Gaussian Arbitrarily Varying Channels

BRIAN HUGHES, MEMBER, IEEE, AND PRAKASH NARAYAN, MEMBER, IEEE

Abstract—**The *arbitrarily varying channel* (AVC) can be interpreted as a model of a channel jammed by an intelligent and unpredictable adversary. We investigate the asymptotic reliability of optimal random block codes on *Gaussian* arbitrarily varying channels (GAVC's). A GAVC is a discrete-time memoryless Gaussian channel with input power constraint P_T and noise power N_e, which is further corrupted by an additive "jamming signal." The statistics of this signal are unknown and may be arbitrary, except that they are subject to a power constraint P_J. We distinguish between two types of power constraints: *peak* and *average*. For peak constraints on the input power and the jamming power we show that the GAVC has a random coding capacity. For the remaining cases in which either the transmitter or the jammer or both are subject to average power constraints, no capacities exist and only λ-capacities are found. The asymptotic error probability suffered by optimal random codes in these cases is determined. Our results suggest that if the jammer is subject only to an average power constraint, reliable communication is impossible at any positive code rate.**

I. Introduction

CONSIDER the following communications channel (cf., Fig. 1) which we call a *Gaussian arbitrarily varying channel* (GAVC). Once each second the transmitter chooses for transmission to the receiver an arbitrary real-valued random variable, say u_i^* at time i, such that the sequence $\{u_i^*\}$ satisfies a power constraint P_T (to be specified later). In transmission this number is corrupted in such a way that it is received as $u_i^* + \eta_{ei}^* + s_i^*$. The elements of the sequence $\{\eta_{ei}^*\}$ are independent zero-mean Gaussian random variables, each having variance N_e. The transmitter and the receiver have no knowledge of the random sequence $\{s_i^*\}$ other than that it satisfies a certain power constraint, say P_J (also to be specified later). The sequence $\{s_i^*\}$ may have arbitrary time-varying possibly non-Gaussian statistics. The goal of the transmitter and receiver is to construct a coding system to reliably convey discrete source data over this channel, knowing *only* N_e, P_T, and P_J.

The discrete memoryless arbitrarily varying channel (AVC) was introduced in a remarkable paper by Blackwell *et al.* [10] (see also Wolfowitz [21] and Csiszár and Körner [12]). We make no attempt to summarize the substantial body of literature on the discrete AVC; for this the reader is referred to [12, ch. 6].

By comparison, GAVC's have received considerably less attention. Blachman [7], [8] has obtained upper and lower bounds on the capacity of a GAVC (using the maximum probability of error concept) when the sequence $\{s_i^*\}$ is allowed to be chosen with foreknowledge of the transmitter's codeword. Ahlswede [1] has determined the capacity of the GAVC when $\{s_i^*\}$ consists of independent Gaussian random variables where the variance changes arbitrarily from one symbol to the next, within some positive range of values. Başar and Wu [4] have investigated the use of essentially the same channel for a different source transmission problem in which the source is a discrete-time memoryless Gaussian source and reliability is measured by mean-square distortion. Dobrushin [13] and later McEliece and Stark [17] have studied what might be called a *Gaussian compound channel* (cf. [12], [21]) that is similar to the GAVC except that $\{s_i^*\}$ is constrained to be a sequence of independent identically distributed (i.i.d.) random variables.

The practical significance of the GAVC is seen as follows. One may view the sequence $\{s_i^*\}$ as selected by an intelligent and unpredictable adversary, namely the *jammer*, whose intent is to disrupt the transmission of the sequence $\{u_i^*\}$ as much as possible. The jammer, like the transmitter, is subject to the natural constraint of finite power but is otherwise free to generate any signal he chooses.

In this paper we study four GAVC's corresponding to two different types of power constraints (peak and average) on the transmitted codeword and on the jamming sequence. Our main results are coding theorems, one for each pair of constraints, characterizing the *asymptotic reliability* that can be achieved by optimal random codes on these channels. We say asymptotic reliability rather than capacity because, as we shall find, these channels generally have no capacity *per se*.

The remainder of this paper is organized as follows. In Section II we introduce the terminology and summarize our results. These results are proved in Section III. Finally, in Section IV we discuss the implications of our results and, in particular, their application to certain jamming problems.

Manuscript received September 17, 1985; revised June 2, 1986. This work was supported in part by the Naval Research Laboratory and the Office of Naval Research under Grants N00014-83-G-0192 and N00014-84-G-0101, the National Science Foundation under Grant ECS-82-0444-9, and the Minta Martin Fund for Aerospace Research from the University of Maryland. This paper was presented in part at the 19th Annual Conference on Information Sciences and Systems, The Johns Hopkins University, March 27–29, 1985.

B. Hughes was with the Electrical Engineering Department of the University of Maryland, College Park, MD 20742. He is now with the Department of Electrical Engineering and Computer Science of The Johns Hopkins University, Baltimore, MD 21218, USA.

P. Narayan is with the Electrical Engineering Department of the University of Maryland, College Park, MD 20742, USA.

IEEE Log Number 8610798.

0018-9448/87/0300-0267$01.00 ©1987 IEEE

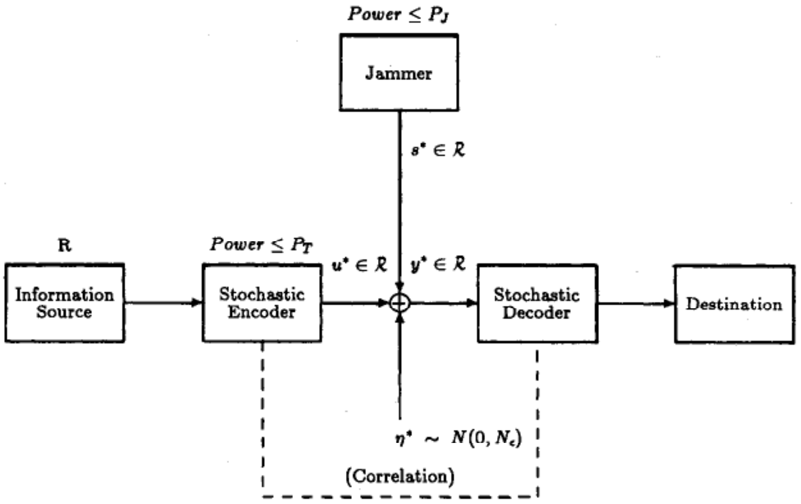

Fig. 1. Gaussian arbitrarily varying channel.

II. Definitions and Results

A *codeword* of length n for the GAVC is a sequence of n real numbers selected by the transmitter, say $\boldsymbol{u} = (u_1, \cdots, u_n)$. Similarly, a *jamming sequence* of length n, denoted by $\boldsymbol{s} = (s_1, \cdots, s_n)$ is a sequence of n real numbers selected by the jammer. These sequences may be thought of geometrically as points in n-dimensional Euclidean space ($\boldsymbol{R}^n$). With this interpretation the output of the GAVC corresponding to the codeword $\boldsymbol{u}$ and the jamming sequence $\boldsymbol{s}$ is

$$\boldsymbol{y}^* = \boldsymbol{u} + \boldsymbol{\eta}_e^* + \boldsymbol{s} \tag{2.1}$$

where $\boldsymbol{\eta}_e^*$ denotes an n-vector of i.i.d. $N(0, N_e)$ random variables.[1]

An (n, M) block code C_n is a system[2]

$$C_n = \{(\boldsymbol{u}_1, D_1), \cdots, (\boldsymbol{u}_M, D_M)\} \tag{2.2}$$

where $\{\boldsymbol{u}_i\}_{i=1}^M$ are codewords of length n and $\{D_i\}_{i=1}^M$ are disjoint (Borel) subsets of $\boldsymbol{R}^n$ called *decoding sets*.

We are interested in the problem of transmitting the output of a given information source generating R bits per second over the GAVC with minimum error probability (to be defined). The goal of the transmitter is to construct a block-coding system of length n that suffers an error probability no greater than this minimum, regardless of the jamming sequence $\boldsymbol{s}$. The goal of the jammer is to inflict the largest possible error probability on any code chosen by the transmitter by an appropriate choice of $\boldsymbol{s}$. For the transmitter a *strategy* to accomplish this goal consists of an $(n, 2^{nR})$ code; a strategy for the jammer is a jamming sequence of length n.

We allow both the transmitter and the jammer the additional flexibility of being able to choose their respective strategies *randomly*. Accordingly, we define an (n, M) *random (block) code*

$$C_n^* = \{(\boldsymbol{u}_1^*, D_1^*), \cdots, (\boldsymbol{u}_M^*, D_M^*)\} \tag{2.3}$$

to be an (n, M) code-valued random variable which satisfies the obvious measurability requirements. A (*random*) *jamming sequence* of length n, with the obvious definition, will be denoted by $\boldsymbol{s}^*$.

Clearly, if no further restrictions are imposed on the random codes and jamming sequences, the problem has an uninteresting solution. The error probability of any fixed positive-rate random code can be made arbitrarily close to $1 - 1/M$ by letting $\boldsymbol{s}^*$ be memoryless zero-mean Gaussian noise of arbitrarily large variance (or power). In practice, however, other restrictions will exist that prevent such trivial solutions. An interesting and natural restriction to investigate is that of placing some kind of *power constraint* on the codewords and the jamming sequences. In this paper we consider two types of power constraints: *peak* and *average*. We say that C_n^* satisfies a *peak input power constraint* (PI) if each codeword lies on or within an n-dimensional sphere (n-sphere) of radius $\sqrt{nP_T}$ almost surely (a.s.), i.e., if for each $1 \le i \le M$ the codeword $\boldsymbol{u}_i^* = (u_{i1}^*, \cdots, u_{in}^*)$ satisfies

$$\frac{1}{n} \sum_{j=1}^{n} u_{ij}^{*2} \le P_T \text{ a.s.} \tag{2.4}$$

This code satisfies an *average input power constraint* (AI) if the power averaged over all codewords has an expectation of at most P_T, i.e., if

$$\boldsymbol{E}\left\{ \frac{1}{nM} \sum_{j=1}^{M} \sum_{i=1}^{n} u_{ij}^{*2} \right\} \le P_T \tag{2.5}$$

where $\boldsymbol{E}\{\cdot\}$ denotes mathematical expectation. We also define two similar power constraints on the random jamming sequence $\boldsymbol{s}^*$. We say that $\boldsymbol{s}^*$ satisfies a *peak jamming power constraint* (PJ) if

$$\frac{1}{n} \sum_{i=1}^{n} s_i^{*2} \le P_J \text{ a.s.} \tag{2.6}$$

[1] Throughout this paper, except where otherwise indicated, asterisks are used as superscripts to denote random variables, bold lower-case letters indicate vectors (or vector-valued mappings) in $\boldsymbol{R}^n$, and $N(\mu, \sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2.

[2] We extend this definition to nonintegral M as follows: by an (n, M) code we mean an (n, M') code where M' is the smallest integer greater than or equal to M.

and an *average jamming power constraint* (AJ) if

$$E\left\{\frac{1}{n}\sum_{i=1}^{n} s_i^{*2}\right\} \leq P_J. \tag{2.7}$$

Two input power constraints (PI or AI) and two jamming power constraints (PJ or AJ) exist, and so a total of four possible combinations of transmitter and jammer power constraints can be considered. We adopt a simple binary nomenclature to describe each case. In the sequel, when we speak of the GAVC $A|B$, we mean the GAVC with input power constraint A (= PI or AI), and jamming power constraint B (= PJ or AJ).

We now specify what is meant by the "error probability" of the code C_n^*. Given a code C_n^* on the GAVC $A|B$ and the jamming sequence s^*, we can in principle calculate the (maximum) probability of error:

$$\lambda(C_n^*, s^*) \equiv \max_{1 \leq i \leq M} \Pr\{u_i^* + \eta_e^* + s^* \in \overline{D}_i^*\} \tag{2.8}$$

where $\overline{D}_i^*$ denotes $\boldsymbol{R}^n - D_i^*$. However, the distribution of s^* is not known in advance to the transmitter and may change from one block to the next in an unpredictable and arbitrary way, subject only to the power constraint B. The smallest error probability *guaranteed* to be achievable by the code C_n^* is the supremum of (2.8) over all s^* that satisfy power constraint B. Therefore, we define the error probability of the code C_n^* by

$$\lambda^B(C_n^*) = \sup_{s^*} \lambda(C_n^*, s^*) \tag{2.9}$$

where the supremum is performed over all B-admissible s^*.

For a given source rate R and constraint pair $A|B$, we ask what is the smallest error probability $\lambda^B(C_n^*)$ that can be achieved by any (n, M) random code C_n^* that satisfies constraint A and conveys R bits per second when n is large? Clearly, this error probability depends on both the rate R and the constraints $A|B$. We say that a pair (R, λ) where $R \geq 0$ and $0 \leq \lambda < 1$ is achievable for the case $A|B$ (*achievable* $A|B$) if for all $\epsilon > 0$ and for all n sufficiently large an (n, M) random code C_n^* exists satisfying constraint A such that

$$\log_2 M \geq n(R - \epsilon) \tag{2.10a}$$

and

$$\lambda^B(C_n^*) \leq \lambda + \epsilon. \tag{2.10b}$$

Let $\boldsymbol{R}_{A|B}$ denote the set of all achievable pairs (R, λ) for the GAVC $A|B$.

Note that if a certain pair (R, λ) is achievable $A|B$, then all pairs (R', λ') such that $R' \leq R$ and $\lambda' \geq \lambda$ are also achievable $A|B$. Consequently, $\boldsymbol{R}_{A|B}$ must be of the form

$$\boldsymbol{R}_{A|B} = \{(R, \lambda) | 0 \leq R \leq C_{A|B}(\lambda), 0 \leq \lambda < 1\} \tag{2.11}$$

where $C_{A|B}(\lambda)$ is a monotone increasing right-continuous function of λ. Thus to characterize $\boldsymbol{R}_{A|B}$ it suffices to determine $C_{A|B}(\lambda)$.

The function $C_{A|B}(\lambda)$ is called the λ-*capacity* of the channel [12], [21]. It can be interpreted as the largest rate of transmission that can be achieved by a random code that suffers an error probability no greater than λ for large n. If $C_{A|B}(\lambda)$ is equal to a constant on $0 \leq \lambda < 1$, say $C_{A|B}$, the latter is called the *capacity* of the channel; otherwise, if $C_{A|B}(\lambda)$ is *not* constant, we say that no capacity exists.[3] Most simple channel models that arise in information theory have a capacity. We will show for certain constraint pairs $A|B$ that GAVC's have no capacity; i.e., $C_{A|B}(\lambda)$ is *not* constant. This interesting and somewhat surprising fact distinguishes GAVC's from discrete AVC's: Blackwell *et al.* [10] have shown that the latter *always* possess a (random coding) capacity.

Recall that our objective is to determine the minimum error probability suffered by large block-length random codes of rate R when used on the GAVC $A|B$. Define this error probability by

$$\lambda^{A|B}(R) \equiv \lim_{\epsilon \to 0^+} \limsup_{n \to \infty} \inf_{C_n^*} \lambda^B(C_n^*) \tag{2.12}$$

where the infimum is over all A-admissible (n, M) random codes such that $M \geq 2^{n(R-\epsilon)}$. It is easy to show that the relationship between $\lambda^{A|B}(R)$ and $C_{A|B}(\lambda)$ is

$$\lambda^{A|B}(R) = \min\{0 \leq \lambda \leq 1 | C_{A|B}(\lambda) \geq R \text{ or } \lambda = 1\}. \tag{2.13}$$

Although it clearly provides the same information about $\boldsymbol{R}_{A|B}$ as $C_{A|B}(\lambda)$ does, $\lambda^{A|B}(R)$ is often easier to interpret.

We now present four theorems that characterize $C_{A|B}(\lambda)$ for each pair of constraints $A|B$, the proofs of which are provided in the next section. We first consider the case in which both the transmitter and jammer are constrained in peak power: GAVC PI|PJ. This channel actually has a capacity that is given by the following familiar formula.

Theorem 1: For the GAVC PI|PJ a (random coding) capacity exists and is given by

$$C_{\mathrm{PI|PJ}}(\lambda) = \hat{C}_{\mathrm{PI|PJ}}$$

for all $0 \leq \lambda < 1$, where

$$\hat{C}_{\mathrm{PI|PJ}} \equiv \frac{1}{2}\log_2\left(1 + \frac{P_T}{N_e + P_J}\right). \tag{2.14}$$

Remark: Blachman [8, p. 53, eq. 10] states (without proof) a similar result.

It is interesting to note that $\hat{C}_{\mathrm{PI|PJ}}$ is identical to the capacity formula of the memoryless Gaussian channel that would be formed if the jammer transmitted a sequence of i.i.d. $N(0, P_J)$ random variables (cf. Wolfowitz [21, theorem 9.2.1])[4,5] We conclude for the GAVC PI|PJ that an intelligent jammer can do no more harm (in the sense of

[3] An alternative (e.g., Csiszár and Körner [12]) definition of capacity (which always exists) is $C_{A|B} \equiv \lim_{\lambda \to 0^+} C_{A|B}(\lambda)$. Our definition is that of Wolfowitz [21].

[4] It is also the formula obtained by Dobrushin [13] for the capacity of the Gaussian *compound channel.*

[5] Note that this Gaussian jamming sequence does not satisfy PJ. It is possible, however, to construct a jamming sequence that does satisfy PJ and that yields nearly the same capacity (cf. proof of Theorem 2).

reducing the achievable region) than Gaussian noise of the same power, regardless of how he distributes his power.

We now change the jamming power constraint from PJ to AJ (i.e., GAVC PI|AJ) and ask whether the foregoing conclusion is still valid. Since bounds on average power are weaker than those on peak power, it is obvious that $\boldsymbol{R}_{\mathrm{PI|AJ}} \subset \boldsymbol{R}_{\mathrm{PI|PJ}}$. However, as the next theorem demonstrates, this inclusion is strict. In fact, we find for this and all remaining cases in which either the transmitter or the jammer or both are subject to average power constraints that no capacity exists, i.e., only λ-capacities are found.

Theorem 2: For the GAVC PI|AJ the (random coding) λ-capacity is

$$G_{\mathrm{PI|AJ}}(\lambda) = \hat{C}_{\mathrm{PI|AJ}}(\lambda)$$

for all $0 \le \lambda < 1$, where

$$\hat{C}_{\mathrm{PI|AJ}}(\lambda) \equiv \frac{1}{2}\log_2\left(1 + \frac{P_T}{N_e + P_J/\lambda}\right). \qquad (2.15)$$

Remark: $\hat{C}_{\mathrm{PI|AJ}}(0)$ is interpreted as 0.

Observe that the expression for $\hat{C}_{\mathrm{PI|AJ}}(\lambda)$ is identical to that of $\hat{C}_{\mathrm{PI|PJ}}$ except that the jamming power appears boosted by a factor that is the reciprocal of the tolerable error probability λ. Some insight into this formula can be gained by stating the result in terms of the error probability suffered by codes of rate R. Theorem 2 states that for increasing n, optimal $(n, 2^{nR})$ random codes satisfying PI suffer an error probability that approaches

$$\lambda^{\mathrm{PI|AJ}}(R) = \hat{\lambda}^{\mathrm{PI|AJ}}(R) \equiv \begin{cases} \dfrac{(4^R - 1)P_J}{P_T - (4^R - 1)N_e}, & R \le \hat{C}_{\mathrm{PI|AJ}}(1) \\ 1, & R > \hat{C}_{\mathrm{PI|AJ}}(1) \end{cases} \qquad (2.16)$$

against an AJ-constrained jammer.

The right side of (2.16) is increasing positive whenever R is positive and for small R becomes asymptotic to $2\ln 2 RP_J/P_T$. The region $\boldsymbol{R}_{\mathrm{PI|AJ}}$ is sketched in Fig. 2. It is apparent that a PI-admissible random code can achieve high reliability (i.e., $\lambda^{\mathrm{AJ}}(C_n^*) \approx 0$) only in the limit as R or P_J/P_T becomes vanishingly small. Evidently, *reliable communication is impossible at any positive source rate.*

We now sketch the basic idea behind (2.16) (or equivalently, Theorem 2); a detailed proof follows in Section III. Let C_n^* be any PI-admissible random code of rate R. Suppose the jammer transmits only jamming sequences that have the following structure. First, a jamming power P^* is selected where P^* is a nonnegative random variable that satisfies $\boldsymbol{E}P^* \le P_J$. Conditioned on the value of P^* the jamming sequence $\boldsymbol{s}^*$ is an i.i.d. $N(0, P^*)$ sequence of length n. (Clearly, this restriction can only *increase* the achievable region.) It is easily verified that $\boldsymbol{s}^*$ satisfies AJ. With this restriction the channel "seen" by the transmitter during any block transmission is a discrete-time Gaussian channel with (unknown) noise power $N_e + P^*$. According to the coding theorem and strong converse for this channel (cf. Wolfowitz [21, theorems 9.2.1–2]), if

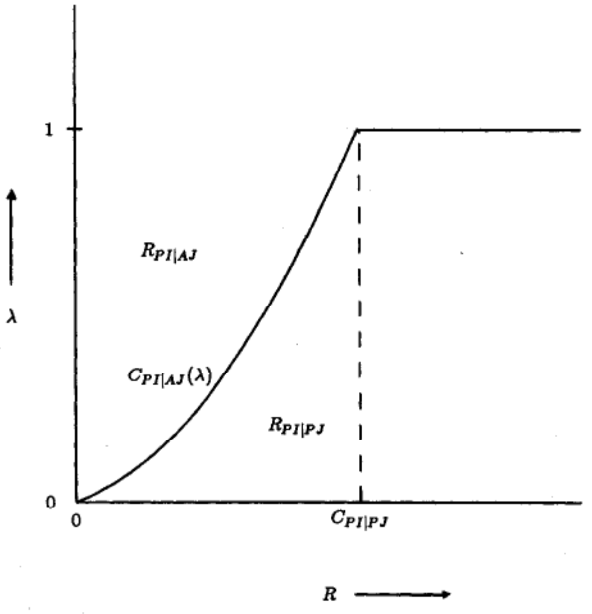

Fig. 2. Achievable regions for GAVC PI/PJ and PI/AJ.

$$R < \frac{1}{2}\log_2\left(1 + \frac{P_T}{N_e + P^*}\right)$$

and n is large, then $\lambda^{\mathrm{AJ}}(C_n^*) \approx 0$ is possible; however, if

$$R > \frac{1}{2}\log_2\left(1 + \frac{P_T}{N_e + P^*}\right),$$

then $\lambda^{\mathrm{AJ}}(C_n^*) \approx 1$ is certain. The jammer must, therefore, choose

$$P^* \ge \frac{P_T}{(4^R - 1)} - N_e$$

to be guaranteed an appreciable error probability, and this power is sufficient to yield an error probability near unity. Therefore, the best codes have an error probability that approximates the probability of this event:

$$\lambda^{\mathrm{AJ}}(C_n^*) \approx \Pr\left\{P^* \ge \frac{P_T}{(4^R - 1)} - N_e\right\}.$$

Finally, the foregoing right-hand expression takes on a maximum value which is equal to the $\hat{\lambda}^{\mathrm{PI|AJ}}(R)$ when P^* is chosen in the following way:

$$\Pr\left\{P^* \approx \frac{P_T}{(4^R - 1)} - N_e\right\} = 1 - \Pr\{P^* = 0\} = \hat{\lambda}^{\mathrm{PI|AJ}}(R).$$

It follows that $\lambda^{\mathrm{AJ}}(C_n^*)$ is not appreciably less than $\hat{\lambda}^{\mathrm{PI|AJ}}(R)$ for large n.

Although we have allowed the jammer foreknowledge of the statistics of the transmitter's random code when selecting a jamming sequence (cf. (2.9)), it turns out that this knowledge is *unnecessary*. Remarkably, the aforementioned jamming sequence does *not* depend on the detailed structure of the code but only on the blocklength n, the source rate R, and the parameters P_T, P_J, and N_e. It is also interesting that this jamming sequence is essentially a *pulsed strategy* (i.e., either "off" or "on" with high peak power). Memoryless pulsed jamming strategies have been shown to maximize the error probability of certain uncoded modulation systems such as binary phase-shift keying (BPSK) (cf. Simon *et al.* [20]). Theorem 2 shows that pulsed jamming sequences *with memory* play a similar role for random block codes on the GAVC.

We have seen from Theorem 2 that an average-power-limited jammer has a tremendous advantage against a peak-power-limited transmitter; in fact, reliable communication is impossible in this case. It is interesting to ask whether the transmitter might similarly gain by varying codeword power against a peak-power-limited jammer as in the case AI|PJ. The next theorem shows that little advantage will be gained.

Theorem 3: For the GAVC with constraints AI|PJ the (random coding) λ-capacity is

$$C_{\mathrm{AI|PJ}}(\lambda) = \hat{C}_{\mathrm{AI|PJ}}(\lambda)$$

for all $0 \le \lambda < 1$, where

$$\hat{C}_{\mathrm{AI|PJ}}(\lambda) \equiv \frac{1}{2}\log_2\left(1 + \frac{P_T/(1-\lambda)}{N_e + P_J}\right). \quad (2.17)$$

The corresponding achievable region is sketched in Fig. 3. We see that if a high error probability can be tolerated, the allowable coding rate is much improved; however, at low error probabilities $C_{\mathrm{AI|PJ}}(\lambda)$ approaches $C_{\mathrm{PI|PJ}}$, and the improvements are negligible. As in the other cases, we can state the result in terms of error probabilities: optimal AI-admissible $(n, 2^{nR})$ random codes suffer an error probability that, for large n, approaches

$$\lambda^{\mathrm{AI|PJ}}(R) = \begin{cases} 0, & R \le \hat{C}_{\mathrm{AI|PJ}}(0) \\ 1 - \dfrac{P_T}{(4^R - 1)(N_e + P_J)}, & R > \hat{C}_{\mathrm{AI|PJ}}(0) \end{cases}. \quad (2.18)$$

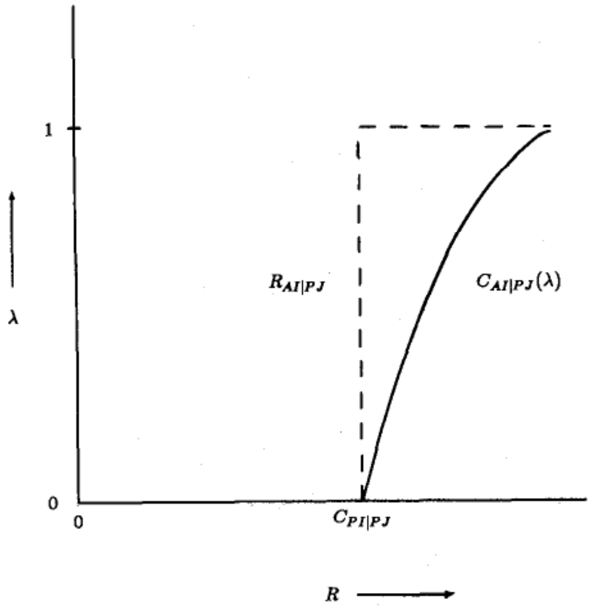

Fig. 3. Achievable region for GAVC AI/PJ.

Thus the rates at which *reliable* communication can occur are the same as in the case PI|PJ. Clearly, codeword power variation offers little improvement to the transmitter.

We now consider the GAVC AI|AJ. As Theorem 3 shows, the additional flexibility offered by the power constraint AI is relatively useless against a peak-power-limited jammer. We now ask if the transmitter might at least reduce the gain of the average-power-limited jammer compared with the GAVC PI|AJ. The next theorem shows that some limited improvement is made.

Theorem 4: For the GAVC with constraints AI|AJ the (random coding) λ-capacity is given by

$$C_{\mathrm{AI|AJ}}(\lambda) = \hat{C}_{\mathrm{AI|AJ}}(\lambda)$$

for all $0 \le \lambda < 1$, where

$$\hat{C}_{\mathrm{AI|AJ}}(\lambda) = \begin{cases} \dfrac{1}{2}\log_2\left(1 + \dfrac{P_T}{N_e + P_J/2\lambda}\right), & 0 \le \lambda \le \lambda_c \\ \dfrac{1}{2}\log_2\left(1 + \dfrac{P_T(1-2\lambda_c)}{(1-\lambda)N_e}\right), & \lambda_c \le \lambda < 1 \end{cases} \quad (2.19a)$$

$$\lambda_c \equiv \frac{P_J}{2N_e}\left(\sqrt{1 + \frac{2N_e}{P_J}} - 1\right)$$

for $N_e > 0$, and

$$\hat{C}_{\mathrm{AI|AJ}} = \begin{cases} \dfrac{1}{2}\log_2\left(1 + \dfrac{2\lambda P_T}{P_J}\right), & 0 \le \lambda < \dfrac{1}{2} \\ \dfrac{1}{2}\log_2\left(1 + \dfrac{P_T}{2(1-\lambda)P_J}\right), & \dfrac{1}{2} \le \lambda < 1 \end{cases} \quad (2.19b)$$

for $N_e = 0$.

Remark: The function (2.19a) tends continuously to (2.19b) as $N_e \to 0$.

The corresponding achievable region is sketched in Fig. 4, with $C_{\mathrm{PI|PJ}}$, $C_{\mathrm{PI|AJ}}(\lambda)$, and $C_{\mathrm{AI|PJ}}(\lambda)$ included for comparison. Optimal $(n, 2^{nR})$ random codes satisfying AI must then, as n grows large, suffer an error probability that approaches $\lambda^{\mathrm{AI|AJ}}(R) = \hat{\lambda}^{\mathrm{AI|AJ}}(R)$ where

$$\hat{\lambda}^{\mathrm{AI|AJ}}(R) = \begin{cases} \dfrac{P_J(4^R - 1)}{2\left(P_T - (4^R - 1)N_e\right)}, & R \le \hat{C}_{\mathrm{AI|AJ}}(\lambda_c) \\ 1 - \dfrac{P_T(1-2\lambda_c)}{(4^R - 1)N_e}, & R > \hat{C}_{\mathrm{AI|AJ}}(\lambda_c) \end{cases} \quad (2.20a)$$

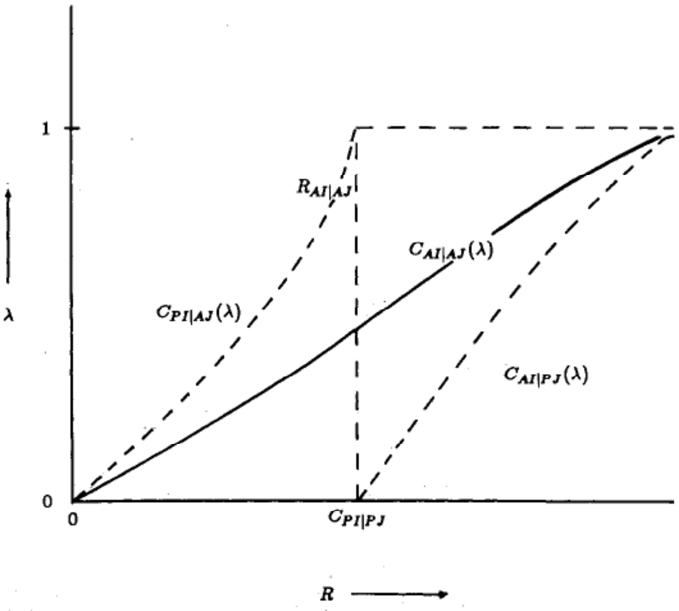

Fig. 4. Achievable region for GAVC AI/AJ (with all other λ-capacities included for comparison).

when $N_e > 0$, and

$$\hat{\lambda}^{\mathrm{AI|AJ}}(R) = \begin{cases} \dfrac{P_J(4^R - 1)}{2P_T}, & R \le \dfrac{1}{2}\log_2\left(1 + \dfrac{P_T}{P_J}\right) \\ 1 - \dfrac{P_T}{2(4^R - 1)P_J}, & R > \dfrac{1}{2}\log_2\left(1 + \dfrac{P_T}{P_J}\right) \end{cases} \tag{2.20b}$$

when $N_e = 0$.

For $R < C_{\mathrm{AI|AJ}}(\lambda_c)$ observe that the error probability is half that of GAVC PI|AJ; however, when $R > C_{\mathrm{AI|AJ}}(\lambda_c)$, the probability of being *correct* $(= 1 - \lambda^{\mathrm{AJ}}(C_n^*))$ is $(1 - 2\lambda_c)(P_J + N_e)/N_E$ times that in the AI|PJ case. $C_{\mathrm{AI|AJ}}(\lambda)$ is, therefore, a compromise between $C_{\mathrm{PI|AJ}}(\lambda)$ and $C_{\mathrm{AI|PJ}}(\lambda)$. As in the PI|AJ case the error probability can be made small only by making R or P_J/P_T small.

An intuitive justification for (2.20a) follows (a rigorous proof is provided in Section III). Suppose as before that the jammer transmits only sequences that have the following structure. Conditioned on the value of P_2^*, s^* is an i.i.d. sequence of $N(0, P_2^*)$ random variables, where P_2^* is a nonnegative random variable that satisfies $EP_2^* \le P_J$. The transmitter constructs a random code C_n^* in the following way. He first selects a random code $\bar{C}_n^*$ of rate R whose average power is no greater than unity, i.e.,

$$E\left\{\frac{1}{nM}\sum_{j=1}^{M}\sum_{i=1}^{n}u_{ij}^{*2}\right\} \le 1,$$

and then to form C_n^*, he multiplies each codeword in $\bar{C}_n^*$ by $\sqrt{P_1^*}$, where P_1^* is an independent nonnegative random variable satisfying $EP_1^* \le P_T$. The performance of this code against s^* is a function of the signal-to-noise power ratio $P_1^*/(P_2^* + N_e)$. As in the earlier argument following Theorem 2, if

$$\frac{P_1^*}{P_2^* + N_e} > (4^R - 1),$$

then $\lambda(C_n^*, s^*)$ can be small; however, if

$$\frac{P_1^*}{P_2^* + N_e} < (4^R - 1),$$

then it is certainly true that $\lambda(C_n^*, s^*) \approx 1$. Therefore, for the best choice of $\bar{C}_n^*$ we have for large n

$$\lambda(C_n^*, s^*) \approx \Pr\left\{P_1^* < (4^R - 1)(P_2^* + N_e)\right\}. \tag{2.21}$$

The optimum error probability thus depends only on the power distribution of the transmitter and jammer. Naturally, the transmitter wants to minimize (2.21) with an appropriate choice of P_1^*, and the jammer wants to maximize it by an effective choice of P_2^*. Therefore, an optimal code suffers the error probability

$$\lambda^{\mathrm{AJ}}(C_n^*) \approx \max_{P_1^*:\, EP_1^* \le P_T}\ \min_{P_2^*:\, EP_2^* \le P_J} \Pr\left\{P_1^* < (4^R - 1)(P_2^* + N_e)\right\}.$$

It can be shown (cf. proof of Theorem 4) that the right side of this equation is equal to the right side of (2.20a).

Finally, we consider the coding problems that result from the imposition of *multiple* constraints. Suppose our code must satisfy some constraint, say A, for some constant P_T and another constraint A' for some constant $P_T' \ne P_T$. Denote this joint constraint by AA'. Similarly, one may define a double constraint BB' on jamming sequences. It is easily checked that the λ-capacities for these more complex coding problems can be constructed from the λ-capacities defined by Theorems 1–4 according to the following simple rules:[6]

$$C_{AA'|B}(\lambda) = \min\left\{\hat{C}_{A|B}(\lambda), \hat{C}_{A'|B}(\lambda)\right\} \tag{2.22a}$$

$$C_{A|BB'}(\lambda) = \max\left\{\hat{C}_{A|B}(\lambda), \hat{C}_{A|B'}(\lambda)\right\}. \tag{2.22b}$$

III. Proofs of Theorems 1–4

We now present some definitions and results that we use in the following proofs. By the *standard* (n, M) *random code* we mean a random code

$$C_n^{s*} \equiv \{(v_1^*, A_1^*), \cdots, (v_M^*, A_M^*)\} \tag{3.1}$$

constructed in the following way.

1) The M random codewords $\{v_1^*, \cdots, v_M^*\}$ are a collection of mutually independent random n-vectors, each of which is uniformly distributed on the n-sphere of radius $\sqrt{n}$; i.e., the probability that v_i^* lies within a certain region on the surface of this n-sphere is proportional to the surface are (or, equivalently, solid angle) of this region.

[6] It is unknown whether $C_{AA'|BB'}(\lambda)$ can similarly be decomposed.

2) The random decoding sets $\{A_i^*\}_{i=1}^M$ are defined by a *strict minimum Euclidean distance* rule, viz.,

$$A_i^* \equiv \{y \in \boldsymbol{R}^n \| y - \boldsymbol{v}_i^*| < |y - \boldsymbol{v}_k^*|, \quad \text{for all } k \neq i, 1 \leq k \leq M\} \tag{3.2}$$

where $|\cdot|$ denotes the usual Euclidean norm on $\boldsymbol{R}^n$. If a tie occurs, the receiver draws no conclusion about the transmitted message (and hence an error occurs).

We make several observations about the random code C_n^{s*}. First, the codewords of C_n^{s*} are clearly PI-admissible for $P_T = 1$; in fact, (2.4) is satisfied with equality (a.s.). Second, since all codewords have equal length (or power), each decoding set in (3.2) is a "flat-sided" cone with vertex at the origin. It follows that the sets $\{A_i^*\}_{i=1}^M$ are also minimum distance decoding sets for every codeword set of the form $\{\sqrt{P}\boldsymbol{v}_1^*, \cdots, \sqrt{P}\boldsymbol{v}_M^*\}$ where $P > 0$. Third, Shannon [19] has considered the use of this random code on the discrete-time additive Gaussian noise channel and has obtained the following result: there exist functions, say $K(R, P)$ and $E(R, P)$, both positive so long as

$$R \equiv \frac{1}{n} \log_2 M < \frac{1}{2} \log_2 (1 + P), \tag{3.3}$$

such that[7]

$$\Pr\{\sqrt{P}\,\boldsymbol{v}_i^* + \boldsymbol{\eta}^* \in \overline{A}_i^*\} \leq K(R, P) \exp\{-nE(R, P)\} \tag{3.4}$$

holds for all $1 \leq i \leq M$ and $n \geq 1$, where η^* denotes an n-vector of i.i.d. $N(0,1)$ random variables. Furthermore, $K(R, P)$ and $E(R, P)$ can be selected so that

$$K(\cdot, P), -E(\cdot, P) \text{ are increasing} \tag{3.5a}$$

and

$$K(R, \cdot), -E(R, \cdot) \text{ are decreasing} \tag{3.5b}$$

for all R and P satisfying (3.3).

We also require an Arimoto-style strong converse [3] for the discrete-time additive Gaussian noise channel with peak input power constraint and the average probability of error concept. The proof of the following result is very similar to the argument leading to [19, eq. 51] (carried out for $R > C$) and is, therefore, omitted. Let

$$C_n^* \equiv \{(\boldsymbol{u}_1^*, D_1^*), \cdots, (\boldsymbol{u}_M^*, D_M^*)\}$$

be any PI-admissible (n, M) random code with $P_T = P$. Functions exist, say $K'(R, P)$ and $E'(R, P)$, which are both positive whenever

$$R \equiv \frac{1}{n} \log_2 M > \frac{1}{2} \log_2 (1 + P) \tag{3.6}$$

such that

$$\frac{1}{M} \sum_{i=1}^{M} \Pr\{\boldsymbol{u}_i^* + \boldsymbol{\eta}^* \in \overline{D}_i^*\} \geq 1 - K'(R, P) \exp\{-nE'(R, P)\} \tag{3.7}$$

holds for all $n \geq 1$. (Note that any lower bound on the average error probability is *a fortiori* a lower bound on the maximum probability of error.) Furthermore, $K'(R, P)$ and $E'(R, P)$ can be selected so that

$$K'(\cdot, P), -E'(\cdot, P) \text{ are increasing} \tag{3.8a}$$

and

$$K'(R, \cdot), -E'(R, \cdot) \text{ are decreasing} \tag{3.8b}$$

for all R and P that satisfy (3.6).

We now present a lemma that forms the kernel of the converses to Theorems 3 and 4. This lemma is of independent interest because it gives a tight lower bound on the average error probability of any code when used on a Gaussian channel in terms of the code's power distribution.

Define for any $\boldsymbol{u} = (u_1, \cdots, u_n) \in \boldsymbol{R}^n$ the quantity

$$P(\boldsymbol{u}) \equiv \frac{1}{n} \sum_{j=1}^{n} u_j^2. \tag{3.9}$$

For any random code C_n^* let $U^*(C_n^*)$ be the random variable that is uniformly distributed on the set $\{\boldsymbol{u}_1^*, \cdots, \boldsymbol{u}_M^*\}$ of codewords of C_n^*; i.e., let k^* be uniformly distributed on the set $\{1, \cdots, M\}$ and independent of C_n^*, and define

$$U^*(C_n^*) \equiv \boldsymbol{u}_{k^*}^*. \tag{3.10}$$

Lemma 1: Let C_n^* be any (n, M) random code and J^* be any nonnegative random variable. Then for all $0 < \epsilon < R \equiv (1/n) \log_2 M$ the following holds:

$$\frac{1}{M} \sum_{i=1}^{M} \Pr\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + \sqrt{J^*}\,\boldsymbol{\eta}^* \in \overline{D}_i^*\} \geq \Pr\{P(U^*(C_n^*)) < (4^{R-\epsilon} - 1)(N_e + J^*)\} - \gamma_n(\epsilon) \tag{3.11}$$

where

$$\gamma_n(\epsilon) \equiv K'(R - \epsilon/2, 4^{R-\epsilon} - 1) \cdot \exp\{-nE'(R - \epsilon/2, 4^{R-\epsilon} - 1)\} + 2^{-n\epsilon/2}. \tag{3.12}$$

Remarks: Note that we do not assume that C_n^* and J^* are independent. The function $\gamma_n(\epsilon)$ depends *only* on n, ϵ, and R and is independent of the random code and the jamming power. In addition, for all $\epsilon > 0$, $\gamma_n(\epsilon) \to 0$ exponentially.

Proof of Lemma 1: Let $0 < \epsilon < R$. Suppose first that C_n^* and J^* are *deterministic*, say $C_n^* = C_n \equiv \{(\boldsymbol{u}_1, D_1), \cdots, (\boldsymbol{u}_M, D_M)\}$ and $J^* = J$. Define the set

$$S_\epsilon(C_n, J) \equiv \{1 \leq i \leq M | P(\boldsymbol{u}_i) < (4^{R-\epsilon} - 1)(N_e + J)\}, \tag{3.13}$$

[7] We have presented Shannon's result in a form that is different from the original statement in [19] but which is convenient for the proofs of the present section. Our form can be obtained from Shannon's "firm" upper bound in [19] by making the substitution indicated in the footnote to Gallager [14, p. 16] and simplifying the resulting bound.

and let $N_\epsilon(C_n, J)$ denote its cardinality. It is immediate that

$$\Pr\left\{P(U^*(C_n)) < (4^{R-\epsilon}-1)(N_e+J)\right\} = \frac{N_\epsilon(C_n, J)}{M}. \tag{3.14}$$

Consider the subcode of C_n that consists of those codewords and decoding sets with indices in $S_\epsilon(C_n, J)$. ,The average error probability of this subcode when $J^* = J$ can be bounded below by the following strong converse (cf. (3.7)) for the Gaussian channel[8]

$$\frac{1}{N_\epsilon(C_n, J)} \sum_{i \in S_\epsilon(C_n, J)} \Pr\left\{\boldsymbol{u}_i + \boldsymbol{\eta}_e^* + \sqrt{J}\,\boldsymbol{\eta}^* \in \overline{D}_i\right\} \ge 1 - K'\left(R_n, 4^{R-\epsilon}-1\right) \cdot \exp\left\{-nE'\left(R_n, 4^{R-\epsilon}-1\right)\right\} \tag{3.15}$$

provided that

$$R_n \equiv \frac{\log_2\left(N_\epsilon(C_n, J)\right)}{n} > R - \epsilon. \tag{3.16}$$

Therefore, the following holds for all C_n, J, ϵ, and R:[9]

$$\begin{aligned}
\frac{1}{N_\epsilon(C_n, J)} &\sum_{i \in S_\epsilon(C_n, J)} \Pr\left\{\boldsymbol{u}_i + \boldsymbol{\eta}_e^* + \sqrt{J}\,\boldsymbol{\eta}^* \in \overline{D}_i\right\} \\
&\ge \left(1 - K'\left(R_n, 4^{R-\epsilon}-1\right) \cdot \exp\left\{-nE'\left(R_n, 4^{R-\epsilon}-1\right)\right\}\right) 1_{[R-\epsilon/2, +\infty)}(R_n) \\
&\overset{a)}{\ge} \left(1 - \gamma_n(\epsilon) + 2^{-n\epsilon/2}\right) 1_{[R-\epsilon/2, +\infty)}(R_n) \\
&\ge 1_{[R-\epsilon/2, +\infty)}(R_n) - \gamma_n(\epsilon) + 2^{-n\epsilon/2}
\end{aligned} \tag{3.17}$$

where $\gamma_n(\epsilon)$ is as defined in (3.12). Step a) is a consequence of (3.8a). We now derive a lower bound to the average error probability of C_n when $J^* = J$:

$$\begin{aligned}
\frac{1}{M} \sum_{i=1}^{M} &\Pr\left\{\boldsymbol{u}_i + \boldsymbol{\eta}_e^* + \sqrt{J}\,\boldsymbol{\eta}^* \in \overline{D}_i\right\} \\
&\ge \frac{1}{M} \sum_{i \in S_\epsilon(C_n, J)} \Pr\left\{\boldsymbol{u}_i + \boldsymbol{\eta}_e^* + \sqrt{J}\,\boldsymbol{\eta}^* \in \overline{D}_i\right\} \\
&\overset{a)}{\ge} \frac{N_\epsilon(C_n, J)}{M} \cdot \left(1_{[R-\epsilon/2, +\infty)}(R_n) - \gamma_n(\epsilon) + 2^{-n\epsilon/2}\right) \\
&\ge \frac{N_\epsilon(C_n, J)}{M} - \gamma_n(\epsilon) + 2^{-n\epsilon/2} - \frac{N_\epsilon(C_n, J)}{M} 1_{(-\infty, R-\epsilon/2)}(R_n) \\
&\ge \frac{N_\epsilon(C_n, J)}{M} - \gamma_n(\epsilon) \\
&\overset{b)}{\ge} \Pr\left\{P(U^*(C_n)) < (4^{R-\epsilon}-1)(N_e+J)\right\} - \gamma_n(\epsilon)
\end{aligned} \tag{3.18}$$

where a) follows from (3.17) and b) from (3.14). Thus the lemma holds for all deterministic C_n and J. To prove (3.11) for random C_n^* and J^*, average (3.18) over the joint distribution of C_n^* and J^*. This completes the proof of Lemma 1.

Proof of Theorem 1:

a) $C_{PI|PJ}(\lambda) \ge \hat{C}_{PI|PJ}$: Let R nonnegative be given and set $M_n = \lfloor 2^{nR} \rfloor$.[10] Define a sequence of (n, M_n) random codes, say $\{C_n^*\}_{n=1}^{\infty}$, in the following way:

$$C_n^* = \left\{(\boldsymbol{u}_1^*, A_1^*), \cdots, (\boldsymbol{u}_{M_n}^*, A_{M_n}^*)\right\} \tag{3.19}$$

where $\boldsymbol{u}_i^* = \sqrt{P_T}\,\boldsymbol{v}_i^*$ and $\{(\boldsymbol{v}_1^*, A_1^*), \cdots, (\boldsymbol{v}_{M_n}^*, A_{M_n}^*)\}$ is the standard (n, M_n) random code defined in (3.1). It is easily verified that C_n^* satisfies PI for each $n \ge 1$. We claim that if

$$R < \hat{C}_{\mathrm{PI|PJ}}, \tag{3.20}$$

then a positive sequence $\{\gamma_n\}_{n=1}^{\infty}$ exists such that

$$\lambda^{\mathrm{PJ}}(C_n^*) \le \gamma_n \tag{3.21}$$

and $\gamma_n \to 0$ as $n \to +\infty$, thereby proving a).

To establish this claim, suppose that (3.20) holds. Let ω^* be an independent random variable that is uniformly distributed on the unit n-sphere and define

$$\sigma_n(l) \equiv \Pr\left\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + l\omega^* \in \overline{A}_i^*\right\} \tag{3.22}$$

for any real number $l \ge 0$. (Clearly, $\sigma_n(\cdot)$ does *not* depend on i.) Let $\boldsymbol{s}^*$ be any jamming sequence that satisfies PJ. The error probability incurred by $\boldsymbol{s}^*$ when message i is sent can be bounded in the following way:

$$\Pr\left\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + \boldsymbol{s}^* \in \overline{A}_i^*\right\} \overset{a)}{=} \boldsymbol{E}\sigma_n(|\boldsymbol{s}^*|) \overset{b)}{\le} \sigma_n\left(\sqrt{nP_J}\right). \tag{3.23}$$

The justification of these steps is as follows. To prove a), let $\boldsymbol{s}$ be any vector in $\boldsymbol{R}^n$, let ω be any unit vector in $\boldsymbol{R}^n$, and let T be any orthogonal transformation on $\boldsymbol{R}^n$ that maps $\boldsymbol{s}$ to $|\boldsymbol{s}|\omega$, i.e.,

$$T\boldsymbol{s} = |\boldsymbol{s}|\omega.$$

Since minimum distance decoding is used (and distances are preserved by T), the following holds:

$$\Pr\left\{\boldsymbol{u}_1^* + \boldsymbol{\eta}_e^* + \boldsymbol{s} \in \overline{A}_1^*\right\} = \Pr\left\{T\boldsymbol{u}_1^* + T\boldsymbol{\eta}_e^* + |\boldsymbol{s}|\omega \in T\overline{A}_1^*\right\}.$$

The sets $\{T\overline{A}_i^*\}_{i=1}^{M}$ remain minimum distance decoding sets for the codewords $\{T\boldsymbol{u}_i^*\}_{i=1}^{M}$, and the distributions of $\{\boldsymbol{u}_i^*\}_{i=1}^{M}$ and $\boldsymbol{\eta}_e^*$ are spherically symmetric and so are unchanged by T. We conclude that

$$\Pr\left\{\boldsymbol{u}_1^* + \boldsymbol{\eta}_e^* + s \in \overline{A}_1^*\right\} = \Pr\left\{\boldsymbol{u}_1^* + \boldsymbol{\eta}_e^* + |\boldsymbol{s}|\omega \in \overline{A}_1^*\right\}$$

for all ω in the ensemble of ω^* from which a) immediately follows. To prove b), it suffices to show that $\sigma_n(\cdot)$ is increasing, since PI implies $|\boldsymbol{s}^*| \le \sqrt{nP_J}$ (a.s.). Let l and $\hat{l}$ be nonnegative numbers so that $0 \le \hat{l} \le l$. Let the random variable m_l^* be defined by

$$m_l^* \equiv |\boldsymbol{\eta}_e^* + l\omega^*|,$$

[8] We interpret the left-hand expression in (3.15) as zero if $N(C_n, J) = 0$.

[9] $1_A(x) \equiv \begin{cases} 1, & x \in A \\ 0, & x \in \overline{A} \end{cases}$.

[10] $\lfloor x \rfloor$ denotes the integer n such that $x - 1 < n \le x$.

and let $F_l(m)$ be its distribution function. It is easy to verify that, conditioned on the occurrence $m_l^* = m$, the random variable $\boldsymbol{\eta}_e^* + l\omega^*$ is uniformly distributed on the n-sphere of radius m; hence its conditional distribution does not depend on l. Therefore, define the quantity

$$\gamma(m) \equiv \Pr\left\{\boldsymbol{u}_1^* + \boldsymbol{\eta}_e^* + l\omega^* \in \bar{A}_1^* | m_l^* = m\right\} \quad (3.24)$$

and note that

$$\sigma_n(l) = \int_0^\infty \gamma(m)\, dF_l(m).$$

Since A_1^* is a set formed by the minimum distance rule, if $m < \hat{m}$, then

$$\boldsymbol{u}_1^* + m\left(\frac{\boldsymbol{\eta}_e^* + l\omega^*}{m_l^*}\right) \in \bar{A}_1^*$$

implies

$$\boldsymbol{u}_1^* + \hat{m}\left(\frac{\boldsymbol{\eta}_e^* + l\omega^*}{m_l^*}\right) \in \bar{A}_1^*,$$

and, consequently, $\gamma(\cdot)$ is monotone increasing. If for each m, $F_l(m)$ is monotone decreasing as a function of l, then using integration by parts and the monotonicity of $\gamma(\cdot)$, we obtain

$$\sigma_n(\hat{l}) = \int_0^\infty \gamma(m)\, dF_{\hat{l}}(m) \le \int_0^\infty \gamma(m)\, dF_l(m) = \sigma_n(l)$$

as desired. It, therefore, only remains to show that

$$F_l(m) \le F_{\hat{l}}(m). \quad (3.25)$$

We have, in fact, a stronger result that implies (3.25):

$$\Pr\left\{|\boldsymbol{\eta}_e^* + l\omega^*|^2 \le m^2 | \omega^* = \omega\right\}$$
$$\le \Pr\left\{|\boldsymbol{\eta}_e^* + \hat{l}\omega^*|^2 \le m^2 | \omega^* = \omega\right\}$$

for all ω. The latter inequality is an immediate consequence of the fact that the distribution of $\boldsymbol{\eta}_e^*$ decreases monotonically and symmetrically with distance from the origin. This completes the proof of (3.23).

Taking the supremum of (3.23) over all $1 \le i \le M$ and $\boldsymbol{s}^*$ satisfying PJ, we obtain the bound

$$\lambda^{\mathrm{PJ}}(C_n^*) \le \sigma_n\left(\sqrt{nP_J}\right). \quad (3.26)$$

It only remains to estimate the right-hand expression in (3.26); this is easily done by relating it to the error probability for the ordinary Gaussian channel. Let $\sqrt{P_J}\,\boldsymbol{\eta}^*$ denote a vector of i.i.d. $N(0, P_J)$ random variables, and let $f(\cdot)$ denote the probability density function of the random variable $m^* \equiv \sqrt{P_J}|\boldsymbol{\eta}^*|$. It is easy to show that

$$\Pr\left\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + \sqrt{P_J}\,\boldsymbol{\eta}^* \in \bar{A}_i^*\right\} = \int_0^\infty \sigma_n(l) f(l)\, dl. \quad (3.27)$$

Since $\sigma_n(l)$ is monotonically increasing, it follows that

$$\sigma_n\left(\sqrt{nP_J}\right) \le \frac{\int_{\sqrt{nP_J}}^\infty \sigma_n(l) f(l)\, dl}{\int_{\sqrt{nP_J}}^\infty f(l)\, dl}$$
$$\le \frac{\Pr\left\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + \sqrt{P_J}\,\boldsymbol{\eta}^* \in \bar{A}_i^*\right\}}{\Pr\left\{|\boldsymbol{\eta}^*| \ge 1\right\}}. \quad (3.28)$$

We now invoke (3.4) (compare (3.20) to (3.3)) to bound the numerator of (3.28) by

$$K(R, P_1)\exp\left\{-nE(R, P_1)\right\} \quad (3.29)$$

where

$$P_b \equiv \frac{P_T}{N_e + P_J/b} \quad (3.30)$$

for all $b > 0$. For large n the denominator of (3.28) approaches 0.5 by the central limit theorem. Thus a positive constant, say c_0, exists such that

$$\Pr\left\{|\boldsymbol{\eta}^*| \ge 1\right\} \ge 1/c_0 > 0 \quad (3.31)$$

for all $n \ge 1$. (In fact, $c_0 = 4$ will suffice.) Combining (3.26), (3.28), (3.29), and (3.31), we conclude that

$$\lambda^{\mathrm{PJ}}(C_n^*) \le c_0 K(R, P_1)\exp\left\{-nE(R, P_1)\right\} \quad (3.32)$$

for all $n \ge 1$. The right side tends to zero as $n \to +\infty$, as desired. This completes the proof of the forward part of Theorem 1.

b) $C_{\mathrm{PI|PJ}}(\lambda) \le \hat{C}_{\mathrm{PI|PJ}}$: Suppose that

$$R > \hat{C}_{\mathrm{PI|PJ}}. \quad (3.33)$$

We claim that a positive sequence $\{\gamma_n\}_{n=1}^\infty$ exists such that

$$\lambda^{\mathrm{PJ}}(C_n^*) \ge 1 - \gamma_n \quad (3.34)$$

is satisfied for all PI-admissible (n, M) random codes C_n^*, where $R \equiv (1/n)\log_2 M$ and $\gamma_n \to 0$ as $n \to +\infty$. Clearly, b) follows from (3.34).

To prove the claim, take $\epsilon > 0$ small enough so that

$$\hat{C}_{\mathrm{PI|PJ}} < \frac{1}{2}\log_2\left(1 + \frac{P_T}{N_e + P_J/(1+\epsilon)}\right) < R, \quad (3.35)$$

and let $C_n^* = \{(\boldsymbol{u}_1^*, D_1^*), \cdots, (\boldsymbol{u}_M^*, D_M^*)\}$ be any (n, M) random code satisfying PI. If the jamming sequence were $\boldsymbol{s}^* = \sqrt{P_J/(1+\epsilon)}\,\boldsymbol{\eta}^*$, then by (3.7) we know that

$$\max_{1 \le i \le M} \Pr\left\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + \sqrt{P_J/(1+\epsilon)}\,\boldsymbol{\eta}^* \in \bar{D}_i^*\right\}$$
$$\ge 1 - K'(R, P_{1+\epsilon})\exp\left\{-nE'(R, P_{1+\epsilon})\right\} \quad (3.36)$$

where $P(\cdot)$ is as defined in (3.30). Unfortunately, $\sqrt{P_J/(1+\epsilon)}\,\boldsymbol{\eta}^*$ does not satisfy PJ; therefore, we define a truncated noise process $\boldsymbol{\eta}_t^*(\epsilon)$ as follows:

$$\boldsymbol{\eta}_t^*(\epsilon) \equiv \begin{cases} \sqrt{P_J/(1+\epsilon)}\,\boldsymbol{\eta}^*, & |\boldsymbol{\eta}^*| \le \sqrt{n(1+\epsilon)} \\ \dfrac{\sqrt{nP_J}}{|\boldsymbol{\eta}^*|}\boldsymbol{\eta}^*, & |\boldsymbol{\eta}^*| \ge \sqrt{n(1+\epsilon)} \end{cases} \quad (3.37)$$

so that $\boldsymbol{\eta}_t^*(\epsilon)$ is clearly admissible under PJ. We can bound the error incurred by $\eta_i^*(\epsilon)$ below as follows:

$$\begin{aligned}\Pr\left\{u_i^* + \eta_e^* + \sqrt{P_J/(1+\epsilon)}\,\eta^* \in \bar{D}_i^*\right\} &= \Pr\left\{u_i^* + \eta_e^* + \sqrt{P_J/(1+\epsilon)}\,\eta^* \Big| |\eta^*| \le \sqrt{n(1+\epsilon)}\right\} \\ &\quad \cdot \Pr\left\{|\eta^*| \le \sqrt{n(1+\epsilon)}\right\} \\ &\quad + \Pr\left\{u_i^* + \eta_e^* + \sqrt{P_J/(1+\epsilon)}\,\eta^* \Big| |\eta^*| > \sqrt{n(1+\epsilon)}\right\} \\ &\quad \cdot \Pr\left\{|\eta^*| > \sqrt{n(1+\epsilon)}\right\} \\ &\le \Pr\left\{u_i^* + \eta_e^* + \eta_i^*(\epsilon) \in \bar{D}_i^*\right\} \\ &\quad + \Pr\left\{|\eta^*| > \sqrt{n(1+\epsilon)}\right\}. \qquad (3.38)\end{aligned}$$

Taking the maximum of (3.38) over all i and substituting (3.36), we conclude that

$$\begin{aligned}\lambda^{\mathrm{PJ}}(C_n^*) &\ge \max_{1\le i\le M} \Pr\left\{u_i^* + \eta_e^* + \eta_i^*(\epsilon) \in \bar{D}_i^*\right\} \\ &\ge 1 - K'(R, P_{1+\epsilon}) \exp\left\{-nE'(R, P_{1+\epsilon})\right\} \\ &\quad - \Pr\left\{|\eta^*| > \sqrt{n(1+\epsilon)}\right\}. \qquad (3.39)\end{aligned}$$

Using the weak law of large numbers, the right side of (3.39) tends to unity as n increases *uniformly* over all codes of rate R, which is the desired result. This completes the proof of the strong converse to Theorem 1.

Proof of Theorem 2:

a) $C_{PI|AJ}(\lambda) \ge \hat{C}_{PI|AJ}(\lambda)$: We retain the notation and results of part a) of the proof of Theorem 1. Let R nonnegative be given. Set $M_n \equiv \lfloor 2^{nR} \rfloor$, and let $\{C_n^*\}_{n=1}^{\infty}$ be the sequence of PI-admissible (n, M_n) random codes introduced in (3.19). We claim that if

$$R < \hat{C}_{\mathrm{PI|AJ}}(\lambda), \qquad (3.40)$$

then a positive sequence $\{\gamma_n\}_{n=1}^{\infty}$ exists such that

$$\lambda^{\mathrm{AJ}}(C_n^*) \le \lambda + \gamma_n \qquad (3.41)$$

and $\gamma_n \to 0$; this implies a).

To prove (3.41), let λ be such that $0 < \lambda < 1$. Suppose that (3.40) holds, and let s^* be any jamming sequence that satisfies AJ. As demonstrated in part a) of the proof of Theorem 1 (cf. (3.32)), for each $1 \le i \le M_n$

$$\Pr\left\{u_i^* + \eta_e^* + s^* \in \bar{A}_1^* \left| \frac{1}{n}\sum_{i=1}^{n} s_i^{*2} \le P_J/\lambda\right.\right\} \le c_0 K(R, P_\lambda) \exp\left\{-nE(R, P_\lambda)\right\} \qquad (3.42)$$

where $P(\cdot)$ is defined in (3.30). Since s^* satisfies AJ, Chebyshev's inequality (cf. [6]) yields

$$\Pr\left\{\frac{1}{n}\sum_{i=1}^{n} s_i^{*2} > P_J/\lambda\right\} \le \lambda. \qquad (3.43)$$

Using (3.42) and (3.43), we can bound the error probability incurred by any s^* satisfying AJ above in the following way:

$$\begin{aligned}\Pr\left\{u_i^* + \eta_e^* + s^* \in \bar{A}_i^*\right\} &= \Pr\left\{u_i^* + \eta_e^* + s^* \in \bar{A}_i^* \left| \frac{1}{n}\sum_{i=1}^{n} s_i^{*2} \le P_J/\lambda\right.\right\} \\ &\quad \cdot \Pr\left\{\frac{1}{n}\sum_{i=1}^{n} s_i^{*2} \le P_J/\lambda\right\} \\ &\quad + \Pr\left\{u_i^* + \eta_e^* + s^* \in \bar{A}_i^* \left| \frac{1}{n}\sum_{i=1}^{n} s_i^{*2} > P_J/\lambda\right.\right\} \\ &\quad \cdot \Pr\left\{\frac{1}{n}\sum_{i=1}^{n} s_i^{*2} > P_J/\lambda\right\} \\ &\le \lambda + c_0 K(R, P_\lambda) \exp\left\{-nE(R, P_\lambda)\right\}. \qquad (3.44)\end{aligned}$$

Taking the supremum over i and AJ-admissible s^*, we have

$$\lambda^{\mathrm{AJ}}(C_n^*) \le \lambda + c_0 K(R, P_\lambda) \exp\left\{-nE(R, P_\lambda)\right\}. \qquad (3.45)$$

The right side of (3.45) decreases to λ as n tends to infinity, thereby proving (3.41) and a). This concludes the proof of the forward part of Theorem 2.

b) $C_{\mathrm{PI|AJ}}(\lambda) \le \hat{C}_{\mathrm{PI|AJ}}(\lambda)$: We now prove that if

$$R > \hat{C}_{\mathrm{PI|AJ}}(\lambda), \qquad (3.46)$$

then a positive sequence $\{\gamma_n\}_{n=1}^{\infty}$ exists so that $\gamma_n \to 0$ as $n \to \infty$ and

$$\lambda^{\mathrm{AJ}}(C_n^*) \ge \lambda(1 - \gamma_n) \qquad (3.47)$$

is satisfied for all PI-admissible (n, M) random codes, where $R \equiv (1/n)\log_2 M$; b) follows from (3.47).

First, let λ be such that $0 < \lambda < 1$. Suppose that a pulsed jamming sequence, say s_λ^*, is defined to be

$$s_\lambda^* \equiv \sqrt{P_J/\lambda}\, Z_\lambda^* \eta^* \qquad (3.48)$$

where η^* is an n-vector of i.i.d $N(0,1)$ random variables and Z_λ^* is a Bernoulli random variable that is independent of η^* and distributed as follows:

$$\Pr\{Z_\lambda^* = 1\} = 1 - \Pr\{Z_\lambda^* = 0\} = \lambda. \qquad (3.49)$$

It is easy to verify that s_λ^* satisfies AJ for all $0 < \lambda \le 1$ and all $n \ge 1$.

Since λ satisfies (3.46) the error probability of C_n^* can be bounded below in the following way:

$$\begin{aligned}\lambda^{\mathrm{AJ}}(C_n^*) &\overset{\text{a)}}{\ge} \max_{1\le i\le M} \Pr\left\{u_i^* + \eta_e^* + s_\lambda^* \in \bar{D}_i^*\right\} \\ &\ge \max_{1\le i\le M} \Pr\left\{u_i^* + \eta_e^* + s_\lambda^* \in \bar{D}_i^* \big| Z_\lambda^* = 1\right\} \\ &\quad \cdot \Pr\{Z_\lambda^* = 1\} \\ &\overset{\text{b)}}{=} \lambda\left(\max_{1\le i\le M} \Pr\left\{u_i^* + \eta_e^* + \sqrt{P_J/\lambda}\,\eta^* \in \bar{D}_i^*\right\}\right) \\ &\overset{\text{c)}}{\ge} \lambda\left(1 - K'(R, P_\lambda) \exp\left\{-nE'(R, P_\lambda)\right\}\right) \\ &\qquad (3.50)\end{aligned}$$

where $P(\cdot)$ is defined in (3.30). These steps are justified in the following way: a) is an immediate consequence of the definition of $\lambda^{\mathrm{AJ}}(\cdot)$, b) follows from (3.48) and (3.49), and c) is a consequence of (3.46) and (3.7).

If (3.46) holds, then

$$K'(R, P_\lambda)\exp\{-nE'(R, P_\lambda)\} \to 0,$$

thus completing the proof of the converse to Theorem 2.

Proof of Theorem 3:

a) $C_{\mathrm{AI|PJ}}(\lambda) \geq \hat{C}_{\mathrm{AI|PJ}}(\lambda)$: Let a nonnegative R be given, and set $M_n = \lfloor 2^{nR} \rfloor$. For any $0 \leq \lambda < 1$ define a sequence of (n, M_n) random codes, say $\{C_n^*(\lambda)\}_{n=1}^{\infty}$, in the following way:

$$C_n^*(\lambda) \equiv \left\{\left(\boldsymbol{u}_1^*(\lambda), A_1^*\right), \cdots, \left(\boldsymbol{u}_{M_n}^*(\lambda), A_{M_n}^*\right)\right\} \tag{3.51}$$

where

$$\boldsymbol{u}_i^*(\lambda) \equiv \sqrt{P_T/(1-\lambda)}\, Z_{1-\lambda}^* \boldsymbol{v}_i^*. \tag{3.52}$$

$Z_{1-\lambda}^*$ is a Bernoulli random variable independent of $\boldsymbol{v}_i^*$ such that

$$\Pr\{Z_{1-\lambda}^* = 1\} = 1 - \Pr\{Z_{1-\lambda}^* = 0\} = 1 - \lambda, \tag{3.53}$$

and $C_n^{s*} = \{(\boldsymbol{v}_1^*, A_1^*), \cdots, (\boldsymbol{v}_M^*, A_{M_n}^*)\}$ is the standard (n, M_n) random code as in (3.1). It is easy to verify that $C_n^*(\lambda)$ satisfies AI for all $0 \leq \lambda < 1$ and all n. We claim that if

$$R < \hat{C}_{\mathrm{AI|PJ}}(\lambda), \tag{3.54}$$

then a positive sequence $\{\gamma_n\}_{n=1}^{\infty}$ exists such that

$$\lambda^{\mathrm{PJ}}(C_n^*(\lambda)) \leq \lambda + \gamma_n \tag{3.55}$$

and $\gamma_n \to 0$; this implies a).

The proof of this claim is in the same spirit as earlier proofs, so we will be brief. Let $\boldsymbol{s}^*$ be any PJ-admissible jamming signal, and suppose λ is such that (3.54) holds. We can then bound the error probability above as follows:

$$\begin{aligned}
&\Pr\{\boldsymbol{u}_i^*(\lambda) + \boldsymbol{\eta}_e^* + \boldsymbol{s}^* \in \bar{A}_i^*\}\\
&\quad = \Pr\left\{\boldsymbol{u}_i^*(\lambda) + \boldsymbol{\eta}_e^* + \boldsymbol{s}^* \in \bar{A}_i^* \middle| Z_{1-\lambda}^* = 0\right\}\\
&\qquad \cdot \Pr\{Z_{1-\lambda}^* = 0\}\\
&\qquad + \Pr\left\{\boldsymbol{u}_i^*(\lambda) + \boldsymbol{\eta}_e^* + \boldsymbol{s}^* \in \bar{A}_i^* \middle| Z_{1-\lambda}^* = 1\right\}\\
&\qquad \cdot \Pr\{Z_{1-\lambda}^* = 1\}\\
&\quad \overset{\text{a)}}{\leq} \lambda + \Pr\left\{\sqrt{P_T/(1-\lambda)}\, \boldsymbol{v}_i^* + \boldsymbol{\eta}_e^* + \boldsymbol{s}^* \in \bar{A}_i^*\right\}\\
&\quad \overset{\text{b)}}{\leq} \lambda + c_0 K(R, P^\lambda)\exp\{-nE(R, P^\lambda)\}
\end{aligned} \tag{3.56}$$

for all i and PJ-admissible $\boldsymbol{s}^*$, where

$$P^\lambda \equiv \frac{P_T/(1-\lambda)}{N_e + P_J}. \tag{3.57}$$

The justification of these steps is as follows: a) results when (3.53) is substituted into the preceding equation, and the first conditional probability is bounded above by one; b) follows from (3.54), (3.32), and the fact that $\boldsymbol{s}^*$ satisfies PJ. Equation (3.56) implies (3.55), thus completing the forward part of Theorem 3.

b) $C_{\mathrm{AI|PJ}}(\lambda) \leq \hat{C}_{\mathrm{AI|PJ}}(\lambda)$: We now prove that a positive sequence $\{\gamma_n\}_{n=1}^{\infty}$ exists such that $\gamma_n \to 0$, and if

$$R > \hat{C}_{\mathrm{AI|PJ}}(\lambda), \tag{3.58}$$

then

$$\lambda^{\mathrm{PJ}}(C_n^*) \geq \lambda - \gamma_n \tag{3.59}$$

is satisfied for all AI-admissible (n, M) random codes, where $R \equiv (1/n)\log_2 M$; this implies b).

To prove this, let

$$C_n^* = \{(\boldsymbol{u}_1^*, D_1^*), \cdots, (\boldsymbol{u}_M^*, D_M^*)\}$$

be any AI-admissible (n, M) random code. Take $\epsilon > 0$ small enough so that

$$R > \frac{1}{2}\log_2\left(1 + \frac{P_T/(1-\lambda)}{N_e + P_J/(1+\epsilon)}\right) + \epsilon > \hat{C}_{\mathrm{AI|PJ}}(\lambda), \tag{3.60}$$

and let $\boldsymbol{\eta}_t^*(\epsilon)$ be the PJ-admissible jamming sequence introduced in (3.37). As in part b) of the proof of Theorem 1, it is easy to show that

$$\begin{aligned}
&\lambda^{\mathrm{PJ}}(C_n^*)\\
&\quad \geq \max_{1 \leq i \leq M} \Pr\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + \boldsymbol{\eta}_t^*(\epsilon) \in \bar{D}_i\}\\
&\quad \geq \max_{1 \leq i \leq M} \Pr\left\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + \sqrt{P_J/(1+\epsilon)}\, \boldsymbol{\eta}^* \in \bar{D}_i\right\}\\
&\qquad - \Pr\left\{|\boldsymbol{\eta}^*| > \sqrt{n(1+\epsilon)}\right\}.
\end{aligned} \tag{3.61}$$

We now use Lemma 1 to obtain a lower bound for the first expression on the right side of (3.61):

$$\begin{aligned}
&\max_{1 \leq i \leq M} \Pr\left\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + \sqrt{P_J/(1+\epsilon)}\, \boldsymbol{\eta}^* \in \bar{D}_i\right\}\\
&\quad \geq \frac{1}{M}\sum_{i=1}^{M} \Pr\left\{\boldsymbol{u}_i^* + \boldsymbol{\eta}_e^* + \sqrt{P_J/(1+\epsilon)}\, \boldsymbol{\eta}^* \in \bar{D}_i^*\right\}\\
&\quad \geq \Pr\{P(U^*(C_n^*)) < (4^{R-\epsilon} - 1)(N_e + P_J/(1+\epsilon))\}\\
&\qquad - \gamma_n(\epsilon)
\end{aligned} \tag{3.62}$$

where $U^*(\cdot)$ is defined by (3.10) and $\gamma_n(\epsilon)$ is as defined in (3.12). If C_n^* satisfies AI, we can bound the first summand below as follows:

$$\begin{aligned}
&\Pr\{P(U^*(C_n^*)) < (4^{R-\epsilon} - 1)(N_e + P_j/(1+\epsilon))\}\\
&\quad \overset{\text{a)}}{\geq} 1 - \frac{P_J}{(4^{R-\epsilon} - 1)(N_e + P_J/(1+\epsilon))}\\
&\quad \overset{\text{b)}}{\geq} \lambda.
\end{aligned} \tag{3.63}$$

The justification of these steps is as follows: a) follows by applying Chebyshev's inequality using $\boldsymbol{E}P(U^*(C_n^*)) \leq P_T$, which is equivalent to AI; b) follows by observing that the right side of a) is an increasing function of R and using (3.60). Equations (3.61)–(3.63) imply (3.59), thereby completing the proof of the converse to Theorem 3.

Proof of Theorem 4:

a) $C_{\mathrm{AI|AJ}}(\lambda) \geq \hat{C}_{\mathrm{AI|AJ}}(\lambda)$*:* For any nonnegative R set $M_n \equiv \lfloor 2^{nR} \rfloor$. Fix $\epsilon > 0$, and define a sequence of AI-admissible (n, M_n) random codes, say

$$C_n^*(\epsilon) \equiv \left\{ \left(\boldsymbol{u}_1^*(\epsilon), A_1^* \right), \cdots, \left(\boldsymbol{u}_{M_n}^*(\epsilon), A_{M_n}^* \right) \right\} \quad (3.64)$$

where

$$\boldsymbol{u}_i^*(\epsilon) \equiv \sqrt{P_0^*(\epsilon)}\, \boldsymbol{v}_i^*. \quad (3.65)$$

$P_0^*(\epsilon)$ is a nonnegative random variable independent of $\{\boldsymbol{v}_1^*, \cdots, \boldsymbol{v}_M^*\}$ that satisfies $\boldsymbol{E}P_0^*(\epsilon) \leq P_T$ and whose distribution will follow. $\hat{C}_n^{s*} = \{(\boldsymbol{v}_i^*, A_i^*)\}_{i=1}^{M_n}$ is the standard (n, M_n) random code. It is easy to verify that $C_n^*(\epsilon)$ satisfies AI for all $\epsilon > 0$ and all n. We claim that if

$$R \leq \hat{C}_{\mathrm{AI|AJ}}(\lambda) - \epsilon, \quad (3.66)$$

then a positive sequence $\{\gamma_n\}_{n=1}^{\infty}$ exists such that

$$\lambda^{\mathrm{AJ}}(C_n^*(\epsilon)) \leq \lambda + \gamma_n \quad (3.67)$$

and $\gamma_n \to 0$; this implies a).

In proving this claim we assume that $N_e > 0$; the proof if $N_e = 0$ is similar. We refer the reader to Theorem 5 of the Appendix and adopt the notation used there. A consequence of this theorem (cf. (A.3)) is that if X_0 is a nonnegative random variable with distribution (A.26b) and v_0 is as defined in (A.26a), then

$$\Pr\{X_0 \geq Y + c\} \geq v_0 \quad (3.68)$$

holds for all independent nonnegative random variables Y that satisfy $\boldsymbol{E}Y \leq b$.

Now make the following substitutions:

$$a = \frac{P_T}{(4^{R+\epsilon} - 1)} \qquad b = P_J \qquad c = N_e,$$

and define $P_0^*(\epsilon)$ in (3.65) by

$$P_0^*(\epsilon) \equiv (4^{R+\epsilon} - 1) X_0.$$

With these substitutions it is easy to verify that

$$\begin{aligned} v_0 &= 1 - \hat{\lambda}^{\mathrm{AI|AJ}}(R + \epsilon) \\ &\overset{\mathrm{a)}}{\geq} 1 - \hat{\lambda}^{\mathrm{AI|AJ}}\left(\hat{C}_{\mathrm{AI|AJ}}(\lambda)\right) \\ &= 1 - \lambda \end{aligned} \quad (3.69)$$

where a) follows from (3.66) and the fact that $\hat{\lambda}^{\mathrm{AI|AJ}}(R)$ is strictly increasing. From (3.68) and (3.69) it follows that if J^* is any nonnegative random variable that satisfies $\boldsymbol{E}J^* \leq P_J$, then

$$\Pr\left\{P_0^*(\epsilon) < (4^{R+\epsilon} - 1)(N_e + J^*)\right\} \leq 1 - v_0 \leq \lambda. \quad (3.70)$$

Let $\boldsymbol{s}^*$ be any AJ-admissible jamming sequence, define $J^* = |\boldsymbol{s}^*|^2/n$ (so that $\boldsymbol{E}J^* \leq P_J$), and set $\hat{\boldsymbol{s}}^* \equiv \boldsymbol{s}^*/\sqrt{J^*}$ when $J^* > 0$ and $\hat{\boldsymbol{s}}^* \equiv 0$ otherwise (so that $|\hat{\boldsymbol{s}}^*| \leq n$ a.s.). In the proof of Theorem 1 (cf. (3.32)) we showed that if $|\hat{\boldsymbol{s}}^*| \leq n$ a.s. and P and J are positive constants, then

$$\begin{aligned} \Pr\Big\{\sqrt{P_0^*(\epsilon)}\, v_i^* + \eta_e^* & \\ + \sqrt{J^*}\,\hat{s}^* \in \bar{A}_i^* \,\big|\, P_0^*(\epsilon) = P,\ J^* = J\Big\}& \\ \leq c_0 K(R, P') \exp\{-nE(R, P')\}& \end{aligned} \quad (3.71)$$

for all $n \geq 1$ provided that

$$P' \equiv \frac{P}{N_e + J} > (4^R - 1).$$

In particular, if

$$\frac{P}{N_e + J} \geq (4^{R+\epsilon} - 1), \quad (3.72)$$

then using (3.5b) we can further upper-bound the right side of (3.71) by

$$\bar{B}_n(R, \epsilon) \equiv c_0 K(R, 4^{R+\epsilon} - 1) \exp\left\{-nE(R, 4^{R+\epsilon} - 1)\right\}. \quad (3.73)$$

Note that $\bar{B}_n(R, \epsilon) \to 0$ for all $\epsilon > 0$. Now define

$$h_n(P, J) \equiv \begin{cases} \bar{B}_n(R, \epsilon), & P \geq (4^{R+\epsilon} - 1)(N_e + J) \\ 1, & \text{otherwise} \end{cases} \quad (3.74)$$

so that $h_n(P, J)$ is an upper bound for the left-hand quantity in (3.71) for *all* P, J, and n. Averaging this bound over the distributions of $C_n^*(\epsilon)$ and J^*, we find that

$$\begin{aligned} \Pr\{\boldsymbol{u}_i^*(\epsilon) + \boldsymbol{\eta}_e^* + \boldsymbol{s}^* \in \bar{A}_i^*\} &= \Pr\left\{\sqrt{P_0^*(\epsilon)}\, \boldsymbol{v}_i^* + \boldsymbol{\eta}_e^* + \sqrt{J^*}\,\hat{\boldsymbol{s}}^* \in \bar{A}_i^*\right\} \\ &\leq \boldsymbol{E}h_n(P_0^*(\epsilon), J^*) \\ &= \bar{B}_n(R, \epsilon) + \left(1 - \bar{B}_n(R, \epsilon)\right) \\ &\quad \cdot \Pr\left\{P_0^*(\epsilon) < (4^{R+\epsilon} - 1)(N_e + J^*)\right\} \\ &\leq \bar{B}_n(R, \epsilon) + \lambda \end{aligned} \quad (3.75)$$

where the last inequality follows from (3.70). Taking the supremum of (3.75) over all i and AJ-admissible $\boldsymbol{s}^*$, we obtain the bound

$$\lambda^{\mathrm{AJ}}(C_n^*(\epsilon)) \leq \bar{B}_n(R, \epsilon) + \lambda \quad (3.76)$$

for all $\epsilon > 0$ satisfying (3.66) and $n \geq 1$. This completes the proof of the forward part of Theorem 4.

b) $C_{\mathrm{AI|AJ}}(\lambda) \leq \hat{C}_{\mathrm{AI|AJ}}(\lambda)$*:* We now prove that for each $\epsilon > 0$ and $R \geq 0$ a positive sequence $\{\gamma_n\}_{n=1}^{\infty}$ exists so that $\gamma_n \to 0$ and

$$\lambda^{\mathrm{AJ}}(C_n^*) \geq \lambda - \gamma_n \quad (3.77)$$

holds for any AI-admissible (n, M) random code C_n^* where $R \equiv (1/n)\log_2 M$ is such that

$$R > \hat{C}_{\mathrm{AJ}}(\lambda) + \epsilon; \quad (3.78)$$

this implies b).

As in part a) of the proof of Theorem 4, we invoke Theorem 5 of the Appendix. This theorem implies that if

Y_0 has the distribution (A.26c) and v_0 is as defined in (A.26a), then

$$\Pr\{X \geq Y_0 + c\} \leq v_0 \tag{3.79}$$

holds for all independent nonnegative random variables X that satisfy $EX \leq a$. Making the substitution

$$a = \frac{P_T}{(4^{R-\epsilon}-1)} \qquad b = P_J \qquad c = N_e$$

and defining

$$J_0^*(\epsilon) \equiv Y_0$$
$$P^* \equiv (4^{R-\epsilon}-1)X,$$

we obtain

$$v_0 = 1 - \hat{\lambda}^{\mathrm{AI|AJ}}(R-\epsilon) \leq 1 - \lambda$$

where we have used (3.78) and the fact that $\hat{\lambda}^{\mathrm{AI|AJ}}(\cdot)$ is strictly increasing. Therefore,

$$\Pr\{P^* < (4^{R-\epsilon}-1)(N_e + J_0^*(\epsilon))\} \geq 1 - v_0 \geq \lambda \tag{3.80}$$

holds for all P^* satisfying

$$EP^* \leq P_T. \tag{3.81}$$

Note that $\sqrt{J_0^*(\epsilon)}\,\eta^*$ is AI-admissible for all $\epsilon > 0$.

Let C_n^* be any (n, M) random code. We may bound the error probability of this code from below as follows:

$$\begin{aligned}\lambda^{\mathrm{AJ}}(C_n^*) &\geq \max_{1\leq i\leq M} \Pr\left\{u_i^* + \eta_e^* + \sqrt{J_0^*(\epsilon)}\,\eta^* \in \bar{D}_i^*\right\}\\ &\geq \frac{1}{M}\sum_{i=1}^{M} \Pr\left\{u_i^* + \eta_e^* + \sqrt{J_0^*(\epsilon)}\,\eta^* \in \bar{D}_i^*\right\}\\ &\overset{a)}{\geq} \Pr\left\{P(U^*(C_n^*)) < (4^{R-\epsilon}-1)(N_e + J_0^*(\epsilon))\right\} - \gamma_n(\epsilon)\\ &\overset{b)}{\geq} \lambda - \gamma_n(\epsilon)\end{aligned} \tag{3.82}$$

where $\gamma_n(\epsilon)$ is as defined in (3.12). The justification of these steps is as follows: a) results by applying Lemma 1; b) follows from (3.80) and the fact that $EP(U^*(C_n^*)) \leq P_T$. This completes the proof of the converse of Theorem 4.

IV. Discussion

Our results show that the asymptotic behavior of GAVC's is qualitatively different from that of discrete AVC's: whereas the latter always have a random coding capacity (cf. Blackwell *et al.* [10]), the former generally have no capacity (except in the case PI|PJ). This is a direct consequence of the imposition of power constraints of the *average* type; in particular, it is not due to the fact that the GAVC has real input and output alphabets. In fact, a discrete AVC with an average cost structure will also generally fail to have a capacity [16].

It remains to determine the corresponding λ-capacities, if they exist, for the GAVC when the transmitter is restricted to *deterministic* codes (i.e., those of the form (2.2)). However, we an make an interesting observation: Blachman [7], [8] has investigated the use of the GAVC PI|PJ (our terminology) when the transmitter uses deterministic codes and the maximum probability of error concept. Here the transmitter's signal (which is deterministic) is constrained by

$$\frac{1}{n}\sum_{i=1}^{n} u_i^2 \leq P_T,$$

and the jammer may transmit any deterministic sequence $\{s_i\}$ that satisfies

$$\frac{1}{n}\sum_{i=1}^{n} s_i^2 \leq P_J.$$

Blachman obtained the following bound on the capacity of this channel: if the CAVC PI|PJ has a deterministic capacity, say $C^d_{\mathrm{PI|PJ}}$, then

$$D^d_{\mathrm{PI|PJ}} \leq C_b \tag{4.1}$$

where

$$C_b = \begin{cases} \dfrac{1}{2}\log\left(\dfrac{P_T}{N_e + P_J}\right), & P_T P_J > (P_J + N_e)^2\\[2ex] \dfrac{1}{2}\log_2\left(1 + \dfrac{\left(\sqrt{P_T} - \sqrt{P_J}\right)^2}{N_e}\right), & P_J^2 < P_T P_J \leq (P_J + N_e)^2\\[2ex] 0, & P_T \leq P_J\end{cases}$$

Comparing (2.14) with the upper bound in (4.1), we find that

$$C_b < C_{\mathrm{PI|PJ}}$$

for all (positive) values of N_e, P_T, and P_J. We conclude that, as has previously been shown for discrete AVC's, the class of random codes outperforms the class of deterministic codes for the GAVC PI|PJ. For the discrete AVC the deterministic coding capacities are known in many special cases. Ahlswede [2], using the average probability of error concept, has shown that the capacity of the discrete AVC is either equal to the random coding capacity, or else it is zero. This method apparently fails for the GAVC, owing to the presence of a cost structure on the allowable channels and encoders.

The coding problems of Section II may be case in an alternative game theoretic formulation. Corresponding to each GAVC, say $A|B$, a family of two-player zero-sum games exist (cf. Blackwell and Girshick [9]) defined as follows. Fix the block length n and the source rate R. The transmitter's (resp. jammer's) *allowable strategies* consist of all $(n, 2^{nR})$ random codes C_n^* (resp. $\boldsymbol{R}^n$-valued random vectors s^*) that satisfy the power constraint A (resp. B). The payoff, when the jammer plays s^* and the transmitter plays C_n^*, is the error probability $\lambda(C_n^*, s^*)$ defined in (2.8). The jammer wants to maximize this probability; the transmitter wants to minimize it. Therefore, they seek strategies that attain the outer extrema in the following programs:

$$\text{transmitter's program: } \bar{\nu}_n \equiv \inf_{C_n^*} \sup_{s^*} \lambda(C_n^*, s^*) \quad (4.2\text{a})$$

$$\text{jammer's program: } \nu_n \equiv \sup_{s^*} \inf_{C_n^*} \lambda(C_n^*, s^*) \quad (4.2\text{b})$$

where the extrema are taken over all allowable s^* and C_n^*. An optimal strategy for the transmitter (resp. jammer), if it exists, is one that attains the outer extrema in the transmitter's (resp. jammer's) program. For any $\epsilon > 0$, *ϵ-optimal strategies* $C_{n\epsilon}^*$ and s_ϵ^* are allowable strategies for which

$$\sup_{s^*} \lambda(C_{n\epsilon}^*, s^*) \le \bar{\nu}_n + \epsilon \quad (4.3)$$

$$\inf_{C_n^*} \lambda(C_n^*, s_\epsilon^*) \ge \nu_n - \epsilon \quad (4.4)$$

where the extrema are taken over all allowable s^* and C_n^*. It is always true that $\nu_n \le \bar{\nu}_n$; if $\nu_n = \bar{\nu}_n$, then the game is said to have a *value* $\nu_{on} \equiv \nu_n = \bar{\nu}_n$.

Equation (4.2a) defined a sequence (in n) of communication games. Başar and Wu [4] have considered games of this type (for the constraints AI|AJ) for a memoryless Gaussian source and for a different cost function, viz., mean-square distortion. For each n they obtain the value of the game and characterize saddle-point strategies for each player. In contrast, we can say little about each game in the sequence; we can, however, say a great deal about the *asymptotic behavior* of the sequence.

Implicit in the proofs of Theorems 1–4 is the following result. The sequences $\{\nu_n\}_{n=1}^{\infty}$ and $\{\bar{\nu}_n\}_{n=1}^{\infty}$ converge and

$$\lim_{n\to+\infty} \nu_n = \lim_{n\to+\infty} \bar{\nu}_n = \hat{\lambda}^{A|B}(R) \quad (4.5)$$

holds for every R and every pair of constraints $A|B$. Thus the sequence of games has an "asymptotic value" equal to $\hat{\lambda}^{A|B}(R)$. Furthermore, for all $\epsilon > 0$, there exist for all sufficiently large n, ϵ-optimal strategies for both the transmitter and the jammer. (Such strategies for the transmitter are explicitly constructed in the forward parts of the proofs in Section III; jamming strategies are constructed in the converse parts.)

Some authors further constrain the jammer to signals of the form

$$s^* = (z_1^*\eta_1^*, \cdots, z_n^*\eta_n^*) \quad (4.6)$$

where $\{\eta_i^*\}_{i=1}^n$ is i.i.d. $N(0,1)$ and $\{z_i^*\}_{i=1}^n$ is a sequence of nonnegative random variables independent of $\{\eta_i^*\}_{i=1}^n$ and subject only to the average power constraint

$$\boldsymbol{E}\left\{\frac{1}{n}\sum_{i=1}^{n} z_i^{*2}\right\} \le P_J.$$

We call this constraint AJG and use the notation GAVC A|AJG to refer to the channel with input constraint A and jamming power constraint AJG. Since AJG is more restrictive than AJ, we must have $\boldsymbol{R}_{A|\mathrm{AJG}} \supset R_{A|\mathrm{AJ}}$. However, the jamming strategies constructed in the converses to Theorems 2 and 4 are all of the form (4.6), so that we must have $\boldsymbol{R}_{A|\mathrm{AJG}} = R_{A|\mathrm{AJ}}$ and, consequently,

$$\lambda^{A|\mathrm{AJG}}(R) = \lambda^{A|\mathrm{AJ}}(R). \quad (4.7)$$

Thus our results remain valid in the special case of Gaussian jammers.

It is especially interesting that the achievable regions of Theorem 2–4 are not determined solely by a simple optimization program involving mutual information as is usually the case in information theory. McEliece and Stark [17] have modeled the conflict between transmitter and jammer when coding is used by a two-player zero-sum game with *mutual information* as the pay-off function (see also [11]). As an example, they considered the channel that we have called the GAVC AI|AJ (for the special case $N_e = 0$) and obtained the following results. Optimal transmission strategies for both players are i.i.d. Gaussian sequences of maximum power and of length n, and the value (or optimal payoff) is

$$\frac{n}{2}\log_2\left(1 + \frac{P_T}{P_J}\right).$$

If the value of the game considered by McEliece–Stark is actually the capacity of the channel (the authors do not assert that it is), then it carries the following interpretation: when n is large and

$$R < \frac{1}{2}\log_2\left(1 + \frac{P_T}{P_J}\right),$$

then $\lambda^{\mathrm{AJ}}(C_n^*) \approx 0$ is possible. In contrast, however, note that the ϵ-optimal strategies for the game AI|AJ in (4.2a) (cf. proof of Theorem 4) are *not* memory-less, and the error probability of any positive rate code is bounded away from zero. It is of considerable interest that these two apparently related games lead to such different results.

An explanation of this disparity between the predictions of these two games lies in the fact that mutual information takes on operational significance only when the block length is large compared to the memory of the channel. The error probability formulation (cf. (4.2a)) allows the jamming memory to equal the block length, whereas the mutual information formulation always assumes that the block length of the code is large compared to the jamming memory. Therefore, the game involving mutual information gives an *a priori* advantage to the transmitter, and it is not surprising that this approach leads to much more optimistic results for the transmitter. We conclude that, at least for GAVC's, one cannot in general attribute a coding significance to games having mutual information as a payoff function.

From a practical viewpoint the results of this paper may be difficult to achieve or may lack meaning for a real jammer. Like the pulse-jamming signals considered by Houston [15] our ϵ-optimal strategies demand high peak power when R is small; unlike Houston's, however, this peak power must be sustained over the block length of the code. When n is large, the average power constraints (AI, AJ) may fail to reflect all the physical constraints that would limit a practical system. As an extreme example let $n \to +\infty$. Then the optimal jamming strategy for the case PI|AJ is of the form $s_i^* \sim N(0, P_J/\rho)$ for *all time* with probability ρ, and $s_i^* = 0$ for *all time* with probability $1-\rho$. One may approach a more realistic situation by considering multiple constraints on the jammer (as discussed in Section II).

Acknowledgment

The authors would like to thank A. Ephremides for a number of helpful discussions and I. Csiszár for several constructive suggestions which led to a considerable simplification of the proofs and an improved presentation of our ideas. This work was carried out while the first author was a Fellow at the Information Technology Division of the Naval Research Laboratory. It is a pleasure to acknowledge the excellent working conditions there and to thank D. McGregor and J. Wieselthier of the Naval Research Laboratory for many stimulating discussions. Finally, the second author is grateful to I. Csiszár for several illuminating discussions while the latter was visiting the University of Maryland in Fall 1985.

Appendix

In this Appendix we study the following two-player zero-sum game (cf. Blackwell and Girshick [9]). Let a, b, and c be real numbers such that $a, b > 0$ and $c \geq 0$. Player I's (respectively, player II's) allowable strategies consist of all nonnegative real-valued random variables X (resp. Y) satisfying $EX \leq a$ (resp. $EY \leq b$).[11] The payoff to player I, when I plays X and II plays Y is

$$\Pr\{X \geq Y + c\}. \tag{A.1}$$

Player I wishes to maximize (A.1); player II wants to minimize it. Therefore, players I and II seek independent strategies that attain the outer extrema in the following programs:

$$\text{program I: } \boldsymbol{\nu} = \sup_{X:\, EX \leq a} \inf_{Y:\, EY \leq b} \Pr\{X \geq Y + c\} \tag{A.2a}$$

$$\text{program II: } \bar{\boldsymbol{\nu}} = \inf_{Y:\, EY \leq b} \sup_{X:\, EX \leq a} \Pr\{X \geq Y + c\}. \tag{A.2b}$$

If a strategy exists that attains the outer extrema for program I (resp. II), it is called an *optimal strategy* for player I (resp. II). It is always true that $\bar{\boldsymbol{\nu}} \geq \boldsymbol{\nu}$; if $\bar{\boldsymbol{\nu}} = \boldsymbol{\nu}$, then the game is said to have a *value* $v_0 = \bar{\boldsymbol{\nu}} = \boldsymbol{\nu}$. A saddle-point solution to this game (if it exists) is a pair of allowable strategies, say (X_0, Y_0), such that

$$\Pr\{X \geq Y_0 + c\} \leq \Pr\{X_0 \geq Y_0 + c\} \leq \Pr\{X_0 \geq Y + c\} \tag{A.3}$$

is satisfied for all allowable (X, Y). The existence of a saddle point is a sufficient condition for a value to exist; in this case we have

$$v_0 = \bar{\boldsymbol{\nu}} = \boldsymbol{\nu} = \Pr\{X_0 \geq Y_0 + c\}, \tag{A.4}$$

and thus X_0 (resp. Y_0) is an optimal strategy for player I (resp. player II).

In this appendix we derive a unique saddle-point solution (A.2a). The special case $a = b = 1$, $c = 0$ has been studied by Bell and Cover [5] in connection with competitive investment and the special case $c = 0$ by McEliece and Rodemich [18] as part of a study of optimal jamming of uncoded multiple frequency shift keying (MFSK). We construct the general solution of (A.2a) from the known solution in the special case $c = 0$. Without many of the complications that arise in the MFSK problem studied in [18], this special case admits a proof that is much simpler than that given in [18]; we present this as follows.

Lemma 2 (Bell–Cover–McEliece–Rodemich): Consider the two-player zero-sum game given by (A.2a) when $c = 0$. This game has a value v_0 and unique saddle-point strategies $X_0 \sim F_0$ and $Y_0 \sim G_0$. These are given in the case $a \geq b$ by[12]

$$v_0 = 1 - \frac{b}{2a} \tag{A.5a}$$

$$F_0(x) = U_{[0,2a]}(x) \tag{A.5b}$$

$$G_0(x) = \left(\frac{b}{a}\right) U_{[0,2a]}(x) + \left(1 - \frac{b}{a}\right) \Delta_0(x); \tag{A.5c}$$

and if $a < b$,

$$v_0 = \frac{a}{2b} \tag{A.5d}$$

$$F_0(x) = \left(\frac{a}{b}\right) U_{[0,2b]}(x) + \left(1 - \frac{a}{b}\right) \Delta_0(x) \tag{A.5e}$$

$$G_0(x) = U_{[0,2b]}(x). \tag{A.5f}$$

Remark: The proof given here is a generalization of Bell and Cover's [5].

[11] In this Appendix we abandon the convention used earlier in the paper that distinguishes random variables with asterisks.

[12] Throughout this Appendix we use the following notation: $X \sim F$ means that the real-valued random variable X has distribution function F. We denote by $U_{[a,b]}(x)$ the distribution function of a random variable that is uniformly distributed on the interval $[a, b]$ and denote by $\Delta_c(x)$ the distribution function of the trivial random variable $X \equiv c$.

Proof: Let $X \sim F$ and $Y \sim G$ be any allowable strategies. Observe that

$$\Pr\{X \geq Y\} = \int_0^\infty G(X)\,dF(x) = 1 - \int_0^\infty F(x-)\,dG(x). \tag{A.6}$$

First consider the case $a \geq b$. Let us show that (X_0, Y_0) satisfies (A.3) when $c = 0$. Using the obvious inequality $U_{[0,d]}(x) \leq x/d$ when $x \geq 0$, we then obtain

$$\begin{aligned}\Pr\{X \geq Y_0\} &= \int_0^\infty G_0(x)\,dF(x)\\ &= \left(1 - \frac{b}{a}\right) + \frac{b}{a}\int_0^\infty U_{[0,2a]}(x)\,dF(x)\\ &\leq \left(1 - \frac{b}{a}\right) + \frac{b}{2a^2}\int_0^\infty x\,dF(x)\\ &\leq 1 - \frac{b}{2a} = v_0.\end{aligned} \tag{A.7}$$

In much the same way, using the right-most equality in (A.6), we can show

$$\Pr\{X_0 \geq Y\} \geq v_0. \tag{A.8}$$

Since $\Pr\{X_0 \geq Y_0\} = v_0$, we conclude that (X_0, Y_0) is a saddle point and v_0 is the value of the game.

To complete the proof in the case $a \geq b$, it only remains to show the uniqueness of F_0 and G_0. First consider G_0. Let $Y_0' \sim G_0'$ be any other random variable such that $EY_0' \leq b$ and

$$\Pr\{X \geq Y_0'\} \leq v_0 \tag{A.9}$$

for all admissible X. Substitution of

1) $X \sim U_{[0,2a]}(x)$

2) $X \sim \left(\frac{\beta}{\alpha+\beta}\right)\Delta_{a-\alpha}(x) + \left(\frac{\alpha}{\alpha+\beta}\right)\Delta_{a+\beta}(x)$

for all $0 \leq \alpha, \beta \leq a$ into (A.9) yields, respectively,

1′) $G_0'(2a) = 1$

2′) $\left(\frac{\beta}{\alpha+\beta}\right)G_0'(a-\alpha) + \left(\frac{\alpha}{\alpha+\beta}\right)G_0'(a+\beta) \leq v_0$

for all $0 \leq \alpha, \beta \leq a$.

We claim that 2′) implies that a line exists, say $l(x)$, that passes through the point (a, v_0) and is such that

$$G_0'(x) \leq l(x) \tag{A.10}$$

for all $x \geq 0$. To prove this claim, define[13]

$$\mu \equiv \max_{0 \leq \beta \leq a} \frac{G_0'(a+\beta) - v_0}{\beta} < +\infty, \tag{A.11}$$

and let $\bar{\beta}$ attain the maximum. Let $l(x)$ be the line through (a, v_0) having slope μ. We know that $G_0'(a) \leq v_0 = l(a)$ (proof: take $\alpha = \beta = 0$ in (2′)). By construction $l(x)$ satisfies (A.10) when $x \geq a$ and passes through the point $(a + \bar{\beta}, G_0'(a + \bar{\beta}))$. Now if

$$G_0'(a-\alpha) \geq l(a-\alpha) \tag{A.12}$$

for some $0 \leq \alpha \leq a$, then α and $\bar{\beta}$ violate 2′). Therefore, to avoid a contradiction, $l(x)$ must satisfy (A.10) for $0 \leq x \leq a$ as well, proving the claim.

We now show that (A.10) implies that $G_0' = G_0$. For any measurable function, say $f(x)$, let ν_f denote the Lebesgue volume of the region in $\mathbf{R}^2$ comprising the points $R_f = \{(x,y) | 0 \leq x \leq 2a, f(x) \leq y \leq 1)\}$. By an elementary fact of probability theory and 1′) we know that

$$\nu_{G_0'} = EY_0' \leq b. \tag{A.13}$$

Equation (A.10) implies that $\nu_{G_0'} \geq \nu_l$ and hence

$$\nu_l \leq b. \tag{A.14}$$

Since $l(0) \geq G_0'(0) \geq 0$, $l(2a) \geq G_0'(2a) = 1$, and $l(a) = v_0$, R_l is a triangular region and $l(0)$ must be such that $0 \leq l(0) \leq 2v_0 - 1$. By elementary geometry we can show that

$$\nu_l = \frac{a(1 - l(0))^2}{2(v_0 - l(0))} \tag{A.15}$$

for all $0 \leq l(0) \leq 2v_0 - 1$. It is easy to show that (A.15) is a strictly decreasing function of $l(0)$ that attains a minimum value of $\nu_l = b$ when $l(0) = 2v_0 - 1$. Therefore, the only line $l(x)$ that passes through (a, v_0) and that does not contradict (A.14) satisfies $l(0) = 2v_0 - 1$, and hence

$$l(x) = \frac{bx}{2a^2} + \left(1 - \frac{b}{a}\right). \tag{A.16}$$

Comparing (A.16) with (A.5c), we see that l equals G_0 for all x such that $0 \leq x \leq 2a$ and $0 \leq l(x) \leq 1$. It follows from (A.10), the nonnegativity of Y_0' and Y_0, and (1) that

$$G_0'(x) \leq G_0(x)$$

for all real x. This implies that $G_0' \equiv G_0$, since if $G_0'(x) < G_0(x)$ for some $0 < x \leq 2a$, then

$$EY_0' = \nu_{G_0'} > \nu_{G_0} = b,$$

a contradiction. We conclude that in the case $a \geq b$, G_0 is unique. The proof that F_0 is unique, and the proofs for the case $a < b$ are similar. This completes the proof of Lemma 2.

We now consider the game (A.2a) when $c > 0$ and show that the solution in this case can be constructed from the known solution for the case $c = 0$. To see this, note that any nonnegative $X \sim F$ can be decomposed in the following way:

$$X = \begin{cases} c + Z, & \text{with probability } p \\ W, & \text{with probability } 1 - p \end{cases} \tag{A.17}$$

where $p = 1 - F(c-)$ and $W \sim L$ and $Z \sim H$ are independent nonnegative real-valued random variables. The distribution functions L and H are given by

$$L(x) = \begin{cases} \dfrac{F(x)}{F(c-)}, & -\infty < x < c \\ 1, & x \geq c \end{cases}$$

if $F(c-) > 0$; otherwise $L(x) = \Delta_0(x)$, and

$$H(x) = \begin{cases} 0, & -\infty < x < 0 \\ \dfrac{F(x+c) - F(c-)}{1 - F(c-)}, & x \geq 0 \end{cases}$$

if $F(c-) < 1$; otherwise $H(x) = \Delta_0(x)$.

[13] The "max" in (A.11) is justified because $(G_0'(a+\beta) - v_0)/\beta$ is upper semicontinuous, and the right-hand inequality is justified because this function is bounded by v_0/a (to prove: take $\alpha = a$ in (2′)).

In terms of the new variables p, Z, and W the cost function (A.1) becomes

$$\begin{aligned}\Pr\{X \geq Y + c\} &= p \Pr\{Z + c \geq Y + c\} \\ &\quad + (1-p)\Pr\{W \geq Y + c\} \\ &= p \Pr\{Z \geq Y\}. \qquad \text{(A.18)}\end{aligned}$$

Clearly, W has no effect on the cost function $\Pr\{X \geq Y + c\}$; only our choice of p and Z influence it. The latter choice is constrained by

$$EX = (1-p)EW + p(c + EZ) \leq a$$

or

$$EZ \leq \frac{a - (1-p)EW}{p} - c$$

so that the widest choice of Z is permitted when $W \equiv 0$ and

$$EZ \leq \frac{a}{p} - c \equiv \hat{a}(p).$$

Using this decomposition, we can reformulate (A.2a) in the following way:

$$\text{program I: } v = \sup_{(p,Z):\, EZ \leq \hat{a}(p)} \; \inf_{Y:\, EY \leq b} p \Pr\{Z \geq Y\} \qquad \text{(A.19a)}$$

$$\text{program II: } \bar{v} = \inf_{Y:\, EY \leq b} \; \sup_{(p,Z):\, EZ \leq \hat{a}(p)} p \Pr\{Z \geq Y\}. \qquad \text{(A.19b)}$$

Games (A.2a) and (A.19a) are equivalent in the following sense: if X_0, p_0, and Z_0 are related as in (A.17), then $\{(p_0, Z_0), Y_0\}$ is a saddle point for (A.19a) *if and only if* (X_0, Y_0) is a saddle point for (A.2a); of course, the resulting values of both games are the same. Therefore, solving (A.19a) is entirely equivalent to solving (A.2a).

Using (A.19a), we can derive the only candidate saddle point for (A.2a) in the following way. Suppose that $\{(p_0, Z_0), Y_0\}$ is a saddle point so that

$$\begin{aligned} p \Pr\{Z \geq Y_0\} &\leq p_0 \Pr\{Z_0 \geq Y_0\} \\ &\leq p_0 \Pr\{Z_0 \geq Y\} \qquad \text{(A.20)}\end{aligned}$$

for all admissible $\{(p, Z), Y\}$. Then, in particular, we have

$$\begin{aligned} p_0 \Pr\{Z \geq Y_0\} &\leq p_0 \Pr\{Z_0 \geq Y_0\} \\ &\leq p_0 \Pr\{Z_0 \geq Y\} \qquad \text{(A.21)}\end{aligned}$$

for all (Z, Y) such that $\{(p_0, Z), Y\}$ is allowable. Ignoring momentarily the trivial possibility that $p_0 = 0$, (A.21) implies that (Z_0, Y_0) must be a saddle point of (A.2a) with constants

$$a' = \hat{a}(p_0) \equiv \frac{a}{p_0} - c \qquad b' = b \qquad c' = 0. \qquad \text{(A.22)}$$

Since (A.5a) gives the unique solution to (A.2a) when $c = 0$, we conclude that (Z_0, Y_0) must have the distributions F_0 and G_0 obtained when the constants (A.22) are substituted into (A.5a). The corresponding value of this game as a function of p_0 is

$$v_0(p_0) \equiv \begin{cases} p_0\left(1 - \dfrac{b}{2\hat{a}(p_0)}\right), & \hat{a}(p_0) \geq b \\[2ex] \dfrac{p_0 \hat{a}(p_0)}{2b}, & \hat{a}(p_0) < b. \end{cases} \qquad \text{(A.23)}$$

We now show that (A.20) fixes a value for p_0 as well. If $\{(p_0, Z_0), Y_0\}$ is a saddle point for (A.19a) and v_0 is the corresponding value, then the left bound in (A.20) implies that

$$v_0 = \max_{0 \leq p \leq 1} v_0(p).$$

Using this, we may explicitly find the only possible saddle point. The following facts will be useful.

1) The maximum of $v_0(p)$ over the range $0 \leq p \leq 1$ is attained uniquely by

$$p_0 = \begin{cases} \dfrac{a}{c}\left(1 - \sqrt{\dfrac{b}{2c + b}}\right), & a \leq c + \dfrac{b}{2}\left[1 + \sqrt{1 + \dfrac{2c}{b}}\right] \\[2ex] 1, & a > c + \dfrac{b}{2}\left[1 + \sqrt{1 + \dfrac{2c}{b}}\right]. \end{cases} \qquad \text{(A.24)}$$

Note that $p_0 \leq a/c$ when $c > 0$.

2) Define $g(p)$ on the interval $0 \leq p \leq a/c$ by

$$g(p) = 1 - \frac{b}{\hat{a}(p)} - \frac{bc}{2\hat{a}^2(p)}.$$

Then $g(p_0) = 0$ if $0 \leq p_0 < 1$, and $g(p_0) \geq 0$ if $p_0 = 1$.

3) $\hat{a}(p_0) \geq b$ for all $a, b > 0$ and $c > 0$, where p_0 is as defined in (A.24).

Therefore, based on facts (1) and (3), Lemma 2, and the foregoing comments, the only possible saddle point for the game (A.19a) is p_0, $Z_0 \sim H_0$, and $Y_0 \sim G_0$, where p_0 is given in (A.24) and

$$H_0(x) = U_{[0, 2\hat{a}(p_0)]}(x) \qquad \text{(A.25a)}$$

$$\begin{aligned} G_0(x) &= \left(\frac{b}{\hat{a}(p_0)}\right) U_{[0, 2\hat{a}(p_0)]}(x) \\ &\quad + \left(1 - \frac{b}{\hat{a}(p_0)}\right) \Delta_0(x). \qquad \text{(A.25b)}\end{aligned}$$

Remark: Note that $a > 0$ implies that $\hat{a}(p_0) \equiv (a/p_0) - c > 0$, so that (A.25b) is always well defined.

H_0 and G_0 are obtained by substituting p_0 into (A.22), substituting the resulting constants into (A.5a), and taking $H_0 \equiv F_0$. The corresponding value of the game is

$$v_0 = \begin{cases} \dfrac{a}{c}\left[1 + \dfrac{b}{c}\left(1 - \sqrt{1 + \dfrac{2c}{b}}\right)\right], & a \leq c + \dfrac{b}{2}\left[1 + \sqrt{1 + \dfrac{2c}{b}}\right] \\[2ex] 1 - \dfrac{b}{2(a - c)}, & a > c + \dfrac{b}{2}\left[1 + \sqrt{1 + \dfrac{2c}{b}}\right]. \end{cases}$$

We have shown that $\{(p_0, Z_0), Y_0\}$ is the only candidate for a saddle point for the game (A.19a); let us now verify that this is indeed a saddle point. Let $\{(p, Z), Y\}$ be any admissible triple, and suppose that $Z \sim H$ and $Y \sim G$. Then

$$\begin{aligned} &p \Pr\{Z \geq Y_0\} \\ &\quad = p \int_0^\infty G_0(x)\, dH(x) \\ &\quad = p\left(1 - \frac{b}{\hat{a}(p_0)}\right) + \frac{bp}{\hat{a}(p_0)} \int_0^\infty U_{[0, 2\hat{a}(p_0)]}(x)\, dH(x) \\ &\quad \leq p\left(1 - \frac{b}{\hat{a}(p_0)}\right) + \frac{bp}{2\hat{a}^2(p_0)} \int_0^\infty x\, dH(x) \\ &\quad \leq p\left(1 - \frac{b}{\hat{a}(p_0)}\right) + \frac{bp\hat{a}(p)}{2\hat{a}^2(p_0)} \\ &\quad = p\left(1 - \frac{b}{\hat{a}(p_0)} - \frac{bc}{2\hat{a}^2(p_0)}\right) + \frac{ba}{2\hat{a}^2(p_0)} \\ &\quad = pg(p_0) + \frac{ba}{2\hat{a}^2(p_0)}. \end{aligned}$$

From fact 2) it follows that $pg(p_0) \leq p_0 g(p_0)$, and therefore,

$$p \Pr\{Z \geq Y_0\} \leq p_0 g(p_0) + \frac{ba}{2\hat{a}^2(p_0)} = p_0\left[1 - \frac{b}{2\hat{a}(p_0)}\right] = v_0.$$

The proof of

$$p_0 \Pr\{Z_0 \geq Y\} \geq v_0$$

for all allowable Y is similar to the proof of Lemma 2 and so is omitted.

We conclude that $\{(p_0, Z_0), Y_0\}$ is the unique saddle-point for (A.19a) and that v_0 is the corresponding value. Recalling the equivalence between the games (A.19a) and (A.2a) when p, Z, and X are related by (A.17) (cf. remarks following (A.19a)), we have, therefore, proved the following theorem.

Theorem 5: Consider the two-player zero sum game given by (A.2a). This game has a value v_0 and unique saddle-point strategies $X_0 \sim F_0$ and $Y_0 \sim G_0$. These are given in Lemma 2 for the case $c = 0$ and for the case $c > 0$ by

$$v_0 = \begin{cases} \dfrac{a}{c}\left[1 + \dfrac{b}{c}\left(1 - \sqrt{1 + \dfrac{2c}{b}}\right)\right], & a \leq c + \dfrac{b}{2}\left[1 + \sqrt{1 + \dfrac{2c}{b}}\right] \\ 1 - \dfrac{b}{2(a-c)}, & a > c + \dfrac{b}{2}\left[1 + \sqrt{1 + \dfrac{2c}{b}}\right] \end{cases} \tag{A.26a}$$

$$F_0(x) = p_0 U_{[0, 2\hat{a}(p_0)]}(x - c) + (1 - p_0)\Delta_0(x) \tag{A.26b}$$

$$G_0(x) = \left(\frac{b}{\hat{a}(p_0)}\right) U_{[0, 2\hat{a}(p_0)]}(x) + \left(1 - \frac{b}{\hat{a}(p_0)}\right)\Delta_0(x) \tag{A.26c}$$

where $\hat{a}(p) = a/p - c$ and

$$p_0 = \begin{cases} \dfrac{a}{c}\left(1 - \sqrt{\dfrac{b}{2c + b}}\right), & a \leq c + \dfrac{b}{2}\left[1 + \sqrt{1 + \dfrac{2c}{b}}\right] \\ 1, & a > c + \dfrac{b}{2}\left[1 + \sqrt{1 + \dfrac{2c}{b}}\right]. \end{cases}$$

Remark: Note that some of the foregoing quantities are indeterminant when $c = 0$. Nevertheless, the saddle-point strategies and the value in (A.26a) tend continuously to those of Lemma 2 as $c \to 0$.

References

[1] R. Ahlswede, "The capacity of a channel with arbitrarily varying Gaussian channel probability functions," in *Trans. 6th Prague Conf. Information Theory, Statistical Decision Functions, and Random Processes*, Sept. 1971, pp. 13–21.

[2] ——, "Elimination of correlation in random codes for arbitrarily varying channels," *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, vol. 44, pp. 159–175, 1978.

[3] S. Arimoto, "On the converse to the coding theorem for discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 357–359, May 1973.

[4] T. Başar and Y. W. Wu., "Solutions to a class of minimax decision problems arising in communications systems," preprint, 1984.

[5] R. M. Bell and T. M. Cover, "Competitive optimality of logarithmic investment," *Math Oper. Res.*, vol. 5, pp. 161–166, 1980.

[6] P. Billingsley, *Probability and Measure*. New York: Wiley, 1979.

[7] N. M. Blachman, "The effect of statistically dependent interference upon channel capacity," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 553–557, Sept. 1962.

[8] ——, "On the capacity of a band-limited channel perturbed by statistically dependent interference," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 48–55, Jan. 1962.

[9] D. Blackwell and M. A. Girshick, *Theory of Games and Statistical Decisions*. New York: Wiley, 1954 (reprinted by Dover, 1979).

[10] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacities of certain channel classes under random coding," *Ann. Math. Statist.*, vol. 31, pp. 558–567, 1960.

[11] J. M. Borden, D. M. Mason, and R. J. McEliece, "Some information-theoretic saddlepoints," *SIAM J. Contr. Opt.*, vol. 23, Jan. 1985.

[12] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[13] R. L. Dobrushin, "Optimal information transmission over a channel with unknown parameters," *Radiotekh. Elektron.*, vol. 4, pp. 1951–1956, 1959.

[14] R. G. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 3–18, Jan. 1965.

[15] S. W. Houston, "Modulation techniques for communication, part 1: Tone and noise jamming performance of spread-spectrum M-ary FSK and 2,4-ary DPSK waveforms," in *Proc. IEEE National Aeronautics and Electronics Conf. (NAECON)*, 1975, pp. 51–58.

[16] B. Hughes and P. Narayan, "Interleaving and channels with unknown memory," presented at the Conf. Information Sciences and Systems, Johns Hopkins University, Baltimore, MD, 1985.

[17] R. J. McEliece and W. E. Stark, "An information-theoretic study of communication in the presence of jamming," in *Proc. IEEE Int. Conf. Communications*, 1981, pp. 45.3.1–45.3.5.

[18] R. J. McEliece and E. R. Rodemich, "A study of optimal abstract jamming strategies vs. noncoherent MFSK," in *MILCOM Record*, Oct. 1983, pp. 1.1–1.6.

[19] C. E. Shannon, "Probability of error for optimal codes in a Gaussian Channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, May 1959.

[20] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread-Spectrum Communications*, vol. 1. Rockville, MD: Computer Science Press, 1985.

[21] J. Wolfowitz, *Coding Theorems of Information Theory*. 3rd ed. Berlin: Springer-Verlag, 1978.