TR-80-001

Soft Decision Decoding on Linear Block Codes

by

Tai-Yang Hwang

This work was supported by National Science Council Grant NSC-69E-0404-01(01).

Institute of Information Science

Academia Sinica

Nankang, Taipei 115

Republic of China

November 1980

Abstract

A new characterization of the soft decision decoding is formulated. By this characterization it is found that the algebraic structure and the combinatorial structure of a linear block code as well as the received channel measurement information all may be utilized to reduce the decoding complexity. Besides, a new upper bound on the minimum distance of linear binary code is also found.

I. Introduction

Decoding linear block codes by using channel measurement information (i.e., soft decision decoding) is known having better performance than hard decision decoding [1]. But it also poses a question — Is it possible to utilize the algebraic structure of linear block codes into the design of a soft decision decoding scheme ? Obviously, the old correlation decoding which works in the code domain does not answer this question at all. Yet some researches did show that it is possible to utilize the code property of being linear to transform the decoding problem into the dual code domain [2] - [3]. The decoding methods thus obtained have complexities proportional to the total number of dual code words. Some researches have shown that there are other structures in the code space which may be utilized to reduce the decoding complexity from correlation decoding, meanwhile still maintain optimal performance [4] - [5]. There are other researches which showed that for those linear codes having specific combinatoral design, such as orthogonalizable codes, much simpler soft decision decoding methods with suboptimum performance are found [6] - [8]. All these pointed out the possibility of employing the algebraic structure of linear codes effectively in designing a soft decision decoding scheme. However, with their diverse treatment about the decoding problem, those mentioned hardly provide us an integral view about the many faces encountered in our decoding attempt.

There is another question concerned in soft decision decoding — Does the channel measurement information itself

play some role in our attempt ro reduce the decoding complexity ? This question never occurs in hard decision decoding. But with channel measurement information which is composed of variant values, we intuitively suspect that two channel measurement information digits of different values would have different reliabilities in deciding their corresponding code word digits, respectively. But how do we formulize it ? And when can we say that a channel measurement information digit is reliable enough to solely decide its corresponding code word digit ?

This paper presents an attempted effort in characterizing the essence of decoding problem. It shows that the decoding problem is generally a nonlinear programming problem. However, it is also seen in this characterization that the algebraic structure of a code, the combinatorial structure of that code, together with the received channel measurement information all play vivid interactive roles. Henceforth this characterization is eligible to provide insight about how to reduce the decoding complexity. Further, more algebraic structure of linear codes are revealed by taking this approach. Therefore endows us another direction to comprehend error-correcting codes.

In Section II of this paper new algebraic proterties concerning code word digit positions are discussed. This discussion further leads to an upper bound on the minimum Hamming distance of linear binary block codes. In Section III a new characterization of the decoding problem is derived. Based on this characterization, Section IV first provides

condition under which a received channel measurement information digit solely decides its corresponding desired code word digit. Then, method of how to utilize this condition to reduce overall decoding complexity is given. Finally, a maximum likelihood decoding scheme is formulated to end Section IV. A detailed example on (17,8) code to illustrate the decoding algorithm is shown in Section V.

Section VI demonstrates that special combinatorial design may still lead to simpler soft decision decoding method, just like majority-logic decoding in hard decision case, Though the performance is suboptimum. At last, a conclusion is contained in Section VII.

II . New Algebraic Properties

Consider an (n,k) linear binary code C over $\{0,1\}^n$ and its dual (n,n-k) code C'. Let $\underline{c}_1,\underline{c}_2,\ldots,\underline{c}_k$ be k linearly independent code words in C over GF(2), then a lemma follows.

Lemma 1 : $\underline{c}' = (c'_1,c'_2,\ldots,c'_n) \in \{0,1\}^n$ is a code word in C' if and only if

$$\underline{c}' \cdot \underline{c}_j \equiv 0 \mod 2 \qquad (1)$$

for j = 1,2,...,k.

Let $\underline{z}_1$ and $\underline{z}_2$ be two nonzero n-tuples in $\{0,1\}^n$. If $\underline{z}_1 \times \underline{z}_2 = \underline{z}_2$ then we say $\underline{z}_1$ is projected by $\underline{z}_2$ or $\underline{z}_1$ covers $\underline{z}_2$. Consider a set I_t of t distinct positions, $I_t = \{i_1,i_2,\ldots,i_t\}$, with $1 \le i_\ell \le n$ for all $\ell = 1,2,\ldots,t$. Let $\underline{u}_t = (u_1,u_2,\ldots,u_n) \in \{0,1\}^n$ be the incidence vector of I_t such that $u_i = 1$ if and only if $i \in I_t$, $1 \le i \le n$. Now we give a definition.

Definition 1 : The t positions in I_t are linearly independent in code C' if and only if $\underline{u}_t$ is not projected by any nonzero code word $\underline{c} \in C$.

Let d be the minimum Hamming distance of C and d' the minimum Hamming distance of C', respectively. Then d and d' are also the minimum weight of C and C', respectively. Because of any $\underline{u}_t$ being projected by a nonzero $\underline{c} \in C$ will have Hamming weight greater than or equal to d, we have the following Theorem.

Theorem 1. For any $t < d$ distinct positions $i_1, i_2, \ldots, i_t$ with $1 \leq i_\ell \leq n$ for $\ell = 1,2,\ldots,t$, these t positions are always linearly independent in C'.

When t positions are linearly independent in C', it also reveals on the dual code word patterns. This is stated in the next theorem.

Theorem 2. If t distinct positions are linearly independent in C', then the 2^{n-k} code words in C' will present all possible 2^t binary patterns on these t positions.

(Proof) By the definition of t positions being linearly independent in C', it is easily seen that the corresponding t columns of a generator matrix H of C' are linearly independent over GF(2). So the theorem follows.

Q.E.D.

Since there are at most 2^{n-k} different patterns shown by all 2^{n-k} dual code words in C' on any $t \geq n-k$ distinct positions, we have next corollary.

Corollary 1. Any set of n-k+1 or more positions are linearly dependent in C'.

By Theorem 2 it is easily seen that for $t \leq n-k$, given any binary pattern of t zeros and ones and given a set of t positions which are linearly independent in C', we can always find at least one dual code word c' which has the same pattern of zeros and ones on those t positions. But when $t = n-k$, there is one and only one such word exists. The n-k linearly positions form an information set of C'[9]. The k positions left form a redundant set of C'. However, the following

theorem relates the respective information sets in C and C'.

Theorem 3. If $i_1, i_2, \ldots, i_{n-k}$ these n-k positions form an information set in C', then the k positions left form an information set in C.

(Proof) Since $i_1, i_2, \ldots, i_{n-k}$ form an information set in C', by Theorem 2 we can find n-k dual code words such that the $(n-k) \times n$ matrix H formed by these n-k words shows an identity matrix on columns $i_1, i_2, \ldots, i_{n-k}$. From H we can find a generator matrix G of dimensions $k \times n$ for code C and G shows an identity matrix on the k positions left. So the theorem follows.

Q.E.D.

For the dual code word $\underline{c}'$ which is specified by a binary pattern assigned on a given information set in C', the values of the k redundant digits in $\underline{c}'$ can be uniquely decided by substituting the n-k known information digits into the k equations in (1) and solve them. The complete $\underline{c}'$ is hereby constructed.

Given any set of n-k positions, how can we quickly find out whether they form an information set in C' ? Obviously it is not pratical to check whether the incidence vector of this n-k positions is projected by any nonzero code word in C or not. One easy way is to assign a nonzero binary pattern on these n-k positions and substitute them into (1). If the k unknowns are then uniquely solved, we know the n-k given positions form an information set in C'; otherwise, when a contradiction occurs or, there are more than one set of

solutions, it is concluded that they do not form an information set in C'. An example at this point would be adequate.

Example. Consider the (7,4) Hamming code C and its (7,3) dual code C'. In C we find four linearly independent code words

$$\underline{c}_1 = (1\ 0\ 1\ 1\ 0\ 0\ 0)$$
$$\underline{c}_2 = (1\ 1\ 1\ 0\ 1\ 0\ 0)$$
$$\underline{c}_3 = (1\ 1\ 0\ 0\ 0\ 1\ 0)$$
$$\underline{c}_4 = (0\ 1\ 1\ 0\ 0\ 0\ 1)$$

and by Lemma 1 a binary vector $\underline{c}' = (c'_1, c'_2, c'_3, c'_4, c'_5, c'_6, c'_7)$ is a dual code word in C' if and only if

$$\begin{aligned} c'_1 \oplus c'_3 \oplus c'_4 &= 0 \\ c'_1 \oplus c'_2 \oplus c'_3 \oplus c'_5 &= 0 \\ c'_1 \oplus c'_2 \oplus c'_6 &= 0 \\ c'_2 \oplus c'_3 \oplus c'_7 &= 0 \end{aligned} \qquad (2)$$

Now to find out whether positions 1,2,4 form an information set in C', we arbitrarily assign $c'_1 = 1$, $c'_2 = 1$, $c'_4 = 0$. Substitute them into (2) we have an unique solution $c'_3 = 1$, $c'_5 = 1$, $c'_6 = 0$, and $c'_7 = 0$. So positions 1,2,4 form an information set in C'.

As for another set of positions 1,2,6, we assign $c'_1 = 1$, $c'_2 = 1$, and $c'_6 = 0$ and substitute them into (2) results

$$
\begin{aligned}
c'_3 \oplus c'_4 &= 1 \\
c'_3 \oplus c'_5 &= 0 \\
0 &= 0 \\
c'_3 \oplus c'_7 &= 1 \ .
\end{aligned}
$$

So we have two sets of solution for (c'_3,c'_4,c'_5,c'_7), this leads to the conclusion that positions 1,2,6 are not linearly independent in C'. Further check finds that the incidence vector of these three positions is projected by code word (1 1 0 0 0 1 0) in C. Morever, if we assigned $c'_1 = 1$, $c'_2 = 0$ and $c'_6 = 0$ we would see the third equation in (2) resulted 1 = 0, and thus a contradiction.

When a given set of n-k positions are not linearly independent in C', by Theorem 1 we can always extract at least d-1 linearly independent positions from them. Basing on these d-1 positions, we can gradually expand the set of linearly independent positions by adding other positions to it one at a time, then test whether this newly added position will cause linear dependency with the existed positions. The next theorem puts an upper bound on how many positions we have to check to find a set of n-k linearly independent positions in C'.

Theorem 4. In any set of n-(d' - 1) positions we can find a set of n-k linearly independent positions in C'.

(Proof) we prove this by contradiction argument. Suppose we could not find an information set, subsequently we could not find 2^{n-k} different patterns on these n - (d' - 1) positions

in all 2^{n-k} dual code words. Took two nonzero dual code words which had the same pattern on these n - (d' - 1) positions, their binary sum would result either an all-zero word or a nonzero dual code word with weight at most d' - 1. Thus constitutes an contradiction.

Q.E.D.

The discussion of n-k positions to form an information set in C' also sheds insight about the minimum distance of C. Based on this observation we have next theorem which provides an upper bound on d.

Theorem 5. For linear binary codes with $k \geq 2$, $\sum_{i=0}^{m} (\lceil d/2^i \rceil - 1) \leq n - k$, where $m = \min(k-1, \lceil \log_2 d \rceil)$.

(Proof) Let I_1 and I_2 be two different sets of d positions in C' with their incidence vectors $\underline{u}_1$ and $\underline{u}_2$ being two different code words of weight d in C, respectively. Then I_1 and I_2 have at most $\lfloor d/2 \rfloor$ components in common. Denote I_{12} the intersection of I_1 and I_2, and U_{12} the union of I_1 and I_2. Then $|U_{12}|$, the cardinal number of U_{12}, is at least $d + \lceil d/2 \rceil$ and $|I_{12}| \leq \lfloor d/2 \rfloor$. The positions in U_{12} are not linearly independent in C', because the incidence vector of U_{12} is projected by $\underline{u}_1$, $\underline{u}_2$, and $\underline{u}_1 \oplus \underline{u}_2$. However, if we take two positions i_1 and i_2 out of U_{12} with $i_1 \in I_{12}$ and $i_2 \in U_{12} - I_{12}$, then all the positions in $U_{12} - \{i_1, i_2\}$ are linearly independent in C', so we have $(d-1) + (\lceil d/2 \rceil - 1) \leq n - k$. If $k \geq 3$ and we take three different sets of d or more positions in C' with their incidence vectors are linearly independent code words in C respectively, then we will obtain $(d-1) + (\lceil d/2 \rceil - 1) + (\lceil d/4 \rceil - 1) \leq n-k$ by same reasoning. Since there are at most

k linearly independent code words in C and $\lceil d/2^i \rceil - 1 = 0$ for $i \geq \lceil \log_2 d \rceil$, the theorem thus follows.

Q.E.D.

To illustrate Theorem 5, numerical examples of its application are presented in Table 1. For comparison, the results of Elias bound [10] are also listed. It is seen that Theorem 5 provides a tighter bound in most cases.

Table 1 . Upper bounds on minimum distance d .

n	k	d_1	d_2
7	3	4	5
15	6	6	9
17	8	6	9
21	11	7	10
23	12	8	7
23	11	8	9
31	16	9	10
31	15	10	13
35	19	10	10
35	16	12	14
41	21	12	12
41	20	12	15
47	24	14	20
47	23	14	14
51	26	15	16
51	25	16	18
63	32	17	20

d_1 upper bound by Theorem 5

d_2 by Elias bound

III. Characterizing the Decoding Problem

Let the transmitted version of a code word $\underline{c} = (c_1, c_2, \ldots, c_n) \in C$ be $\underline{c}^* = (c_1^*, c_2^*, \ldots, c_n^*) \in C^*$, where $c_i^* = (-1)^{c_i}$. Thus C^* is over $\{+1,-1\}^n$ and the group $\{C, \oplus\}$ is isomorphic to $\{C^*, \times\}$, where $\oplus$ denotes modulo 2 addition and $\times$ denotes component-by-component multiplication.

Assume without loss of generality that the n-k positions $k+1, k+2, \ldots, n$ are information positions in C'. In other words, we can find n-k dual code words $\underline{c}'_1, \underline{c}'_2, \ldots, \underline{c}'_{n-k}$ to form a parity-check matrix H of the following form

$$H = \begin{pmatrix} \underline{c}'_1 \\ \underline{c}'_2 \\ \vdots \\ \underline{c}'_{n-k} \end{pmatrix} = \left(\begin{array}{ccc|ccccc} c'_{11} & c'_{12} & \cdots & c'_{1k} & 1 & & & \\ c'_{21} & c'_{22} & \cdots & c'_{2k} & & 1 & & \\ \vdots & & & \vdots & & & \ddots & \\ c'_{(n-k)1} & \cdots & & c'_{(n-k)k} & & & & 1 \end{array}\right)$$

$$= \left[P \mid I_{n-k} \right]$$

By this parity-check matrix the relationship between C^* and C' is stated in the next lemma.

Lemma 2 For any $\underline{c}^* = (c_1^*, c_2^*, \ldots, c_k^*, c_{k+1}^*, \ldots, c_n^*) \in C^*$

$$c_{k+j}^* = \prod_{i=1}^{k} (c_i^*)^{c'_{ji}}, \quad j = 1, 2, \ldots, n-k. \tag{3}$$

Further, we can write from H a generator matrix G of C which is composed of k linearly independent code words $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$ in C,

$$G = \begin{pmatrix} \underline{c}_1 \\ \underline{c}_2 \\ \vdots \\ \underline{c}_k \end{pmatrix} = \left(\begin{array}{ccccc|ccc} 1 & 0 & 0 & \dots & 0 & c'_{11} & \dots & c'_{(n-k)1} \\ 0 & 1 & 0 & \dots & 0 & c'_{12} & \dots & c'_{(n-k)2} \\ \vdots & & & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & c'_{1k} & \dots & c'_{(n-k)k} \end{array}\right) \quad (4)$$

Obviously, the first k positions are information positions in C.

Now consider a code word $\underline{c}^*$ being sent through a time-discrete memoryless channel with all the code words of C^* are equiprobable. The received word $\underline{r} = (r_1, r_2, \dots, r_n)$ is the real sum of $\underline{c}^*$ and an error vector $\underline{e} = (e_1, e_2, \dots, e_n)$ with $e_i \in R$. A maximum likelihood decoder will find a code word $\underline{c}^*_m$ which maximzes the probability of $\underline{c}^*$ given $\underline{r}$, $P_r(\underline{c}^* \mid \underline{r})$. Define the bit log likelihood ratio of r_i to be

$$\phi_i = \ell n \left\{ \frac{P_r(r_i \mid 1)}{P_r(r_i \mid -1)} \right\}, \quad i = 1, 2, \dots, n.$$

Then $\underline{\phi} = (\phi_1, \phi_2, \dots, \phi_n)$ is the channel measurement information vector of $\underline{r}$. According to [4 , Theorem 5], $\underline{c}^*_m$ maximized $P_r(\underline{c}^* \mid \underline{r})$ if and only if $\underline{c}^*_m$ is the nearest code word to $\underline{\phi}$. Or, in other words, $\underline{c}^*_m$ maximized $P_r(\underline{c}^* \mid \underline{r})$ if and only if $\underline{c}^*_m$ maximizes

$$\underline{\phi} \cdot \underline{c}^* = \phi_1 c^*_1 + \phi_2 c^*_2 + \dots + \phi_n c^*_n \quad (5)$$

for all $\underline{c}^* \in C^*$.

Equation (5) does not take the structure of linear codes directly into consideration. However, by substituting (3) into (5) we have

$$\underline{\phi} \cdot \underline{c}^* = \sum_{i=1}^{k} \phi_i c_i^* + \sum_{j=1}^{n-k} \phi_{k+j} \prod_{i=1}^{k} (c_i^*)^{c'_{ji}}$$

$$= A(c_1^*, c_2^*, \ldots, c_k^*) \qquad (6)$$

which is an function of k information digits $c_1^*, c_2^*, \ldots c_k^*$ in C*. So the decoding problem is now transformed to find a k-tuple $(c_{m1}^*, c_{m2}^*, \ldots, c_{mk}^*) = \underline{x}_m$ which maximizes (6) under the restrictions that $(c_{mi}^*)^2 = 1$ for i = 1,2,...,k. The desired $\underline{c}_m^*$ is then decided by substituting the k-tuple $\underline{x}_m$ into (3) and solving for $c_{m(k+1)}^*, \ldots, c_{mn}^*$.

Maximizing (6) under the restrictions that $(c_i^*)^2 = 1$ is clearly a nonlinear programming problem [11], which is known to be NP-complete [12]. In other words, the best general algorithm known, which is correlation decoding, has complexity 2^k. It is done by substituting all possible binary patterns of $\{+1,-1\}^k$ into (6) and comparing the results, without considering the algebraic structures of linear codes. Further, correlation decoding is not concerned with the values of ϕ_i. However, we shall see in the following that the special form of (6) (which is due to linear codes), the combinatorial structures of C(and C'), and the values of ϕ_i all may be utilized under special conditions to reduce the decoding complexity.

Denote $\underline{x}_1 = (-c_{m1}^*, c_{m2}^*, \ldots, c_{mk}^*)$, $\underline{x}_2 = (c_{m1}^*, -c_{m2}^*, \ldots, c_{mk}^*)$,

..., and $\underline{x}_k = (c^*_{m1}, c^*_{m2}, \ldots, -c^*_{mk})$. So $\underline{x}_i$ is different from $\underline{x}_m$ by the ith bit. Assumed that $\underline{x}_m$ maximizes (6), we have $A(\underline{x}_m) > A(\underline{x}_i)$ for $i = 1,2,\ldots,k$. Or we can write

$$A(\underline{x}_m) - A(\underline{x}_i) = 2\left(\phi_i c^*_{mi} + \sum_{j=1}^{n-k} c'_{ji}\, \phi_{k+j} \prod_{\ell=1}^{k} (c^*_{m\ell})^{c'_{j\ell}}\right) \tag{7}$$

$$= 2\, \underline{c}_i \cdot (\underline{\phi} \times \underline{c}^*_m) \tag{8}$$

$$> 0$$

where $\underline{c}_i$ were defined in (4) and $i = 1,2,\ldots,k$. Examing (7) and (8) it is found that each nonzero term in them involves c^*_{mi}, so we define

$$A_i \equiv \phi_i + \sum_{j=1}^{n-k} c'_{ji}\, \phi_{k+j} \prod_{\substack{\ell=1 \\ \ell \neq i}}^{k} (c^*_\ell)^{c'_{j\ell}}, \quad i = 1,2,\ldots,k,$$

and the next theorem is derived.

Theorem 6 A set of k necessary conditions for an $(c^*_1, c^*_2, \ldots, c^*_k)$ to maximize (6) is

$$c^*_i \; A_i > 0 \tag{9}$$

for $i = 1,2,\ldots,k$.

Now we can state a characterization of the decoding problem.

Characterization of Decoding

The decoding problem is to find a set of k information digits $(c^*_{m1}, c^*_{m2}, \ldots, c^*_{mk})$ in C* such that these digits

(1) satisfy $c^*_i A_i > 0$;

(2) maximizes $A(c^*_1, c^*_2, \ldots, c^*_k)$.

Several remarks are needed here to illustrate what we just derived. The first is, (8) conforms to [4 , Theorem 1] from a different approach. Actually if we consider all 2^k-1 binary patterns which are different from $\underline{x}_m$, we shall obtain [4 , Theorem 1]. The second remark is that from (8) it is clearly seen that each A_i has at least d nonzero terms. In the third, c^*_{mi} may be decided by A_i, since their product must be greater than zero. Finally it is interesting to point out the similarity in form between A_i and the $F_0(\underline{r})$ defined in [6, Eq.(1)].

A tempting decoding procedure seen from Theorem 6 is started by randomly selecting an initial $\underline{x}$ and then check whether $\underline{x}$ satisfies the k inequalities in (9). If there is an $c^*_i A_i < 0$, then change c^*_i to $-c^*_i$ and go back to test (9) again. Eventually it will stop at a $\underline{x}'_m$ which satisfies (9) and $\underline{x}'_m$ is taken as the desired output. Though this decoding procedure would have hardware complexity k simply, it unfortunately does not always do maximum likelihood decoding. Due to the fact that two or more binary k-tuples may satisfy (9) simultaneously.

However, Theorem 6 does suggest a maximum likelihood decoding algorithm which has complexity 2^{n-k}. To illustrate this let us go back to (6). In (6) we found $\prod_{i=1}^{k}(c_i^*)^{c'_{ji}}$ have value either 1 or -1, for j = 1,2,.., n-k. So totally there are 2^{n-k} combinations for these n-k products. For each one of them we can decide a set of $(c_1^*,c_2^*,\ldots,c_k^*)$ which satisfies (9) as well as the assigned values of $\prod_{i=1}^{k}(c_i^*)^{c'_{ji}}$, j= 1,2,...,n-k. The desired output $\underline{x}_m$ is obtained by comparing these 2^{n-k} sets of $(c_1^*,c_2^*,\ldots,c_k^*)$.

Yet we are not contented with decoding algorithm of complexity 2^{n-k}. Our characterization of the decoding problem provides us more insight about when and how we can reduce the decoding complexity. This is discussed in the next section.

IV. Decoding

Take Theorem 6 into consideration again, we are aware of two facts: 1) Not all binary k-tuples in $\{+1,-1\}^k$ may satisfy (9). For example, if $(c_1^*, c_2^*, \ldots, c_k^*)$ satisfies (9) then $(-c_1^*, c_2^*, \ldots, c_k^*)$, $(c_1^*, -c_2^*, \ldots, c_k^*)$, $\ldots$, $(c_1^*, c_2^*, \ldots, -c_k^*)$ do not satisfy (9); 2) Each of $\phi_1, \phi_2, \ldots, \phi_k$ appears only once in one inequality in (9), respectively. And each such ϕ_i is combined with c_i^* closely. These two facts offer incentive to build even sophisticated decoding algorithm. Let us consider the second fact first, which leads to the following theorem.

Theorem 7. If

$$|\phi_i| > \sum_{j=1}^{n-k} c'_{ji} |\phi_{k+j}| \qquad (10)$$

for some i, $1 \le i \le k$, then c_{mi}^* is uniquely determined.

(Proof) By equations (7) and (9) it is clearly seen that if (10) is true then the only solution for (9) is $c_{mi}^* = 1$ when $\phi_i > 0$ and $c_{mi}^* = -1$ when $\phi_i < 0$.

Q.E.D.

Theorem 7 provides the condition that a received ϕ_i solely determines c_{mi}^* and thereby demonstrates that $\underline{\phi}$ may be employed to reduce the decoding complexity under certain circumstances. If there are several c_{mi}^* determined in this way, undoubtedly the decoding problem of finding $\underline{x}_m$ is greatly reduced. By then we might be able to decide $\underline{x}_m$ by testing all possible patterns of the undetermined c_{mj}^*. So the problem that we are interested now is : How to make (10) happen for as many i's as possible?

Obviously, a necessary condition for (10) to be true is that $|\phi_i| > c'_{ji}|\phi_{k+j}|$ for all $j = 1,2,\ldots, n-k$. This gives us a hint that in order to fully take the advantage of (10), do not fix the n-k information positions in C' which were used in (3) to define the n-k equations. Alternatively, the n-k information positions in C' which would be chosen should have the absolute values of their corresponding ϕ_j as small as possible, due to the fact that those ϕ_j would appear in the right hand side of (10). When this is done, the remaining k ϕ_i's which appear in the left part of the k inequalities in (10) respectively will have greater absolute values. Thus (10) will have better chance to be realized.

Therefore at the receiving end we first select n-k positions according to the n-k components in $\underline{\phi}$ which have the smallest absolute values. If they are verified as information positions in C', n-k equations of the form in (3) can be constructed according to Theorem 2. Subsequently the k inequalities in (9) are found with each of the k largest absolute value components in $\underline{\phi}$ appears in exactly one inequality. And it gives (10) most chance to be realized.

If the n-k originally selected positions are not information positions in C', then by Theorem 1 we can first take the d-1 positions which have smallest absolute ϕ_j components. We then gradually expand this set of linearly independent positions in C' by adding other positions to it one at a time, starting with the one which has the next smallest absolute ϕ_j. Then check whether this newly added position will cause linear dependency with the existed positions. By Theorem 4 at most n-(d-1)-(d'-1) positions have to be searched to form an information set in C'. combined with the d-1 originally selected positions. By these n-k positions in the information set, n-k equations of the form in (3) can be written with these n-k linearly independent positions appear at the left hand side of (3). Subsequently k inequalities of the form in (9) can be written. Importantly, by Theorem 4 we know that in these k inequalities there are t inequalities having the t largest absolute ϕ_i's appear in each of them, with $d'-1 \leq t < k$. So (10) may be realized for at least t ϕ_i's. We conclude this result in next theorem.

Theorem 8 . It is always possible to find an information set I_{n-k} in C', where I_{n-k} contains the d-1 positions which having smallest absolute ϕ_j's, and I_{n-k} does not contain the d'-1 positions which having the largest absolute ϕ_i s.

Another factor which will affect the outcome of (10) is the number of nonzero terms at the right hand side of (10) (or the number of nonzero c'_{ji} for each i). We would like to keep $\sum_{j=1}^{n-k} c'_{ji}$ as small as possible, though $\sum_{j=1}^{n-k} c'_{ji} \geq d - 1$. However, since our n-k selected information positions in C' are changing,

there is no general answer to this problem. At this moment we could only wish $\sum_{j=1}^{n-k} c'_{ji}$ is as small as possible, no matter what information set in C' is selected.

Now assume n - t desired digits, where $0 \le t \le k$, are decided by employing Theorem 7 on (9) and let them be $c^*_{m(t+1)}, c^*_{m(t+2)}, \ldots, c^*_{mk}$, without loss of generality. So $c^*_{m1}, c^*_{m2}, \ldots, c^*_{mt}$ are still to be decided and they must satisfy (9) for $i = 1,2,\ldots,t$. Substituting $c^*_{m(t+1)}, \ldots, c^*_{mk}$ into the A_i's in these t inequalities we have

$$A_i = \phi_i + \sum_{j=1}^{n-k} c'_{ji} \left(\phi_{k+j} \prod_{\ell=t+1}^{k} (c^*_\ell)^{c'_{j\ell}} \right) \prod_{\substack{\ell=1 \\ \ell \neq i}}^{t} (c^*_\ell)^{c'_{j\ell}} \qquad (11)$$

for $i = 1,2,\ldots, t$. Here we note that

$$\prod_{\ell=t+1}^{k} (c^*_\ell)^{c'_{j\ell}}$$

is already decided by the n - t known digits. So if we have some $\underline{c}'_j$ such that $c'_{j\ell} = 0$ for all $1 \le \ell \le t$ except $\ell = i$, $1 \le i \le t$, we have

$$\phi_{k+j} \prod_{\ell=t+1}^{k} (c^*_\ell)^{c'_{j\ell}}$$

as a constant and it can be added to ϕ_i such that Theorem 7 may be employed again to find even more c^*_{mi}. By the same reason, such newly found c^*_{mi} may further help finding more desired digits. So the originally found $c^*_{m(t+1)}, \ldots, c^*_{mk}$ may trigger

a chain reaction to have most of the desired digits found.

However, we may also encounter the situation that Theorem 7 only helped us finding few c^*_{mi}'s. By then we are left with t inequalities of the form in (9) which are to be satisfied and t is not a small number. To find $c^*_{m1}, c^*_{m2}, \ldots, c^*_{mt}$ we note that though there may be more than one t-tuple in $\{+1,-1\}^t$ which satisfies those t inequalities, we also aware that not all t-tuplesin $\{+1,-1\}^t$ can satisfy them. Therefore one decoding approach that may have less complexity is, managing to find all $(c^*_1, c^*_2, \ldots, c^*_t)$ which satisfy the required t inequalities and then find the desired one from them by comparing their effects on (6). To realize this, a strategy is : Find a solution for some $c^*_i A_i > 0$, say $c^*_t A_t > 0$, by assigning values to unknown components in A_t as well as c^*_t and substitute this solution to rest inequalities. Repeating this at most t times we may decide a solution for all t inequalities or, if there is a contradiction occurs in the middle of this process, we can conclude that no solution exists by previous assignments. After this we go back one inequality and find another solution for that $c^*_i A_i > 0$ and proceed as described. Doing this iteratively we will be able to find all solutions of the required t inequalities.

We note that in above process Theorem 7 still helps us to reduce the decoding complexity. To illustrate this we know in a $c^*_i A_i > 0$ with A_i of the form in (11), A_i can have 2^{u_i} different values by assigning either +1 or -1 to those unknown

$$\prod_{\substack{\ell=1 \\ \ell \neq i}}^{t} (c_\ell^*)^{c'_{j\ell}} \quad , \quad j = 1,2,\ldots,n-k \text{ and } c'_{ji} \neq 0$$

where $u_i = \sum_{j=1}^{n-k} c'_{ji}$, and the value of c_i^* with respect to A_i is determined subsequently. So generally there are 2^{u_i} solutions for a $c_i^* A_i > 0$ and each one of them has to be checked against the rest inequalities. But suppose there is a $c'_{vi} \neq 0$, $1 \leq v \leq n-k$, and

$$|\phi_i| + |\phi_{k+v}| > \sum_{\substack{j=1 \\ j \neq v}}^{n-k} |c'_{ji}\phi_{k+j}| \quad ,$$

then we may assign

$$\prod_{\substack{\ell=1 \\ \ell \neq i}}^{t} (c_\ell^*)^{c'_{v\ell}}$$

a value s_v, $s_v = +1$ or -1, such that

$$\operatorname{sgn}(\phi_i) = \operatorname{sgn}(\phi_{k+v} s_v \prod_{\ell=t+1}^{k} (c_\ell^*)^{c'_{v\ell}})$$

and by Theorem 7 the solution

$$(c_i^* = \operatorname{sgn}(\phi_i), \prod_{\substack{\ell=1 \\ \ell \neq i}}^{t} (c_\ell^*)^{c'_{j\ell}} = s_i)$$

suffices to represent $2^{u_i - 1}$ solutions of $c_i^* A_i > 0$.

In conclusion, a soft-decision decoding scheme is given in the following : For a received $\underline{r}$,

(1) Calculate $\underline{\phi}$;

(2) Choose an information set I_{n-k} in C' according to the component absolute values in $\underline{\phi}$, such that for those positions in I_{n-k} their corresponding components in $\underline{\phi}$ having absolute values as small as possible ;

(3) Construct a parity check matrix H such that for the n - k positions in I_{n-k} their corresponding columns in H form an identity matrix ;

(4) Write out n - k equations of the form in (3) by this H ;

(5) Write out k inequalities of the form in (9) which are necessary conditions to be satisfied ;

(6) Find all possible solutions to those k inequalities (with the help of Theorem 7) ;

(7) Find the unique desired solution ;

(8) Find the desired code word.

To illustrate this decoding scheme, a detailed example on (17,8) code is provided in next section.

V. Example

Consider the (17,8;6) code C and its dual (17,9;5) code C'. We have the parity check matrix

$$H = \begin{pmatrix} \underline{c}_1' \\ \underline{c}_2' \\ \underline{c}_3' \\ \underline{c}_4' \\ \underline{c}_5' \\ \underline{c}_6' \\ \underline{c}_7' \\ \underline{c}_8' \\ \underline{c}_9' \end{pmatrix} = \left(\begin{array}{cccccccc|ccccccccc} 1&0&0&1&1&1&0&0 & 1 & & & & & & & & \\ 0&1&0&0&1&1&1&0 & & 1 & & & & & & & \\ 0&0&1&0&0&1&1&1 & & & 1 & & & & & & \\ 1&0&0&0&1&1&1&1 & & & & 1 & & & & & \\ 1&1&0&1&1&0&1&1 & & & & & 1 & & & & \\ 1&1&1&1&0&0&0&1 & & & & & & 1 & & & \\ 1&1&1&0&0&1&0&0 & & & & & & & 1 & & \\ 0&1&1&1&0&0&1&0 & & & & & & & & 1 & \\ 0&0&1&1&1&0&0&1 & & & & & & & & & 1 \end{array}\right)$$

$$= \left[\, P \mid I_{n-k} \,\right],$$

and the generator matrix

$$G = \begin{pmatrix} \underline{c}_1 \\ \underline{c}_2 \\ \underline{c}_3 \\ \underline{c}_4 \\ \underline{c}_5 \\ \underline{c}_6 \\ \underline{c}_7 \\ \underline{c}_8 \end{pmatrix} = \left(\begin{array}{ccccccccc|cccccccc} 1 & & & & & & & & & 1&0&0&1&1&1&1&0&0 \\ & 1 & & & & & & & & 0&1&0&0&1&1&1&1&0 \\ & & 1 & & & & & & & 0&0&1&0&0&1&1&1&1 \\ & & & 1 & & & & & & 1&0&0&0&1&1&0&1&1 \\ & & & & 1 & & & & & 1&1&0&1&1&0&0&0&1 \\ & & & & & 1 & & & & 1&1&1&1&0&0&1&0&0 \\ & & & & & & 1 & & & 0&1&1&1&1&0&0&1&0 \\ & & & & & & & 1 & & 0&0&1&1&1&1&0&0&1 \end{array}\right)$$

$$= \left[\, I_k \mid P^T \,\right].$$

So by Lemma 1 $\underline{c}' = (c'_1, c'_2, \ldots, c'_{17})$ is a code word in (17,9) code if and only if

$$
\begin{aligned}
c'_1 \oplus c'_9 \oplus c'_{12} \oplus c'_{13} \oplus c'_{14} \oplus c'_{15} &= 0 \\
c'_2 \oplus c'_{10} \oplus c'_{13} \oplus c'_{14} \oplus c'_{15} \oplus c'_{16} &= 0 \\
c'_3 \oplus c'_{11} \oplus c'_{14} \oplus c'_{15} \oplus c'_{16} \oplus c'_{17} &= 0 \\
c'_4 \oplus c'_9 \oplus c'_{13} \oplus c'_{14} \oplus c'_{16} \oplus c'_{17} &= 0 \\
c'_5 \oplus c'_9 \oplus c'_{10} \oplus c'_{12} \oplus c'_{13} \oplus c'_{17} &= 0 \\
c'_6 \oplus c'_9 \oplus c'_{10} \oplus c'_{11} \oplus c'_{12} \oplus c'_{15} &= 0 \\
c'_7 \oplus c'_{10} \oplus c'_{11} \oplus c'_{12} \oplus c'_{13} \oplus c'_{16} &= 0 \\
c'_8 \oplus c'_{11} \oplus c'_{12} \oplus c'_{13} \oplus c'_{14} \oplus c'_{17} &= 0
\end{aligned}
\qquad (12)
$$

Assume the receiving $\underline{\phi}$ = (0.91,0.12,-1.2,-0.05,-0.08,1.25, -0.89,-1.5,-0.22,0.14,0.35,-0.56,0.43,0.62,-0.10,1.51,-0.85). We first take positions {2,4,5,9,10,11,12,13,15} since they have the n-k=9 smallest absolute ϕ_i's in $\underline{\phi}$. Arbitrarily assign $c'_2 = c'_4 = c'_{10} = c'_{11} = c'_{13} = 0$ and $c'_5 = c'_9 = c'_{12} = c'_{15} = 1$ and substitute them into (12) we find contradiction This means the nine chosen positions are not linearly independent in C'. Now since d = 6, we take positions {2,4,5,10,15} they have the d-1 smallest absolute ϕ_i's. Starting from these 5 positions we finally find four positions {9,11,12,14} combined to form an information set I_9 in C'. Assigning $c'_2 = 1$, $c'_4 = c'_5 = c'_9 = c'_{10} = c'_{11} = c'_{12} = c'_{14} = c'_{15} = 0$. Substituting them into (12), we obtain $c'_1 = c'_3 = c'_7 = c'_{13} = c'_{17} = 1$, $c'_6 = c'_8 = c'_{16} = 0$. Therefore we have a new dual code word $\underline{c}'$ = (1 1 1 0 0 0 1 0 0 0 0 0 1 0 0 0 1). Similarly we can find

8 other dual code words which together with the one just found form an 9×9 identity matrix on the 9 positions in I_9. Therefore we have a new parity check matrix

$$H_1 = \begin{pmatrix}
1&1&1&0&0&0&1&0&0&0&0&0&1&0&0&0&1\\
1&0&0&1&0&0&0&0&0&0&0&0&1&0&0&1&1\\
1&0&1&0&1&0&0&1&0&0&0&0&1&0&0&1&0\\
1&0&1&0&0&1&0&1&1&0&0&0&0&0&0&0&1\\
0&0&0&0&0&1&0&1&0&1&0&0&0&0&0&1&1\\
0&0&1&0&0&1&1&1&0&0&1&0&0&0&0&0&0\\
0&0&1&0&0&1&1&0&0&0&0&1&1&0&0&1&0\\
1&0&0&0&0&0&1&1&0&0&0&0&0&1&0&1&0\\
0&0&0&0&0&1&1&0&0&0&0&0&1&0&1&0&1
\end{pmatrix}$$

By Lemma 2 nine equations are written,

$$\begin{aligned}
c_2^* &= c_1^* \; c_3^* \; c_7^* \; c_{13}^* \; c_{17}^*\\
c_4^* &= c_1^* \; c_{13}^* \; c_{16}^* \; c_{17}^*\\
c_5^* &= c_1^* \; c_3^* \; c_8^* \; c_{13}^* \; c_{16}^*\\
c_9^* &= c_1^* \; c_3^* \; c_6^* \; c_8^* \; c_{17}^*\\
c_{10}^* &= c_6^* \; c_8^* \; c_{16}^* \; c_{17}^* \qquad (13)\\
c_{11}^* &= c_3^* \; c_6^* \; c_7^* \; c_8^*\\
c_{12}^* &= c_3^* \; c_6^* \; c_7^* \; c_{13}^* \; c_{16}^*\\
c_{14}^* &= c_1^* \; c_7^* \; c_8^* \; c_{16}^*\\
c_{15}^* &= c_6^* \; c_7^* \; c_{13}^* \; c_{17}^*
\end{aligned}$$

And the decoding problem is transformed to find a $(c_1^*, c_3^*,$

$c_6^*, c_7^*, c_8^*, c_{13}^*, c_{16}^*, c_{17}^*$) which maximizes

$$
\begin{aligned}
A = {} & 0.91c_1^* - 1.2c_3^* + 1.25c_6^* - 0.89c_7^* - 1.5c_8^* + 0.43c_{13}^* \\
& + 1.51c_{16}^* - 0.85c_{17}^* + 0.12c_1^* c_3^* c_7^* c_{13}^* c_{17}^* \\
& - 0.05c_1^* c_{13}^* c_{16}^* c_{17}^* - 0.08c_1^* c_3^* c_8^* c_{13}^* c_{16}^* \\
& - 0.22c_1^* c_3^* c_6^* c_8^* c_{17}^* + 0.14c_6^* c_8^* c_{16}^* c_{17}^* \\
& + 0.35c_3^* c_6^* c_7^* c_8^* - 0.56c_3^* c_6^* c_7^* c_{13}^* c_{16}^* \\
& + 0.62c_1^* c_7^* c_8^* c_{16}^* - 0.10c_6^* c_7^* c_{13}^* c_{17}^* .
\end{aligned}
$$

By Theorem 6 , $(c_1^*, c_3^*, c_6^*, c_7^*, c_8^*, c_{13}^*, c_{16}^*, c_{17}^*)$ must satisfy the following eight inequalites :

$$
\begin{aligned}
c_1^* \; & (0.91 + 0.12c_3^* c_7^* c_{13}^* c_{17}^* - 0.05c_{13}^* c_{16}^* c_{17}^* \\
& - 0.08c_3^* c_8^* c_{13}^* c_{16}^* - 0.22c_3^* c_6^* c_8^* c_{17}^* \\
& + 0.62c_7^* c_8^* c_{16}^*) > 0 \\
c_3^* \; & (-1.2 + 0.12c_1^* c_7^* c_{13}^* c_{17}^* - 0.08c_1^* c_8^* c_{13}^* c_{16}^* \\
& - 0.22c_1^* c_6^* c_8^* c_{17}^* + 0.35c_6^* c_7^* c_8^* \\
& - 0.56c_6^* c_7^* c_{13}^* c_{16}^*) > 0 \\
c_6^* \; & (1.25 - 0.22c_1^* c_3^* c_8^* c_{17}^* + 0.14c_8^* c_{16}^* c_{17}^* \\
& + 0.35c_3^* c_7^* c_8^* - 0.56c_3^* c_7^* c_{13}^* c_{16}^* \\
& - 0.10c_7^* c_{13}^* c_{17}^*) > 0 \\
c_7^* \; & (-0.89 + 0.12c_1^* c_3^* c_{13}^* c_{17}^* + 0.35c_3^* c_6^* c_8^* \\
& - 0.56c_3^* c_6^* c_{13}^* c_{16}^* + 0.62c_1^* c_8^* c_{16}^* \\
& - 0.10c_6^* c_{13}^* c_{17}^*) > 0 \\
c_8^* \; & (-1.5 - 0.08c_1^* c_3^* c_{13}^* c_{16}^* - 0.22c_1^* c_3^* c_6^* c_{17}^* \\
& + 0.14c_6^* c_{16}^* c_{17}^* + 0.35c_3^* c_6^* c_7^* \\
& + 0.62c_1^* c_7^* c_{16}^*) > 0 \\
c_{13}^* \; & (0.43 + 0.12c_1^* c_3^* c_7^* c_{17}^* - 0.05c_1^* c_{16}^* c_{17}^* \\
& - 0.08c_3^* c_8^* c_{13}^* c_{16}^* - 0.56c_3^* c_6^* c_7^* c_{16}^*
\end{aligned}
$$

$$c_{16}^* \; (1.51 - 0.05c_1^* \; c_{13}^* \; c_{17}^* - 0.08c_1^* \; c_3^* \; c_8^* \; c_{13}^* + 0.14c_6^* \; c_8^* \; c_{17}^* - 0.56c_3^* \; c_6^* \; c_7^* \; c_{13}^* + 0.62c_1^* \; c_7^* \; c_8^* \;) > 0$$

$$c_{17}^* \; (-0.85 + 0.12c_1^* \; c_3^* \; c_7^* \; c_{13}^* - 0.05c_1^* \; c_{13}^* \; c_{16}^* - 0.22c_1^* \; c_3^* \; c_6^* \; c_8^* + 0.14c_6^* \; c_8^* \; c_{16}^* - 0.10c_6^* \; c_7^* \; c_{13}^* \;) > 0 \; .$$

Examine these nine inequalities and we find by Theorem 7 that $c_8^* = -1$, $c_{16}^* = 1$, and $c_{17}^* = -1$. Substitute them into the above inequalities we further find that $c_6^* = 1$ and we have only four inequalities left,

$$c_1^* \; (0.91 - 0.12c_3^* \; c_7^* \; c_{13}^* + 0.05c_{13}^* + 0.08c_3^* \; c_{13}^* - 0.22c_3^* - 0.62c_7^* \;) > 0 \qquad (14)$$

$$c_3^* \; (-1.2 - 0.12c_1^* \; c_7^* \; c_{13}^* + 0.08c_1^* \; c_{13}^* - 0.22c_1^* - 0.35c_7^* - 0.56c_7^* \; c_{13}^* \;) > 0 \qquad (15)$$

$$c_7^* \; (-0.89 - 0.12c_1^* \; c_3^* \; c_{13}^* - 0.35c_3^* - 0.56c_3^* \; c_{13}^* - 0.62c_1^* + 0.1c_{13}^* \;) > 0 \qquad (16)$$

$$c_{13}^* \; (0.43 - 0.12c_1^* \; c_3^* \; c_7^* + 0.05c_1^* + 0.08c_1^* \; c_3^* - 0.56c_3^* \; c_7^* + 0.1c_7^* \;) > 0 \; . \qquad (17)$$

Since the leading component in (15) has the largest absolute value among the four inequalities, we take (15) into consideration first. Consider the two possible values of $c_7^* \; c_{13}^* = 1$ or -1. When $c_7^* \; c_{13}^* = 1$, by Theorem 7 we have $c_3^* = -1$. Substitute these values into (14) we have $c_1^* = 1$. Subsequently by (16) we have $c_7^* = -1$ and so $c_{13}^* = -1$ (since we assumed $c_7^* \; c_{13}^* = 1$). Check with (17), this

solution ($c_1^* = 1$, $c_3^* = -1$, $c_7^* = -1$, $c_{13}^* = -1$) satisfies it.

When $c_7^* \; c_{13}^* = -1$, c_3^* can not be determined now. But we consider $c_7^* = 1$ and $c_7^* = -1$, respectively. When $c_7^* = 1$, we have by Theorem 7 that $c_3^* = -1$. Also by the assumption that $c_7^* \; c_{13}^* = -1$ we have $c_{13}^* = -1$. Substitute these into (14), we find $c_1^* = 1$. However, the resulted ($c_1^* = 1$, $c_3^* = -1$, $c_7^* = 1$, $c_{13}^* = -1$) does not satisfy (17). So when $c_7^* \; c_{13}^* = -1$, we can only assume $c_7^* = -1$.

Now with $c_7^* \; c_{13}^* = -1$ and $c_7^* = -1$, we have $c_{13}^* = 1$. Substitute them into (15) we have

$$c_3^* \; (-0.31 - 0.02 c_1^* \;) > \; 0 \; .$$

Obviously $c_3^* = -1$. By (14), we also find $c_1^* = 1$. However, the resulted ($c_1^* = 1$, $c_3^* = -1$, $c_7^* = -1$, $c_{13}^* = 1$) does not satisfy (17).

Conclusively, there is only one solution ($c_1^* = 1$, $c_3^* = -1$, $c_6^* = 1$, $c_7^* = -1$, $c_8^* = -1$, $c_{13}^* = -1$, $c_{16}^* = 1$, $c_{17}^* = -1$). By (13), we find $c_2^* = 1$, $c_4^* = 1$, $c_5^* = -1$, $c_9^* = -1$, $c_{10}^* = 1$, $c_{11}^* = -1$, $c_{12}^* = -1$, $c_{14}^* = 1$, and $c_{15}^* = -1$. So

$$\underline{c}_m^* = (1, \; 1 \; , \; -1 \; , \; 1 \; , \; -1 \; , \; 1 \; , \; -1 \; , \; -1 \; , \; -1 \; , \; 1 \; ,$$
$$-1 \; , \; -1 \; , \; -1 \; , \; 1 \; , \; -1 \; , \; 1 \; , \; -1 \;) \; ,$$

which is different from the hard decision vector of $\underline{\phi}$ by three components.

VI Combinatorial Design and Decoding

In previous sections it is shown that our characterization of the decoding problem has led us to new maximum likelihood decoding algorithm. In addition, we will see in this section that suboptimum decoding performance, such as generalized minimum distance (GMD) decoding [1], can also be reached by this line of thought, provided that special combinatorial design exists in the code.

We first introduce some notations which are generally used in GMD decoding [1],[7]. Define $\underline{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ and

$$\alpha_i = \begin{cases} +1 , & \text{if } T < \phi_i , \\ \phi_i / T , & \text{if } -T \le \phi_i \le T , \\ -1 , & \text{if } \phi_i < -T , \end{cases}$$

where T is some positive threshold. Thus, except for the hard-limiting at each end, the bit log-likelihood ratios are preserved in $\underline{\alpha}$. Assume α_M is a component in $\underline{\alpha}$ which has the largest absolute value. If $\alpha_M \neq 0$, $\underline{\alpha}$ can be redefined as

$$\underline{\beta} = \underline{\alpha} / |\alpha_M| = (\alpha_1/|\alpha_M| , \alpha_2/|\alpha_M| , \ldots, \alpha_n/|\alpha_M|).$$

Now the ith component of $\underline{\beta}$ satisfies $-1 \le |\beta_i| \le 1$, $1 \le i \le n$, and at least one component of $\underline{\beta}$ has its absolute value equal to one.

[7, Theorem 1] states that for any $\underline{\beta}$, there is at most one code word $\underline{c}^*$ such that

$$\underline{\beta} \cdot \underline{c}^* > n - d \qquad (18)$$

is satisfied. And any decoding method which successfully finds this $\underline{c}^*$ is credited with doing GMD decoding.

Define $\underline{z} = (z_1, z_2, \ldots, z_n)$ and

$$z_i = \operatorname{sgn}(\beta_i) = \begin{cases} +1, & \text{if } \beta_i > 0 , \\ -1, & \text{if } \beta_i \leq 0 . \end{cases}$$

The following lemmas reveal important facts for GMD decoding.

Lemma 3 If there is an $\underline{c}_m^*$ which satisfies (18), then $\underline{z} \times \underline{c}_m^*$ have at most d - 1 components less than or equal to zero.

Lemma 4. If $\underline{\beta} \cdot \underline{c}_m^* > n - d$, then take any d or more components from $\underline{\beta} \times \underline{c}_m^*$ their sum will be greater than zero.

The proofs of these two lemmas follow naturally from the fact that $|\beta_i| \leq 1$ for $i = 1,2,\ldots,n$, and the assumption that $\underline{\beta} \times \underline{c}_m^* > n - d$.

Let $T = \{i_1, i_2, \ldots, i_t\}$ be a set of t indices with $0 \leq t < d$, $1 \leq i_\ell \leq n$, and $1 \leq \ell \leq t$ such that $\beta_{i_\ell} c^*_{m i_\ell} \leq 0$ for those $i_\ell \in T$. Also let $T' = \{j_1, j_2, \ldots, j_{d-t}\}$ be a set of d - t indices, $j_\ell \notin T$ for $1 \leq j_\ell \leq n$ and $1 \leq \ell \leq d - t$, such that $\beta_{j_\ell} c^*_{m j_\ell}$ is among the d - t smallest terms within the n - t positive $\beta_i c^*_{m\,i}$'s, $i \in \{1,2,\ldots,n\} - T$. Therefore by Lemma 4 we have

$$X \equiv \sum_{j \in T'} \beta_j \, c^*_{mj} + \sum_{i \in T} \beta_i \, c^*_{mi} > 0$$

Now suppose the $(n - k) \times k$ matrix P defined in the parity check matrix H is a balanced incomplete block design [13] with

$$\sum_{i=1}^{k} c'_{ji} = r , \quad j = 1,2,\ldots,n - k ,$$

$$\sum_{j=1}^{n-k} c'_{ji} = s, \quad i = 1,2,\ldots,k$$

and $\lambda = 1$. So P is an $(k, n-k, r, s, \lambda = 1)$ – configuration and next theorem gives the minimum distance of C.

Theorem 9 If P is an $(k, n-k, r, s, \lambda = 1)$ – configuration, then $d = s + 1$.

(Proof) we first see by (4) that the Hamming weights of $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_k$ are $s + 1$, respectively, so $d \leq s + 1$. By checking H we find that since $\lambda = 1$, any set of s columns are linearly independent, so $d \geq s + 1$. Therefore $d = s + 1$.

Q.E.D.

Now for the $\underline{c}^*_m \in C^*$ which satisfies (18), $\underline{c}^*_m$ maximizes $\underline{\beta} \cdot \underline{c}^*$ for all $\underline{c}^* \in C^*$. So $\underline{c}^*_m$ also satisfies

$$c^*_{mi} \cdot (\beta_i + \sum_{j=1}^{n-k} c'_{ji}\beta_{k+j} \prod_{\substack{\ell=1 \\ \ell \neq i}}^{k} (c^*_{m\ell})^{c'_{j\ell}}) > 0 \qquad (19)$$

for $i = 1,2,\ldots,k$, by Theorem 6. To find c^*_{mi} from $\underline{\beta}$, we let

$$\gamma_{ji} \equiv \min_{\substack{1 \leq \ell \leq k \\ \ell \neq i}} (|\beta_{k+j}|, |(\beta_\ell)^{c'_{j\ell}}|)$$

for each $j = 1,2,\ldots,n-k$, $i = 1,2,\ldots,k$, when $c'_{ji} \neq 0$. And define $\gamma_{ji} = 0$ for those $c'_{ji} = 0$. Now for each $i = 1,2,\ldots,k$, we define

$$B_i \equiv \beta_i + \sum_{j=1}^{n-k} c'_{ji}\gamma_{ji} z_{k+j} \prod_{\substack{\ell=1 \\ \ell \neq i}}^{k} (z_\ell)^{c'_{j\ell}}$$

and consider

$$c^*_{mi} B_i = c^*_{mi}\beta_i + \sum_{j=1}^{n-k} c'_{ji}\gamma_{ji} c^*_{mi} z_{k+j} \prod_{\substack{\ell=1 \\ \ell\neq i}}^{k} (z_\ell)^{c'_j} \qquad (20)$$

which in form is similar to the left hand side of (19). There are d terms in (20), and since $\lambda = 1$ we know there are at most t terms in (20) are less than or equal to zero, $0 \leq t < d$. Therefore it is obvious that

$$c^*_{mi} B_i \geq X > 0 \quad , \; i = 1,2,\ldots,k \; .$$

A decoding rule is hereby formulated.

Decoding Rule 1. Decode

$$\hat{c}^*_{mi} = 1 \quad \text{if } B_i > 0 \; ,$$

$$\hat{c}^*_{mi} = -1 \quad \text{if } B_i \leq 0 \; ,$$

for $i = 1,2,\ldots,k$.

We know that when $\underline{\beta} \cdot \underline{c}^*_m > n - d$ this decoding rule will provide correct estimates of c^*_{m1}, c^*_{m2}, ..., c^*_{mk}, and $\underline{c}^*_m$ then follows.

VII. Conclusion

The main purpose of this paper is to introduce an integral approach to soft-decision decoding. The newly proposed characterization of the decoding problem seems to meet this goal successfully. Accordingly, it is found that algebraic structure and combinatorial structure of linear block codes do affect decoding complexity. However, it is the channel measurement information that plays a key role in pointing out when and how to utilize the internal structure of codes. It is surprising to see that channel measurement information did not hinder our effort to reduce the decoding complexity as we first thought it would be. After all if we know more we may be able to do more. Channel measurement information serves exactly this way.

At this moment it is unknown yet whether some traditionally important algebraic structure of linear block codes, such as the roots of generator polynomial of a code, will or will not contribute any in this line of decoding approach. It is though suspected that some such structure may affect the combinatorial design on P and thus lead to simpler but suboptimum decoding methods.

Acknowledgment

The author wishes to thank Mr. S. U. Guam for compiling Table 1 and many helpful discussions.

References

1 G. D. Forney, Jr., "Generalized minimum distance decoding," IEEE Trans. Inform. Theory, vol.IT-12, pp.125-131, Apr. 1966; also Concatenated Codes. Cambridge, MA: M.I.T., 1966, ch. 3.

2 C. R. P. Hartmann and L. D. Rudolph, "An optimum symbol-by-symbol decoding rule for linear codes," IEEE Trans. Inform. Theory, vol.IT-22, pp.514-517, Sept. 1976.

3 J. K. Wolf, "Efficient maximum likelihood decoding of linear block codes using a trellis," IEEE Trans. Inform. Theory, vol.IT-24, pp.76-80, Jan. 1978.

4 T.-Y. Hwang, "Decoding linear block codes for minimizing word error rate," IEEE Trans. Inform. Theory, vol.IT-25, pp.733-737, Nov. 1979.

5 T.-Y. Hwang, "Efficient optimal decoding of linear block codes," IEEE Trans. Inform. Theory, vol.IT-26, pp.603-606, Sept. 1980.

6 L. D. Rudolph, C. R. P. Hartmann, T.-Y. Hwang, and N. Q. Duc, "Algebraic analog decoding of linear binary codes," IEEE Trans. Inform. Theory, vol.IT-25, pp.430-440, July 1979.

7 C. C. Yu and D. J. Costello, Jr., "Generalized minimum distance decoding algorithms for Qary output channels," IEEE Trans. Inform. Theory, vol.IT-26, pp.238-243, Mar. 1980.

8 C. E. Sundberg, "One-step majority-logic decoding with symbol reliability information," IEEE Trans. Inform. Theory, vol.IT-21, pp.236-242, Mar. 1975.

9 R. J. McEliece, The Theory of Information and Coding, Reading, MA : Addison-Wesley, 1977.

10 W. W. Peterson and E. J. Weldon, Jr., Error-Correcting Codes, second ed., Cambridge, MA : M.I.T., 1972.

11 G. Hadley, Nonlinear and Dynamic Programming, Readings, MA: Addison-Wesley, 1964.

12 M. R. Garey and D. S. Johnson, Computers and Intractability, San Francisco, CA : Freeman, 1979.

13 C. L. Liu, Introduction to Combinatorial Mathematics, New York, NY : McGraw-Hill, 1968.