On Metrics Matched to the Discrete† Memoryless Channel

by GÉRALD SÉGUIN

Department of Electrical Engineering, Royal Military College, Kingston, Ontario, Canada, K7L 2W3

ABSTRACT: *A sequence of metrics $\{D_N\}$ is said to be additive and matched to a discrete memoryless channel (DMC) if D_N is the sum on its coordinates of N single letter metrics and if the maximum likelihood decoder for sequences of length N is a minimum D_N-distance decoder. Necessary and sufficient conditions on the transition probabilities of a DMC for the existence of a sequence of additive metrics matched to it are given. In the case of the binary channel these are shown to be equivalent to the channel being symmetric. Explicit transition probabilities are given for a large class of ternary DMCs with an associated sequence of additive matched metrics. The problem solved here may be considered a generalization of the problem of finding the DMCs matched to the Lee metric solved by Chiang and Wolf in* 1971 (**2**).

Nomenclature

A	channel alphabet
A	size of A
A^N	set of N-tuples over A
C	block code over A
P_N	transition probability measure for sequences of length N
D	a metric on A^N
D_N	a metric on A^N
p	transition probability measure
P	transition probability matrix
$\in$	belongs to
$>$	greater than
$\geq$	greater than or equal to
$<$	less than
$\leq$	less than or equal to
log	logarithm to the base e.

I. Introduction

In a discrete communication system the time axis is subdivided into successive time slots each of duration T s. During any one of these time slots a signal $S_i(t)$, taken from a finite set $\{S_0(t), S_1(t), \ldots, S_{M-1}(t)\}$, is transmitted over a

† Part of this paper was presented at the 1977 International Symposium on Information Theory, Cornell University, Ithaca, New York.

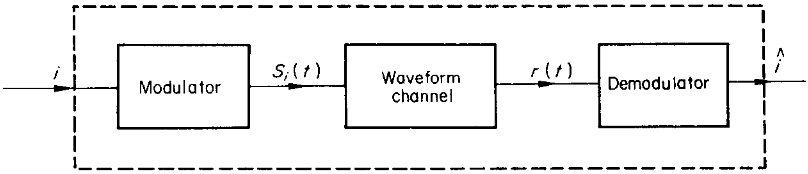


FIG. 1. Discrete channel.

waveform channel. At the receiver end the demodulator receives a signal $r(t)$ [a corrupted version of $S_i(t)$] from which it must decide the value of "i". This being the case we usually coalesce the modulator, the waveform channel and the demodulator into one block called a discrete channel (see Fig. 1)

The discrete channel is characterized by an input alphabet A which we may take to be the canonical alphabet $\{0, 1, \ldots, A-1\}$, an output alphabet, which in this paper we take to be also A but in general is an augmented version of A, and a set of transition probabilities $p(j/i)$, j, $i \in A$. If the noise in the waveform channel is independent from one time slot to another [for example in the additive white Gaussian noise channel (**3**)] then the resulting discrete channel is memoryless. In this case if $x=[x(0), x(1), \ldots, x(N-1)]$ and $y=[y(0), y(1), \ldots, y(N-1)]$ are two N-tuples over A then $p_N(y/x) = \prod_0^{N-1} p[y(n)/x(n)]$. The study of discrete memoryless channels (DMC) occupies a large part of the literature in Information and Communication Theory [see for example the recent text by Viterbi and Omura (**4**)].

The object of a communication system is to transmit information reliably from a source to a destination. This can be achieved, as shown by Shannon (**5**), through coding. We transmit then, not individual letters from A, but N-tuples x over A, i.e. elements from A^N. In this case not all x terms from A^N are allowed but only those x terms belonging to a subset $C=\{x_0, x_1, \ldots, x_{M-1}\}$ of A^N, called a code , are allowed to be transmitted. In this case the output of the DMC is an N-tuple y over A and from this y the decoder must decide which x_i was transmitted. If x was transmitted and $\hat{x}$ is decoded then the probability of error $P(E)$ is simply $Pr(\hat{x} \neq x)$ and the objective is to find a code C and a decoder which will make $P(E)$ small. The situation just described is depicted in Fig. 2.

For most channels one may restrict the decoder to be a maximum likelihood decoder (MLD), i.e. a decoder which decodes y into x_k where k is that value of i which maximizes

$$p_N(y/x_i). \tag{1}$$

If the codewords x_i are equally likely to be transmitted then the MLD is known to be optimum (**4**).

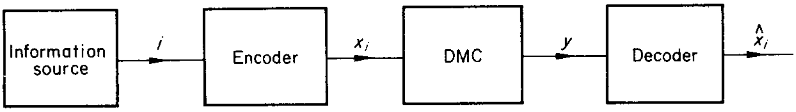


FIG. 2. System with coding.

A minimum D-distance decoder (MDD) is one which decodes y into x_k where k is that value of i which minimizes

$$D(y, x_i) \tag{2}$$

and where D is a metric on A^N i.e. D is real valued function such that:

(1) $D(x, y) \geq 0$ with equality if and only if $x = y$
(2) $D(x, y) = D(y, x)$, $\forall x, y \in A^N$
(3) $D(x, z) \leq D(x, y) + D(y, z)$, $\forall x, y, z \in A^N$.

If there exists a metric D for which the minimum D-distance decoder makes the same decision as the MLD, i.e. if

$$D(x, y) < D(x, z) \quad \text{if and only if} \quad p_N(x/y) > p_N(x/z), \tag{3}$$

then D is said to be matched (more precisely matched at N) to the DMC in question. Actually our definition differs slightly from that given by Massey (**1**) and is what Chiang and Wolf (**2**) call strictly matched. For example, the Hamming metric is matched to the binary symmetric channel provided its cross-over probability, ε, is less than $\frac{1}{2}$.

Some of the advantages which accrue from having a metric D matched to a DMC are the following. Since we have a metric D (or a distance function) then we can associate with a given code C its minimum distance $D_m = \min D(x, y)$ minimized over all $x, y \in C$, $x \neq y$. We then have at our disposition the bounds of Hamming and Gilbert which relate D_m and the rate of the code. Secondly, we can use bounded discrepancy decoding (**6**), hence reducing the complexity of the decoder (over that of MDD) at the cost of an increase in $P(E)$. Lastly, it may be possible to develop algorithms for constructing codes with a certain guaranteed minimum distance. For these reasons and possibly others it is of interest to know under what conditions a DMC has a metric matched to it.

Chiang and Wolf (**2**) have determined the class of symmetric DMC for which the Lee metric is matched. The Lee distance between the integers i and j, $0 \leq i \leq A-1$, $0 \leq j \leq A-1$, is simply $\delta_L(i, j) = \min(|i-j|, A-|i-j|)$ and the Lee distance between two N-tuples x and y is $D_L(x, y) = \sum_0^{N-1} \delta_L[x(n), y(n)]$. The Lee metric was introduced by Lee (**7**) in 1958.

Before proceeding further we make the following definitions. First a sequence of metrics $\{D_N\}$ is said to be matched to a DMC if D_N is matched to the said channel at N for $N = 1, 2, \ldots$. A metric D on A^N is said to be additive if

$$D(x, y) = \sum_{n=1}^{N} \delta_n[x(n), y(n)]$$

where δ_n is a metric on A.

In this paper we establish a set of necessary and sufficient conditions on the transition probabilities of a DMC for the existence of a sequence of additive metrics matched to it. This is a generalization of the work of Chiang and Wolf since first we do not constrain the channel to be symmetric and secondly we do not limit the metric to be the Lee metric.

II. Matched Metrics

First we will show that we can limit ourselves to permutation invariant metrics which we define as follows.

Definition 1. A metric D on A^N is said to be permutation invariant if

$$D(\sigma x, \sigma y) = D(x, y) \tag{4}$$

for every permutation σ on the N objects $\{1, 2, \ldots, N\}$ and for every $x, y \in A^N$ where by σx we mean $x[\sigma(1)], x[\sigma(2)], \ldots, x[\sigma(N)]$. For example, the Hamming metric is permutation invariant. We have the following.

Theorem I. There exists a metric matched to a DMC at N if and only if there exists a permutation invariant metric matched to the same DMC at N.

Proof: Let D and D' be two metrics matched to a DMC at N and let $D'' = D + D'$. Then $D''(x, z) < D''(y, z)$ implies that $D(x, z) < D(y, z)$ or $D'(x, z) < D'(y, z)$ hence that $p_N(z/x) > p_N(z/y)$. Conversely, if $p_N(z/x) > p_N(z/y)$, then $D(x, z) < D(y, z)$ and $D'(x, z) < D'(y, z)$, hence $D''(x, z) < D''(y, z)$. Therefore D'' is matched to the DMC in question at N; i.e. the sum of two matched metrics is a matched metric. Let D be a metric matched to a DMC at N and define, for a permutation σ on the N objects $\{1, 2, \ldots, N\}$ D_σ by

$$D_\sigma(x, y) = D(\sigma x, \sigma y). \tag{5}$$

It is an easy matter, and a consequence of the fact that the channel is memoryless, to show that D_σ is a metric matched to the DMC in question at N. Suppose then that there exists a metric D matched to a DMC at N. Then D', defined by

$$D'(x, y) = \sum_\sigma D_\sigma(x, y) \tag{6}$$

where in (6) σ runs over all permutations, is, by our previous discussion, a metric matched to the channel in question at N. But

$$\begin{aligned} D'(\tau x, \tau y) &= \sum_\sigma D_\sigma(\tau x, \tau y) = \sum_\sigma D(\sigma\tau x, \sigma\tau y) \\ &= \sum_\sigma D(\sigma x, \sigma y) = D'(x, y) \end{aligned}$$

since as σ runs over all permutations on N objects so does $\sigma\tau$ for any permutation τ on N objects. Hence D' is a permutation invariant metric matched to the said channel. The converse is obvious. *Q.E.D.*

We will first limit ourselves to sequences of metrics $\{D_N\}$ of the form

$$D_N(x, y) = \sum_{n=1}^{N} \delta[x(n), y(n)] \tag{7}$$

where δ is a metric on A and then deduce the general case for sequences of additive metrics from this one.

Let α be an A by A real matrix whose (i, j) entry we denote by $\alpha(i, j)$, $i=0, 1, \ldots, A-1$; $j=0, 1, \ldots, A-1$. Define on A^N the function d by

$$d(x, y)=\sum_{n=1}^{N} \alpha[x(n), y(n)]. \tag{8}$$

Then it is quite obvious that d will be a metric if and only if:

(1) $\alpha(i, j) \geq 0$ with equality if and only if $i=j$,
(2) $\alpha(i, j)=\alpha(j, i)$ for every $i, j \in A$,
(3) $\alpha(i, j)+\alpha(j, k) \geq \alpha(i, k)$ for every $i, j, k \in A$.

It is easy to see that (8) may be written as

$$d(x, y)=\sum_{i=0}^{A-1} \sum_{j=0}^{A-1} n_{i,j}(x, y)\alpha(i, j) \tag{9}$$

where $n_{i,j}(x, y)$ is the number of values of n for which $x(n)=i$ and $y(n)=j$. For example, if $A=2$ and if

$$\alpha=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{10}$$

then d as given by (8) is the usual Hamming metric. Let $p(j/i)$ be the probability that j is received given that i was transmitted and assume moreover that $p(j/i)>0$ for all i and j in A. Then

$$p_N(y/x_m)=\prod_{n=1}^{N} p[y(n)/x_m(n)], \tag{11}$$

which may be expressed as

$$p_N(y/x_m)=\prod_{i=0}^{A-1} \prod_{j=0}^{A-1} [p(j/i)]^{n_{i,j}(x_m, y)} \tag{12}$$

where $n_{i,j}(x_m, y)$ is the number of values of n for which $x_m(n)=i$ and $y(n)=j$. Taking logarithms to the base e of both sides of (12) obtains

$$\log p_N(y/x_m)=\sum_i \sum_j n_{i,j}(x_m, y) \log p(j/i). \tag{13}$$

Define $n_j(y)$ by

$$n_j(y)=\sum_i n_{i,j}(x_m, y), \tag{14}$$

$j \in A$. The quantity $n_j(y)$ is simply the number of times that the letter j appears in y. We have that

$$n_{j,j}(x_m, y)=n_j(y)-\sum_{i \neq j} n_{i,j}(x_m, y), \tag{15}$$

which when substituted into (13) and after simple manipulations obtains

$$\log p_N(y/x_m)=\sum_j n_j(y) \log p(j/j)-\sum_i \sum_j n_{i,j}(x_m, y) \log \frac{p(j/j)}{p(j/i)}. \tag{16}$$

Hence a MLD decodes y into x_k where k is that value of m which minimizes

$$\sum_i \sum_j n_{i,j}(x_m, y) \log \frac{p(j/j)}{p(j/i)}. \tag{17}$$

If we define the A by A matrix α by

$$\alpha(i,j) = \log \frac{p(j/j)}{p(j/i)}, \tag{18}$$

then (17) reduces to

$$d_N(x_m, y) = \sum_{n=1}^{N} \alpha[x_m(n), y(n)]. \tag{19}$$

By our previous observation d_N will be a metric on A^N if and only if α satisfies the three conditions listed immediately following (8). We have therefore proven the sufficiency part of the following theorem.

Theorem II. Consider a DMC with input and output alphabets equal to A and with transition probabilities $p(j/i) > 0$ for every i and j in A. There exists a sequence of metrics D_N of the form

$$D_N(x, y) = \sum_{n=1}^{N} \delta[x(n), y(n)],$$

where δ is a metric on A, matched to the said channel if and only if:

(1) $p(j/j) \geq p(j/i)$ with equality if and only if $i = j$,

(2) $\dfrac{p(j/j)}{p(j/i)} = \dfrac{p(i/i)}{p(i/j)}$ for every i, j in A,

(3) $\dfrac{p(j/j)}{p(j/i)} \dfrac{p(k/k)}{p(k/j)} \geq \dfrac{p(k/k)}{p(k/i)}$ for every i, j, k in A.

Moreover if the above conditions are satisfied, then

$$D_N(x, y) = \sum_{n=1}^{N} \log \frac{p[y(n)/y(n)]}{p[y(n)/x(n)]}, \tag{20}$$

$N = 1, 2, \ldots,$ is such a sequence.

To establish the necessary part of Theorem II suppose

$$D_N(x, y) = \sum_{n=1}^{N} \delta[x(n), y(n)],$$

δ a metric on A, is a sequence of metrics matched to a certain DMC. Let α and d_N be defined by (18) and (19) respectively. If $\alpha(i, j) < 0$ for some i and j in A, then with $N = 1$ and the code $\{i, j\}$, it follows that if j is received then the MLD will decode it into i, but the minimum D_1-distance decoder will decode it into j. This contradicts the assumption that D_1 is matched to the DMC at 1. Hence $\alpha(i, j) \geq 0$ for every i, j in A. It is equally easy to show that $\alpha(i, j) = 0$ if and only if $i = j$. Next, suppose $\alpha(i, j) < \alpha(j, i)$ for some $i, j \in A$. Consider the code

$\{x_0=(i, i),\ x_1=(j, j)\}$ and let $y=(i, j)$ be received. Then $d_2(x_0, y)=\alpha(i, j)<\alpha(j, i)=d_2(x_1, y)$ and consequently the MLD decodes y into x_0. On the other hand $D_2(x_0, y)=D_2(x_1, y)=\delta(i, j)$ and so the minimum D_2-distance decoder can be made to decode y into x_1 contradicting the fact that it is matched at 2. It therefore follows that $\alpha(i, j)=\alpha(j, i)$ for every i and j in A. Finally, suppose that for some i, j and k, necessarily distinct, in A, $\alpha(i, j)+\alpha(j, k)<\alpha(i, k)$. Let $N=2$ and consider the code $\{x_0=(i, j),\ x_1=(j, i)\}$ and suppose that $y=(j, k)$ is received. Then $d_2(x_0, y)=\alpha(i, j)+\alpha(j, k)<\alpha(i, k)=d_2(x_1, k)$ and so the MLD decodes y into x_0. However,

$$D_2(x_0, y)=\delta(i, j)+\delta(j, k)\geq\delta(i, k)=D_2(x_1, y),$$

and so the minimum D_2-distance decoder can be made to decode y into x_1 contradicting the fact that it is matched at 2. Hence $\alpha(i, j)+\alpha(j, k)\geq\alpha(i, k)$ for every i, j and k in A. This completes the proof of the necessary part of Theorem II. Note, and we will use this fact shortly, that we have only used the fact that D_1 is matched at 1 and D_2 is matched at 2.

We turn to the general case of a sequence of additive metrics $D_N(x, y)=\sum_{n=1}^{N}\delta_n^{(N)}[x(n), y(n)]$ where $\delta_n^{(N)}$ is a metric on A, $N=1, 2, \ldots$. Clearly the conditions of Theorem II are sufficient to assure the existence of such a sequence of metrics matched to a DMC. To prove the converse, suppose that there exists a sequence of additive metrics $\{D_N\}$ matched to a certain DMC. We may, according to Theorem I, assume that D_N is permutation invariant for $N=1, 2, \ldots$. Let $D_2(x, y)=\delta_1[x(1), y(1)]+\delta_2[x(2), y(2)]$ and let σ be the only non-trivial permutation on two objects, then

$$\begin{aligned}\delta_1[x(1), y(1)]+\delta_2[x(2), y(2)]&=D_2(x, y)=D_2(\sigma x, \sigma y)\\&=\delta_1[x(2), y(2)]+\delta_2[x(1), y(1)].\end{aligned}$$

Define δ on A by $\delta=\delta_1+\delta_2$ and let $D_2'(x, y)=\delta[x(1), y(1)]+\delta[x(2), y(2)]$, then

$$\begin{aligned}D_2'(x, y)&=\delta_1[x(1), y(1)]+\delta_2[x(1), y(1)]+\delta_1[x(2), y(2)]+\delta_2[x(2), y(2)]\\&=2D_2(x, y)\end{aligned}$$

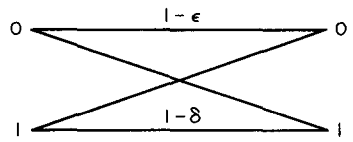
which is clearly matched to the DMC in question. It is an easy matter to show that $D_1'=\delta$ is matched to the said channel at 1. Hence $D_1'(x, y)=\delta[x(1), y(1)]$ and

$$D_2'(x, y)=\sum_{n=1}^{2}\delta[x(n), y(n)]$$

are matched to the said channel at 1 and 2 respectively. By the remark following the proof of the necessity part of Theorem II it follows that the conditions of Theorem II must hold. We have therefore proven the following.

Theorem III. Consider a DMC with input and output alphabets equal to A and with transition probabilities $p(j/i)>0$ for every i and j in A. Then there exists a sequence of additive metrics matched to this channel if and only if the conditions of Theorem II are satisfied. Moreover if those conditions are satisfied then (20) specifies a sequence of additive metrics matched to the said channel.

Example 1. Consider the binary channel as shown.



Condition 1 of Theorem II has it that $1-\delta>\varepsilon$ and $1-\varepsilon>\delta$ which is equivalent to $\varepsilon+\delta<1$. From Condition 2 we have $1-\delta/\varepsilon=1-\varepsilon/\delta$ which reduces to $\varepsilon=\delta$. Hence there exists a sequence of additive metrics matched to the binary DMC if and only if it is symmetric and the cross-over probability ε is less than $\frac{1}{2}$. In this case the metric given by (20) becomes

$$D_N(x, y)=\left(\log\frac{1-\varepsilon}{\varepsilon}\right)d_H(x, y)$$

where d_H is the Hamming distance. The quantity $\log 1-\varepsilon/\varepsilon$ may be dropped to leave the Hamming metric. The fact that the Hamming metric is matched to the BSC is known (**1**).

To show that the conditions of Theorem II do not always lead to a symmetric channel we will consider the case of the ternary channel. Before proceeding to this case it will be expedient to introduce a matrix β which we define by

$$\beta(i, j)=p(j/i)/p(j/j) \quad i, j\in A. \tag{21}$$

The conditions of Theorem II in terms of β are simply:

(1) $0<\beta(i, j)\leq 1$ with equality if and only if $i=j$,
(2) β is symmetric,
(3) $\beta(i, j)\beta(j, k)\leq\beta(i, k)$ for every i, j, k in A.

Since it is clear that

$$\sum_j \beta(i, j)p(j/j)=1 \quad i\in A \tag{22}$$

then this imposes the further restriction on β that there exists an N-tuple q such that

$$0<q(j)<1, \quad j\in A, \quad \text{and} \quad \beta q^t=(1, 1, \ldots, 1)^t \tag{23}$$

where t stands for transpose.

Example 2. Consider for a second example the ternary DMC. In this case we do not attempt a complete description of the transition probabilities which meet the conditions of Theorem II. Instead we describe a class of β matrices whose associated channel does satisfy the conditions of Theorem II. Since β

must be symmetric, we may set

$$\beta=\begin{bmatrix}1 & a & b\\ a & 1 & c\\ b & c & 1\end{bmatrix} \tag{24}$$

where $a, b, c \in (0, 1)$. To meet Condition 2 we must have

$$ac \leq b, \quad ab \leq c \quad \text{and} \quad bc \leq a. \tag{25}$$

The determinant of β is calculated to be

$$\det \beta = 1 + 2abc - (a^2 + b^2 + c^2). \tag{26}$$

We now restrict ourselves to the case

$$\det \beta > 0. \tag{27}$$

In this case the inverse of β is

$$\beta^{-1} = \frac{1}{\det \beta}\begin{bmatrix}1-c^2 & bc-a & ac-b\\ bc-a & 1-b^2 & ab-c\\ ac-b & ab-c & 1-a^2\end{bmatrix}. \tag{28}$$

By (23) we must have a q such that

$$\beta^{-1}\begin{bmatrix}1\\ 1\\ 1\end{bmatrix} = q^t \tag{29}$$

where $0 < q(j) < 1$, $j = 0, 1, 2$. In order to see when this will be the case subtract $\det \beta$ from the sum of the elements of the first row of (28) to obtain

$$1 - c^2 + bc - a + ac - b - 1 - 2abs + a^2 + b^2 + c^2$$

which reduces to

$$(a-1)(a-bc) + (b-1)(b-ac). \tag{30}$$

We already have that $a-1<0$, $b-1<0$ and $a-bc \geq 0$, $b-ac \geq 0$, and one of the last inequalities must be strict, therefore (30) is always strictly less than zero, hence

$$q(0) = \frac{(1-c^2)+(bc-a)+(ac-b)}{\det \beta} < 1$$

and will be greater than zero if

$$(1-c^2)+(bc-a)+(ac-b) > 0$$

which reduces to

$$1 + c > a + b. \tag{31}$$

Similarly if $1+b > a+c$ and $1+a > b+c$, then $0 < q(j) < 1$ for $i = 1, 2$.

Let P denote the transition probability matrix whose (i, j) entry is by definition $p(j/i)$, then we collect the above results as:

Theorem IV. For the ternary DMC with transition probability matrix

$$P=\frac{1}{\det\beta}\begin{bmatrix}(1-c)(1+c-a-b) & a(1-b)(1+b-a-c) & b(1-a)(1+a-b-c)\\ a(1-c)(1+c-a-b) & (1-b)(1+b-a-c) & c(1-a)(1+a-b-c)\\ b(1-c)(1+c-a-b) & c(1-b)(1+b-a-c) & (1-a)(1+a-b-c)\end{bmatrix}$$

where:

(1) $\det\beta = 1+2abc-(a^2+b^2+c^2)>0$,
(2) $ac\leq b,\ ab\leq c,\ bc\leq a$,
(3) $1+c>a+b,\ 1+b>a+c,\ 1+a>b+c$,

the function defined by (20) gives a sequence of metrics matched to the channel in question.

As a specific example of Theorem IV, consider the β matrix

$$\beta=\begin{bmatrix}1 & 1/2 & 1/4\\ 1/2 & 1 & 1/8\\ 1/4 & 1/8 & 1\end{bmatrix}.$$

It satisfies all the conditions of Theorem IV. Its determinant is 45/64. Computing its inverse and then solving for $(p(0/0), p(1/1), p(2/2))$ according to (29) we obtain

$$(21/45,\ 2/3,\ 12/15).$$

The corresponding transition probability matrix is

$$\begin{bmatrix}21/45 & 1/3 & 1/5\\ 21/90 & 2/3 & 1/10\\ 21/180 & 1/12 & 4/5\end{bmatrix},$$

and we note that this does not correspond to a symmetric channel as defined in Gallager (**3**).

III. Conclusions

In this paper we have given necessary and sufficient conditions on the transition probabilities of a DMC for the existence of a sequence of additive matched metrics. We have done this for channels with the same input and output alphabet A and such that $p(j/i)>0$ for every i and j in A. It is not necessary for the input and output alphabets to be same, all that is needed is that the number of inputs be equal to the number of outputs, the label we place on these is immaterial.

The work presented in this paper can be pursued in at least two directions. First, it would be nice to remove the restriction that the number of inputs and outputs be the same. Second, it would be interesting to see what happens if we remove the restriction that the metrics be additive.

References

(**1**) James L. Massey, "Notes on Coding Theory" Waltam Research Centre, General Telephone and Electronics Lab. Inc., Waltam, Mass., 1969

(**2**) J. Chung-Yaw Chiang and Jack K. Wolf, "On Channels and Codes for the Lee Metric", Information and Control, Academic Press Inc., New York, Vol. 19, No. 2, pp. 159–173, Sept. 1971.

(**3**) Robert G. Gallager, "Information Theory and Reliable Communication", p. 94, John Wiley and Sons, New York, 1968.

(**4**) Andrew Viterbi and J. Omura, "Principles of Digital Communication and Coding", McGraw Hill, New York, 1979.

(**5**) C. E. Shannon, "A mathematical theory of communication", *Bell Syst. Tech. J.*, Vol. 27, pp. 379–423 and pp. 623–656, 1948.

(**6**) A. D. Wyner, "Capabilities of bounded discrepancy decoding", *Bell Syst. Tech. J.*, Vol. 44, pp. 1061–1122, 1965.

(**7**) C. Y. Lee, "Some properties of non-binary error-correcting codes", *IEEE Trans. Information Theory*, IT-4, pp. 77–82, 1958.