Invited Article

Constructing High-Rate QC-LDPC Codes with Large-Girth based on Shortened Array Codes

Xu Chen, Francis C. M. Lau

Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong

Correspondence: Francis C. M. Lau, encmlau@polyu.edu.hk

Manuscript communication: received 12 August 2011, revised 15 October 2011

Abstract– **In this paper, we aim at constructing high-rate quasi-cyclic low-density parity-check (QC-LDPC) codes with girth-10 based on shortened array codes. Our first contribution is the derivation of analytic results on the maximum number of columns for shortened array codes of different girths. Then, inspired by the analysis, we propose a code construction method for column-weight-three codes. We further compare the minimum length and the error performance of the column-weight-three codes constructed by the proposed algorithm and those found by the conventional greedy construction algorithm. We show that the proposed method is more effective than the conventional greedy algorithm in the sense that the minimum length of the codes constructed using the proposed method to achieve different code rates is comparative or much shorter than those constructed using the greedy construction.**

Keywords– **Additive white Gaussian noise, code construction, girth, high-rate quasi-cyclic low-density parity-check codes, QC-LDPC codes.**

The work described in this paper was partially supported by a grant from the RGC of the Hong Kong SAR, China (Project No. PolyU 521809).
Part of this paper has been published in the proceedings of 2011 International Conference on Advanced Technologies for Communications and has been given the Best Paper Award.

1 Introduction

Quasi-cyclic low-density parity-check (QC-LDPC) codes form a class of structured LDPC codes with the parity-check matrices consisting of circulant permutation sub-matrices [1, 2]. It has attracted much interest in research because the quasi-cyclic structure facilitates the encoder and decoder implementations [3, 4]. The application of QC-LDPC codes to optical communications has been extensively investigated recently, see [5] and references therein. One of the challenges in the next generation optical communication systems is to find channel codes with low redundancy along with extremely low error floor.

There are mainly two ways of alleviating the error floor problem. One way is to improve the conventional sum-product decoding algorithm [6, 7]. The other way is to construct codes avoiding the detrimental combinatorial structures such as the trapping sets with short cycles [8] and absorbing sets [9] contributing to the error floor. Constructing large-girth codes is one promising solution to achieve the objective [10–12]. However, how to systematically construct *high-rate* QC-LDPC codes with a girth of ten or higher is still an open problem [13].

There have been some works on constructing QC-LDPC codes of large girth. The shortened array codes has been first investigated in [12], where only a subset of columns in array codes are retained to avoid the "cycle-governing" equations corresponding to various cycle lengths. In [14], a scheme combining shortened array codes and the Chinese remainder theorem has been proposed to construct QC-LDPC codes of large girth. However, previous works lack the theoretical support on the application of the shortened array codes to high-rate code construction. For shortened array codes, the maximum number of columns that remain after shortening plays an essential role in the range of the applicable code rate. The larger the number of remaining columns is, the higher the code rate is achieved. In [12], the maximum number of columns retained to avoid certain cycle conditions has been analyzed, but the achievable code rate of shortened array codes of girth-ten has not been fully characterized.

In this paper, we derive a general lower bound of the maximum size of retaining columns for shortened array codes and upper bounds of that for codes with girth-eight, girth-ten or above. The theoretical analysis provides us much insight, leading to a construction method for column-weight-three shortened array codes. We further compare our method with the conventional greedy algorithm [12]. We show that the proposed code construction method is effective in constructing column-weight-three shortened array codes in the sense that it can achieve code rates ranging from 0.727 to 0.842 with a comparable or much shorter code length.

The rest of the paper is organized as follows. Section 2 reviews the basics of shortened array codes and the previously established results. Section 3 presents our analysis on the shortened array codes. Section 4 focuses on the special case of column-weight-three shortened array codes. A code construction method and some results are presented in the same section. Section 5 concludes the paper.

2 Review of Array Codes

The general form for the parity-check matrix of an array code [15] is represented by

$$\mathbf{H} = \begin{bmatrix} \mathbf{I} & \mathbf{P}^{a_0 \cdot 1} & \cdots & \mathbf{P}^{a_0 \cdot (p-1)} \\ \mathbf{I} & \mathbf{P}^{a_1 \cdot 1} & \cdots & \mathbf{P}^{a_1 \cdot (p-1)} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{I} & \mathbf{P}^{a_{r-1} \cdot 1} & \cdots & \mathbf{P}^{a_{r-1} \cdot (p-1)} \end{bmatrix}, \quad (1)$$

where p denotes the number of columns and is a prime number; r is the number of block rows with $1 \leq r \leq p$; a_g's are distinct numbers with $0 \leq a_g \leq p-1$, for $g = 0, \ldots, r-1$; $\mathbf{I}$ is the identity matrix and $\mathbf{P}$ is a $p \times p$ circulant permutation matrix defined as

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}. \quad (2)$$

Since each column of $\mathbf{H}$ has a weight of r and each row has a weight of p, the code rate R is lower bounded by $R \geq 1 - r/p$. Throughout the paper, we call a_g the block-row indices, where $0 \leq g \leq r-1$; and h the block-column indices, where $0 \leq h \leq p-1$.

According to [1, Theorem 2.1], a cycle of length $2k$ associated with a sequence of circulant permutation matrices $\mathbf{P}^{a_{g_1}h_1}, \mathbf{P}^{a_{g_1}h_2}, \mathbf{P}^{a_{g_2}h_2}, \ldots, \mathbf{P}^{a_{g_k}h_k}, \mathbf{P}^{a_{g_k}h_1}$ exists if and only if

$$(a_{g_1} - a_{g_k})h_1 + (a_{g_2} - a_{g_1})h_2 + \cdots + (a_{g_k} - a_{g_{k-1}})h_k \equiv 0 \pmod{p}, \quad (3)$$

where the block-row indices a_{g_1} to a_{g_k} and the block-column indices h_1 to h_k of the permutation matrices satisfy $a_{g_l} \neq a_{g_{l+1}}$ and $h_l \neq h_{l+1}$, for $l = 1, 2, \ldots k-1$, $a_{g_k} \neq a_{g_1}$, and $h_k \neq h_1$. Thus, in order to avoid cycles of length eight, the selection of the block-row indices $\{a_0, a_1, \ldots, a_{r-1}\}$ in a shortened array code must take into account the following property [12].

Inevitability of cycle-eight due to block-row indices: For $a_{g_1}, a_{g_2}, a_{g_3}, a_{g_4} \in \{a_0, a_1, \ldots, a_{r-1}\}$, cycle-eight must exist if the following equality holds,

$$a_{g_1} + a_{g_3} - a_{g_2} - a_{g_4} = 0 \pmod{p}$$
$$\text{subject to } a_{g_1} \neq a_{g_2}, a_{g_1} \neq a_{g_4}, a_{g_2} \neq a_{g_3}, a_{g_3} \neq a_{g_4}. \quad (4)$$

Once the set of block-row indices $\{a_0, a_1, \ldots, a_{r-1}\}$ has been selected, the block-column indices can be regarded as variables and the cycle-governing equations in the form of (3) can be written as

$$\sum_{i=1}^{k} c_i x_i \equiv 0 \pmod{p}, \quad (5)$$

where the coefficient $c_i = a_{g_1} - a_{g_k}$ if $i = 1$ and $c_i = a_{g_i} - a_{g_{i+1}}$ if $i \neq 1$. Hence, the equation $\sum_{i=1}^{k} c_i = 0$ is always satisfied. An example showing the systems of cycle-governing equations corresponding to different sets of block-row indices is given in Table I. Any solution $\mathbf{x} = (x_1, x_2, \cdots, x_k)$ to (5) with $x_i \in \{0, 1, \cdots, p-1\}$ and $x_i \neq x_j$ for all $i \neq j$ is referred to as a *proper solution* over Z_p, where Z_p is the ring of integers modulo p. In the construction of shortened array codes, we aim to find a subset of block-column indices $\mathcal{S}(p;\Omega) \subseteq \{0, 1, \cdots, p-1\}$ such that $\mathcal{S}(p;\Omega)$ contains no proper solutions to a system of cycle-governing equations denoted by Ω.

Definition 1 *An equation* $\sum_{i=1}^{k} c_i x_i \equiv 0 \pmod{p}$ *over* Z_p *is of type* (l, m) *if* l *coefficients are positive and* m *are negative with* $l + m = k$. *Note that type* (l, m) *is equivalent to type* (m, l).

For example, all the cycle-six governing equations involve three variables, and they are of type (1,2), i.e.,

$$\Omega : (c_{m,1} + c_{m,2})x_3 \equiv c_{m,1}x_1 + c_{m,2}x_2 \pmod{p}, m = 1, \ldots, |\Omega|. \quad (6)$$

We further denote by $s(p;\Omega)$ the maximum number of columns retained in a shortened array code. Intuitively, $s(p;\Omega)$ is closely related to the range of the code rate that shortened array codes can achieve. In [12], the authors have characterized $s(p;\Omega)$ for a certain set of cycle-governing equations. In particular, they have derived a lower bound of $s(p;\Omega)$ for shortened array codes of girth-eight by slightly modifying the Behrend's construction method [16].

Theorem 1 *(Lower Bound of* $s(p;\Omega)$ *for shortened array codes of girth-eight [12]) Let* Ω *denote the system of cycle-six governing equations in the form of (6) and let* $V = \max_{m \in \{1,\ldots,|\Omega|\}} \{c_{m,2} + c_{m,1}\}$. *Then* $s(p;\Omega)$ *is lower-bounded by*

$$s(p, \Omega) \geq \gamma_1 p e^{-\gamma_2 \sqrt{\log p} - \frac{1}{2}\log\log p}(1 + o(1)),$$

where $\gamma_1 = V^2\sqrt{\frac{1}{2}\log V}$, $\gamma_2 = 2\sqrt{2\log V}$ *and* $o(1)$ *is a term vanishing as* $p \to \infty$.

However, how $s(p;\Omega)$ scales for shortened array codes of different girths has not been derived. In the following, we will first derive a general lower bound of $s(p;\Omega)$ for shortened array codes of various girths. Then we derive upper bounds of $s(p;\Omega)$ for codes with girth-eight, and girth-ten and above, respectively.

Table I
Cycle-governing equations for a column weight $r = 3$ and different block-row indices.

Block-row indices	Cycle-six governing equations (computed over Z_p)	Cycle-eight governing equations (computed over Z_p)
$r = 3$ $\{a_0, a_1, a_2\} = \{0, 1, 3\}$	$3x_1 - 2x_2 - x_3 = 0$	$3x_1 - 3x_2 - x_3 + x_4 = 0$ $3x_1 - 3x_2 - 2x_3 + 2x_4 = 0$ $x_1 + x_2 - x_3 - x_4 = 0$ $2x_1 + x_2 - x_3 - 2x_4 = 0$ $4x_1 - 3x_2 - x_3 = 0$ $2x_1 - x_2 - x_3 = 0$ $5x_1 - 3x_2 - 2x_3 = 0$
$r = 3$ $\{a_0, a_1, a_2\} = \{0, 1, 4\}$	$4x_1 - 3x_2 - x_3 = 0$	$5x_1 - 4x_2 - x_3 = 0$ $2x_1 - x_2 - x_3 = 0$ $7x_1 - 4x_2 - 3x_3 = 0$ $3x_1 - 2x_2 - x_3 = 0$ $4x_1 - 4x_2 - x_3 + x_4 = 0$ $4x_1 - 4x_2 - 3x_3 + 3x_4 = 0$ $x_1 + x_2 - x_3 - x_4 = 0$ $3x_1 + x_2 - x_3 - 3x_4 = 0$
$r = 3$ $\{a_0, a_1, a_2\} = \{0, 1, 5\}$	$5x_1 - 4x_2 - x_3 = 0$	$4x_1 - 4x_2 - x_3 + x_4 = 0$ $5x_1 - 4x_2 - 5x_3 + 4x_4 = 0$ $x_1 + x_2 - x_3 - x_4 = 0$ $5x_1 + x_2 - x_3 - 5x_4 = 0$ $4x_1 - 3x_2 - x_3 = 0$ $2x_1 - x_2 - x_3 = 0$ $9x_1 - 4x_2 - 5x_3 = 0$ $6x_1 - x_2 - 5x_3 = 0$
$r = 3$ $\{a_0, a_1, a_2\} = \{0, 1, 6\}$	$6x_1 - 5x_2 - x_3 = 0$	$5x_1 - 5x_2 - x_3 + x_4 = 0$ $6x_1 - 5x_2 - 6x_3 + 5x_4 = 0$ $x_1 + x_2 - x_3 - x_4 = 0$ $6x_1 + x_2 - x_3 - 6x_4 = 0$ $5x_1 - 4x_2 - x_3 = 0$ $2x_1 - x_2 - x_3 = 0$ $11x_1 - 5x_2 - 6x_3 = 0$ $7x_1 - x_2 - 6x_3 = 0$
$r = 3$ $\{a_0, a_1, a_2\} = \{0, 2, 5\}$	$5x_1 - 3x_2 - 2x_3 = 0$	$3x_1 - 2x_2 - x_3 = 0$ $2x_1 - x_2 - x_3 = 0$ $8x_1 - 5x_2 - 3x_3 = 0$ $7x_1 - 2x_2 - 5x_3 = 0$ $5x_1 - 2x_2 - 5x_3 + 2x_4 = 0$ $3x_1 - 2x_2 - 3x_3 + 2x_4 = 0$ $x_1 + x_2 - x_3 - x_4 = 0$ $5x_1 + 3x_2 - 5x_3 - 3x_4 = 0$
$r = 3$ $\{a_0, a_1, a_2\} = \{0, 2, 7\}$	$7x_1 - 5x_2 - 2x_3 = 0$	$5x_1 - 2x_2 - 5x_3 + 2x_4 = 0$ $7x_1 - 2x_2 - 7x_3 + 2x_4 = 0$ $x_1 + x_2 - x_3 - x_4 = 0$ $7x_1 + 5x_2 - 7x_3 - 5x_4 = 0$ $5x_1 - 3x_2 - 2x_3 = 0$ $2x_1 - x_2 - x_3 = 0$ $9x_1 - 7x_2 - 2x_3 = 0$ $12x_1 - 7x_2 - 5x_3 = 0$

3 Analysis of the Shortened Array Codes

3.1 General Lower Bound of $s(p;\Omega)$

A lower bound of $s(p;\Omega)$ for shortened array codes with different girths can be obtained using the following theorem.

Theorem 2 (General Lower Bound) *Given a system of cycle-governing equations Ω up to cycle-2k, there exists a sequence $s_1, s_2, \cdots, s_n$ with $1 = s_1 < s_2 < \cdots < s_n \leq |\Omega|k(n-1)^{k-1}$ such that $\mathcal{S}(p;\Omega) = \{s_1, s_2, \cdots, s_n\}$ does not contain proper solutions to Ω. Then $s(p;\Omega)$ for shortened array codes of girth-$2(k+1)$ is lower-bounded by $s(p;\Omega) \geq p^{\frac{1}{k-1}}(|\Omega|k)^{-\frac{1}{k-1}}$.*

Proof: The proof is based on the idea of the greedy construction of $\mathcal{S}(p;\Omega)$ specified in [17, Theorem 2.1] and we sketch the proof here.

Obviously we can choose $s_1 = 1$. Assume that a sequence $s_1, s_2, \cdots, s_{n-1}$ has been chosen such that the sequence does not contain any proper solutions to Ω. Now we want to find a s_n such that adding it to the sequence would not create any proper solutions to Ω. Since the cycle-2k equations involve at most k variables, we want to find s_n satisfying

$$c_{m,i_0}s_n \not\equiv -\sum_{1\leq i\leq k, i\neq i_0} c_{m,i}x_i \pmod{p}, \quad \forall i_0 = 1, \cdots, k, \forall m = 1, \cdots, |\Omega|, \qquad (7)$$

where for all $1 \leq i \leq k$ and $i \neq i_0$, x_i are distinct integers and $x_i \in \{s_1, \cdots, s_{n-1}\}$. For a fixed i_0 and m, x_i $(1 \leq i \leq k, i \neq i_0)$ can be taken from at most $\binom{n-1}{k-1}(k-1)!$ possible values and hence the constraint function (7) excludes no more than $\binom{n-1}{k-1}(k-1)!$ values for s_n. Considering that there are k possible values of i_0 and $|\Omega|$ possible values of m, the cycle-governing equations exclude at most $k|\Omega|\binom{n-1}{k-1}(k-1)!$ possible values for s_n. Since $k|\Omega|\binom{n-1}{k-1}(k-1)! < k|\Omega|(n-1)^{k-1}$, we can always find a $s_n < k|\Omega|(n-1)^{k-1}$ so that adding it to the set would not create proper solutions.

Following this process, we can extend the set $\mathcal{S}(p;\Omega) = \{s_1, s_2, \cdots, s_n\}$ to the extent that $p^{\frac{1}{k-1}}(|\Omega|k)^{-\frac{1}{k-1}} \leq n \leq p^{\frac{1}{k-1}}(|\Omega|k)^{-\frac{1}{k-1}} + 1$. Hence, $s(p;\Omega) \geq n \geq p^{\frac{1}{k-1}}(|\Omega|k)^{-\frac{1}{k-1}}$. ∎

Note that Theorem 2 provides an explicit code construction method of choosing $\mathcal{S}(p;\Omega)$ and a general lower bound of $s(p;\Omega)$ for shortened array codes of various girths, but the bound is not tight in general. For example, for the case of girth-eight codes, Theorem 2 gives a lower bound scaling at $\Theta(\sqrt{p})$ which is loose compared with the result given in Theorem 1.

3.2 $s(p;\Omega)$ for Shortened Array Codes of Girth-Eight

A lower bound of $s(p;\Omega)$ for shortened array codes of girth-eight has been derived in [12]. Here we will derive an upper bound of $s(p;\Omega)$.

Theorem 3 *(Upper Bound of $s(p;\Omega)$ for shortened array codes of girth-eight) Let Ω denote the system of cycle-six governing equations in the form of (6). Then $s(p;\Omega)$ is upper-bounded by $s(p;\Omega) \leq p(\log\log p)^{-c(v)}$, where $c(v) = 2^{-2^{v+10}}$ and $v = \min_{m\in\{1,\ldots,|\Omega|\}}\{c_{m,2} + c_{m,1}\}$.*

Proof: Let $v_m = c_{m,1} + c_{m,2}$ and $\mathcal{S}(p;\Omega)$ be the set of column indices avoiding the proper solutions to Ω. Then $\mathcal{S}(p;\Omega)$ must not contain an arithmetic progression of length $v_m + 1$. We prove this by contradiction, showing that it would contain proper solutions to (6).

Assume $\mathcal{S}(p;\Omega)$ contains an arithmetic progression of length $v_m + 1$. Then we can take three variables from $\mathcal{S}(p;\Omega)$ such that $x_3 = x_1 + c_{m,2}d$ and $x_2 = x_1 + (c_{m,1} + c_{m,2})d$, where d is the common difference of successive members in the arithmetic progression. Obviously x_1, x_2 and x_3 are distinct integers that satisfy the equation $(c_{m,1} + c_{m,2})x_3 \equiv c_{m,1}x_1 + c_{m,2}x_2$. It contradicts the fact that $\mathcal{S}(p;\Omega)$ avoids the proper solutions.

By making use of [18, Theorem 1.3] that any subset of $\{1, \ldots, p\}$ of size not less than $p(\log\log p)^{-c(v_m)}$ contains an arithmetic progression of length $v_m + 1$, where $c(v_m) = 2^{-2^{v_m+10}}$, we can prove that $s(p;\Omega) \leq p(\log\log p)^{-c(v_m)}$. Since the inequality holds for all $m = 1, 2, \ldots, |\Omega|$, taking $v = \min_{m\in\{1,\cdots,|\Omega|\}}\{c_{m,2} + c_{m,1}\}$ yields the tightest upper bound for $s(p;\Omega)$ as stated in the theorem. ∎

The upper bound of $s(p;\Omega)$ derived above is of almost the same order as the lower bound of $s(p;\Omega)$ in Theorem 1 except for the sublinear terms. In particular, by combining both theorems we can obtain the asymptotic order of $s(p;\Omega)$ for girth-eight shortened array codes as $\lim_{p\to\infty} \log s(p;\Omega)/\log p = 1$.

3.3 $s(p;\Omega)$ for Shortened Array Codes with a Girth of Ten and Above

In this section, we consider $s(p;\Omega)$ for shortened array codes of girth-ten or higher.

Definition 2 (Symmetric Equations over Z_p) *An equation is symmetric over Z_p if the number of variables k is even and the equation can be arranged as the following form*

$$c_1x_1 + \cdots + c_{k/2}x_{k/2} \equiv c_1x_{k/2+1} + \cdots + c_{k/2}x_k \ (mod\ p). \quad (8)$$

In the following, we obtain an upper bound of the set of proper solutions to symmetric equations over Z_p by slightly modifying [17, Theorem 3.2],.

Lemma 1 *Denote by $\mathcal{A}$ the set in which there is no proper solution to a symmetric equation over Z_p in k variables, then we have $|\mathcal{A}| \leq \sqrt{k(k-1)p/2}$ and hence $|\mathcal{A}| = O(\sqrt{p})$.*

Proof: Consider a set $\mathcal{A} \subseteq \{0, 1, \ldots, p-1\}$ and the variables $x_i \in \mathcal{A}$, for $i = 1, \ldots, k$. Let $t(n)$ be the number of solutions of $c_1x_1 + \cdots + c_{k/2}x_{k/2} = n(\text{mod } p)$. Then $\sum_{n=0}^{p-1} t(n) = |\mathcal{A}|^{k/2}$ and the total number of solutions to (8) is $\sum_{n=0}^{p-1} t^2(n)$.

Consider the solutions to (8) in which $x_i = x_j$ for certain $1 \leq i < j \leq k$, the number of such solutions is not more than $|\mathcal{A}|^{k-2}$. Taking into account the $\binom{k}{2}$ possible choices i and j, the total number of solutions containing at least two identical values is not more than $\binom{k}{2}|\mathcal{A}|^{k-2}$. For a set $\mathcal{A}$ containing only non-proper solutions (i.e., no proper solutions) to (8) over Z_p, it must satisfy $\sum_{n=0}^{p-1} t^2(n) \leq \binom{k}{2}|\mathcal{A}|^{k-2}$. On the other hand, we can apply the Cauchy-Schwarz inequality to the lower bound $\sum_{n=0}^{p-1} t^2(n)$ and obtain $p\sum_{n=0}^{p-1} t^2(n) \geq \left(\sum_{n=0}^{p-1} t(n)\right)^2 = |\mathcal{A}|^k$. Therefore, we can obtain an upper bound of $|\mathcal{A}|$ as $|\mathcal{A}| \leq \sqrt{(k-1)p/2}$. ∎

Since the cycle-eight governing equations must include at least one symmetric equation involving four variables in the form of $x_1 + x_2 \equiv x_3 + x_4 \ (\text{mod } p)$, applying $k = 4$ to Lemma 1 yields an upper bound of $s(p;\Omega)$.

Corollary 1 *For shortened array codes of girth-ten or higher, $s(p;\Omega)$ is upper-bounded by $s(p;\Omega) \leq \sqrt{6p}$ and $s(p;\Omega) = O(\sqrt{p})$.*

4 Construct Column-Weight-Three Shortened Array Codes with Girth-Ten and Their Performance

For the case of girth-ten codes, Theorem 2 gives a lower bound scaling as $\Theta(p^{\frac{1}{3}})$ while Corollary 1 gives an upper bound scaling as $\Theta(p^{\frac{1}{2}})$. Therefore, either one

Algorithm 1 Proposed Code Construction

Let $V = \max_{m=1,2,\cdots,|\Omega|}\{c_{m,1} + c_{m,2}\}$.
for the largest prime number q with $q < \sqrt{p}/V$ **do**
1) Construct a set $\mathcal{X} = \{x' + Vqx : 0 \leq x \leq q-1, x' = x^2 \pmod q\}$.
2) Initialize $\mathcal{S}(p;\Omega)$ as the largest subset of $\mathcal{X}$ that avoids proper solutions to cycle-governing equations in four variables.[1]
3) Update $\mathcal{S}(p;\Omega)$ by sequentially adding the integers in $[0, p-1]\backslash\mathcal{S}(p;\Omega)$ that would not create proper solutions to Ω.
4) Repeat 3) until all integers have been exhausted.
end for

bound or both bounds are loose. In the following, we will show that when the shortened array codes have a column weight of three, the lower bound given in Theorem 2 is loose and we will derive a much tighter lower bound for this case. Although this is merely a special case of array codes, it is useful in searching for large-girth high-rate QC-LDPC codes [12, 13].

Theorem 4 *Given a column-weight-three array code, $s(p;\Omega)$ for shortened array codes of girth-ten is lower bounded by $s(p;\Omega) \geq \beta\sqrt{p}e^{-\alpha\sqrt{\log p}}$ for some positive β and α.*

Proof: Please refer to the Appendix. ■

According to Theorem 4, $\lim_{p\to\infty} s(p;\Omega)/p^{\frac{1}{2}-\epsilon} = \infty$, for any $\epsilon > 0$. It can be observed that this lower bound is almost at the same order of the upper bound given in Corollary 1 and thus is much tighter than that given in Theorem 2.

Constructing high-rate QC-LDPC codes of girth-ten or higher is a challenging problem. In Algorithm 1, we propose a code construction algorithm for such a purpose. Then, we compare the minimum p required to achieve the same code rate R using our proposed method and that using the greedy construction method [12].

The results are shown in Table II. It can be seen that our proposed code construction can achieve the same code rate using either a comparable p or a much smaller p compared with the greedy construction method. For example, for a code rate of $R = 0.750$, our proposed method finds a minimum p value of $p_{\text{proposed}} = 1319$ and the greedy algorithm obtains a minimum value of $p_{\text{greedy}} = 1439$. When $R = 0.786$, we obtain $p_{\text{proposed}} = 1787$ and $p_{\text{greedy}} = 2179$. We further compare the error performance of the codes constructed when the code rate equals 0.750 and 0.769. The additive white Gaussian noise (AWGN) channel is assumed and a maximum of 50 iterations are used in the belief propagation decoding process [19]. Figure 1 presents the bit error rate (BER) performance of girth-ten codes constructed using the proposed method and that constructed using the greedy method. We observe that the codes constructed using the two different methods produce similar bit error performance.

[1] $\mathcal{X}$ is guaranteed to avoid proper solutions to all cycle-governing equations in three variables as proved in the Appendix.

Algorithm 2 Random Code Construction

Initilization: Set a maximum iteration number M and a target code rate R.
Iteration: Set the iteration number to 1. Repeat the following steps until the iteration number reaches M or the code rate achieves R.
1) Initialize $S(p;\Omega)$ by randomly choosing three distinct integers from $[0, p-1]$ that avoid proper solutions to the cycle-governing equations in three variables.
2) Update $S(p;\Omega)$ by randomly adding integers in $[0, p-1]\backslash\mathcal{S}(p;\Omega)$ that would not create proper solutions to Ω.
3) Update the code rate to be$(|S(p;\Omega)| -$ number of block rows$)/|S(p;\Omega)|$.
4) Repeat 2) and 3) until all integers have been exhausted.
5) If the code rate R has been achieved, stop; otherwise, increment the iteration number and go to 1).

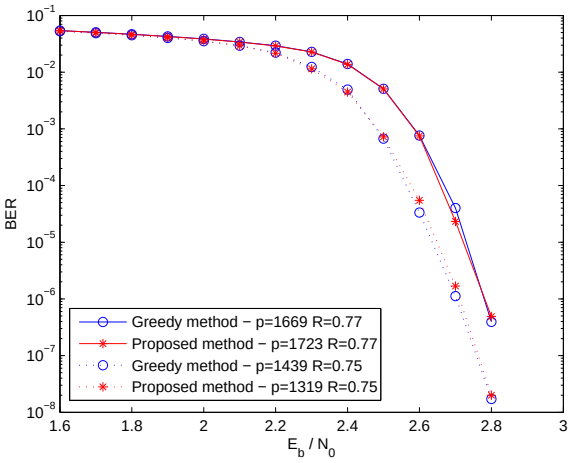

Figure 1. Bit error rate of the array code constructed using the proposed method and that constructed using the greedy algorithm. The additive white Gaussian noise (AWGN) channel is assumed and the maximum number of decoding iteration is 50.

In order to further evaluate the effectiveness of our proposed code construction algorithm, we compare it with the random code construction method shown in Algorithm 2. This method is based on heuristic computer search. Given a submatrix size p and a maximum iteration number M, the random code construction updates $S(p;\Omega)$ by randomly adding integers in $\{0,\ldots,p-1\}\backslash\mathcal{S}(p;\Omega)$ that would not create proper solutions to Ω. It stops when all the integers in $\{0,\ldots,p-1\}$ have been exhausted. The same process repeats for M iterations and it outputs the maximum achievable code rate. In Table II, we show the achievable code rate using the random construction method with $M = 500$ iterations. Moreover, p is chosen as those found by the greedy construction method and our proposed method.

The results have indicated that (i) increasing the value of p may not produce a higher code rate (first case); (ii) the achievable code rate may be further enhanced by the random construction method (second case); and (iii) the achievable code rate is not enhanced in most cases when the smaller one between p_{greedy} and p_{proposed} is used in the random construction method.

Table II
COMPARISON OF THE MINIMUM p REQUIRED TO CONSTRUCT GIRTH-TEN SHORTENED ARRAY CODES FOR DIFFERENT CODE RATES R USING THE GREEDY CODE CONSTRUCTION [12] (p_{greedy}) AND OUR PROPOSED METHOD (p_{proposed}). THE RANDOM CONSTRUCTION METHOD INVOLVES 500 ITERATIONS. THE COLUMN WEIGHT OF THE CODES IS $r = 3$ AND THE BLOCK-ROW INDICES IS $\{0, 1, 3\}$.

Number of Columns	Code Rate	Minimum p required	Random Code Construction
11	$R = 0.727$	$p_{\text{proposed}} = 911$ $p_{\text{greedy}} = 1039$	$p = p_{\text{proposed}} : R = 0.727$ $p = p_{\text{greedy}} : R = 0.727$
12	$R = 0.750$	$p_{\text{proposed}} = 1319$ $p_{\text{greedy}} = 1439$	$p = p_{\text{proposed}} : R = 0.769$ $p = p_{\text{greedy}} : R = 0.769$
13	$R = 0.769$	$p_{\text{proposed}} = 1723$ $p_{\text{greedy}} = 1669$	$p = p_{\text{proposed}} : R = 0.786$ $p = p_{\text{greedy}} : R = 0.769$
14	$R = 0.786$	$p_{\text{proposed}} = 1787$ $p_{\text{greedy}} = 2179$	$p = p_{\text{proposed}} : R = 0.786$ $p = p_{\text{greedy}} : R = 0.800$
15	$R = 0.800$	$p_{\text{proposed}} = 2579$ $p_{\text{greedy}} = 2797$	$p = p_{\text{proposed}} : R = 0.800$ $p = p_{\text{greedy}} : R = 0.813$
16	$R = 0.813$	$p_{\text{proposed}} = 2999$ $p_{\text{greedy}} = 2971$	$p = p_{\text{proposed}} : R = 0.824$ $p = p_{\text{greedy}} : R = 0.813$
17	$R = 0.824$	$p_{\text{proposed}} = 3449$ $p_{\text{greedy}} = 3407$	$p = p_{\text{proposed}} : R = 0.833$ $p = p_{\text{greedy}} : R = 0.824$
18	$R = 0.833$	$p_{\text{proposed}} = 3823$ $p_{\text{greedy}} = 4079$	$p = p_{\text{proposed}} : R = 0.833$ $p = p_{\text{greedy}} : R = 0.842$
19	$R = 0.842$	$p_{\text{proposed}} = 4493$ $p_{\text{greedy}} = 4861$	$p = p_{\text{proposed}} : R = 0.842$ $p = p_{\text{greedy}} : R = 0.850$

5 Conclusion

We have derived a general lower bound of the maximum number of columns retained for the shortened array codes. We have further derived the corresponding upper bounds for the shortened array codes with girth-eight, girth-ten or above. Such upper bounds are closely related to the range of code rate that the shortened array codes can achieve. Based on the our derivations, we have proposed a method of constructing high-rate QC-LDPC of girth-ten with column-weight-three. We have shown that the proposed method is more effective than the conventional greedy algorithm in the sense that the minimum length of the codes constructed using our proposed method to achieve different code rates is comparative or much shorter than those constructed using the greedy construction method. Moreover, codes constructed using the proposed algorithm produce similar error performance as those constructed using the greedy construction method.

Appendix
Proof of Theorem 4

The proof of Theorem 4 is based on Erdös-Turán's construction in [20]. We first show that the integers constructed using Erdös-Turán's method can avoid proper solutions to equations in three variables. Equations in four variables can be further transformed into a set of equations in three variables. Then based on Theorem 1, an explicit lower bound of $s(p;\Omega)$ can be obtained.

Given a system of cycle-six and cycle-eight governing equations Ω, we divide it into two disjoint sets Ω_1 and Ω_2 with $\Omega = \Omega_1 \cup \Omega_2$, where Ω_1 denotes the set of equations in three variables and Ω_2 denotes the set of equations in four variables. It can be seen that Ω_1 only contains equations of type (2,1). It can also be verified from (3) that in the case of three block rows, Ω_2 *only* contains symmetric equations for column-weight-3 array codes. Thus Ω_1 and Ω_2 can be described as follows:

$$\Omega_1 : \ (c_{m,1} + c_{m,2})X_3 \equiv c_{m,1}X_1 + c_{m,2}X_2 \pmod{p} \quad m = 1, 2, \ldots, |\Omega_1| \tag{9}$$

$$\Omega_2 : \ c_{m,1}X_1 + c_{m,2}X_2 \equiv c_{m,1}X_3 + c_{m,2}X_4 \pmod{p} \quad m = |\Omega_1| + 1, |\Omega_1| + 2, \ldots, |\Omega| \tag{10}$$

Let $V = \max_{m=1,2,\ldots,|\Omega|}\{c_{m,1} + c_{m,2}\}$ and take a largest prime q such that $q < \sqrt{p}/V$. Define a set $\mathcal{X} = \{x' + Vqx : 0 \le x \le q-1, x' = x^2 \pmod{q}\}$, which is based on Erdös-Turán construction [20]. Then $|\mathcal{X}| = q$ and $\mathcal{X} \subset [0, Vq^2)$.

Lemma 2 *$\mathcal{X}$ does not contain proper solutions to the set of equations in three variables Ω_1 as defined in (9).*

Proof: Given any equation in three variables $(c_{m,1} + c_{m,2})X_3 \equiv c_{m,1}X_1 + c_{m,2}X_2 \pmod{p}$, we take distinct $X_i \in \mathcal{X}$ and denote it as $X_i = x'_i + Vqx_i$ $(i = 1, 2, 3)$. Since both sides of the congruence are in the range of $[0, p)$, they are equal, i.e., $(c_{m,1} + c_{m,2})X_3 = c_{m,1}X_1 + c_{m,2}X_2$. We can further obtain

$$(c_{m,1} + c_{m,2})x'_3 + Vq(c_{m,1} + c_{m,2})x_3 = c_{m,1}x'_1 + c_{m,2}x'_2 + Vq(c_{m,1}x_1 + c_{m,2}x_2). \tag{11}$$

Hence, we have $(c_{m,1} + c_{m,2})x'_3 \equiv c_{m,1}x'_1 + c_{m,2}x'_2 \pmod{Vq}$. Since both sides of the congruence

are in the range of $[0, Vq)$, they are equal. By substituting $x'_i = x_i^2 \pmod q$ $(i = 1, 2, 3)$, we have

$$(c_{m,1} + c_{m,2})x_3^2 \equiv c_{m,1}x_1^2 + c_{m,2}x_2^2 \pmod q \tag{12}$$

and further based on (11), we can obtain

$$(c_{m,1} + c_{m,2})x_3 = (c_{m,1}x_1 + c_{m,2}x_2). \tag{13}$$

By first taking square on both sides of (13) and multiplying both sides of (12) by $(c_{m,1} + c_{m,2})$, then combining these two quantities, we have $c_{m,1}c_{m,2}(x_1 - x_2)^2 \equiv 0 \pmod q$. Consequently, it implies $x_1 = x_2$ and we can conclude that the Erdös-Turán construction can avoid proper solutions to all equations in three variables. ■

Lemma 2 implies that we only need to consider the impact of $\mathcal{X}$ on the solutions to Ω_2, which is summarized as the following Lemma.

Lemma 3 *$\mathcal{X}$ will equivalently transform a set of symmetric equations Ω_2 into a collection of equations of type (3,1).*

Proof: Given any symmetric equation $c_{m,1}X_1 + c_{m,2}X_2 \equiv c_{m,1}X_3 + c_{m,2}X_4 \pmod p$ in the set Ω_2, we take $X_i \in \mathcal{X}$ and denote it as $X_i = x'_i + Vqx_i$, $i = 1, 2, 3, 4$. Following the same process in proving Lemma 2, we obtain

$$c_{m,1}x_1 + c_{m,2}x_2 = c_{m,1}x_3 + c_{m,2}x_4 \tag{14}$$

$$(x_1 - x_2)^2 \equiv (x_3 - x_4)^2 \pmod q. \tag{15}$$

There are two cases concerning (15) as follows,

Case 1: $x_1 - x_2 = x_3 - x_4 \pmod q$

Combining (14) yields $x_1 = x_3$ and $x_2 = x_4$, which is not a proper solution.

Case 2: $x_1 - x_2 = x_4 - x_3 \pmod q$

Without loss of generality, we assume $c_{m,1} \geq c_{m,2}$ for all $m = |\Omega_1| + 1, |\Omega_1| + 2, \ldots, |\Omega|$. Combining (14) yields

$$2c_{m,1}x_1 \equiv (c_{m,1} - c_{m,2})x_2 + (c_{m,1} + c_{m,2})x_4 \pmod q \tag{16}$$

$$(c_{m,1} + c_{m,2})x_1 \equiv (c_{m,1} - c_{m,2})x_3 + 2c_{m,2}x_4 \pmod q. \tag{17}$$

Consequently, the set of equations in Ω_2 are transformed into a collection of equations of the form as (16) and (17), which can be regarded as a collection of equations of type (3,1) as

$$\Omega'_2: \quad b_m x_1 \equiv \sum_{i=1}^{3} c'_{m,i}x_{i+1} \pmod q \qquad m = 1, 2, \ldots, 2 \times |\Omega_2|, \tag{18}$$

where $b_m = \sum_{i=1}^{3} c'_{m,i}$ and $c'_{m,i} \geq 0$ for all i and m. According to [12, Theorem 11], we know there exists a set $\mathcal{S}(p;\Omega) \subset \mathcal{X}$ such that $\mathcal{S}(p;\Omega)$ does not contain proper solutions to Ω'_2 and $|\mathcal{S}(p;\Omega)| \geq qe^{-\alpha\sqrt{\log q}}$ for a certain α. Considering that there exists a prime q with $\beta\sqrt{p} \leq q \leq \sqrt{p}/V$ for a certain $\beta < 1/V$, we conclude that $s(p;\Omega)$ can be lower-bounded by $s(p;\Omega) \geq \beta\sqrt{p}e^{-\alpha\sqrt{\log p}}$. ■

References

[1] M. P. C. Fossorier, "Quasicyclic low-density parity-check codes from circulant permutation matrices," *IEEE Transactions on Information Theory*, vol. 50, no. 8, pp. 1788–1793, Aug. 2004.

[2] R. M. Tanner, D. Sridhara, A. Sridharan, T. E. Fuja, and D. J. Costello, Jr., "LDPC block and convolutional codes based on circulant matrices," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 2966–2984, Dec. 2004.

[3] Z. Li, L. Chen, L. Zeng, S. Lin, and W. Fong, "Efficient encoding of quasi-cyclic low-density parity-check codes," *IEEE Transactions on Communications*, vol. 53, no. 11, pp. 1973–1973, Nov. 2005.

[4] W. M. Tam, F. C. M. Lau, and C. K. Tse, "A class of QC-LDPC codes with low encoding complexity and good error performance," *IEEE Communications Letters*, vol. 14, no. 2, pp. 169–171, Feb. 2010.

[5] I. B. Djordjevic, M. Arabaci, and L. L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 27, no. 16, pp. 3518–3530, Aug. 2009.

[6] N. Varnica, M. P. C. Fossorier, and A. Kavcic, "Augmented belief propagation decoding of low-density parity check codes," *IEEE Transactions on Communications*, vol. 55, no. 7, pp. 1308–1317, July 2007.

[7] Z. Zhang, L. Dolecek, B. Nikolic, V. Anantharam, and M. J. Wainwright, "Lowering LDPC error floors by post-processing," in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM 2008)*, New Orleans, LO, 30 Nov.–4 Dec. 2008, pp. 1–6.

[8] X. Zheng, F. C. M. Lau, and C. K. Tse, "Constructing short-length irregular LDPC codes with low error floor," *IEEE Transactions on Communications*, vol. 58, no. 10, pp. 2823–2834, Oct. 2010.

[9] Z. Zhang, L. Dolecek, B. Nikolic, V. Anantharam, and M. Wainwright, "GEN03-6: Investigation of error floors of structured low-density parity-check codes by hardware emulation," in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM 2006)*, San Francisco, CA, 27 Nov.–1 Dec. 2006, pp. 1–6.

[10] Y. Wang, J. S. Yedidia, and S. C. Draper, "Construction of high-girth QC-LDPC codes," in *Proc. 5th International Symp. on Turbo Codes and Related Topics*, Lausanne, 1–5 Sep. 2008, pp. 180–185.

[11] L. Lan, L. Zeng, Y. Y. Tai, L. Chen, S. Lin, and K. Abdel-Ghaffar, "Construction of quasi-cyclic LDPC codes for AWGN and binary erasure channels: A finite field approach," *IEEE Transactions on Information Theory*, vol. 53, no. 7, pp. 2429–2458, July 2007.

[12] O. Milenkovic, N. Kashyap, and D. Leyba, "Shortened array codes of large girth," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3707–3722, Aug. 2006.

[13] I. B. Djordjevic, L. Xu, T. Wang, and M. Cvijetic, "Large girth low-density parity-check codes for long-haul high-speed optical communications," in *Proc. Conf. Optical Fiber communication/National Fiber Optic Engineers Conf. (OFC/NFOEC 2008)*, San Diego, CA, 24-28 Feb. 2008, pp. 1–3.

[14] X. Jiang and M. H. Lee, "Large girth quasi-cyclic LDPC codes based on the Chinese remainder theorem," *IEEE Communications Letters*, vol. 13, no. 5, pp. 342–344, May 2009.

[15] J. L. Fan, "Array codes as low-density parity-check codes," in *Proc. 2nd International Symp. on Turbo Codes and Related Topics*, Brest, France, Sep. 2000, pp. 545–546.

[16] F. A. Behrend, "On sets of integers which contain no three terms in arithmetical progression." *Proc. National Academy of Science of the United States of America*, vol. 32, no. 12, pp. 331–332, Dec. 1946.

[17] I. Z. Ruzsa, "Solving a linear equation in a set of integers I," *Acta Arithmetica*, vol. LXV, no. 3, pp. 259–282, 1993.

[18] W. T. Gowers, "A new proof of Szemerédi's theorem," *Geometric and Functional Analysis*, vol. 11, no. 3, pp. 465–588, Aug. 2001.
[19] X. Zheng, F. C. M. Lau, C. K. Tse, and S. C. Wong, "Techniques for improving block error rate of LDPC decoders," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 2006)*, Island of Kos, 21–24 May 2006.
[20] P. Erdös and P. Turán, "On a problem of Sidon in additive number theory and some related problems," *Journal London of Mathematical Society*, vol. s1-16, no. 4, pp. 212–215, 1941.

Xu Chen received his B.E. degree from Sun Yat-sen (Zhongshan) University, China, in 2007 and his M.S. degree from Purdue University, USA, in 2009. From 2009 to 2011, he was a research assistant at the Department of Electronic and Information Engineering, the Hong Kong Polytechnic University. He is currently working towards his Ph.D. degree at Northwestern University, USA. His research interests include coding theory, optimization and cooperative communications.

Francis C.M. Lau received the BEng (Hons) degree in electrical and electronic engineering and the PhD degree from King's College London, University of London, UK, in 1989 and 1993, respectively.

He is a Professor and Associate Head at the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong. He is also a senior member of IEEE. He is the co-author of *Chaos-Based Digital Communication Systems* (Heidelberg: Springer-Verlag, 2003) and *Digital Communications with Chaos: Multiple Access Techniques and Performance Evaluation* (Oxford: Elsevier, 2007). He is also a co-holder of two US patent, one pending US patent and one pending international patent. He has published over 200 papers. His main research interests include channel coding, cooperative networks, wireless sensor networks, chaos-based digital communications, applications of complex-network theories, and wireless communications.

He served as an associate editor for *IEEE Transactions on Circuits and Systems II* in 2004–2005 and *IEEE Transactions on Circuits and Systems I* in 2006–2007. He was also an associate editor of *Dynamics of Continuous, Discrete and Impulsive Systems, Series B* from 2004 to 2007 and was a co-guest editor of *Circuits, Systems and Signal Processing* for the special issue "Applications of Chaos in Communications" in 2005. He is currently a guest associate editor of *International Journal and Bifurcation and Chaos.*