Geometric quasi-cyclic low density parity check codes Discrete Mathematics Days 2024*

Simeon Ball[†1] and Tomas Ortega[‡2]

[1]Dept. of Mathematics, Universitat Politecnica Catalunya, 08034 Barcelona
[2]Dept. of Electrical Engineering and Computer Science, University of California, Irvine Irvine, CA 92697, USA.

Abstract

In this talk we present families of quasi-cyclic LDPC codes derived from the quasi-cyclic representation of the point-line incidence matrix of the classical finite generalised quadrangles. We detail how to explicitly calculate quasi-cyclic generator and parity check matrices for classical finite generalised quadrangles codes of length up to 400000. These codes cover a wide range of transmission rates, are easy and fast to implement and perform close to Shannon's limit.

1 Introduction

In many modern communication systems Low Density Parity Check (LDPC) codes are used. LDPC codes are those codes for which the number of ones in the check matrix is very small compared to the size of the matrix. A quasi-cyclic LDPC check matrix H can be described by a block size b, sometimes called the lifting degree or lifting factor, and a $(m/b) \times (n/b)$ matrix $\mathrm{H}^{\mathrm{rep}}$, whose entries are subsets H_{ij} of $\{1, \ldots, b\}$, where $i \in \{1, \ldots, (m/b)\}$ and $j \in \{1, \ldots, (n/b)\}$. Typically the subset H_{ij} is empty which corresponds to the $b \times b$ zero matrix in H in the (i, j) cell. A singleton subset $\mathrm{H}_{ij} = \{r\}$ indicates that in the (i, j) cell we have a copy of the $b \times b$ identity matrix shifted r bits (cyclically) to the right. A larger subset will involve a superposition of such shifts of the identity matrix. This representation of the quasi-cyclic LDPC check matrix H allows one to implement decoding algorithms, such as the sum-product algorithm, in an efficient manner.

The Tanner graph Γ is the bipartite graph with stable sets of size m and n, where there is a correspondence between the edges in Γ and a one entry in the matrix H. The decoding algorithms mentioned in the previous paragraph work well if the girth of Γ, the length of the shortest cycle, is large, and decode quickly if Γ has low diameter [6]. The diameter is the maximum distance between any two vertices. These conflicting objectives are optimised when the girth is twice the diameter. The graphs Γ which achieve this bound are the incidence matrices H of a generalised polygon. The rows of H are indexed by the points of the polygon and the columns are indexed by the lines, or vice-versa, where there is a one entry in the matrix H if and only if the point indexing the column and the line indexing the row are incident in the geometry. Finite generalised polygons have diameter 3, 4, 6 or 8, see [3], and are respectively called, projective planes, generalised quadrangles, generalised hexagons and generalised octagons. The LDPC code used in IEEE 802.3 standard (2048,1723) LDPC code for the 10-G Base-T Ethernet, is a quasi-cyclic LDPC code from an affine plane over $\mathbb{F}_{32}$ (a projective plane over the field of 32 elements with a line deleted) which has block size $b = 64$, length $n = 2048$

*The full version of this work can be found in [1] and [2]. This research is supported by the Spanish Ministry of Science, Innovation and Universities grant PID2020-113082GB-I00 funded by MICIU/AEI/10.13039/501100011033.

†Email: simeon.michael.ball@upc.edu.
‡Email: tomaso@uci.edu.

and dimension $k = 1723$, see [7, Example 10.5]. The LDPC code used in the NASA Landsat Data Continuation is a quasi-cyclic LDPC code from a 3-dimensional affine space (a projective space with a plane deleted) which has block size $b = 511$, length $n = 8176 = 16b$ and dimension $k = 7154 = 14b$, see [7, Example 10.10]. In this talk, we will describe how to efficiently employ quasi-cyclic LDPC codes derived from classical generalised quadrangles. These codes are fast and efficient, can be extremely long, and perform favourably compared to commercially used codes with similar parameters.

2 Quasi cyclic generator and check matrices

It was proven in [5] that the classical generalised quadrangles, of which there are six types, have a quasi-cyclic representation. However, up until now, no description of these quasi-cyclic representations was known. Here, we detail a simple, explicit description of a quasi-cyclic representation for all the classical generalised quadrangles. Using this representation, we describe how to employ the corresponding quasi-cyclic LDPC code in an efficient manner. In four of the six types, we do not take the entire quadrangle but a carefully chosen large sub-structure, which allows us to increase the size of the blocks b. It is advantageous to have a large block size since this allows the implementation of significantly longer codes. As evidenced in the proof of Shannon's theorem, the implementation of long codes brings the performance of the code close to Shannon's limit.

Once we have described how to construct the quasi-cyclic representation of the check matrix H in a purely algebraic manner, we can compute H for classical generalised quadrangles LDPC codes of length up to 400000. These computations were performed on a standard laptop and one can compute quasi-cyclic check matrices and generator matrices for longer codes with more computational power. The complexity of these computations for each quadrangle is detailed in Table 1.

Let C denote the binary linear code whose check matrix is H. We shall refer to C as the *full code*. Efficient encoding can be implemented if one can find a generator matrix for the code in quasi-cyclic form, see [4]. However, such a generator matrix for the full code C does not generally exist, so we take a large subcode C' of C for which there is a generator matrix G in quasi-cyclic form. We will call C' the *implementable code*. A $k \times n$ generator matrix is in standard form if it has $k \times k$ submatrix which is an identity matrix. For each quadrangle and each q, we compute a generator matrix G in standard quasi-cyclic form for the implementable code C'. Note that H is also a check matrix for C'. The matrix G can be described by a $(k/b) \times ((n-k)/b)$ matrix $\mathrm{P}^{\mathrm{rep}}$, where the (i,j) entry of $\mathrm{P}^{\mathrm{rep}}$ is P_{ij}, a vector of $\{0,1\}^b$. Replacing each first row, with the full circulant $b \times b$ matrix one obtains a matrix P, where

$$\mathrm{G} = (\mathrm{P} \mid \mathrm{id}). \tag{1}$$

Here, id denotes the $k \times k$ identity matrix.

Given the matrix $\mathrm{P}^{\mathrm{rep}}$, a shift-register-adder-accumulator (SRAA) circuit with a b-bit feedback shift register can be implemented to calculate each block of b parity check bits of the encoded codeword, see [4, Figure 1]. In series, this gives an encoding circuit of $(n-k)/b$ SRAA circuits with a total of $2(n-k)$ flip-flops, $n-k$ AND gates, and $n-k$ two-input XOR gates. The encoding is completed in a time proportional to $n-k$, see [4, Figure 2]. An encoder which completes in $n-k$ clock cycles and k/b feedback shift registers, each with b flip-flops can be implemented when the circuits are put in parallel, see [4, Figure 3].

Thus, we have an efficient encoding and decoding of very long quasi-cyclic LDPC codes whose performance is close to Shannon's limit.

In the talk we will give a simple algebraic description of the quasi-cyclic representations which allow the construction of $\mathrm{H}^{\mathrm{rep}}$ for codes of extraordinary length. In some cases it is feasible to calculate $\mathrm{H}^{\mathrm{rep}}$ for codes of length 10^7. One can calculate $\mathrm{P}^{\mathrm{rep}}$, and thus an explicit generator matrix in standard quasi-cyclic form for codes of length up to 400000 and transmission rates covering a wide spectrum of possible rates.

Code	increased block size b	block size	approx length n	complex -ity $\mathrm{H}^{\mathrm{rep}}$	complex -ity $\mathrm{P}^{\mathrm{rep}}$	min. dist.	approx. rate
$W(3,q)$	q^2+1	q^2+1 (q even) $\frac{1}{2}(q^2+1)$ (q odd)	q^3	$O(q^4)$	$O(q^9)$	$\geqslant 2q$	$1-q^{-0.286}$ (q even) 0.5 (q odd)
$W(3,q)$ dual	q^2+1	q^2+1 (q even) $\frac{1}{2}(q^2+1)$ (q odd)	q^3	$O(q^4)$	$O(q^9)$	$\geqslant 2q$	$1-q^{-0.286}$ (q even) 0.5 (q odd)
$Q(5,q)$	q^3+1	q^2-q+1 ($q=0,1 \bmod 3$) $\frac{1}{3}(q^2-q+1)$ ($q=2 \bmod 3$)	q^5	$O(q^6)$	$O(q^{13})$	$\geqslant 2q$	$1-q^{-1}$
$Q(5,q)$ dual	q^3+1	q^2-q+1 ($q=0,1 \bmod 3$) $\frac{1}{3}(q^2-q+1)$ ($q=2 \bmod 3$)	q^4	$O(q^6)$	$O(q^{14})$	$\geqslant q^3$	q^{-1}
$H(4,q^2)$	$\frac{q^5+1}{q+1}$	$\frac{q^5+1}{q+1}$	q^8	$O(q^{11})$	$O(q^{22})$	$\geqslant 2q^2$	$1-q^{-1}$
$H(4,q^2)$ dual	$\frac{q^5+1}{q+1}$	$\frac{q^5+1}{q+1}$	q^7	$O(q^{11})$	$O(q^{23})$	$\geqslant q^5$	q^{-1}

Table 1: The block size, length, and complexity of constructing $\mathrm{H}^{\mathrm{rep}}$ and $\mathrm{P}^{\mathrm{rep}}$.

3 Quasi-cyclic LDPC codes from classical generalised quadrangles.

Although there are six types of classical generalised quadrangles, these come in pairs, where one is the dual of the other. To obtain the dual quadrangle one switches the role of the points and the lines. In practical terms this is achieved by replacing the check matrix H by its transpose. Thus, we only need to describe how to construct H for three of the six types. These are labelled in Table 1 as $W(3,q)$ (the symplectic quadrangle), $Q(5,q)$ (the elliptic quadrangle) and $H(4,q^2)$ (the Hermitian quadrangle).

We will describe how to find these quasi-cyclic representations by consider the geometries as subsets of certain field extensions. This will lead us to increased block sizes which are detailed in Table 1. I will also present data on how these codes perform with respect to Shannon's bound. It is not possible to simulate performance of codes of very long length without using a field programmable gate array. The very longest codes, for example the quasi-cyclic LDPC codes arising from Q(5,13) with $n = 371462$ and rate $R = 0.9172$ were implemented using a field programmable gate array. It was seen empirically that these codes work exceedingly well with low complexity decoding algorithms which require just a few iterations. This indicates that they may have a use in storing large amounts of data, where fast and reliable decoding can be employed on retrieval.

References

[1] S. Ball and T. Ortega, Quasi-cyclic LDPC codes based on generalised quadrangles, Patent Cooperation Treaty, International Application No. PCT/EP2023/062797, World Intellectual Property Organization WO/2023/218050 (2023).

[2] S. Ball and T. Ortega, Practical implementation of geometric quasi-cyclic low density parity check codes, preprint, 2024.

[3] W. Feit and G. Higman, The nonexistence of certain generalised polygons, *J. Algebra*, **1** (1964) 114–131.

[4] Z. Li, L. Chen, L. Zeng and S. Lin, Efficient encoding of quasi-cyclic low-density parity-check codes, *IEEE Trans. Inform. Theory*, **54** (2006) 71–81.

[5] Z. Liu and D. Pados, LDPC codes from generalised polygons, *IEEE Trans. Inform. Theory*, **51** (2005) 3890–3898.

[6] G. A. Margulis, Explicit constructions of graphs without short cycles and low-density codes, *Combinatorica*, **2** (1982) 71–78.

[7] W. E. Ryan and S. Lin, *Channel codes: Classical and Modern*, Cambridge University Press, 2009.