Towards Efficient Encoding Using Polynomials

David Mitchell, Joint Research Institute for Signal & Image Processing, The University of Edinburgh
Liam O'Carroll, School of Mathematics and Maxwell Institute for Mathematical Sciences, The University of Edinburgh
Norbert Goertz, Joint Research Institute for Signal & Image Processing, The University of Edinburgh
{David.Mitchell, L.O'Carroll, Norbert.Goertz}@ed.ac.uk

Abstract

We present an efficient new method of obtaining a generator matrix G from certain types of parity check matrices with a defective cyclic block structure. This novel approach describes parity check and generator matrices in terms of polynomials. Moreover, by using this polynomial algebra we have found efficient ways to implement the scheme. In addition, this method is as such interesting as it allows us to convert H into G without a systematic encoder in between (i.e., there is no diagonal subpart in the output). This alone is interesting as normally G would be dense if we were to form it from the given H by Gaussian elimination.

1 Introduction

We begin by introducing the notion of generator and parity check polynomials. To begin, we define a cyclic code.

Definition. A linear block code C of length n over some field F is called **cyclic** if for any codeword $(c_0 c_1 \dots c_{n-1}) \in C$ then we also have $(c_{n-1} c_0 c_1 \dots c_{n-2}) \in C$.

Let C be a cyclic block code over a field F with block length n. There exists a unique monic polynomial $g(t)$ over the field F such that $g(t)$ *generates* C, denoted by $C =< g(t) >$. By this it is meant that C is viewed as a subcode of the factor ring $R_n = F[t]/(t^n - 1)$ (considered as a vector space over F) and in R_n the code C is generated as an ideal by $g(t)$ [1, pp. 188-190], and hence we call $g(t)$ a *generator polynomial* of C. Let $g(t) = g_0 + g_1 t + \dots + g_{n-k} t^{n-k}$ where $0 < k < n$. Then the generator matrix G for the code C is formed using $g(t)$ in the following way:

$$G = \begin{bmatrix} g_0 & g_1 & \cdots & g_{n-k} & & & \\ & \ddots & \ddots & & & \ddots & \\ & & g_0 & g_1 & & \cdots & g_{n-k} \end{bmatrix}_{k\times n} . \qquad (1)$$

Continuing in this vein, set $h(t) = (t^n - 1)/g(t)$. We call $h(t)$ the *check polynomial*. Suppose $h(t) = h_0 + h_1 t + \dots + h_k t^k$. Then it can be shown [1, pp. 194-196] that a parity check matrix H for C is given by:

$$H = \begin{bmatrix} & & h_k & \dots & h_1 & h_0 \\ & \cdot^{\cdot^{\cdot}} & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \\ h_k & \dots & h_1 & h_0 & & \end{bmatrix}_{(n-k)\times n} .$$

Note that $g_{n-k} = h_k = 1$, $g_0 = h_0 = 1$. This implies the code C has dimension[1] k and the dual code $C^{\perp}$ has dimension $n-k$.

1.1 Aims and Structure

We wish to use the properties of cyclic codes where it may not seem appropriate to do so. As discussed above, we saw that to use the polynomials $g(t)$ and $h(t)$, their product must be $t^n - 1$, where n is the block length of the code. Often, and in particular when dealing with LDPC codes (see, e.g., [2]), we have a parity check matrix which appears to have cyclic structure or be built of blocks with this structure, where the corresponding polynomial $h(t)$ does not satisfy the condition[2] $h(t) | t^n - 1$ for such a code length. We call such a parity check matrix *defective*.

In addition to this, in order to have better codes (see again [2]), we wish to work with a parity check matrix made up of multiple blocks

$$H = \begin{bmatrix} H_1 & H_2 & \dots & H_s \end{bmatrix}^{\mathrm{T}},$$

where each H_i is apparently the parity check matrix of a cyclic code, the caveat 'apparently' indicating that H_i is defective (or may be defective) in the above sense. Of course, to have cyclic blocks would be ideal. However we must find a way of dealing with the issue of the blocks being defective. As the parity check matrix H is made up of multiple blocks, we have multiple parity check polynomials which will most likely not divide the relevant $t^n - 1$.

Our novel approach is as follows: we begin by considering a parity check matrix made up of a single block in cyclic form, yet which is defective. Next we introduce an extension method which efficiently finds a generator matrix even when the check polynomial does not divide the corresponding $t^n - 1$. Then, in Section 3, an approach is detailed for a parity check matrix made up of two blocks in cyclic form. Section 4 introduces a general method of dealing with an arbitrary number of blocks.

2 The Extension Method

In this section we consider the case where we have a parity check matrix H that 'looks cyclic'

$$H = \begin{bmatrix} & h(t) \\ \swarrow & \end{bmatrix}_{(n-k)\times n} ,$$

[1] When a code has 2^k codewords we refer to k as the dimension of the code [1, p. 6].

[2] By the notation $h(t)|(t^n - 1)$ it is meant that $h(t)$ divides $t^n - 1$ without remainder. If this does not hold, we write $h(t) \nmid (t^n - 1)$. Since there is no danger of ambiguity we abreviate notation by $h(t)|t^n - 1$.

however, $h(t) \nmid t^n - 1$. The proposed method of 'extension' outputs a generator matrix for the code using polynomial division and the truncation of an interim matrix. This can be implemented efficiently using a Computer Algebra package such as *Maple*. The method begins by finding an $n^* > n$ such that $h(t)$ divides $t^{n^*} - 1$.

We must first show that such an n^* exists and how to find it. Later we will see that finding and working with a large n^* is not necessary, but it is essential to show its existence. That this is indeed the case is the first new result shown in Section 2.1 using two standard facts. We can then use this to prove the main new Theorem in Section 2.2. In Section 2.3, we propose a practical refinement of this Theorem 4 that enables it to be implemented efficiently.

This optimization allows us to obtain the desired generator matrix directly via polynomial division, so avoiding the theoretical values displayed in Sections 2.1 and 2.2. In the case of a single defective block, standard methods using Gaussian elimination are also easy to implement due to the cyclic structure of the matrix. However, when we consider stacked blocks (Sections 3 and 4) this new method is a very efficient alternative.

2.1 Finding a Suitable n^*

Suppose $h(t)$ is an irreducible polynomial of degree $k > 1$ over GF(2). Then we create an extension field GF(2^k) in the usual way [3]. This process produces an extending element α which is a root of $h(t)$. If we throw away the zero element from this field and consider GF$(2^k)\backslash\{0\}$, we are left with an abelian group under multiplication. We now use:

Theorem 1 (Lagrange [4]) *The order*[3] *of an element g of a finite group G divides the order $|G|$ of G.*

So $g^{|G|} = 1$. Hence g is a root of the polynomial $f(t) = t^{|G|} - 1$. Let $G = \text{GF}(2^k)\backslash\{0\}$, so that $|G| = 2^k - 1$. Suppose $g = \alpha$ where α is a root of $h(t)$. Since $f(\alpha) = 0$, it follows that $h(t)|f(t)$.

So for an irreducible polynomial $h(t)$ of degree $k > 1$ over GF(2), we can immediately write down an $n^* = 2^k - 1$ such that $h(t)|t^{n^*} - 1$. If $h(t)$ is not irreducible, we proceed by splitting the polynomial into irreducible parts. We will see that there is a systematic way of choosing n^* which depends on the structure of $h(t)$. To avoid trivialities we assume that $t \nmid h(t)$, this avoids zero columns in H. Let us first recall a required result on the divisors of certain polynomials.

Theorem 2 [1, pp. 99–106] *Over any field, $x^s - 1|x^r - 1$ if and only if $s|r$.*

If we can reduce $h(t)$ there are two possible cases:

Case (1)*: No repeated irreducible factors*

In this case, suppose $h(t) = q_1(t)q_2(t)\dots q_s(t)$, where all the $q_i(t)$ are distinct and irreducible over GF(2). Here we apply the above procedure to each of the s factors. Thus for each factor we observe that $q_i(t)|t^{n_i^*} - 1$ for a particular n_i^*. Now, using Theorem 2, we observe

$$t^{n_i^*} - 1|t^{n^*} - 1 \text{ if and only if } n_i^*|n^*.$$

Thus if we set $n^* = \text{lcm}(n_1^*, n_2^*, \dots, n_s^*)$, $h(t)|t^{n^*} - 1$, as each of the distinct factors $q_i(t)$ of $h(t)$ divide $t^{n^*} - 1$.

Case (2)*: Repeated irreducible factors*

If there are repeated factors, then $h(t) = q_1^{r_1}(t)q_2^{r_2}(t)\dots q_s^{r_s}(t)$ where all the $q_i(t)$ are distinct and irreducible and at least one of the $r_i \geq 2$. We observe that when working modulo 2

$$(t^k - 1)^2 = t^{2k} + 1.$$

Suppose we write $t^k - 1 = u_1(t)u_2(t)\dots u_p(t)$, where each of the $u_i(t)$ is irreducible over GF(2). Then

$$\begin{aligned}(t^k - 1)^2 &= t^{2k} + 1 = & u_1^2(t)u_2^2(t)\dots u_p^2(t),\\ (t^k - 1)^4 &= t^{4k} + 1 = & u_1^4(t)u_2^4(t)\dots u_p^4(t),\\ &\vdots\end{aligned}$$

In general $t^{2^i k} + 1 = u_1^{2^i}(t)u_2^{2^i}(t)\dots u_p^{2^i}(t)$. So if $q_i(t)$ is a factor of $t^k - 1$ then $q_i^{r_i}(t)$ will be a factor of $t^{2^i k} + 1$ if $2^i \geq r_i$.

Thus the approach in this case is as follows. For each of the irreducible factors $q_i(t)$ we find an n_i^* (in the usual way) such that $q_i(t)|t^{n_i^*} - 1$ as in case (1). Set $k = \text{lcm}(n_1^*, n_2^*, \dots, n_s^*)$. Let $r = \max(r_1, r_2, \dots, r_s)$. Using the above rule we find the least such integer c so that $2^c \geq r$. Finally we set $n^* = k \cdot 2^c$. Then $h(t)|t^{n^*} - 1$. Hence we have proved our first new result.

Theorem 3 *In the above notation there exists a positive integer n^* such that $h(t)|t^{n^*} - 1$.*

2.2 Finding a Generator Matrix in the Defective Case

Now we know that an n^* exists, we can state an algorithm to obtain a generator matrix:

1. we begin with a defective parity check matrix H for a code of block length n, in the sense that H 'looks' cyclic (with corresponding check polynomial $h(t)$) but $h(t) \nmid t^n - 1$;
2. we find a larger n^* such that $h(t)|t^{n^*} - 1$;
3. we then calculate $g(t) = (t^{n^*} - 1)/h(t)$;
4. the generator matrix associated to $g(t)$ is truncated from the left so as to have n columns. This new matrix is a generator matrix of the original code.

[3] Recall for any element g from a group G if there exists a positive integer m such that $g^m = e$ (where e is the identity element from the group) then the smallest such positive integer is called the *order* of g. If no such m exists we say g is of *infinite order*. The order of a group G is simply the number of elements in the group: this is denoted by $|G|$.

Theorem 4 *The matrix so formed is a generator matrix of the original code C.*

The proof of this can be split into three parts. Throughout we will denote the check polynomial by $h(t) = h_0 + h_1 t + \ldots + h_k t^k$ with $h_k = h_0 = 1$, where $0 < k < n$.
Claim 1. The matrix so formed has the appropriate rank and number of columns to be a generator matrix for C.
Proof. Suppose the parity check matrix H is a $(n-k) \times n$ matrix with cyclic shift structure. This matrix is necessarily of full rank, since H has a backwards echelon structure with $h_k \neq 0$. Let G be a generator matrix for C. Then G is a $k \times n$ matrix with rank k.
Note that the degree of the parity check polynomial is k, as we require $n-k$ rows in H. Using $t^{n^*} - 1$, we calculate the generator polynomial $g(t)$ of the corresponding cyclic code as $g(t) = (t^{n^*} - 1)/h(t)$. Thus the degree of the generator polynomial $g(t)$ of this extended matrix is $n^* - k$. Recall from Section 1 that the corresponding generator matrix G' is formed by 'sliding' the coefficients of $g(t)$ diagonally downwards as to to fill up the matrix (as in (1)). Thus we calculate

$$\text{row rank } G' = n^* - (n^* - k) = k = \text{ row rank } G.$$

Now set G'' to be G' with the first n^*-n columns removed. Then G'' is an $k \times n$ matrix. The final necessary condition is that G'' must be of full rank. Recall the generator polynomial necessarily has non-zero coefficient g_{n^*-k}. This creates an echelon structure in the cyclic generator matrix G'. We note that the truncation length n is greater than k (from construction of the check matrix), thus the echelon structure of the cyclic generator matrix G' is preserved in this new matrix G''. □
Claim 2. $G''H^T = 0$.
Proof. Let H' be the $(n^* - k) \times n^*$ parity check matrix of the cyclic code of block length n^* corresponding to the parity check polynomial $h(t)$. So

$$H' = \begin{bmatrix} 0 & & h(t) \\ & \swarrow & \\ & & \end{bmatrix}_{(n^*-k)\times n^*},$$

Then $G'H'^T = 0_{(k\times(n^*-k))}$ by construction. Now create a new matrix H'' which is equal to the first $n-k$ rows of H',

$$H'' = \begin{bmatrix} 0_{(n-k)\times(n^*-n)} & H \end{bmatrix}_{(n-k)\times n^*}.$$

Observe that $G'H''^T = 0_{k\times(n-k)}$.
By construction, the first $n^* - n$ columns of H'' are all zero. Thus they are irrelevant in the calculation. Remove the first $n^* - n$ columns from G' and H'', making G'' and H''' respectively. Observe that

$$H''' = \begin{bmatrix} & h(t) \\ \swarrow & \\ & \end{bmatrix}_{(n-k)\times n} = H.$$

Thus, $G''H'''^T = G''H^T = 0_{k\times n}$. □
Claim 3. Rowspace G'' = rowspace G.
Proof. As $G''H^T = 0$ and $GH^T = 0$ we know that each word in G'' and G is perpendicular to each word in H. Thus

$$\begin{aligned} \text{rowspace } G'' &\subseteq \text{ null space } H, \\ \text{rowspace } G &\subseteq \text{ null space } H. \end{aligned}$$

The null-space of the $(n-k) \times n$ parity check matrix H (of full rank) must be of dimension k. As both G and G'' have rank k it follows that rowspace G'' = rowspace G. □
A Toy Example. Suppose $h(t) = t^3 + t + 1$ and $n = 5$. We observe that $h(t)$ is irreducible over GF(2) and that $h(t) \nmid t^5 - 1 (\text{mod } 2)$, though the corresponding defective parity check matrix H looks cyclic:

$$H = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}_{2\times 5}.$$

We find that $n^* = 7$ so $h(t) | t^7 - 1 (\text{mod } 2)$. Then the generator polynomial for the extended cyclic code is

$$g(t) = (t^7 - 1)/(t^3 + t + 1) = t^4 + t^2 + t + 1.$$

So the extended matrices are

$$\begin{aligned} H' &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}_{4\times 7} \quad \text{and} \\ G' &= \left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{array} \boxed{\begin{array}{ccccc} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{array}}\right]_{3\times 7}. \end{aligned}$$

Thus by construction $G'H'^T = 0_{3\times 4}$. Recall that the next step is to set H'' equal to the first $n-k$ rows of H'

$$H'' = \begin{bmatrix} 0_{2\times 2} & H \end{bmatrix}_{2\times 7}.$$

Clearly now $G'H''^T = 0_{3\times 2}$. The final step is to observe there will be no contribution from the first $n^* - n = 2$ columns. Removing these columns from G' results in G'' (which is the boxed area of G'). It should now be evident that $G''H^T = 0_{3\times 2}$ and hence G'' is a generator matrix for the code because of an argument using rank (as above), or by direct calculation here.

2.3 Optimizing the Method

It was noted earlier that desired n^* could be very large and hence polynomial division would be very time-consuming. Depending on the desired block length n, we can optimize the method to save on calculations. The generator matrix of the extended code is formed by sliding the coefficients of the generator polynomial diagonally downwards (as in (1)). Then, as we are truncating this extended matrix, some of the coefficients may be surplus to requirements. This is best observed in an example.

Example. Let $h(t) = t^{34} + t^6 + t^3 + 1$. We observe $h(t)$ is reducible; explicitly,

$$\begin{aligned} h(t) &= (t^7 + t^5 + t^4 + t^3 + 1)(t^6 + t^5 + t^4 + t + 1) \\ &\quad (t + 1)(t^{20} + t^{18} + t^{10} + t^9 + t^6 + t^5 + t^4 + t^2 + 1). \end{aligned}$$

The factors here are all irreducible. Note that there are no repeated factors, hence $h(t)$ is a *Case* (1) polynomial. We calculate

$$n^* = \text{lcm}(2^7 - 1, 2 - 1, 2^6 - 1, 2^{20} - 1) = 2796549525.$$

Suppose $n = 37$. (Typically we will choose a block length and this will define the rank of our generator matrix, since the degree of $h(t)$ is given.) In this case, as the degree of our check polynomial is $k = 34$, H is a 3×37 matrix and G will be a 34×37 generator matrix. However we observe G' looks as follows, where the entries are the coefficients of $g(t)$:

$$G' = \left[\begin{array}{cc} \star & \boxed{\begin{matrix} g_{n^*-37} & \cdots & g_{n^*-2} & g_{n^*-1} \\ g_{n^*-38} & g_{n^*-37} & \cdots & g_{n^*-2} \\ \vdots & \vdots & & \vdots \\ g_{n^*-70} & g_{n^*-69} & \cdots & g_{n^*-34} \end{matrix}} \end{array} \right],$$

where

$$\star = \begin{pmatrix} g_0 & g_1 & & \cdots & & g_{n^*-39} & g_{n^*-38} \\ 0 & g_0 & g_1 & & \cdots & & g_{n^*-39} \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ 0 & \cdots & 0 & g_0 & g_1 & \cdots & g_{n^*-71} \end{pmatrix}.$$

Observe that the boxed section is the generator matrix G that we want. Thus the majority of the matrix is irrelevant. So to write down G we only require the coefficients of $t^{n^*-1}, \ldots, t^{n^*-70}$ from the generator polynomial $g(t)$. The division to obtain $g(t)$ looks like

$$\begin{aligned} g(t) &= \frac{t^{n^*} - 1}{t^{34} + t^6 + t^3 + 1} \\ &= t^{n^*-34} + t^{n^*-62} + t^{n^*-65} + t^{n^*-68} + t^{n^*-90} + \ldots. \end{aligned}$$

Thus there are only 4 non-zero coefficients to fill in to obtain the boxed area of G', since coefficients of terms of degree $n^* - 71$ or less are in the non-boxed section of G'. For the purpose of obtaining the necessary number of non-zero coefficients we can work with a smaller value of n^*, denoted $\tilde{n}$. Using this smaller value of n^* will result in a remainder which can be discarded. So we will have

$$t^{\tilde{n}} - 1 = q(t)h(t) + r(t), \tag{2}$$

where the polynomial $q(t)$ has the correct coefficients to form the desired generator matrix G (the boxed part of G'). There is a limit to how small we can choose $\tilde{n}$ to be so that $q(t)$ still has the correct coefficients. We observe in our example we require $\tilde{n} \geq 70$ to include the coefficient of $t^{\tilde{n}-70}$. Observe that (2) in this case becomes

$$t^{70} - 1 = (t^{36} + t^8 + t^5 + t^2)(t^{34} + t^6 + t^3 + 1) + t^{14} + t^8 + t^2 + 1.$$

Thus $q(t) = t^{36} + t^8 + t^5 + t^2$. Recall the non-zero coefficients from G' are $t^{n^*-34}, t^{n^*-62}, t^{n^*-65}$ and t^{n^*-68}. Observe these coefficents match up when substituting $n^* = 70$. So for $\tilde{n} = 70$

$$G = \begin{bmatrix} q_{33} & \cdots & q_{68} & q_{69} \\ q_{32} & q_{33} & \cdots & q_{68} \\ \vdots & \vdots & & \vdots \\ q_0 & q_1 & \cdots & q_{34} \end{bmatrix}.$$

Thus the optimized procedure established above can be summarised in the following new Theorem.

Theorem 5 *In the above notation given a block length n and a check polynomial of degree k then $\tilde{n} = n + k - 1$.*

3 Finding a Generator Matrix in the Case of Two Defective Blocks

3.1 Setup

Consider a parity check matrix of the form

$$H = \begin{bmatrix} H_1 & H_2 \end{bmatrix}^T, \tag{3}$$

where H_1 and H_2 are in 'cyclic' form, but defective. By that it is meant that the submatrices appear as the form specified in Section 1 but the associated parity check polynomials do not both divide $t^n - 1$, where n is the number of columns of H. Let C_1 and C_2 be the codes formed by the null space of the matrices H_1 and H_2 respectively. Then we define the concatenated code [5]:

$$C_1 \dotplus C_2 = \{(c_1 : c_2) \,|c_1 \in C_1 \text{ and } c_2 \in C_2\}.$$

We define the modulo 2 addition of two codes

$$C_1 \oplus C_2 = \{c_1 + c_2 \text{ mod } 2 \,|c_1 \in C_1 \text{ and } c_2 \in C_2\}.$$

Consider the linear mapping $(c_1|c_2) \to c_1 \oplus c_2$ of the concatenated code $C_1 \dotplus C_2$ to the modulo-2 sum of its two component codes. Using the conservation law of dimensions [6] we have

$$\dim(C_1 \dotplus C_2) = \dim(C_1 \oplus C_2) + \dim(C_1 \cap C_2). \tag{4}$$

Notice that if $C_1 \cap C_2 = \{0 \ldots 0\}$ then $\dim(C_1 \dotplus C_2) = \dim(C_1 \oplus C_2)$. Let H_1 and H_2 have corresponding generator matrices G_1 and G_2. We are interested in the null space of the parity check matrix (3); this is the intersection of the null spaces C_1 and C_2. To get a nice result on this code we will see that we can use the method of extension to bring H into a form we can usefully work with.

We have that C_1 and C_2 are the rowspaces of G_1 and G_2 respectively. It is easily seen that

$$C_1 \oplus C_2 = \text{rowspace} \begin{bmatrix} G_1^T & G_2^T \end{bmatrix}^T.$$

Using the method of extension, we know that for both the check polynomials $h_1(t)$ and $h_2(t)$ associated with the blocks H_1 and H_2 there exists an integer n_i^* such that $h_i(t)|t^{n_i^*}-1$, $i=1,2$. We again suppose that $t \nmid h_i(t)$, $i=1,2$. Let $n^* = \text{lcm}\,(n_1^*, n_2^*)$, then $h_i(t)|t^{n^*}-1$, $i=1,2$. We extend (3) to an $((n-k_1)+(n-k_2))\times n^*$ matrix, which consists of two cyclic blocks of size $(n^*-k_i^*)\times n^*$, $i=1,2$, where k_i is the degree of polynomial $h_i(t)$. Now we can calculate the generator polynomials $g_1(t)$ and $g_2(t)$ associated with the two extended blocks and form the respective generator matrices in the usual way

$$G_i' = \begin{bmatrix} g_i(t) & & \\ & \searrow & \\ & & \end{bmatrix}_{k_i\times n^*} .$$

Let the extended codes formed from these matrices be denoted C_1' and C_2' respectively. Note that the sum of two codes is again cyclic and has generator polynomial the greatest common divisor of the individual generator polynomials.

We now introduce notation for the truncation operation. Suppose a code C' of length $n^* > n$ is truncated from the left to be of length n. We denote the new code C (the right hand side of the original code) by $C = C'|_{\text{RHS}}$. Following the work done in Section 2, it should be evident that truncating the sum of extended codes (from the left) to be of length n will result in the sum of the required codes. That is,

$$(C_1' \oplus C_2')|_{\text{RHS}} = C_1 \oplus C_2.$$

Of course we are interested in the intersection of the codes C_1 and C_2. We have

$$(C_1' \cap C_2')|_{\text{RHS}} \subseteq C_1'|_{\text{RHS}} \cap C_2'|_{\text{RHS}} . \qquad (5)$$

This is easily seen by choosing a codeword c that exists in the intersection $C_1' \cap C_2'$. Suppose we split c into left and right parts, so $c = l : r$ with the part r of length n. Then

$$l : r \in C_1' \cap C_2' \Rightarrow r \in C_1'|_{\text{RHS}} \cap C_2'|_{\text{RHS}} .$$

Note that if we can show equality in the dimensions of the two sides of (5) then the two spaces must be equal because one is contained in the other. So we proceed by examining the dimensions in (5). We immediately get the inequality

$$\begin{aligned} \dim\big((C_1' \cap C_2')|_{\text{RHS}}\big) \quad &\leq \quad \dim\big(C_1'|_{\text{RHS}} \cap C_2'|_{\text{RHS}}\big) \\ &= \dim\,(C_1 \cap C_2)\,. \end{aligned} \qquad (6)$$

We can write the left hand side of (6) as

$$\begin{aligned} &\dim\big((C_1' \cap C_2')|_{\text{RHS}}\big) \\ &= \dim\,(C_1' \cap C_2') \text{ for a suitably large}^4\ n\,, \\ &= \dim\,(C_1' \dotplus C_2') - \dim\,(C_1' \oplus C_2') \text{ from } (4), \\ &= \dim\,(C_1') + \dim\,(C_2') - \dim\,(C_1' \oplus C_2')\,, \end{aligned}$$

where in the last line we use the fact that the dimension of the concatenated codes is equal to the sum of the individual dimensions of the codes [5]. Using this property and (4) we now write the right-hand side of (6) as

$$\dim\,(C_1 \cap C_2) = \dim\,(C_1) + \dim\,(C_2) - \dim\,(C_1 \oplus C_2)\,.$$

As the rank of the generator matrices of the original and the extended codes are equal, the dimensions of the codes must be equal ($\dim(C_i) = \dim(C_i')$). The upshot is that we can now re-write (6) as

$$\dim\,(C_1' \oplus C_2') \leq \dim\,(C_1 \oplus C_2)\,. \qquad (7)$$

Now, considering the modulo 2 sum of two codes as rowspaces of stacked generator matrices, we observe

$$\begin{aligned} C_1 \oplus C_2 \quad &= \quad \text{rowspace} \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} = \text{rowspace} \begin{bmatrix} G_1' \\ G_2' \end{bmatrix}_{\text{RHS}} \\ &= \quad (C_1' \oplus C_2')|_{\text{RHS}}\,. \end{aligned}$$

So as $(C_1' \oplus C_2')|_{\text{RHS}} = C_1 \oplus C_2$ their dimensions must be equal. Thus (7) becomes

$$\dim\,(C_1' \oplus C_2') \leq \dim\big((C_1' \oplus C_2')|_{\text{RHS}}\big)\,.$$

Note that by design the truncation (block) length n is larger than the degree of the check polynomials. Thus truncation leaves the echelon structure intact and we must have equality. Recall equality here implies that the spaces in (5) are equal!

3.2 Fundamental Method

Given two polynomials $h_1(t)$ and $h_2(t)$ the method is:

- calculate n^* so that $h_1(t), h_2(t)|t^{n^*}-1$;
- solve $g_i(t) = (t^{n^*}-1)/h_i(t)$ for $i=1,2$;
- solve $g(t) = \text{lcm}(g_1(t), g_2(t))$. This polynomial generates the intersection of the extended codes;
- truncation from the left gives us the desired code, providing n is greater than the degree of $\text{lcm}(h_1(t), h_2(t))$.

3.3 Optimizing the Method

This process could be fairly time consuming if n^* is large. If we work with polynomials of high degree this method would probably not be beneficial. So we look to alternative means. We aim to get the least common multiple of the two generator polynomials $g_1(t)$ and $g_2(t)$ (possibly of very high degree). Using properties of cyclic codes and the g.c.d. we write

$$\text{lcm}(g_1(t), g_2(t)) = (t^{n^*}-1)/\text{gcd}(h_1(t), h_2(t)).$$

[4] We require that n is larger than the degree of the least common multiple of the check polynomials $h_1(t)$ and $h_2(t)$.

Observe that this is now in the form of our original optimization problem. We solve (2) for $q(t)$ using $h(t) = \gcd(h_1(t), h_2(t))$ where $\tilde{n}$ is the number of rows plus the number of columns of our desired generator matrix minus one.

An Illustrative Example. Let

$$\begin{aligned} h_1(t) &= 1 + t + t^2 + t^5 + t^6 + t^7 + t^8 + t^{11}, \\ h_2(t) &= 1 + t^2 + t^4 + t^9 + t^{10} + t^{12} + t^{14} + t^{15}, \end{aligned}$$

We calculate the $\text{lcm}(h_1(t), h_2(t))$ and the $\gcd(h_1(t), h_2(t))$. The degree of the l.c.m. is 16, thus we must choose $n > 16$; let $n = 32$. We know G will be a 10×32 matrix thus $\tilde{n} = 41$. Now we can solve (2) for $q(t)$:

$$t^{41} - 1 = q(t)(1 + t^2 + t^3 + t^4 + t^6 + t^8 + t^9 + t^{10}) + r(t),$$

and we find that

$$\begin{aligned} q(t) = &1 + t^3 + t^4 + t^8 + t^{10} + t^{13} + t^{14} + t^{15} + t^{19} + t^{21} \\ &+ t^{22} + t^{23} + t^{25} + t^{28} + t^{30} + t^{31}. \end{aligned}$$

These coefficients are entered into the matrix in the following way (with notation as before):

$$G = \begin{bmatrix} q_9 & \cdots & q_{39} & q_{40} \\ q_8 & q_9 & \cdots & q_{39} \\ \vdots & \vdots & & \vdots \\ q_0 & q_1 & \cdots & q_{31} \end{bmatrix}.$$

The rowspace of this matrix G is the code we desire.

4 An Arbitrary Number of Parity Check Blocks

The procedure given in the previous section can be extended to deal with an arbitrary number of these defective blocks $H_1, \ldots, H_s$. Suppose we wish to have a parity check matrix

$$H = \begin{bmatrix} H_1 & H_2 & \ldots & H_s \end{bmatrix}^T,$$

defined by the s polynomials $h_1(t), \ldots, h_s(t)$. Given a block length n such that

$$n > \text{degree}(\text{lcm}(h_1(t), h_2(t), \ldots, h_s(t))),$$

the polynomial given by

$$\text{lcm}(g_1(t), \ldots, g_s(t)) = \frac{(t^{n^*} - 1)}{\gcd(h_1(t), \ldots, h_s(t))}$$

generates the intersection of the extended codes. Let $h(t) = \gcd(h_1(t), h_2(t), \ldots, h_s(t))$ and $k = \text{degree}(h(t))$. Then to form the $k \times n$ generator matrix G associated with our original matrix H we solve (2) for $q(t)$ with $\tilde{n} = k + n - 1$. Then the coefficients of the polynomial $q(t) = q_0 + q_1 t + \ldots + q_{\tilde{n}} t^{\tilde{n}}$ form G as

$$G = \begin{bmatrix} q_{k-1} & q_k & \cdots & q_{\tilde{n}-1} & q_{\tilde{n}} \\ q_{k-2} & q_{k-1} & \cdots & q_{\tilde{n}-2} & q_{\tilde{n}-1} \\ \vdots & \vdots & & \vdots & \vdots \\ q_0 & q_1 & \cdots & q_{n-2} & q_{n-1} \end{bmatrix}.$$

From a practical point of view, the G that is created is non-systematic and we do not need Gaussian elimination, which is an advantage when we are given an LDPC parity check matrix with a large block length. Moreover, G has a polynomial description which allows for an implemenation of the encoder by simple shift registers rather than by a costly matrix multiplication. When the parity check matrix can be brought into the form stated above, the result is a very efficient alternative to the method presented in [7].

5 Conclusion

In this paper we have provided an new method of obtaining a generator matrix G from certain types of parity check matrices consisting of a defective cyclic block structure. The generator matrices are described in terms of polynomials and by using polynomial algebra we have found efficient ways to implement the scheme. There is a large class of practical parity check matrices, e.g., those of LDPC codes, that are candidates for this method as we can use row and column operations to obtain the necessary form consisting of cyclic "blocks". Further, by careful choice of polynomials we are free to choose any length and rate code we may desire, and the density of G depends solely on it's corresponding polynomial.

Acknowledgement

The authors acknowledge the support of the Scottish Funding Council for the Joint Research Institute with the Heriot-Watt University which is a part of the Edinburgh Research Partnership.

References

[1] F. J. Macwilliams & N. J. A. Sloane, *The Theory Of Error-Correcting Codes*, Elsevier Science Publishers, 1977.

[2] S. Lin & D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, Second Edition, Prentice Hall, pp. 851–951, 2004.

[3] P. B. Garrett, *The Mathematics of Coding Theory : Information, Compression, Error Correction, and Finite Fields*, Pearson/Prentice Hall, 2004.

[4] R. B. J. T. Allenby, *Rings, Fields and Groups: An Introduction to Abstract Algebra*, Second Edition, Butterworth-Heinemann, pp. 211–212, 1991.

[5] A. Betten, M. Braun, H. Fripertinger, A. Kerber, A. Kohnert, A. Wassermann, *Error-Correcting Linear Codes: Classification by Isometry and Applications*, Springer, p. 97, 2006.

[6] R. B. J. T. Allenby, *Linear Algebra*, Butterworth-Heinemann, pp. 152–153, 1997.

[7] T. J. Richardson, R. L. Urbanke, *Efficient Encoding of Low-Density Parity-Check Codes*, IEEE Transcations on Information Theory, vol. 47, pp. 638–656, Feb. 2001.