Algebraic Constructions of Quasi-cyclic LDPC Codes – Part I: For AWGN and Binary Random Erasure Channels*

Lan Lan, Lingqi Zeng, Ying Y. Tai, Lei Chen, Shu Lin,
and Khaled Abdel-Ghaffar

Department of Electrical and Computer Engineering,
University of California, Davis, CA 95616
{squash, lqzeng, yytai, leichen, shulin,
ghaffar}@ece.ucdavis.edu

Abstract. This paper is the first part of a sequence of two papers that present algebraic constructions of quasi-cyclic LDPC codes for AWGN, binary random and burst erasure channels. In this paper, a class of quasi-cyclic LDPC codes for both AWGN and binary random erasure channels is constructed based on finite fields and special vector representations of finite field elements.

1 Introduction

LDPC codes, discovered by Gallager in 1962 [1], were rediscovered and shown to form a class of Shannon capacity approaching codes in the late 1990's [2, 3]. Ever since their rediscovery, design, construction, decoding, efficient encoding, and applications of these codes in digital communication and storage systems have become focal points of research. Many methods for constructing these codes have been proposed. Based on the methods of construction, LDPC codes can be classified into two general categories: (1) random-like codes [4, 5] that are generated by computer search based on certain design guidelines and required structural properties of their Tanner graphs [6]; and (2) structured codes that are constructed based on algebraic and combinatorial tools [7, 8, 9, 10, 11, 12, 13, 14, 15, 16].

Most of the proposed constructions of LDPC codes are for the AWGN channel, however only a few of them for other types of channels. In this and a succeeding papers, we present two algebraic methods for constructing quasi-cyclic (QC) LDPC codes for AWGN, binary random and burst erasure channels. QC-LDPC codes have encoding advantage over the other types of LDPC codes. They can be encoded with simple shift-registers with linear complexity [17]. It has been shown that well designed QC-LDPC codes decoded with iterative decoding perform very well over the AWGN channel and close to the Shannon theoretical limit [10, 14]. In this and next papers, we show that well designed QC-LDPC codes decoded with iterative decoding also perform well over binary random and burst erasure channels.

* This research was supported by NASA under the Grant NNG05GD13G and NSF under the Grant CCR-0117891 and ECS-0121469.

M. Fossorier et al. (Eds.): AAECC 2006, LNCS 3857, pp. 255–264, 2006.
© Springer-Verlag Berlin Heidelberg 2006

A binary regular LDPC code [1] is defined as the null space of a sparse parity-check matrix $\mathbf{H}$ over GF(2) with the following structural properties: (1) each row has constant weight ρ; (2) each column has constant weight γ; (3) no two rows (or two columns) have more than one 1-component in common; and (4) both ρ and γ are small compared with the code length. $\mathbf{H}$ is said to be (γ,ρ)-*regular* and the code given by the null space of $\mathbf{H}$ is called a (γ,ρ)-regular LDPC code. Property (3) is referred to as the *column-row* (*RC*) *constraint*. The RC-constraint ensures that: (1) the minimum distance of the code is at least $\gamma+1$; and (2) the Tanner graph of the code is *free* of cycles of length 4 [7]. An LDPC code is said to be *irregular* if its parity-check matrix has *varying column weights and/or varying row weights*. A QC-LDPC code is given by the null space of an *array of sparse circulants* [7, 10, 14].

The performance of an LDPC code decoded with iterative decoding is measured by its bit-error probability, block-error probability, error-floor and rate of decoding convergence, collectively. Structured LDPC codes in general have a lower error-floor which is important in digital communication and storage systems, where very low error rates are required. Structured LDPC codes with large minimum distances can be constructed much easier than computer generated random-like LDPC codes.

The performance of an LDPC code over the AWGN channel with iterative decoding depends on a number of code structural properties besides its minimum distance. One such structural property is the *girth* of the code that is defined as the length of the shortest cycle in the code's Tanner graph. For an LDPC code to perform well over the AWGN channel with iterative decoding, its Tanner graph must not contain short cycles. The shortest cycles that affect code performance the most are cycles of length 4. Therefore, cycles of length 4 must be prevented in LDPC code construction for the AWGN channel. For an LDPC code to perform well over the binary random erasure channel, its Tanner graph must also be free of cycles of length 4 [18, 19].

2 LDPC Codes for the Binary Random Erasure Channel

For transmission over the binary random erasure channel, a symbol, 0 or 1, is either correctly received with probability $1-p$ or erased with probability p (called *erasure probability*), and there is no transmission error. The output of the binary random erasure channel consists of three symbols, 0, 1, and ?, where the symbol "?" denotes a transmitted symbol being erased, called an *erasure*. Suppose a codeword $\boldsymbol{x} = (x_0, x_1, \ldots, x_{n-1})$ from a binary code $\mathcal{C}$ of length n is transmitted and $\boldsymbol{y} = (y_0, y_1, \ldots, y_{n-1})$ is the corresponding received sequence. Let $\mathcal{E} = \{j_1, j_2, \ldots, j_t\}$ be the set of locations in $\boldsymbol{y}$ with $0 \leq j_1 < j_2 < \ldots < j_t < n$, where the transmitted symbols are being erased. Let $[n] \triangleq \{0, 1, \ldots, n-1\}$. Define $\bar{\mathcal{E}} \triangleq [n] \backslash \mathcal{E}$. Then $\bar{\mathcal{E}}$ is the set of locations in $\mathbf{y}$ where the transmitted symbols are correctly received, i.e., $y_i = x_i$ for $i \in \bar{\mathcal{E}}$. The set $\mathcal{E}$ displays the *pattern* of erased symbols in $\mathbf{y}$ and is called an *erasure pattern*. Decoding $\boldsymbol{y}$ is to determine the value of each erasure in $\mathcal{E}$. An erasure pattern $\mathcal{E}$ is said to be *recoverable* (*resolvable* or *correctable*) if the value of each erasure in $\mathcal{E}$ can be uniquely determined.

Consider an LDPC code $\mathcal{C}$ of length n given by the null space of a $J \times n$ sparse matrix $\mathbf{H}$. Then a binary n-tuple $\boldsymbol{x} = (x_0, x_1, \ldots, x_{n-1})$ is a codeword in $\mathcal{C}$ if and only if $\boldsymbol{x}\mathbf{H}^T = \mathbf{0}$. Suppose a codeword $\boldsymbol{x}$ is transmitted and $\boldsymbol{y} = (y_0, y_1, \ldots, y_{n-1})$ is the corresponding received sequence. Let $\mathcal{E} = \{j_1, j_2, \ldots, j_t\}$ be the erasure pattern contained in $\boldsymbol{y}$. Let $\mathbf{H}_{\mathcal{E}}$ be the submatrix that consists of the columns of $\mathbf{H}$ corresponding to the locations of the erased symbols given in $\mathcal{E}$ and $\mathbf{H}_{\bar{\mathcal{E}}}$ be the submatrix that consists of the columns of $\mathbf{H}$ corresponding to the locations of the correctly received code symbols in $\bar{\mathcal{E}}$. Let $\boldsymbol{y}_{\mathcal{E}}$ denote the subsequence that consists of the erased symbols in $\boldsymbol{y}$ and $\boldsymbol{y}_{\bar{\mathcal{E}}}$ denote the subsequence that consists of the known symbols in $\boldsymbol{y}$ at the locations given in $\bar{\mathcal{E}}$. The symbols in $\boldsymbol{y}_{\mathcal{E}}$ are unknown. For $\boldsymbol{y}$ to be a codeword in $\mathcal{C}$, we must have $\boldsymbol{y}\mathbf{H}^T = \mathbf{0}$. This constraint can be put in the form:

$$\boldsymbol{y}_{\mathcal{E}}\mathbf{H}^T_{\mathcal{E}} = \boldsymbol{y}_{\bar{\mathcal{E}}}\mathbf{H}^T_{\bar{\mathcal{E}}}. \tag{1}$$

The right-hand side of (1) is known and can be computed from $\boldsymbol{y}_{\bar{\mathcal{E}}}$ and $\mathbf{H}_{\bar{\mathcal{E}}}$. The left-hand side of this equation (or $\boldsymbol{y}_{\bar{\mathcal{E}}}$) is unknown. Then decoding $\boldsymbol{y}$ is to solve (1). An iterative method for solving (1) was proposed in [18].

Let $\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_J$ be the rows of $\mathbf{H}$. For $1 \leq i \leq J$, let $\boldsymbol{h}_i = (h_{i,0}, h_{i,1}, \ldots, h_{i,n-1})$. Then a codeword $\boldsymbol{y} = (y_0, y_1, \ldots, y_{n-1})$ must satisfy the condition, $s_i \triangleq y_0 h_{i,0} + y_1 h_{i,1} + \ldots + y_{n-1} h_{i,n-1} = 0$ for $1 \leq i \leq J$, which is called a *check-sum*. The code symbol y_j is said to be checked by the sum s_i if $h_{i,j} = 1$, i.e., y_j is included in the sum s_i. Then y_j can be determined from other code bits that are checked by $\boldsymbol{h}_i$ as follows:

$$y_j = \sum_{k=0, k \neq j}^{n-1} y_k h_{i,k}. \tag{2}$$

For each erased position j_l in an erasure pattern $\mathcal{E} = \{j_1, j_2, \ldots, j_t\}$ with $1 \leq l \leq t$, if there exists a row $\boldsymbol{h}_i$ in $\mathbf{H}$ that checks *only* the erased symbol y_{j_l} and not any of the other $t-1$ erased symbols in $\mathcal{E}$, then it follows from (2) that the value of each erased symbol in $\mathcal{E}$ can be determined by the correctly received symbols in $\bar{\mathcal{E}}$ as follows:

$$y_{j_l} = \sum_{k \in \bar{\mathcal{E}}} y_k h_{i,k}. \tag{3}$$

Such an erasure pattern is said to be *resolvable* in *one step* (or *one iteration*). However, there are erasure patterns that are not resolvable in one step but resolvable in *multiple steps* iteratively. Given an erasure pattern $\mathcal{E}$, we first determine the values of those erased symbols that can be resolved in one step using (3). Then we remove the known erased symbols from $\mathcal{E}$. This results in a new erasure pattern $\mathcal{E}_1$ of smaller size. Next we determine the values of erased symbols in $\mathcal{E}_1$ that are resolvable using (3). Removing the known erased symbols from $\mathcal{E}_1$, we obtain an erasure pattern $\mathcal{E}_2$ of size smaller than that of $\mathcal{E}_1$. We repeat the above process iteratively until either all the erased symbols in $\mathcal{E}$ are resolved or an erasure pattern $\mathcal{E}_m$ is obtained such that no erasure in $\mathcal{E}_m$ can be resolved using (3). In the latter case, some erasures can not be recovered.

The above decoding process is iterative in nature and can be formulated as an algorithm [18]. To initialize the decoding process, we first set $k = 0$ and $\mathcal{E}_0 = \mathcal{E}$. Then we execute the following steps iteratively:

(1) Determine $\mathcal{E}_k$. If $\mathcal{E}_k$ is empty, stop decoding, otherwise go to Step 2.
(2) Form $\mathbf{H}_{\mathcal{E}_k}$, $\mathbf{H}_{\bar{\mathcal{E}}_k}$, $\boldsymbol{y}_{\mathcal{E}_k}$, and $\boldsymbol{y}_{\bar{\mathcal{E}}_k}$.

(3) Compute $\boldsymbol{y}_{\bar{\varepsilon}_k}\mathbf{H}_{\bar{\varepsilon}_k}^T$.
(4) Find the rows in $\mathbf{H}_{\varepsilon_k}$ such that each contains only one 1-component. Determine the erasures in $\mathcal{E}_k$ that are checked by these rows. Determine the values of these erasures by application of (3) and go to Step 5. If there is no row in $\mathbf{H}_{\varepsilon_k}$ that contains only one 1-component, stop decoding.
(5) Remove the erasures resolved at the step 4 from $\mathcal{E}_k$. Set $k = k+1$ and go to Step 1.

If decoding stops at Step 1, all the erasures in the erasure pattern $\mathcal{E}$ are resolved and the decoding is successful. If decoding stops at Step 4, some erasures in $\mathcal{E}$ can not be recovered.

The performance measure of an LDPC code over the binary random erasure channel is the error probability. Di et. al. [18] have derived the *threshold* for regular LDPC codes with given Tanner degree distribution pair (γ, ρ) (or column and row weight pair of a (γ,ρ)-regular parity-check matrix) using the above iterative decoding algorithm. The threshold is a small probability $\varepsilon(\gamma, \rho)$ associated with an ensembles of regular LDPC codes whose Tanner graphs have degree distribution pair (γ,ρ). The implication of threshold $\varepsilon(\gamma, \rho)$ is as follows: over all binary random erasure channels with erasure probability p smaller than $\varepsilon(\gamma, \rho)$, information can be reliably transmitted by using a sufficiently long LDPC code with degree distribution pair (γ, ρ). Reliable transmission of information is not possible if the erasure probability p is larger than the threshold $\varepsilon(\gamma, \rho)$.

The performance of an LDPC code over the binary random erasure channel is determined by the *stopping sets* of its Tanner graph $\mathcal{T}$ [18]. Let $\mathcal{V}$ be a set of variable nodes in $\mathcal{T}$ and $\mathcal{S}$ be the set of check nodes in $\mathcal{T}$ such that each check node in $\mathcal{S}$ is connected to at least one variable node in $\mathcal{V}$. The nodes in $\mathcal{S}$ are called the neighbors of the nodes in $\mathcal{V}$. A set $\mathcal{V}$ of variable nodes is called a stopping set of $\mathcal{T}$ if each check node in the neighbor check set $\mathcal{S}$ of $\mathcal{V}$ is connected to at least two nodes in $\mathcal{V}$. If an erasure pattern $\mathcal{E}$ corresponds to a stopping set in the Tanner graph of an LDPC code, then a checksum that checks an erasure in $\mathcal{E}$ also checks *at least one other erasure* in $\mathcal{E}$. As a result, no erasure in $\mathcal{E}$ can be determined with Eq. (3) (or Eq. (1)) and $\mathcal{E}$ is an irrecoverable erasure pattern.

A set $\mathcal{Q}$ of variable nodes in $\mathcal{T}$ may contain many stopping sets. It is clear that: (1) the union of two stopping sets in $\mathcal{Q}$ is also a stopping set in $\mathcal{Q}$; and (2) the union of all the stopping sets in $\mathcal{Q}$ gives the maximum stopping set in $\mathcal{Q}$. A set $\mathcal{V}_{ssf}$ of variable nodes in $\mathcal{T}$ is said to be *stopping-set-free* (SSF) if it does not contain any stopping set. The following theorem [18] characterizes the significance of stopping sets for correcting erasures: Suppose an LDPC code $\mathcal{C}$ is used for correcting erasures using iterative decoding. Let **y** be a received sequence that contains an erasure pattern $\mathcal{E}$. Then the erasures contained in the maximum stopping set of $\mathcal{E}$ cannot be recovered. This theorem says that any erasure pattern $\mathcal{E}$ is recoverable if it is SSF.

Let $\mathcal{B}$ be a stopping set of minimum size in the Tanner graph of an LDPC code, called a *minimal stopping set* (not unique). If the code symbols corresponding to the variable nodes in $\mathcal{B}$ are being erased, it follows from the above theorem that $\mathcal{B}$ forms an irrecoverable erasure pattern of minimum size. Therefore, for random erasure correction with iterative decoding, it is desired to construct codes with largest possible minimal stopping sets in their Tanner graphs. A good LDPC code for erasure correction must have no or very few small stopping sets. A stopping set always contains cycles. In [19], it has been proved that the size of a minimal stopping set of a Tanner graph with girth 4 is two. The size of a minimal stopping set of a Tanner graph with girth 6 is $\gamma + 1$ and is 2γ for girth 8, where γ is the degree of a variable node (or the column weight of the parity-check matrix of the code). Hence for iterative decoding of an LDPC code over the binary random erasure channel, the most critical cycles in the code's Tanner graph are cycles of length 4. Therefore, in code construction for the binary random erasure channel, cycles of length 4 must be avoided in the Tanner graph of a code. It is proved in [20] that for a code with minimum distance d_{min}, it must contain a stopping set of size d_{min}. Therefore, in the construction of a code for erasure correction, we need to keep its minimum distance large. For a regular LDPC code whose parity-check matrix has column weight γ and satisfies the RC-constraint, the size of a minimal stopping set in the code's Tanner graph is at least $\gamma + 1$.

3 A Class of QC-LDPC Codes Constructed Based on Finite Fields

Consider the Galois field GF(q) where q is a power of a prime. Let α be a primitive element of GF(q). Then $\alpha^{-\infty} \triangleq 0, \alpha^0 = 1, \alpha, \ldots, \alpha^{q-2}$ form all the elements of GF(q) and $\alpha^{q-1} = 1$. The $q-1$ nonzero elements of GF(q) form the multiplicative group of GF(q) under the multiplication operation. For each nonzero element α^i with $0 \leq i \leq q-2$, we form a $(q-1)$-tuple over GF(2), $\mathbf{z}(\alpha^i) = (z_0, z_1, \ldots, z_{q-2})$, whose components correspond to the $q-1$ nonzero elements of GF(q), where the ith component $z_i = 1$ and all the other $q-2$ components are equal to 0. This $(q-1)$-tuple $\mathbf{z}(\alpha^i)$ is referred to as the *location vector* of α^i with respective to the multiplicative group of GF(q). We call $\mathbf{z}(\alpha^i)$ the location-vector of α^i. The location-vectors of two different nonzero elements of GF(q) are different. The location vector of the 0 element of GF(q) is defined as the all-zero $(q-1)$-tuple, $(0, 0, \ldots, 0)$. Let β be a nonzero element in GF(q), then the location-vector $\mathbf{z}(\alpha\beta)$ of $\alpha\beta$ is the *cyclic-shift* (*one place to the right*) of the location-vector $\mathbf{z}(\beta)$ of β. Form a $(q-1) \times (q-1)$ matrix $\mathbf{A}$ over GF(2) with the location-vectors of $\beta, \alpha\beta, \ldots, \alpha^{q-2}\beta$ as rows. Then $\mathbf{A}$ is a *circulant permutation matrix*.

Form the following $(q-1) \times (q-1)$ matrix over GF(q):

$$\mathbf{M} = \begin{bmatrix} \mathbf{w}_0 \\ \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_{q-2} \end{bmatrix} = \begin{bmatrix} \alpha^0 - 1 & \alpha - 1 & \cdots & \alpha^{q-2} - 1 \\ \alpha - 1 & \alpha^2 - 1 & \cdots & \alpha^{q-1} - 1 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{q-2} - 1 & \alpha^{q-1} - 1 & \cdots & \alpha^{2(q-2)} - 1 \end{bmatrix}. \tag{4}$$

Matrix $\mathbf{M}$ has the following structural properties: (1) any two rows (or two columns) differ in all positions; (2) all the entries in a row (or a column) are different elements in GF(q); and (3) each row (or column) contains one and only one zero element.

Lemma 1. *For $0 <= i, j, k, l < q-1$ with $i \neq j$, the two $(q-1)$-tuples $\alpha^k \mathbf{w}_i$ and $\alpha^l \mathbf{w}_j$ can not have more than one position with identical components, i.e., they differ in at least $q-2$ positions.*

Proof. Suppose there are two different positions, say s and t with $0 \leq s, t < q-1$, where $\alpha^k \mathbf{w}_i$ and $\alpha^l \mathbf{w}_j$ have identical components. Then $\alpha^k(\alpha^{i+s} - 1) = \alpha^l(\alpha^{j+s} - 1)$

and $\alpha^k(\alpha^{i+t}-1)=\alpha^l(\alpha^{j+t}-1)$. These two equalities imply that either $i=j$ or $s=t$ that contradicts the assumptions that $i\neq j$ and $s\neq t$. This proves the theorem.

For each row $\mathbf{w}_i$ of $\mathbf{M}$ given by (4) with $0\leq i<q-1$, we form the following $(q-1)\times(q-1)$ matrix over GF(q) with $\mathbf{w}_i, \alpha\mathbf{w}_i, \ldots, \alpha^{q-2}\mathbf{w}_i$ as rows:

$$\mathbf{M}_i=\begin{bmatrix}\mathbf{w}_i\\ \alpha\mathbf{w}_i\\ \vdots\\ \alpha^{q-2}\mathbf{w}_i\end{bmatrix}=\begin{bmatrix}\alpha^i-1 & \alpha^{i+1}-1 & \cdots & \alpha^{i+q-2}-1\\ \alpha(\alpha^i-1) & \alpha(\alpha^{i+1}-1) & \cdots & \alpha(\alpha^{i+q-2}-1)\\ \vdots & \vdots & \ddots & \vdots\\ \alpha^{q-2}(\alpha^i-1) & \alpha^{q-2}(\alpha^{i+1}-1) & \cdots & \alpha^{q-2}(\alpha^{i+q-2}-1)\end{bmatrix}. \quad (5)$$

We label the column of $\mathbf{M}_i$ from 0 to $q-2$. We readily see that: (1) any two rows differ in every position, except the $(q-1-i)$th position, where they both have the 0 element of GF(q); and (2) the $q-1$ entries of each column of $\mathbf{M}_i$ form the $q-1$ nonzero elements of GF(q), except the entries of the $(q+1-i)$th column, which are all zeros.

Replacing each entry in $\mathbf{M}_i$ by its location-vector, we obtain a $(q-1)\times(q-1)^2$ matrix over GF(2), $\mathbf{B}_i=[\mathbf{A}_{i,0}\mathbf{A}_{i,1}\ldots\mathbf{A}_{i,q-2}]$, which consists of a row of $q-1$ $(q-1)\times(q-1)$ square submatrices, where $\mathbf{A}_{i,j}$ is formed with the location-vectors of the $q-1$ entries of the jth column of $\mathbf{M}_i$, $\alpha^{i+j}-1, \alpha(\alpha^{i+j}-1), \ldots, \alpha^{q-2}(\alpha^{i+j}-1)$, as rows. All the submatrices of $\mathbf{B}_i$ are $(q-1)\times(q-1)$ circulant permutation matrices, except $\mathbf{A}_{i,q-1-i}$, which is a $(q-1)\times(q-1)$ zero matrix. All the circulant permutation matrices in $\mathbf{B}_i$ are different. Form the following $(q-1)\times(q-1)$ array of $(q-1)\times(q-1)$ circulant permutation and zero matrices:

$$\mathbf{H}=\begin{bmatrix}\mathbf{B}_0\\ \mathbf{B}_1\\ \vdots\\ \mathbf{B}_{q-2}\end{bmatrix}=\begin{bmatrix}\mathbf{A}_{0,0} & \mathbf{A}_{0,1} & \cdots & \mathbf{A}_{0,q-2}\\ \mathbf{A}_{1,0} & \mathbf{A}_{1,1} & \cdots & \mathbf{A}_{1,q-2}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{A}_{q-2,0} & \mathbf{A}_{q-2,1} & \cdots & \mathbf{A}_{q-2,q-2}\end{bmatrix}, \quad (6)$$

which is a $(q-1)^2\times(q-1)^2$ matrix over GF(2) with both column and row weight $q-2$. It follows from Lemma 1 and the structural properties of matrices $\mathbf{M}$ and $\mathbf{M}_i$ that $\mathbf{H}$ satisfies the RC-constraint.

For any pair of positive integers with $1\leq\gamma,\rho<q$, let $\mathbf{H}(\gamma,\rho)$ be a $\gamma\times\rho$ subarray of $\mathbf{H}$. $\mathbf{H}(\gamma,\rho)$ is a $\gamma(q-1)\times\rho(q-1)$ matrix over GF(2) which also satisfies the RC-constraint. If $\mathbf{H}(\gamma,\rho)$ does not contain zero submatrices of $\mathbf{H}$, it has constant column and row weights γ and ρ, respectively. The null space of $\mathbf{H}(\gamma,\rho)$ gives a (γ,ρ)-regular QC-LDPC code $\mathcal{C}_{qc}$ of length $\rho(q-1)$, rate at least $(\rho-\gamma)/\rho$ and minimum distance at least $\gamma+1$, whose Tanner graph has a girth of at least 6. Since $\mathbf{H}(\gamma,\rho)$ is an array of permutation matrices, no odd number of columns can be added to the zero column vector $\mathbf{0}$, and hence the minimum distance of $\mathcal{C}_{qc}$ must be even. Consequently, the minimum distance of $\mathcal{C}_{qc}$ is at least $\gamma+2$ for even γ and $\gamma+1$ for odd γ. Since the girth of the Tanner graph of $\mathcal{C}_{qc}$ is at least 6, the size of a minimal stopping set in the Tanner graph is at least $\gamma+1$ [19]. If $\mathbf{H}(\gamma,\rho)$ contains some zero submatrices of $\mathbf{H}$, then $\mathbf{H}(\gamma,\rho)$ has two column weights, $\gamma-1$ and γ, and possibly two row weights $\rho-1$ and ρ. In this case, the null space of $\mathbf{H}(\gamma,\rho)$ gives a near-regular QC-LDPC code with minimum distance at least γ for even γ and at least $\gamma+1$ for odd γ. The size of a

minimal stopping set in the code's Tanner graph is either at least γ or $\gamma + 1$. The above construction gives a class of QC-LDPC codes whose Tanner graph have girth at least 6.

4 An Example

In the following, we use an example to illustrate the method of construction of QC-LDPC codes described in Section III and to demonstrate the performances of a code over the AWGN and binary random erasure channels. For the AWGN channel, the code is decoded with the *sum-product algorithm* (SPA) [3,4,7], and its performance is compared with the Shannon limit. For the binary random erasure channel, the code is decoded with the iterative decoding algorithm given in Section II and its performance is compared with the threshold $\epsilon(\gamma, \rho)$ for the degree pair (γ,ρ) of its Tanner graph. We set the maximum number of decoding iterations to 100. We also assume BPSK signaling.

Let GF(73) be the field for code construction. Using this field, we can construct a 72×72 array $\mathbf{H}$ of 72×72 circulant permutation and zero matrices. Set $\gamma = 6$ and $\rho = 72$. We take a 6×72 subarray $\mathbf{H}(6, 72)$ from array $\mathbf{H}$ (the first 6 rows of submatrices of $\mathbf{H}$). Each of the first 6 columns of submatrices of $\mathbf{H}(6, 72)$ contains a single 72×72 zero matrix. Hence $\mathbf{H}(6, 72)$ is a 432×5184 matrix over GF(2) with constant row weight 71 and two column weights, 5 and 6. The null space of $\mathbf{H}(6, 72)$ gives a $(5184, 4752)$ QC-LDPC code with rate 0.917. The performance and the rate of decoding convergency of this code over the AWGN channel are shown in Figure 1. We see that the decoding of this code converges very fast. At the BER of 10^{-6}, the performance gap between 5 iterations and 100 iterations is within 0.2dB. At BER of

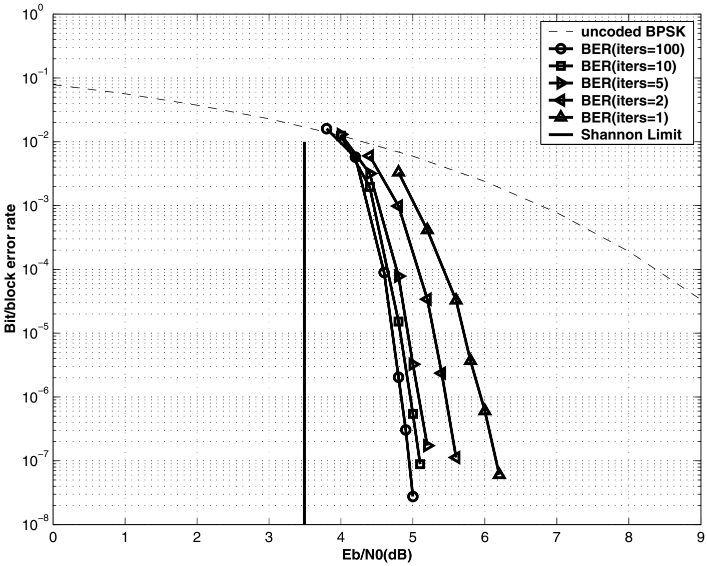

Fig. 1. Performance and the rate of decoding convergence of the (5184,4752) QC-LDPC code given in Section 4 over the AWGN channel

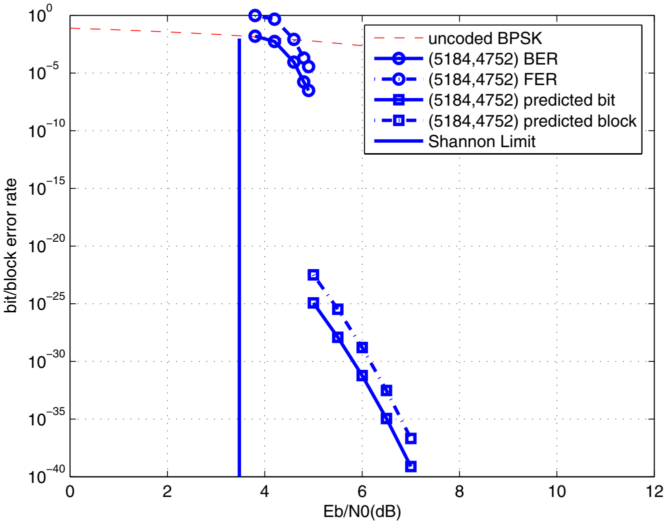

Fig. 2. Estimated error floor of the (5184,4752) QC-LDPC code given in Section 4 over the AWGN channels

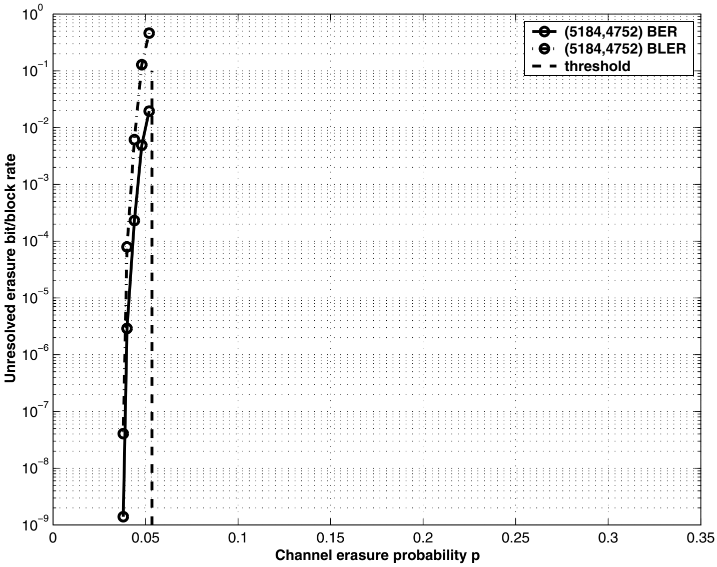

Fig. 3. Performance of the (5184,4752) QC-LDPC code given in Section 4 over the binary random erasure channel

10^{-6} with 100 iterations, the code performs 1.3 dB from the Shannon limit. The error-floor of this code is estimated below the BER of 10^{-25} and the block-error rate (BLER) of 10^{-22} as shown in Figure 2 (using the method given in [21]). The estimated minimum distance of this code is 19. The error performance of this code for the binary random erasure channel is shown in Figure 3. At the BER of 10^{-6}, the code performs 0.002 from the threshold $\epsilon(6, 72) = 0.0528$. Figures 1 to 3 demonstrates that the (5184,4752) QC-LDPC code constructed based on GF(73) performs well on both the AWGN and binary erasure channels.

5 Conclusion

In this paper, we have presented a method for constructing a class of QC-LDPC codes based on finite fields and location-vector representations of finite field elements. The Tanner graphs of the codes in this class have girth of at least 6. For a given finite field, a family of QC-LDPC codes with various lengths, rates, minimum distances and sizes of minimal stopping sets can be constructed. The proposed construction of QC-LDPC codes may be regarded parallel to the construction of BCH codes [22]. A QC-LDPC code was constructed to show that it performs very well over both the AWGN and binary random erasure channels with iterative decoding. It has a very low error-floor. In a succeeding paper, we will use the RC-constrained arrays of circulant permutation matrices constructed based on finite fields together with a masking technique to construct QC-LDPC codes for AWGN, binary random and burst erasure channels.

References

1. R. G. Gallager, "Low density parity-check codes," *IRE Trans. Inform. Theory*, IT-8, pp. 21-28, Jan. 1962.
2. D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity-check codes," *Electro. Lett.*, vol. 32, pp. 1645-1646, Aug. 1996.
3. D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, pp. 399-432, Mar. 1999.
4. T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 599-618, Feb., 2001.
5. T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of capacity-approaching low desnsity parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 619-637, Feb., 2001.
6. R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory*, vol. 27, pp. 533-547, Sept. 1981.
7. Y. Kou, S. Lin, and M. Fossorier, "Low density parity-check codes based on finite geometries: a discovery and new results," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2711-2736, Nov. 2001.
8. S. J. Johnson and S. R. Weller, "Codes for iterative decoding from partial geometries," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 236-247, Feb. 2003.
9. I. Djurdjevic, J. Xu, K. Abdel-Ghaffar and S. Lin, "Construction of low-density parity-check codes based on Reed-Solomon codes with two information symbols," *IEEE Commun. Lett.*, vol. 7, no. 7, pp. 317-319, July 2003.
10. L. Chen, J. Xu, I. Djurdjevic, and S. Lin, "Near Shannon limit quasi-cyclic low-density parity-check codes," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1038-1042, July 2004.
11. B. Vasic and O. Milenkovic, "Combinatorial construction of low density parity-check codes for iterative decoding," *IEEE Trans. Inform. Theory*, vol 50, no. 6, pp. 1156-1176, June 2004.
12. H. Tang, J. Xu, Y. Kou, S. Lin, and K. Abdel-Ghaffar, "On algebraic construction of Gallager and circulant low density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp 1269-1279, June 2004.

13. B. Ammar, B. Honary, Y. Kou, J. Xu, and S. Lin, “Construction of low density parity-check codes based on balanced imcomplete block designs,” *IEEE Trans. Inform. Theory*, vol. 50, vo. 6, pp. 1257-1268, June 2004.
14. L. Chen, L. Lan, I. Djurdjevic, S. Lin, K. Abdel-Ghaffar, “An algebraic method for constructing quasi-cyclic LDPC codes,” *Proc. Int. Symp. Inform. Theory and Its Applications*, ISITA2004, pp. 535-539, Parma, Italy, Oct. 10-13, 2004.
15. H. Tang, J. Xu, S. Lin, and K. Abdel-Ghaffar, “Codes on finite geometries,” *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 572-596, Feb. 2005.
16. J. Xu, L. Chen, L. -Q. Zeng, L. Lan, and S. Lin, “Construction of low-density parity-check codes by superposition,” *IEEE Trans. Commun.*, vol.53, no. 2, pp. 243-251, Feb. 2005.
17. Z. -W. Li, L. Chen, S. Lin, W. Fong, and P. -S. Yeh, “Efficient encoding of quasi-cyclic low-density parity-check codes,” *IEEE Trans. Commun.*, (accepted) 2005.
18. C. Di, D. Proietti, I. E. Teletar, T. j. Richarson, and R. L. Urbanke, “Finite length analysis of low-density parity-check codes on binary erasure channel,” *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1570-1579, June 2002.
19. A. Orlitsky, R. L. Urbanke, K. Viswanathan, and J. Zhang, “Stopping sets and the girth of Tanner graph,” *Proc. Int. Symp. Inform. Theory*, p.2, Lausanne, Switzerland, June 30-July 5, 2002.
20. T. Tian, C. Jones, J. D. Villasensor, and R. D. Wesel,”Construction of irregular LDPC codes with low error floors,” *IEEE ICC’03 Int. Conf. on Commun.*, pp. 3125-3129, 2003.
21. T. Richardson, “Error floors of LDPC codes,”*Proc. Allerton Conf. on Communication, Control and Computing*, pp. 1426-1435, Monticello, IL., Oct. 2003.
22. S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, 2nd edition, Prentice Hall, Upper Saddle River, NJ., 2004.