with the capacity 0.461(bits) and the optimal OPD being ${}^t\bar{q} = (0.354, 0.323, 0.323)$, where "$t$" indicates a transposed vector. The set Γ consists of 30 simplicial subchannels with dimension two. The algorithm determines the set Π as the convex-linear combination of three vectors below:

$$\left.\begin{aligned} {}^t\bar{p}^1 &= (0.424, 0.203, 0.373, \ 0 \ , 0, 0, \ 0 \ ), \\ {}^t\bar{p}^2 &= (\ 0 \ , 0.373, 0.203, 0.424, 0, 0, \ 0 \ ), \\ {}^t\bar{p}^3 &= (\ 0 \ , 0.203, 0.373, \ 0 \ , 0, 0, 0.424). \end{aligned}\right\} \tag{5}$$

Finally, the relation between the theorem and the algorithm must be emphasized. The theorem suggests that as far as the capacity and the optimal IPD's are concerned the DMC should most effectively be characterized by those composite simplicial subchannels. The algorithm, however, proposes nothing but a practical device that can easily be implemented using the theorem. It is possible that the simplicial structure inherent in the DMC will play an essential role in current problems regarding multiple-terminal channels.

Acknowledgment

The author appreciates the valuable comments and encouragement given by Professor S. Arimoto. The author is also greatly indebted to Doctor K. Kamoi for his suggestions on the presentation of the manuscript.

References

[1] B. Meister and W. Oettli, "On the capacity of a discrete, constant channel," *Inform. Contr.*, vol. 11, pp. 341–351, 1967.
[2] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[3] S. Arimoto, *Information Theory*. Tokyo: Kyoritsu-Shuppan, 1976 (in Japanese).
[4] ——, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 14–20, Jan. 1972.
[5] R. E. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 460–473, July 1972.

The Gaussian Test Channel with an Intelligent Jammer

TAMER BAŞAR, SENIOR MEMBER, IEEE

***Abstract*—The problem of transmitting a sequence of identically distributed independent Gaussian random variables through a Gaussian memoryless channel with a given input power constraint, in the presence of an intelligent jammer, is considered. The jammer taps the channel and feeds back a signal, at a given energy level, for the purpose of jamming the transmitting sequence. Under a square-difference distortion measure which is to be maximized by the jammer and to be minimized by the transmitter and the receiver, this correspondence obtains the complete set of optimal (saddle-point) policies. The solution is essentially unique, and it is structurally different in three different regions in the parameter space, which are determined by the signal-to-noise ratios and relative magnitudes of the noise variances. The best (maximin) policy of the jammer is either to choose a linear function of the measurement he receives through channel-tapping, or to choose, in addition (and additively), an independent Gaussian noise sequence, depending on the region where the parameters lie. The optimal (minimax) policy of the transmitter is to amplify the input sequence to the given power level by a linear transformation, and that of the receiver is to use a Bayes estimator.**

Manuscript received June 29, 1981. This work was supported in part by the Joint Services Electronics Program under Contract N00014-79-C-0424 and in part by the Office of Naval Research under Contract N00014-81-K-0014. This work was presented in part at the IEEE International Symposium on Information Theory, Les Arcs, France, June 1982.
The author is with the Coordinated Science Laboratory, University of Illinois, 1101 W. Springfield Ave., Urbana, IL 61801.

I. Introduction and Problem Description

The communication system depicted in Fig. 1 represents an extended version of the so-called Gaussian test channel (cf. [1]), which also includes an intelligent jammer who has access to a (possibly) noise-corrupted version of the signal to be transmitted through a Gaussian channel. More specifically, a Gaussian random variable[1] of zero mean and unit variance [denoted $u \sim N(0,1)$] is to be transmitted through a Gaussian channel with input energy constraint c^2, and additive noise $(w = w_1 + w_2)$ with total noise variance $\xi = \xi_1 + \xi_2$. Let the transmitter strategy be denoted $\gamma(\cdot)$, which is an element of the space Γ_t of real-valued Borel measurable functions satisfying the power constraint $E\{[\gamma(u)]^2\} \le c^2$. The jammer has access to a noise-corrupted version of

$$x \triangleq \gamma(u) + w_1, \tag{1}$$

denoted

$$y = x + v, \tag{2}$$

where $v \sim N(0, \sigma)$, all random variables (u, w_1, w_2, and v) are statistically independent, and $\xi_1 \ge 0$, $\xi_2 \ge 0$, and $\sigma \ge 0$. Based on the observed value of y, the jammer feeds back a second-order random variable $\nu = \beta(y)$ to the channel, so that the input to the receiver is now

$$z = x + \nu + w_2. \tag{3}$$

The random variable ν is correlated with y, but it is not necessarily determined through a deterministic transformation on y [i.e., $\beta(\cdot)$ is in general a random mapping]; furthermore, it satisfies the energy constraint $E[\nu^2] \le k^2$. Let us denote the class of all associated probability measures μ for the jammer by M_j. Finally, the receiver applies a Borel-measurable transformation $\delta(\cdot)$ on its input z, so as to produce an estimate $\hat{u}$ of u, by minimizing the square-difference distortion measure

$$R(\gamma, \delta, \mu) = \int_{-\infty}^{\infty} E\{[\delta(z) - u]^2 \mid \nu\} \, d\mu(\nu). \tag{4}$$

Denote the class of all Borel-measurable mappings $\delta(\cdot)$, to be used as an estimator for u, by Γ_r. Then, the transmitter and the receiver seek to minimize R by a proper choice of $\gamma \in \Gamma_t$ and $\delta \in \Gamma_r$, respectively, and the jammer seeks to maximize the same quantity by his choice of $\mu \in M_j$. Since there is a complete conflict of interests in this communication problem, an "optimal" transmitter-receiver-jammer policy would be the *saddle-point solution* ($\gamma^* \in \Gamma_t$, $\delta^* \in \Gamma_r$, $\mu^* \in M_j$) satisfying the set of inequalities

$$R(\gamma^*, \delta^*, \mu) \le R(\gamma^*, \delta^*, \mu^*) \le R(\gamma, \delta, \mu^*), \qquad \forall \gamma \in \Gamma_t, \ \delta \in \Gamma_r, \ \mu \in M_j. \tag{5}$$

The maximin policy μ^* is also known as a *least-favorable probability measure* for the jammer [3].

In this correspondence, we verify existence and "essential" uniqueness (up to the sign) of the saddle-point solution, and determine the corresponding policies explicitly and in analytic form. The main result is presented in the next section and, in particular, in Theorem 1. The structure of the solution is different in three different regions of the parameter space: in one of these regions the solution is trivial, and in the other two regions (which

[1] This single variable can be replaced with a sequence of independent identically distributed Gaussian random variables, without altering the results of this correspondence.

0018-9448/83/0100-0152$01.00

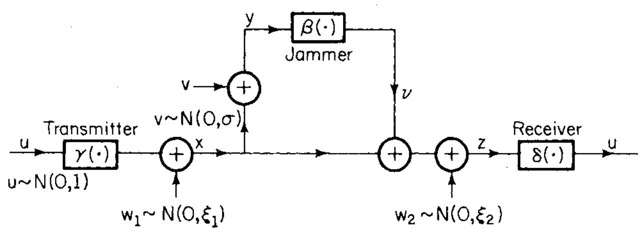

Fig. 1. The Gaussian test channel with an intelligent jammer.

are covered by Theorem 1) the saddle-point policy for the transmitter is to amplify the input signal to the maximum power level through a linear transformation. The saddle-point policy for the jammer is to choose a Gaussian random variable (or a sequence of independent identically distributed Gaussian random variables, if the input is also a sequence) which is correlated with the input signal; the nature of this correlation turns out to be different in the two regions of interest. For the receiver, the optimal policy is to use a Bayes estimator. The proof of this result, which is given in Section II, is rather involved, and in places it requires some rather intricate arguments, but it is essentially a proof of the "verification" type.

Section III of the paper includes some discussion on special cases and on some related results in the literature. The Appendix provides proofs of two lemmas which are used in the derivation in Section II.

II. Derivation of the Saddle-Point Solution

In this section, we obtain the saddle-point solution of the problem formulated in Section I, for all values of the parameters, $c \geq 0$, $k \geq 0$, $\xi_1 \geq 0$, $\xi_2 \geq 0$, $\sigma \geq 0$. There exists, however, a region in this parameter space, in which the problem is trivial, in the sense that the jammer has the power to do the best he can possibly do, by cancelling out the signal component $\gamma(x)$ in the received signal z. Specifically, consider region

R1 $$k^2 \geq c^2 + \xi_1 + \sigma,$$

where the deterministic feedback policy

$$\beta^*(y) = -y \tag{6}$$

is feasible for the jammer, which leads to $z = w_2 - v$ and thereby to

$$\delta^*(z) = 0, \tag{7}$$

resulting in a maximum distortion level of

$$R(\gamma, \delta^*, \beta^*) = 1.$$

Note that the choice of any specific coding strategy is irrelevant here, since they all lead to the same maximum distortion level, under (6) and (7). Hence, for this special case, the pair (δ^*, β^*) as given by (6)–(7) constitutes a (trivial) saddle-point solution (and the only one) for any choice of $\gamma \in \Gamma_t$.

Leaving this "uninteresting" case aside, we henceforth restrict our analysis to parameter region

R2 $$k^2 < c^2 + \xi_1 + \sigma,$$

which we further decompose into two subregions characterized by additional constraints

R3 $$k^2 - \frac{(c^2 + \xi_1)k}{(c^2 + \xi_1 + \sigma)^{1/2}} + \xi_2 > 0,$$

and

R4 $$k^2 - \frac{(c^2 + \xi_1)k}{(c^2 + \xi_1 + \sigma)^{1/2}} + \xi_2 \leq 0.$$

The complete solution to the problem is now provided in Theorem 1 below, after introducing some notation and terminology.

Preliminary Notation for Theorem 1: Introduce the scalar parameters λ and t by

$$\lambda = -k/(c^2 + \xi_1 + \sigma)^{1/2}, \tag{8a}$$

$$t = 1 - \left\{(k^2 + \xi_2)^2 (c^2 + \xi_1 + \sigma)/\left[k^2 (c^2 + \xi_1)^2\right]\right\}, \tag{8b}$$

and let η denote a Gaussian random variable with mean zero and variance tk^2, i.e.,

$$\eta \sim N(0, tk^2) \tag{9}$$

whenever $t \geq 0$.

Theorem 1: In region *R2*, the communication problem admits two saddle-point solutions $(\gamma^*, \delta^*, \mu^*)$ and $(-\gamma^*, -\delta^*, \mu^*)$, where

i) $$\gamma^*(u) = cu, \tag{10}$$

ii) μ^* is the Gaussian probability measure associated with the random variable

$$v = \beta^*(y) = \begin{cases} \lambda y, & \text{in } R3, \\ \lambda(1-t)^{1/2} y + \eta, & \text{in } R4, \end{cases} \tag{11}$$

where $t \in [0, 1]$ in *R4*, and $\eta \sim N(0, tk^2)$.

iii) δ^* is the Bayes estimator for u under the least favorable distribution μ^*, computed as

$$\delta^*(z) = \begin{cases} \left\{c(1+\lambda)/\left[(1+\lambda)^2(c^2+\xi_1) + \lambda^2\sigma + \xi_2\right]\right\} z, \\ \qquad \text{in } R3, \\ \left[c/\left(c^2 + \xi_1\right)\right] z, \qquad \text{in } R4. \end{cases} \tag{12}$$

Proof: The proof proceeds in two steps. We first establish validity of the right-hand side (RHS) inequality of (5) when μ^* is determined by (11), and then prove the left-hand side (LHS) inequality of (5) when γ^* and δ^* are given by (10) and (12), respectively. Finally we discuss the "essential uniqueness" property of the saddle-point solution.

A. The RHS Inequality

Region R2 ∩ R3: Suppose that μ^* is determined by (11) and the parameter values lie in region *R2 ∩ R3*. Then, the RHS inequality of (5) dictates a combined coding–decoding problem, with the channel output (equivalently, receiver input) being (from (3))

$$z = (1+\lambda)\gamma(u) + (1+\lambda)w_1 + \lambda v + w_2,$$

where $0 < (1+\lambda) < 1$ from (8a), since we are in region *R2*. Let $\tilde{\gamma}(u) \triangleq (1+\lambda)\gamma(u)$. Then, the problem we face is the Gaussian test channel of Fig. 2 with square-difference distortion, Gaussian channel noise [with mean zero and variance $(1+\lambda)^2\xi_1 + \lambda^2\sigma + \xi_2$], and channel-input energy constraint

$$E\{[\tilde{\gamma}(u)]^2\} \leq (1+\lambda)^2 c^2.$$

It is well-known that this problem admits a linear solution (cf. [1]), which is that the best coding scheme is to amplify the input u to the maximum available power level, i.e., $\tilde{\gamma}^*(u) = c(1+\lambda)u$, $c > 0$, and to choose the quadratic distortion minimizer δ as the Bayes estimator

$$\begin{aligned} \delta^*(z) &= E[u \mid z] \\ &= \left\{(1+\lambda)c/\left[(1+\lambda)^2(c^2+\xi_1) + \lambda^2\sigma + \xi_2\right]\right\} z, \end{aligned}$$

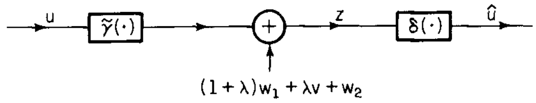

Fig. 2. The reduced Gaussian channel used in verifying the RHS inequality in region $R2 \cap R3$.

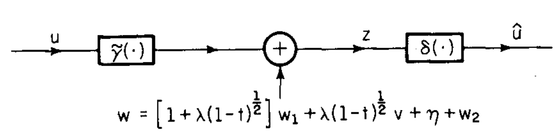

Fig. 3. The reduced Gaussian channel used in verifying the RHS inequality in region $R2 \cap R4$.

which is precisely (12) in the region $R2 \cap R3$. Moreover, since $1+\lambda>0$,

$$\gamma^*(u)=\frac{1}{1+\lambda}\bar{\gamma}^*(u)=cu,$$

which is the same as (10). Hence, we have established the validity of the RHS of (5), for the solution presented in Theorem 1, in the region $R2 \cap R3$. Note that, yet another possible coding policy for the transmitter would be

$$\gamma^*(u)=-cu, \qquad c>0,$$

i.e., amplification by a negative factor, but since this leads to the same minimum distortion level, we adopt the convention of choosing only the positive amplification factor and call such a solution "*essentially unique*."

Region $R2 \cap R4$: Now suppose that the parameter values lie in the region $R2 \cap R4$. A similar reasoning as above again leads to a Gaussian test channel (of Fig. 3) with the channel-input energy constraint being

$$E\{[\bar{\gamma}(u)]^2\} \le c^2[1+\lambda(1-t)^{1/2}]^2,$$

where $\bar{\gamma}(u) \triangleq [1+\lambda(1-t)^{1/2}]\gamma(u)$. The Gaussian channel noise has its mean zero and variance

$$\operatorname{var}(w)=\left[1+\lambda(1-t)^{1/2}\right]^2\xi_1+\lambda^2(1-t)\sigma+tk^2+\xi_2.$$

Substituting for λ and t from (8a) and (8b), respectively, we can evaluate the latter expression to be

$$\operatorname{var}(w)=\xi_1-\xi_2-k^2-\frac{(k^2+\xi_2)^2c^2}{(c^2+\xi_1)^2}+\frac{2(k^2+\xi_2)c^2}{(c^2+\xi_1)}.$$

Likewise, the input power constraint can be written as

$$\operatorname{var}(\bar{\gamma}(u)) \le c^2[c^2+\xi_1-\xi_2-k^2]^2/(c^2+\xi_1)^2 \triangleq m$$

and furthermore $m+\operatorname{var}(w)=c^2+\xi_1+\xi_2-k^2$.

Again using the well-known result for Gaussian test channels, we obtain the essentially unique solution to be

$$\bar{\gamma}^*(u)=c\left[1+\lambda(1-t)^{1/2}\right]u,$$

$$\delta^*(z)=E[u\,|\,z]=\frac{E[uz]}{\operatorname{var} z}z$$

$$=\frac{m^{1/2}}{m+\operatorname{var}(w)}z=\frac{c}{c^2+\xi_1}z,$$

with the latter expression verifying (12) in region $R4$. Furthermore, since $1+\lambda(1-t)^{1/2}$ is nonsingular (in fact, it lies in $(0,1]$ under the assertion that $t\in[0,1]$),

$$\gamma^*(u)=\bar{\gamma}^*(u)/\left[1+\lambda(1-t)^{1/2}\right]=cu$$

which verifies (10), also in the region $R4$. This then completes the proof of parts i) and iii) of the Theorem, under the assertions that a least favorable distribution μ^* for ν exists, as given by (11), and $t\in[0,1]$ in $R4$. The former assertion is verified next, in the sequel, and the latter one is verified in Lemma 2 in the Appendix.

B. The LHS Inequality

Region $R2 \cap R3$: Suppose now that γ^* and δ^* are given by (10) and (12), respectively, and the parameter values lie in $R2 \cap R3$. Then, the LHS inequality of (5) dictates the following optimization problem for the jammer:

$$\max_{\mu\in M_j}\int_{-\infty}^{\infty}E\{[\alpha z-u]^2\,|\,\nu\}\,d\mu(\nu), \tag{13}$$

where

$$z=cu+w_1+\nu+w_2,$$

$$\alpha=c(1+\lambda)/\left[(1+\lambda)^2(c^2+\xi_1)+\lambda^2\sigma+\xi_2\right]. \tag{14}$$

Note that (13) can also be written as

$$\max_{\mu\in M_j}\int_{-\infty}^{\infty}E_y\{E\{[\alpha z-u]^2\,|\,\nu,y\}\}\,d\mu(\nu)$$

$$=\max_{\mu\in M_j}E_y\left\{\int_{-\infty}^{\infty}E\{\alpha^2z^2-2\alpha zu\,|\,\nu,y\}\,d\mu(\nu\,|\,y)\right\}+1,$$

where $y=cu+w_1+v$.

Furthermore, since w_2 is independent of u, w_1, and ν, u is independent of w_1, and they both have zero mean, the latter expression can be simplified to

$$\max_{\mu\in M_j}E_y\left\{\int_{-\infty}^{\infty}E\{\alpha^2\nu^2+\alpha^2 2\nu(cu+w_1)\right.$$

$$+\alpha^2(c^2u^2+w_1^2+w_2^2)-2\alpha cu^2$$

$$\left.-2\alpha\nu u\,|\,\nu,y\}\,d\mu(\nu\,|\,y)\right\}+1$$

$$=\max_{\mu\in M_j}E_y\left\{\int_{-\infty}^{\infty}[\alpha^2\nu^2-2\nu\pi(y)]\,d\mu(\nu\,|\,y)\right\}$$

$$+(\alpha c-1)^2+\alpha^2(\xi_1+\xi_2)$$

where

$$\pi(y)=\alpha E[u\,|\,y]-\alpha^2cE[u\,|\,y]-\alpha^2E[w_1\,|\,y]$$

$$=\{[\alpha(1-\alpha c)c/(c^2+\xi_1+\sigma)]$$

$$-[\alpha^2\xi_1/(c^2+\xi_1+\sigma)]\}y$$

$$=[-\alpha^2+\alpha(c+\alpha\sigma)/(c^2+\xi_1+\sigma)]y\triangleq py. \tag{15}$$

Hence, we may confine attention to the maximization problem

$$J=\max_{\mu\in M_j}E_y\left\{\int_{-\infty}^{\infty}[\alpha^2\nu^2-2\nu\pi(y)]\,d\mu(\nu\,|\,y)\right\}, \tag{16a}$$

which is in fact invariant under the transformation $c\to -c$, and is therefore also the maximization problem (for the jammer) corresponding to the pair $(-\gamma^*,-\delta^*)$. Now, by the Cauchy-Schwartz inequality (cf. [2]), (16a) can be bound above by

$$J\le\max_{\mu\in M_j}\left\{\int_{-\infty}^{\infty}\alpha^2\nu^2\,d\mu(\nu)\right.$$

$$\left.+2\left[\int_{-\infty}^{\infty}\nu^2\,d\mu(\nu)\right]^{1/2}|\,E_y\{|\pi(y)|^2\}\,|^{1/2}\right\}$$

and since $\mu \in M_j$, this can further be bound above by

$$J \leq \alpha^2 k^2 + 2k \left| E_y\{|\pi(y)|^2\} \right|^{1/2}. \tag{16b}$$

But, provided that

$$E_y\{[|\pi(y)|^2]\} \neq 0,$$

this upper bound is attained *uniquely* if we choose, in (16a), $\mu(\nu \mid y)$ to be the one-point conditional probability measure corresponding to the strategy

$$\nu = \beta^*(y) = -\left[k/\left[E_y\{|\pi(y)|^2\}\right]^{1/2}\right]\pi(y), \tag{17}$$

which may be verified by direct substitution of (17) into (16a) and by comparing the resulting expression with the upper bound (16b). Now, what remains to be shown is that (17) is equivalent to (11) in region $R2 \cap R3$, and that $E_y\{|\pi(y)|^2\} > 0$. Lemma 1 in the Appendix proves that the coefficient p of y in (15) is in fact positive in the region $R2 \cap R3$, and hence the latter requirement is readily fulfilled. Furthermore, since

$$\frac{-k}{\left[E_y\{|\pi(y)|^2\}\right]^{1/2}}\pi(y) = \frac{-kp}{|p|\,|\operatorname{var}(y)|^{1/2}}y = \lambda y,$$

from (8a) and the property that $p > 0$, the former requirement is also satisfied. This then completes the verification of the LHS inequality of (5), and thereby verification of the theorem, for the region $R2 \cap R3$.

Region $R2 \cap R4$: We now verify the LHS inequality of (5) when the parameters belong to the region $R2 \cap R4$. What replaces the maximization problem (13) in this case is

$$\max_{\mu \in M_j} \int_{-\infty}^{\infty} E\left\{\left[\frac{c}{c^2 + \xi_1}z - u\right]^2 \middle| \nu\right\} d\mu(\nu),$$

which can be rewritten as (through some straightforward manipulations)

$$\frac{c^2}{(c^2 + \xi_1)^2}\left\{\max_{\mu \in M_j} E_y\left[\int_{-\infty}^{\infty}\left\{\nu^2 - 2\nu E\left[\frac{\xi_1}{c}u - w_1 \mid y\right]\right\} d\mu(\nu \mid y)\right] + \xi_1 + \xi_2 + \frac{\xi_1^2}{c^2}\right\}, \tag{18}$$

which is an expression that is invariant under the transformation $c \to -c$. Note that

$$E\left[\frac{\xi_1}{c}u - w_1 \mid y\right] = \frac{\xi_1}{c}\cdot\frac{c}{c^2 + \xi_1 + \sigma}y - \frac{\xi_1}{c^2 + \xi_1 + \sigma}y = 0,$$

and therefore the maximizing solution is any probability measure μ, with the property

$$\int_{-\infty}^{\infty} \nu^2\, d\mu(\nu) = k^2.$$

Let us now investigate whether μ^*, determined by (11), is one such measure in region $R2 \cap R4$. Towards this end, it suffices to show that

$$\operatorname{var}\left[\lambda(1-t)^{1/2}y + \eta\right] = k^2$$

and

$$t \in [0, 1].$$

The latter is shown in Lemma 2, in the Appendix. For the former, simply note that, because of independence and the fact that $\eta \sim N(0, tk^2)$,

$$\begin{aligned}\operatorname{var}\left[\lambda(1-t)^{1/2}y + \eta\right] &= \lambda^2(1-t)\operatorname{var}(y) + \operatorname{var}(\eta)\\ &= \lambda^2(1-t)\left(c^2 + \xi_1 + \sigma\right) + tk^2\\ &= k^2(1-t) + tk^2 = k^2,\end{aligned}$$

thus establishing the desired result. As a parenthetical remark, we should mention that the estimator (12) in region $R2 \cap R4$ may also be viewed as an *equalizer* decision rule (see [3]) since the conditional (on μ) risk function corresponding to it is a constant on ∂M_j, the boundary of M_j. (Note that in this interpretation, elements of ∂M_j are the decision variables of the jammer, and we have to introduce probability measures on ∂M_j.) Hence, the minimax (saddle-point) property of δ^* in $R2 \cap R4$ can also be verified (with γ^* fixed, as given) by resorting to a well-known property of equalizer decision rules when they are also Bayes with respect to a least favorable probability measure, which in this case is the one-point distribution on ∂M_j that selects the Gaussian random variable $\lambda(1-t)^{1/2}y + \eta$. (See [4], [5].) The proof given here seems to be more suited to the problem under consideration since i) it does not require additional probability measures to be defined on M_j, and ii) it also establishes the optimality of γ^*.

To *recapitulate*, we have verified existence of a saddle-point solution (10)–(12) for the communication problem under consideration, in the parameter region $R2$. The analysis also readily leads to the conclusion that in addition to (10)–(12), the triple $(-\gamma^*, -\delta^*, \mu^*)$ also provides a saddle-point solution, naturally leading to the same saddle-point value for R. The question now arises as to whether other saddle-point equilibria exist. In region $R2 \cap R3$, there is clearly no other saddle point, since the maximization problem (16a) (which corresponds to both (γ^*, δ^*) and $(-\gamma^*, -\delta^*)$) admits a unique solution, thereby eliminating the possibility of multiple saddle-point policies for the jammer. (Otherwise, interchangeability property of saddle points (cf. [6]) would lead to a contradiction.) In the remaining part of $R2$, i.e., $R2 \cap R4$, however, the issue is more subtle. Since the maximization problem (18) is invariant under different choices of probability measures from ∂M_j, the LHS inequality of (5) clearly does not admit a unique solution—in fact, all second-order probability measures with first moment zero and second moment equal to k^2 constitute a solution. But, for any one of these to constitute a saddle-point policy for the jammer, it has to be in equilibrium with (γ^*, δ^*), because of the interchangeability property of saddle-point equilibria. This further implies that, with γ^* fixed as given, δ^* has to be Bayes with respect to that least-favorable distribution. Since δ^* is a linear estimator and all random variables are Gaussian, this requires the chosen element of ∂M_j to be a Gaussian probability measure, and some further analysis reveals that (11) is in fact the only such element. □

Some of the expressions derived in the proof of Theorem 1 now lead to the following Corollary which gives the saddle-point values in different regions.

Corollary 1: The saddle-point value (R^*) of $R(\gamma, \delta, \mu)$ in different regions is given as follows.

$R1$: $R^* = 1$;

$$R2 \cap R3: \quad R^* = (\alpha c - 1)^2 + \alpha^2\left(\xi_1 + \xi_2 + k^2\right) + 2kp^2\left(c^2 + \xi_1 + \sigma\right);$$

and

$$R2 \cap R4: \quad R^* = \frac{c^2}{\left(c^2 + \xi_1\right)^2}\left(k^2 + \xi_1 + \xi_2 + \frac{\xi_1^2}{c^2}\right),$$

where α and p are defined by (14) and (15), respectively. □

III. Discussion of Some Special Cases, and Concluding Remarks

The general solution to the communication problem of Fig. 1 has the property that it is structurally different in the two regions of interest, with the dividing "line" between these two regions being a hyperplane determined by the allowable power levels for the transmitter and jammer, and the noise intensities in the main channel and the jammer's wiretap link. In particular, if the transmitter's allowable power level (c^2) is larger than that of the jammer (k^2), we stay in region $R2$, and if this difference is sufficiently large the jammer's maximin policy is an additive mixture of a linear transformation on his measurement and an independent Gaussian random variable, whereas if the difference is small it is more likely (depending on the values of other parameters) that his maximin policy will be only a linear transformation on his measurement.

If the wiretapping channel noise variance (σ) is sufficiently large, the parameter region is $R2 \cup R3$, and hence the optimum strategy for the jammer is a linear policy—which may seem, at first sight, to be somewhat counter-intuitive, since the information contained in y (concerning u) is quite unreliable. However, some scrutiny reveals that the jammer, in fact, uses this noisy measurement as a source of noise in order to jam the transmission channel. This makes particular sense in the limiting case $\sigma \to \infty$, when the optimal jamming policy is to choose μ^* as a Gaussian distribution with mean zero and variance k^2, which should be independent of the transmitter output. This conclusion for the limiting case corroborates a result obtained in [7] in a somewhat different context. More specifically, this recent reference addresses the problem of obtaining optimal policies in the presence of jamming, when jammer's policies (considered as random variables) are forced to be independent of the transmitter outputs, and the loss function (to be minimaximized) is taken as the mutual information between the transmitter output and the receiver input. In this framework, McEliece and Stark solve in [7], as an application of their general approach, the communication problem depicted in Fig. 1, but without the tapping channel, and arrive at the conclusion that the least-favorable distribution for ν is a Gaussian distribution. Hence, the two seemingly different problems (with different loss functions—square-difference distortion and mutual information) admit the same saddle-point solution in the presence of an independent jammer strategy. (This equivalence can in fact be verified directly by making use of some inequalities of Shannon [8] on mutual information.) But, this equivalence does not directly extend to the communication system considered in this correspondence, and derivation of the saddle-point solution of the communication system of Fig. 1 when the loss function is taken as the mutual information between u and z remains today as a challenging problem.[2]

There exist quite a few results in the literature on worst case designs, wherein the Gaussian distribution has been proven to be the least-favorable distribution (such as the cases of entropy maximization [9], Fisher-information minimization [10], or minimax estimation problems [11], [12]), and the present paper adds to this list a new class of problems not considered heretofore. We should note, however, that if the input sequence in Fig. 1 is vector-valued or the number of channels is more than one, the saddle-point solution will no longer be linear-Gaussian (i.e., the solution of this correspondence does not carry over to the vector case), since the counterpart of the Gaussian test channel does not admit a simple linear coding scheme in the vector case [13].

Appendix

In this appendix, we provide proofs for the two lemmas which were used in the proof of Theorem 1 in Section II.

[2]For some related results see references [14]–[15], which were recently brought to the attention of the author by Dr. Blackman.

Lemma 1: $p = -\alpha^2 + [\alpha(c + \alpha\sigma)/(c^2 + \xi_1 + \sigma)] > 0$, in $R2 \cap R3$, where α is defined by (14).

Proof: Through straightforward substitution from (14) and (8a), and some manipulations,

$$\begin{aligned} p &= \frac{\alpha^2}{c^2 + \xi_1 + \sigma}\left[\frac{c}{\alpha} + \sigma - c^2 - \xi_1 - \sigma\right] \\ &= \frac{\alpha^2}{c^2 + \xi_1 + \sigma}\Big[(1 + \lambda)(c^2 + \xi_1) \\ &\quad + \frac{\lambda^2\sigma}{1 + \lambda} + \frac{\xi_2}{1 + \lambda} - c^2 - \xi_1\Big] \\ &= \frac{\alpha^2}{(1 + \lambda)(c^2 + \xi_1 + \sigma)} \\ &\quad \cdot \left[\lambda^2(c^2 + \xi_1 + \sigma) + \lambda(c^2 + \xi_1) + \xi_2\right] \\ &= \frac{\alpha^2}{(1 + \lambda)(c^2 + \xi_1 + \sigma)} \\ &\quad \cdot \left[k^2 - \frac{k(c^2 + \xi_1)}{(c^2 + \xi_1 + \sigma)^{1/2}} + \xi_2\right] \\ &> 0 \end{aligned}$$

since $1 + \lambda > 0$ in $R2$, and the last multiplicative term is positive in $R3$. □

Lemma 2: $t \in [0, 1]$ in $R2 \cap R4$, where t is defined by (8b).

Proof: Starting with the inequality that determines $R4$,

$$k^2 + \xi_2 \le \frac{(c^2 + \xi_1)}{(c^2 + \xi_1 + \sigma)^{1/2}} k$$

$$\Leftrightarrow (k^2 + \xi_2)(c^2 + \xi_1 + \sigma)^{1/2} \le (c^2 + \xi_1)k,$$

and squaring both sides

$$(k^2 + \xi_2)^2(c^2 + \xi_1 + \sigma) \le (c^2 + \xi_1)^2 k^2$$

we arrive at

$$\frac{(k^2 + \xi_2)^2(c^2 + \xi_1 + \sigma)}{(c^2 + \xi_1)^2 k^2} \le 1.$$

Since this latter expression is equal to $1 - t$(from (8b)), and it is also positive, the desired property follows. □

References

[1] R. G. Gallager, *Information Theory and Reliable Communication.* New York: John Wiley, 1968, pp. 475–482.
[2] D. G. Luenberger, *Optimization by Vector Space Methods.* New York: John Wiley, 1969, pp. 47, 78–82.
[3] T. S. Ferguson, *Mathematical Statistics, A Decision Theoretic Approach.* New York: Academic, 1967.
[4] E. L. Lehmann, *Notes on the Theory of Estimation.* Berkeley: Univ. of California, 1966.
[5] A. Wald, *Statistical Decision Functions.* New York: Wiley, 1950.
[6] G. Owen, *Game Theory.* Philadelphia: Saunders, 1968, p. 13.
[7] R. J. McEliece and W. E. Stark, "An information theoretic study of communication in the presence of jamming," in *Proc. 1981 IEEE Int. Conf. Commun.*, Denver, CO, June 15–18, 1981.
[8] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.
[9] A. M. Kagan, Yu. V. Linnik, and C. R. Rao, *Characterization Problems in Mathematical Statistics.* New York: Wiley, 1973.
[10] T. Başar, "Optimum Fisherian information for multivariate distributions," *Ann. Statist.*, vol. 5, no. 6, pp. 1240–1244, 1977.
[11] T. Başar and M. Mintz, "On a minimax estimate for the mean of a normal random vector under a generalized quadratic loss function," *Ann. Statist.*, vol. 1, no. 1, pp. 127–138, 1973.

[12] ——, "Minimax estimation under generalized quadratic loss," *Proc. 1971 IEEE Conf. Decision Contr.*, Miami, FL, pp. 456–461, Dec. 1971.
[13] Y. C. Ho, M. P. Kastner, and E. Wong, "Teams, signaling, and information theory," *IEEE Trans. Automat. Contr.*, vol. AC-23, no. 2, pp. 305–312, 1978.
[14] N. M. Blackman, "On the capacity of a band-limited channel perturbed by statistically dependent interference," *IRE Trans. Inform. Theory*, vol. IT-8, no. 1, pp. 48–55, Jan. 1962.
[15] ——, "The effect of statistically dependent interference upon channel capacity," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 53–57, Sept. 1962.

The Sampling Window

JOHN J. KNAB, MEMBER, IEEE

Abstract—**The concept of the sampling window is introduced for the central interpolation of finite energy band-limited functions. The sampling window does not increase the rate of convergence of the truncation error series, as do various convergence factors, but does significantly reduce the truncation-error bound.**

I. Introduction

The cardinal series was first introduced by Whittaker [1] for the interpolation of finite energy band-limited functions. Later, convergence factors were introduced [2], [3] to increase the rate of convergence of the cardinal series and thus reduce the truncation-error bound. This correspondence introduces the concept of a sampling window which reduces the truncation-error bound but does not increase the rate of convergence of the cardinal series. In many cases, however, this new truncation-error bound is less than bounds when using a convergence factor.

We define the function class $E(\omega_0)$ as containing all functions having energy E and being bandlimited to ω_0. That is, if $f(t)$ is in $E(\omega_0)$ then

$$f(t) = \frac{1}{2\pi}\int_{-\omega_0}^{\omega_0} F(\omega)e^{i\omega t}\,d\omega$$

for some $F(\omega)$ in $L_2(-\omega_0, \omega_0)$, i.e., $\int_{-\omega_0}^{\omega_0} |F(\omega)|^2\,d\omega = 2\pi E < \infty$. We also define the function class $M(\omega_0)$ as containing all functions bounded by M and being bandlimited to ω_0. If $f(t)$ is in $M(\omega_0)$, then for complex z, $f(z)$ is of exponential type ω_0 [4]. It should be noted that any function in $E(\omega_0)$ is necessarily in $M(\omega_0)$ with $M = \sqrt{\omega_0 E/\pi}$ since if $f(t)$ is in $E(\omega_0)$ then $|f(t)| \leq \sqrt{\omega_0 E/\pi}$ [5, p. 213], [6].

II. Sampling Theorem Using Window

We first quote the well-known sampling theorem.

Theorem 1: If $f(t)$ is in $E(\omega_B)$, then

$$f(t) = \sum_{k=-\infty}^{\infty} f(k\tau)\frac{\sin \omega_B(t-k\tau)}{\omega_B(t-k\tau)}, \tag{1}$$

where $\tau = \pi/\omega_B$. A simple proof is given by Brown [7]. This series (1) is often called the cardinal series or the Shannon sampling theorem [8].

We now prove a theorem which introduces the concept of a sampling window.

Theorem 2: If $f(t)$ is in $E(\omega_0)$, δ is any real number between zero and unity, i.e., $0 \leq \delta < 1$, and $h \triangleq \pi(1-\delta)/\omega_0$, then

$$f(t) \triangleq \sum_{k=-\infty}^{\infty} f(kh)\frac{\sin \frac{\omega_0}{1-\delta}(t-kh)}{\frac{\omega_0}{1-\delta}(t-kh)}\psi(t-kh) \tag{2}$$

for any $\psi(t)$ in $M(\Omega)$ where $\Omega = \delta\omega_0/(1-\delta)$ and $\psi(0) = 1$.

Proof: For any fixed real x, $f(t)\psi(x-t)$ has finite energy and by the well-known Paley–Wiener theorem [5, p. 103] is bandlimited to $\omega_0 + \Omega = \omega_0/(1-\delta)$. Hence by Theorem 1, with $\omega_B = \omega_0/(1-\delta)$ and $\tau = \pi(1-\delta)/\omega_0$, we get

$$f(t)\psi(x-t) = \sum_{k=-\infty}^{\infty} f(k\tau)\psi(x-k\tau)\frac{\sin \frac{\omega_0}{1-\delta}(t-k\tau)}{\frac{\omega_0}{1-\delta}(t-k\tau)}.$$

Setting $x = t$ and $h = \tau = \pi(1-\delta)/\omega_0$ we get the final result (2). The function $\psi(t)$ will be called the sampling window. Note that $\psi(t)$ is not forced to approach zero for large $|t|$ as was a requirement for $\psi(t)$ in all previous developments along these lines [2], [3].

It is not possible in practice to reconstruct $f(t)$ using either (1) or (2), since an infinite number of samples would be needed. The truncation error for central interpolation is defined as

$$e_N(t) \triangleq f(t) - \hat{f}(t) \qquad \text{for } |t/h| \leq 1/2 \tag{3}$$

with

$$\hat{f}_N(t) \triangleq \sum_{k=-N}^{N} f(kh)\frac{\sin \frac{\omega_0}{1-\delta}(t-kh)}{\frac{\omega_0}{1-\delta}(t-kh)}\psi(t-kh). \tag{4}$$

Hence

$$|e_N(t)| = \left|\sum_{|k|=N+1}^{\infty} f(kh)\frac{\sin \frac{\omega_0}{1-\delta}(t-kh)}{\frac{\omega_0}{1-\delta}(t-kh)}\psi(t-kh)\right|. \tag{5}$$

In all previous works, $\psi(t)$ was chosen so that the convergence of the error series (5) was increased; therefore, $\psi(t)$ was called a convergence factor. For example, the self-truncating (ST) convergence factor [2] has the form

$$\psi(t) = \left[\frac{\sin \frac{\delta\omega_0}{(1-\delta)m}t}{\frac{\delta\omega_0}{(1-\delta)m}t}\right]^m,$$

where m is a positive integer. Hence $\psi(t)$ for this case is $O(t^{-m})$. The approximate prolate (AP) convergence factor [3] was later introduced in an attempt to match the number of retained samples in (4) in order to minimize the bound for (5). The AP convergence factor is given by

$$\psi(t) = \frac{\sin\left(\pi N\delta\sqrt{(t/Nh)^2 - 1}\right)}{(\sinh \pi N\delta)\sqrt{(t/Nh)^2 - 1}}.$$

As can be seen, this $\psi(t)$ is $O(t^{-1})$.

III. The Sampling Window

The window we seek is one that minimizes the bound of (5). To this end, we seek the window $\psi(t)$ which satisfies

$$\min_{\psi \in M(\Omega)} \max_{|t| \geq T} |\psi(t)|, \tag{6}$$

Manuscript received November 19, 1980; revised July 21, 1981.
The author is with the Fairchild Space and Electronics Co., 20301 Century Blvd., Germantown, MD 20874.