consider two different types of power constraints on the channel input X: an average total power constraint

$$E[X_1^2 + X_2^2] \leq P \qquad (2)$$

and an individual per-subchannel power constraint

$$E[X_i^2] \leq P_i, \quad i = 1, 2. \qquad (3)$$

    The transmitter has two independent messages M_0 and M_1, where M_0 is a common message intended for all three receivers and M_1 is a private message intended only for receiver 1. The capacity region $\mathcal{C}(P)$ is given by the set of

nonnegative rate pairs (R_0, R_1) that can be achieved by any coding scheme under the average total power constraint (2). Likewise, the capacity region $\mathcal{C}(P_1, P_2)$ is given by the set of nonnegative rate pairs (R_0, R_1) that can be achieved by any coding scheme under the individual per-subchannel power constraint (3).

    From the channel model (1), it is clear that $X - Y_1 - Y_2$ forms a Markov for any distribution on the channel input X. In this case, a single-letter characterization of the capacity region was obtained in [4, Prop. 2] and is given by the set of nonnegative rate tuples (R_0, R_1) such that

$$
\begin{aligned}
R_0 &\leq I(U_1; Y_{21}) + I(U_2; Y_{22}) \\
R_0 &\leq I(V_1; Y_{31}) \\
R_1 &\leq I(X_1; Y_{11}|U_1) + I(X_2; Y_{12}|U_2) \\
R_0 + R_1 &\leq I(V_1; Y_{31}) + I(X_1; Y_{11}|V_1) + I(X_2; Y_{12}|U_2)
\end{aligned}
\qquad (4)
$$

for some joint distributions on (U_1, V_1, X_1) and (U_2, X_2) such that $U_1 - V_1 - X_1$ forms a Markov chain. The main goal of this paper is to evaluate the rate region (4) for the specific product Gaussian model (1) under both average total and individual per-subchannel power constraints.

III. MAIN RESULTS

    The main result of the paper is an explicit characterization of the capacity region of the product Gaussian broadcast channel (1) under the individual per-subchannel power constraint (3), summarized in the following theorem.

    Theorem 1: The capacity region $\mathcal{C}(P_1, P_2)$ of the three-receiver product Gaussian broadcast channel (1) under the individual per-subchannel power constraint (3) is given by the set of nonnegative rate tuple (R_0, R_1) such that

$$
\begin{aligned}
R_0 &\leq C\left(\frac{P_1 - Q_1}{Q_1 + N_1 + N_2 + N_3}\right) + C\left(\frac{P_2 - Q_2}{Q_2 + N_4 + N_5}\right) \\
R_0 &\leq C\left(\frac{P_1}{N_1}\right) \\
R_1 &\leq C\left(\frac{Q_1}{N_1 + N_2}\right) + C\left(\frac{Q_2}{N_4}\right) \\
R_0 + R_1 &\leq C\left(\frac{P_1}{N_1}\right) + C\left(\frac{Q_2}{N_4}\right)
\end{aligned}
\qquad (5)
$$

for some $0 \leq Q_1 \leq P_1$ and $0 \leq Q_2 \leq P_2$, where $C(x) := \frac{1}{2}\log(1 + x)$.

    As a corollary, we have the following characterization of the capacity region of the product Gaussian broadcast channel (1) under the average total power constraint (2).

    Corollary 1: The capacity region $\mathcal{C}(P)$ of the three-receiver product Gaussian broadcast channel (1) under the average total power constraint (2) is given by the set of nonnegative rate tuple (R_0, R_1) such that

$$
\begin{aligned}
R_0 &\leq C\left(\frac{Q_3}{Q_1 + N_1 + N_2 + N_3}\right) + C\left(\frac{Q_4}{Q_2 + N_4 + N_5}\right) \\
R_0 &\leq C\left(\frac{Q_1 + Q_3}{N_1}\right) \\
R_1 &\leq C\left(\frac{Q_1}{N_1 + N_2}\right) + C\left(\frac{Q_2}{N_4}\right) \\
R_0 + R_1 &\leq C\left(\frac{Q_1 + Q_3}{N_1}\right) + C\left(\frac{Q_2}{N_4}\right)
\end{aligned}
\qquad (6)
$$

for some $Q_i \geq 0$, $i = 1, 2, 3, 4$, and $Q_1 + Q_2 + Q_3 + Q_4 \leq P$.

Proof: This is a simple consequence of Theorem 1 and the well-known fact that

$$\mathcal{C}(P) = \bigcup_{P_1+P_2 \leq P} \mathcal{C}(P_1, P_2).$$

■

The rest of the paper is devoted to the proof of Theorem 1.

IV. PROOF OF THEOREM 1

The achievability of the rate region (5) follows from that of (4) by setting $X_i = U_i + W_i$ for $i = 1, 2$ and $V_1 = X_1$, where U_i and W_i are two independent Gaussian variables with zero means and variances $P_i - Q_i$ and Q_i, respectively. (Note that for such a choice of (U_1, V_1, X_1), $U_1 - V_1 - X_1$ forms a trivial Markov chain.) We therefore concentrate on proving the converse part of the theorem.

To prove the converse part of the theorem, we shall need the following extremal entropy inequality which first appeared in [6, Th. 8].

Lemma 1 ([6]): Let P and μ be two nonnegative real numbers, and let Z_1, Z_2 be two Gaussian variables with zero means and variances N_1 and N_2, respectively. Assume that $0 < N_1 \leq N_2$. If there exists a nonnegative real number P^* satisfying

$$(P^* + N_1)^{-1} + M_1 = \mu(P^* + N_2)^{-1} + M_2$$
$$M_1 P^* = 0$$
$$M_2(P - P^*) = 0$$

for some nonnegative real numbers M_1 and M_2, then

$$h(X + Z_1|U) - \mu h(X + Z_2|U)$$
$$\leq \frac{1}{2}\log 2\pi e(P^* + N_1) - \frac{\mu}{2}\log 2\pi e(P^* + N_2)$$

for any (X, U) independent of (Z_1, Z_2) and such that $E[X^2] \leq P$.

We are now ready to prove the converse part of the theorem. Consider proof by contradiction. Let (R_0^o, R_1^o) be an *achievable* rate pair that lies *outside* the rate region (5). From [7], we have $R_0^o \leq R_0^{max}$ where

$$R_0^{max} := \min\left\{C\left(\frac{P_1}{N_1+N_2+N_3}\right) + C\left(\frac{P_2}{N_4+N_5}\right), C\left(\frac{P_1}{N_1}\right)\right\}.$$

Note that when $R_1^o = 0$, R_0^{max} can be achieved by letting $Q_1 = Q_2 = 0$ in (5). Thus, we may assume that $R_1^o > 0$ and write $R_1^o = R_1^* + \delta$ for some $\delta > 0$, where R_1^* is given by

$$
\begin{aligned}
\max \quad & R_1 \\
\text{s.t.} \quad R_0^o &\leq C\left(\frac{P_1-Q_1}{Q_1+N_1+N_2+N_3}\right) + C\left(\frac{P_2-Q_2}{Q_2+N_4+N_5}\right) \\
R_1 &\leq C\left(\frac{Q_1}{N_1+N_2}\right) + C\left(\frac{Q_2}{N_4}\right) \\
R_0^o + R_1 &\leq C\left(\frac{P_1}{N_1}\right) + C\left(\frac{Q_2}{N_4}\right) \\
Q_1 &\leq P_1 \\
Q_2 &\leq P_2 \\
-Q_1 &\leq 0 \\
-Q_2 &\leq 0.
\end{aligned}
$$

Let (R_1^*, Q_1^*, R_2^*) be an optimal solution to the above optimization problem. Then, (R_1^*, Q_1^*, R_2^*) must satisfy the Karush-Kuhn-Tucker (KKT) conditions [5] as shown in the top of next page, where T_i, $i = 1, 2, 3, 4$, and M_i, $i = 1, 2, 3, 4$, are nonnegative Lagrangian multipliers. From the KKT conditions (9)–(12), we have

$$
\begin{aligned}
&(T_1 + T_3)R_0^o + R_1^o \\
&= (T_1 + T_3)R_0^o + R_1^* + \delta \\
&= (T_1 + T_3)R_0^o + (T_2 + T_3)R_1^* + \delta \\
&= T_1 R_0^o + T_2 R_1^* + T_3(R_0^o + R_1^*) + \delta \\
&= T_1\left[C\left(\frac{P_1 - Q_1^*}{Q_1^* + N_1 + N_2 + N_3}\right) + \right. \\
&\quad \left. C\left(\frac{P_2 - Q_2^*}{Q_2^* + N_4 + N_5}\right)\right] + T_2\left[C\left(\frac{Q_1^*}{N_1 + N_2}\right) + \right. \\
&\quad \left. C\left(\frac{Q_2^*}{N_4}\right)\right] + T_3\left[C\left(\frac{P_1}{N_1}\right) + C\left(\frac{Q_2^*}{N_4}\right)\right] + \delta \\
&= T_1\left[C\left(\frac{P_1 - Q_1^*}{Q_1^* + N_1 + N_2 + N_3}\right) + \right. \\
&\quad \left. C\left(\frac{P_2 - Q_2^*}{Q_2^* + N_4 + N_5}\right)\right] + T_2 C\left(\frac{Q_1^*}{N_1 + N_2}\right) + \\
&\quad T_3 C\left(\frac{P_1}{N_1}\right) + C\left(\frac{Q_2^*}{N_4}\right) + \delta.
\end{aligned}
\tag{17}
$$

On the other hand, by the KKT condition (9) and the assumption that (R_0^o, R_1^o) is achievable, we have

$$
\begin{aligned}
&(T_1 + T_3)R_0^o + R_1^o \\
&= (T_1 + T_3)R_0^o + R_1^o \\
&= (T_1 + T_3)R_0^o + (T_2 + T_3)R_1^o \\
&= T_1 R_0^o + T_2 R_1^* + T_3(R_0^o + R_1^o) \\
&\leq T_1\left[I(U_1; Y_{21}) + I(U_2; Y_{22})\right] + \\
&\quad T_2\left[I(X_1; Y_{11}|U_1) + I(X_2; Y_{12}|U_2)\right] + \\
&\quad T_3\left[I(V_1; Y_{31}) + I(X_1; Y_{11}|V_1) + I(X_2; Y_{12}|U_2)\right] \\
&= T_1 h(Y_{21}) + T_1 h(Y_{22}) + T_3 h(Y_{31}) - \\
&\quad [h(Y_{11}|X_1) + h(Y_{12}|X_2)] + \\
&\quad [T_2 h(Y_{11}|U_1) - T_1 h(Y_{21}|U_1)] + \\
&\quad [h(Y_{12}|U_2) - T_1 h(Y_{22}|U_2)] + \\
&\quad T_3[h(Y_{11}|V_1) - h(Y_{31}|V_1)]
\end{aligned}
\tag{18}
$$

for some joint distributions on (U_1, V_1, X_1) and (U_2, X_2) such that $U_1 - V_1 - X_1$ forms a Markov chain and $E[X_i^2] \leq P_i$ for $i = 1, 2$.

The terms on the right-hand side of the above equation can be further bounded/evaluated as follows.

1) It is well known [1] that Gaussian maximizes differential entropy for a given power, so we have

$$
\begin{aligned}
h(Y_{21}) &\leq \tfrac{1}{2}\log 2\pi e(P_1 + N_1 + N_2 + N_3) \\
h(Y_{22}) &\leq \tfrac{1}{2}\log 2\pi e(P_2 + N_4 + N_5) \\
h(Y_{31}) &\leq \tfrac{1}{2}\log 2\pi e(P_1 + N_1).
\end{aligned}
\tag{19}
$$

2) The channel inputs (X_1, X_2) are independent of the

$$T_2(Q_1^* + N_1 + N_2)^{-1} + M_1 = T_1(Q_1^* + N_1 + N_2 + N_3)^{-1} + M_2 \tag{7}$$

$$(Q_2^* + N_4)^{-1} + M_3 = T_1(Q_2^* + N_4 + N_5)^{-1} + M_4 \tag{8}$$

$$T_2 + T_3 = 1 \tag{9}$$

$$T_1 R_0^o = T_1 \left[C \left(\frac{P_1 - Q_1^*}{Q_1^* + N_1 + N_2 + N_3} \right) + C \left(\frac{P_2 - Q_2^*}{Q_2^* + N_4 + N_5} \right) \right] \tag{10}$$

$$T_2 R_1^* = T_2 \left[C \left(\frac{Q_1^*}{N_1 + N_2} \right) + C \left(\frac{Q_2^*}{N_4} \right) \right] \tag{11}$$

$$T_3(R_0^o + R_1^*) = T_3 \left[C \left(\frac{P_1}{N_1} \right) + C \left(\frac{Q_2^*}{N_4} \right) \right] \tag{12}$$

$$M_1 Q_1^* = 0 \tag{13}$$

$$M_2(P_1 - Q_1^*) = 0 \tag{14}$$

$$M_3 Q_2^* = 0 \tag{15}$$

$$M_4(P_2 - Q_2^*) = 0 \tag{16}$$

Gaussian noise $(Z_1, Z_2, Z_3, Z_4, Z_5)$, so we have

$$h(Y_{11}|X_1) = h(Z_1 + Z_2) = \tfrac{1}{2} \log 2\pi e (N_1 + N_2)$$
$$h(Y_{12}|X_1) = h(Z_4) = \tfrac{1}{2} \log 2\pi e N_4. \tag{20}$$

3) Putting together the KKT conditions (7), (13) and (14), we have

$$T_2(Q_1^* + N_1 + N_2)^{-1} + M_1 = T_1(Q_1^* + N_1 + N_2 + N_3)^{-1} + M_2$$
$$M_1 Q_1^* = 0$$
$$M_2(P_1 - Q_1^*) = 0$$

where M_1, M_2, T_1 and T_2 are nonnegative real numbers. By Lemma 1[1], we have

$$T_2 h(Y_{11}|U_1) - T_1 h(Y_{21}|U_1)$$
$$= T_2 h(X_1 + Z_1 + Z_2|U_1) - T_1 h(X_1 + Z_1 + Z_2 + Z_3|U_1)$$
$$\leq \frac{T_2}{2} \log 2\pi e (Q_1^* + N_1 + N_2) - \frac{T_1}{2} \log 2\pi e (Q_1^* + N_1 + N_2 + N_3). \tag{21}$$

4) Similarly, putting together the KKT conditions (8), (15) and (16), we have

$$(Q_2^* + N_4)^{-1} + M_3 = T_1(Q_2^* + N_4 + N_5)^{-1} + M_4$$
$$M_3 Q_2^* = 0$$
$$M_4(P_2 - Q_2^*) = 0$$

where M_3, M_4 and T_1 are nonnegative real numbers.

Again, by Lemma 1, we have

$$h(Y_{12}|U_2) - T_1 h(Y_{22}|U_2)$$
$$= h(X_2 + Z_4|U_2) - T_1 h(X_2 + Z_4 + Z_5|U_2)$$
$$\leq \frac{1}{2} \log 2\pi e (Q_2^* + N_4) - \frac{T_1}{2} \log 2\pi e (Q_2^* + N_4 + N_5). \tag{22}$$

5) Finally, note that

$$h(Y_{11}|V_1) - h(Y_{31}|V_1)$$
$$= h(X_1 + Z_1 + Z_2|V_1) - h(X_1 + Z_1|V_1)$$
$$= I(Z_2; X_1 + Z_1 + Z_2|V_1)$$
$$= h(Z_2) - h(Z_2|X_1 + Z_1 + Z_2, V_1) \tag{23}$$
$$\leq h(Z_2) - h(Z_2|X_1 + Z_1 + Z_2, V_1, X_1) \tag{24}$$
$$= h(Z_2) - h(Z_2|Z_1 + Z_2, V_1, X_1)$$
$$= h(Z_2) - h(Z_2|Z_1 + Z_2) \tag{25}$$
$$= I(Z_2; Z_1 + Z_2)$$
$$= \frac{1}{2} \log 2\pi e (N_1 + N_2) - \frac{1}{2} \log 2\pi e N_1 \tag{26}$$

where (23) is due to the independence of Z_2 and V_1; (24) is due to the fact that conditioning reduces differential entropy [1]; and (25) is due to the independence of (Z_1, Z_2) and (V_1, X_1).

Substitute (19)–(22) and (26) into (18). With some rearranging of terms, we may obtain

$$(T_1 + T_3)R_0^o + R_1^o$$
$$\leq T_1 \left[C \left(\frac{P_1 - Q_1^*}{Q_1^* + N_1 + N_2 + N_3} \right) + C \left(\frac{P_2 - Q_2^*}{Q_2^* + N_4 + N_5} \right) \right] + T_2 C \left(\frac{Q_1^*}{N_1 + N_2} \right) + T_3 C \left(\frac{P_1}{N_1} \right) + C \left(\frac{Q_2^*}{N_4} \right) \tag{27}$$

Note that $\delta > 0$, so this is a contradiction to (17). We therefore

---

[1]If $T_2 = 0$, we have either $T_1 = 0$ or $Q_1^* = 0$. In either case, inequality (21) holds trivially.

conclude that any achievable rate pair (R_0^o, R_1^o) must also be *inside* the rate region (5). This completes the proof of the converse part of the theorem.

ACKNOWLEDGMENT

REFERENCES

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

[2] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 60–64, Jan. 1977.

[3] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 2, pp. 197–207, Mar. 1973.

[4] C. Nair and A. El Gammal, "The capacity region of a class of three-receiver broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4479–4493, Oct. 2009.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[6] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.

[7] D. Blackwell, L. Breiman, and A. Thomasian, "The capacity of a class of channels," *Ann. Math. Statist.*, vol. 30, no. 4, pp. 1229–1241, Dec. 1959.