The Optimum Linear Modulator for a Gaussian Source Used with a Gaussian Channel

By RANDOLPH J. PILC

(Manuscript received June 12, 1969)

The optimum linear modulator and demodulator which provide transmission of a gaussian vector source through an additive gaussian vector channel are derived in this paper. The measure of performance that is used is the transmission distortion, which is defined here as the mean square error between the source output and the decoder output. It is assumed that the source and channel are mutually independent but that correlations can exist among the components of each. The performance of the best linear system is then compared with the distortion shown by Shannon to be theoretically obtainable when no functional constraint is imposed at the modulator other than an energy constraint. Although the precise form of this optimum modulator is not known for general gaussian vector sources and channels, it is known to be nonlinear and to require arbitrarily long coding block lengths. However, it is a commonly held notion that when the source and channel dimensionalities are equal the optimum modulator is linear and requires a block length of only one. It is shown here that this belief is incorrect except in very particular situations which are described. Some relations between the optimum linear modulator-demodulator pair and Shannon's test channel are discussed, and an example is included which shows that the nonoptimality of linear devices can be quite small.

I. INTRODUCTION

We are concerned here with the transmission of a gaussian vector source over an additive gaussian vector channel. The mean square difference between the source and decoder outputs is used to measure the transmission distortion in the system and is, therefore, attempted to be minimized in the design of the encoder and decoder. In this design the encoder is constrained to present only a limited energy to the channel, thus constraining the transmission capacity of the system.[1] It is because the transmission capacity of the system is limited in this way that the given gaussian vector source cannot be transmitted with arbitrarily small error.

The distortion which necessarily must exist in the system is prescribed by Shannon's rate-distortion theory.[2] This theory states that when the transmission rate in a system is limited to R, the transmission of the source must include an average distortion of at least d_R, which in general is a function of the source statistics and the distortion measure. The theory further states that the distortion level d_R is attainable with some modulator-demodulator pair. Unfortunately, the precise form of this modulator and demodulator is not known in general, except that it is nonlinear[3,4] and that it requires the use of arbitrarily long coding block lengths.[2]

Since the nonlinearity of the optimum encoder is probably a very complex twisting of the source space locus within the channel input space, the implementation of the optimum encoder, even if it were known, would be extraordinarily complex. Of course, the long coding block length requirement does nothing to help the situation. For these reasons we study in this paper the optimum linear transmission system, restricting both the encoder and decoder to be linear operators. Such a system uses a block length of only one and is very simple to implement. (It is later shown that increasing the block length does not improve the performance.)

The degradation in performance with the use of the optimum linear system is found by comparing the resulting distortion to that of the optimum nonlinear system as found by Shannon. Contrary to popular belief, the best linear system does not provide the minimum attainable distortion, *even* when source and channel dimensionalities are equal, except in very particular situations that are described. However, in many cases the difference is small. At the end of the paper we discuss some relations between the optimum linear modulator-demodulator pair and Shannon's test channel.[2]

II. THE LINEAR TRANSMISSION SYSTEM

The system considered is shown in Fig. 1. The N_s dimensional zero-mean source vector w is linearly modulated by A to form the input to the N_c dimensional additive gaussian noise channel. We assume the noise vector n to be independent of w. The linear demodulator B extracts from the received vector y an estimate $\hat{w}$ of the source

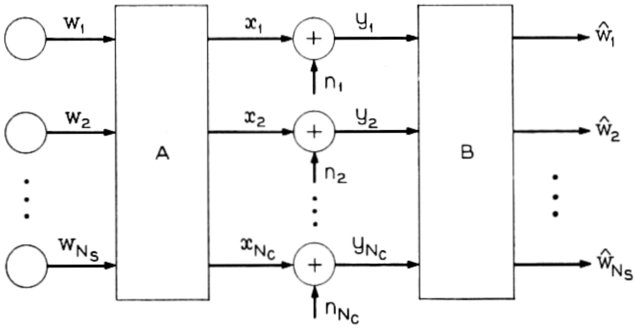

Fig. 1 — The linear system.

which is presented to the user. In summary

$$\hat{w} = By = B(x + n) = B(Aw + n). \tag{1}$$

The measure of distortion in the system is taken to be the sum of mean-square errors between the components of w and $\hat{w}$, that is

$$d = E[| w - \hat{w} |^2] = E\left[\sum_{i=1}^{N_s} (w_i - \hat{w}_i)^2 \right]. \tag{2}$$

The modulation matrices, A and B, are sought which minimize this distortion, their choice subject only to an average channel input energy constraint,

$$S_T = E\left[\sum_{i=1}^{N_c} x_i^2 \right] = \sum_{i=1}^{N_c} \text{Var } x_i , \tag{3}$$

$$\leqq S_0 , \tag{4}$$

which obviously will be met with equality in the optimum system.

It is well known that the minimum mean square error estimate of any quantity (here the source vector w) based on the observation of a second quantity (here the channel output vector y) is the conditional expected value of the first given the second.[4] Further, the average error made with such an estimate is the conditional variance of the first given the second. Therefore, we have

$$\begin{aligned} \hat{w}_i &= E(w_i \mid y); \qquad i = 1, 2, \cdots , N_s \\ d &= \sum_{i=1}^{N_s} \text{Var } (w_i \mid y). \end{aligned} \tag{5}$$

The required conditional density $p(w|y)$ can be found from

$$p(w) = k_1 \exp [-\tfrac{1}{2} w^t \Phi_w^{-1} w]$$

and

$$p(y \mid w) = k_2 \exp [-\tfrac{1}{2}(y - Aw)^t \Phi_n^{-1}(y - Aw)]$$

by application of Bayes rule. The result is

$$p(w \mid y) = k_3 \exp [-\tfrac{1}{2}(w - \hat{w})^t \Phi_{w|y}^{-1}(w - \hat{w})]$$

with

$$\Phi_{w|y}^{-1} = A^t \Phi_n^{-1} A + \Phi_w^{-1} \tag{6}$$

and

$$\hat{w}^t = y^t \Phi_n^{-1} A \Phi_{w|y} . \tag{7}$$

From these equations we have one immediate result, that is, the optimum demodulator matrix is given in terms of A by

$$B = \Phi_{w|y} A^t \Phi_n^{-1}. \tag{8}$$

If we now rewrite equations (5) and (3) as

$$d = \text{trace } \Phi_{w|y} \tag{9}$$

$$S_T = \text{trace } \Phi_x \tag{10}$$

we can restate our problem as that of finding the matrix A which minimizes the trace of $\Phi_{w|y}$ subject to a constrained maximum trace of Φ_x .

III. THE SOLUTION UNDER CERTAIN ASSUMPTIONS

We first restrict our attention to systems in which the source and channel dimensionalities are equal, $N_s = N_c = N$, and in which the correlation matrices Φ_w and Φ_n are diagonal. From equation (6) we have

$$\Phi_w \Phi_{w|y}^{-1} = \Phi_w A^t \Phi_n^{-1} A + I \tag{11}$$

and from equation (1) that $\Phi_x = A \Phi_w A^t$ and $\Phi_y = \Phi_x + \Phi_n$, which provides

$$\Phi_y \Phi_n^{-1} = A \Phi_w A^t \Phi_n^{-1} + I. \tag{12}$$

Noting that Φ_y enters these equations in a more symmetric way than dose Φ_x , we recast the energy constraint in equation (10) to be in terms of the received energy at the channel output. This energy equals

$$S_R = E\left[\sum_{i=1}^{N_c} y_i^2\right] = \sum_{i=1}^{N_c} \text{Var } y_i$$

$$= \text{trace } \Phi_y$$

$$= \text{trace } \Phi_x + \text{trace } \Phi_n$$

which, if trace $\Phi_n \equiv N_0$, is constrained to satisfy

$$S_R \leqq S_0 + N_0 . \tag{13}$$

3.1 *The Proof that the Optimum Modulator Matrix is Diagonal*

If we denote the characteristic polynomial of a matrix M in the variable λ by

$$\text{c.p. } [M, \lambda] = \det (M - \lambda I)$$

and state that M_i is square, we can use the following two matrix properties:[5]

$$(i) \qquad \text{c.p. } [M_1 M_2 , \lambda] = \text{c.p. } [M_2 M_1 , \lambda] \tag{14}$$

$$(ii) \qquad \text{c.p. } [M_1 , \lambda] = \text{c.p. } [M_1 + I, \lambda - 1] \tag{15}$$

to conclude from equations (11) and (12) that

$$\text{c.p. } [\Phi_w \Phi_{w|y}^{-1} , \lambda] = \text{c.p. } [\Phi_y \Phi_n^{-1}, \lambda]. \tag{16}$$

It is this equation which provides the important relations among the correlation matrices in the system.

We note that the set of matrix pairs $\Phi_{w|y}$, Φ_y which are consistent with equation (16) include many pairs which do not satisfy both equations (11) and (12) for any given A. The latter equations of course specify the relations among $\Phi_{w|y}$ and Φ_y which must exist in the communication problem under consideration. Nevertheless, we will work with equation (16) to perform the optimization and then show that the solutions for $\Phi_{w|y}$ and Φ_y can be realized with some modulator matrix A and, therefore, are consistent with the more restrictive equations (11) and (12).

Equation (14) and the assumed diagonal form of Φ_w and Φ_n allows us to rewrite equation (16) as

$$\text{c.p. } [\Phi_w^{\frac{1}{2}} \Phi_{w|y}^{-1} \Phi_w^{\frac{1}{2}} , \lambda] = \text{c.p. } [\Phi_n^{-\frac{1}{2}} \Phi_y \Phi_n^{-\frac{1}{2}}, \lambda].$$

As Φ_w and Φ_n are system constants not under the control of the user, any specification of Φ_y completely determines the roots of $\Phi_n^{-\frac{1}{2}} \Phi_y \Phi_n^{-\frac{1}{2}}$, which we denote by $\{\alpha_i\}$, $i = 1, 2, \cdots , N$. The roots of $\Phi_w^{-\frac{1}{2}} \Phi_{w|y} \Phi_w^{-\frac{1}{2}}$

are also determined and are equal to $\{\alpha_i^{-1}\}$. Our claim now is that among all matrices Φ with roots $\{\alpha_i^{-1}\}$, the one which produces the minimum trace of $\Phi_{w|y} = \Phi_w^{\frac{1}{2}}\Phi\Phi_w^{\frac{1}{2}}$ is diagonal.

If φ_{ij} are used to denote the elements of Φ, the trace of $\Phi_{w|y}$ equals

$$\text{trace } \Phi_{w|y} = \sum_{i=1}^{N} \sigma_i^2 \varphi_{ii} .$$

At this point we impose, without loss of generality, that the variances σ_i^2 be ordered such that $\sigma_1^2 \geqq \sigma_2^2 \geqq \cdots \geqq \sigma_N^2$. Since the minimum trace of $\Phi_{w|y}$ is sought, clearly the diagonal elements φ_{ii} should correspondingly satisfy $\varphi_{11} \leqq \varphi_{22} \leqq \cdots \leqq \varphi_{NN}$. This presents no restriction on Φ as a simultaneous interchange of rows and columns produces no change in the characteristic equation of Φ.

Now consider any nondiagonal candidate for the desired Φ. In particular, let $\varphi_{mk} = \varphi_{km}$, $m > k$, be nonzero. Because the submatrix

$$\Phi(km) = \begin{bmatrix} \varphi_{kk} & \varphi_{km} \\ \varphi_{mk} & \varphi_{mm} \end{bmatrix}$$

is itself a correlation matrix, it can be diagonalized by some orthogonal matrix T such that

$$\Phi'(km) = T\Phi(km)T^t = \begin{bmatrix} \varphi'_{kk} & 0 \\ 0 & \varphi'_{mm} \end{bmatrix}.$$

From (14) it is known that the characteristic polynomials of $\Phi(km)$ and $\Phi'(km)$ are equal. The trace and determinant of each are therefore equal. It follows that $\varphi'_{kk} = \varphi_{kk} - c$ and $\varphi'_{mm} = \varphi_{mm} + c$; $c > 0$, or that the larger diagonal element is increased and that the smaller one is decreased.

The diagonalization of the submatrix $\Phi(km)$ within Φ can be effected by an orthogonal matrix Q which contains T in the appropriate submatrix position and identity matrix elements in the other positions:

$$q_{ij} = t_{ij} ; \quad (i, j) = (k, k), (k, m), (m, k), (m, m)$$

$$q_{ij} = \delta_{ij} ; \quad \text{other } (i, j).$$

We then have $\Phi' = Q\Phi Q^t$ with only the elements in Φ' in rows and columns k and m changed from those in Φ. If Φ' is used to generate a new correlation matrix $\Phi'_{w|y} = \Phi_w^{\frac{1}{2}}\Phi'\Phi_w^{\frac{1}{2}}$, we have

$$\text{tr } \Phi'_{w|y} = \sum_{i=1}^{N} \sigma_i^2 \varphi'_{ii} = \sum_{i=1}^{N} \sigma_i^2 \varphi_{ii} - c(\sigma_k^2 - \sigma_m^2)$$

$$= \text{tr } \Phi_{w|} \quad - c(\sigma_k^2 - \sigma_m^2)$$

$$\leqq \text{tr } \Phi_{w|y} , \tag{17}$$

which establishes the claim of this section. That is, any nondiagonal correlation matrix Φ with roots $\{\alpha_i^{-1}\}$ conjectured as providing a minimum trace correlation matrix $\Phi_w^{\frac{1}{2}}\Phi\Phi_w^{\frac{1}{2}} = \Phi_{w|y}$ can be improved upon by Φ'. The desired matrix for Φ is therefore diagonal and equal to

$$\Phi = [\alpha_i^{-1}\delta_{ij}] \tag{18}$$

with the corresponding form of $\Phi_{w|y}$ equal to

$$\Phi_{w|y} = [\sigma_i^2\alpha_i^{-1}\delta_{ij}]. \tag{19}$$

It follows that among all matrices $\Phi_{w|y}$ consistent with equation (16) with any given Φ_y , the one with minimum trace is diagonal.

An identical argument yields the symmetric conclusion. That is, for any specified $\Phi_{w|y}$ the matrix Φ_y with minimum trace among those consistent with equation (16) is also diagonal and equal to

$$\Phi_y = [\sigma_{ni}^2\alpha_i\delta_{ij}]. \tag{20}$$

The argument assumes only that the noise variances are ordered $\sigma_{n1}^2 \leqq \sigma_{n2}^2 \leqq \cdots \leqq \sigma_{nN}^2$.

We can now state that the minimization of the trace of $\Phi_{w|y}$ over all pairs $\Phi_{w|y}$, Φ_y which satisfy equation (16) and the constraint equation (13) is obtained with a pair of diagonal matrices parametrically related as in equations (19) and (20). Any pair not so related can be altered, one matrix at a time, to decrease either the error (trace $\Phi_{w|y}$) or the received energy (trace Φ_y). Although we have worked with pairs $\Phi_{w|y}$, Φ_y consistent with equation (16) rather than the smaller set satisfying equations (11) and (12), the solution forms for $\Phi_{w|y}$ and Φ_y are still valid as they do satisfy these equations.

The modulator matrix which produces the correlation matrices $\Phi_{w|y}$ and Φ_y in the optimum form can be found from either equation (11) or (12) to be

$$A = \left[\frac{\sigma_{ni}}{\sigma_i}(\alpha_i - 1)^{\frac{1}{2}}\delta_{ij}\right]. \tag{21}$$

Equations (12), (14), and (15) and the fact that $\Phi_n^{-\frac{1}{2}}A\Phi_wA^t\Phi_n^{-\frac{1}{2}}$ has nonnegative roots (it is a correlation matrix) can be used to show that $\alpha_i \geqq 1$, $i = 1, 2, \cdots , N$ which guarantees that the elements of A are real. It remains to solve for the set of roots $\{\alpha_i\}$ which provides the desired optimization.

3.2 *The Optimum Diagonal Modulator Matrix*

In terms of the set $\{\alpha_i\}$, the distortion which is to be minimized is given by

$$d = \text{trace } \Phi_{w|y} = \sum_{i=1}^{N} \sigma_i^2 \alpha_i^{-1}$$

and the received energy constraint by

$$S_R = \text{trace } \Phi_y = \sum_{i=1}^{N} \sigma_{ni}^2 \alpha_i \leqq S_0 + N_0 .$$

A further constraint is that $\alpha_i \geqq 1$, $i = 1, 2, \cdots, N$. As the set of permissible α_i's is a convex set and the functions $d(\alpha_i)$ and $S_R(\alpha_i)$ are convex functions, the Kuhn–Tucker theorem is applicable.[6] This states that at the point of minimization:

$$\frac{\partial}{\partial \alpha_i}\left[d + \frac{1}{\lambda^2} S_R\right] \begin{array}{ll} = 0 & \text{if} \quad \alpha_i > 1 \\ < 0 & \text{if} \quad \alpha_i = 1. \end{array}$$

Therefore we have

$$-\sigma_i^2 \alpha_i^{-2} + \frac{1}{\lambda^2} \sigma_{ni}^2 \begin{array}{ll} = 0 & \text{if} \quad \alpha_i > 1 \\ < 0 & \text{if} \quad \alpha_i = 1 \end{array}$$

or

$$\alpha_i = \max \left[\left(\frac{\sigma_i}{\lambda \sigma_{ni}}\right), 1\right]. \tag{22}$$

It has already been observed that $\alpha_1 \geqq \alpha_2 \geqq \cdots \geqq \alpha_N$ and that $\alpha_i = 1$ corresponds to $a_{ii} = 0$ or no transmission of the ith source component. If we let N' denote the last α_i strictly greater than one we have the following solution for the optimum modulator matrix

$$A = \begin{bmatrix} \dfrac{\sigma_{ni}}{\sigma_i}\left(\dfrac{\sigma_i}{\lambda \sigma_{ni}} - 1\right)^{\frac{1}{2}} \delta_{ij} & 0 \\ 0 & 0 \end{bmatrix}; \qquad 1 \leqq i, j \leqq N'. \tag{23}$$

The solution for the distortion in the optimum linear system follows directly from equation (19):

$$d = \sum_{i=1}^{N'} \lambda \sigma_i \sigma_{ni} + \sum_{i=N'+1}^{N} \sigma_i^2 , \tag{24}$$

as does the solution for the total received energy from equation (20):

$$S_R = \sum_{i=1}^{N'} \frac{1}{\lambda} \sigma_i \sigma_{ni} + \sum_{i=N'+1}^{N} \sigma_{ni}^2 . \tag{25}$$

In these equations, the parameter λ is chosen to satisfy the constraint in equation (13) with equality. It should be remembered in the solution for λ that N' is a function of λ, being equal to the largest value of i for which $\sigma_i/\sigma_{ni} \geqq \lambda$. For completeness, we give the optimum demodulator matrix:

$$B = \begin{bmatrix} \lambda\left(\frac{\sigma_i}{\lambda\sigma_{ni}} - 1\right)^{\frac{1}{2}} \delta_{ij} & 0 \\ 0 & 0 \end{bmatrix}; \qquad 1 \leqq i, j \leqq N'. \tag{26}$$

IV. ELIMINATION OF THE ASSUMPTIONS

4.1 *A Source and Channel with Nonindependent Components*

We now consider systems in which Φ_w and Φ_n are not diagonal. Let P and R be the orthogonal matrices which respectively diagonalize these two correlation matrices, that is, $\Phi_{w'} = P\Phi_w P^t$ and $\Phi_{n'} = R\Phi_n R^t$ with $\Phi_{w'}$ and $\Phi_{n'}$ diagonal. Using the previous results, we can find the optimum modulator matrix A' in the primed system containing the correlation matrices $\Phi_{w'}$ and $\Phi_{n'}$. Now consider the use of the modulator matrix $A = R^t A' P$ in the system with Φ_w and Φ_n. From equation (6) and $\Phi_y = A\Phi_w A^t + \Phi_n$, it can be easily shown that using A' in the primed system and A in the unprimed system each produces the same distortion and uses the same energy. Consequently, A must be the optimum matrix in the unprimed system. If it is not, and A_0 is better, $A_0' = RA_0P^t$ would be a better choice than A' for modulator in the primed system contrary to A' being optimum.

4.2 *Nonequal Source and Channel Dimensionality*

When $N_s \neq N_c$, we can appropriately modify either the source or channel to restore the equality. For example, when $N_s < N_c$, $N_c - N_s$ source components of arbitrarily small variance, say ϵ, are added to the original source vector. The optimum modulator is then found as a function of ϵ by the previous method, and finally the limit taken as ϵ goes to zero. Similarly, when $N_c < N_s$, $N_s - N_c$, channel components of arbitrarily large noise variance, say $1/\epsilon$, are added to the original channel, the optimum modulator found, and the limit taken as ϵ goes to zero. We have seen that whenever either the source has a component with small variance or the channel has a component with large noise variance, the number of source components actually transmitted, N', is smaller than N. Since the optimum modulator matrix is diagonal, N' is also the number of channel components actually used. Therefore, the limiting modulator form in both of the above situations is attained for a nonzero value of ϵ, say ϵ_1. This modulator form is then optimum for all $\epsilon < \epsilon_1 \neq 0$.

V. COMPARISON OF OPTIMUM LINEAR AND NONLINEAR MODULATORS

In 1959 C. E. Shannon introduced a relation between d_R, the minimum attainable transmission distortion of a source, and R, the total information rate used in transmission.[2] This relation involves only the source statistics and the distortion measure in use. From it one is able to conclude that any channel with capacity R can be used to transmit the source with a transmission distortion arbitrarily close to d_R. One need only use a "sufficiently complex" encoder and decoder.

Another part of rate-distortion theory is the idea of a "test channel." Associated with each point on the rate-distortion curve, (d_R, R), is such a test channel which has the significance that among all channels that transmit the source at a rate equal to R, it provides the minimum transmission distortion d_R. Therefore, if there exist pre- and post-operators which can transform a given capacity R channel into the test channel for the source at (d_R, R), these operators must be optimum. An obvious necessary condition for this transformation, which is not always met, is that the capacity of the test channel at (d_R, R) be equal to R.

For a gaussian source with variance σ^2 and squared difference distortion, Shannon has found[2] both the rate distortion expression, $d_R = \sigma^2 e^{-2R}$ and the test channel:

$$\begin{array}{c} w \leftarrow \oplus \leftarrow \hat{w}. \\ \uparrow \\ n \end{array} \tag{27}$$

In this reverse channel, $\hat{w}$ and n are independent gauss variables with respective variances $\sigma^2 - d_R$ and d_R. It can be shown that this channel is identical to the forward channel:

$$\begin{array}{ccccc} w \rightarrow & \otimes & \rightarrow & \oplus & \rightarrow \hat{w} \\ & \uparrow & & \uparrow & \\ & A_1 & & n & \end{array} \tag{28}$$

with $A_1 = (\sigma^2 - d_R)/\sigma^2$, $\sigma_n^2 = A_1 d_R$, and the independence between w and n. A similar form is given by Gallager in Ref. 7. Still another form of the test channel is:

$$w \longrightarrow \otimes \longrightarrow \oplus \longrightarrow \otimes \longrightarrow \hat{w} \qquad (29)$$
$$\uparrow \quad\; \uparrow \quad\; \uparrow$$
$$A \quad\; n \quad\; B$$

with $A^2 = (\sigma^2 - d_R)\sigma_n^2/\sigma^2 d_R$, $B^2 = (\sigma^2 - d_R)d_R/\sigma^2\sigma_n^2$, and n *any* given additive gaussian noise.

Now consider a single dimensional gaussian channel of capacity R. Since the received energy S_R is accordingly restricted to $\sigma_n^2 \exp(2R)$, we have from equations (23) through (26) that the optimum linear operators are

$$a_{11}^2 = \frac{\sigma_n^2}{\sigma^2}\left(\frac{\sigma}{\lambda\sigma_n} - 1\right) = \frac{\sigma_n^2}{\sigma^2}\left(\frac{\sigma^2 - d}{d}\right)$$

$$b_{11}^2 = \lambda^2\left(\frac{\sigma}{\lambda\sigma_n} - 1\right) = \frac{d(\sigma^2 - d)}{\sigma^2\sigma_n^2}$$

$$\lambda = \frac{d}{\sigma\sigma_n} = \frac{\sigma\sigma_n}{S_R}.$$

Note that the distortion d equals $\sigma^2 \exp(2R)$, and that a_{11} and b_{11} agree precisely with the test channel parameters in (29). Therefore, we can conclude that in this case the operators in equations (23) and (26) are optimum, even outside the linear class.

The rate-distortion curve and the test channel for gaussian vector sources can also be found from Shannon's results. The results for the N-dimensional source with variances $\sigma_1^2, \sigma_2^2, \cdots, \sigma_N^2$ are (we continue to assume that $\sigma_1^2 \geqq \sigma_2^2 \geqq \cdots \geqq \sigma_N^2$):

$$d_R = N\left\{e^{-2R}\prod_{i=1}^{N}\sigma_i^2\right\}^{1/N}; \qquad 0 \leqq d_R \leqq N\sigma_N^2$$

$$= \sigma_N^2 + (N-1)\left\{e^{-2R}\prod_{i=1}^{N-1}\sigma_i^2\right\}^{1/N-1};$$

$$N\sigma_N^2 \leqq d_R \leqq \sigma_N^2 + (N-1)\sigma_{N-1}^2$$

$$= \sigma_N^2 + \sigma_{N-1}^2 + (N-2)\left\{e^{-2R}\prod_{i=1}^{N-2}\sigma_i^2\right\}^{1/N-2};$$

$$\sigma_N^2 + (N-1)\sigma_{N-1}^2 \leqq d_R \leqq \sigma_N^2 + \sigma_{N-1}^2 + (N-2)\sigma_{N-2}^2$$

$$\vdots$$

$$= \sigma_N^2 + \cdots + \sigma_2^2 + \sigma_1^2 e^{-2R};$$

$$\sigma_N^2 + \cdots + 2\sigma_2^2 \leqq d_R \leqq \sigma_N^2 + \cdots + \sigma_1^2$$

$$= \sigma_N^2 + \cdots + \sigma_1^2 ; \qquad R = 0. \tag{30}$$

This expression can also be applied to a gaussian vector source with correlated components if the variances σ_i^2 are interpreted as those in the diagonalized correlation matrix $\Phi_{w'} = P\Phi_w P^t$. The test channel for $N > 1$ is the product of elementary test channels given in (29) with

$$A = A_1, A_2, \cdots, A_N,$$

$$A_i^2 = \frac{(\sigma_i^2 - d_i)}{\sigma_i^2 d_i} \sigma_{ni}^2,$$

$$d_i = \min(\sigma_i^2, d_R),$$

$$\sigma_n^2 = \sigma_{n1}^2, \sigma_{n2}^2, \cdots, \sigma_{nN}^2 = \text{any noise vector.}$$

Let us now presume that the vector channel provided for use has the additive noise variances given by the vector σ_n^2 and is constrained to have an output energy level equal to S_R. This equivalently specifies the channel capacity as

$$R = \max_{S_{Ri}} \sum_{i=1}^{N} \frac{1}{2} \log \frac{S_{Ri}}{\sigma_{ni}^2} \tag{31}$$

with

$$S_{Ri} = \max(S, \sigma_{ni}^2)$$

and S adjusted to have $\sum S_{Ri} = S_R$. The comparison between the minimum attainable transmission distortion using linear transmitter and receiver operators (equation 24) and using unrestricted transmitter and receiver operators (equation 30) now reveals that contrary to the single dimension case, when $N > 1$ the linear operators are *not*, in general, optimum. The only exception is when both the vectors σ^2 and σ_n^2 are uniform. Some intuition as to why the single and multi-dimensional cases are different might be provided by the following.

The test channel at (d_R, R), for example, the one including the noise vector σ_n^2 in its form, is a result of a minimization of mutual information under a distortion constraint. It does not, therefore, necessarily divide the total energy presented to the gaussian vector channel in a way which uses this channel to capacity. Since this channel, by definition, transmits information at a rate equal to R, its total capacity is (except for the special case noted previously) strictly greater than R. Consequently, when the same additive noise channel, σ_n^2, is to be used for transmission but is stated to have a capacity of only R, it cannot be transformed into the test channel by any pre- and postoperators.

The impossiblity of such a transformation can also be observed by noting that the allowed total input energy on the given capacity R channel is restricted to a lower level than present on the test channel The uniqueness of the test channel, which is formed with linear operators, and the continuity of both the mutual information and distortion with the modulator matrix then precludes the possibility of attaining the test channel's performance with the given capacity R channel and linear operators.

One could argue that the comparison to this point is not fair in that Shannon allows modulators and demodulators that operate on blocks of letters, whereas the results in equations (23), (24), and (25) were derived using a coding block length of one. However, the previous results show that the optimum linear modulator does not mix independent source components before presentation to the channel, assuming the channel has already been rotated in N-space so as to have independent noise components. Neither does it cross-couple sets of source components having no cross dependence when presentation is to a channel with sets of noise components of equal respective dimensionalities also having no cross dependence. Therefore, if successive source and channel (vector) events are independent, and their dimensionalities filled out to be equal by adding either zero variance source components or infinite variance noise components, there is no memory introduced by the optimum linear modulator among elements of the encoded block. The consequence is that the distortion and the energy are only scaled by the block length in use.

VI. AN EXAMPLE

We cite here just one example which shows that at least in many cases the performance of the optimum linear modulator- demodulator pair compares favorably with that theoretically obtainable with more complex operators. We take $\sigma_1 = \sigma_2 = 1$, $\sigma_{n1} = a$, $\sigma_{n2} = ae^{2\varphi}$ and use a and φ as parameters that generate a set of different channels. To better compare the two performances, we fix the channel capacity at C which in turn fixes Shannon's minimum attainable distortion at $d_C = 2e^{-C}$. The total allowed received energy is thus specified according to equation (31).

Upon solution for λ and d in equations (24) and (25) we have the following expression for the ratio between the distortion obtainable with linear operators and that theoretically attainable:

$$\frac{d(\varphi)}{d_C} = \begin{cases} \cosh^2 \varphi & ; \quad 0 \leqq \varphi \leqq \tfrac{1}{2}C \\ \dfrac{\cosh^2 \varphi}{\cosh(2\varphi - C)} & ; \quad \tfrac{1}{2}C \leqq \varphi \leqq C \\ \cosh C & ; \quad C \leqq \varphi. \end{cases}$$

We illustrate this function for several different values of capacity in Fig. 2. At $\varphi = 0$ (where both the vectors σ^2 and σ_n^2 are uniform) we see that $d(0) = d_C$ indicating the optimality of the linear modulator and demodulator for this case. Using a term introduced in Ref. 8, we can therefore say that when $\varphi = 0$ the source and channel are "matched." As φ increases, the source-channel mismatch increases and the nonoptimality of linear operators also increases. As the figure illustrates, the nonoptimality ratio, $d(\varphi)/d_C$, can be quite large when both the channel capacity is high and the additive noise vector is highly skewed in variance. However, over a significant region of interest,

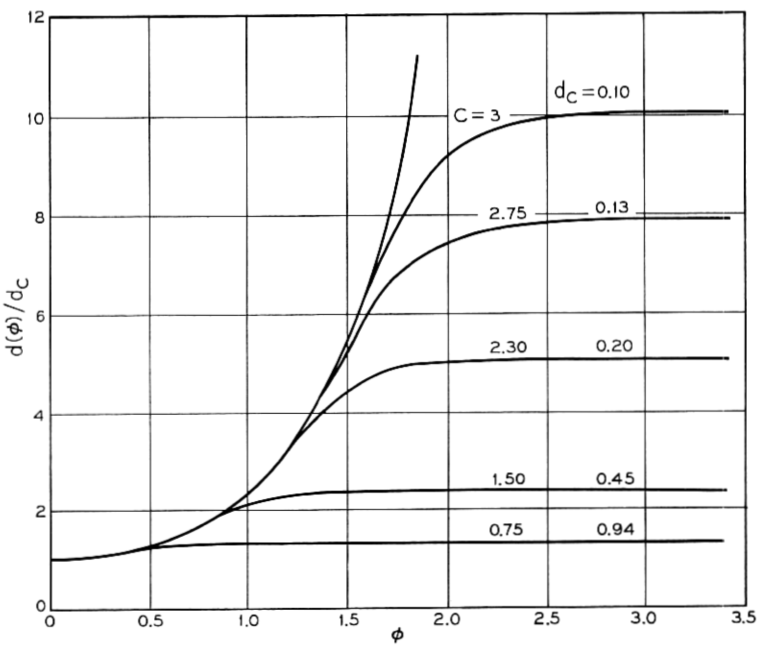

Fig. 2 — The linear system nonoptimality for $N = 2, \sigma_1 = \sigma_2 = 1, \sigma_{n1} = 1, \sigma_{n2} = \exp 2\phi$.

$\varphi \leqq 1$ (reflecting a noise component variance ratio of about 50), the nonoptimality ratio is small.

VII. SUMMARY

In this paper we have derived the optimum linear modulator and demodulator for the transmission of a gaussian vector source through an additive gaussian vector channel. It was found that when both the source and channel components are independent, both the modulator and demodulator matrices are diagonal. This specifies the separate amplification, transmission, and decoding of each source component. When both the source and channel components are correlated, the optimum modulator matrix was found to be the cascade of three matrices: (i) the orthogonal matrix which diagonalizes the source correlation matrix, (*ii*) the optimum modulator matrix which transmits this newly formed independent component source over the independent component additive noise channel which is formed by (*iii*) the orthogonal transformation matrix that diagonalizes the noise correlation matrix. We have found that in general the best linear system does not provide a distortion as small as that stated by Shannon to be attainable with a channel of the same capacity. The only exception is when both the source and channel noise variance vectors are uniform. The non-optimality of linear modulators and demodulators can be quite large in some cases but, in many other situations, can be small enough to justify the use of these very simple operators.

REFERENCES

1. Fano, R. M., *Transmission of Information*, New York: John Wiley, 1961.
2. Shannon, C. E., "Coding Theorems for a Discrete Source with a Fidelity Criterion," IRE Nat. Conv. Record, Part 4 (1959), pp. 142–163.
3. Shannon, C. E., "Communication in the Presence of Noise," Proc. IRE, *37*, No. 1 (January 1949), pp. 10–21.
4. Wozencraft, J. M., and Jacobs, I. M., *Principles of Communication*, New York: John Wiley, (1965).
5. Hoffman, K., and Kunze, R., *Linear Algebra*, Englewood Cliffs, New Jersey: Prentice-Hall, 1961.
6. Kuhn, H. W., and Tucker, A. W., "Nonlinear Programming," Proc. 2nd Berkeley Symp. Math. Stat. and Prob., 1951, pp. 481–492.
7. Gallager, R. G., *Information Theory and Reliable Communication*, New York: John Wiley, 1968.
8. Pilc, R. J., "Coding Theorems for Discrete Source-Channel Pairs," Ph.D. thesis, Department of Electrical Engineering, M.I.T., Cambridge, Massachusetts, (February 1967).