Box Codes of Lengths 48 and 72

G. Solomon[1] and Y. Jin[2]

A self-dual code of length 48, dimension 24, with Hamming distance essentially equal to 12 is constructed here. There are only six codewords of weight 8. All the other codewords have weights that are multiples of 4 and have a minimum weight equal to 12. This code may be encoded systematically and arises from a strict binary representation of the (8,4;5) Reed–Solomon (RS) Code over GF(64). The code may be considered as six interrelated (8,7;2) codes. The Mattson–Solomon representation of the cyclic decomposition of these codes and their parity sums are used to detect an odd number of errors in any of the six codes. These may then be used in a correction algorithm for hard or soft decision decoding.

A (72,36;15) box code was constructed from a (63,35;8) cyclic code. The theoretical justification is presented herein.

A second (72,36;15) code is constructed from an inner (63,27;16) Bose–Chaudhuri–Hocquenghem (BCH) code and expanded to length 72 using box code algorithms for extension. This code was simulated and verified to have a minimum distance of 15 with even weight words congruent to 0 modulo 4. The decoding for hard and soft decision is still more complex than the first code constructed above.

Finally, an (8,4;5) RS Code over GF(512) in the binary representation of the (72,36;15) box code gives rise to a (72,36;16) code with nine words of weight 8, and all the rest have weights $\geq$ 16.*

I. Codes of Length 48

The self-dual (48,24;12) Quadratic Residue Code had a history of difficulty and complexity in decoding for five errors algebraically as well as decoding for soft decision. This led us to apply the techniques of box codes as successfully developed for Golay Codes to rate 1/2 codes of length 48. See [1,2]. Subcodes of dimension 23 and Hamming distance 12 were easily found. In addition, the box structure gave parity information to detect odd errors in rows that simplify decoding procedures.

The attempt to avoid the six codewords of weight 8 in the natural box code construction yielded two self-dual

[1] Independent consultant to the Communications Systems Research Section.

[2] Student at the California Institute of Technology, Pasadena, California.

(48,24;12) codes [1]. Upon closer examination of computer simulation, these codes were found to contain 42 words of weight 8.

In [1], the two codes constructed were designed to be self-dual. The (48,23;12) systematic subcodes of each were easily found. The 24th dimension in each was more elaborately constructed with the proviso that odd parities of the rows were induced to be used as tools in an erasure-error correcting decoding procedure. A search of the codeword weights' structure indicated the presence of 42 words of weights 8 and 40 in both these codes. The remaining nonzero words were of minimum weight 12. There exists a straight systematic construction of the Reed–Solomon (RS) (8,4;5) Code over $GF(64)$ for the 24th dimension given below, still using the particular binary representation in [1], which yields only six codewords of weights 8 and 40. This gives a box code with even parity on the rows. So for a low signal-to-noise ratio, this code and the previously constructed codes of dimension 48, rate 1/2, are effectively of minimum distance 12. The decoding procedure for soft decision mentioned in [1] is still applicable and preferred over any current soft decoding of the (48,24;12) Quadratic Residue Code.

In [2], a code of length 72 and distance 15 was constructed specifically to have simplified soft decoding. The (72,35;16) subcode was constructed with even parity on the nine rows in a nonsystematic manner as a subcode of the Reed–Solomon (8,4;5) Code over $GF(512)$. The 36th dimension was constructed to give odd parity on the rows and yield a code of minimum distance 15. The full code was designed to have a systematic encoding. This code, however, upon investigation, was found to have a very small number of words of length 11.

To meet this emergency, a new (72,36;15) box code is constructed here with rows of even or odd parity, and so it possesses, perhaps, a simple hard decision 7–8 error correcting procedure. This code has been simulated and verified to have a minimum distance of 15 and even weight words congruent to 0 modulo 4.

II. (8,4;5) RS Code Over *GF* (64)

Represent the Reed–Solomon (8,4;5) Code over $GF(64)$ in binary using the particular normal basis in [1]. One can generate a rate 1/2 self-dual code of length 48 and dimension 24 with weights that are multiples of 4.

This binary representation of the RS (8,4;5) Code over $GF(64)$ yields six (8,7;2) codewords whose decomposition via Mattson–Solomon into two cyclic code components and a constant component looks like (6,4;3) and (6,2;5) RS Codes over $GF(8)$ and a (6,6;1) binary code, respectively.

In particular, let γ be a root of the polynomial $f(x) = x^6 + x^5 + x^4 + x + 1$, where γ is a primitive generator of the 63 roots of unity. Represent the elements of $GF(64)$ in the normal representation using the roots of $f(x)$. The roots are γ^j; $j \in J$; $J = \{1, 2, 4, 8, 16, 32\}$.

NOTE: For this particular choice of $f(x)$, we have

$$\mathrm{Tr}\gamma^j = 1; \quad j \in J; \quad J = \{1, 2, 4, 8, 16, 32\}$$

$$\mathrm{Tr}(\gamma^i \gamma^k) = 0; \quad i \neq k; \quad i, k \in J$$

Let β be a root of the polynomial $g(x) = x^3 + x^2 + 1$. β is an element of $GF(8)$, a subfield of $GF(64)$, and $\beta = \gamma^9$.

A. Encoding

Now use the recursion or check polynomial $h(x) = \prod_{i=0}^{3}(x + \beta^i)$ to generate an extended (8,4;5) RS Code over $GF(64)$. This means that the initial shift register contains four elements in $GF(64)$ expressed as coefficients in the normal representation above. The cyclic portion of the code is of length 7, and the overall parity symbol is the eighth dimension. Represent the binary code as components $\mathrm{Tr}(P(x)\gamma^j)$; $j = 1, 2, 4, 8, 16, 32$.

The general Mattson–Solomon (M–S) polynomial of a codeword $\mathbf{a}$, similar to the Golay codeword over $GF(8)$, is $P_{\mathbf{a}}(x) = C_0 + C_1 x + C_2 x^2 + C_3 x^3$ where $C_i \in GF(64)$ for $0 \leq i \leq 3$ and $x \in GF(8)$.

Encode the codeword in its cyclic portion. The extended codeword $\mathbf{a}$ expressed in terms of the M–S polynomial is

$$\mathbf{a} = (P_{\mathbf{a}}(\beta^i); 0 \leq i \leq 6, P_{\mathbf{a}}(0))$$

Writing the codewords in binary using the normal basis γ^j, $j \in J$ above, there are six binary codewords of length 8:

$$\mathrm{Tr}(P(x)\gamma^j); \quad j = 1, 2, 4, 8, 16, 32$$

where $\mathrm{Tr}a$ denotes the value in $GF(2)$ given by the Trace of an element $a \in GF(64)$:

$$\mathrm{Tr}a = a + a^2 + a^4 + a^8 + a^{16} + a^{32}$$

Consider one of the six binary words in its Mattson–Solomon setting,

$$\mathrm{Tr}(P_{\mathbf{a}}(x)\gamma^j) = \mathrm{Tr}((C_0 + C_1x + C_2x^2 + C_3x^3)\gamma^j)$$

$$= \mathrm{Tr}(C_0\gamma^j) + \mathrm{Tr}'[(C_1x + C_2x^2 + C_3x^3)\gamma^j$$

$$+ ((C_1x + C_2x^2 + C_3x^3)\gamma^j)^8]$$

$$\mathrm{Tr}'a = a + a^2 + a^4; \quad a \in GF(8)$$

Set $C_0 = 0$ temporarily, as this does not affect the arguments to follow.

$$\mathrm{Tr}(P(x)\gamma^j) = \mathrm{Tr}'[(C_1\gamma^j + (C_1\gamma^j)^8 + (C_2\gamma^j)^4$$

$$+ (C_2\gamma^j)^{32})x + ((C_3\gamma^j)^2 + (C_3\gamma^j)^{16})x^6]$$

Lemma: The coefficient of x is a (6,4:3) code over $GF(8)$. The coefficient of x^3 is a (6,2;5) code over $GF(8)$. The code is indexed by the values of $\gamma^j; j \in J = \{1, 2, 4, 8, 16, 32\}$.

Proof: The set $\gamma^j; j \in J = \{1, 2, 4, 8, 16, 32\}$ is a linear independent set and thus can take zero values only one less than the number of terms in the coefficient of x, x^6. The term $\mathrm{Tr}(C_0\gamma^j)$ in the code's expression when $\mathrm{Tr}C_0 = 0$, the constant terms, forms a (6,5;2) binary code.

Theorem: The RS Code determined by codewords with M–S polynomials $P_{\mathbf{a}}(x); \mathrm{Tr}C_0 = 0$ forms a (48,23;12) binary code with weight multiples of 4.

Proof: The multiple of 4 property of the weights follows, using the Solomon–McEliece Γ_2 Formula.

$$\mathrm{Tr}(P(x)\gamma^j) = \mathrm{Tr}'[(C_1\gamma^j + (C_1\gamma^j)^8 + (C_2\gamma^j)^4$$

$$+ (C_2\gamma^j)^{32})x + ((C_3\gamma^j)^2 + (C_3\gamma^j)^{16})x^6]$$

where Tr is defined in $GF(64)$ and Tr′ is defined in $GF(8)$.

Now

$$\Gamma_2(\mathrm{Tr}P(x)\gamma^j) = \mathrm{Tr}'(C_1C_3^2\gamma^{3j} + C_1^8C_3^2\gamma^{10j} + C_1C_3^{16}\gamma^{17j}$$

$$+ C_1^8C_3^{16}\gamma^{24j} + C_2^{32}C_3^2\gamma^{34j} + C_2^4C_3^2\gamma^{6j}$$

$$+ C_2^{32}C_3^{16}\gamma^{48j} + C_2^4C_3^{16}\gamma^{20j})$$

$$= \mathrm{Tr}(C_1C_3^2\gamma^{3j} + C_1^8C_3^2\gamma^{10j} + C_2^4C_3^2\gamma^{6j}$$

$$+ C_2^4C_3^{16}\gamma^{20j})$$

and therefore $\sum_{j\in J}\Gamma_2(\mathrm{Tr}P(x)\gamma^j) = 0$.

Recall that the normal basis was chosen so that $\mathrm{Tr}\gamma^j = 1; j \in J\mathrm{Tr}(\gamma^i\gamma^k) = 0;\ i \neq k;$ and $i, k \in J$.

It has been demonstrated that the binary weight of any codeword in the RS Code above is a multiple of 4. Since the symbol distance of the code is greater than 5, we have narrowed the weights down to $8, 12, 16, 20, \cdots, 40$.

III. Structure of the Code

Using the same arguments given in [1], the minimum weight of the code for $\mathrm{Tr}C_0 = 0$ is equal to 12. However, note that for each $C_0 = \gamma^i;\ i = 1, 2, 4, 8, 16, 32$ and $C_i = 0;$ and $i = 1, 2, 3$, one obtains a codeword of weight 8. We proved that these six are the only words of weight 8. A counting argument on the weights would do the same. Since all words have weight multiples of 4, the code is self-dual.

IV. (72,36;15) Code

In [2], an alternate (72,36;15) box code was constructed from the (63,35;8) cyclic code, generated by the check polynomial $f(x) = \prod f_i(x); i = 1, 3, 5, 7, 9,\ 13, 21$ where $f_i(x)$ is a polynomial irreducible over $GF(2)$ with β^i as a root where β is a primitive 63rd root of unity. We now present the theoretical justification.

Place the codewords in the usual 9×7 box code matrices corresponding to their values $7i + 9j$ (mod 63) for $0 \leq i \leq 8,\ 0 \leq j \leq 6$. Let $z = xy$ where $x^7 = 1, y^9 = 1, x = \beta^{9j}$, and $y = \beta^{7i}$. Indexing the rows by y, the M–S polynomial for each row y is

$$P_y(x) = \mathrm{Tr}(C_1 z + C_3 z^3 + C_5 z^5 + C_7 z^7 + C_{13} z^{13})$$

$$+ C_9 z^9 + C_{18} z^{18} + C_{36} z^{36} + C_{21} z^{21} + C_{42} z^{42}$$

$$= C_{21} y^3 + C_{21}^2 y^6 + \mathrm{Tr}(C_7 y^7)$$

$$+ \mathrm{Tr}'[(C_9^4 + C_1 y + C_1^8 y^8)x + (C_5^4 y^2 + C_5^{32} y^7$$

$$+ C_3^{16} y^3 + C_3^2 y^6 + C_{13} y^4 + C_{13}^8 y^5)x^6]$$

where Tr is defined in $GF(64)$ and Tr′ is defined in $GF(8)$.

Thus the coefficient of x is a (9,3;7) code over $GF(8)$, and the coefficient of x^6 is a (9,6;4) code over $GF(8)$.

Construct an eighth column on the nine rows by the usual parity rule. The eighth column will have the same Γ_2 value as the original (63,35;8) code. Then we immediately have the following lemma:

Lemma 1: The extended box code is a (72,35;16) code.

Proof: Consider the Solomon–McEliece Formula. The sum Γ_2 over the nine rows and eight columns gives 0, showing that the weight of every codeword is a multiple of 4. The properties of the coefficients of x and x^6 for the subcode or dimension 27 imply the minimum weight of the entire code to be 16.

Now adjoin a vector of all ones to the original 9×7 matrix setting. This will make the rows have odd parity and will complement the column sums. It is easy to show that all odd-weight codewords have weights of the form $4m - 1$. We will prove that the minimum code distance is 15.

The degree of the Mattson–Solomon polynomial for the entire 63 length code is 56; the next highest degree is 52. From this and the properties of the coefficients of x and x^6, one can easily see that the weight of the inner cyclic codeword is less than or equal to 54. If the inner weight is 54, the nine-row weight patterns 6 6 6 6 6 6 6 6 6, 7 6 6 6 6 6 6 6 5, and 7 7 6 6 6 6 6 6 4 could generate codewords of weights less than 15. If the inner weight is 52, the weight pattern 6 6 6 6 6 6 6 6 4 could generate codewords of weights less than 15.

Lemma 2: For the original cyclic (63,35;8) code, none of the weight patterns above are possible.

Proof: Weight pattern 6 6 6 6 6 6 6 6 6 gives the sum Γ_2 to be 1. This is impossible.

Let

$$P(y) = C_9^4 + C_1 y + C_1^8 y^8$$

$$Q(y) = C_5^4 y^2 + C_5^{32} y^7 + C_3^{16} y^3 + C_3^2 y^6 + C_{13} y^4 + C_{13}^8 y^5$$

then

$$\Sigma_y P(y)Q(y) = 0$$

$$\deg(P^6(y) + Q(y)) = 7$$

If the weight pattern is 7 6 6 6 6 6 6 6 5, then $\Sigma_y P(y)Q(y) = 1 + \alpha \neq 0$ for some $\alpha \neq 1$; $\alpha \in GF(8)$.

If the weight patterns are 6 6 6 6 6 6 6 6 4 or 7 7 6 6 6 6 6 6 4, the polynomial $P^6(y) + Q(y)$ has eight zeros. Then $P^6(y) = Q(y)$. But $P(y)Q(y) = 0$ and $P^6(y) = Q(y)$ cannot give weight 4 for any row indexed by y.

Theorem: The box code is a (72,36;15) code, where the even-weight subcode is a (72,35;16) code with all codewords having weights of the form $4m$, and the odd-weight subcode is a (72,35;15) code with all codewords having weights of the form $4m - 1$. **QED**

V. (72,36;15) Alternate Code

One can construct the Bose–Chaudhari–Hocquenghem (BCH) (63,27;16) code generated by the check polynomial $f(x) = \prod f_i(x), i = 1, 3, 5, 9, 11$. The cyclic decomposition in the box code setting yields a (9,5;5) code over $GF(8)$ for the coefficient of x and a (9,4;6) code over $GF(8)$ for the coefficient of x^6. This does extend to a (72,36;15) code, too. In fact, this code has been simulated and verified to have a minimum distance of 15 with even weight words congruent to 0 modulo 4. If we try all possibilites for the check polynomial $g(x) = \prod f_i(x)$; $i = 7, 21$, which totals to 256 codewords, we are left with an inner BCH code that can algebraically correct 7 errors. This leaves soft decoding still very complex and unworkable.

Note that because of the Mattson–Solomon decomposition here, one may correct five errors easily by hard decision, but six or more take more trials. Similarly, a soft

decision would require $8^4 = 2^{12}$ trials. This requires more trials for both hard and soft decisions than the alternate code mentioned above in [2]. Note the advantage that the (9,3;7) code over $GF(8)$, the coefficient of x^6, has over the (9,4;6) code over $GF(8)$, the coefficient of x.

VI. (8,4;5) RS Code Over $GF(512)$

The (8,4;5) RS Code over $GF(512)$ in the binary representation of [2] gives rise to a systematic (72,36;16*) code with nine words of weight 8, and all the rest have weights ≥ 16. The normal basis consists of $\gamma^i; i = 2^j; 0 \leq j \leq 8$ with γ a root of $f(x) = x^9 + x^8 + x^6 + x^5 + x^4 + x + 1$. The proof that there are no words of weight 12 is a simple counting argument. We prove there are no words of weight 60 in the code of dimension 35 given by $C_0 = 0$.

Represent the elements of $GF(512)$ in the normal representation using the roots of $f(x)$. The roots are $\gamma^j; j \in J$; and $J = \{1, 2, 4, 8, 16, 32, 64, 128, 256\}$.

For this particular choice of $f(x)$, we have

$$\mathrm{Tr}\gamma^j = 1; \quad j \in J; \quad J = \{1, 2, 4, 8, 16, 32, 64, 128, 256\}$$

$$\mathrm{Tr}(\gamma^i \gamma^k) = 0; \quad i \neq k; \quad i, k \in J$$

Represent the binary code as components $\mathrm{Tr}(P(x)\gamma^i)$; $i = 1, 2, 4, 8, 16, 32, 64, 128, 256$, giving nine words of length 8.

Let β be a root of the polynomial $g(x) = x^3 + x^2 + 1$. β is an element of $GF(8)$, a subfield of $GF(512)$ and $\beta = \gamma^{73}$.

A. Encoding

Now use the recursion or check polynomial $h(x) = \prod_{i=0}^{3}(x + \beta^i)$ to generate an extended (8,4;5) RS code over $GF(512)$. This means that the initial shift register contains four elements in $GF(512)$ expressed as coefficients in the normal representation above. The cyclic portion of the code is of length 7, the overall parity symbol; the eighth dimension is the usual sum over the seven symbols.

In the (72.36;16*) binary representation of the RS (8,4;5) Code over $GF(512)$, any codeword with the coeffiecients of x and x^6 nonzero has a minimum weight of 16. When these are zero, then clearly there are only nine words of weight 8, which come from the encoding of the symbol γ^j; $j \in J$; $J = \{1, 2, 4, 8, 16, 32, 64, 128, 256\}$. A similar proof would argue that there are only six words of weight 8 in the binary representation of the (48,24;12*) RS Code over $GF(64)$.

To prove there are no words of weight 12, a counting argument notes that there are no words of weight 60 in the even-weight codes, where $\mathrm{Tr}C_0 = 0$. Words of weight 60 must possess a weight distribution over the nine words in any permutation of 888666666. This implies that three rows are zero and six rows are nonzero with weights 6 or $\Gamma_2 = 1$, to say the least. However, in [2], we note that this cannot be. The cyclic coefficients are (9,3;7) and (9,6;4) codes over $GF(8)$, so there are at least seven rows with $\Gamma_2 = 1$. **QED**

References

[1] G. Solomon, "Self-Dual (48,24;12) Codes," *The Telecommunications and Data Acquisition Progress Report 42-111, vol. July–September 1992*, Jet Propulsion Laboratory, Pasadena, California, pp. 75–79, November 15, 1992.

[2] G. Solomon, "A (72,36;15) Box Code," *The Telecommunications and Data Acquisition Progress Report 42-112, vol. October–December 1992*, Jet Propulsion Laboratory, Pasadena, California, pp. 19–21, February 15, 1993.