On the Tanner Cycle Distribution of QC-LDPC Codes from Polynomial Parity-Check Matrices

Anthony Gómez-Fonseca*, Roxana Smarandache*†, and David G. M. Mitchell‡
Departments of *Mathematics and †Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA
{agomezfo, rsmarand}@nd.edu
‡Klipsch School of Electrical and Computer Engineering, New Mexico State University, Las Cruces, NM 88003, USA
dgmm@nmsu.edu

2023 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-7554-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISIT54713.2023.10207005

***Abstract*—In this paper, we present an efficient strategy to enumerate the number of k-cycles, $g \leq k < 2g$, in the Tanner graph of a quasi-cyclic low-density parity-check (QC-LDPC) code with girth g using its polynomial parity-check matrix H. This strategy works for both (n_c, n_v)-regular and irregular QC-LDPC codes. In this approach, we note that the mth power of the polynomial adjacency matrix can be used to describe walks of length m in the protograph and can therefore be sufficiently described by the matrices $B_m(H) \triangleq \left(HH^{\mathsf{T}}\right)^{\lfloor m/2 \rfloor} H^{(m \mod 2)}$, where $m \geq 0$. For example, in the case of QC-LDPC codes based on the $3 \times n_v$ fully-connected protograph, the complexity of determining the number of k-cycles, $\mathcal{N}_k$, for $k = 4, 6$ and 8, is $O(n_v^2 \log(N))$, $O(n_v^2 \log(n_v) \log(N))$ and $O(n_v^4 \log^4(n_v) \log(N))$, respectively. The complexity, depending logarithmically on the lifting factor N, gives our approach, to the best of our knowledge, a significant advantage over previous works on the cycle distribution of QC-LDPC codes.**

I. Introduction

Low-density parity-check (LDPC) codes form a class of error-correcting codes that were discovered by Gallager [1] in the early 1960s and that have been shown to be capacity-approaching. Because of this, members of this class are now part of many industry standards, including those developed by the Consultative Committee for Space Data System (CCSDS) [2]. The subclass of quasi-cyclic LDPC (QC-LDPC) codes is attractive for both implementation and analysis purposes since its members can be described in a compact and simple way [3], [4].

The simple structure of QC-LDPC codes, and the graph representation of an LDPC code in general, plays a fundamental role in determining the performance of the code under iterative decoding algorithms. In fact, the girth [5], together with the number of short cycles [6], and other graphical structures composed of short cycles, such as trapping sets [7], are important parameters to measure the iterative decoding performance of the code. As a consequence, and for a long time, researchers have been actively trying to find ways to not only reduce but eliminate, when possible, all the short cycles in a graph in an attempt to improve the performance of the corresponding code.

It is well-known that enumerating the k-cycles in a general graph is hard [8], [9]. Consequently, a lot of effort has been dedicated to reduce the complexity of solving these problems. Several algorithms have been designed for cycle enumeration, but their complexities are, unfortunately, dependent on the number of vertices, the number of edges, and even the number of cycles, which may itself increase exponentially with the number of vertices.

Let $G = (V, E)$ be a bipartite graph with girth g. In [10], a message-passing algorithm for counting short cycles in a graph is presented. This algorithm is capable of counting k-cycles, with $g \leq k \leq 2g - 2$, in the case of bipartite graphs, with complexity growing as $O(g|E|^2)$. In [11], a matrix of size $2|E| \times 2|E|$, called the directed edge matrix, is constructed and used to count the number of short cycles. This strategy requires to calculate the trace of the kth power of this matrix or, equivalently, the eigenvalues of the kth power of the adjacency matrix. Such an approach has complexity $O(|E|^3)$ and becomes prohibitively high with an increase in the size of the Tanner graph. This approach is also analyzed in [12] in the case of QC-LDPC codes. The complexity of this approach is reduced from $O(N^3|E_b|^3)$ to $O(N|E_b|^3)$, where N is the lifting factor and $|E_b|$ is the number of edges in the protograph, by exploiting the circulant structure to compute the eigenvalues as in [13].

II. Definitions, notation and background

Let $\mathcal{C}$ be a QC-LDPC code, either (n_c, n_v)-regular or irregular, with block length $n_v N$ based on the $n_c \times n_v$ protograph [14] described by the matrix $B = (b_{ij})_{n_c \times n_v}$, where b_{ij} is a nonnegative integer for $i \in [n_c]$ and $j \in [n_v]$, and where $[l] \triangleq \{0, 1, \ldots, l-1\}$. Then $\mathcal{C}$ can be described by a (scalar) parity-check matrix $H = (H_{ij})_{n_c \times n_v}$, where each H_{ij}, for $i \in [n_c]$ and $j \in [n_v]$, is a summation of b_{ij} $N \times N$ circulant permutation matrices if b_{ij} is nonzero, and the $N \times N$ all-zero matrix if $b_{ij} = 0$. Graphically, this operation is equivalent to taking an N-fold graph cover, or *lifting*, of the protograph. Here, N is called the lifting factor (alternatively, lifting degree, or degree of the graph cover). We use the terms *image* and *inverse image* to refer to the projection from the graph cover to the protograph and the mapping from the protograph to the graph cover, respectively.

Let x^r denote the $N \times N$ circulant permutation matrix obtained by circularly shifting to the left, by r positions modulo N, the entries of the $N \times N$ identity matrix I. For simplicity in the notation, let $p_{ij}(x)$ be the polynomial representation

This material is based upon work supported by the National Science Foundation under Grant Nos. CCF-2145917, CNS-2148358, HRD-1914635, and OIA-1757207. A. G. F. thanks the support of the GFSD (formerly NPSC) and Kinesis-Fernández Richards fellowships.

of H_{ij}, where $p_{ij}(x) = \sum_{l=0}^{N-1} a_l x^l$ and $a_l \in \{0,1\}$ for all $l \in [N]$. Each polynomial $p_{ij}(x)$ has b_{ij} nonzero terms. Then we can rewrite the parity-check matrix H, using the polynomial representation, as $H = (p_{ij})_{n_c \times n_v}$.

From the parity-check matrix H, we construct a bipartite graph $G = (V, E)$, called a Tanner graph [15], by considering H as its biadjacency matrix. The set V is the set of vertices (or nodes) and E is the set of edges, and their cardinalities are denoted by $|V|$ and $|E|$, respectively. Denote the vertices of G by v_a, for $a = 0, 1, 2, \ldots, |V|-1$, and the edges by e_b, for $b = 0, 1, 2, \ldots, |E| - 1$. Each edge e_b has the form $e_b = (v_a, v_c)$, for some $v_a, v_c \in V$, and the vertices v_a and v_c are called the endpoints of e. A (directed) walk W of length m in the graph G is an alternating sequence $W = v_0 e_1 v_1 e_2 \cdots v_{m-1} e_m v_m$ of vertices and edges such that $e_l = (v_{l-1}, v_l) \in E$ for all $1 \leq l \leq m$. The first vertex appearing in the alternating sequence, v_0, is called the base point of W. A walk W is said to be closed if the two endpoints are the same, this is, when $v_0 = v_m$. A closed walk W is backtrackless if $e_l \neq e_{l+1}$ for all $1 \leq l \leq m-1$. A backtrackless closed walk W is tailless if $e_m \neq e_1$, and W is called, in this case, a TBC walk. A cycle is a closed walk W having distinct vertices, except for the endpoints, and if its alternating sequence has k edges in it, then we call W a k-cycle.

The adjacency matrix $A = (A_{ij})$ is the symmetric binary matrix with $A_{ij} = 1$ if $(v_i, v_j) \in E$, and $A_{ij} = 0$ otherwise. After some reordering of the vertices, if necessary, we can write A, for either the scalar or polynomial representation of H, in the compact expression

$$A = \begin{bmatrix} 0 & H \\ H^{\mathsf{T}} & 0 \end{bmatrix}, \tag{1}$$

where H^{T} denotes the transpose of H. The powers of A, and in particular the matrices

$$B_t(H) \triangleq \left(HH^{\mathsf{T}}\right)^{\lfloor t/2 \rfloor} H^{(t \mod 2)}, \quad t \geq 0, \tag{2}$$

give information about the walks [16]. It is not difficult to see that, for any nonnegative integer t, we have

$$A^{2t} = \begin{bmatrix} B_{2t}(H) & 0 \\ 0 & B_{2t}(H^{\mathsf{T}}) \end{bmatrix}, \tag{3}$$

and

$$A^{2t+1} = \begin{bmatrix} 0 & B_{2t+1}(H) \\ B_{2t+1}(H^{\mathsf{T}}) & 0 \end{bmatrix}. \tag{4}$$

Theorem 1 ([17])**.** *If $A^m = ((A^m)_{ij})$ is the mth power of the adjacency matrix A, then the entry $(A^m)_{ij}$ is equal to the number of walks of length m between the vertices v_i and v_j.*

Consider the triangle operator $\triangle$ introduced in [16]. For two nonnegative integers e and f, define $d = e \triangle f \triangleq 1$ if $e \geq 2$ and $f = 0$, and $d = e \triangle f \triangleq 0$ otherwise. This definition can be extended to matrices component-wise.

Theorem 2 ([16], [17])**.** *A Tanner graph of an LDPC code with parity-check matrix H has girth$(H) > 2l$ if and only if*

$$B_m(H) \triangle B_{m-2}(H) = 0 \quad \textit{for} \quad 2 \leq m \leq l. \tag{5}$$

The kth power of the scalar adjacency matrix A of the Tanner graph can be used to determine the number of k-walks between any two vertices, as we have seen in Theorem 1. The kth power of the polynomial version of the adjacency matrix, however, does not help us to count the number of k-walks between any two vertices of the Tanner graph, but, as we will see, can be used to describe the edges traversed in a k-walk between any two vertices in the protograph. For example, if A is the polynomial version of the adjacency matrix (1), then

$$(A^2)_{ij}(x) = \sum_{l=0}^{n_c+n_v-1} A_{il}(x) A_{lj}(x), \tag{6}$$

and every term of the polynomial $(A^2)_{ij}(x)$ is a product of the form $x^{c_{il}} x^{c_{lj}} = x^{c_{il}+c_{lj}}$, where $x^{c_{il}}$ and $x^{c_{lj}}$ come from the polynomials $A_{il}(x)$ and $A_{lj}(x)$, respectively. Each one of the two circulants $x^{c_{il}}$ and $x^{c_{lj}}$ corresponds to a unique edge in the protograph, and the order in which they appear in the product is the order used to traverse the walk in the protograph. The exponent $c_{il} + c_{lj}$, in consequence, corresponds to the two edges traversed from vertex v_i to vertex v_l to vertex v_j in the protograph. In the same way, every term of $(A^3)_{ij}(x)$ is a product of the form $x^{c_{il}} x^{c_{lk}} x^{c_{kj}} = x^{c_{il}+c_{lk}+c_{kj}}$, and the exponent $c_{il}+c_{lk}+c_{kj}$ corresponds to the three edges traversed in the protograph from vertex v_i to vertex v_l to vertex v_k to vertex v_j. In general, every term of $(A^m)_{ij}(x)$ is of the form $x^{c_{il_1}} x^{c_{l_1 l_2}} \cdots x^{c_{l_m j}} = x^{c_{il_1}+c_{l_1 l_2}+\cdots+c_{l_m j}}$, and each one of them corresponds to a walk of length m and the specific order in which it is traversed, which is nicely described by the way the matrix multiplication in (3) and (4) is performed. This allows us to state a polynomial version of Theorem 1.

Theorem 3. *If $A^m = ((A^m)_{ij}(x))$ is the mth power of the polynomial adjacency matrix A, then every term of the polynomial $(A^m)_{ij}(x)$ is of the form $x^{c_{il_1}} x^{c_{l_1 l_2}} \cdots x^{c_{l_m j}}$ and corresponds to a walk of length m between the vertices v_i and v_j in the protograph.*

Definition 4. *The exponent $c_{il_1} + c_{l_1 l_2} + \cdots + c_{l_m j}$ corresponding to the product $x^{c_{il_1}} x^{c_{l_1 l_2}} \cdots x^{c_{l_m j}}$ in Theorem 3 is called a permutation shift.* □

If $x^{c_{il_1}} x^{c_{l_1 l_2}} \cdots x^{c_{l_m j}}$ and $x^{c'_{il_1}} x^{c'_{l_1 l_2}} \cdots x^{c'_{l_m j}}$ are two terms of the polynomial $(A^m)_{ij}(x)$ describing two m-walks between vertices v_i and v_j in the protograph, then the combination

$$x^{c_{il_1}} x^{c_{l_1 l_2}} \cdots x^{c_{l_m j}} x^{-c'_{l_m j}} \cdots x^{-c'_{l_1 l_2}} x^{-c'_{il_1}}$$

of the first walk and the reversal of the second one describes a closed $(2m)$-walk that starts and ends at the vertex v_i, and that has the vertex v_j midway. Hence, the entries $(A^m)_{ij}(x)$ of the power A^m describe all the m-walks in the protograph and can be used to count certain cycles in the Tanner graph. The strategy of counting cycles in the Tanner graph presented in this paper requires to keep track of TBC walks in the protograph.

Lemmas 5 and 6, based on some results from [18], are useful to study both the images of cycles in the Tanner graph and the inverse images of TBC walks in the protograph.

Lemma 5 ([19], [20])**.** *Let $\tilde{G}$ be an N-fold graph cover of the protograph G. Let W be a k-walk in G starting at vertex v and ending at vertex v', and having edge sequence $e_1, e_2, \ldots, e_k$ with associated circulant permutation matrices $x^{s_1}, x^{s_2}, \ldots, x^{s_k}$. Then the permutation shift s that maps $\tilde{v}$, the inverse image of v in $\tilde{G}$, to $\tilde{v'}$, the inverse image of v' in $\tilde{G}$, through the walk $\tilde{W}$ is given by*

$$s = \sum_{i=0}^{k-1} (-1)^i s_{i+1} \mod N. \tag{7}$$

We denote by $\mathbb{Z}_N$ the additive group of integers modulo N. For any element $a \in \mathbb{Z}_N$, the order of a is the smallest integer m such that $a^m = m \cdot a = 0$.

Lemma 6 ([19])**.** *Let $\tilde{G}$ be an N-fold graph cover of the protograph G and let W' be a k-cycle in $\tilde{G}$. Then W' is projected onto a TBC walk W of length k/m, where $m \geq 1$ is the order of the permutation shift of W in $\mathbb{Z}_N$.*

We combine the following lemma with our analysis of TBC walk to count cycles in the Tanner graph.

Lemma 7 ([21])**.** *Let $\tilde{G}$ be an N-fold graph cover of the protograph G and let W be a closed k-walk in G. Then the inverse image of W in $\tilde{G}$ is the union of N/m closed (km)-walks, where $m \geq 1$ is the order of the permutation shift of W in $\mathbb{Z}_N$.*

We extend the result in Lemma 7, stated for closed k-walks, to TBC walks of length k.

Theorem 8. *Let $\tilde{G}$ be an N-fold graph cover of the protograph G and let W be a TBC walk of length k in G. Then the inverse image of W in $\tilde{G}$ is the union of N/m TBC walks of length km, where $m \geq 1$ is the order of the permutation shift of W in $\mathbb{Z}_N$.*

The following lemma explains why we restrict our analysis to k-cycles with $k < 2g$, where g is the girth of the Tanner graph.

Lemma 9 ([10])**.** *Let G be a graph with girth g. Then the set of TBC walks of length k coincides with the set of k-cycles if $k < 2g$.*

III. Counting cycles: a general protograph

The equivalence of closed walks is an important notion in this work.

Definition 10. *Two closed walks W_1 and W_2 are said to be equivalent if one can be obtained from the other by a change of base point, a change in direction, or both.* □

In the following definition, we introduce a set whose cardinality is used in the formulas for the number of k-cycles, $\mathcal{N}_k$, in the Tanner graph.

Definition 11. *Let H be a polynomial parity-check matrix and let N be the lifting factor. For integers $d \geq 0$ and $f \geq 1$, we denote by $W(d, f)$ the set of all nonequivalent TBC walks of length d in the protograph having permutation shift of order f in $\mathbb{Z}_N$.* □

Remark 12. Notice that the construction of the set $W(d, f)$ in Definition 11 depends on both the protograph and the lifting factor N. In algebra, the additive group $\mathbb{Z}_N$ has order N and the order of every element is a divisor of N. Hence, if s is not a divisor of N, the set $W(d, f)$ is empty independently of the selection of the length d. For example, if $N = 4$, no element in $\mathbb{Z}_4$ has order 3 because 3 does not divide 4, so the set $W(d, 3)$ is empty for any value of d. However, even if f does divide N, there are instances where the set $W(d, f)$ is automatically empty. If the protograph is the $n_c \times n_v$ fully-connected (all-ones), it is not possible to obtain a TBC walk of length 4 from the double traversal of a walk of length 2, forcing $W(2, f)$ to be empty. If the protograph is a multiedge graph, then it is possible to have a nonempty set $W(2, f)$. □

The following theorem gives the number of k-cycles in the Tanner graph using the walks described by the entries of the polynomial parity-check matrix H. Its proof, and the proof of other results in the paper, are omitted for space constraints.

Theorem 13. *Let H be the polynomial parity-check matrix of a protograph-based QC-LDPC code having girth g and let k be an even integer with $2 \leq k < 2g$. Let*

$$D(k) = \{d \mid d \text{ divides } k, d \geq 2, d \text{ even}\}$$

and, for any $d \in D(k)$, let $W(d, k/d)$ denote the set of nonequivalent TBC walks of length d having permutation shift of order k/d in $\mathbb{Z}_N$. Then the number $\mathcal{N}_k$ of k-cycles, $k < 2g$, in the corresponding Tanner graph G with parity-check matrix H is given by

$$\mathcal{N}_k = \sum_{d \in D(k)} \left(|W(d, k/d)| - \sum_{\substack{d' \in D(k) \\ d' \mid d,\ d' < d}} |W(d', k/d')| \right) \cdot \frac{N}{k/d}, \tag{8}$$

where N is the lifting factor.

IV. Counting cycles: all-ones protograph

The equation (8) in Theorem 13 gives the number of k-cycles in the Tanner graph of an arbitrary QC-LDPC code. Since the Tanner graph is a graph cover of the protograph, and $N \geq 1$, the former has, at least, the same number of vertices and edges as the latter. In practice, we restrict $N > 1$, so counting TBC walks in the protograph, instead of directly counting cycles in the Tanner graph, represents a reduction in the number of computations required to determine $\mathcal{N}_k$. In this section, we focus on the fully-connected (all-ones) protograph and discuss how to determine $\mathcal{N}_k$ with a strategy that has complexity logarithmic on the lifting factor N. This section provides an efficient way to find the cardinality of the sets $W(d, f)$, for integers $d \geq 0$ and $f \geq 1$. For space

requirements, we show how to calculate $\mathcal{N}_k$, with $k = 4, 6, 8$, using our strategy, but the approach works for any $k < 2g$.

First, consider the following definition.

Definition 14. *A multiset (shortened to mset) is a generalization of a set in which elements are allowed to repeat. The number of times an element occurs in a multiset is called its multiplicity. A repetition in a multiset $S = \{s_1, s_2, \ldots, s_\sigma\}$ is an equality $s_{i_0} = s_{i_1}$ with $i_0 < i_1$. If $T = \{t_1, t_2, \ldots, t_\tau\}$ is another multiset, a repetition between the multisets S and T is an equality $s = t$ for some $s \in S$ and $t \in T$.* □

Let $\mathcal{C}$ be a QC-LDPC code with parity check matrix H given by

$$H = \begin{bmatrix} x^{h_0} & x^{h_1} & \cdots & x^{h_{n_v-1}} \\ x^{i_0} & x^{i_1} & \cdots & x^{i_{n_v-1}} \\ x^{j_0} & x^{j_1} & \cdots & x^{j_{n_v-1}} \end{bmatrix}. \tag{9}$$

By Theorem 2, as discussed in [22], girth$(H) > 4$ if and only if all the elements in each one of the three msets $\{h_l - i_l \mid l \in [n_v]\}$, $\{h_l - j_l \mid l \in [n_v]\}$ and $\{i_l - j_l \mid l \in [n_v]\}$ are distinct. If one of these msets has a repetition, some 4-cycles appear in the Tanner graph and the exact amount of them is calculated in the following theorem.

Theorem 15. *Let H be as in* (9)*. A repetition in any of the following three msets $A_1 = \{h_l - i_l \mid l \in [n_v]\}$, $A_2 = \{h_l - j_l \mid l \in [n_v]\}$ and $A_3 = \{i_l - j_l \mid l \in [n_v]\}$ lifts to exactly N 4-cycles in the Tanner graph. The total number of 4-cycles in the Tanner graph, $\mathcal{N}_4$, is given by*

$$\mathcal{N}_4 = |W(4,1)| \cdot N, \tag{10}$$

and

$$|W(4,1)| = \sum_{m=1}^{3} \mathcal{R}_{A_m}, \tag{11}$$

where $\mathcal{R}_{A_m}$ is the number of repetitions $\alpha_l = \alpha_{l'}$ in A_m.

Example 16. Let H be the polynomial parity-check matrix given by

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & x & x^2 & x^3 & x^4 \\ 1 & x^2 & x & x^5 & x^7 \end{bmatrix}.$$

This matrix H has girth 4 for lifting factor $N = 5$. Calculating the three msets in Theorem 15 over $\mathbb{Z}_5$, we obtain the msets $\{0, 4, 3, 2, 1\}$, $\{0, 3, 4, 0, 3\}$ and $\{0, 4, 1, 3, 2\}$, Notice that there are two repetitions in the second mset, so there are two elements in $W(4,1)$, specifically $W(4,1) = \{h_0 - j_0 + j_3 - h_3, h_1 - j_1 + j_4 - h_4\}$. Hence, the number of 4-cycles $\mathcal{N}_4$ in the Tanner graph is $\mathcal{N}_4 = 2 \cdot 5 = 10$.

If we take $N = 10$, then the parity-check matrix H has girth 6. To confirm that there is no 4-cycle in H, we calculate the three msets in Theorem 15 over $\mathbb{Z}_{10}$ and we obtain the msets $\{0, 9, 8, 7, 6\}$, $\{0, 8, 9, 5, 3\}$ and $\{0, 9, 1, 8, 7\}$. Since there is no repetition in these msets, $\mathcal{N}_4 = 0$. □

Remark 17. Theorem 15 was used to calculate the number of elements in $W(4,1)$, but the strategy can be simply modified to count the number of elements in $W(4,2)$. In this case, we are not targeting repetitions in the three msets $\{h_l - i_l \mid l \in [n_v]\}$, $\{h_l - j_l \mid l \in [n_v]\}$ and $\{i_l - j_l \mid l \in [n_v]\}$; instead, we are looking for two elements α_l and $\alpha_{l'}$ with $\alpha_l \neq \alpha_{l'}$, coming from the same mset, such that $2 \cdot (\alpha_l - \alpha_{l'}) = 0$ in $\mathbb{Z}_N$. For example, for $l \neq l'$, let $\alpha_l, \alpha_{l'} \in \{h_l - i_l \mid l \in [n_v]\}$ be such that $\alpha_l = h_l - i_l$, $\alpha_{l'} = h_{l'} - i_{l'}$ and $\alpha_l \neq \alpha_{l'}$. Then $\alpha_l - \alpha_{l'} = h_l - i_l + i_{l'} - h_{l'}$ describes a TBC walk of length 4. The double traversal of this TBC walk has permutation shift given by $h_l - i_l + i_{l'} - h_{l'} + h_l - i_l + i_{l'} - h_{l'}$ and is an element of $W(4,2)$ if it is 0 in $\mathbb{Z}_N$. The same strategy is applied to the other two msets. This approach is used to compute $W(4,f)$, for any f, by requiring $f \cdot (\alpha_l - \alpha_{l'}) = 0$ in $\mathbb{Z}_N$. □

By Theorem 2, as discussed in [22], girth$(H) > 6$ if and only if, for $m \in [n_v]$, all the elements in each one of the msets

$$\{h_l - i_l + i_m, h_l - j_l + j_m \mid l \in [n_v], l \neq m\},$$
$$\{i_l - h_l + h_m, i_l - j_l + j_m \mid l \in [n_v], l \neq m\},$$
$$\{j_l - h_l + h_m, j_l - i_l + i_m \mid l \in [n_v], l \neq m\},$$

are distinct. In the following theorem, we rewrite these conditions in a way that is helpful to count 6-cycles in the Tanner graph.

Theorem 18. *Let H be as in* (9) *and, for $m \in [n_v]$ and $l \in [n_v] \backslash \{m\}$, consider the following msets*

$$A_{1,m} = \{(l, h_l - i_l + i_m)\}, \qquad B_{2,m} = \{(l, i_l - j_l + j_m)\},$$
$$A_{2,m} = \{(l, h_l - j_l + j_m)\}, \qquad C_{1,m} = \{(l, j_l - h_l + h_m)\},$$
$$B_{1,m} = \{(l, i_l - h_l + h_m)\}, \qquad C_{2,m} = \{(l, j_l - i_l + i_m)\}.$$

For $l, l' \in [n_v]$, let $(l, \alpha_l) \in A_{1,m}$ and $(l', \alpha_{l'}) \in A_{2,m}$ be such that $\alpha_l = \alpha_{l'}$. Then the repetition $\alpha_l = \alpha_{l'}$ lifts to exactly N 6-cycles in the Tanner graph if $l \neq l'$. The same result follows for the pairs $B_{1,m}, B_{2,m}$ and $C_{1,m}, C_{2,m}$. Moreover, one of these pairs, running over all $m \in [n_v]$, is sufficient to describe all 6-cycles in H. Hence, the total number of 6-cycles in the Tanner graph, $\mathcal{N}_6$, is given by

$$\mathcal{N}_6 = |W(6,1)| \cdot N, \tag{12}$$

and

$$|W(6,1)| = \sum_{m \in [n_v]} \mathcal{R}_{A_{1,m}, A_{2,m}}, \tag{13}$$

where $\mathcal{R}_{A_{1,m}, A_{2,m}}$ is the number of repetitions $\alpha_l = \alpha_{l'}$ between the msets $A_{1,m}$ and $A_{2,m}$.

Example 19. Let H be as in Example 16 and take $N = 5$. Since H has girth 4, we can also use our strategy to calculate 6-cycles in the Tanner graph. Following Theorem 18, $|W(6,1)| = 16$ and we conclude that the number of 6-cycles, $\mathcal{N}_6$, in the Tanner graph is given by $\mathcal{N}_6 = 16 \cdot 5 = 80$. □

Definition 20. *Let W be a walk in the $n_c \times n_v$ fully-connected protograph. If W has length $k = 2m$ and its permutation shift is given by $h_{\alpha_1\beta_1} - h_{\alpha_2\beta_1} + h_{\alpha_3\beta_2} - h_{\alpha_4\beta_2} + \cdots + h_{\alpha_{k-1}\beta_m} - h_{\alpha_k\beta_m}$, then we use the shorthand $[h_{\alpha_1}, h_{\alpha_2}, \ldots, h_{\alpha_k}]_{\beta_1, \beta_2, \ldots, \beta_m}$. If W has length $k = 2m + 1$*

and its permutation shift is given by $h_{\alpha_1\beta_1} - h_{\alpha_2\beta_1} + h_{\alpha_3\beta_2} - h_{\alpha_4\beta_2} + \cdots + h_{\alpha_{k-1}\beta_m} - h_{\alpha_k\beta_m} + h_{\alpha_{k+1}\beta_{m+1}}$*, then we use the shorthand* $[h_{\alpha_1}, h_{\alpha_2}, \ldots, h_{\alpha_k} | h_{\alpha_{k+1}}]_{\beta_1,\beta_2,\ldots,\beta_m|\beta_{m+1}}$. □

By Theorem 2, girth$(H) > 8$ if and only if the permutation shifts

$$[h,i,i,h,h,i,i,h]_{u,v,v',u'}, \qquad [h,j,j,i,i,j,j,h]_{u,v,v',u'},$$
$$[h,i,i,h,h,j,j,h]_{u,v,v',u'}, \qquad [h,i,i,j,j,i,i,h]_{u,v,v',u'},$$
$$[h,j,j,h,h,j,j,h]_{u,v,v',u'}, \qquad [i,j,j,i,i,j,j,i]_{u,v,v',u'},$$

with $u \neq v$, $v \neq v'$, $v' \neq u'$ and $u' \neq u$, are all nonzero. In the following theorem, we rewrite these conditions in a way that is helpful to count 8-cycles in the Tanner graph.

Theorem 21. *Let H be as in* (9) *and, for $u, v \in [n_v]$, $u \neq v$, consider the following msets*

$$A_1 = \{(u,v,[h,i,i,h]_{u,v})\}, \qquad D_1 = \{(u,v,[h,j,j,i]_{u,v})\},$$
$$A_2 = \{(u,v,[h,i,i,h]_{u,v})\}, \qquad D_2 = \{(u,v,[h,j,j,i]_{u,v})\},$$
$$B_1 = \{(u,v,[h,i,i,h]_{u,v})\}, \qquad E_1 = \{(u,v,[h,i,i,j]_{u,v})\},$$
$$B_2 = \{(u,v,[h,j,j,h]_{u,v})\}, \qquad E_2 = \{(u,v,[h,i,i,j]_{u,v})\},$$
$$C_1 = \{(u,v,[h,j,j,h]_{u,v})\}, \qquad F_1 = \{(u,v,[i,j,j,i]_{u,v})\},$$
$$C_2 = \{(u,v,[h,j,j,h]_{u,v})\}, \qquad F_2 = \{(u,v,[i,j,j,i]_{u,v})\}.$$

For $u, v, u', v' \in [n_v]$, let $(u, v, \alpha_{u,v}) \in A_1$ and $(u', v', \alpha_{u',v'}) \in A_2$ be such that $\alpha_{u,v} = \alpha_{u',v'}$. Then this repetition $\alpha_{u,v} = \alpha_{u',v'}$ lifts to a set of 8-cycles in the Tanner graph if $u \neq u'$ and $v \neq v'$. The same result follows for the other five pairs. Moreover, these six pairs are sufficient to describe all 8-cycles. The total number of 8-cycles in the Tanner graph, $\mathcal{N}_8$, is given by

$$\mathcal{N}_8 = |W(4,2)| \cdot N/2 + (|W(8,1)| - |W(4,2)|) \cdot N, \quad (14)$$

and

$$|W(8,1)| = \mathcal{R}^*_{A_1,A_2} + \frac{1}{2}\mathcal{R}_{B_1,B_2} + \mathcal{R}^*_{C_1,C_2} + \frac{1}{2}\mathcal{R}_{D_1,D_2} + \frac{1}{2}\mathcal{R}_{E_1,E_2} + \mathcal{R}^*_{F_1,F_2}, \quad (15)$$

where $\mathcal{R}^*_{X_1,X_2}$ is given by

$$\mathcal{R}^*_{X_1,X_2} = \frac{1}{2}\binom{\text{number of repetitions from } X_1, X_2 \text{ with } u' = v \text{ and } v' = u}{} + \frac{1}{4}\binom{\text{number of repetitions from } X_1, X_2 \text{ otherwise}}{}, \quad (16)$$

$\mathcal{R}_{X_1,X_2}$ is the number of repetitions $\alpha_{u,v} = \alpha_{u',v'}$ between the msets X_1 and X_2, and the coefficient of each $\mathcal{R}_{X_1,X_2}$ is equal to the reciprocal of the number of equivalent walks for the corresponding TBC walk pattern.

Example 22. Let H be the parity-check matrix of the $[155, 64, 20]$ Tanner code given by

$$H = \begin{bmatrix} x & x^2 & x^4 & x^8 & x^{16} \\ x^5 & x^{10} & x^{20} & x^9 & x^{18} \\ x^{25} & x^{19} & x^7 & x^{14} & x^{28} \end{bmatrix}. \quad (17)$$

TABLE I
TIME TAKEN TO COUNT THE NUMBER OF k-CYCLES, $\mathcal{N}_k$, FOR H IN EXAMPLE 22 FOR LIFTING FACTOR N USING OUR APPROACH.

N	k = 4	6	8	10	12
5	84.4 μs	252 μs	–	–	–
10	83.4 μs	244 μs	–	–	–
15	86.6 μs	246 μs	–	–	–
20	86.0 μs	247 μs	2.85 ms	5.01 ms	–
25	86.8 μs	246 μs	2.78 ms	4.96 ms	–
31	85.5 μs	246 μs	2.77 ms	5.18 ms	63.1 ms
50	85.8 μs	244 μs	2.97 ms	5.06 ms	64.2 ms
75	85.1 μs	247 μs	2.57 ms	5.08 ms	64.4 ms
100	84.6 μs	252 μs	2.82 ms	5.02 ms	65.0 ms
125	85.7 μs	246 μs	2.77 ms	4.97 ms	64.3 ms
150	86.7 μs	244 μs	2.97 ms	5.02 ms	64.6 ms
175	83.1 μs	244 μs	2.88 ms	5.03 ms	63.9 ms
200	87.7 μs	247 μs	2.93 ms	5.06 ms	64.1 ms
500	88.0 μs	254 μs	3.03 ms	5.25 ms	65.5 ms
1000	87.5 μs	257 μs	3.04 ms	5.16 ms	65.0 ms

Then H has girth 8 for $N = 31$. Following Theorem 21, we should construct the six pair of msets. Some computations show that $|W(8,1)| = 15$ and we conclude that the number of 8-cycles, $\mathcal{N}_8$, in the Tanner graph is given by $\mathcal{N}_8 = |W(8,1)| \cdot N = 15 \cdot 31 = 465$. □

V. COMPLEXITY

In Theorem 15, to count the number of 4-cycles in the Tanner graph, we need to construct the three msets $\{h_l - i_l \mid l \in [n_v]\}$, $\{h_l - j_l \mid l \in [n_v]\}$ and $\{i_l - j_l \mid l \in [n_v]\}$, and check for repetitions in each one of them. Since each mset has n_v elements, it is sufficient to do $\frac{(n_v-1)n_v}{2}$ comparisons in each one of them. This implies that the complexity of determining $\mathcal{N}_4$ is $O(n_v^2 \log(N))$. Following a similar argument, the complexity of determining $\mathcal{N}_6$ is $O(n_v^2 \log(n_v) \log(N))$ and the complexity of determining $\mathcal{N}_8$ is $O(n_v^4 \log^4(n_v) \log(N))$.

To show how fast we can calculate the number of k-cycles, $\mathcal{N}_k$, in the Tanner graph of a QC-LDPC code, we include Table I. There, we provide the time taken to count the number of k-cycles using our algorithms. The computations were done using SageMath [23] in a MacBook Pro (13-inch, 2018, Four Thunderbolt 3 Ports) with a 2.3 GHz Quad-Core Intel Core i5 processor and 16 GB 2133 MHz LPDDR3 of memory.

VI. CONCLUDING REMARKS

This paper discusses an efficient strategy to count cycles in the Tanner graph of arbitrary QC-LDPC codes. We use some results on graph covers involving the images of cycles in the Tanner graph and the inverse images of tailless backtrackless closed walks in the protograph to provide closed formulas for the number of k-cycles, $\mathcal{N}_k$, by just taking into account repetitions in some msets constructed from the matrices $B_m(H)$. Our strategy has been shown to reduce the complexity of determining $\mathcal{N}_k$, giving our approach a significant advantage over previous works on the cycle distribution of QC-LDPC codes.

REFERENCES

[1] R. G. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, pp. 21–28, 1962.

[2] CCSDS, "TC Synchronization and Channel Coding. Issue 4. Recommendation for Space Data System Standards (Blue Book), CCSDS 231.0-B-4. Washington, D.C.: CCSDS," Jul. 2021.

[3] Z. Li, L. Chen, L. Zeng, S. Lin, and W. H. Fong, "Efficient encoding of quasi-cyclic low-density parity-check codes," *IEEE Transactions on Communications*, vol. 54, no. 1, pp. 71–81, Jan. 2006.

[4] Z. Wang and Z. Cui, "A memory efficient partially parallel decoder architecture for quasi-cyclic LDPC codes," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 15, no. 4, pp. 483–488, Apr. 2007.

[5] Y. Mao and A. H. Banihashemi, "A heuristic search for good low-density parity-check codes at short block lengths," *IEEE International Conference on Communications Proceedings*, vol. 1, pp. 41–44, Jun. 2001.

[6] T. R. Halford and K. M. Chugg, "An algorithm for counting short cycles in bipartite graphs," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 287–292, Jan. 2006.

[7] M. Karimi and A. H. Banihashemi, "Efficient algorithm for finding dominant trapping sets of LDPC codes," *IEEE Transactions on Information Theory*, vol. 58, no. 11, pp. 6942–6958, Nov. 2012.

[8] J. Flum and M. Grohe, "The parameterized complexity of counting problems," *IEEE Symposium on Foundations of Computer Science Proceedings*, pp. 538–547, 2002.

[9] ——, "The parameterized complexity of counting problems," *SIAM Journal on Computing*, vol. 33, no. 4, pp. 892–922, 2004.

[10] M. Karimi and A. H. Banihashemi, "Message-passing algorithms for counting short cycles in a graph," *IEEE Transactions on Communications*, vol. 61, no. 2, pp. 485–495, Feb. 2013.

[11] H. M. Stark and A. A. Terras, "Zeta functions of finite graphs and coverings," *Advances in Mathematics*, vol. 121, pp. 124–165, 1996.

[12] M. Karimi and A. H. Banihashemi, "Counting short cycles of quasi cyclic protograph LDPC codes," *IEEE Communications Letters*, vol. 16, no. 3, pp. 400–403, Mar. 2012.

[13] G. J. Tee, "Eigenvectors of block circulant and alternating circulant matrices," *Research Letters in the Information and Mathematical Sciences*, vol. 8, pp. 123–142, 2005.

[14] J. Thorpe, "Low-density parity-check (LDPC) codes constructed from protographs," *Jet Propulsion Laboratory Pasadena, CA, INP Progress Report 42-154*, pp. 42–154, Aug. 2003.

[15] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 533–547, Sept. 1981.

[16] X. Wu, X. You, and C. Zhao, "A necessary and sufficient condition for determining the girth of quasi-cyclic LDPC codes," *IEEE Transactions on Communications*, vol. 56, no. 6, pp. 854–857, June 2008.

[17] J. McGowan and R. Williamson, "Loop removal from LDPC codes," *IEEE Information Theory Workshop Proceedings*, pp. 230–233, 2003.

[18] J. L. Gross and T. W. Tucker, *Topological Graph Theory*. New York: Wiley, 1987.

[19] R. Asvadi, A. H. Banihashemi, and M. Ahmadian-Attari, "Design of irregular quasi-cyclic protograph codes with low error floors," *IEEE International Symposium on Information Theory Proceedings*, pp. 908–912, 2011.

[20] ——, "Lowering the error floor of LDPC codes using cyclic liftings," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2213–2224, Apr. 2011.

[21] C. A. Kelley, "On codes designed via algebraic lifts of graphs," *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1254–1261, 2008.

[22] R. Smarandache and D. G. M. Mitchell, "Necessary and sufficient girth conditions for Tanner graphs of quasi-cyclic LDPC codes," *IEEE International Symposium on Information Theory Proceedings*, pp. 380–385, July 2021.

[23] The Sage Developers, "SageMath, the Sage Mathematics Software System," 2018. [Online]. Available: https://www.sagemath.org