NFORMATION AND CONTROL 37, 150–177 (1978)

Optimal Encodings of Linear Block Codes for Unequal Error Protection

LARRY A. DUNNING

Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27607

AND

W. E. ROBBINS

Department of Computer Science, North Carolina State University, Raleigh, North Carolina 27607

It is possible for a linear block code to provide more protection for selected message positions than is guaranteed by the minimum distance of the code. The protection provided a message position can be measured by associating a number with that position called its separation. The separation of a message position measures the protection provided to that position in a manner analogous to that in which the minimum distance of a code measures the protection provided the entire message. This paper proves that any fixed linear block code has an encoding which is optimal with respect to the error protection provided the individual message positions. More precisely, among those encodings of the code for which the separations associated with the message positions are arranged in nondecreasing order, there is at least one which simultaneously maximizes all the separations associated with the message positions. A procedure is given which may be used to construct optimal encodings for linear codes of small dimension. When the Hamming metric is employed, the procedure builds a generator matrix which is as sparse as possible for the given code. At each iteration the procedure adds a row to a partially constructed generator matrix. A code word of minimum weight is chosen for this purpose—subject to the restriction that the rows of the generator matrix must be linearly independent. A more general result is that any generator matrix which is as sparse as possible induces an optimal encoding of its row space. A similar result holds when the Lee metric is used to model a channel. Theorems dealing with cyclic codes and product codes are developed. Under suitable restrictions, an optimal generator matrix for a cyclic code may be formed by concatenating the generator matrices of the minimal ideals which are contained in it. When the Hamming metric is employed, an optimal generator matrix for a product code may be obtained by taking the Kronecker product of optimal generator matrices for the component codes.

0019-9958/78/0372-0150$02.00/0

1. Introduction and Preliminaries

We shall restrict our attention to (n, k) block codes. Let $F \triangleq GF(q)$ be any finite field where q is a prime power. Throughout this paper, an (n, k) *code* will be a subset of F^n of cardinality q^k where $k \leqslant n$. By an *encoding* of a code, C, we mean any bijection $\eta: F^k \to C$. If C is a vector subspace of F^n, then it is said to be a linear code. In this case we will say that a $k \times n$ matrix with entries from F is a generator matrix for C if its rows form a basis for C. There is a natural one-to-one correspondence between the linear encodings and the generator matrices of a linear code. Every generator matrix, G, induces a linear encoding, L, defined by the formula,

$$L(m) \triangleq mG \qquad \forall m \in F^k,$$

where the message vector, m, and all vectors throughout the paper are identified with row matrices. For most applications of block codes, it is sufficient to study codes without reference to their encodings. However, this has not always been the case.

The construction of codes in which some message positions might be provided protection against a greater number of errors than others has been considered by several authors (Masnick and Wolf, 1967; Gore and Kilgus, 1971; Kilgus and Gore, 1972a; Mandelbaum, 1972). Masnick and Wolf (1967) proved that cyclic codes in systematic form provide equal error protection for every information digit. A nonsystematic cyclic code which provides one "information digit" protection against errors, beyond that guaranteed by the minimum distance of the code, was exhibited by Gore and Kilgus (1971). Thus, it became apparent that the protection against error afforded individual message positions depends not only on the code used, but also upon the encoding used. A direct means of establishing this result is to inspect the mappings η_1, η_2: $GF(2)^2 \to GF(2)^4$ given in Table I. η_1 and η_2 are two different encodings for the same code. Given a received word containing at most a single error, one can determine whether the code word originally transmitted was of the form $\eta_1(m_1, 0)$ or of the form $\eta_1(m_1, 1)$. Thus, the encoding, η_1, allows determination of the second message bit, m_2, despite any single error. However, consideration of the received word 1000 shows that the encoding η_2 fails to protect either message bit against all single errors.

TABLE I

$\eta_i(m_1, m_2)$	$\eta_1(m_1, 0)$	$\eta_1(m_1, 1)$	$\eta_2(m_1, 0)$	$\eta_2(m_1, 1)$
$m_1 = 0$	0000	0111	0000	0111
$m_1 = 1$	1100	1011	1011	1100

One place where *unequal-error-protection* codes were expected to find application was in the transmission of digital telemetry data. Here it may be desirable to give high order digits more protection than low order digits. Calculated data for several such codes so employed were given by Kilgus and Gore (1972b). Recently, several papers (Crimmins, 1976; Crimmins *et al.* 1969; Crimmins and Horowitz, 1970; Redinbo, 1976; Redinbo and Wolf, 1974; and Wolf and Redinbo, 1974) studying the *mean-square-error* protection afforded numeric data by block coding schemes have appeared. The approach in these papers is not to construct codes, but rather to find optimal encoding[1] and decoding schemes for a fixed linear code. Crimmins *et al.* (1969) gave a restricted formulation of the problem in which each encoding of a binary linear code generates a decoding scheme in a prescribed manner. They gave a procedure for finding linear encodings, which are optimal in the set of all encodings, linear and nonlinear, of the fixed binary linear code under consideration.

Our purpose is to investigate the encodings of a fixed linear code. However, we shall use the unequal-error-protection approach to evaluate and compare these encodings instead of the mean-square-error evaluation. The mean-square-error evaluation method of Crimmins *et al.* (1969) associates a nonnegative real number with each encoding. Since each code has only finitely many possible encodings, one of the encodings must have mean-square-error as small (good) as possible for the code. Thus, it is immediate that every code has an encoding which is optimal with respect to the mean-square-error evaluation. Using the unequal-error-protection approach we will prove that optimal encodings exist for linear codes. In doing this a procedure will be found for obtaining an encoding which is optimal in the set of all encodings, linear and nonlinear. This result parallels that of Crimmins *et al.* (1969), and the procedure found is similar to theirs. Further, when the encodings of a linear code are evaluated using a measure of unequal-error-protection based on either the Hamming or the Lee metric, the procedure will yield a linear encoding which is optimal among all encodings of the fixed linear code under consideration. In these cases, any generator matrix which has minimal Hamming or Lee weight, respectively, among all generator matrices for its row space, induces an encoding which is optimal for its row space.

Masnick and Wolf (1967) assign each information position an *error protection level.* Under this scheme, if an information position has error protection level, f, and not more than f errors occur in the reception of a code word, then the original value of the position in question can be determined correctly even though it may be impossible to determine the entire code word correctly. Instead of using this generalization of the error correcting capability of a code, we employ

[1] When numeric data are encoded using a 1–1 mapping (e.g., Crimmins *et al.*, 1969) from $\{0, 1, \ldots, 2^k - 1\}$ onto a code, we identify these integers with their binary representations (following Mitryayev, 1963) to obtain an equivalent encoding.

a generalization of the minimum distance of a block code. Given an encoding of a block code, for each message position we will define an associated *separation*, which is related to its error protection level in the same manner that the minimum distance of a block code is related to its error correcting capability. Encodings which we find to be optimal will necessarily be optimal with respect to their error protection levels.

Block codes may be used to detect errors, correct errors, fill in erasures, or combinations of these things. Fortunately, one parameter, minimum distance, suffices to measure the capabilities of a block code regardless of the type of protection desired—provided that one stays within the list given. However, different decoding algorithms are used depending on the task at hand. Given a particular encoding, the separation associated with a message position measures the capability of a block code to detect errors which may cause that position to be in error, determine that position despite errors, determine that position despite erasures, or combinations of these things in an analogous manner. The decoding algorithms, given later, differ very little from those used when all positions receive the same protection. Depending on the types of protection desired, the message positions may be decoded separately or as a unit. Treating the message as a unit will not necessarily preclude giving different positions varying degrees of protection.

As was mentioned before, the protection provided to the message positions depends upon the encoding as well as the code. The generator matrices and encodings of interest are frequently nonsystematic. That is, the message positions may not appear explicitly in the code words. We will *not* be able to make reference to "the information positions" of the code word. Before an encoding function can be used, an inverse mapping must be constructed for use as a part of the decoding rule. When we speak of choosing an optimal encoding, we will also be choosing decoding rules which will depend both on the encoding and on the type of error protection desired for each message position.

In order to handle the Hamming and Lee metrics simultaneously, we will develop results with respect to a function, $w: F^n \to \mathbb{R}$, which has the property that the function $d: F^n \times F^n \to \mathbb{R}$ given by

$$d(x, y) \triangleq w(x - y)$$

is a metric. Such a function, w, will be called a *weight function*. We will have occasion to refer to the Hamming weight function specifically and will denote it by h. One can easily verify that necessary and sufficient conditions for a function, $w: F^n \to \mathbb{R}$, to be a weight function are that for all $x, y \in F^n$

(i) $w(x) = 0$ if and only if $x = \overline{0_n}$,

(ii) $w(x) = w(-x)$,

(iii) $w(x + y) \leqslant w(x) + w(y)$,

where $\overline{0_n}$ denotes the zero vector in F^n.

Suppose $X \subseteq F^n$. It will be convenient to abbreviate $w[\Phi] = +\infty$, else

$$w[X] \triangleq \min_{x \in X} w(x).$$

This is not to be confused with the usual conventions for extending point functions to sets, i.e., for example, $w(X) \triangleq \{w(x): x \in X\}$.

Since we will be dealing with linear codes, which are the row space of their generator matrices, it is desirable to develop some notational devices to assist in arguments involving the rows of a matrix. Given a $k \times n$ matrix M with entries in F, we denote the entry in row i and column j by M_{ij}, the ith row (vector) by $M_{i\cdot}$ and the jth column (vector) by $M_{\cdot j}$. The set of all rows of M is denoted by

$$M_r \triangleq \{M_{1\cdot}, \ldots, M_{k\cdot}\}.$$

For any function, $f: F^n \to S$, where S is any set, define

$$f_r(M) \triangleq \begin{pmatrix} f(M_{1\cdot}) \\ \vdots \\ f(M_{k\cdot}) \end{pmatrix}.$$

Thus, $f_r(M) \in S^k$ is a vector whose ith component is found by applying f to the ith row of M.

Our main line of argument will require only some elementary knowledge of linear algebra. When listed, vectors will always be enclosed in parentheses and matrices in brackets. In our discussion of product codes, some properties of the (left) Kronecker product of matrices will be required. In particular, recall that

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

We shall take Kronecker products over both F and $\mathbb{R}$ and will denote the respective operators by $\otimes_F$ and $\otimes_{\mathbb{R}}$.

The Hamming weight function, h, applied to a matrix will count the number of nonzero entries in that matrix. The Lee weight function applied to a matrix will add the Lee weights of its entries. It is easy to show that given two vectors $a \in F^k$, $b \in F^l$ their Hamming weights are related by

$$h(a \otimes_F b) = h(a)h(b). \tag{1}$$

This result may be used to prove that, given any two matrices A, B with entries in F,

$$h_r(A \otimes_F B) = h_r(A) \otimes_{\mathbb{R}} h_r(B). \tag{2}$$

We will denote the span operator by $\langle \cdot \rangle$. Given a *set* S of vectors taken from a finite vector space, V, over the field F,

$$\langle S \rangle = \left\{ \sum_{s \in S} \alpha_s s : \alpha \in F^S \right\}$$

is the subspace consisting of all linear combinations of elements of S. $\langle S \rangle$ is the smallest subspace of V containing S. According to convention, $\langle \Phi \rangle = \{\bar{0}\}$ is the subspace consisting of only the zero vector. Set brackets may be omitted when taking the span of a list of vectors, e.g., $\langle a, b, c \rangle = \langle \{a, b, c\} \rangle$.

2. The Minimum Separation Approach

One possible expression for the minimum distance, d, of a code, C, with encoding $\eta: F^k \to C$ is

$$d = w[\{c - c' : c \neq c'\}] = w[\{\eta(m) - \eta(m') : m \neq m'\}]$$

where c, c' range over C and m, m' over F^k. We make a definition that is syntactically similar.

Definition 1. Given an encoding, η, of an (n, k) code, C, the *separation vector* of η with respect to a weight function, w, is denoted by $S_w(\eta) \in \mathbb{R}^k$ and is defined by:

$$S_w(\eta)_i \triangleq w[\{\eta(m) - \eta(m') : m_i \neq m_i'\}] \qquad (i = 1, \ldots, k), \tag{3}$$

where m and m' range over F^k.

When no confusion will result, subscripted references to weight functions may be dropped. If L is the encoding induced by the generator matrix, G, then we may write $S(G)$ or $S_w(G)$ instead of $S_w(L)$. We will always subscript h for emphasis when a separation vector is taken with respect to the Hamming weight function. It is easily shown that the minimum distance, d, and any separation vector, $S(\eta)$, of a code are related by $d = \min\{S(\eta)_1, \ldots, S(\eta)_k\}$. The ith component of the separation vector is used to guarantee protection for the ith message position. Since correct determination of all message positions is equivalent to correct determination of the entire message, the last equation should not be unexpected.

Theorem 1. *Given an encoding, η, for an (n, k) code,*

(a) *η allows detection of any error pattern of w-weight not more than d at the ith message position if and only if $d < S(\eta)_i$.*

(b) *η allows correction of any error pattern of w-weight not more than t at the ith message position if $2t < S(\eta)_i$.*

(c) *η allows correction of any error pattern of w-weight not more than t and simultaneous detection of any error pattern of w-weight not more than $d \geq t$ at the ith message position if $t + d < S(\eta)_i$.*

Proof of (b). Suppose that $m \in F^k$ is a message and that $c \triangleq \eta(m) \in C$ is transmitted through a channel. Let $e \in F^n$ be any error pattern of w-weight not more than t. If c is perturbed by e in passage through the channel, then the received word is given by $r \triangleq c + e$. Consider the following decoding procedure.

Maximum likelihood decoding procedure. (1) Find any code word c' which is as close to r (with respect to the metric induced by w) as any other code word.

(2) Set $m' = \eta^{-1}(c')$ and guess $m_i = m_i'$.

Clearly $w(c' - r) \leqslant t$ since c' must be at least as close to r as c is. Denote $m' \triangleq \eta^{-1}(c')$. Now, m_i is given correctly by $m_i = m_i'$. If this were not the case, then a contradiction arises since using Eq. (3) we obtain

$$S(\eta)_i \leqslant w(\eta(m) - \eta(m')) = w(c - c') < w(c - r) + w(r - c') \leqslant t + t < S(\eta)_i .$$

Proof of (c). Suppose that a message $m \in F^k$ is encoded to give a code word $c \triangleq \eta(m)$ which is perturbed by an error pattern $e \in F^n$ to yield a received word $r \triangleq c + e$. Consider the following decoding procedure (Wyner, 1965).

Bounded distance decoding procedure. (1) Find any code word c' which is as close to r as any other code word.

(2) If $w(c' - r) \leqslant t$ find $m' \triangleq \eta^{-1}(c')$, guess $m_i = m_i'$, and stop.

(3) Otherwise, declare that the ith message position has a detected error.

This is an extension of the maximum likelihood decoding procedure used in the proof of (b). Since $2t \leqslant t + d < S(\eta)_i$, we already know that m_i will be determined correctly by this scheme whenever $w(e) \leqslant t$. We must show that if $w(e) \leqslant d$, then the algorithm will either determine m_i correctly or declare a detected error. In fact, assume to the contrary that $w(e) \leqslant d$ and $m_i' \neq m_i$ is computed at step (2). Then, using Eq. (3), a contradiction arises since

$$S(\eta)_i \leqslant w(\eta(m) - \eta(m')) \leqslant w(c - r) + w(r - c') \leqslant d + t < S(\eta)_i .$$

Proof of (a). Sufficiency may be obtained by taking $t = 0$ in (c). For the necessity let $m, m' \in F^k$ satisfy $m_i \neq m_i'$ and $S(\eta)_i = w(\eta(m) - \eta(m'))$. Taking $c \triangleq \eta(m')$, $e \triangleq \eta(m) - \eta(m')$, and $r \triangleq c + e$, we see that $d < S(\eta)_i$. Q.E.D.

Let η be an encoding of an (n, k) code, C. For the Hamming weight function, part (b) of the last theorem implies that for any $i \in \{1,..., k\}$ the ith message digit is protected against t errors whenever $2t + 1 \leqslant S_h(\eta)_i$. Thus, the "error protection level" associated with the ith position is lower bounded by, f,

$$f \triangleq \left\lfloor \frac{S_h(\eta)_i - 1}{2} \right\rfloor, \tag{4}$$

where $\lfloor \cdot \rfloor$ denotes the greatest integer function. In fact, the "error protection level" associated with the ith position is given exactly by (4); i.e., the value of the ith message position can be determined despite any occurrence of f errors, but not despite any occurrence of $f+1$ errors. This relates our evaluation system to that of Masnick and Wolf (1967).

To prove the reverse inequality, consider any $i \in \{1,\ldots, k\}$. There exist $m, m' \in F^k$ satisfying $m_i \neq m_i'$ and

$$S_h(\eta)_i = w(\eta(m) - \eta(m')).$$

Note that f is the largest integer satisfying $2f+1 \leqslant S_h(\eta)_i$. Set

$$J \triangleq \{j: \eta(m)_j \neq \eta(m')_j,\ 1 \leqslant j \leqslant n\},$$

and choose $E \subseteq J$ satisfying $|E| = f+1 \leqslant |J| = S_h(\eta)_i$ where $|\cdot|$ signifies cardinality. Define $e, e' \in F^n$ by

$$e_j \triangleq \begin{cases} 0, & j \notin E, \\ \eta(m')_j - \eta(m)_j, & j \in E, \end{cases} \qquad e_j' \triangleq \begin{cases} \eta(m)_j - \eta(m')_j, & j \in J \backslash E, \\ 0, & j \notin J \backslash E, \end{cases}$$

where $\backslash$ denotes set difference. Since $2(f+1) \geqslant S_h(\eta)_i$ it follows that the number of nonzero components of e' is at most $f+1$. Now, suppose that the received word, r, given by

$$r \triangleq \eta(m) + e = \eta(m') + e'$$

is encountered. Given only that no more than $f+1$ errors occurred, there are at least two possibilities.

(1) The original message was m and error pattern e occurred ($f+1$ errors).

(2) The original message was m' and error pattern e' occurred ($\leqslant f+1$ errors).

The values of the ith message positions, m_i and m_i', in these cases are different, which completes the proof.

The procedure(s) given in the proof of Theorem 1 can be used to guarantee protection for individual-message positions even when an encoding–decoding scheme handles the entire message as a unit. We note two special cases. Suppose C to be a linear (n, k) code with generator matrix, G, and parity check matrix, H. Since G has maximal row rank, there exists an $n \times k$ matrix, G^-, which is a right inverse of G and satisfies $G \cdot G^- = I_k$ where I_k is the $k \times k$ identity matrix. G^- gives us a representation of the inverse of the encoding induced by G. Now, the scheme depicted in Fig. 1 will either give an error indication or the correct value for the ith message position in its output whenever $w(e) < S(G)_i$; and this statement is true for each $i \in \{1,\ldots, k\}$. For each $s \in F^{n-k}$ set

$$V_s \triangleq \{v \in F^n: vH^t = s\},$$

and choose $l(s) \in V_s$ satisfying

$$w(l(s)) = w[V_s].$$

Thus, for each syndrome, s, we have chosen a minimum w-weight coset leader, $l(s)$, of its associated coset, V_s. Now, the scheme depicted in Fig. 2 determines a reconstructed message

$$m' \triangleq (r - l(rH^t))\, G^{-}.$$

The value of the ith position of the reconstructed message is correct whenever $2w(e) < S(G)_i$, and this is true for each $i \in \{1,..., k\}$.

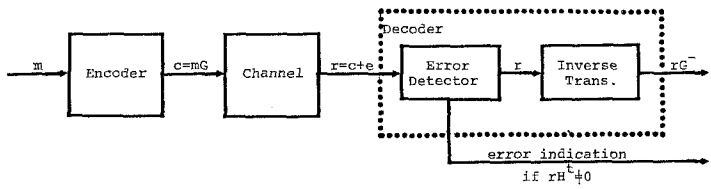

FIG. 1. An error detection scheme for linear codes.

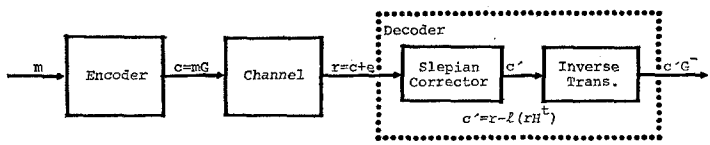

FIG. 2. A minimum distance decoding scheme for linear codes.

While other results in this direction are possible, we shall be content to state one more. Its proof may be constructed from the proof of the parallel result for ordinary minimum distance by generalizing in the same manner as we did for the previous theorem.

THEOREM 2. *Given an encoding, η, for an (n, k) code, C, η allows determination of the* ith *message position despite any simultaneous occurrence of not more than t errors and not more than e erasures if and only if $2t + e + 1 \leqslant S_h(\eta)_i$.*

If $\eta: F^k \to C$ is an encoding and $\phi: \{1,..., k\} \to \{1,..., k\}$ is a permutation, then define another encoding, η_ϕ, for C by $\eta_\phi(m) \triangleq \eta(m \circ \phi^{-1})$ where $\circ$ denotes composition of functions; i.e., for $m = (m_1,..., m_k) \in F^k$,

$$\eta_\phi(m_{\phi(1)},..., m_{\phi(k)}) = \eta(m_1,..., m_k). \tag{5}$$

Now, given any $i \in \{1,..., k\}$ it follows that (for any given weight function)

$$\begin{aligned} S(\eta_\phi)_i &= w[\{\eta_\phi(m) - \eta_\phi(m') : m_i \neq m_i'\}] \\ &= w[\{\eta(m \circ \phi^{-1}) - \eta(m' \circ \phi^{-1}) : m_i \neq m_i'\}] \\ &= w[\{\eta(m) - \eta(m') : (m \circ \phi)_i \neq (m' \circ \phi)_i\}] \\ &= S(\eta)_{\phi(i)}, \end{aligned}$$

where m and m' range over F^k in the above expressions. We have shown

$$S(\eta_\phi) = S(\eta) \circ \phi. \tag{6}$$

Thus, the separation vectors associated with η and η_ϕ differ only by a permutation of coordinates. We shall regard such pairs of encodings as being equivalent. This discussion will be incorporated into a formal result after a prerequisite definition.

Given $x \in \mathbb{R}^k$, we define $x^* \in \mathbb{R}^k$ to be the vector obtained from x by permuting its coordinates to obtain a nondecreasing vector. For any $x \in \mathbb{R}^k$, there exists a permutation $\phi: \{1,..., k\} \to \{1,..., k\}$ such that $x^* = x \circ \phi^{-1}$ and

$$x_1^* \leqslant x_2^* \leqslant \cdots \leqslant x_k^* \tag{7}$$

where x_i^* is interpreted as $(x^*)_i$. Thus, $x = x^*$ if and only if x is nondecreasing.

PROPOSITION 1. *Given an encoding, η, of an (n, k) code, C, there exists another encoding, ξ, of the same code such that*

$$S(\xi) = S(\xi)^* = S(\eta)^*.$$

If $\overline{0_n} \in C$ then ξ may be chosen so as to satisfy $\xi(\overline{0_k}) = \overline{0_n}$.

Proof. Given η, choose a permutation $\phi: \{1,..., k\} \to \{1,..., k\}$ satisfying $S(\eta)^* = S(\eta) \circ \phi$. We have already shown that the encoding η_ϕ defined by $\eta_\phi(m) \triangleq \eta(m \circ \phi^{-1})$ and satisfying (5) has separation vector

$$S(\eta_\phi) = S(\eta) \circ \phi = S(\eta)^*.$$

If $\overline{0_n} \notin C$, set $\xi \triangleq \eta_\phi$. If $\overline{0_n} \in C$, then set $z \triangleq \eta_\phi^{-1}(\overline{0_n})$; and define ξ by

$$\xi(m) \triangleq \eta_\phi(z - m) \qquad \forall m \in F^k.$$

Clearly $\xi(\overline{0_k}) = \eta_\phi(z - \overline{0_k}) = \eta_\phi(z) = \overline{0_n}$. The separation vectors associated with η_ϕ and ξ are again the same since for $i = 1,..., k$,

$$\begin{aligned} S(\xi)_i &= w[\{\xi(m) - \xi(m') : m_i \neq m_i'\}] \\ &= w[\{\eta_\phi(z - m) - \eta_\phi(z - m') : m_i \neq m_i'\}] \\ &= w[\{\eta_\phi(m) - \eta_\phi(m') : (z - m)_i \neq (z - m')_i\}] \\ &= w[\{\eta_\phi(m) - \eta_\phi(m') : m_i \neq m_i'\}] \\ &= S(\eta_\phi)_i = S(\eta)_i^*, \end{aligned}$$

where m and m' range over F^k in the above expressions. In either case, $\xi \triangleq \eta_\phi$ or $\xi(m) \triangleq \eta_\phi(z - m)$; our proof is complete with the observation that

$$S(\xi) = S(\eta_\phi) = S(\eta)^* = (S(\eta)^*)^* = S(\xi)^*. \qquad \text{Q.E.D.}$$

Some simplification of the expression for the separation vector is possible for linear encodings. Suppose G is a generator matrix for an (n, k) code C, then for each $i = 1,..., k$

$$\begin{aligned} S(G)_i &= w[\{mG - m'G: m_i \neq m_i'\}] \\ &= w[\{(m - m')G: (m - m')_i \neq 0\}] \\ &= w[\{mG: m_i \neq 0\}] = w[C\backslash\{mG: m_i = 0\}] \\ &= w[C\backslash\langle G_r\backslash G_i.\rangle]; \end{aligned}$$

where we have written "$G_i.$" for the singleton set $\{G_i.\}$ and m and m' range over F^k in each expression. If $S(G)$ is nondecreasing, then we can show that for each $i \in \{1,..., k\}$

$$S(G)_i = w[\{mG: m_j \neq 0 \text{ for some } j \geqslant i\}].$$

That $S(G)_i$ is not less than the right-hand expression follows from $\{mG: m_i \neq 0\} \subseteq \{mG: m_j \neq 0 \text{ for some } j \geqslant i\}$. The reverse inequality is true since if $m \in F^k$ and $m_j \neq 0$ for some $j \geqslant i$, then $S(G)_i \leqslant \cdots \leqslant S(G)_j \leqslant w(mG)$. Another expression for the right-hand side is $w[C\backslash\langle G_1., ..., G_{(i-1)}.\rangle]$. We have shown the following proposition.

PROPOSITION 2. *Given a generator matrix, G, for a linear (n, k) code C, the following expressions are equal for any fixed $i \in \{1,..., k\}$:*

(a) $S(G)_i$,

(b) $W[\{mG: m_i \neq 0\}]$,

(c) $w[C\backslash\langle G_r\backslash G_i.\rangle]$.

If $S(G)$ is nondecreasing, then the following expression is equal to each of the preceding for any fixed $i \in \{1,..., k\}$:

(d) $w[C\backslash\langle G_1., ..., G_{(i-1)}.\rangle]$.

The following corollary is an immediate consequence of either Proposition 2, (a) = (c) or of Eqs. (5) and (6).

COROLLARY 1. *Let $\phi: \{1,..., k\} \to \{1,..., k\}$ be a permutation. If G is a $k \times n$ generator matrix, then*

$$S\left(\begin{bmatrix} G_{\phi(1)}. \\ \vdots \\ G_{\phi(k)}. \end{bmatrix}\right) = \begin{pmatrix} S(G)_{\phi(1)} \\ \vdots \\ S(G)_{\phi(k)} \end{pmatrix} = S(G) \circ \phi.$$

Before turning to our main results we shall develop one more result in which the separation vector plays a role analogous to that usually played by the minimum distance of a code. Let A be a generator matrix for a linear (n_1, k_1) code, C_1, over $GF(q)$, and let B be a generator matrix for a linear (n_2, k_2) code, C_2, over $GF(q)$. Then the Kronecker product

$$A \otimes_F B \triangleq \begin{bmatrix} A_{11}B & \cdots & A_{1n_1}B \\ \vdots & & \vdots \\ A_{k_11}B & \cdots & A_{k_1n_1}B \end{bmatrix},$$

of A and B forms a generator matrix for a linear $(n_1 \cdot n_2, k_1 \cdot k_2)$ code C. C depends only upon C_1 and C_2, not upon the particular choice of A and B, and is called the Kronecker product of the codes C_1 and C_2 (e.g., see Section 1.5 of Blake and Mullin, 1975). Another form of the following theorem was stated by Kilgus (1971) for the case in which both C_1 and C_2 are majority logic decodable.

THEOREM 3. *Let A be a generator matrix for a linear (n_1, k_1) code and let B be a generator matrix for a linear (n_2, k_2) code, both over the same finite field, F. Then,*

$$S_h(A \otimes_F B) = S_h(A) \otimes_{\mathbb{R}} S_h(B).$$

Proof. For convenience denote $k = k_1 \cdot k_2$. Given $v \in F^k$ or $v \in \mathbb{R}^k$, define

$$f(v) = \begin{bmatrix} v_1 & \cdots & v_{k_2} \\ v_{k_2+1} & \cdots & v_{2k_2} \\ \vdots & & \vdots \\ v_{k-k_2+1} & \cdots & v_k \end{bmatrix}.$$

Given any $m \in F^k$, the reader may verify that

$$m(A \otimes_F B) = ((A^t f(m)B)_1, \ldots, (A^t f(m)B)_{n_1}).$$

Thus

$$h(m(A \otimes_F B)) = h(A^t f(m)B).$$

Now it may be verified that

$$f(S_h(A \otimes_F B)_{ij} = h[\{A^t MB: M_{ij} \neq 0\}],$$

where M ranges over all $k_1 \times k_2$ matrices with entries in F; and that

$$f(S_h(A) \otimes_{\mathbb{R}} S_h(B))_{ij} = S_h(A)_i S_h(B)_j,$$

for i, j satisfying $1 \leqslant i \leqslant k_1$, $1 \leqslant j \leqslant k_2$. We will prove that

$$S_h(A)_i \cdot S_h(B)_j \leqslant h[\{A^t MB: M_{ij} \neq 0\}], \tag{8}$$

which will imply that $f(S_h(A) \otimes_{\mathbb{R}} S_h(B)) \leqslant f(S_h(A \otimes_F B))$ and hence that $S_h(A) \otimes_{\mathbb{R}} S_h(B) \leqslant S_h(A \otimes_F B)$.

At this point, we have merely converted the statement of the problem to the direct product representation. Note that the rows of A^tMB are always code words in C_2, and the columns of $A^tMB = [(MB)^tA]^t$ are always code words in C_1. We shall now parallel the proof usually employed (e.g., see Theorem 5.3 of Peterson and Weldon, 1972) to show that the minimum distance of the direct product code is the product of the minimum distances of the component codes.

Let integers i, j satisfying $1 \leqslant i \leqslant k_1$ and $1 \leqslant j \leqslant k_2$ be given. Suppose a $k_1 \times k_2$ matrix M with entries in F is given and $M_{ij} \neq 0$. Since the jth position of $M_{i\cdot}$ is nonzero, it follows that $(M_{i\cdot})B = (MB)_{i\cdot}$ is a code word in C_2 with at least $S_h(B)_j$ nonzero entries. Abbreviate $l \triangleq S_h(B)_j$ and let $(MB)_{ij(1)}, \ldots, (MB)_{ij(l)}$ be any $S_h(B)_j$ nonzero entries of $(MB)_{i\cdot}$. Now, for $k \in \{1, \ldots, l\}$, $((MB)^t)_{j(k)\cdot}$ has nonzero entry $((MB)^t)_{j(k)i} = (MB)_{ij(k)}$ in the ith position. Thus, $[((MB)^t)_{j(k)\cdot}] \cdot A = [(MB)^tA]_{j(k)\cdot}$ is a code word in C_1 with at least $S_h(A)_i$ nonzero entries. At least $S_h(A)_i$ of the rows of $(MB)^tA$ have at least $S_h(B)_j$ nonzero entries; i.e., $(MB)^tA$ has at least $S_h(A)_i \cdot S_h(B)_j$ nonzero entries. It follows that

$$h(A^tMB) = h((A^tMB)^t) = h((MB)^tA) \geqslant S_h(A)_i \cdot S_h(B)_j .$$

Since M was an arbitrary $k_1 \times k_2$ matrix with entries in F, except for the stipulation $M_{ij} \neq 0$; we have shown (8).

For the reverse inequality let any integer $l \in \{1, \ldots, k\}$ be given. There exist unique positive integers $i \in \{1, \ldots, k_1\}$ and $j \in \{1, \ldots, k_2\}$ such that $l = (i-1)k_2 + j$. Let $a \in F^{k_1}$ and $b \in F^{k_2}$ be chosen so that $a_i \neq 0$, $b_j \neq 0$ and $S_h(A)_i = h(aA)$, $S_h(B)_j = h(bB)$. It follows that $(a \otimes_F b)_l \neq 0$. The inequality is at hand. For arbitrary $l \in \{1, \ldots, k\}$, we apply (1) to obtain

$$\begin{aligned}
(S_h(A) \otimes_{\mathbb{R}} S_h(B))_l &= S_h(A)_i \cdot S_h(B)_j \\
&= h(aA) \cdot h(bB) \\
&= h(aA \otimes_F bB) \\
&= h((a \otimes_F b) \cdot (A \otimes_F B)) \\
&\geqslant h[\{m(A \otimes_F B) : m_l \neq 0\}] \\
&= S_h(A \otimes_F B)_l
\end{aligned}$$

where m ranges over F^k. Q.E.D.

3. Optimal Encodings

We shall impose the usual partial order on $\mathbb{R}^k$; i.e., given $x, y \in \mathbb{R}^k$, we will write $x \leqslant y$ if and only if $x_i \leqslant y_i$ for $i = 1, \ldots, k$. Given a set of vectors $A \subseteq \mathbb{R}^k$, $a \in A$ will be said to be *Gale optimal in A* (cf. p. 277 of Lawler, 1976) provided

that given any other member $b \in A$, there exists a permutation $\phi: \{1,..., k\} \to \{1,..., k\}$ such that

$$a \circ \phi \geqslant b.$$

In our notation this may be formulated as $a \in A$ is Gale optimal in A if and only if for all $b \in A$

$$a^* \geqslant b^*.$$

With this definition at hand we are prepared to give meaning to the term "optimal."

DEFINITION 2(a). Let C be a code and let $\mathscr{E}$ denote the set of all encodings for C. $\eta \in \mathscr{E}$ will be said to be an *optimal encoding* (for C) (with respect to the weight function w) if and only if $S_w(\eta)$ is Gale optimal in $S_w(\mathscr{E})$.

DEFINITION 2(b). If C is linear and $\mathscr{G}$ is the set of all generator matrices for C, we will say that $G \in \mathscr{G}$ is an *optimal generator matrix* (for C) (with respect to the weight function w) if and only if $S_w(G)$ is Gale optimal in $S_w(\mathscr{G})$.

To show η to be an optimal encoding, it will suffice to show that for all $\xi \in \mathscr{E}$

$$S_w(\eta)^* \geqslant S_w(\xi)^*.$$

To show that G is an optimal generator matrix it will suffice to show that for all $A \in \mathscr{G}$

$$S_w(G)^* \geqslant S_w(A)^*.$$

It is evident that $\eta \in \mathscr{E}$ is an optimal encoding of the code, C, if and only if $S(\eta)^*$ is the greatest element of the finite partially ordered set $S(\mathscr{E})^* = \{S(\xi)^*: \xi \in \mathscr{E}\}$. Since a partially ordered set has at most one greatest element, it follows that if $\eta \in \mathscr{E}$ is an optimal encoding and $\xi \in \mathscr{E}$, then ξ is an optimal encoding if and only if $S(\eta)^* = S(\xi)^*$. If C is linear, we may apply a similar argument to show that given an optimal generator matrix $G \in \mathscr{G}$ and any other generator matrix $A \in \mathscr{G}$, then A is an optimal generator matrix if and only if $S(A)^* = S(G)^*$. Suppose C is a linear code, η is an optimal encoding of C, and G is an optimal generator matrix for C; then $S(G)^* \leqslant S(\eta)^*$ since $S(\mathscr{G})^* \subseteq S(\mathscr{E})^*$. However, we will find that equality may not hold.

When C is linear we will be able to find optimal generator matrices for C. We will then give conditions on C which will guarantee that every optimal generator matrix for C induces an optimal encoding of C. From our results, we will then deduce that every linear code has an optimal encoding. However, we will be able to give an example of a linear code, $C \subseteq F^n$, whose optimal encodings all fail to be linear over F. The best one can do is to guarantee the existence of an optimal endoding which is linear over the prime subfield of F. We now proceed with the first step of this development.

Given a linear code C and a weight function w, for each $\rho \in \mathbb{R} \cup \{\infty\}$, denote

$$C_{w^\rho} \triangleq \{c \in C: w(c) < \rho\}. \tag{9}$$

Now suppose $\rho \in \mathbb{R}$ and G is a generator matrix for the linear code C. Clearly $C_{w^\rho} \subseteq \langle G_r \rangle$. Further, if X, $Y \subseteq G_r$ such that $C_{w^\rho} \subseteq \langle X \rangle$ and $C_{w^\rho} \subseteq \langle Y \rangle$, then

$$C_{w^\rho} \subseteq \langle X \rangle \cap \langle Y \rangle = \langle X \cap Y \rangle$$

follows since G_r is linearly independent. Thus, there exists a smallest subset, G_{w^ρ}, of G_r such that $C_{w^\rho} \subseteq \langle G_{w^\rho} \rangle$: i.e., $\langle C_{w^\rho} \rangle \subseteq \langle G_{w^\rho} \rangle$ where

$$G_{w^\rho} \triangleq \bigcap \{X \subseteq G_r : C_{w^\rho} \subseteq \langle X \rangle\}. \tag{10}$$

When no confusion will result, the subscripted references to weight functions may be omitted. We relate the sets defined in (9) and (10) to the separation vector associated with G in the following lemma.

LEMMA 1. *Given a $k \times n$ generator matrix, G, for all $i \in \{1,..., k\}$, $\rho \in \mathbb{R}$*

$$S(G)_i \geqslant \rho \qquad \textit{iff} \quad G_{i\cdot} \notin G^\rho.$$

Proof. Fix $i \in \{1,..., k\}$. If $G_{i\cdot} \notin G^\rho$, then $C^\rho \subseteq \langle G^\rho \rangle \subseteq \langle G_r \backslash G_{i\cdot} \rangle$ implies

$$S(G)_i = w[C \backslash \langle G_r \backslash G_{i\cdot} \rangle] \geqslant w[C \backslash C^\rho] \geqslant \rho.$$

For the reverse implication, suppose $G_{i\cdot} \in G^\rho$. Assume, for a moment, that $C^\rho \subseteq \langle G_r \backslash G_{i\cdot} \rangle$. Since $C^\rho \subseteq \langle G^\rho \rangle$ this would imply that

$$C^\rho \subseteq \langle G^\rho \rangle \cap \langle G_r \backslash G_{i\cdot} \rangle = \langle G^\rho \cap (G_r \backslash G_{i\cdot}) \rangle = \langle G^\rho \backslash G_{i\cdot} \rangle.$$

However, this cannot be since G^ρ is the minimal subset of G_r satisfying $C^\rho \subseteq \langle G^\rho \rangle$. Thus, the assumption is in error, and $C^\rho \nsubseteq \langle G_r \backslash G_{i\cdot} \rangle$ lest the assumption be implied. We have shown

$$C^\rho \cap (C \backslash \langle G_r \backslash G_{i\cdot} \rangle) \neq \Phi.$$

Hence,

$$S(G)_i = w[C \backslash \langle G_r \backslash G_{i\cdot} \rangle] < \rho. \qquad \text{Q.E.D.}$$

We now employ this lemma to obtain a basic result for linear codes. Note that $w(C) = \{w(c): c \in C\} \neq w[C]$.

THEOREM 4. *A necessary and sufficient condition for a generator matrix, G, to be an optimal generator matrix for its row space, C, is that for each $\rho \in w(C)$ there exists $X \subseteq G_r$ such that $\langle C^\rho \rangle = \langle X \rangle$.*

Sufficiency. Let G satisfy the condition; i.e., for each $\rho \in w(C)$ there exists $X \subseteq G_r$ such that $\langle C^\rho \rangle = \langle X \rangle$. Suppose G is not an optimal generator matrix; then there must exist a generator matrix A for C such that $S(G)^* \not\geqslant S(A)^*$. By appealing to Corollary 1, we may assume that $S(A) = S(A)^*$. Now there exists a smallest integer, i, having the property that $S(G)^* < S(A)_i$. Set $\rho \triangleq S(A)_i$. Let $X \subseteq G_r$ such that $\langle C^\rho \rangle = \langle X \rangle$. Since $C^\rho \subseteq \langle X \rangle$, $G^\rho \subseteq X$. Thus, $\langle G^\rho \rangle \subseteq \langle X \rangle = \langle C^\rho \rangle$; i.e., $\langle G^\rho \rangle = \langle C^\rho \rangle$. According to Proposition 2(d) it must be true that $C^\rho \subseteq \langle A_1, \ldots, A_{(i-1)} \rangle$ and hence that $\langle C^\rho \rangle \subseteq \langle A_1, \ldots, A_{(i-1)} \rangle$. Let $\phi: \{1, \ldots, k\} \to \{1, \ldots, k\}$ be a permutation such that $S(G)^* = S(G) \circ \phi$. Applying (7)

$$S(G)_1^* \leqslant \cdots \leqslant S(G)_i^* < S(A)_i = \rho,$$

and it follows that

$$S(G)_{\phi(1)} \leqslant \cdots \leqslant S(G)_{\phi(i)} < S(A)_i = \rho.$$

According to Lemma 1 this implies that $G_{\phi(1)}, \ldots, G_{\phi(i)} \in G^\rho$. Thus, we have shown

$$\langle G_{\phi(1)}, \ldots, G_{\phi(i)} \rangle \subseteq \langle C^\rho \rangle \subseteq \langle A_1, \ldots, A_{(i-1)} \rangle,$$

and it follows that

$$i \leqslant \dim\langle C^\rho \rangle \leqslant i - 1.$$

A contradiction has been reached and hence our assumption was false; i.e., G is an optimal generator matrix.

Necessity. Suppose G is an optimal generator matrix for C and that $\rho \in w(C)$. Let $\{a_1, \ldots, a_l\}$ be a basis for $\langle C^\rho \rangle$ where $l \triangleq \dim\langle C^\rho \rangle$. Extend this to a basis, $\{a_1, \ldots, a_l, a_{l+1}, \ldots, a_k\}$ of C. Define a generator matrix, A, for C by

$$A \triangleq \begin{bmatrix} a_1 \\ \vdots \\ a_l \\ a_{l+1} \\ \vdots \\ a_k \end{bmatrix}.$$

From Lemma 1 we see that, since $a_{l+1}, \ldots, a_k \notin A^\rho = \{a_1, \ldots, a_l\}$, $S(A)_{l+1}, \ldots, S(A)_k \geqslant \rho$. It is easily seen that this implies $S(A)_{l+1}^*, \ldots, S(A)_k^* \geqslant \rho$. Since $S(G)^* \geqslant S(A)^*$, it follows that $S(G)_{l+1}^*, \ldots, S(G)_k^* \geqslant \rho$. Let $\phi: \{1, \ldots, k\} \to \{1, \ldots, k\}$ be a permutation such that $S(G)^* = S(G) \circ \phi$. Now, it follows that $S(G)_{\phi(l+1)}, \ldots, S(G)_{\phi(k)} \geqslant \rho$. Hence, by Lemma 1, $G_{\phi(l+1)}, \ldots, G_{\phi(k)} \notin G^\rho$. Thus,

$$G^\rho \subseteq G_r \backslash \{G_{\phi(l+1)}, \ldots, G_{\phi(k)}\} = \{G_{\phi(1)}, \ldots, G_{\phi(l)}\}.$$

Now, $\langle C^\rho \rangle \subseteq \langle G^\rho \rangle \subseteq \langle G_{\phi(1)}, \ldots, G_{\phi(l)} \rangle$; and since $\dim\langle C^\rho \rangle = l = \dim\langle G_{\phi(1)}, \ldots, G_{\phi(l)} \rangle$, it follows immediately that

$$\langle C^\rho \rangle = \langle G^\rho \rangle. \qquad \text{Q.E.D.}$$

If $\rho, \lambda \in w(C)$ and $\rho \leqslant \lambda$, then $\langle C^\rho \rangle \subseteq \langle C^\lambda \rangle$. Using this observation an optimal generator matrix for a linear code C could be constructed in the following manner. Start with Φ as a basis for $\langle C^0 \rangle$. Having found a basis for $\langle C^\rho \rangle$ ($\rho \in w(C)$), if $\langle C^\rho \rangle \neq C$, then set $\lambda = w[\{c \in C: w(c) > \rho\}]$ and extend the basis of $\langle C^\rho \rangle$ to a basis for $\langle C^\lambda \rangle$. Repetition of the last step will eventually result in a basis of C ($\lambda = +\infty$ is possible). Form a generator matrix with the members of the basis just found as its rows; any generator matrix so formed will be an optimal generator matrix. Later, we will use a different procedure to construct optimal generator matrices, and we will state an existence theorem at that time. For now, we will be content with the next corollary.

The following corollary makes the construction of optimal generator matrices for cyclic codes easy when certain conditions are met. Let an (n, k) cyclic code, C, be given with n and q relatively prime and parity check polynomial, $h(X)$, having the complete factoring $h(X) = h_1(X) \cdot h_2(X) \cdots h_l(X)$ over F. The minimal ideals (e.g., see Section 1.7 of Blake and Mullin, 1975) of C are themselves cyclic codes with generator polynomials $(X^n - 1)/h_i(X)$, $i = 1, \ldots, l$. The generator matrices of these minimal ideals will be used to form an optimal generator matrix for C. However, the reader should note that the theorem will not say which message positions receive extra protection or how much protection any position receives.

COROLLARY 2. *Let C be a cyclic (n, k) code over $F \triangleq GF(q)$ with n and q relatively prime. Suppose w satisfies*

$$w(c_1, \ldots, c_n) = w(c_n, c_1, \ldots, c_{n-1})$$

for all code words $c \in C$; and let $M_1, \ldots, M_l$ be generator matrices for the minimal ideals (of the algebra of polynomials over F modulo $X^n - 1$) contained in C. Then,

$$G \triangleq \begin{bmatrix} M_1 \\ \vdots \\ M_l \end{bmatrix}$$

is an optimal generator matrix for C.

Proof. Given any $\rho \in w(C)$, $\langle C^\rho \rangle$ is easily shown to be a cyclic code since it has a set of generators, $C^\rho \triangleq \{c \in C: w(c) < \rho\}$ which is closed under cyclic rotation. Hence, $\langle C^\rho \rangle$ is a direct sum of the minimal ideals contained in it. Since $\langle C^\rho \rangle \subseteq C$ these minimal ideals are also contained in C. Define a subset, X, of the rows of G by

$$X \triangleq \{G_{i\cdot}: G_{i\cdot} \in \langle C^\rho \rangle\}.$$

Now, we complete the proof by showing that

$$\langle C^{\rho} \rangle = \langle X \rangle.$$

Clearly $\langle X \rangle \subseteq \langle C^{\rho} \rangle$. On the other hand if $\mathcal{M}$ is any minimal ideal satisfying $\mathcal{M} \subseteq \langle C^{\rho} \rangle$, then there exists an integer i, $1 \leqslant i \leqslant l$, such that M_i is a generator matrix for $\mathcal{M}$. $(M_i)_r \subseteq \mathcal{M} \subseteq \langle C^{\rho} \rangle$ so $(M_i)_r \subseteq X$, which, in turn, implies $\mathcal{M} \subseteq \langle X \rangle$. Since $\langle X \rangle$ contains all the minimal ideals of $\langle C^{\rho} \rangle$, it follows that $\langle C^{\rho} \rangle \subseteq \langle X \rangle$. Q.E.D.

In particular this result holds when w represents either the Hamming weight function or the Lee weight function.

We have given a necessary and sufficient condition which can be employed to determine which generator matrices of a linear code are optimal generator matrices. We now employ this result in developing a sufficient condition which will make the existence of optimal generator matrices for all linear codes even more apparent.

First, a definition is needed. Definition 3 and Definition 4, given later, are an adaptation of the terminology of Massey *et al.* (1973).

DEFINITION 3. A generator matrix G of a linear (n, k) code will be said to be *monotonically weight retaining* (with respect to the weight function w) if and only if for $i = 1, \ldots, k$

$$w(G_{i\cdot}) = w[C \backslash \langle G_{1\cdot}, \ldots, G_{(i-1)\cdot} \rangle]. \tag{11}$$

The terminology "monotonically weight retaining" is suggestive. Witness the following observations. Let C be the row space of a $k \times n$ monotonically weight retaining generator matrix G. Consider any linear combination, c, of the rows of G,

$$c \triangleq \alpha_1 G_{1\cdot} + \cdots + \alpha_k G_{k\cdot},$$

where $\alpha_1, \ldots, \alpha_k \in F$. If $\alpha_l \neq 0$ for some l where $1 \leqslant l \leqslant k$, then $c \notin \langle G_{1\cdot}, \ldots, G_{(l-1)\cdot} \rangle$, and hence

$$w(c) \geqslant w[C \backslash \langle G_{1\cdot}, \ldots, G_{(l-1)\cdot} \rangle] = w(G_{l\cdot}).$$

Thus, c retains (at least) the weight of G_l. Now, for each integer, i, satisfying $1 \leqslant i < k$, we have $C \backslash \langle G_{1\cdot}, \ldots, G_{(i-1)\cdot} \rangle \supseteq C \backslash \langle G_{1\cdot}, \ldots, G_{i\cdot} \rangle$ and hence

$$w(G_{i\cdot}) = w[C \backslash \langle G_{1\cdot}, \ldots, G_{(i-1)\cdot} \rangle] \leqslant w[C \backslash \langle G_{1\cdot}, \ldots, G_{i\cdot} \rangle] = w(G_{(i+1)\cdot}).$$

Thus $w_r(G)$ is nondecreasing. We have shown the following lemma.

LEMMA 2. *If G is a monotonically weight retaining generator matrix, then $w_r(G)$ is nondecreasing.*

We characterize those generator matrices we shall show to be optimal generator matrices in the following theorem.

THEOREM 5. *Let C be a linear (n, k) code and let G be a generator matrix for C. Then, the following statements are equivalent.*

(i) *Given any other generator matrix, A, for C*

$$w(G_{1\cdot}) + \cdots + w(G_{k\cdot}) \leqslant w(A_{1\cdot}) + \cdots + w(A_{k\cdot}).$$

(ii) $w_r(G) = S(G)$.

(iii) *G may be obtained from a monotonically weight retaining generator matrix, A, for C by permuting its rows.*

(iv) *If A is any generator matrix for C obtained from G by permuting its rows, then $w_r(A)$ is nondecreasing if and only if A is monotonically weight retaining.*

Proof. We shall show

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iv)} \Rightarrow \text{(iii)} \Rightarrow \text{(i)}.$$

(i) *implies* (ii): We shall show the contrapositive. Suppose that G is a generator matrix for C such that (ii) does not hold for G. There exists an integer, $i \in \{1,..., k\}$ such that $w(G_{i\cdot}) \neq S(G)_i$. Since it is always the case that $S(G)_i = w[C\backslash\langle G_r\backslash G_{i\cdot}\rangle] \leqslant w(G_{i\cdot})$, it follows that $S(G)_i < w(G_{i\cdot})$. Let $v \in C\backslash\langle G_r\backslash G_{i\cdot}\rangle$ such that $w(v) = S(G)_i$. It is clear that

$$\begin{aligned} &w(G_{1\cdot}) + \cdots + w(G_{k\cdot}) \\ &\quad > w(G_{1\cdot}) + \cdots + w(G_{(i-1)\cdot}) + w(v) + w(G_{(i+1)\cdot}) + \cdots + w(G_{k\cdot}). \end{aligned}$$

But $v \in C$, and it is easy to verify that $G_{1\cdot}, ..., G_{(i-1)\cdot}, v, G_{(i+1)\cdot}, ..., G_{k\cdot}$ are linearly independent. Thus,

$$\begin{bmatrix} G_{1\cdot} \\ \vdots \\ G_{(i-1)\cdot} \\ v \\ G_{(i+1)\cdot} \\ \vdots \\ G_{k\cdot} \end{bmatrix}$$

is a generator matrix for C. We have shown that G does not satisfy (i).

(ii) *implies* (iv): Suppose A is a matrix obtained from G by permuting the rows of G, where G satisfies (ii). Then, that A satisfies (ii) follows easily from Corollary 1. Now, suppose that $w_r(A)$ is nondecreasing. It is clear that $S(A)_1 \leqslant \cdots \leqslant S(A)_k$, and hence (a) = (d) of Proposition 2 applies to yield

$$w(A_{i\cdot}) = S(A)_i = w[C\backslash\langle A_{1\cdot}, ..., A_{(i-1)\cdot}\rangle]$$

for $i = 1,..., k$. Thus, A is monotonically weight retaining. We have shown that whenever $w_r(A)$ is nondecreasing, A is monotonically weight retaining. The reverse implication does not depend on (ii) and was shown as Lemma 2.

(iv) *implies* (iii): Let A be a matrix obtained from G by permuting the rows of G in a manner which will ensure that $w_r(A)$ is nondecreasing. By (iv) A will be monotonically weight retaining. G may be obtained from A by permuting the rows of A with the inverse of the permutation used to obtain A from G.

(iii) *implies* (i): Suppose G satisfies (iii). G will satisfy (i) if and only if a given matrix differing from G by only a permutation of its rows satisfies (i). It follows that we may assume, without loss of generality, that G itself is monotonically weight retaining. The implication is shown by contradiction. Suppose that there exists a generator matrix, A, for C such that

$$w(G_{1\cdot}) + \cdots + w(G_{k\cdot}) \nleqslant w(A_{1\cdot}) + \cdots + w(A_{k\cdot}).$$

We may assume that $w_r(A)$ is nondecreasing, else we could permute the rows of A to make this true. Let i be the smallest integer, $1 \leqslant i \leqslant k$, such that $w(G_{i\cdot}) > w(A_{i\cdot})$. Set $\rho \triangleq w(G_{i\cdot})$ since

$$w(A_{1\cdot}) \leqslant \cdots \leqslant w(A_{i\cdot}) < w(G_{i\cdot}) = \rho;$$

it follows that $\{A_{1\cdot},..., A_{i\cdot}\} \subseteq \langle C^\rho \rangle$. On the other hand from (11)

$$\rho = w(G_{i\cdot}) = w[C \backslash \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle]. \tag{12}$$

Now $C^\rho \subseteq C$ so

$$C^\rho \subseteq \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle \cup (C \backslash \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle).$$

We claim that $C^\rho \subseteq \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle$ lest $C^\rho \cap (C \backslash \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle) \neq \Phi$. To see why this is impossible let $c \in C^\rho \cap (C \backslash \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle)$, then $w(c) < \rho$ by (9) and

$$w[C \backslash \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle] \leqslant w(c) < \rho \tag{13}$$

since $c \in C \backslash \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle$. But (12) and (13) cannot both be true. Thus, $C^\rho \subseteq \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle$, and we have shown

$$\langle A_{1\cdot},..., A_{i\cdot} \rangle \subseteq \langle C^\rho \rangle \subseteq \langle G_{1\cdot},..., G_{(i-1)\cdot} \rangle.$$

A contradiction has been found: $i \leqslant \dim \langle C^\rho \rangle \leqslant i - 1$. Q.E.D.

In view of statements (ii) and (iii) of the preceding theorem, the following definition seems warranted.

DEFINITION 4. A generator matrix will be said to be *weight retaining* if and only if it satisfies the equivalent statements of Theorem 5.

Our sufficient condition for optimality of a generator matrix is that it be weight retaining. We formulate this as a corollary to Theorem 4.

COROLLARY 3. *Every weight retaining generator matrix is an optimal generator matrix for its row space.*

Proof. Suppose G is a weight retaining generator matrix for a linear (n, k) code, C. We desire to show that G is an optimal generator matrix. By statement (iii) of Theorem 5, there exists another matrix, A, such that the rows of A are a permutation of the rows of G and A is monotonically weight retaining. From Corollary 1 we know that $S(A)^* = S(G)^*$ and hence that A is an optimal generator matrix if and only if G is an optimal matrix. We will complete the proof by applying Theorem 5 to show that A is an optimal generator matrix. Let $\rho \in w(C)$ and set

$$X \triangleq \{A_{i.}: w(A_{i.}) < \rho\}.$$

Clearly $\langle X \rangle \subseteq \langle C^\rho \rangle$. For the reverse inclusion we need only show that $C^\rho \subseteq \langle X \rangle$. Since A is monotonically weight retaining, $w_r(A)$ is nondecreasing by Lemma 2. Thus, there exists an integer, $l \in \{1,..., k\}$, such that $X = \{A_{1.},..., A_{l.}\}$. If $l = k$ then $\langle X \rangle = C$, a trivial case. If $l < k$, then

$$w[C \backslash \langle X \rangle] = w(A_{(l+1).}) \geqslant \rho$$

and Eq. (9) imply the desired inclusion. Q.E.D.

Several remarks are in order. Corollary 2 gives an easy method for finding optimal generator matrices for cyclic codes under certain conditions. Unfortunately, Corollary 2 gives little information about the separation vectors associated with these generator matrices. However, when a generator matrix is guaranteed to be an optimal generator matrix by means of the preceding corollary, its separation vector is easily calculated by employing statement (ii) of Theorem 5. When w represents either the Hamming weight function or the Lee weight function, statement (i) of Theorem 5 implies that the weight retaining matrices for a particular linear code are precisely those generator matrices for that code whose weight is as small as possible. In this case the preceding corollary guarantees that any generator matrix whose weight is as small as possible, for a fixed linear code, is an optimal generator matrix.

If A and B are weight retaining and have the same row space, then $w_r(A)^* = w_r(B)^*$ follows from statement (ii) of Theorem 5 and the last corollary. In fact, suppose C is a linear (n, k) code and $\mathscr{G}$ denotes the set of all generator matrices for C. Then, it is not difficult to show that G satisfies statement (i) of Theorem 5 if an only if $w_r(G)^*$ is the least element of the finite partially ordered set $w_r(\mathscr{G})^*$.

Statement (i) of Theorem 5 used in conjunction with Corollary 3 provides an

easy proof of the existence of optimal generator matrices. Among the finitely many generator matrices of a fixed linear (n, k) code there must be at least one, call it G, such that $w(G_{1.}) + \cdots + w(G_{k.})$ is at a minimum. By statement (i), G is weight retaining; and by Corollary 3, it is an optimal generator matrix. Monotonically weight retaining generator matrices are weight retaining by statement (iii) of Theorem 5, and hence, they are optimal generator matrices by Corollary 3. We will give a procedure for obtaining a monotonically weight retaining generator matrix of any linear code. Thus, we will have an alternate, and more constructive, existence proof. The procedure might be termed a "greedy" algorithm (e.g., see pp. 267, 275–277 of Lawler, 1976).

Procedure. Given a linear (n, k) code, C, to find a $k \times n$ optimal monotonically weight retaining generator matrix, G, for C:

1. Set $i = 1$.
2. Choose $v \in C \backslash \langle G_{1.}, \ldots, G_{(i-1).} \rangle$ such that
$$w(v) = w[C \backslash \langle G_{1.}, \ldots, G_{(i-1).} \rangle].$$
3. Set $G_{i.} = v$.
4. If $i < k$ then replace i by $i + 1$ and go to step 2, else stop.

One may also use the first existence proof to guarantee the existence of optimal generator matrices that are monotonically weight retaining by applying statement (iii) of Theorem 5. In either case, we have shown the following theorem.

THEOREM 6. *Every linear code has at least one optimal generator matrix with respect to any given weight function. In fact, a monotonically weight retaining optimal generator matrix may be found for any fixed linear code.*

In the next theorem we deal with the more general problem of finding optimal encodings for linear codes. When proving the result it will be convenient to have the following lemma at our disposal.

LEMMA 3. *Let* X, $Y \subseteq F^n \triangleq GF(q)^n$ *and suppose* $\overline{0_n} \in Y$. *If either of the conditions,*

(a) *for all* $\alpha \in F$, $\quad Y + \alpha X \subseteq Y$,

or

(b) q *is prime and* $\quad Y + X \subseteq Y$

holds, then $\langle X \rangle \subseteq Y$.

Proof. We first show case (a). Suppose that for an integer, $l \geqslant 0$, $\alpha_1 x_1 + \cdots + \alpha_l x_l \in Y$ whenever $\alpha_1, \ldots, \alpha_l \in F$ and $x_1, \ldots, x_l \in X$. If $\alpha_1, \ldots, \alpha_{l+1} \in F$ and $x_1, \ldots, x_{l+1} \in X$, then $\alpha_1 x_1 + \cdots + \alpha_{l+1} x_{l+1} \in Y + \alpha_{l+1} X \subseteq Y$. The inductive hypoth-

esis is true for $l = 0$ since $\overline{0_n} \in Y$. Now, we show case (b) follows from case (a). If q is prime and $\alpha \in F$ then $\alpha = \sum_1^i 1$ for some positive positive integer i. $Y + (\sum_1^i 1)X \subseteq Y$ for all positive integers i follows from $Y + X \subseteq Y$ by an induction. Q.E.D.

THEOREM 7. *Let C be a linear (n, k) code. With respect to a given weight function w, any optimal generator matrix for C induces an optimal encoding of C provided that either of the conditions*

(a) $w(\alpha c) = w(c)$ *for all nonzero* $\alpha \in F$ *and all* $c \in C$

or

(b) q *is prime*

holds.

Proof. It will suffice to show that a particular optimal generator matrix for C induces an optimal encoding. Applying Theorem 6, let G be a monotonically weight retaining optimal generator matrix for C. We desire to show that, under certain conditions, $S(G)^* \geqslant S(\eta)^*$ for any encoding, η, of C. In view of Proposition 1, we may assume η is an encoding of C satisfying $S(\eta) = S(\eta)^*$ and $\eta(\overline{0_k}) = \overline{0_n}$. We will show $S(G) \geqslant S(\eta)$ and

$$S(G)^* = w_r(G)^* = w_r(G) = S(G) \geqslant S(\eta) = S(\eta)^*$$

will follow from this and Lemma 2.

We suppose that $S(G) \not\geqslant S(\eta)$ and search for a contradiction. In this case there exists a smallest positive integer, $l \in \{1, \ldots, k\}$, such that $S(G)_l < S(\eta)_l$. Set $\rho \triangleq S(\eta)_l$; and set

$$X \triangleq C^\rho, \qquad Y \triangleq \{\eta(m) : m_i = 0 \ \forall i \geqslant l\}.$$

We now show that $Y + X \subseteq Y$. Let $y \in Y$ and $x \in X$ and suppose $y + x \notin Y$. Since $y + x \in C$, there exists $\hat{m} \in F^k$ satisfying $\eta(\hat{m}) = y + x$. For some $i \geqslant l$, $\hat{m}_i \neq 0$ since $y + x \notin Y$. The contradiction is evident;

$$S(\eta)_i = w[\{\eta(m) - \eta(m') : m_i \neq m_i'\}] \leqslant w(y - (y + x)) = w(x) < S(\eta)_l \leqslant S(\eta)_i.$$

Thus, $y + x \in Y$ and $Y + X \subseteq Y$.

We now wish to show $\langle X \rangle \subseteq Y$. $\overline{0_n} \in Y$ since $\eta(\overline{0_k}) = \overline{0_n}$. We argue cases (a) and (b) separately.

(a) In this case $\alpha X \subseteq X$ for all $\alpha \in F$ ($\alpha = 0$ is trivial).

Hence, $Y + \alpha X \subseteq Y + X \subseteq Y$ for all $\alpha \in F$. $\langle X \rangle \subseteq Y$ follows from Lemma 3(a).

(b) We already have $Y + X \subseteq Y$ and can apply Lemma 3(b) to obtain $\langle X \rangle \subseteq Y$.

Since $w(G_{1\cdot}) \leqslant \cdots \leqslant w(G_{l\cdot}) = S(G)_l < S(\eta)_l = \rho$ it follows that $\{G_{1\cdot}, \ldots, G_{l\cdot}\} \subseteq X$. We have shown

$$\langle G_{1\cdot}, \ldots, G_{l\cdot} \rangle \subseteq \langle X \rangle \subseteq Y.$$

Taking cardinalities, we have $q^l \leqslant |\langle X \rangle| \leqslant q^{l-1}$, a contradiction. Q.E.D.

Condition (a) of Theorem 7 holds for the Hamming weight function, while condition (b) holds for the Lee weight function which is not defined for nonprime fields. We give an example to show that the conditions (a) and (b) cannot be removed.

TABLE II

α	$\hat{w}(\alpha)$	$\hat{\eta}(\alpha)$
0	0	(0, 0)
1	3	(X, X)
X	1	(X, 0)
$X+1$	2	(0, X)

EXAMPLE 1. We represent the elements of $GF(4)$ by their canonical representatives 0, 1, X, and $X+1$ taken from $GF(2)[X]/(X^2+X+1)$. Let $C \triangleq GF(4)^2$. We define functions $\hat{w}$: $GF(4) \to \mathbb{R}$ and $\hat{\eta}$: $GF(4) \to C$ by means of Table II. A weight function, w: $C \to \mathbb{R}$, is defined by

$$w(\alpha, \beta) \triangleq \hat{w}(\alpha) + \hat{w}(\beta).$$

It may be verified directly that the encoding, η: $GF(4)^2 \to C$, defined by

$$\eta(\alpha, \beta) \triangleq \hat{\eta}(\alpha) + X \cdot \hat{\eta}(\beta) \qquad \text{(components mod } X^2 + X + 1)$$

has associated separation vector $S_w(\eta) = (1, 2)$. On the other hand, there is a generator matrix, G, of C defined by

$$G \triangleq \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix}$$

which is w-weight retaining (and hence an optimal generator matrix), and yet $S_w(G) = (1, 1)$. Thus, no generator matrix for C induces an optimal encoding for C with respect to the weight function w.

When either the Hamming metric or the Lee metric is employed, Theorem 7 guarantees that the minimum weight generator matrices, which were found to be optimal generator matrices earlier, induce optimal encodings. For the Hamming metric a generator matrix of minimum weight is merely a generator matrix which has as few nonzero entries as possible, i.e., one that is as sparse as possible. We now state this formally.

COROLLARY 4. *Let C be linear (n, k) code.*

(a) *If the Hamming metric is employed, then any generator matrix, G, for C which is as sparse as possible induces an optimal encoding of C and $S_h(G) = h_r(G)$.*

(b) *If F is a prime field and the Lee metric is employed, then any generator matrix, G, for C which is of minimal Lee weight induces an optimal encoding of C and $S_l(G) = l_r(G)$ where l denotes the Lee weight function.*

Let $w: F^n \to \mathbb{R}$ be a weight function, and suppose that an (n, k) code $C \subseteq F^n$ is linear over the prime subfield $P \triangleq GF(p)$ of $F \triangleq GF(p^e)$. Identify F^n and P^{ne} as isomorphic vector spaces over P. $w: P^{ne} \to \mathbb{R}$ remains a weight function. By Theorems 6 and 7(b) there is an optimal (P-linear) encoding $\eta: P^{ke} \to C$ for the (ne, ke) P-linear code C. Assume without loss of generality that $S(\eta) = S(\eta)^*$.

Now, let $\alpha: F \to P^e$ be any bijection, and define a bijection $\gamma. F^k \to P^{ke}$ by

$$\gamma(m) \triangleq (\alpha(m_1),..., \alpha(m_k))$$

for all $m \in F^k$. Define $\hat{\eta}: F^k \to C$ by $\hat{\eta} \triangleq \eta \circ \gamma$. The separation vector of $\hat{\eta}$ is given by

$$S(\hat{\eta})_i = \min\{S(\eta)_j: (i-1)e + 1 \leqslant j \leqslant ie\}, \qquad i = 1,..., k.$$

$S(\hat{\eta}) = S(\hat{\eta})^*$ is inherited. We claim that $\hat{\eta}$ is an optimal encoding of the (n, k) code C.

In fact suppose $\hat{\xi}: F^k \to C$ is an encoding such that $S(\hat{\eta}) \not\geqslant S(\hat{\xi})^* = S(\hat{\xi})$. Define $\xi: P^{ke} \to C$ by

$$\xi \triangleq \hat{\xi} \circ \gamma^{-1} \qquad (\hat{\xi} = \xi \circ \gamma).$$

Let $l \in \{1,..., k\}$ be the smallest integer such that $S(\hat{\eta})_l < S(\hat{\xi})_l$. Denote $t \triangleq (l-1)e + 1$. Let $j \in \{t,..., ke\}$. There exists a unique $i \in \{l,..., k\}$ such that $(i-1)e + 1 \leqslant j \leqslant ie$. Now,

$$S(\xi)_j \geqslant \min\{S(\xi)_h: (i-1)e + 1 \leqslant h \leqslant ie\} = S(\hat{\xi})_i \geqslant S(\hat{\xi})_l.$$

Since $S(\xi)_j \geqslant S(\hat{\xi})_l$ for all $j \geqslant t$, it follows that $S(\xi)_j^* \geqslant S(\hat{\xi})_l$ for all $j \geqslant t$; and in particular $S(\xi)_t^* \geqslant S(\hat{\xi})_l$. Thus,

$$S(\eta)_t^* = S(\eta)_t = \min\{S(\eta)_t,..., S(\eta)_{le}\} = S(\hat{\eta})_l < S(\hat{\xi})_l \leqslant S(\xi)_t^*.$$

We have shown that $\eta: P^{ke} \to C$ is not optimal, a contradiction. $\hat{\eta}: F^k \to C$ must be optimal.

We have shown that every code which is linear over its prime subfield has an optimal encoding. Further some optimal encoding is linear over the prime subfield. We shall be content to record the following:

COROLLARY 5. *Given a fixed linear code and weight function, there is an optimal encoding for the linear code which is linear over the prime subfield.*

The term "linear" cannot, in general, be deleted from the preceding corollary.

EXAMPLE 2. The code, C, shown in Table III is an example of a nonlinear code with no optimal encodings.

TABLE III

i	c_i	$\eta_1^{-1}(c_i)$	$\eta_2^{-1}(c_i)$
1	0000000	0000	0000
2	1000000	0010	0100
3	1100000	0100	1000
4	1111000	0110	0001
5	1011000	1000	0101
6	0011000	1010	1001
7	0001100	1100	0010
8	1001100	1110	0110
9	0100011	0001	1100
10	1100111	0011	1101
11	1101011	0101	1010
12	0101111	0111	1110
13	1110011	1001	0011
14	0110111	1011	0111
15	0111011	1101	1011
16	1111111	1111	1111

The encodings η_1 and η_2 given by Table III have separation vectors $S_h(\eta_1) = (1, 1, 1, 3)$ and $S_h(\eta_2) = (1, 1, 2, 2)$, respectively. Let $e_1, \ldots, e_4$ be the usual basis of $GF(2)^4$ and set $X = \{c_1, c_2, c_3\}$, $Y = \{c_4, c_5, c_6\}$, $Z = \{c_7, c_8\}$. Suppose η is an encoding of C such that $S_h(\eta)_3, S_h(\eta)_4 \geqslant 2$. Then, one can show that $\{X, Y, Z, \{c_9\}, \ldots, \{c_{16}\}\}$ is a partition of C which is a refinement of the partition $P \triangleq \{\eta(\langle e_1, e_2\rangle + v): v \in \langle e_3, e_4\rangle\}$. It is clear that X, Y, and Z must be contained in distinct members of P. From this and $h(X - Y) = h(Y - Z) = h(Z - X) = 2$, one can deduce that $S_h(\eta)_3 = S_h(\eta)_4 = 2$. Proposition 1 and the preceding argument may be used to show that no encoding, η, of C satisfies $S_h(\eta)^* \geqslant (1, 1, 2, 3)$.

With the next theorem, we complete our study of product codes. Again, we restrict our attention to the Hamming weight function. The proof of the theorem will rely heavily on the following lemma.

LEMMA 4. *Let $x, y \in \mathbb{R}^k$ and $u, v \in \mathbb{R}^l$ such that $\overline{0_k} \leqslant x^* \leqslant y^*$, $\overline{0_l} \leqslant u^* \leqslant v^*$, $y \neq \overline{0_k}$, and $v \neq \overline{0_l}$, where k, l are positive integers. Under these conditions*

$$(x \otimes_{\mathbb{R}} u)^* = (y \otimes_{\mathbb{R}} v)^* \qquad \textit{iff } x^* = y^* \text{ and } u^* = x^*.$$

Proof. For the sufficiency, $x^* = y^*$ and $u^* = v^*$ imply there exist permutation matrices P and Q such that $x = yP$ and $u = vQ$. We see that

$$x \otimes_{\mathbb{R}} u = yP \otimes_{\mathbb{R}} vQ = (y \otimes_{\mathbb{R}} v)(P \otimes_{\mathbb{R}} Q).$$

$P \otimes_{\mathbb{R}} Q$ is a permutation matrix, so $(x \otimes_{\mathbb{R}} u)^* = (y \otimes_{\mathbb{R}} v)^*$.

For the necessity, assume $x^* \neq y^*$ the case $u^* \neq v^*$ being similar. It follows easily that $0 \leqslant x\overline{1_k}^t < y\overline{1_k}^t$ and that $0 \leqslant u\overline{1_l}^t \leqslant v\overline{1_l}^t$ where $\overline{1_k}$, $\overline{1_l}$ are vectors of all ones. This implies

$$x\overline{1_k}^t \cdot u\overline{1_l}^t < y\overline{1_k}^t \cdot v\overline{1_l}^t.$$

Now,

$$\begin{aligned}(x \otimes_{\mathbb{R}} u)^* \cdot \overline{1_{kl}}^t &= (x \otimes_{\mathbb{R}} u) \cdot \overline{1_{kl}}^t \\ &= (x \otimes_{\mathbb{R}} u)(\overline{1_k}^t \otimes_{\mathbb{R}} \overline{1_l}^t) = x\overline{1_k}^t \otimes u\overline{1_l}^t = x\overline{1_l}^t \cdot u\overline{1_l}^t.\end{aligned}$$

Similarly, one may show $(y \otimes_{\mathbb{R}} v)^* \cdot \overline{1_{kl}}^t = y\overline{1_k}^t \cdot v\overline{1_l}^t$. If $(x \otimes_{\mathbb{R}} u)^* = (y \otimes_{\mathbb{R}} v)^*$ then we contradict the previously derived strict inequality since

$$x\overline{1_k}^t \cdot u\overline{1_l}^t = (x \otimes_{\mathbb{R}} u)^* \cdot \overline{1_{kl}}^t = (y \otimes_{\mathbb{R}} v)^* \cdot \overline{1_{kl}}^t = y\overline{1_k}^t \cdot v\overline{1_l}^t. \qquad \text{Q.E.D.}$$

THEOREM 8. *Suppose A and B are both generator matrices over a common field F. With respect to the Hamming weight function, $A \otimes_F B$ is an optimal generator matrix for its row space if and only if A and B are both optimal generator matrices for their row spaces.*

Proof. Applying Theorem 6 let $\hat{A}$ and $\hat{B}$ be weight retaining generator matrices for the row spaces of A and B, respectively. Recall from (2) that $h_r(\hat{A} \otimes_F \hat{B}) = h_r(\hat{A}) \otimes_{\mathbb{R}} h_r(\hat{B})$. Using Theorems 3 and 5 we see that $\hat{A} \otimes_F \hat{B}$ is weight retaining since

$$h_r(\hat{A} \otimes_F \hat{B}) = h_r(\hat{A}) \otimes_{\mathbb{R}} h_r(\hat{B}) = S_h(\hat{A}) \otimes_{\mathbb{R}} S_h(\hat{B}) = S_h(\hat{A} \otimes_F \hat{B}).$$

Now A, B, and $A \otimes_F B$ are optimal generator matrices if and only if

$$S_h(A)^* = S_h(\hat{A})^*, \quad S_h(B)^* = S_h(\hat{B})^*, \quad \text{and } S_h(A \otimes_F B)^* = S_h(\hat{A} \otimes_F \hat{B})^*,$$

respectively. We wish to apply Lemma 4 to complete the proof. Our correspondences will be $x \leftarrow S_h(A)$, $y \leftarrow S_h(A)$, $u \leftarrow S_h(B)$, and $v \leftarrow S_h(B)$. It is clear that $\overline{0} \leqslant S_h(A)^* \leqslant S_h(\hat{A})^*$, $\overline{0} \leqslant S_h(B)^* \leqslant S_h(\hat{B})^*$, and that $S_h(\hat{A}) \neq \overline{0}$ and $S(\hat{B}) \neq \overline{0}$. Q.E.D.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge many helpful discussions with Professor Jack Levine and with M. J. Niccolai. We wish to thank an unknown referee for his careful reading of the manuscript and for pointing out several errors of omission.

RECEIVED: January 22, 1976; REVISED: June 24, 1977

REFERENCES

BLAKE, I. F., AND MULLIN, R. C. (1975), "The Mathematical Theory of Coding," Academic Press, New York.

CRIMMINS, T. R. (1976), On encoding and decoding maps for group codes, *IEEE Trans. Inform. Theory* **IT-22**, 763–764.

CRIMMINS, T. R., AND HOROWITZ, H. M. (1970), Mean-square error optimum coset leaders for group codes, *IEEE Trans. Inform. Theory* **IT-16**, 429–432.

CRIMMINS, T. R., HOROWITZ, H. M., PALERMO, C. J., AND PALERMO, R. V. (1969), Minimization of mean-square error for data transmitted via group codes, *IEEE Trans. Inform. Theory* **IT-15**, 72–78.

GORE, W. C., AND KILGUS, C. C. (1971), Cyclic codes with unequal error protection, *IEEE Trans. Inform. Theory* **IT-17**, 214–215.

KILGUS, C. C. (1971), "A Class of Cyclic Unequal Error Protection Codes," Ph. D. Dissertation, Department of Electrical Engineering, The Johns Hopkins Univ., Baltimore, Maryland.

KILGUS, C. C., and GORE, W. C. (1972a), A class of cyclic unequal-error-protection codes, *IEEE Trans. Inform. Theory* **IT-18**, 687–690.

KILGUS, C. C., AND GORE, W. C. (1972b), Root-mean-square error in encoded digital telemetry, *IEEE Trans. Communications* **COM-20**, 315–320.

LAWLER, E. L. (1976), "Combinatorial Optimization, Networks and Matroids," Holt Rinehart and Winston, New York.

MANDELBAUM, D. (1972), Unequal-error-protection codes derived from difference sets, *IEEE Trans. Inform. Theory* **IT-18**, 686–687.

MASNICK, B., AND WOLF, J. K. (1967), On linear unequal error protection codes, *IEEE Trans. Inform. Theory* **IT-13**, 600–607.

MASSEY, J. L., COSTELLO, D. J., Jr., AND JUSTESEN, J. (1973), Polynomial weights and code constructions, *IEEE Trans. Inform. Theory* **IT-19**, 101–110.

MITRYAYEV, YE. V. (1963), Transmission of telemetry data by means of group codes, *Radio Eng. Electron. (USSR)* **8**, 935–940.

PETERSON, W. W., AND WELDON, E. J., Jr. (1972), "Error-Correcting Codes," 2nd ed., MIT Press, Cambridge, Mass.

REDINBO, G. R. (1976), The optimum mean-square decoding of general group codes, *Inform. Contr.* **31**, 341–363.

REDINBO, G. R., AND WOLF, G. A. (1974), On minimum mean-square error linear block codes when the data have q-adic weighting, *Inform. Contr.* **26**, 154–177.

WOLF, G. A., AND REDINBO, G. R. (1974), The optimum mean-square estimate for decoding binary block codes, *IEEE Trans. Inform. Theory* **IT-20**, 344–351.

WYNER, A. D. (1965), Capabilities of bounded discrepancy decoding, *Bell System Tech. J.* **44**, 1061–1122.