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 (GAVCs). A GAVC is a discrete-time memoryless Gaussian channel with input power constraint $P_T$ and noise power $N_s$, 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 $\lambda$-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 to transmit 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 $u_i^* + n_i^* + s_i^*$. The elements of the sequence $\{n_i^*\}$ are independent zero-mean Gaussian random variables, each having variance $N_s$. 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 $N_s$, $P_T$, 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.
II. DEFINITIONS AND RESULTS

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

\[
y^* = u + \eta^* + s
\]

where \( \eta^* \) 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\(^2\)

\[
C_n = \{(u_1, D_1), \ldots, (u_M, D_M)\}
\]

where \( \{u_i\}_{i=1}^M \) are codewords of length \( n \) and \( \{D_i\}_{i=1}^M \) are disjoint (Borel) subsets of \( \mathbb{R}^n \) 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 \( s \). 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 \( s \). For the transmitter a strategy to accomplish this goal consists of an \((n, 2^n)\) 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

\[
C_n^* = \{(u_1^*, D_1^*), \ldots, (u_M^*, D_M^*)\}
\]

(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 \( s^* \).

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 \( s^* \) 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 \((n\text{-sphere})\) of radius \( \sqrt{n} P_T \) almost surely (a.s.), i.e., if for each \( 1 \leq i \leq M \) the codeword \( u_i^* = (u_{i1}^*, \ldots, u_{in}^*) \) satisfies

\[
\frac{1}{n} \sum_{j=1}^{n} u_{ij}^2 \leq P_T \text{ a.s.}
\]

(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

\[
E\left(\frac{1}{nM} \sum_{j=1}^{M} \sum_{i=1}^{n} u_{ij}^2\right) \leq P_T
\]

(2.5)

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

\[
\frac{1}{n} \sum_{i=1}^{n} s_{i}^2 \leq P_j \text{ a.s.}
\]

(2.6)
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 \(B\) (= PI 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:
\[
\lambda(C_n, s^*) \equiv \max_{1 \leq i \leq M} \Pr \left\{ \eta_i^* + \xi_i^* + s^* \in \overline{D}_i^* \right\} \tag{2.8}
\]
where \(\overline{D}_i^*\) denotes \(R^* - D_i^*\). 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 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 \((R, \lambda)\) 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 \(R_{A|B}\) denote the set of all achievable pairs \((R, \lambda)\) for the GAVC \(A|B\).

Note that if a certain pair \((R, \lambda)\) is achievable \(A|B\), then all pairs \((R', \lambda')\) such that \(R' \leq R\) and \(\lambda' \geq \lambda\) are also achievable \(A|B\). Consequently, \(R_{A|B}\) must be of the form
\[
R_{A|B} = \{ (R, \lambda) | 0 \leq R \leq C_{A|B}(\lambda), 0 \leq \lambda < 1 \} \tag{2.11}
\]
where \(C_{A|B}(\lambda)\) is a monotone increasing right-continuous function of \(\lambda\). Thus to characterize \(R_{A|B}\) it suffices to determine \(C_{A|B}(\lambda)\).

The function \(C_{A|B}(\lambda)\) is called the \(\lambda\)-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 \(\lambda\) for large \(n\). If \(C_{A|B}(\lambda)\) is equal to a constant on \(0 \leq \lambda < 1\), say \(C_{A|B}^\ast\), 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
\[
\lambda^A(R) = \lim_{\epsilon \to 0^+} \sup_{n \to \infty} \inf_{C_n} \lambda^B(C_n) \tag{2.12}
\]
where the infimum is over all \(A\)-admissible \((n, M)\) random codes such that \(M > 2^{n(R-\epsilon)}\). It is easy to show that the relationship between \(\lambda^A(R)\) and \(C_{A|B}(\lambda)\) is
\[
\lambda^A(R) = \min \{ 0 \leq \lambda \leq 1 | C_{A|B}(\lambda) \geq R \text{ or } \lambda = 1 \}. \tag{2.13}
\]

Although it clearly provides the same information about \(R_{A|B}\) as \(C_{A|B}(\lambda)\) does, \(\lambda^A(R)\) 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
\[
C_{PI|PJ}(\lambda) = \hat{C}_{PI|PJ}
\]
for all \(0 \leq \lambda < 1\), where
\[
\hat{C}_{PI|PJ} = \frac{1}{2} \log_2 \left( 1 + \frac{P_T}{N_s + P_J} \right). \tag{2.14}
\]

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

It is interesting to note that \(\hat{C}_{PI|PJ}\) 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]).\(^4\) 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., Csiszar and Körner [12]) definition of capacity (which always exists) is \(C_{A|B} = \lim_{n \to \infty} C_{A|B}(\lambda)\). 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 \( \lambda_{\text{PI|AJ}} < \lambda_{\text{P|P}} \). 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 \( \lambda \)-capacities are found.

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

\[
\lambda_{\text{PI|AJ}}(R) = \frac{1}{2} \log_2 \left( 1 + \frac{P_T}{N_e + P_J} \right) \tag{2.15}
\]

**Remark:** \( \lambda_{\text{PI|AJ}}(0) \) is interpreted as 0.

Observe that the expression for \( \lambda_{\text{PI|AJ}}(\lambda) \) is identical to that of \( \lambda_{\text{P|P}} \) except that the jamming power appears boosted by a factor that is the reciprocal of the tolerable error probability \( \lambda \). 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 \( (n,2^R) \) random codes satisfying PI suffer an error probability that approaches

\[
\lambda_{\text{PI|AJ}}(R) = \lambda_{\text{PI|AJ}}(R) = \left\{ \begin{array}{ll}
(4^R - 1) P_J & 
\text{if } R \leq \hat{C}_{\text{PI|AJ}}(1) \\
1 & 
\text{if } R > \hat{C}_{\text{PI|AJ}}(1)
\end{array} \right.
\]

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 \( 2 \ln 2 R P_J / P_T \). The region \( \lambda_{\text{PI|AJ}} \) is sketched in Fig. 2. It is apparent that a PI-admissible random code can achieve high reliability (i.e., \( \lambda_{\text{PI|AJ}}(C_n^*) \) 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 \( EP^* \leq P_J \). Conditioned on the value of \( P^* \) the jamming sequence \( s^* \) 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 \( s^* \) 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

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

and \( n \) is large, then \( \lambda_{\text{PI|AJ}}(C_n^*) \) is 0 is possible; however, if

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

then \( \lambda_{\text{PI|AJ}}(C_n^*) \) is 1 is certain. The jammer must, therefore, choose

\[
P^* > \frac{P_T}{(4^R - 1) - N_e}
\]

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 \( 2 \ln 2 R P_J / P_T \). The region \( \lambda_{\text{PI|AJ}} \) is sketched in Fig. 2. It is apparent that a PI-admissible random code can achieve high reliability (i.e., \( \lambda_{\text{PI|AJ}}(C_n^*) \) 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 \( EP^* \leq P_J \). Conditioned on the value of \( P^* \) the jamming sequence \( s^* \) 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 \( s^* \) 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

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

and \( n \) is large, then \( \lambda_{\text{PI|AJ}}(C_n^*) \) is 0 is possible; however, if

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

then \( \lambda_{\text{PI|AJ}}(C_n^*) \) is 1 is certain. The jammer must, therefore, choose

\[
P^* > \frac{P_T}{(4^R - 1) - N_e}
\]

against an AJ-constrained jammer.
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) \( \lambda \)-capacity is

\[
C_{\text{AI|PJ}}(\lambda) = \hat{C}_{\text{AI|PJ}}(\lambda)
\]

for all \( 0 \leq \lambda < 1 \), where

\[
\hat{C}_{\text{AI|PJ}}(\lambda) = \frac{1}{2} \log_2 \left( 1 + \frac{P_T/(1-\lambda)}{N_e + P_j} \right). \tag{2.17}
\]

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 \( C_{\text{AI|PJ}}(\lambda) \) approaches \( C_{\text{PJ|PJ}} \), and the improvements are negligible. As in the other cases, we can state the result in terms of error probabilities: optimal AI-admissible \( (n, 2^n \lambda) \) random codes suffer an error probability that, for large \( n \), approaches

\[
\lambda_{\text{AI|PJ}}(R) = \begin{cases} 
0, & \text{if } R \leq \hat{C}_{\text{AI|PJ}}(0) \\
\frac{P_T}{(4^N - 1)(N_e + P_j)}, & \text{if } R > \hat{C}_{\text{AI|PJ}}(0).
\end{cases}
\]

Thus the rates at which reliable communication can occur are the same as in the case PJ|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 PJ|AJ. The next theorem shows that some limited improvement is made.

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

\[
C_{\text{AI|AJ}}(\lambda) = \hat{C}_{\text{AI|AJ}}(\lambda)
\]

for all \( 0 \leq \lambda < 1 \), where

\[
\hat{C}_{\text{AI|AJ}}(\lambda) = \begin{cases} 
\frac{1}{2} \log_2 \left( 1 + \frac{P_T}{N_e + P_j/2\lambda} \right), & 0 \leq \lambda \leq \lambda_c \\
\frac{1}{2} \log_2 \left( 1 + \frac{P_T(1 - 2\lambda_c)}{(1 - \lambda)N_e} \right), & \lambda_c \leq \lambda < 1
\end{cases}
\]

for \( N_e > 0 \), and

\[
\hat{C}_{\text{AI|AJ}}(\lambda) = \begin{cases} 
\frac{1}{2} \log_2 \left( 1 + \frac{2\lambda P_T}{P_j} \right), & 0 \leq \lambda < \frac{1}{2} \\
\frac{1}{2} \log_2 \left( 1 + \frac{P_T}{2(1 - \lambda)P_j} \right), & \frac{1}{2} \leq \lambda < 1
\end{cases}
\]

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 \( C_{\text{PJ|PJ}}, C_{\text{AI|AJ}}(\lambda), \) and \( C_{\text{AI|PJ}}(\lambda) \) included for comparison. Optimal \( (n, 2^n \lambda) \) random codes satisfying AI must then, as \( n \) grows large, suffer an error probability that approaches \( \lambda_{\text{AI|AJ}}(R) = \hat{\lambda}_{\text{AI|AJ}}(R) \) where

\[
\lambda_{\text{AI|AJ}}(R) = \begin{cases} 
\frac{P_T(4^N - 1)}{2(P_T - (4^N - 1)N_e)}, & R \leq \hat{C}_{\text{AI|AJ}}(\lambda_c) \\
\frac{P_T(1 - 2\lambda_c)}{(4^N - 1)N_e}, & R > \hat{C}_{\text{AI|AJ}}(\lambda_c)
\end{cases}
\]

Fig. 3. Achievable region for GAVC AI|PJ.
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 \( \lambda(C_n^*, s^*) = 1 \). Therefore, for the best choice of \( C_n^* \) we have for large \( n \)
\[ \lambda(C_n^*, s^*) = \Pr \{ P_1^* < (4^R - 1)(P_2^* + N_e) \}. \] (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
\[ \lambda^M(C_n^*) \]
\[ \approx \max_{P_1^*: EP_1^* < P_T} \min_{P_2^*: EP_2^* < P_T} \Pr \{ P_1^* < (4^R - 1)(P_2^* + N_e) \}. \]

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 \( P_T' \neq P_T \). Denote this joint constraint by \( AA' \). Similarly, one may define a double constant \( BB' \) on jamming sequences. It is easily checked that the \( \lambda \)-capacities for these more complex coding problems can be constructed from the \( \lambda \)-capacities defined by Theorems 1–4 according to the following simple rules:

\[ C_{AA'B}^*(\lambda) = \min \left\{ \hat{C}_{A'B}^*(\lambda), \hat{C}_{A'B}^*(\lambda) \right\} \] (2.22a)
\[ C_{AB'B}^*(\lambda) = \max \left\{ \hat{C}_{AB}^*(\lambda), \hat{C}_{AB}^*(\lambda) \right\}. \] (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
\[ C_n^* = \{(e_1^*, A_1^*), \ldots, (e_M^*, A_M^*)\} \] (3.1)
constructed in the following way.

1) The \( M \) random codewords \( \{e_1^*, \ldots, e_M^*\} \) is 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 \( e^*_1 \) 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.

\[ * \] It is unknown whether \( C_{AA'B}^*(\lambda) \) can similarly be decomposed.
2) The random decoding sets \( \{ A_i^* \}_{i=1}^M \) are defined by a strict minimum Euclidean distance rule, viz.,
\[
A_i^* = \{ y \in \mathbb{R}^n \mid |y - v_i^*| < |y - v_{i+1}^*| \}
\]
for all \( k \neq i, 1 \leq k \leq M \) \( (3.2) \)
where \(| \cdot | \) denotes the usual Euclidean norm on \( \mathbb{R}^n \). 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 \( C_i^* \). First, the codewords of \( C_i^* \) 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 of the form \( (\sqrt{P} v_1^*, \ldots, \sqrt{P} v_M^*) \) 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\(^7\)
\[
\Pr \{ \sqrt{P} v_i^* + \eta^* \in A_i^* \} \leq K(R, P) \exp \{-nE(R, P)\} \tag{3.4}
\]
holds for all \( 1 \leq i \leq M \) and \( n \geq 1 \), where \( \eta^* \) 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
\[
C_n^* = \{ (u_1^*, D_1^*), \ldots, (u_M^*, D_M^*) \}
\]
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
\[
\frac{1}{M} \sum_{i=1}^M \Pr \{ u_i^* + \eta^* \in D_i^* \} \geq 1 - K'(R, P) \exp \{-nE'(R, P)\} \tag{3.7}
\]
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 \( u = (u_1, \ldots, u_n) \in \mathbb{R}^n \) the quantity
\[
P(u) \equiv \frac{1}{n} \sum_{j=1}^n u_j^2. \tag{3.9}
\]
For any random code \( C_n^* \) let \( U^*(C_n^*) \) be the random variable that is uniformly distributed on the set \( \{ u_1^*, \ldots, u_n^* \} \) of codewords of \( C_n^* \); i.e., let \( k^* \) be uniformly distributed on the set \( \{ 1, \ldots, M \} \) and independent of \( C_n^* \), and define
\[
U^*(C_n^*) = u_{k^*}. \tag{3.10}
\]

Lemma 1: Let \( C_n^* \) be any \( (n, M) \) random code and \( J^* \) be any nonnegative random variable. Then for all \( 0 < \epsilon < R = (1/n) \log_2 M \) the following holds:
\[
\frac{1}{M} \sum_{i=1}^M \Pr \{ u_i^* + \eta^* + \sqrt{J^*} \eta^* \in D_i^* \} \geq \Pr \{ P(U^*(C_n^*)) < (4^{R-\epsilon} - 1)(N_\epsilon + J^*) \}
\]
\[
- \gamma_n(\epsilon) \tag{3.11}
\]
where
\[
\gamma_n(\epsilon) = K'(R - \epsilon/2, 4^{R-\epsilon} - 1)
\]
\[\cdot \exp \{-nE'(R - \epsilon/2, 4^{R-\epsilon} - 1)\} + 2^{-n^4/2}. \tag{3.12}\]

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

Proof of Lemma 1: Let \( 0 < \epsilon < R \). Suppose first that \( C_n^* \) and \( J^* \) are deterministic, say \( C_n^* = C_n = \{ (u_1, D_1), \ldots, (u_M, D_M) \} \) and \( J^* = J \). Define the set
\[
S_n(C_n, J) = \{ 1 \leq i \leq M \mid P(u_i) < (4^{R-\epsilon} - 1)(N_\epsilon + J) \}, \tag{3.13}
\]
and let \( N_c(C_n, J) \) denote its cardinality. It is immediate that

\[
Pr \left\{ P(U^s(C_n)) \leq (4^R - 1)(N_c + J) \right\} = \frac{N_c(C_n, J)}{M}.
\]

(3.14)

Consider the subcode of \( C_n \) that consists of those codewords and decoding sets with indices in \( S(C_n, J) \). 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\(^8\)

\[
\frac{1}{N_c(C_n, J)} \sum_{i \in S(C_n, J)} Pr \left\{ u_i + \eta_i^* + \sqrt{J} \eta^* \in \bar{D}_i \right\} \\
\geq 1 - K'(R_n, 4^{R - \epsilon} - 1) \cdot \exp \left\{ -nE'(R_n, 4^{R - \epsilon} - 1) \right\}
\]

(3.15)

provided that

\[
R_n = \log_2 \left( \frac{N_c(C_n, J)}{n} \right) > R - \epsilon.
\]

(3.16)

Therefore, the following holds for all \( C_n, J, \epsilon, \) and \( R \):\(^9\)

\[
\frac{1}{N_c(C_n, J)} \sum_{i \in S(C_n, J)} Pr \left\{ u_i + \eta_i^* + \sqrt{J} \eta^* \in \bar{D}_i \right\} \\
\geq (1 - K'(R_n, 4^{R - \epsilon} - 1) \cdot \exp \left\{ -nE'(R_n, 4^{R - \epsilon} - 1) \right\} 1_{(R - \epsilon) / 2, +\infty}(R_n) \\
a) \geq (1 - \gamma_n(\epsilon) + 2^{-\epsilon}) 1_{(R - \epsilon) / 2, +\infty}(R_n) \\
\geq 1_{(R - \epsilon) / 2, +\infty}(R_n) - \gamma_n(\epsilon) + 2^{-\epsilon / 2}
\]

(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 \( J^* = J \):

\[
\frac{1}{M} \sum_{i=1}^{M} Pr \left\{ u_i^* + \eta_i^* + \sqrt{J} \eta^* \in \bar{D}_i \right\} \\
\geq \frac{1}{M} \sum_{i \in S(C_n, J)} Pr \left\{ u_i + \eta_i^* + \sqrt{J} \eta^* \in \bar{D}_i \right\}
\]

\[
a) \geq \frac{N_c(C_n, J)}{M} - \gamma_n(\epsilon) + 2^{-\epsilon / 2}
\]

(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) \( C_{PI/JL}(\lambda) \geq C_{PI/JL} \): Let \( R \) nonnegative be given and set \( M_n = [2^n R] \).\(^9\) Define a sequence of \( (n, M_n) \) random codes, say \( \{ C_n^* \}_{n=1}^{\infty} \), in the following way:

\[
C_n^* = \left\{ (u_1^*, A_1^*), \cdots, (u_{M_n}^*, A_{M_n}^*) \right\}
\]

(3.19)

where \( u_i^* = \sqrt{P_n \eta_i^*} \) and \( \{ (v_1^*, A_1^*), \cdots, (v_{M_n}^*, A_{M_n}^*) \} \) is the standard \( (n, M_n) \) random code defined in (3.1). It is easily verified that \( C_n^* \) satisfies PI for each \( n \geq 1 \). We claim that if

\[
R < C_{PI/JL} \]

(3.20)

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

\[
\lambda^P(C_n^*) = \gamma_n
\]

(3.21)

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

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

\[
\sigma_n(i) = Pr \left\{ u_i^* + \eta_i^* + i \omega^* \in \bar{A}_i^* \right\}
\]

(3.22)

for any real number \( I \geq 0 \). (Clearly, \( \sigma_n(\cdot) \) does not depend on \( i \)) Let \( s^* \) be any jamming sequence that satisfies PI. The error probability incurred by \( s^* \) when message \( i \) sent can be bounded in the following way:

\[
Pr \left\{ u_i^* + \eta_i^* + s^* \in \bar{A}_i^* \right\} \leq \frac{1}{M} \sum_{i=1}^{M} \sigma_n(i) \leq \frac{1}{M} \sum_{i=1}^{M} \sigma_n(i) \sqrt{P_n}
\]

(3.23)

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

\[
T_s = |s| \omega.
\]

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

\[
Pr \left\{ u_i^* + \eta_i^* + s \in \bar{A}_i^* \right\} = Pr \left\{ u_i^* + T \eta_i^* + |s| \omega \in T \bar{A}_i^* \right\}
\]

(3.24)

The sets \( \{ T \bar{A}_i^* \}_{i=1}^{M} \) remain minimum distance decoding sets for the codewords \( \{ T u_i^* \}_{i=1}^{M} \), and the distributions of \( \{ u_i^* \}_{i=1}^{M} \) and \( \eta_i^* \) are symmetrically similar and so are unchanged by \( T \). We conclude that

\[
Pr \left\{ u_i^* + \eta_i^* + s \in \bar{A}_i^* \right\} = Pr \left\{ u_i^* + \eta_i^* + |s| \omega \in \bar{A}_i^* \right\}
\]

for all \( \omega \) in the ensemble of \( \omega^* \) from which a) immediately follows. To prove b), it suffices to show that \( \alpha_n(\cdot) \) is increasing, since PI implies \( |s| \leq \sqrt{P_n} \) (a.s.). Let \( \bar{l} \) and \( \bar{l} \) be nonnegative numbers so that \( 0 \leq \bar{l} \leq l \). Let the random variable \( m^* \) be defined by

\[
m^* = |\eta_i^* + l \omega^*|
\]

(3.25)

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

\(^9\) \( l_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \in \overline{A} \end{cases} \)

\(^{10}\) \( |x| \) denotes the integer \( n \) such that \( x - 1 < n \leq x \).
and let $F_i(m)$ be its distribution function. It is easy to verify that, conditioned on the occurrence $m^*_i - m$, the random variable $\eta^*_i + l\omega^*$ is uniformly distributed on the $n$-sphere of radius $m$; hence its conditional distribution does not depend on $l$. Therefore, define the quantity

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

and note that

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

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

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

implies

$$u^*_1 + \tilde{m} \left( \frac{\eta^*_1 + l\omega^*}{m^*_1} \right) \in \tilde{A}^*_1,$$

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

$$\sigma_n(l) = \int_0^\infty \gamma(m) \, dF_i(m) \leq \int_0^\infty \gamma(m) \, dF_i(m) = \sigma_n(l)$$

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

$$F_i(m) \leq F_i(m). \quad (3.25)$$

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

$$\Pr \left\{ |\eta^*_1 + l\omega^*|^2 \leq m^2 | \omega^* = \omega \right\} \leq \Pr \left\{ |\eta^*_1 + l\omega^*|^2 \leq m^2 | \omega^* = \omega \right\}$$

for all $\omega$. The latter inequality is an immediate consequence of the fact that the distribution of $\eta^*_1$ decreases monotonically and symmetrically with distance from the origin. This completes the proof of (3.23).

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

$$\lambda^{\text{PI}}(C^*_s) \leq \sigma_n(\sqrt{n}P_f). \quad (3.26)$$

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 $\sqrt{P_f} \eta^*$ denote a vector of i.i.d. $N(0, P_f)$ random variables, and let $f(\cdot)$ denote the probability density function of the random variable $m^* \equiv \sqrt{P_f} \eta^*$. It is easy to show that

$$\Pr \left\{ u^*_1 + \eta^*_1 + \sqrt{P_f} \eta^* \in \tilde{A}^*_1 \right\} = \int_0^\infty \sigma_n(l) f(l) \, dl. \quad (3.27)$$

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

$$\sigma_n(\sqrt{n}P_f) \leq \int_0^\infty \sigma_n(l) f(l) \, dl \leq \frac{\Pr \left\{ u^*_1 + \eta^*_1 + \sqrt{P_f} \eta^* \in \tilde{A}^*_1 \right\}}{\Pr \{ |\eta^*_1| \geq 1 \}}. \quad (3.28)$$

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

$$P_b = \frac{P_T}{N_e + P_f/b} \quad (3.30)$$

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

$$\Pr \{ |\eta^*_1| \geq 1 \} \geq 1/c_0 > 0 \quad (3.31)$$

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

$$\lambda^{\text{PI}}(C^*_s) \leq c_0 K(R, P_1) \exp \left\{ -nE(R, P_1) \right\} \quad (3.32)$$

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

b) $C^{\text{PI}}(\lambda) \leq \tilde{C}^{\text{PI}}(\lambda)$. Suppose that

$$R > \tilde{C}^{\text{PI}}(\lambda). \quad (3.33)$$

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

$$\lambda^{\text{PI}}(C^*_s) > 1 - \gamma_n \quad (3.34)$$

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

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

$$\tilde{C}^{\text{PI}}(\lambda) < \frac{1}{2} \log_2 \left( 1 + \frac{P_T}{N_e + P_f/(1 + \epsilon)} \right) < R, \quad (3.35)$$

and let $C^*_n = \{(u^*_1, D^*_1), \ldots, (u^*_M, D^*_M)\}$ be any $(n, M)$ random code satisfying PJ. If the jamming sequence were $s^* = \sqrt{P_f/(1 + \epsilon)} \eta^*$, then by (3.7) we know that

$$\max_{1 \leq i \leq M} \Pr \left\{ u^*_1 + \eta^*_1 + \sqrt{P_f/(1 + \epsilon)} \eta^* \in \tilde{D}^*_1 \right\} \geq 1 - K'(R, P_{11}) \exp \left\{ -nE'(R, P_{11}) \right\} \quad (3.36)$$

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

$$\eta^*_\epsilon(\epsilon) = \begin{cases} \sqrt{P_f/(1 + \epsilon)} \eta^*, & |\eta^*| \leq \sqrt{n(1 + \epsilon)} \\ \frac{\sqrt{n}P_f}{|\eta^*|} \eta^*, & |\eta^*| > \sqrt{n(1 + \epsilon)} \end{cases} \quad (3.37)$$

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

\[
\Pr \left\{ u^*_t + \eta^*_t + \sqrt{P_j/(1+\epsilon)} \eta^*_t \in \bar{D}_t^* \right\} \\
= \Pr \left\{ u^*_t + \eta^*_t + \sqrt{P_j/(1+\epsilon)} \eta^*_t | |\eta^*_t| \leq \sqrt{n}(1+\epsilon) \right\} \\
\quad \cdot \Pr \left\{ |\eta^*_t| \leq \sqrt{n}(1+\epsilon) \right\} \\
+ \Pr \left\{ u^*_t + \eta^*_t + \sqrt{P_j/(1+\epsilon)} \eta^*_t | |\eta^*_t| > \sqrt{n}(1+\epsilon) \right\} \\
\quad \cdot \Pr \left\{ |\eta^*_t| > \sqrt{n}(1+\epsilon) \right\} \\
\leq \Pr \left\{ u^*_t + \eta^*_t + \sqrt{P_j/(1+\epsilon)} \eta^*_t \in \bar{D}_t^* \right\} \\
+ \Pr \left\{ |\eta^*_t| > \sqrt{n}(1+\epsilon) \right\}.
\] (3.38)

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

\[
\lambda^pi(C_n^*) \geq \max_{1 \leq i \leq M} \Pr \left\{ u^*_t + \eta^*_t + \eta^*_t(\epsilon) \in \bar{D}_t^* \right\} \\
\geq 1 - K'(R, P_{1+n}) \exp \left\{ -nF'(R, P_{1+n}) \right\} \\
\quad - \Pr \left\{ |\eta^*_t| > \sqrt{n}(1+\epsilon) \right\}.
\] (3.39)

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) \( C_{pi}(\lambda) \geq \bar{C}_{pi}(\lambda) \): We retain the notation and results of part a) of the proof of Theorem 1. Let \( R \) nonnegative be given. Set \( M_n = [2^nR] \), 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

\[
R < \bar{C}_{pi}(\lambda),
\] (3.40)

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

\[
\lambda^{AJ}(C_n^*) \leq \lambda + \gamma_n
\] (3.41)

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

To prove (3.41), let \( \lambda \) 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 \( 1 \leq i \leq M_n \)

\[
\Pr \left\{ u^*_i + \eta^*_i + s^* \in \bar{D}_i^* \right\} \leq \frac{1}{n} \sum_{i=1}^{n} s^*_i^2 \leq P_j/\lambda
\]

\[
\leq c_0 K(R, P_{\lambda}) \exp \left\{ -nE(R, P_{\lambda}) \right\}
\] (3.42)

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

\[
\Pr \left\{ \frac{1}{n} \sum_{i=1}^{n} s^*_i^2 > P_j/\lambda \right\} \leq \lambda.
\] (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:

\[
\Pr \left\{ u^*_i + \eta^*_i + s^* \in \bar{D}_i^* \right\} \\
= \Pr \left\{ u^*_i + \eta^*_i + s^* \in \bar{D}_i^* \right\} \frac{1}{n} \sum_{i=1}^{n} s^*_i^2 \leq P_j/\lambda
\]

\[
\quad \cdot \Pr \left\{ \frac{1}{n} \sum_{i=1}^{n} s^*_i^2 \leq P_j/\lambda \right\} \\
+ \Pr \left\{ u^*_i + \eta^*_i + s^* \in \bar{D}_i^* \right\} \frac{1}{n} \sum_{i=1}^{n} s^*_i^2 > P_j/\lambda
\]

\[
\quad \cdot \Pr \left\{ \frac{1}{n} \sum_{i=1}^{n} s^*_i^2 > P_j/\lambda \right\}
\]

\[
\leq \lambda + c_0 K(R, P_{\lambda}) \exp \left\{ -nE(R, P_{\lambda}) \right\}.
\] (3.44)

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

\[
\lambda^{AJ}(C_n^*) \leq \lambda + c_0 K(R, P_{\lambda}) \exp \left\{ -nE(R, P_{\lambda}) \right\}.
\] (3.45)

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

b) \( C_{pi}(\lambda) \leq \bar{C}_{pi}(\lambda) \): We now prove that if

\[
R > \bar{C}_{pi}(\lambda),
\] (3.46)

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

\[
\lambda^{AJ}(C_n^*) \geq \lambda(1 - \gamma_n)
\] (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 \( \lambda \) be such that \( 0 < \lambda < 1 \). Suppose that a pulsed jamming sequence, say \( s^*_n \), is defined to be

\[
s^*_n = \sqrt{P_j/\lambda} Z^*_n \eta^*
\] (3.48)

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

\[
\Pr \{ Z^*_n = 1 \} = 1 - \Pr \{ Z^*_n = 0 \} = \lambda.
\] (3.49)

It is easy to verify that \( s^*_n \) satisfies AJ for all \( 0 < \lambda \leq 1 \) and all \( n \geq 1 \).

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

\[
\lambda^{AJ}(C_n^*) \geq \max_{1 \leq i \leq M} \Pr \left\{ u^*_i + \eta^*_i + s^*_i \in \bar{D}_i^* \right\}
\]

\[
\geq \max_{1 \leq i \leq M} \Pr \left\{ u^*_i + \eta^*_i + s^*_i \in \bar{D}_i^* \right\} \Pr \{ Z^*_i = 1 \}
\]

\[
= \lambda \left( \max_{1 \leq i \leq M} \Pr \left\{ u^*_i + \eta^*_i + \sqrt{P_j/\lambda} \eta^*_i \in \bar{D}_i^* \right\} \right)
\]

\[
\geq \lambda \left( 1 - K'(R, P_{\lambda}) \exp \left\{ nE'(R, P_{\lambda}) \right\} \right)
\] (3.50)
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 \( \lambda^W(\cdot) \), b) follows from (3.48) and (3.49), and c) is a consequence of (3.46) and (3.7).

If (3.46) holds, then

\[
K(R, P_\alpha) \exp \{-nE(R, P_\alpha)\} \to 0,
\]

thus completing the proof of the converse to Theorem 2.

**Proof of Theorem 3:**

a) \( C_{\text{AIP}}(\lambda) \geq \tilde{C}_{\text{AIP}}(\lambda) \): Let a nonnegative \( R \) be given, and set \( M_n = [2^{nR}] \). 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:

\[
C_n^*(\lambda) \equiv \{(u_1^*(\lambda), A_1^*), \ldots, (u_{M_n}^*(\lambda), A_{M_n}^*)\}
\]

(3.51)

where

\[
u_1^*(\lambda) = \sqrt{P_T/(1 - \lambda)} Z_{1-\lambda}.
\]

(3.52)

\( Z_{1-\lambda} \) is a Bernoulli random variable independent of \( \nu_1^* \) such that

\[
\Pr\{Z_{1-\lambda} = 1\} = 1, \Pr\{Z_{1-\lambda} = 0\} = \lambda
\]

(3.53)

and \( C_n^* = \{(\nu_1^*, A_1^*), \ldots, (\nu_{M_n}^*, A_{M_n}^*)\} \) is the standard \((n, M_n)\) random code as in (3.1). It is easy to verify that \( C_n^*(\lambda) \) satisfies Al for all \( 0 \leq \lambda < 1 \) and all \( n \). We claim that if

\[
R < \tilde{C}_{\text{AIP}}(\lambda),
\]

(3.54)

then a positive sequence \( \gamma_n \to 0 \) exists such that

\[
\lambda^W(C_n^*(\lambda)) \leq \lambda + \gamma_n
\]

(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 \( s^* \) be any PJ-admissible jamming signal, and suppose \( \lambda \) is such that (3.54) holds. We can then bound the error probability above as follows:

\[
\Pr\{u_i^*(\lambda) + \eta_i^* + s^* \in A_i^*\} = \Pr\{u_i^*(\lambda) + \eta_i^* + s^* \in A_i^*|Z_{1-\lambda} = 0\}
\]

\[
\cdot \Pr\{Z_{1-\lambda} = 0\}
\]

\[
+ \Pr\{u_i^*(\lambda) + \eta_i^* + s^* \in A_i^*|Z_{1-\lambda} = 1\}
\]

\[
\cdot \Pr\{Z_{1-\lambda} = 1\}
\]

\[
\leq \lambda + \Pr\{\sqrt{P_T/(1 - \lambda)} \eta_i^* + \eta_i^* + s^* \in A_i^*\}
\]

\[
\leq \lambda + c_0 K(R, P^\lambda) \exp \{-nE(R, P^\lambda)\}
\]

(3.56)

for all \( i \) and PJ-admissible \( s^* \), where

\[
P^\lambda \equiv \frac{P_T/(1 - \lambda)}{N_e + P_J}.
\]

(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 \( s^* \) satisfies PJ. Equation (3.56) implies (3.55), thus completing the forward part of Theorem 3.

b) \( C_{\text{Al+}}(\lambda) \leq \tilde{C}_{\text{AIP}}(\lambda) \): We now prove that a positive sequence \( \{\gamma_n\}_{n=1}^{\infty} \) exists such that \( \gamma_n \to 0 \), and if

\[
R > \tilde{C}_{\text{AIP}}(\lambda),
\]

(3.58)

then

\[
\lambda^W(C_n^*) \geq \lambda - \gamma_n
\]

(3.59)

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

To prove this, let

\[
C_n^* = \{(u_1^*, D_1^*), \ldots, (u_M^*, D_M^*)\}
\]

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

\[
R > \frac{1}{2} \log_2 \left[ 1 + \frac{P_T/(1 - \lambda)}{N_e + P_J/(1 + \epsilon)} \right] + \epsilon > \tilde{C}_{\text{Al+}}(\lambda),
\]

(3.60)

and let \( \eta^*(\epsilon) \) 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

\[
\lambda^W(Pu_n^*) \geq \max_{1 \leq i \leq M} \Pr\{u_i^* + \eta_i^* + \eta^*(\epsilon) \in D_i\}
\]

\[
> \max_{1 \leq i \leq M} \Pr\{u_i^* + \eta_i^* + \sqrt{P_J/(1 + \epsilon)} \eta^* \in D_i\}
\]

\[
- \Pr\{|\eta^*| > \sqrt{n(1 + \epsilon)}\}.
\]

(3.61)

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

\[
\max_{1 \leq i \leq M} \Pr\{u_i^* + \eta_i^* + \sqrt{P_J/(1 + \epsilon)} \eta^* \in D_i\}
\]

\[
\geq \frac{1}{M} \sum_{i=1}^{M} \Pr\{u_i^* + \eta_i^* + \sqrt{P_J/(1 + \epsilon)} \eta^* \in D_i\}
\]

\[
\geq \Pr\{P(U^*(C_n^*)) < (4^{R - \epsilon} - 1)(N_e + P_J/(1 + \epsilon))\}
\]

\[
- \gamma_n(\epsilon)
\]

(3.62)

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

\[
\Pr\{P(U^*(C_n^*)) < (4^{R - \epsilon} - 1)(N_e + P_J/(1 + \epsilon))\}
\]

\[
\geq 1 - \frac{P_J}{(4^{R - \epsilon} - 1)(N_e + P_J/(1 + \epsilon))}
\]

(3.63)

The justification of these steps is as follows: a) follows by applying Chebyshev’s inequality using \( E\lambda^W(U^*(C_n^*)) \leq P_T \), which is equivalent to Al; 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) $C_{\text{AI}}(\lambda) \geq C_{\text{AI}}(\lambda)$: For any nonnegative $R$ set $M_n = \lfloor 2^n \lambda \rfloor$. Fix $\epsilon > 0$, and define a sequence of AI-admissible $(n, M_n)$ random codes, say

$$
C_n^*(\epsilon) = \left\{ (u_n^*(\epsilon), A_n^*) \right\}
$$

(3.64) where

$$
u_n^*(\epsilon) = (P_0^*(\epsilon) v_n^*).
$$

(3.65)

$P_0^*(\epsilon)$ is a nonnegative random variable independent of $\{v_1^*, \ldots, v_M^*\}$ that satisfies $EP_0^*(\epsilon) \leq P_T$ and whose distribution will follow. $\hat{C}_n^* = (u_n^*(\epsilon), A_n^*)_{n=1}^{M_n}$ 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

$$
R \leq \hat{C}_{\text{AI}}(\lambda) - \epsilon,
$$

(3.66)

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

$$
\lambda^\text{AI}(C_n^*(\epsilon)) \leq \lambda \gamma_n
$$

(3.67)

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

In proving this claim we assume that $N_0 > 0$; the proof if $N_0 = 0$ 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

$$
\Pr\{X_0 \geq Y + c\} \geq v_0
$$

(3.68)

holds for all independent nonnegative random variables $Y$ that satisfy $EY \leq b$.

Now make the following substitutions:

$$
a = \frac{P_T}{(4R + \epsilon - 1)} \quad b = P_J \quad c = N_0,
$$

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

$$
P_0^*(\epsilon) = (4R + \epsilon - 1)X_0.
$$

With these substitutions it is easy to verify that

$$
u_0 = 1 - \hat{\lambda}^\text{AI}(R + \epsilon)
$$

$$
a) \geq 1 - \hat{\lambda}^\text{AI}(\hat{C}_{\text{AI}}(\lambda))
$$

$$
= 1 - \lambda
$$

(3.69)

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

$$
\Pr\{P_0^*(\epsilon) < (4R + \epsilon - 1)(N_0 + J^*)\} \leq 1 - \nu_0 \leq \lambda
$$

(3.70)

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

$$
\Pr\{\sqrt{P_0^*(\epsilon)} v_1^* + \eta^* + \sqrt{J^*} \delta^* \in \hat{A}^*_n\}
$$

$$
+ \sqrt{J^*} \delta^* \in \hat{A}^*_n\}
$$

$$
< c_0 K(R, P) \exp\{-nE(R, P')\}
$$

(3.71)

for all $n \geq 1$ provided that

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

In particular, if

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

(3.72)

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

$$
\bar{B}_n(R, \epsilon) = c_0 K(R, 4R + \epsilon - 1) \exp\{-nE(R, 4R + \epsilon - 1)\}.
$$

(3.73)

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

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

(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

$$
\Pr\{u_n^*(\epsilon) + \eta_n^* + s^* \in \hat{A}^*_n\}
$$

$$
= \Pr\{\sqrt{P_0^*(\epsilon)} v_1^* + \eta_n^* + \sqrt{J^*} \delta^* \in \hat{A}^*_n\}
$$

$$
\leq E h_n(P_0^*(\epsilon), J^*)
$$

$$
= \bar{B}_n(R, \epsilon) + (1 - \bar{B}_n(R, \epsilon))
$$

$$
\Pr\{P_0^*(\epsilon) < (4R + \epsilon - 1)(N_e + J^*)\}
$$

$$
\leq \bar{B}_n(R, \epsilon) + \lambda
$$

(3.75)

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

$$
\lambda^\text{AI}(C_n^*(\epsilon)) \leq \bar{B}_n(R, \epsilon) + \lambda
$$

(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) $C_{\text{AI}}(\lambda) \leq C_{\text{AI}}(\lambda)$: 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

$$
\lambda^\text{AI}(C_n^*) \geq \lambda - \gamma_n
$$

(3.77)

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

$$
R > \hat{C}_{\text{AI}}(\lambda) + \epsilon;
$$

(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 $\nu_0$ is as defined in (A.26a), then

$$\Pr \{ X \geq Y_0 + \epsilon \} \leq \nu_0$$

(3.79)

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

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

and defining

$$J_0^*(\epsilon) = Y_0$$

$$J^* = (4^{R-\epsilon} - 1) X,$$

we obtain

$$\nu_0 = 1 - \frac{1}{\lambda^{\text{MAX}}(\cdot)} \leq 1 - \lambda$$

where we have used (3.78) and the fact that $\lambda^{\text{MAX}}$ is strictly increasing. Therefore,

$$\Pr \{ J^* < (4^{R-\epsilon} - 1)(N_e + J_0^*(\epsilon)) \} \geq 1 - \nu_0 \geq \lambda$$

(3.80)

holds for all $J^*$ satisfying

$$\gamma^* \leq P_T.$$  

(3.81)

Note that $\sqrt{J_0^*(\epsilon)} \gamma^*$ is 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:

$$\lambda^M(C^*_n) \geq \max_{1 \leq i \leq M} \Pr \{ u^*_i + \gamma^* + \sqrt{J_0^*(\epsilon)} \gamma^* \in \tilde{D}_i^* \}$$

$$\geq \frac{1}{M} \sum_{i=1}^{M} \Pr \{ u^*_i + \gamma^* + \sqrt{J_0^*(\epsilon)} \gamma^* \in \tilde{D}_i^* \}$$

$$\geq \Pr \{ U^*(C^*_n)) < (4^{R-\epsilon} - 1)(N_e + J_0^*(\epsilon)) \} - \gamma_n(\epsilon)$$

(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 $\gamma^* \leq P_T$. 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 AVCs: whereas the latter always have a random coding capacity (cf. Blackwell et al. [10]), the former generally have no capacity (except in the case P\{\|\}). 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 $\lambda$-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 can make an interesting observation: Blachman [7], [8] has investigated the use of the GAVC P\{\|\} (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

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

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

$$\frac{1}{n} \sum_{i=1}^{n} s^2_i \leq P_j.$$  

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

$$D^\gamma_{\{\|\}} \geq C_b$$

(4.1)

where

$$C_b = \begin{cases} \frac{1}{2} \log \left( \frac{P_T}{N_e + P_j} \right), & P_T P_j > (P_j + N_e)^2 \\ \frac{1}{2} \log \left( 1 + \frac{(\sqrt{P_T} - \sqrt{P_j})^2}{N_e} \right), & P_j^2 < P_T \leq (P_j + N_e)^2 \\ 0, & P_T \leq P_j \end{cases}$$
Comparing (2.14) with the upper bound in (4.1), we find that

\[ C_b < C_{PI|PJ} \]

for all (positive) values of \( N_r, P_r, \) and \( P_f \). 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 \( \text{A}|\text{B} \), 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. \( R^n \)-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:

transmitter's program: \( \bar{\nu}_n = \inf \sup_{C^*_n, s^*} \lambda(C^*_n, s^*) \) \hspace{1cm} (4.2a)

jammer's program: \( \nu_n = \sup \inf_{s^* \in C^*_n} \lambda(C^*_n, s^*) \) \hspace{1cm} (4.2b)

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 \( \varepsilon > 0 \), \( \varepsilon \)-optimal strategies \( C^*_n \) and \( s^* \) are allowable strategies for which

\[ \sup_{s^*} \lambda(C^*_n, s^*) \leq \bar{\nu}_n + \varepsilon \] \hspace{1cm} (4.3)

\[ \inf_{s^* \in C^*_n} \lambda(C^*_n, s^*) \geq \nu_n - \varepsilon \] \hspace{1cm} (4.4)

where the extrema are taken over all allowable \( s^* \) and \( C^*_n \). It is always true that \( \nu_n \leq \bar{\nu}_n \); if \( \nu_n = \bar{\nu}_n \), then the game is said to have a \textit{value} \( \nu_n = \bar{\nu}_n \).

Equation (4.2a) defined a sequence (in \( n \)) of communication games. Bäsar and Wu [4] have considered games of this type (for the constraints \( \text{A}|\text{AJ} \)) 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 \textit{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

\[ \lim_{n \to +\infty} \nu_n = \lim_{n \to +\infty} \bar{\nu}_n = \hat{\lambda}^{\text{A|B}}(R) \] \hspace{1cm} (4.5)

holds for every \( R \) and every pair of constraints \( \text{A}|\text{B} \). Thus the sequence of games has an "asymptotic value" equal to \( \hat{\lambda}^{\text{A|B}}(R) \). Furthermore, for all \( \varepsilon > 0 \), there exist for all sufficiently large \( n \), \( \varepsilon \)-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

\[ s^* = (z_1^*, \eta_1^*, \ldots, z_{n^*}^*, \eta_{n^*}^*) \] \hspace{1cm} (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

\[ \frac{1}{n} \sum_{i=1}^{n} z_i^* \leq P_f. \]

We call this constraint \( \text{A|AJ} \) and use the notation GAVC \( \text{A|AJ} \) to refer to the channel with input constraint \( \text{A} \) and jamming power constraint \( \text{AJ} \). Since \( \text{AJ} \) is more restrictive than \( \text{A} \), we must have \( R_{\text{A|AJ}} \leq R_{\text{A|AJ}} \). 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 \( R_{\text{A|AJ}} = R_{\text{A|AJ}} \) and, consequently,

\[ \lambda^{\text{A|AJ}}(R) = \hat{\lambda}^{\text{A|AJ}}(R). \] \hspace{1cm} (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 \textit{mutual information} as the pay-off function (see also [11]). As an example, they considered the channel that we have called the GAVC \( \text{A|AJ} \) (for the special case \( N_r = 0 \)) 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

\[ \frac{n}{2} \log_2 \left( 1 + \frac{P_r}{P_f} \right). \]

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

\[ R < \frac{1}{2} \log_2 \left( 1 + \frac{P_T}{P_f} \right), \]

then \( \lambda^{\text{A|AJ}}(C^*_n) \approx 0 \) is possible. In contrast, however, note that the \( \varepsilon \)-optimal strategies for the game \( \text{A|AJ} \) 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 (A1, AJ) may fail to reflect all the physical constraints that would limit a practical system. As an extreme example let n → +∞. Then the optimal jamming strategy for the case Pr|AJ is of the form s* ~ N(0, P2/ρ) for all time with probability ρ, and s* = 0 for all time with probability 1 − ρ. 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. Csizár 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. Csizár 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 ≥ 0. Player I’s (respectively, player II’s) allowable strategies consist of all nonnegative real-valued random variables X (resp. Y) satisfying EX ≤ a (resp. EY ≤ b).[11] The payoff to player I, when I plays X and II plays Y is

$$\Pr \{ X \geq Y + c \}. \quad (A.1)$$

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:

$$\text{program I: } v = \sup_{X : EX \leq a} \inf_{Y : EY \leq b} \Pr \{ X \geq Y + c \} \quad (A.2a)$$

$$\text{program II: } \bar{v} = \inf_{Y : EY \leq b} \sup_{X : EX \leq a} \Pr \{ X \geq Y + c \}. \quad (A.2b)$$

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 \( \bar{v} \geq v \); if \( \bar{v} = v \), then the game is said to have a value \( v_0 = \bar{v} = v \). A saddle-point solution to this game (if it exists) is a pair of allowable strategies, say \( (X_0, Y_0) \), such that

$$\Pr \{ X \geq Y_0 + c \} \leq \Pr \{ X_0 \geq Y_0 + c \} \quad (A.3)$$

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

$$v_0 = \bar{v} = v = \Pr \{ X_0 \geq Y_0 + c \}, \quad (A.4)$$

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 \( X_0 = F_0 \) and \( Y_0 = G_0 \). These are given in the case \( a = b \) by

$$v_0 = 1 - \frac{b}{2a} \quad (A.5a)$$

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

$$G_0(x) = \left\{ \begin{array}{ll}
\frac{b}{a} U_{[0,2a]}(x) + \left( 1 - \frac{b}{a} \right) \Delta_0(x) ; & \text{if } a \leq b, \\
\frac{a}{2b} & \text{if } a > b.
\end{array} \right. \quad (A.5c)$$

and if \( a < b \),

$$v_0 = \frac{a}{2b} \quad (A.5d)$$

$$F_0(x) = \left\{ \begin{array}{ll}
\frac{a}{b} U_{[0,2b]}(x) + \left( 1 - \frac{a}{b} \right) \Delta_0(x) ; & \text{if } a \leq b, \\
\frac{a}{2b} & \text{if } a > b.
\end{array} \right. \quad (A.5e)$$

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

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

[11] 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 \( \Delta_0(x) \) 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

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

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 $U_{[0,1]}(x) \leq x/d$ when $x \geq 0$, we then obtain

$$\Pr \{ X \geq Y_0 \} = \int_0^\infty G_0(x) \, dF(x)$$

$$\begin{align*}
&= \left(1 - \frac{b}{a}\right) + \frac{b}{a} \int_0^\infty U_{[0,2a]}(x) \, dF(x) \\
&\leq \left(1 - \frac{1}{a}\right) + \frac{2b}{a^2} \int_0^\infty x \, dF(x) \\
&\leq 1 - \frac{1}{a} = \nu_0.
\end{align*} \quad (A.7)$$

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

$$\Pr \{ X_0 \geq Y \} \geq \nu_0. \quad (A.8)$$

Since $\Pr \{ X_0 \geq Y_0 \} = \nu_0$, we conclude that $(X_0, Y_0)$ is a saddle point and $\nu_0$ 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

$$\Pr \{ X \geq Y_0' \} \leq \nu_0. \quad (A.9)$$

for all admissible $X$. Substitution of

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

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

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

$1') \quad G_0'(2a) = 1$

$2') \quad \left( \frac{\beta}{a + \beta} \right) G_0'(a - \alpha) + \left( \frac{\alpha}{\alpha + \beta} \right) G_0'(a + \beta) \leq \nu_0$

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 $(a, \nu_0)$ and is such that

$$G_0'(x) \leq l(x). \quad (A.10)$$

for all $x \geq 0$. To prove this claim, define\(^\dagger\)

$$\mu = \max_{0 \leq \beta \leq a} \left( \frac{G_0'(a + \beta) - \nu_0}{\beta} \right) < +\infty, \quad (A.11)$$

and let $\beta$ attain the maximum. Let $l(x)$ be the line through $(a, \nu_0)$ having slope $\mu$. We know that $G_0(a) \leq \nu_0 = l(a)$ (proof: take $a = \beta = 0$ in $(2')$). By construction $l(x)$ satisfies (A.10) when $x \geq a$ and passes through the point $(a + \beta, G_0(a + \beta))$. Now if

$$G_0'(a - \alpha) \geq l(a - \alpha) \quad (A.12)$$

for some $0 \leq \alpha \leq a$, then $a$ and $\beta$ 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 $\nu_f$ denote the Lebesgue volume of the region $R_f = \{(x, \nu) | 0 \leq x \leq 2a, f(x) \leq y \leq 1\}$. By an elementary fact of probability theory and $1'$ we know that

$$\nu_0 = EY_0' \leq b. \quad (A.13)$$

Equation (A.10) implies that $\nu_0' \geq \nu_f$ and hence

$$\nu_f \leq b. \quad (A.14)$$

Since $l(0) \geq G_0'(0) \geq 0$, $l(2a) \geq G_0'(2a) = 1$, and $l(a) = \nu_0$, $R_f$ is a triangular region and $l(0)$ must be such that $0 \leq l(0) \leq 2 \nu_0 - 1$. By elementary geometry we can show that

$$\nu_f = \frac{a(1 - l(0))}{2(2 \nu_0 - l(0))}. \quad (A.15)$$

for all $0 \leq l(0) \leq 2 \nu_0 - 1$. It is easy to show that (A.15) is a strictly decreasing function of $l(0)$ that attains a minimum value of $\nu_f = b$ when $l(0) = 2 \nu_0 - 1$. Therefore, the only line $l(x)$ that passes through $(a, \nu_0)$ and that does not contradict (A.14) satisfies $l(0) = 2 \nu_0 - 1$, and hence

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

Comparing (A.16) with (A.5e), 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 $G_0' = G_0$, since if $G_0'(x) < G_0(x)$ for some $0 < x < 2a$, then

$$EY_0' = \nu_0' \geq \nu_0' = b,$n

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 $X - F$ can be decomposed in the following way:

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

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

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

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

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

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

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

\[ \]
In terms of the new variables \( p, Z \), and \( W \) the cost function (A.1) becomes
\[
\Pr \{ X \geq Y + c \} = p \Pr \{ Z + c \geq Y + c \}
+ (1 - p) \Pr \{ W \geq Y + c \}
= p \Pr \{ Z > Y \}. \tag{A.18}
\]
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 \( W = 0 \) and
\[
EZ \leq \frac{a}{p} - c = \hat{a}(p).
\]
Using this decomposition, we can reformulate (A.2a) in the following way:

program I: \( v = \sup_{(p, Z)} \inf_{Y : EY \leq b} \frac{p \Pr \{ Z \geq Y \}}{\hat{a}(p)} \) \tag{A.19a}

program II: \( \bar{v} = \inf_{Y : EY \leq b} \sup_{(p, Z)} \frac{p \Pr \{ Z \geq Y \}}{\hat{a}(p)} \) \tag{A.19b}

Games (A.2a) and (A.19a) are equivalent in the following sense: if \( X_0, p_0, Z_0 \) and \( Y_0 \) 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
\[
p \Pr \{ Z \geq Y_0 \} \leq p_0 \Pr \{ Z_0 \geq Y_0 \}
\leq p_0 \Pr \{ Z_0 \geq Y \} \tag{A.20}
\]
for all admissible \( (p, Z, Y) \). Then, in particular, we have
\[
p_0 \Pr \{ Z \geq Y_0 \} \leq p_0 \Pr \{ Z_0 \geq Y_0 \}
\leq p_0 \Pr \{ Z_0 \geq Y \} \tag{A.21}
\]
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
\[
a' = \hat{a}(p_0) = \frac{a}{p_0} - c \quad b' = b \quad c' = 0. \tag{A.22}
\]
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
\[
v_0(p_0) = \begin{cases} 
\frac{p_0}{2b \hat{a}(p_0)} & , \hat{a}(p_0) > b \\
\frac{p_0 \hat{a}(p_0)}{2b} & , \hat{a}(p_0) < b.
\end{cases} \tag{A.23}
\]
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
\[
v_0 = \max_{0 \leq p \leq 1} v(p).
\]
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
\[
p_0 = \begin{cases} 
\frac{a}{c} \left( 1 - \sqrt{\frac{b}{2c + b}} \right) , & a \leq c + \frac{b}{2} \left( 1 + \sqrt{1 + \frac{2c}{b}} \right) \\
1, & a > c + \frac{b}{2} \left( 1 + \sqrt{1 + \frac{2c}{b}} \right).
\end{cases} \tag{A.24}
\]

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) \( \hat{a}(p_0) \geq h \) 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
\[
H_0(x) = U_{[0,2\hat{a}(p_0)]}(x) \tag{A.25a}
\]
\[
G_0(x) = \left( \frac{b}{\hat{a}(p_0)} \right) U_{[0,2\hat{a}(p_0)]}(x)
+ \frac{1}{\hat{a}(p_0)} \Delta_0(x). \tag{A.25b}
\]

Remark: Note that \( a > 0 \) implies that \( \hat{a}(p_0) = (a/p_0) - c > 0 \), 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 \( H_0 = F_0 \). The corresponding value of the game is
\[
v_0 = \begin{cases} 
\frac{a}{c} \left[ 1 + \frac{b}{c} - \sqrt{1 + \frac{2c}{b}} \right] , & a \leq c + \frac{b}{2} \left( 1 + \sqrt{1 + \frac{2c}{b}} \right) \\
1 - \frac{b}{2(a - c)} , & a > c + \frac{b}{2} \left( 1 + \sqrt{1 + \frac{2c}{b}} \right).
\end{cases} \tag{A.26}
\]

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 \( Z = H \) and \( Y = G \). Then
\[
p \Pr \{ Z \geq Y_0 \}
= \int_0^{\infty} G_0(x) \ dH(x)
= p \left( 1 - \frac{b}{\hat{a}(p_0)} \right) + \frac{bp}{2\hat{a}^2(p_0)} \int_0^{\infty} U_{[0,2\hat{a}(p_0)]}(x) \ dH(x)
\leq p \left( 1 - \frac{b}{\hat{a}(p_0)} \right) + \frac{bp}{2\hat{a}^2(p_0)} \int_0^{\infty} x \ dH(x)
\leq p \left( 1 - \frac{b}{\hat{a}(p_0)} \right) + \frac{bp\hat{a}(p)}{2\hat{a}^2(p_0)}
= p \left( 1 - \frac{b}{\hat{a}(p_0)} - \frac{bc}{2\hat{a}^2(p_0)} \right) + \frac{ba}{2\hat{a}^2(p_0)}.
\]
From fact 2) it follows that $pg(p_0) \leq p_0g(p_0)$, and therefore,
\[
p \Pr \{ Z \geq Y \} \leq p_0g(p_0) + \frac{ba}{2a^2(p_0)}
- p_0\left[ \frac{b}{2a(p_0)} \right] - v_0.
\]
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 $(\rho_0, Z_0, Y_0)$ 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
\[
v_0 = \begin{cases}
  a \left[ 1 + c \left( 1 + \frac{2c}{b} \right) \right], \\
  1 - \frac{b}{2(a-c)},
\end{cases}
\]
\[
F_0(x) = p_0U_{[0,2\hat{a}(p_0)]}(x - c) + (1 - p_0)\Delta_0(x) \quad \text{(A.26b)}
\]
\[
G_0(x) = \left( \frac{b}{\hat{a}(p_0)} \right) U_{[0,2\hat{a}(p_0)]}(x) + \left( 1 - \frac{b}{\hat{a}(p_0)} \right) \Delta_0(x) \quad \text{(A.26c)}
\]
where $\hat{a}(p) = a/p - c$ and
\[
\rho_0 = \begin{cases}
  a \left[ 1 - \frac{b}{2c + b} \right], & a \leq c + \frac{b}{2} \left[ 1 + \frac{2c}{b} \right] \\
  1, & a > c + \frac{b}{2} \left[ 1 + \frac{2c}{b} \right]
\end{cases}
\]

**Remark:** Note that some of the foregoing quantities are inde terminant 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