1 BINARY EXTENDED GOLAY CODE

The Binary Extended Golay Code, simply referred to in this as the ‘Golay Code’, is a linear code of length 24 and dimension 12. The G24 Code takes in a 12 bit binary message and produces a 24 bit codeword. The minimum Hamming distance between G24 codewords is 8, meaning the code can correct up to 3 errors.

The Golay Code generator matrix G, shown below, is the augmented matrix consisting of a 12x12 identity matrix and a 12x12 parity check matrix. With the exception of the first row, the parity check matrix is cyclic, meaning all rows are the result of a 1-cycle permutation of the row above. Since the Golay Code is linear, its generator matrix consists of basis codewords, and the multiplication of this matrix by a column vector will produce another codeword.

$$G = \left[\begin{array}{cccccccccccc|ccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & \cdots & 1
\end{array}\right]$$

While a linear code with the parameters described (length 24, dimension 12, minimum distance 8) was hypothesized to exist, the following is proof that the Golay Code satisfies these parameters, specifically showing that all of its codewords have a minimum distance of 8:

2 PROOF OF THE GOLAY CODE MINIMUM DISTANCE

2.1 MINIMUM DISTANCE OF A LINEAR CODE

For a linear code, it can be proven that the weight of a given codeword is equal to the codeword’s minimum distance(1). This means that, to show that all G24 codewords have a minimum distance of 8, it is sufficient to instead show that all codewords have a minimum weight of 8.

2.2 THE WEIGHT OF THE PRODUCT OF ANY TWO CODEWORDS IS EVEN

It can be shown through computation that the weight of the product of any two basis codewords is even. Furthermore, since the code is linear, any two codewords x, y can be written as a sum of the basis codewords (where a_i and c_i are some coefficient 0 or 1).

x = $a_1b_1 + \ldots a_{12}b_{12}$

y = $c_1b_1 + \ldots + c_{12}b_{12}$
xy = $(a_1b_1)(c_1b_1) + \ldots + (a_1b_1)(c_{12}b_{12}) + \ldots + (a_{12}b_{12})(c_1b_1) + \ldots + (a_{12}b_{12})(c_{12}b_{12})$
xy = $(b_1b_1)(a_1c_1) + \ldots + (b_1b_{12})(a_1c_{12}) + \ldots + (b_{12}b_1)(a_{12}c_1) + \ldots + (b_{12}b_{12})(a_{12}c_{12})$

Since the product of any two basis codewords is even and each term in the expansion of the product of x and y contains the product of basis codewords, the product of x and y is a sum of even terms, which must also be even. Therefore, the weight of any two arbitrary codewords is even.

2.3 THE WEIGHT OF ANY CODEWORD IS DOUBLY EVEN

The formula for the weight of a codeword formed from two basis codewords is $w(b_1 + b_2) = w(b_1) + w(b_2) - 2(b_1b_2)$. Since we know the weight of a basis codeword is 8 and the weight of any two codewords multiplied together is even, this means $w(b_1 + b_2)$ is 16 minus some even number times 2, which means a factor of 4 can be pulled from the expression. This case can be generalized to the weight of the sum of any number of basis vectors, meaning all possible codewords are divisible by 4.

2.4 THERE ARE NO CODEWORDS OF WEIGHT 4

Let L be the leftmost digits of the codeword produced by the identity matrix (indices 1-12) and let R be the rightmost digits of the codeword produced by the parity check matrix (indices 13-24). In order for a codeword x to have a weight of 4, and since $w(x) = w(L)+w(R)$, one of the following cases must occur:

CASE 1: $w(L) = 0, w(R) = 4$
This case is impossible. Since L is produced by multiplication with the identity matrix, the only vector with the property w(L) = 0 is the zero vector itself, meaning w(R) would also equal 0.

CASE 2: $w(L) = 1, w(R) = 3$
If w(L) = 1, the codeword begins with a row of the identity matrix, and is therefore a row of the generator matrix. By simple observation of G, we can confirm there are no basis codewords of weight 3, making this case impossible.

CASE 3: $w(L) = 2, w(R) = 2$
Similar to case 2, if w(L) = 2, then the codeword is the linear combination of two rows of the generator matrix. Through computation, we can confirm that adding two rows of the generator matrix will never produce a codeword with w(R) = 2.

CASE 4: $w(L) = 3, w(R) = 1$
Similar to case 2, if w(R) = 1, then the codeword is a row of the reversed augmented matrix $P|I$ (rather than $I|P$). No basis codeword belonging to P has a

weight of 3, so this case is impossible.

CASE 5: $w(L) = 4, w(R) = 0$
Similar to case 1, w(R) = 0 only if the codeword is the zero vector, which would necessitate that w(L) = 0 as well.

Since all of these cases proved to be impossible, we can conclude that there are no codewords of weight 8.

2.5 CONCLUSION

Following the steps of the proof, we can see from section 2.3 that the weight of any codeword is divisible by four, and from section 2.4, there exist no codewords of weight 4. This means that the weight of all codewords (excluding the zero vector) is a minimum of 8. Because this is a linear code, the weight of a codeword is equal to the minimum distance of that codeword, so the minimum distance of all Golay codewords is 8.

3 DECODING

While there are many methods for decoding Golay codewords, the method implemented in the example code is syndrome decoding. The parity check matrix of the code, H, is designed with the property that any codeword c, when multiplied with H, will produce 0. This means that a codeword containing errors can easily be identified, since it will produce some nonzero vector when multiplied by H. This syndrome can be used to identify generally which bits contain an error, regardless of what the received message was. Upon identifying the bits containing an error, this ‘error pattern’ is added back to the message to flip the bits that had an error.

4 SOURCES

- Hill, Raymond. ”A First Course in Coding Theory”. Clarendon Press, 1986.
- Truong, T.K et al. ”Decoding of 1/2-Rate (24,12) Golay Codes”. NASA, 1989.
- ”Linear Block Codes: Encoding and Syndrome Decoding”. MIT OpenCourseWare, 2012.
- ”The Hidden Geometry of Error-Free Communication”. YouTube, uploaded by Another Roof, 2023. https://www.youtube.com/watch?v=Tmx-v4FiP6I