SPARSE CODES DERIVED FROM GRAPHS

Shahrooz Janbaz[1], Zahra Taheri[2], Ali Zaghian[3]

In this paper, we present a method to construct column weight two Low-Density Parity-Check (LDPC) codes (namely, Cycle codes) from arbitrary graphs and we obtain a new class of girth twelve LDPC codes from complete graphs. Also, we use the incidence matrix of bi-regular bipartite graphs to construct new sparse codes and we prove that some classes of these codes are self-orthogonal. The weight distribution of the later codes are obtained. Also, a conjecture about the covering radius of these codes are presented with some computational evidences. This conjecture partially solved for special class of these codes and we posed an interesting problem in the conclusion. At the end of this paper, the performances of constructed codes are simulated on Additive White Gaussian Noise (AWGN) channels.

MSC2010: 94B05, 94C15

Keywords: Sparse Code, Tanner Graph, Cycle Code, Weight Distribution, Covering Radius.

1. Introduction

Low-density parity-check ($LDPC$) codes, introduced by Gallager [1], are linear block codes with sparse parity-check matrices and implementable decoders. These codes provide near-capacity performance on a large set of data-transmission and data-storage channels.

$LDPC$ codes with column weight two have their minimum distance increasing logarithmically with code size [1]. There are a lot of methods for construction of $LDPC$ codes based on graphs, finite geometry and design theory, one may see [2] and the references therein. Recently, cage graph are used to construct column weight two $LDPC$ codes with wide range of girth[3]. Also, in [4] the authors constructed some new cycle $LDPC$ codes based on Tanner graphs of $LDPC$ codes. One advantage of column weight two $LDPC$ codes are their simple mathematical structure that allow to compute the parameters of codes. Also, there are some methods such as superposition on a larger finite field to obtain non-binary good $LDPC$ codes from column weight two $LDPC$ codes. There are several applications for column weight

[1]Corresponding author, Malek-Ashtar university of Technology, Email: shahrooz.janbaz@sci.ui.ac.ir

[2]Email: z.taheri@math.iut.ac.ir

[3]Email: a_zaghian@mut-es.ac.ir

two $LDPC$ codes. For example, in [5],[6] and [7] column weight two codes are investigated for disk storage because of their low complexity (few edge connection).

In this paper, graph theory is used to introduce some methods for constructing column weight two $LDPC$ codes. A new class of girth twelve $LDPC$ codes is obtained from complete graphs. Also, we use the incidence matrix of a Tanner graph to construct a new sparse code. The weight distribution and covering radius of a class of later codes are computed and it is proved that some classes of these codes are self-orthogonal. We conjectured about the covering radius of a class of these codes and some evidences are given to confirm this conjecture. Finally, the simulation results of the constructed codes on $AWGN$ channels are given.

2. **Preliminaries**

Let $\mathbb{F}_2 = \{0, 1\}$ denotes the finite field with 2 elements. A binary regular $LDPC$ code is defined as the null space of a parity-check matrix H over $\mathbb{F}_2$ with the following structural properties: 1) each row has constant weight ρ; 2) each column has constant weight γ. This parity-check matrix, H, is said to be (ρ, γ)-regular and the code given by the null space of H is called a (ρ, γ)-regular $LDPC$ code. An $LDPC$ code is said to be *irregular* if its parity-check matrix has multiple column weights and/or multiple row weights.
The parity-check matrix H must be sparse, i.e, the number of 1s in H must be much fewer than the total number of entries of H. Because of this property, $LDPC$ codes are known as sparse codes. Let H be an $m \times n$ parity-check matrix of an $LDPC$ code $\mathcal{C}$. We say that $\mathcal{C}$ is an $[n, k, d]$-code when n is the length of the code $\mathcal{C}$, the number of columns of H, k is the dimension of the code $\mathcal{C}$, equal to $n - rank(H)$, and $d = d_{min}(\mathcal{C})$ is the minimum distance of the code $\mathcal{C}$. It is known that if d is the minimum distance of the code $\mathcal{C}$, then each $d-1$ columns of H are independent but there is at least one set of d columns of H that is dependent.

The performance of an $LDPC$ code with iterative decoding depends on the number of structural properties of the code besides its minimum distance. One such structural property is the *girth* of the code which is defined as the length of the *shortest cycle* in the code's Tanner graph. The cycles that affect code performance the most are cycles of length 4. For codes whose Tanner graphs contain these short cycles, messages exchanged in iterative decoding become correlated after two iterations, and decoding either does not converge or converges slowly. Therefore, cycles of length 4 must be prevented in code construction. This is the case in almost every method of constructing $LDPC$ codes that has been proposed.

In the sequel, we give some definitions and theorems that will be used in other sections.

Definition 2.1. [8] *A subset S of code bits forms a stopping set if each equation that involves the bits in S involves two or more of them. In the context of the Tanner graph, S is a stopping set if all of its neighbours are connected to S at least twice.*

Definition 2.2. [9] *Let $\mathcal{C}$ be a code over $\mathbb{F}_2$ of length n. We say that a vector in $\mathbb{F}_2^n$ is ρ-covered by $\mathcal{C}$ if it has distance ρ or less from at least one codeword in $\mathcal{C}$. In this terminology the covering radius $Cr(\mathcal{C})$ of $\mathcal{C}$ is the smallest integer ρ such that every* *vector in* $\mathbb{F}_2^n$ *is* ρ*-covered by* $\mathcal{C}$*. Equivalently,*

$$Cr(\mathcal{C}) = \max_{x\in\mathbb{F}_2^n} \min_{c\in\mathcal{C}} d(x,c),$$

where $d(x,c)$ *denotes the distance between vectors* x *and* c*.*

Definition 2.3. [10] *If* $\mathcal{C}$ *is a linear code of length* n *over* $\mathbb{F}_2$ *with parity-check matrix* H*, the dual code of* $\mathcal{C}$ *is denoted by* $\mathcal{C}^{\perp}$ *and is defined as*

$$\mathcal{C}^{\perp} = \{\mathbf{x} \in \mathbb{F}_2^n \mid \mathbf{x}\cdot\mathbf{y} = 0 \text{ for all } \mathbf{y} \in \mathcal{C}\},$$

where $\mathbf{x}\cdot\mathbf{y}$ *denotes the Euclidean inner product of two vectors* $\mathbf{x}$ *and* $\mathbf{y}$*.*
Also, for the parity-check matrix H*, we have* $\mathcal{C} = H^{\perp}$*, where*

$$H^{\perp} = \{\mathbf{x} \in \mathbb{F}_2^n \mid H\mathbf{x}^t = 0\}.$$

We say that a code $\mathcal{C}$ *is self-orthogonal if* $\mathcal{C} \subseteq \mathcal{C}^{\perp}$*, and that* $\mathcal{C}$ *is self-dual if* $\mathcal{C} = \mathcal{C}^{\perp}$*.*

Theorem 2.4. [10, Theorem 2.1] *If* $G = [\mathbf{I}_k \mid A]$ *is a generator matrix for binary linear code* $\mathcal{C}$ *in standard form, then* $H = [A^t \mid \mathbf{I}_{n-k}]$ *is a parity-check matrix for* $\mathcal{C}$*.*

Suppose n and k are positive integers, then the number of k-combinations of n elements is denoted by $C(n,k)$. For any k, $0 \le k \le n$, we have $C(n,k) = \frac{n!}{k!(n-k)!}$ and for $k > n$, it is defined to be zero. The weight distribution (or weight spectrum) of a code of length n specifies the number of codewords of each possible weight $0, 1, \ldots, n$.

Theorem 2.5. [10] *Let* $\mathcal{C}$ *be an* $[n,k,d]$ *code over* $\mathbb{F}_2$ *with weight distribution* $(A_i \mid 0 \le i \le n)$ *and let the weight distribution of* $\mathcal{C}^{\perp}$ *be* $(B_i \mid 0 \le i \le n)$*. Then for* $0 \le v \le n$*, we have Mac Williams equation as follows*

$$\sum_{j=0}^{n-v} C(n-j,v)A_j = 2^{k-v}\sum_{j=0}^{v} C(n-j,n-v)B_j.$$

Theorem 2.6. [10] *Let* $\mathcal{C}$ *be a binary linear code. Then*

1) *If* $\mathcal{C}$ *is self-orthogonal and has a generator matrix each of whose rows has weight divisible by* 4*, then every codeword of* $\mathcal{C}$ *has weight divisible by* 4*.*
2) *If every codeword of* $\mathcal{C}$ *has weight divisible by* 4*, then* $\mathcal{C}$ *is self-orthogonal.*

3. New Method to Construct *LDPC* codes

In this section we construct a new parity-check matrix of an *LDPC* code by using the incidence matrix of a graph. Our notations are standard and mainly taken from [8] and [10].

Let $G = (V,E)$ be a connected graph with the vertex set $V = \{v_1, v_2, \ldots, v_m\}$ and the edge set $E = \{e_1, e_2, \ldots, e_n\}$. We denote the incidence matrix of the graph G by $I(G)$, that is an $m \times n$ binary matrix with row labels V and column labels E such that its ij-th entry is 1 iff the vertex v_i be an endpoint of the edge e_j. We consider this matrix, $I(G)$, as a parity-check matrix of an *LDPC* code derived from G. We denote the Tanner graph of this parity-check matrix by $T(G)$, which has the check nodes $V = \{v_1, \ldots, v_m\}$ and the variable nodes $E = \{e_1, e_2, \ldots, e_n\}$. Note that the Tanner graph of this parity-check matrix is bipartite with partitions V, E, where each vertex in E has degree 2. Also, we know that if G be a Tanner graph with check nodes $X = \{x_1, \ldots, x_s\}$ and variable nodes $Y = \{y_1, \ldots, y_t\}$, then by our method the resulting $LDPC$ code has the check nodes $X \cup Y$ and the variable nodes E. Note that each bipartite graph can be seen as a Tanner graph of a parity-check matrix.

In the following lemma, we will see some elementary properties of Tanner graph $T(G)$, for the given graph G.

Lemma 3.1. *Let $G = (V, E)$ be a connected graph and D_G, $deg_G(v)$ and $g(G)$ denote the diameter of G, degree of vertex v and the girth of G, respectively. Then the Tanner graph $T(G)$ has the following properties*

1) $|V(T(G))| = |V(G)| + |E(G)|$,
2) $|E(T(G))| = 2|E(G)|$,
3) $D_{T(G)} = 2D_G$,
4) $deg_{T(G)}(v) = \begin{cases} deg_G(v) & \text{if } v \in V(G) \\ 2 & \text{otherwise} \end{cases}$,
5) $g(T(G)) = 2g(G)$.

Proof. By the definition of Tanner graph and incidence matrix of G, all vertices and edges of G are the vertices of $T(G)$, so the item (1) is clear. For proving (2), it suffices to see that the total number of edges in $T(G)$ are twice of the number of edges of G, since each variable node of $T(G)$ has degree 2. Items (3) and (5) are clear, since if we delete the variable nodes in $T(G)$, the corresponding edges of these variable nodes still remain, the remaining graph is G. The item (4) is easy and is left as an exercise. □

For better understanding of the construction method, we give an example with some details.

Example 3.2. *Let K_3 be the complete graph with vertex set $\{v_1, v_2, v_3\}$ and edge set $\{e_1, e_2, e_3\}$. Then the graph K_3, the incidence matrix $I(K_3)$, its Tanner graph $T(K_3)$, and the incidence matrix of this Tanner graph $I(T(K_3))$ are shown in the following:*

We can generalize our previous example and obtain a family of well structured *LDPC* codes. Let $\mathbf{1}_n = [1, 1, \ldots, 1]_{1\times n}$, $\mathbf{0}_n = [0, 0, \ldots, 0]_{1\times n}$ and $\mathbf{I}_n$ be the $n \times n$ identity matrix.

Theorem 3.3. *Let $n \geq 3$ be a positive integer and K_n be the complete graph with n vertices. Then we have*

$$I(K_n) = \begin{bmatrix} \mathbf{0}_{\frac{(n-1)(n-2)}{2}} & \mathbf{1}_{n-1} \\ I(K_{n-1}) & \mathbf{I}_{n-1} \end{bmatrix}.$$

Also, the LDPC code that is derived from $I(K_n)$, $\mathcal{C}_n = I(K_n)^{\perp}$, is the $[C(n, 2), C(n, 2) - n + 1, 3]$-code.

Proof. The graph K_n has n vertices and $C(n, 2)$ edges, so $I(G)$ is an $n \times C(n, 2)$ matrix. Each column has weight 2 and each row has weight $n - 1$, since each edge belongs to two vertices and each vertex belongs to $n-1$ edges. So, $I(K_n)$ is an $(2, n-1)$-regular *LDPC* code. Now, suppose the vertex set of K_n be $\{v_1, v_2, \ldots, v_n\}$. It is easy to see that by deleting the vertex v_n and its adjacent edges form K_n, we obtain the complete graph K_{n-1}. Now, suppose we have the matrix $I(K_{n-1})$. We add new top row with label v_n to $I(K_{n-1})$ and new $n - 1$ columns for its corresponding edges in K_n at the end of $I(K_{n-1})$. By this method, the structure of $I(K_n)$ can be constructed recursively and is the same as presented in theorem. The column weight of $I(K_n)$ is 2, so its rank is $n-1$. Therefore, $\mathcal{C}_n = I(K_n)^{\perp}$ is an $[C(n, 2), C(n, 2) - n + 1, 3]$-code, since the girth of $T(K_n)$ is six and the minimum distance of derived code is half the girth of its Tanner graph. This completes our proof. □

As seen, the girth of the $T(K_n)$ is equal to six. We can repeat the above construction to obtain twelve girth code. But before doing this, we need the next theorem to obtain good structure for the new following codes. Note that for a matrix A, we denote its transpose by A^t.

Theorem 3.4. *Let G be a Tanner graph of parity-check matrix H with the check nodes $\{a_1, a_2, \ldots, a_m\}$, the variable nodes $\{b_1, b_2, \ldots, b_k\}$, and the edge set $\{e_1, e_2, \ldots, e_n\}$. Suppose $I(G)$ is the incidence matrix of the graph G. Then, there are two matrices B and C such that $I(G) = [B\ C]^t$ and $BC^t = H$.*

Proof. Let B and C be two matrices with check nodes $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_k\}$, respectively. Suppose their common variable nodes are the edges of graph G. By this notation, it is clear that $I(G) = [B\ C]^t$. Also, B and $\mathcal{C}$ are $m \times n$ and $k \times n$ matrices, respectively. So BC^t and H have the same size. Let t_{ij} be the ij-th entry of the matrix $I(G)$. We know that $t_{ij} = 1$ iff a_i is one of the endpoints of the edge e_j. But by our construction, the product of the i-th row of B and the j-th column of C^t is 1 iff $a_i b_j$ is an edge of $T(G)$ iff a_i is one of the endpoint of edge e_j. But it is equivalent to the definition of parity-check matrix H and we have $BC^t = H$. □

In the following, we consider $I(T(K_n))$ and its Tanner graph which give us a twelve girth *LDPC* code. Also, we use Theorem 3.4 to find the structure of this parity-check matrix.

By a suitable labelling of the vertices of $T(K_n)$, based on notations in previous theorem, we obtain a well structured parity-check matrix $I(T(K_n))$. Actually, we have

$$I(T(K_n)) = \begin{bmatrix} M_n^1 \\ M_n^2 \end{bmatrix}, \tag{1}$$

where $M_n^1 = \mathbf{I}_n \otimes \mathbf{1}_{n-1}$ and M_n^2 can be constructed recursively, as follows: Let, see Example 3.2,

$$I(T(K_3)) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

So, we have

$$M_3^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Let $B_n = [b_{ij}]$ be a $(n-2) \times 2C(n-1, 2)$ matrix where $b_{ij} = 1$ when $j = i(n-1)$. For i, $1 \le i \le n-2$, let A_i be the submatrix of M_{n-1}^2 of size $C(n-1, 2) \times (n-2)$ such that it begins from the column $i(n-2)+1$. Then we have

$$M_n^2 = \left[\begin{array}{ccccc:ccc} & & \mathbf{0}_{C(n-2,2),2(n-1)} & & & & & \\ & & & & & A_1^* & \dots & A_{n-2}^* \\ \mathbf{0}_{n-2,n-1} & & & \mathbf{I}_{n-2} & \mathbf{0}_{n-2,1} & & & \\ 1 & 0 & \dots & 0 & 1 & & & \\ \mathbf{0}_{n-2,1} & & \mathbf{I}_{n-2} & & \mathbf{0}_{n-2,n-1} & & B_n & \end{array}\right],$$

where

$$A_i^* = \begin{bmatrix} A_i & & \mathbf{0}_{C(n-1,2),1} \\ & \mathbf{0}_{1,n-1} & \end{bmatrix}.$$

In the next theorem, we give some properties of the later constructed codes.

Theorem 3.5. *Let $I(T(K_n))$ be the parity-check matrix of code $\mathcal{C}_n$. Then we have:*
1) $\mathcal{C}_n$ is an $[2C(n,2), C(n-1,2)-2, 6]$-code,
2) $\mathcal{C}_n = [\boldsymbol{I}_{C(n-1,2)} | M_{n-1}^2 (M_{n-1}^1)^t]$.

Proof. It is clear that the number of columns in $I(T(K_n))$ is $n(n-1) = 2C(n,2)$. Also the number of its rows is $n + C(n,2)$ with one dependent row, so the dimension of $\mathcal{C}_n$ is $2C(n,2) - C(n+1,2) - 1 = C(n-1,2) - 2$. Also, the column weight of $I(T(K_n))$ is two and its girth is 12. So $d_{min}(\mathcal{C}_n) = \frac{12}{2} = 6$. This completes the proof of part one of theorem. Now, by the structure of $I(T(K_n))$ and some elementary row and column operations, the second part of theorem is obvious. □

Recall that a stopping set S is a subset of variable nodes, such that all neighbours of the variable nodes in S are connected to S at least twice. The size of the stopping set S is the cardinality of S (for an equivalent definition see Definition 2.1). In the following lemma, we determine the structure of the stopping sets of our codes.

Note that, when we say the cycle (stopping set) distribution of a code is $\sum_{i\geq 1} c_i x^i$, it means that we have c_i cycles (stopping sets) with length (size) i.

Lemma 3.6. *Let G be a Tanner graph with cycle distribution $\sum_{i\geq 2} c_{2i} x^{2i}$. Then the cycle and stopping set distribution of $T(G)$, Tanner graph of $I(G)$, are $\sum_{i\geq 2} c_{2i} x^{4i}$ and $\sum_{i\geq 2} c_{2i} x^i$, respectively.*

Proof. Suppose there exist an l-cycle in G. Since G is bipartite l must be even. From the definition of $I(G)$, we obtain a $2l$-cycle in its corresponding Tanner graph, $T(G)$, and it is one to one correspondence. Furthermore, since the degree of each variable node in $T(G)$ is 2, for every $2l$-cycle in $T(G)$ there exists a stopping set with size l, and vice versa. Thus we completed the proof. □

Remark 3.7. *By Lemma 3.6, one can easily obtain the cycle and the stopping set distributions of the Tanner graph of $I(T(K_n))$.*

Recall that the weight of a codeword in a code $\mathcal{C}$ is the number of non-zero elements in this codeword. Let A_w denote the number of codewords in $\mathcal{C}$ with the weight w. In the following remark, we give some facts about the amount of A_w, for some w, in the code $\mathcal{C}_n = I(T(K_n))^\perp$.

Remark 3.8. *Let A_w be the number of codewords with weight w in the code $\mathcal{C}_n = I(T(K_n))^\perp$. By a simple calculation, one can see that $A_{i(n-1)} \geq C(n,i)$, $A_{2i} \geq C(C(n,2),i)$ and $A_{n-1+2i} \geq nC(C(n-1,2),i)$.*

4. **Incidence Matrix of the Tanner Graph as a Sparse Code**

In the previous section, we introduced a method of construction of *LDPC* codes from an arbitrary graph with the incidence matrix of an arbitrary graph. We know that the subdivision of any connected graph G is a bipartite graph. By this evidence, we are motivated to construct sparse codes from arbitrary bi-regular bipartite graphs. As we know, all Tanner graphs of $(\rho,\gamma)-$regular *LDPC* codes are bi-regular bipartite graphs. It can be interesting that we modify the relation between the *LDPC* code from a Tanner graph and the resulting sparse code by its incidence matrix. Note that, all bi-regular bipartite graph can be seen as a Tanner graph of an *LDPC* code and vice versa.

Let G be a Tanner graph of a $(\rho,\gamma)-$regular *LDPC* code. As it is done in the previous section, we use the incidence matrix of G as a sparse matrix of the new code and we denote it by $I(G)$. In the following, we give two different examples which show the interesting behaviours of our construction. The Tanner Graphs of Gallager parity-check matrix and MacKay Neal parity-check matrix are bi-regular and have the same degree sequences. But these two graphs are not isomorphic as a graph theory point of view. The resulting codes by our construction also are not equivalent.

Example 4.1. *Let G be the Tanner graph of length 12 $(4,3)$-regular Gallager parity-check matrix[2]. Then $I(G)$ generates $[36,20,3]$-code that is sparse. Also, the code $I(G)^\perp$ is a $[36,16,4]$-code.*

Example 4.2. *Let G be the Tanner graph of length* 12 $(4,3)$*-regular MacKay Neal parity-check matrix*[2]*. Then $I(G)$ generates* $[36,20,3]$*-code that is sparse. Also, the code $I(G)^{\perp}$ is a* $[36,16,4]$*-code.*

As it can be seen, the parameters of these two codes are the same, but these codes are not equivalent. It seems that non-isomorphic bipartite graphs with the same degree sequences generate non-equivalent sparse codes.

In the following, we give a method to standardize our construction for better seeing the properties of the matrix $I(G)$.

Let G be a Tanner graph of $(\rho,\gamma)-$regular $LDPC$ code with check nodes and variable nodes $\{a_1,a_2,\dots,a_k\}$ and $\{b_1,b_2,\dots,b_n\}$, respectively. Let for every i, $1\le i\le k$, $\deg(a_i)=\rho$ and for every j, $1\le j\le n$, $\deg(b_j)=\gamma$. It is easy to see that $k\rho=n\gamma$ and $I(G)$ has $k\rho$ columns. Let the labels of the columns of $I(G)$ be $e_1,e_2,\dots,e_{k\rho}$ and the labels of the rows of $I(G)$ be $a_1,a_2,\dots,a_k,b_1,b_2,\dots,b_n$. We correspond to any check nodes a_i, the edges $e_{(i-1)\rho+1},e_{(i-1)\rho+2},\dots,e_{(i-1)\rho+\rho}$. Then we choose integers $n_1,n_2,\dots,n_\gamma$ such that $1\le n_1<n_2<\ldots<n_\gamma\le n\gamma$ and $n_i=n_{i-1}+n$. Now we consider the edges $e_{m_{1j}},e_{m_{2j}},\dots,e_{m_{\gamma j}}$ where $m_{lj}=n_l+(j-1)$ $(\text{mod } n\gamma)$, and relate these edges to vertex b_j. Note that the obtained matrix is unique up to isomorphism, since $I(G)$ is independent from the choice of integers $n_1,n_2,\dots,n_\gamma$.

The weight distribution of a code $\mathcal{C}$ is the set $\{<w,A_w>\mid\ 0\le w\le n\}$, where n is the length of the code $\mathcal{C}$ and $\langle w,A_w\rangle$ denotes the weight enumerator of this code. It is known that in general determining the weight distribution of a code is very difficult.

Remark 4.3. *It is easy to see that* $I(G)=[H_1,H_2]^t$*, where* $H_1=\boldsymbol{I}_k\otimes\boldsymbol{1}_\rho$*,* $H_2=[A\ (\boldsymbol{1}_{\gamma-1}\otimes\boldsymbol{I}_n)\ B]$ *and* $[B\ A]=\boldsymbol{I}_n$*.*

Theorem 4.4. *Let in Tanner graph G we have $\rho=n$ (and so $\gamma=k$). Then* $I_{k,\rho}:=I(G)=\begin{bmatrix}\boldsymbol{I}_k\otimes\boldsymbol{1}_\rho\\ \boldsymbol{1}_k\otimes\boldsymbol{I}_\rho\end{bmatrix}$ *that is equivalent with*

$$\begin{bmatrix} & \vdots & \boldsymbol{1}_{(\rho-1)(k-1)}\\ \boldsymbol{I}_{k+\rho-1} & \vdots & \boldsymbol{1}_{k-1}\otimes\boldsymbol{I}_{\rho-1}\\ & \vdots & \boldsymbol{I}_{k-1}\otimes\boldsymbol{1}_{\rho-1}\end{bmatrix}$$

and the weight distribution of this code is given in Table 1.

Proof. By the above remark and some elementary row and column operations, one can easily obtain the equivalent code of $I_{k,\rho}$. Let the first row, next $\rho-1$ rows, and the last $k-1$ rows of equivalent matrix of $I_{k,\rho}$ are blocks B_1, B_2 and B_3, respectively. Since the blocks B_1, B_2 and B_3 are well-structured, we can obtain the weight of the different combinations of rows by a simple calculation. For example if we choose i rows from block B_2, and j rows from block B_3, then we have a vector with weight $i(k-1)+i$ from block B_2 and a vector with weight $(\rho-1)j+j$ from block B_3 with ij common ones. So the summation of these two vectors has weight $ki+\rho j-ij$. But by the multiplication principle, the number of such vectors is $C(\rho-1,i)C(k-1,j)$. With a similar method, we obtain the weight distribution of this code that is given in Table 1. □

TABLE 1. Weight distribution of $I_{k,\rho}$ when $\rho = n$, where $1 \leq i \leq \rho-1$, $1 \leq j \leq k-1$.

w	A_w
$(\rho-1)(k-1)+1$	1
ki	$C(\rho-1,i)$
ρj	$C(k-1,j)$
$(\rho-i-1)(k-1)+i+1$	$C(\rho-1,i)$
$(k-j-1)(\rho-1)+j+1$	$C(k-1,j)$
$ki+\rho j-ij$	$C(\rho-1,i)C(k-1,j)$
$(k-j-1)(\rho-i-1)+i+j+1$	$C(\rho-1,i)C(k-1,j)$

Remark 4.5. *Let $\mathcal{C}_1 = \boldsymbol{I}_n \otimes \boldsymbol{1}_m$, $\mathcal{C}_2 = \boldsymbol{1}_m \otimes \boldsymbol{I}_n$ and let $Cr(\mathcal{C})$ denote the covering radius of the code $\mathcal{C}$. One can see that these two codes are equivalent and so the covering radius of these codes is $n\lfloor \frac{m}{2} \rfloor$. Also we have $Cr(\mathcal{C}_1^{\perp}) = Cr(\mathcal{C}_2^{\perp}) = n$ [9].*

Corollary 4.6. *Let $I_{k,\rho} = \begin{bmatrix} \boldsymbol{I}_k \otimes \boldsymbol{1}_\rho \\ \boldsymbol{1}_k \otimes \boldsymbol{I}_\rho \end{bmatrix}$ and $Cr(I_{k,\rho})$ be the covering radius of $I_{k,\rho}$. Then we have $Cr(I_{k,\rho}) \leq Min\{k\lfloor \frac{\rho}{2} \rfloor, \rho\lfloor \frac{k}{2} \rfloor, (\rho-1)(k-1)\}$. Also, if $\rho \geq k \geq 3$ then $Cr(I_{k,\rho}) \leq k\lfloor \frac{\rho}{2} \rfloor$.*

Proof. By Definition 2.2, since $\mathbf{I}_k \otimes \mathbf{1}_\rho$ and $\mathbf{1}_k \otimes \mathbf{I}_\rho$ are subcodes of $I_{k,\rho}$, the result is clear. Also, when $\rho \geq k \geq 3$, the minimum of the set in the above corollary is $k\lfloor \frac{\rho}{2} \rfloor$ and this completes the proof. □

Theorem 4.7. *Suppose $\langle w, A_w \rangle$ denotes the weight enumerator, where A_w is the number of codewords with weight w. Let $\rho = 2$, $n_1 = 1$, and $k \geq 2$ in Theorem 4.4. Then we have*

i) $I_{k,2} = \begin{bmatrix} \boldsymbol{I}_k \otimes \boldsymbol{1}_2 \\ \boldsymbol{1}_k \otimes \boldsymbol{I}_2 \end{bmatrix}$ *that its standard form is* $\begin{bmatrix} & \cdot & \boldsymbol{J}_{2,k-1} \\ \boldsymbol{I}_{k+1} & \cdot & \\ & \cdot & \boldsymbol{I}_{k-1} \end{bmatrix}$, *where $\boldsymbol{J}_{m,n}$ denotes a all one matrix with size $m \times n$. Also $I_{k,2}$ is a $[2k, k+1, 2]$-code.*

ii) *$\mathcal{C}(k) := I_{k,2}^{\perp}$ is a $[2k, k-1, 4]$ doubly even weight self-orthogonal code. Moreover $\mathcal{C}(k) = [\boldsymbol{1}_2 \otimes \boldsymbol{I}_{k-1}\ \boldsymbol{J}_{k-1,2}]$ with standard form $[I_{k-1}\ I_{k-1}\ J_{k-1,2}]$.*

iii) *The covering radius of $\mathcal{C}(k)$, $Cr(\mathcal{C}(k))$, is k and the weight distribution of $\mathcal{C}(k)$ is*
$\langle 4i, C(k-1, 2i-1) + C(k-1, 2i) \rangle = \langle 4i, C(k, 2i) \rangle$ for $0 \leq i \leq [\frac{k}{2}]$.

Proof. By Theorem 4.4 and Table 1, the structure and weight distribution of $I_{k,2}$ is clear and the minimum distance of $I_{k,2}$ is 2. Now by Theorem 2.4, the structure of $\mathcal{C}(k)$ is clear. By Theorem 2.5 or the structure of $\mathcal{C}(k)$, the weight distribution of $\mathcal{C}(k)$ is $\langle 4i, C(k, 2i) \rangle$, so the minimum distance of $\mathcal{C}(k)$ is 4 and it is a doubly even weight code. Moreover by Theorem 2.6, $\mathcal{C}(k)$ is self-orthogonal. Let $f_{2k} = [1010 \ldots 10]$ be a vector of length $2k$. By induction on k, we prove the hypothesis about covering radius of $\mathcal{C}(k)$. In the case $k = 2$, it is easy to see that the covering radius is 2. Now suppose that $Cr(\mathcal{C}(k)) = k$. For the code $\mathcal{C}(k+1)$, we have $f_{2(k+1)} = [f_{2k-2}1010]$. By induction, the vector f_{2k-2} has the distance $k-1$ from code $\mathcal{C}(k-1)$. But the vector [1010] has the distance 2 with all the combinations of rows related to the last four columns of $\mathcal{C}(k+1)$. So the covering radius of $\mathcal{C}(k+1)$ is $k-1+2=k+1$ and it completes the proof. □

Conjecture 4.8. *Let $k, \rho \in \mathbb{N}$ and $k, \rho \geq 2$. Then we have $Cr(I_{k,\rho}^{\perp}) = k$.*

For some random integers k and ρ, the covering radius of $I_{k,\rho}$ are computed with Magma Algebra System [11] and the results are summarized in Table 2. It can be seen that the results confirm our conjecture.

TABLE 2. Covering radius of $I_{k,\rho}^{\perp}$ for random parameters k and ρ

k	ρ	Cr	k	ρ	Cr
10	3	10	5	8	5
12	5	12	10	15	10
13	8	13	12	12	12
30	15	30	25	35	25

5. Simulation Results

In this section, we simulate the bit error rate of constructed codes on $AWGN$ channel. The decoding algorithm is Belief propagation algorithm which is a type of message passing algorithm. We simulated these codes by using the package of $LDPC$ analysis which is available in [12]. By Figure 1, we can see that the performance of $I(T(K_n))$ is better than the one of $I(K_n)$. Also the rate of $I(K_n)$ is better than the one of $I(T(K_n))$. We can explain this differences by the better girth and minimum distances of $I(T(K_n))$ than $I(K_n)$. From the Figure 2, it is easily seen that when k and ρ are close, the performance of $I_{k,\rho}$ will be better, when dB is greater than 5. We know that the girth and minimum distance of $I_{k,\rho}$ are fixed for different values of k and ρ. Therefore, we can explain these results with the increase of the rate of $I_{k,\rho}$ when k and ρ are close.

6. Conclusion

In this paper, we introduced some new methods to construct column-weight two $LDPC$ codes based on graphs. Also, we determined some structural properties of some classes of these new codes. We found a class of self-orthogonal codes that are suitable for constructing a new class of quantum codes. We gave a conjecture with some evidences about the covering radius of a class of these new codes. An interesting question that can be studied further on is: If G_1 and G_2 be two non-isomorphic bipartite graphs with the same degree, is it true that $I(G_1)$ and $I(G_2)$ are not equivalent?

REFERENCES

[1] R. G. Gallager, Low-Density parity-check Codes, Cambridge, MA: MIT Press, 1963.
[2] S. J. Johnson, Introducing Low-Density Parity-Check Codes, School of Electrical Engineering and Computer Science, The University of Newcastle, Australia.
[3] G. Exoo, R. Jajcay, Dynamic Cage Survey, Electron. J. Combin., DS16, pp. 155, Jul, 2013.

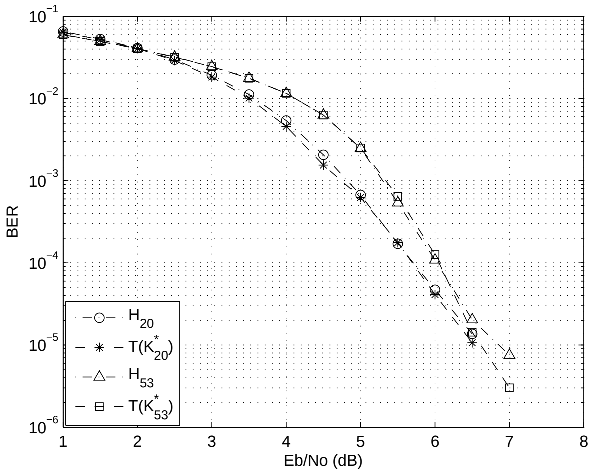

FIGURE 1. Error performances of $I(K_n) = H_n$ and $I(T(K_n)) = T(K_n^*)$ for $n = 20, 53$.

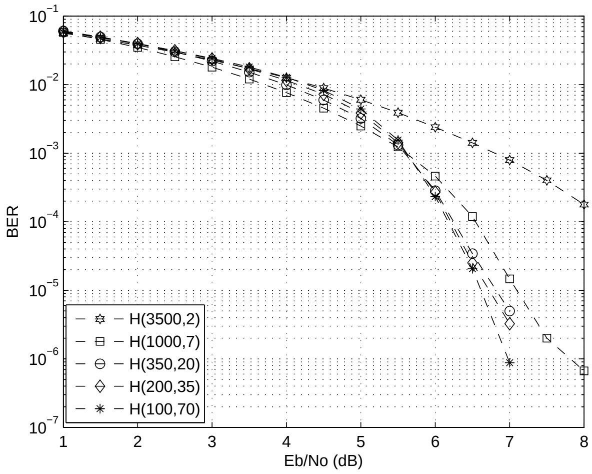

FIGURE 2. Error performances of $I_{k,\rho} = H(k, \rho)$ for some values of k and ρ.

[4] X. Tao, L. Zheng, W. Liu, D. Liu, Recursive Design of High Girth $(2, K)$ LDPC Codes from (k, k)-LDPC Codes, Electron. J. Combin., DS16, pp. 155, Jul, 2013.
[5] H. Song, J. Liu, V. Kumar, Low Complexity *LDPC* Codes for Partial Response Channels. IEEE Globecom. 2, 1294-1299, 2002.
[6] H. Song, J. Liu, V. Kumar, Large Girth Cycle Codes for Partial Response Channels., IEEE Transactions on Magnetics., 40(4), 2362-2371, 2004.
[7] H. Song, V. Kumar, Low-Density Parity-Check Codes for Partial Response Channels. IEEE Signal Processing Magazine. 21, 56-66, 2004.
[8] W. E. Ryan, S. Lin, Channel Codes Classical and Modern, Cambridge University Press, New York, 2009.
[9] W.C. Huffman, V. Pless, Fundamentals of Error-Correcting Codes, Cambridge University Press, New York, 2003.

[10] V. Pless, W.C. Huffman, Handbook of Coding Theory, Volume I, Elsevier Science B.V, 1998.
[11] University of Sydney, Magma Algebraic Software, Department of Mathematics, http://www.magma.maths. usyd.edu.au/magma, 2005.
[12] R.M. Neal, Software for Low Density Parity Check Codes, http://www.cs.toronto.edu / radford/ftp/$LDPC$-2006-02-08/index.html, 2006.