sensors

MDPI

*Article*

# Constructing Grith Eight GC-LDPC Codes Based on the GCD FLRM Matrix with a New Lower Bound

**Kun Zhu and Hongwen Yang ***

School of Information and Communication, Beijing University of Posts and Telecommunications, Beijing 100876, China
* Correspondence: yanghong@bupt.edu.cn

**Abstract:** By connecting multiple short, local low-density parity-check (LDPC) codes with a global parity check, the globally coupled (GC) LDPC code can attain high performances with low complexities. The typical design of a local code is a quasi-cyclic (QC) LDPC for which the code length is proportional to the size of circulant permutation matrix (CPM). The greatest common divisor (GCD)-based full-length row multiplier (FLRM) matrix is constrained by a lower bound of CPM size to avoid six length cycles. In this paper, we find a new lower bound for the CPM size and propose an algorithm to determine the minimum CPM size and the corresponding FLRM matrix. Based on the new lower bound, two methods are proposed to construct the GC-QC-LDPC code of grith 8 based on the GCD based FLRM matrix. With the proposed algorithm, the CPM size can be 45% less than that given by sufficient condition of girth 8. Compared with the conventional GC-LDPC construction, the codes constructed with the proposed method have improved performance and are more flexible in code length and code rate design.

**Keywords:** full-length row multiplier matrix; greatest common divisor; globally coupled LDPC; large girth



**Citation:** Zhu, K.; Yang, H. Constructing Grith Eight GC-LDPC Codes Based on the GCD FLRM Matrix with a New Lower Bound. *Sensors* **2022**, *22*, 7335. https://doi.org/10.3390/s22197335

Academic Editor: Davy P. Gaillot

Received: 3 August 2022
Accepted: 22 September 2022
Published: 27 September 2022



## 1. Introduction

Channel coding has always been one of the most important underlying technologies in communication systems. Capacity-achieving codes such as turbo codes, low-density parity-check (LDPC) codes, polar codes, and quantum codes have been proposed together with various near-maximum likelihood(ML)-decoding algorithms such as belief-propagation, list-decoding, guessing random additive noise decoding (GRAND), etc. [1–3]. LDPC codes have attracted much attention for its low decoding complexity and excellent performance [4–7]. The structure of quasi-cyclic (QC) LDPC code is relatively simple, so it is beneficial for hardware implementation. Therefore, QC-LDPC code plays an important role in many communication protocols, including wireless sensor networks [8–10]. Since regular LDPC codes have lower error-floor [11] and irregular LDPC codes can be easily obtained by transforming the well-designed regular codes [12,13], the construction of regular full-ank QC-LDPC codes is a very popular topic [14–18].

The full-length row multiplier (FLRM) matrix is a typical regular matrix used in construction of QC-LDPC codes. The FLRM matrix consists of multiple circulant permutation matrices (CPMs) [15–18] arranged in $J$ rows and $L$ columns. The exponent matrix of FLRM matrix is derived from the product of row coefficient vector and column indexes. Girth in Tanner graph is one of the most important characters of LDPC codes for it determines the minimum distance of the code. In addition, short cycles will produce trapping sets, stopping sets, and absorption sets, resulting in heavy error-floor and performance degradation [19]. Hence, many scholars devote themselves to improving the girth of the LDPC code [20–22]. It has been shown in [17] that the girth of FLRM codes cannot exceed eight, since there exist cycles of length eight regardless of the CPM's size. In [16], the greatest

common divisor (GCD) scheme was proposed to construct the grith 8 FLRM matrix. As the general form of CPM, the affine permutation matrices (APMs) can also adopt the GCD scheme to construct LDPC code of girth 8 [23]. Recently, the authors of [18] optimized the row coefficient's vector to find the lower bound of CPM size for each pair of $(J, L)$ and further confirmed the validity of the GCD construction method for the FLRM matrix with girth 8.

A globally coupled (GC) LDPC code is a new type of coupled LDPC code proposed by Juane Li [24] where multiple short local LDPC codes are connected with a set of global check nodes. As a result, GC-LDPC can effectively realize a long code with multiple short component codes, avoiding the construction of a completely new longer LDPC code [25]. The GC-LDPC code is adept in correcting erasures clustered in bursts and performs greatly under additive white Gaussian noise (AWGN) channels and binary erasure channels (BECs). With local/global two-phase decoding, the decoders of the component LDPC codes are reusable [26]. In [27], the Reed–Solomon code-based construction of GC-LDPC code was proposed. In [26], the array dispersion was applied to scale the GC-LDPC code, which makes the design of the GC-LDPC code more flexible in code-length selection. With the Reed–Solomon-Like construction [28], local codes and global coupling part of GC-LDPC code can be constructed separately. In addition, the protograph-based GC-LDPC code [29] and tail-biting GC-LDPC code [30] perform well. In [26,31], GC-LDPC codes were used for NAND Flash, and the relative independent structure and local/global two-phase decoding can reduce the critical path and decoding latency. Recently, ref. [25] proposed the free-ride coding to realize an implicit GC-LDPC code, and parallel encoding and efficient decoding algorithms were proposed based on the unique structure of the GC-LDPC code.

Since the existing GC-LPDC code cannot achieve girth 8, in this paper, we consider the construction of a GC-LDPC code based on the FLRM matrix. The main contributions of this paper are as follows:

1. We find a new lower bound of CPM size $P$ to achieve girth 8 for the GCD-based FLRM matrix and propose an algorithm that can output the minimum $P$ and the corresponding FLRM matrix for the given $J, L$.
2. The two new methods for constructing the GC-LDPC code is proposed based on the FLRM matrix with a new lower bound.
3. We find that the performance of GC-LPDC is more sensitive to the number of six length cycles than the girth.

The finding of this study is particularly meaningful for the code designer for they will have more freedom in choosing $P$, code length, or code rate. In addition, the simulation results show that the code constructed with proposed method has improved performances than the code constructed with existing methods.

The sections of this paper are organized as follows. In Section 2, we discuss cycles in the FLRM matrix and the new lower bound of CPM's size $P$. An algorithm is proposed to find the smallest $P$ and the corresponding FLRM matrix for a given $(J, L)$. In Section 3, two code construction methods are proposed for the GC-LDPC code based on the GCD-based FLRM matrix. In Section 4, we show the simulation results, and we conclude our paper in Section 5.

## 2. New Lower Bound of CPM Size

The parity check matrix of the QC-LPDC is generated through a CPM of size $P$ and an exponent matrix of size $J \times L$. The elements of the exponent matrix $\mathbf{E}$ is the product of the row and column coefficients:

$$\mathbf{E} = \begin{bmatrix} a_0 b_0 & a_0 b_1 & \cdots & a_0 b_{L-1} \\ a_1 b_0 & a_1 b_1 & \cdots & a_1 b_{L-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{J-1} b_0 & a_{J-1} b_1 & \cdots & a_{J-1} b_{L-1} \end{bmatrix}, \quad (1)$$

where $a_i$, and $b_i$ are integers with $0 \leq a_0 < a_1 < \cdots < a_{J-1}$ and $0 \leq b_0 < b_1 < \cdots < b_{L-1}$. The row coefficient vector is $\boldsymbol{a} = (a_0, a_1, \cdots, a_{J-1})$, and the column coefficients vector is $\boldsymbol{b} = (b_0, b_1, \cdots, b_{L-1})$. The column coefficients vector of the FLMR matrix is $\mathbf{b} = (0, 1, 2, \cdots, L-1)$. The maximum girth of FLRM codes is eight [17]. The GCD method [16] is an efficient framework for constructing FLRM codes, attaining girth eight.

The girth of GC-LDPC code is essentially the length of the shortest cycle in $\mathbf{E}$. A cycle $\mathbf{W} = \{W_0, W_1, ..., W_{l-1}\}$ in matrix $\mathbf{E}$ is a sequence of the elements of $\mathbf{E}$ such that [32]

$$\sum_{i=0}^{l/2-1} (W_{2i} - W_{2i+1}) = 0 (\mathrm{mod}\, P), \tag{2}$$

where $l \in \{4, 6, 8, \cdots\}$ is the length of cycle.

A length of four cycles in $\mathbf{E}$ forms a $2 \times 2$ sub-matrix of $\mathbf{E}$. Let $\mathbf{E}_4^{'}$ be a $2 \times 2$ sub-matrix of $\mathbf{E}$, which consists of two distinct rows and two distinct columns of $\mathbf{E}$:

$$\mathbf{E}_4^{'} = \begin{bmatrix} a_i b_x & a_i b_y \\ a_j b_x & a_j b_y \end{bmatrix}, \tag{3}$$

where $0 \leq i < j < J$ and $0 \leq x < y < L$. $\mathbf{E}_4'$ can form a length of four cycles if its elements satisfy (2). This is equivalent to $\det(\mathbf{E}_4') = 0 (\mathrm{mod}\, P)$. Hence, the condition for $\mathbf{E}$ with no four-length cycle is that all $2 \times 2$ submatrices of $\mathbf{E}$ are non-singular. This condition can be equivalently expressed as follows:

$$(a_j - a_i)(b_y - b_x) \neq 0 (\mathrm{mod}\, P), \tag{4}$$

for all $0 \leq i < j < J$, and $0 \leq x < y < L$.

A length of six cycles lies in a $3 \times 3$ sub-matrix of $\mathbf{E}$. Consider the following:

$$\mathbf{E}_6^{'} = \begin{bmatrix} a_i b_x & a_i b_y & a_i b_z \\ a_j b_x & a_j b_y & a_j b_z \\ a_k b_x & a_k b_y & a_k b_z \end{bmatrix}, \tag{5}$$

where $0 \leq i < j < k < J, 0 \leq x < y < z < L$. A path of length 6 in $\mathbf{E}_6'$ satisfying (2) will define a cycle of length 6. Similarly to (4), if we define the following:

$$\begin{aligned} S_1 &= (a_k - a_i)(b_y - b_x) + (a_k - a_j)(b_z - b_y) \\ S_2 &= (a_k - a_i)(b_y - b_x) + (a_j - a_i)(b_z - b_y) \\ S_3 &= (a_k - a_i)(b_z - b_y) + (a_k - a_j)(b_y - b_x) \\ S_4 &= (a_k - a_i)(b_z - b_y) + (a_j - a_i)(b_y - b_x) \\ S_5 &= (a_k - a_j)(b_y - b_z) + (a_j - a_i)(b_y - b_x) \\ S_6 &= (a_k - a_j)(b_y - b_x) + (a_j - a_i)(b_y - b_z), \end{aligned} \tag{6}$$

then any of $S_1, S_2, \cdots, S_6$ being zero (module $P$) indicates the existence of $P$ cycles of length 6. The possible paths of these cycles are illustrated in Figure 1. These cycles can be classified into two types: the 'L' type paths correspond to $S_1, S_2, S_3, S_4$, and the 'X' type paths correspond to $S_5, S_6$.

Since $a_i < a_j < a_k$ and $b_x < b_y < b_z$, all $S_1, S_2, S_3, S_4$ are positive and are upper bounded. A sufficiently large $P$ can guarantee $S_i (\mathrm{mod} P) = S_i$ or $S_i \neq 0 (\mathrm{mod} P)$ and, thus, avoid the type 'L' length 6 cycle.

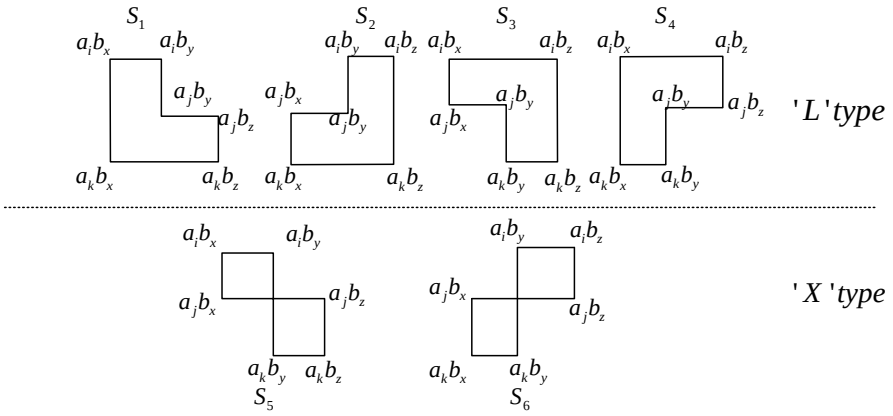



**Figure 1.** Possible paths of length 6 cycles corresponding to (6).

Even with infinite $P$, 'X' type cycles may exist for $S_5$, or $S_6$ may equal to zero. Let $n_a = a_j - a_i$, $m_a = a_k - a_j$, $n_b = b_y - b_x$, and $m_b = b_z - b_y$. $S_5$ and $S_6$ can be rewritten as follows.

$$S_5 = -m_a \cdot m_b + n_a \cdot n_b, \tag{7a}$$
$$S_6 = m_a \cdot n_b - n_a \cdot m_b. \tag{7b}$$

For the 'X' type length 6 cycles, the condition $S_5 = 0$ or $S_6 = 0$ can be equivalently expressed as the difference ratio as follows.

$$\frac{n_a}{m_a} = \frac{m_b}{n_b}, \quad \text{if} \quad S_5 = 0, \tag{8a}$$
$$\frac{n_a}{m_a} = \frac{n_b}{m_b}, \quad \text{if} \quad S_6 = 0. \tag{8b}$$

The condition for the girth 8 FLRM matrix can be derived from the analysis above, and the conclusion is summarized as (Lemmas 1 and 2 of [16])

$$(n_a + m_a) / \gcd(n_a, m_a) \geq L, \tag{9a}$$
$$P \geq (a_{J-1} - a_0)(L-1) + 1. \tag{9b}$$

Equation (9a) is the necessary and sufficient condition for avoiding the type 'X' cycle of length 6. However, (9b) is only a sufficient condition to avoid type 'L' cycle of length 6.

The type 'L' cycle of length 6 exists if and only if any of $\{S_1, S_2, S_3, S_4\}$ equals $0 (\mathrm{mod} P)$. For the sake of simplicity, we use $P^*_{\min}$ to denote the lower bound in (9b).

$$P^*_{\min} = (a_{J-1} - a_0)(L-1) + 1. \tag{10}$$

However, (9b) is only a sufficient condition for avoiding the length 6 cycle. It is possible that, for some $P < P^*_{\min}$, all $P < P^*_{\min}$ and all $S_i \neq 0 (\mathrm{mod} P)$, $i \in \{1, 2, \cdots, 6\}$ for all $3 \times 3$ submatrices. Thus, the real lower bound is given by the following.

$$P^\dagger_{\min} = \underset{P}{\mathrm{argmin}} \left\{ S_i = 0 (\mathrm{mod} P),\ \forall i, \forall \mathbf{E}_6^{'} \right\}. \tag{11}$$

With the new lower bound $P^\dagger_{\min}$, we can extend condition (9b) to the sufficient and necessary condition, as stated in the following theorem.

**Theorem 1.** *The Tanner graph corresponds to the FLRM matrix* **E** *and has no 'L' type cycle of length 6 if and only if the following is the case:*

$$P \geq P^{*}_{\min} = (a_{J-1} - a_0)(L-1) + 1, \tag{12a}$$

or

$$P^{\dagger}_{\min} \leq P < P^{*}_{\min} \text{ and } S_i(\text{mod})P \neq 0, \forall i, \mathbf{E}^{'}_{6} \tag{12b}$$

*for all triples* $(a_i, a_j, a_k)$, $0 \leq i < j < k < J$.

We propose the Algorithm 1 shown in the next page to find the minimum. The input of Algorithm 1 is the the size of FLRM matrix. The algorithm starts from constructing an FLRM matrix **E** using the method stated in [16] and **E** satisfies (9a). This matrix has no length 6 cycle with $S_5 = 0$ or $S_6 = 0$. Then, we calculate path metrics defined in (6) for all $3 \times 3$ submatrices and record the results in a accumulated vector **Sum** where the $e^{th}$ element, $\text{Sum}(e)$, is the number of length 6 paths for which $S_i = e$ for some $i \in \{1, 2, \cdots, 6\}$. An example for **Sum** is illustrated in Figure 2 with $J = 5, L = 20$.

**Algorithm 1** Construct **E** with the minimum $P = P^{\dagger}_{\min}$

**Require:** $J, L$
**Ensure:** **E** and $P^{\dagger}_{\min}$

construct $\mathbf{E}^{'}$ satisfing (9a) as [18];
Initialize **Sum** $= \mathbf{0}$; $P^{\text{init}} = a_{J-1} \times (L-1)$
**for** $(\forall i, j, k, x, y, z;\ 0 \leq i < j < k < J, 0 \leq x < y < z < L)$ **do**
  For each $3 \times 3$ submatrix, calculate path metric $S_i, i = 1, 2, \cdots, 6$ with (6);
  **for** $(i = 0 : 6)$ **do**
    $\text{Sum}(S_i) + +$;
  **end for**
**end for**
**for** $(e = \max\{a_{J-1}, L-1\} : P^{\text{init}})$ **do**
  Candidate = True
  **for** $(r = 1 : P^{\text{init}}/e)$ **do**
    **if** $\text{Sum}(e \times r) \neq 0$ **then**
      Candidate = False; break;
    **end if**
  **end for**
  **if** Candidate = True **then**
    $\mathbf{E}_{\text{temp}} = \mathbf{E}^{'}\ (\text{mod})e$
    **for** $(\forall i, j, x, y,\ 0 \leq i < j < J,\ 0 \leq x < y < L)$ **do**
      For each $2 \times 2$ submatrix, calculate determinant $D = \det\left(\mathbf{E}^{(\mathbf{i},\mathbf{j},\mathbf{x},\mathbf{y})}_{\mathbf{temp},\mathbf{4}}\right)$
      **if** $D = 0(\text{mod})e$ **then**
        Candidate = False; break;
      **end if**
    **end for**
  **end if**
  **if** Candidate = True **then**
    $P^{\dagger}_{\min} = e$; break;
  **end if**
**end for**
$\mathbf{E} = \mathbf{E}_{\text{temp}}$

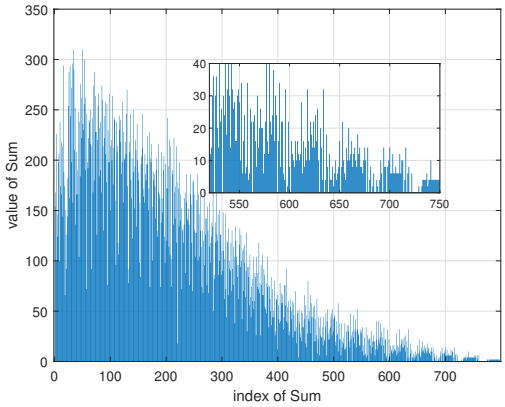



**Figure 2.** An example for the value of **Sum** for $J = 5, L = 20$.

If the Tanner graph has no length 6 cycle under CPM size $P$, then there will be no path metric such that $S_i = 0(\mathrm{mod}P)$, or $S_i = rP$, for $i \in \{1, 2, \cdots, 6\}$ and $r \in \{1, 2, \cdots\}$. This is reflected in **Sum** as $\mathrm{Sum}(P) = \mathrm{Sum}(2P) = \mathrm{Sum}(3P) = \cdots = 0$. In Algorithm 1, the lines 11∼15 check this condition. Note that if the FLRM matrix contains no length 6 cycle, it does not mean that it contains no length 4 cycle. Thus, in lines 17∼23, we check (4) for all $2 \times 2$ submatrices. The algorithm searches the indices of **Sum** in range $\max\{a_{J-1}, L-1\} \leq e < P^*_{\min}$. The new lower bound $P^{\dagger}_{\min}$ is the minimum index of the zero elements of **Sum**.

$$P^{\dagger}_{\min} = \underset{\substack{\max\{a_{J-1}, L-1\} \leq e < P^*_{\min}, \\ 1 \leq r < P^*_{\min}/e}}{\operatorname{argmin}} \{\mathrm{Sum}(e \times r) = 0\}. \tag{13}$$

It is possible that the Tanner graph corresponding to the FLRM matrix has only 'X' type cycles or only 'L' type cycles. Combining Theorems 1 and 2, the properties concerning these cases are summarized as the following theorem:

**Theorem 2.** *The cycles in the Tanner graph corresponding to FLRM matrix* **E** *have the following properties:*

1. *If $P \geq P^*_{\min}$ and there exists triple $(a_i, a_j, a_k)$, $0 \leq i < j < k < J$ such that $(n_a + m_a)/\gcd(n_a, m_a) < L$, then all cycles of length 6 are type 'X' cycle.*
2. *If $P < P^*_{\min}$ and $(n_a + m_a)/\gcd((n_a, m_a) \geq L$ for all triples $(a_i, a_j, a_k)$, $0 \leq i < j < k < J$, then the number of 6 length cycles equals to $Sum(P) \times P$.*

For index $e$, $\mathrm{Sum}(e)$ is the number of length 6 cycles under CPM size $P = e$. It can be seen from Figure 2 that, as the CPM size increase, the number of length 6 cycles decrease rapidly. Latter in Section IV, we can see that number of length 6 cycles has critical influences on the performance.

Some results on the new lower bound $P^{\dagger}_{\min}$ are shown in Tables 1 and 2, where $P_N$ denotes the number of valid $P$ below the original lower bound $P^*_{\min}$ in [16]. From these Tables, it can be seen that the new lower bound is significantly smaller than the original one. In cases $J = 5$ and $L = 13$, the new lower bound $P^{\dagger}_{\min}$ is about 40% lower than $P^*_{\min}$. In the case of $J = 6$, the reduction is about 45% for most values of $L$. The reduction ratio becomes even larger when $J$ and $L$ increase.

**Table 1.** (5, $L$) girth-eight FLRM code with minimum CPM size.

| $L$ | $a_0a_1a_2a_3a_4$ | $P^*_{\min}$ Lower Bound in [18] | $P^{\dagger}_{\min}$ New Lower Bound | $P_N$ Number of $P$ Below $P^*_{\min}$ |
|---|---|---|---|---|
| 5 | 0, 1, 5, 11, 12 | 49 | 49 | 0 |
| 6 | 0, 1, 8, 9, 14 | 71 | 63 | 1 |
| 7 | 0, 2, 7, 11, 16 | 97 | 67 | 5 |
| 8 | 0, 1, 8, 11, 18 | 127 | 111 | 4 |
| 9 | 0, 2, 9, 13, 20 | 161 | 103 | 9 |
| 10 | 0, 1, 10, 11, 23 | 208 | 143 | 8 |
| 11 | 0, 1, 11, 18, 23 | 231 | 165 | 8 |
| 12 | 0, 1, 14, 15, 26 | 287 | 221 | 10 |
| 13 | 0, 2, 13, 17, 28 | 337 | 199 | 21 |
| 14 | 0, 1, 14, 17, 30 | 391 | 285 | 19 |
| 15 | 0, 1, 17, 18, 32 | 449 | 368 | 16 |
| 16 | 0, 1, 16, 23, 33 | 496 | 407 | 15 |
| 17 | 0, 1, 17, 22, 35 | 561 | 357 | 21 |
| 18 | 0, 1, 18, 23, 37 | 630 | 529 | 21 |
| 19 | 0, 1, 19, 32, 39 | 703 | 595 | 21 |
| 20 | 0, 1, 20, 23, 42 | 799 | 525 | 41 |

**Table 2.** (6, $L$) girth-eight FLRM code with minimum CPM size.

| $L$ | $a_0a_1a_2a_3a_4a_5$ | $P^*_{\min}$ Lower Bound in [18] | $P^{\dagger}_{\min}$ New Lower Bound | $P_N$ Number of $P$ Below $P^*_{\min}$ |
|---|---|---|---|---|
| 6 | 0, 2, 7, 11, 16, 18 | 91 | 63 | 7 |
| 7 | 0, 2, 7, 11, 16, 18 | 108 | 67 | 8 |
| 8 | 0, 1, 8, 11, 18, 19 | 134 | 134 | 0 |
| 9 | 0, 2, 9, 13, 20, 22 | 176 | 103 | 13 |
| 10 | 0, 1, 10, 11, 23, 24 | 217 | 217 | 0 |
| 11 | 0, 1, 11, 14, 24, 25 | 251 | 251 | 0 |
| 12 | 0, 2, 13, 17, 28, 30 | 331 | 199 | 24 |
| 13 | 0, 2, 13, 17, 28, 30 | 361 | 199 | 26 |
| 14 | 0, 1, 14, 17, 30, 31 | 404 | 315 | 1 |
| 15 | 0, 2, 15, 19, 32, 34 | 477 | 259 | 34 |
| 16 | 0, 5, 16, 23, 34, 39 | 586 | 357 | 49 |
| 17 | 0, 1, 17, 22, 38, 39 | 625 | 399 | 4 |
| 18 | 0, 1, 18, 23, 40, 41 | 698 | 501 | 6 |

**Table 2.** *Cont.*

| $L$ | $a_0a_1a_2a_3a_4$ | $P^*_{\min}$ Lower Bound in [18] | $P^{\dagger}_{\min}$ New Lower Bound | $P_N$ Number of $P$ Below $P^*_{\min}$ |
|---|---|---|---|---|
| 19 | 0, 2, 19, 23, 40, 42 | 757 | 403 | 53 |
| 20 | 0, 1, 20, 23, 42, 43 | 818 | 693 | 1 |

High-rate LDPC codes require sufficiently large $L$. The lower bound $P^*_{\min}$ in [16,18] is roughly in line with $L^2 \times \lfloor (J-1)/2 \rfloor$ since $a_{2i} > L \times i$ and $a_{J-1}$ are very close to $a_{J-2}$ for even $J$. The code length $N = P \times L \approx L^3 \times \lfloor (J-1)/2 \rfloor$ grows with $L^3$. With the new lower bound $P^{\dagger}_{\min}$, the limitation on $N$ can be greatly reduced, which is significant for the construction of high-rate girth-8 QC-LDPC with small FLRM matrix sizes.

Note that when $J \leq 4$, we have $P^*_{\min} = P^{\dagger}_{\min}$ for all $L$. This is because $a_{J-1} - a_0$ is small; hence, the summations of the paths are enough to iterate over all available values under $(a_{J-1} - a_0)(L-1) + 1$.

## 3. Construct Girth-8 GC-LDPC with New Lower Bound

In general, the parity check matrix of the GC-LPDC has the following structure:

$$\mathbf{H}_{\text{gc}} = \left[\begin{array}{cccc} \mathbf{H}_L^0 & & & \\ & \mathbf{H}_L^1 & & \\ & & \ddots & \\ & & & \mathbf{H}_L^{t-1} \\ \hline & & \mathbf{H}_G & \end{array}\right], \tag{14}$$

where $\mathbf{H}_L^i, i = 0, 1, \cdots, t-1$ is the local parity check matrix of $i$th local codeword, and $\mathbf{H_G}$ is the global parity check that connects all local codes together. The GC-LDPC codeword can be decoded globally with $\mathbf{H}_{\text{gc}}$, or it firstly decodes each local codeword with $\mathbf{H}_L^i$ and then decodes the entire codeword with $\mathbf{H}_G$. Figure 3 illustrates the Tanner graph of GC-LDPC code with a general structure. The circle/square symbol indicates the variable/check nodes, respectively. From Figure 3, it can be seen more intuitively that local codes update and exchange information through the global check node.

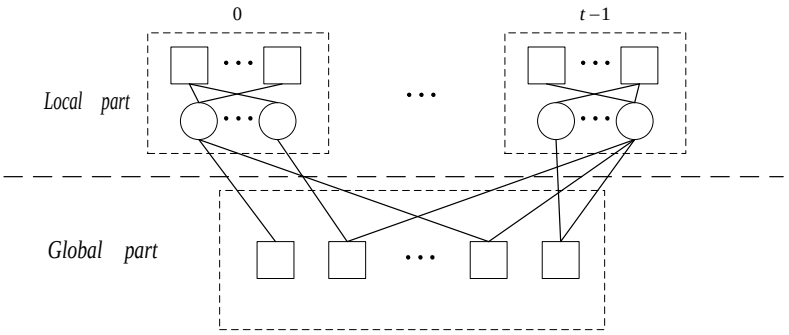



**Figure 3.** GC-LDPC's Tanner graph.

### *3.1. Review of the Construction of GC-LDPC Code*

The GC-LDPC code is generally derived from the well-designed $J \times L$ regular matrix with girth six. The designed matrix is then deformed into the GC-LDPC code using displacement, masking, and dispersion.

In [24], the GC-LDPC code is constructed on GF($q$), and the matrices are provided by the following:

$$\mathbf{B_1} = \begin{bmatrix} \alpha^0 - 1 & \alpha - 1 & \cdots & \alpha^{q-3} - 1 & \alpha^{q-2} - 1 \\ \alpha^{q-2} - 1 & \alpha^0 - 1 & \cdots & \alpha^{q-4} - 1 & \alpha^{q-3} - 1 \\ \vdots & \vdots & \ddots & \vdots & \\ \alpha - 1 & \alpha^2 - 1 & \cdots & \alpha^{q-2} - 1 & \alpha^0 - 1 \end{bmatrix},$$
$$\mathbf{B_2} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \beta & \cdots & \beta^{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \beta^{p-1} & \cdots & (\beta^{p-1})^{p-1} \end{bmatrix}, \tag{15}$$

where $\alpha$ is a primitive element over GF$(q)$, $\beta = \alpha^e$, and$q - 1 = pe$. Any $2 \times 2$ submatrix of $\mathbf{B_1}$ and $\mathbf{B_2}$ is non-singular, which guarantees that $\mathbf{B_1}$ and $\mathbf{B_2}$ are free of length 4 cycles.

A Reed–Solomon code has been used in [27] to construct the GC-LDPC code on GF($2^s$). Let $\alpha$ be the primitive element of GF($2^s$), $2^s - 1 = c \times n$, $c$ is the primitive factor of $2^s - 1$, and $\gamma = \alpha^c$. The elements of vector $\mathbf{s} = (1, \gamma, \gamma^2, ..., \gamma^{n-1})$ are the cyclic elements over GF$(2^s)$ with order $n$. Since the $n$'s smallest prime factor is $p_s > d$, $\mathbf{B}_{RS}(d, n)$ given by the following:

$$\mathbf{B}_{RS}(d, n) = \begin{bmatrix} 1 & \gamma & \cdots & \gamma^{n-1} \\ 1 & \gamma^2 & \cdots & (\gamma^2)^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \gamma^d & \cdots & (\gamma^d)^{n-1} \end{bmatrix} \tag{16}$$

and it satisfies the $2 \times 2$ constraint (4). The construction methods above are based on a matrix satisfying the $2 \times 2$ constraint, and then we obtain all local codes and global parts using segmentation, masking, and re-organization.

The Reed-Solomon-Like construction is proposed in [28] which makes the design of GC-LDPC code more flexible. The local and global codes can be individually designed as follows:

$$\mathbf{RS}(a, b, d) = \begin{bmatrix} \gamma^0 & \gamma^a & \gamma^{2a} & \cdots & \gamma^{a(b-1)} \\ \gamma^0 & \gamma^{2a} & \gamma^{4a} & \cdots & \gamma^{2a(b-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma^0 & \gamma^{ad} & \gamma^{2ad} & \cdots & \gamma^{da(b-1)} \end{bmatrix}, \tag{17}$$

where $a$ denotes the scaling factor, and $b$ and $d$ are the numbers of columns and rows of $\mathbf{RS}(a, b, d)$.

To avoid the length 4 cycles between global and local parts, the number of rows of local codes and global check should be constrained to $d_{\mathrm{L}} < \left\lfloor \frac{b_{\mathrm{L}}}{a_{\mathrm{L}}} \right\rfloor$ and $d_{\mathrm{G}} < \min\{a_{\mathrm{L}}\}$, respectively, where the subscript L and G indicate that the parameter belongs to local codes or global part. With the same set of parameters, the codes derived from above methods have a similar performance. For this reason, only the Reed–Solomon-Like method is used as the comparison scheme since this construction method is more flexible.

### *3.2. GCD-Based GC-LDPC Code*

The GC-LDPC can be obtained through the matrices designed above with some operations without changing the property of the cycle. In the following, we will apply the GCD method to construct girth eight GC-LDPC codes.

**Construction 1.** *We generate the girth 8 FLRM matrix through Algorithm 1 for the code with column weight L. The number of rows, $d_{\mathrm{L}}^i, i \in [0, t)$, and $d_{\mathrm{G}}$ is set in accordance to the code rate's requirement. The first $d_{\mathrm{LMax}}$ rows of the matrix are split into $\mathbf{E}_{\mathrm{L}}^i$, where $d_{\mathrm{LMax}} = \max(d_{\mathrm{L}}^i) \leq L - d_{\mathrm{G}}$. Finally, $\mathbf{E}_L^i$ is rearranged into diagonal forms, as shown in Figure 4.*

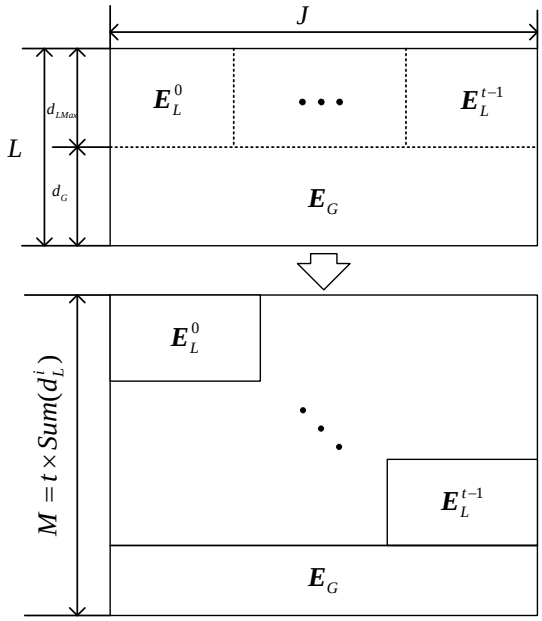



**Figure 4.** Illustration of construction 1.

With Construction 1, the code rate is given by $r_1 \simeq 1 - \sum(d_{\mathrm{L}}^i)/J$. Increasing $r_1$ requires an increase in $J$. Since the lower bound of $P$ increases rapidly with the increase in $J$, the code length $N = P_{\min} \times J$ will soon become unacceptable. To address this problem, we provide Construction 2 as follows.

**Construction 2.** *In Construction 2, the FLRM matrix generated via Algorithm 1 is split into two parts, namely* $\mathbf{E}_{\mathrm{G}}$ *and* $\mathbf{E}_{\mathrm{L}}^{\mathrm{Max}}$*, with size the as* $1 \times J$ *and* $(L-1) \times J$*, respectively. Then, the two sub-matrices are copied t times and interleaved along columns, as shown in Figure 5II. According to the requirements of each sub-matrix, each interleaved copy is cut, and then global and local sub-matrices are obtained as illustrated in Figure 5II. At last, these matrices are rearranged into the form as illustrated in Figure 5III.*

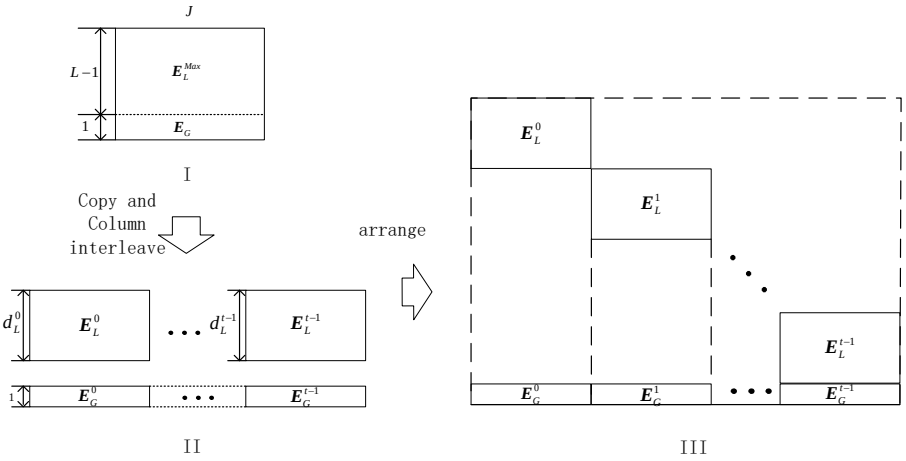



**Figure 5.** Illustration of Construction 2.

The rate of this code is $r_2 = \sum(d_{\mathrm{L}}^i)/(t \times J)$. In case $d_{\mathrm{L}}^i = L - 1$, the code rate is $r_2 = [t(L-1)+1]/(t \times J)$. The global part is limited to having a $d_{\mathrm{G}} = 1$ row to ensure that no new cycles occur in the global part.

If necessary, the column cyclic mask can be used in the construction to reduce the column's weight. Let $\mathbf{m} = (m_0, ..., m_{J-1})^{\mathrm{T}}$ be the binary cyclic mask vector. The elements of cyclic mask matrix $\mathbf{M}$ are given by $M_{i,j} = m_{(i+j+c)(\mathrm{mod}\, J)}, i \in [0, J), j \in [0, L)$ where $c$ is the randomly selected initial offset. For example, if all local codes have the same size of $5 \times 8$, the weight of $\mathbf{m}$ is $M_n = 1$; then, with the randomly generated $\mathbf{m} = (0, 0, 1, 0, 0)^{\mathrm{T}}$ and $c = 0$, the mask matrix is constructed as follows:

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{18}$$

where the entries with value '1' indicate the positions to be masked in local codes $\mathbf{H}_{\mathrm{L}}^{i}$ of the same positions.

The focus of the code's construction is to achieve girth 8. For the decoding of LDPC codes, the short cycles will lead to inappropriate messages passing among nodes or error propagation. Moreover, these short cycles can form the local structure of trapping set and stopping set. All these issues will cause performance degradation. Therefore, the LDPC code should be carefully designed to avoid short cycles.

## 4. Simulation Results

In this section, we use simulations to verify the proposed GC-LDPC code. All codes in this section were simulated in the AWGN channel with BPSK modulations. The min-sum algorithm with a scaling factor 0.55 is used in decoding, and the maximum iteration number of local/global decoding is 50.

We also use the extrinsic information transfer (EXIT) chart to compare the conventional GC-LDPC code and the proposed codes. The mutual information is given by the following:

$$I = 1 - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(u-\sigma^2/2)^2}{2\sigma^2}} \log\left(1 + \mathrm{e}^{-u}\right) \mathrm{d}u, \tag{19}$$

where $\sigma^2$ is the variance of the LLR value.

### *4.1. Comparison of Grith 8 and Grith 6 GC-LDPC Codes*

In the following, the proposed code is referred to as code1. We use notation code 2, code3, and code4 to denote the baseline codes used for the comparison. The details of these codes are as follows.

code1: The construction of code1 begins with Algorithm 1, which outputs the matrix $\mathbf{E}$ of size $J \times L = 4 \times 20$ and the minimum CPM size of $P = P_{\min} = 400$. Then, we perform Construction 2 with the number of local matrices as $t = 2$ and the size of local matrices as $3 \times 20$. The girth of code1 is 8, the size of the exponential matrix is $7 \times 40$, the code length is $N_1 = 16{,}000$, and the code rate is $r_1 = 0.829$. The column weight of local matrices is $w_{\mathrm{L}} = 3$, and the column weight of the entire matrix is $w_{\mathrm{w}} = 4$.

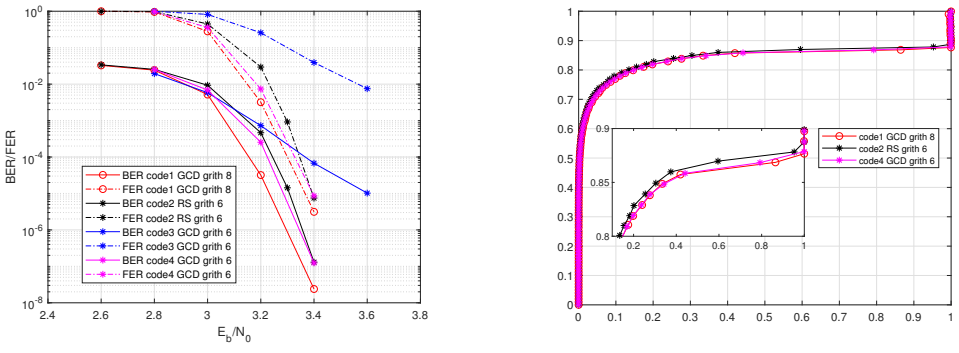
code2: This code is constructed with a Reed–Solomon-Like scheme [28]. $\mathbf{E}$ comprises three RS$(3, 131, 3)$ local matrices and one RS$(1, 393, 1)$ global matrix. Then, the code was slightly changed to match the parameters with code1. Finally, the size of exponential matrix is $7 \times 40$, the CPM size as 393, the code length is $N_2 = 15{,}720$, and the code rate is $r_2 = 0.834$.

code3: This code is constructed to observe the relationship between the performance of GC-LDPC code and the number of length 6 cycles. We construct the grith 6 FLRM matrix $\mathbf{B}$ with the vector of row indices $\boldsymbol{a} = (0, 1, 2, 3)$ and the vector of column indices $\boldsymbol{b} = (0, 1, \ldots, 19)$. The CPM size is $P = 400$, the size of exponential matrix is $7 \times 40$, the code rate is $r_3 = 0.825$, and the code length is $N_3 = 16{,}000$. According to Theorem 2 , this code has no length 4 cycle and no 'L' type length 6 cycle. All $3 \times 3$ sub-matrices involve a large

number of length 6 cycles of type 'X'. According to the statistics data obtained in simulation, there is a total of $1.08 \times 10^7$ length 6 cycles in code3.

code4: The matrix **E** of code4 is constructed with the GCD method as in code1. The CPM size is $P = 393$, the code rate is $r_4 = 0.825$, and the code length is $N_4 = 15720$. This code only has type 'L' length 6 cycles. It was observed from the results of Sum in Algorithm 1 that the code has $3.2 \times 10^3$ length 6 cycles.

Figure 6a compares simulated bit error rate (BER) and frame error rate (FER) performance of code1∼code4. From this figure, we can see that the proposed code1 has the best performance, code4 is in the second position, and code3 is the worst. Note that code1, code3, and code4 are all constructed with the GCD method. The difference mainly lies in the number of length 6 cycles, which is 0, $1.08 \times 10^7$, and $3.2 \times 10^3$ for code1, code3, and code4, respectively. The comparison in Figure 6a indicates that the performance of the GC-LDPC code is seriously affected by the number of length 6 cycles. The reason that code3 has the worst performance is that this code contains more length 6 cycles than other codes. It is worth noting that although the girth of code1 is 8, while the girth of code2 and code3 is 6, the performance of code1 is only slightly improved compared to the other two. The EXIT charts of codes 1, 2, 4 also indicate that the GCD-based GC-LDPC and Reed–Solomn-Like GC-LDPC have nearly the same performance. An interest observation from Figure 6 is that the dominant factor affecting the error rate's performance is the number of length 6 cycles rather than the girth. Therefore, the designer does not have to guarantee girth 8. The more important concern should be the number of length 6 cycles. In other words, it may not be necessary to select a $P$ corresponds to the zero element of **Sum** in Algorithm 1. A $P$ value for which $\sum_r(\text{Sum}(P \times r))$ is relatively small may be sufficient.




(**a**) (**b**)

**Figure 6.** Comparison of the GC-LDPC codes with different numbers of 6-cycles. (**a**) BER/FER performances. (**b**) EXIT chart.

### *4.2. GC-LDPC Code with Minimum Size CPM*

In this section, we consider some GC-LDPC codes for which GCD-based FLRM matrices are generated with $L$ and $\boldsymbol{a}$ selected from Tables 1 and 2. The code is constructed with Construction 2 mentioned in the last section, and the column cyclic mask is used to mitigate the column's weight.

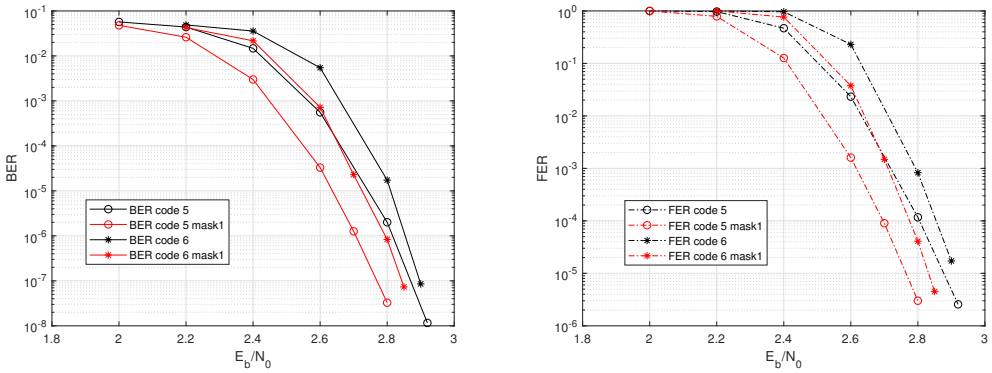
case 1 $J = 5$

In this case, we consider two codes, code5 and code6, with parameters listed in Table 3.

**Table 3.** The parameters of code 5 and 6.

| | $J \times L$ | $P$ | $N$ | $r$ | $t$ |
|---|---|---|---|---|---|
| *code*5 | $5 \times 14$ | 285 | 11970 | 0.692 | 3 |
| *code*6 | $5 \times 16$ | 407 | 19536 | 0.731 | 3 |

In the AWGN channel, the best range of column weight is $w_c = [3, 4]$. We use the column cyclic mask of $M_n = 1$ to reduce $w_c$ to 3. The masked codes are denoted as code5 mask1 and code6 mask1. The code rate of code5 mask1 is $r = 0.69$, and it is $r = 0.729$ for code6 mask1. Figure 7 indicates that both the original codes and the masked codes have shown good BER/FER performances without an error floor. Note that code 5 has short lengths but improved performances than code 6. This is because code 5 has a smaller code rate. With a code length as large as code 5, the performance is dominated by the code rate.

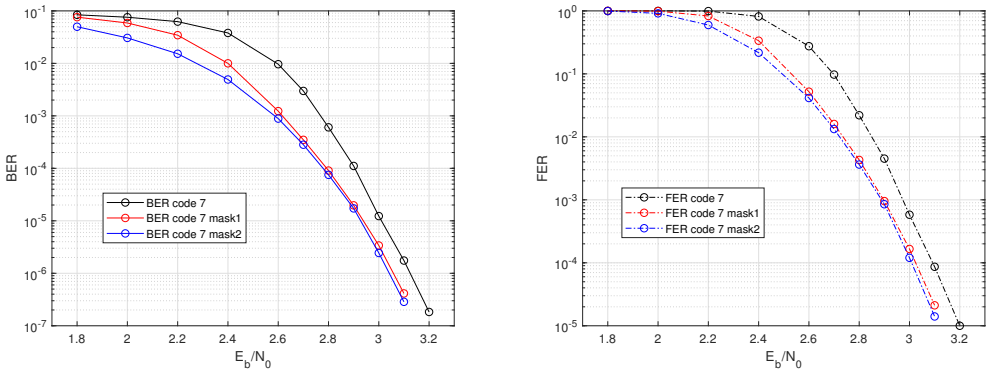



(**a**) (**b**)

**Figure 7.** Performance of code5, code6, code5 mask1, and code6 mask1 for $J = 5$. (**a**) BER. (**b**) FER.

case 2 $J = 6$

In this case, the $6 \times 13$ FLRM matrix is used to construct code7 with Construction 2. The CPM size is $P = 199$, the number of the local codes is $t = 3$, the code length is $N = 7761$, and the code rate is $r = 0.591$. The column cyclic mask of weight $M_n = 1$ or 2 is applied, and the masked codes are denoted as code7 mask1 and code7 mask2. The code rate of masked codes are both $r = 0.59$. The simulated error rate performance is shown in Figure 8. This results indicates that the GC-LPDC codes with different column weight requirements can be obtained using the method of GCD-based FLRM matrices and column cyclic masks, and these codes have good performances without ab error floor.




(**a**) (**b**)

**Figure 8.** Performance of code7, code7 mask1, and code7-mask2 for $J = 6$. (**a**) BER. (**b**) FER.

## 5. Conclusions

In this paper, we show that the lower bound of CPM's size can be even lower than what we known from the previous literature. An algorithm is proposed to find the minimum size of CPM for the GCD-based FLRM matrix. Based on this algorithm, two construction methods were proposed to construct girth 8 GC LPDC codes. In addition, we find that the dominant factor that affects the performance is the number of length 6 cycles rather than the girth. With the proposed algorithm, the CPM size can be 45% less than that given by a

sufficient condition of girth 8. Compared with the conventional GC-LDPC construction, the codes constructed with the proposed method have better performances and are more flexible in code length and code rate design.

**Author Contributions:** Conceptualization, K.Z.; Methodology, K.Z.; software, K.Z.; Formal analysis, K.Z.; Investigation, K.Z.; Writing—original draft preparation, K.Z.; Writing—review & editing, H.Y. Supervision, H.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Panteleev, P.; Kalachev, G. Quantum LDPC codes with almost linear minimum distance. *IEEE Trans. Inf. Theory* **2021**, *68*, 213–229.
2. Duffy, K.R.; Li, J.; Mdard, M. Capacity-achieving guessing random additive noise decoding. *IEEE Trans. Inf. Theory* **2019**, *65*, 4023–4040.
3. Niu, K.; Kai, C. CRC-aided decoding of polar codes. *IEEE Commun. Lett.* **2012**, *16*, 1668–1671.
4. Takayuki, N.; Motohiko, I. LDPC Codes for Communication Systems: Coding Theoretic Perspective. *IEICE Trans. Commun.* **2022**, *E105-B*, 894–905.
5. Shao, S.; Hailes, P.; Wang, T,Y.; Wu, J.Y.; Maunder, R.G.; Al-Hashimi, B.M.; Hanzo, L. Survey of turbo, LDPC and polar decoder ASIC implementations. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2309–2333.
6. Kun, Z.; Zhanji, W. Comprehensive Study on CC-LDPC, BC-LDPC and Polar Code. In Proceedings of the 2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), Seoul, Korea, 6–9 April 2020; pp. 1–6.
7. Arora, K.; Singh, J.; Randhawa, Y.S. A survey on channel coding techniques for 5G wireless networks. *Telecommun. Syst.* **2020**, *73*, 637–663.
8. Gao, C.; Liu, S.; Jiang, D.; Chen, L. Constructing LDPC Codes with Any Desired Girth. *Sensors* **2021**, *21*, 2012.
9. Meng, J.; Zhao, D.; Zhang, L. Design and Analysis of Non-Binary LDPC-CPM System for Hybrid Check Matrix Construction Algorithm of WSN. *Sensors* **2018**, *18*, 2418.
10. Vladimir, L.P.; Milos, M.M.; Dragomir, M.E.M.; Dragomir, M.E.M.; Andreja, R. Flexible High Throughput QC-LDPC Decoder With Perfect Pipeline Conflicts Resolution and Efficient Hardware Utilization. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2020**, *67*, 5454–5467.
11. Daniel, J.C.; Lara D.; Thomas E.F.; Jorg, K.; David, G.M.M.; Roxana, S. Spatially coupled sparse codes on graphs: Theory and practice. *IEEE Commun. Mag.* **2014**, *52*, 168–176.
12. Lu, Q.; Fan, J.; Sham, C.W.; Tam, W.M.; Lau, F.C. A 3.0 Gb/s Throughput Hardware-Efficient Decoder for Cyclically-Coupled QC-LDPC Codes. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2016**, *63*, 134–145.
13. Yang, L.; Ying, L. Design of masking matrix for QC-LDPC codes. In Proceedings of the 2013 IEEE Information Theory Workshop (ITW), Seville, Spain, 9–13 September 2013; pp. 1–5.
14. Alireza, T.; Alireza, B.; Alireza, S. Symmetrical Constructions for Regular Girth-8 QC-LDPC Codes. *IEEE Trans. Commun.* **2017**, *65*, 14–22.
15. Inseon, K.; Hong, S. Some New Constructions of Girth-8 QC-LDPC Codes for Future GNSS. *IEEE Commun. Lett.* **2021**, *25*, 3780–3784.
16. Guo, H.Z.; Rong, S.; Xin, M.W. Construction of girth-eight QC-LDPC codes from greatest common divisor. *IEEE Commun. Lett.* **2013**, *17*, 369–372.
17. Jian, H.Z.; Guo, H.Z. Deterministic girth-eight QC-LDPC codes with large column weight. *IEEE Commun. Lett.* **2014**, *18*, 656–659.
18. Guo, H.Z.; Yu, L.H.; Yi, F.; De, F.R. Relation Between GCD Constraint and Full-Length Row-Multiplier QC-LDPC Codes with Girth Eight. *IEEE Commun. Lett.* **2021**, *25*, 2820–2823.
19. Ali, D.; Amir, H.B. From Cages to Trapping Sets and Codewords: A Technique to Derive Tight upper Bounds on the Minimum Size of Trapping Sets and Minimum Distance of LDPC Codes. *IEEE Trans. Inf. Theory* **2019**, *65*, 2062–2074.
20. Ghaffar, R.; Mohanmmad, G. Edge coloring of graphs with applications in coding theory. *China Commun.* **2021**, *18*, 181–195.
21. Mohammad, G.;Ghaffar, R. Large Girth Column-Weight Two and Three LDPC Codes. *IEEE Commun. Lett.* **2014**, *18*, 1671–1674.
22. Mohammad, G.; Mehdi, S.; Ghaffar, R. Column-Weight Three QC LDPC Codes with Girth 20. *IEEE Commun. Lett.* **2013**, *17*, 1439–1442.
23. Mohammad, G.; Masoumeh, A. Explicit APM-LDPC codes with girths 6, 8, and 10. *IEEE Signal Process Lett.* **2017**, *24*, 741–745.
24. Juane, L.; Shu, L.; Khaled, A.G.; William, R.; Daniel, C. Globally coupled LDPC codes. In Proceedings of the 2016 Information Theory and Applications Workshop (ITA), La Jolla, CA, USA, 31 January–5 February 2016; pp. 1–10.
25. Xiao, M.; Qian, F.W.; Mang, G.X.; Sui, H.C. Implicit Globally-Coupled LDPC Codes Using Free-Ride Coding. In Proceedings of the 2022 IEEE Wireless Communications and Networking Conference (WCNC), Austin, TX, USA, 10–13 April 2022; pp. 1117–1122.

26. Yen, C.L.; Hsie, C.C.; Shu, L. Scalable Globally-Coupled Low-Density Parity Check Codes. In Proceedings of the 2018 IEEE 10th International Symposium on Turbo Codes Iterative Information Processing (ISTC), Hong Kong, China, 3–7 December 2018; pp. 1–5.
27. Juane, L.; Keke, L.; Shu, L.; Khaled, A.G. Reed-Solomon based globally coupled quasi-cyclic LDPC codes. In Proceedings of the 2017 Information Theory and Applications Workshop (ITA), San Diego, CA, USA, 12–17 February 2017; pp. 1–10.
28. Yen, C.L.; Hsie, C.C.; Shu, L. Generalized Globally-Coupled Low-Density Parity-Check Codes. In Proceedings of the 2018 IEEE Information Theory Workshop (ITW), Guangzhou, China, 25–29 November 2018; pp. 1–5.
29. Ji, Z.; Bao, M.B.; Min, Z.; Shuang, Y.L.; Huaan, L. Protograph-Based Globally-Coupled LDPC Codes over the Gaussian Channel With Burst Erasures. *IEEE Access* **2019**, *7*, 153853–153868.
30. Ji, Z.; Bao, M.B.; Shuang, Y.L.; Min, Z.; Huaan, L. Tail-Biting Globally-Coupled LDPC Codes. *IEEE Trans. Commun.* **2019**, *67*, 8206–8219.
31. Yen, C.L.; Chien, L.; Hsie, C.C.; Shu, L. A (21150, 19050) GC-LDPC Decoder for NAND Flash Applications. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2019**, *66*, 1219–1230.
32. Bocharova, I.E.; Hug, F.; Johannesson R.; Kudryashov, B.D.; Satyukov, R.V. Searching for Voltage Graph-Based LDPC Tailbiting Codes with Large Girth. *IEEE Trans. Inf. Theory* **2012**, *58*, 2265–2279.