Quasi-Cyclic LDPC Codes with Parity-Check Matrices of Column Weight Two or More for Correcting Phased Bursts of Erasures

Xin Xiao, Bane Vasić, *Fellow, IEEE,* Shu Lin, *Life Fellow, IEEE,* Juane Li, and Khaled Abdel-Ghaffar, *Senior Member, IEEE*

***Abstract*—In his pioneering work on LDPC codes, Gallager dismissed codes with parity-check matrices of weight two after proving that their minimum Hamming distances grow at most logarithmically with their code lengths. In spite of their poor minimum Hamming distances, it is shown that quasi-cyclic LDPC codes with parity-check matrices of column weight two have good capability to correct phased bursts of erasures which may not be surpassed by using quasi-cyclic LDPC codes with parity-check matrices of column weight three or more. By modifying the parity-check matrices of column weight two and globally coupling them, the erasure correcting capability can be further enhanced. Quasi-cyclic LDPC codes with parity-check matrices of column weight three or more that can correct phased bursts of erasures and perform well over the AWGN channel are also considered. Examples of such codes based on Reed-Solomon and Gabidulin codes are presented.**

***Index Terms*—Erasure correction, Gabidulin code, global coupling, Golomb ruler, LDPC code, phased burst, quasi-cyclic code, Reed-Solomon code.**

I. Introduction

With the rediscovery of low-density parity-check (LDPC) codes by the turn of the century, researchers have recognized that LDPC codes perform well over the binary erasure channel (BEC) that causes the value of a transmitted bit to be lost, in addition to their superior performance over the AWGN channel [1]. A simple "peeling" algorithm that can be applied to a sparse parity-check matrix of an LDPC code to correct erasures was proposed early on. The algorithm may not correct all erasures that can be corrected by an optimal maximum-likelihood (ML) decoder. However, for long LDPC codes, it is very difficult to determine the capability of an ML decoder to correct erasures let alone implement such a decoder. Motivated by potential applications of LDPC codes in storage systems and communication over fading channels, researchers investigated the capability of LDPC codes to correct erasure bursts, and in particular one long burst of erasures [2]–[6].

The work of B. Vasić is funded in part by the NSF under grants SaTC-1813401, CIF-1855879, and ECCS/CCSS-2027844. The material in this paper was presented in part at the Information Theory and Applications (ITA) Workshop, 2019.

X. Xiao and B. Vasić are with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85712, USA (e-mail: 7xinxiao7@email.arizona.edu; vasic@email.arizona.edu). J. Li is with Micron Technology Inc., San Jose, CA 95131, USA (e-mail: jueli@ucdavis.edu). S. Lin and K. Abdel-Ghaffar are with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA (e-mail: shulin@ucdavis.edu; ghaffar@ucdavis.edu).

In this paper we consider binary quasi-cyclic (QC) LDPC codes, with parity-check matrices which are $m \times n$ arrays of circulant permutation matrices (CPMs) of size $t \times t$. These are the most widely known, studied, and used QC-LDPC codes. A codeword $\mathbf{v}$ in such a code can be written as $\mathbf{v} = (\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{n-1})$, where $\mathbf{v}_j$, $0 \leq j < n$, is a sequence of t bits which we call a *section*. We assume that such a codeword is transmitted over a channel that causes multiple *phased bursts* of erasures. By a phased burst of erasures we mean that all the erasures affect one and only one of the sections $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$. Each section may correspond, for example, to a part of a large file stored at a node in a distributed storage system. Losses in part of a file stored at a node can then be modeled as a phased burst of erasures in which all the erasures are confined to a section.

First, we notice that no QC-LDPC code with a parity-check matrix composed of CPMs can correct two "solid" phased bursts, i.e., all bits in a section are erased. Therefore, the best we can hope for is to correct pairs of mutually "semi-solid" phased bursts of erasures in which all the bits in the two phased bursts are erased except for one. We show that a QC-LDPC code with parity-check matrix of column weight two, i.e., composed of just two row blocks of CPMs, if properly designed, can correct any pair of such phased bursts. Ge and Xia call such a parity-check matrix *ultra sparse* [7]. We demonstrate that codes with parity-check matrices of column weight two which can correct two phased bursts except for one bit have the highest possible dimension among all codes with this correction capability. We also prove that the peeling algorithm when applied to such parity-check matrices of column weight two can correct all erasures that can be corrected by an ML decoder. This means that all stopping sets of such parity-check matrices are nonzero codewords. Basically, our analysis of QC-LDPC codes with parity-check matrices of column weight two is rather comprehensive as we determine, for all such codes, their dimensions, minimum Hamming distances, and their capabilities to correct phased bursts. We also show how to add extra rows to such parity-check matrices in order to correct any two solid phased bursts, without the exception of one bit. We also present a technique to globally couple the codes in order to correct long phased bursts of erasures. Since QC-LDPC codes with parity-check matrices of column weight two have poor performance over AWGN channels, and practical channels rarely only cause erasures, we propose methods for constructing parity-check matrices with

column weight three or more which are natural extensions of the parity-check matrices with column weight two. The constructions are related to Reed-Solomon and Gabidulin codes [8],[9].

It should be noted that correcting erasures can be accomplished using other codes. For example, a Hamming code can correct two erasures. To construct a code of length $N = nt$ to correct two phased bursts of erasures of length t, we can interleave t copies of a (possibly shortened or lengthened) Hamming code of length n. As each copy has about $\log_2 n$ redundant bits, the total redundancy in the interleaved code is about $t \log_2 n$. We present codes that have the same erasure correcting capability but with about $2t$ redundant bits regardless of the value of n as long as $n \leq t$. In the practical applications we envision, the values of n and t are large, and our codes are more efficient than schemes using interleaving. If $n \leq 2^t$, then a (possibly shortened or lengthened) Reed-Solomon code of length n and dimension $n-2$ over GF(2^t), in which each symbol represents a section as a binary vector of length t, can be used to correct two phased bursts of erasures. The number of redundant bits is $2t$ which is comparable to the number of redundant bits in our codes. However, correcting erasures using such Reed-Solomon code requires computations over GF(2^t). For values of t in the hundreds, this may not be feasible. On the other hand, the peeling algorithm applied to our codes requires only simple computations over GF(2) regardless of the value of t.

This paper is organized as follows. The notation for burst erasures, QC-LDPC codes, and their parity-check matrices with some basic results are presented in Section II. Section III covers QC-LDPC codes with parity-check matrices of column weight two and Section IV extends this to column weights more than two. The paper is concluded in Section V. For smooth reading, all proofs are relegated to appendices.

II. Preliminaries

A. Correcting Bursts of Erasures

We consider transmission over a binary erasure channel (BEC) in which a transmitted bit is either received correctly or erased. The decoder knows exactly the set of indices, $\mathcal{J}$, of the erased bits. To be able to recover the values of the erased bits, a binary linear code is used. An (N, K) binary linear code is the K-dimensional null space of an $M \times N$ binary matrix $\mathbf{H}$, for some integer $M \geq N - K$. This matrix is a *parity-check matrix* for the code, the rank of which is $\mathrm{rank}(\mathbf{H}) = N - K$, which we call the *redundancy* of the code. For any codeword $\mathbf{v}$, we have $\mathbf{v}\mathbf{H}^{\mathsf{T}} = \mathbf{0}$ where computations are over GF(2), T denotes transpose, and $\mathbf{0}$ is the all-zero M-tuple. Suppose a codeword is transmitted over the channel and e erasures occur in the bits indexed by $\mathcal{J}$. Then, a maximum-likelihood (ML) decoder [10],[11] can recover the erased bits if and only if the code does not have any nonzero codeword in which the indices of all the 1's are confined to $\mathcal{J}$. In this case, we say that the erasures are *recoverable* by the ML decoder. By considering the values of the erased bits to be unknowns in the codeword $\mathbf{v}$, these unknowns can be determined from $\mathbf{v}\mathbf{H}^{\mathsf{T}} = \mathbf{0}$ which is a system of M parity equations. A necessary condition for this system to be solvable is that $N - K \geq e$. Codes meeting this bound with equality are said to be *optimal* for correcting the erasures specified by $\mathcal{J}$. Although a code in general has many parity-check matrices, its ability to correct erasures does not depend on the choice of $\mathbf{H}$ to solve for the unknowns in the equation $\mathbf{v}\mathbf{H}^{\mathsf{T}} = \mathbf{0}$. However, if e is large, say in the hundreds, then solving this system of equations may be computationally intensive.

In 2001, Luby *et al.* [1] came up with a simple decoding algorithm to correct erasures. The algorithm is applied to a particular parity-check matrix of the code and its success depends on this matrix. Although the algorithm may not be able to recover all erasures recoverable by the ML decoder, it is quite simple as it allows the recovery of the erased bits one by one. The peeling algorithm works as follows. If there is a parity equation that checks only one unknown erasure, then the erased value can be determined from that equation by an XOR operation and the number of unknowns is then reduced by one. Next, if another parity equation is found that checks only one of the remaining unknowns, then that unknown can be determined and the number of unknowns is further reduced by one. This may continue until all erasures are recovered or until no equation is found that checks only one unknown erasure in which case decoding fails. The set of erased positions at this stage forms a *stopping set* [10]. Although there is no universal term to identify this algorithm in the coding literature, some researchers call it figuratively the *peeling algorithm* [12], a term which we will adopt. The peeling algorithm was initially developed for randomly constructed LDPC codes and applied to their sparse parity-check matrices. The randomness makes it hard to develop erasure decoding algorithms that exploit the structure of the codes. On the other hand, the sparseness helps in having parity equations involving a small number of terms for which the peeling algorithm is most effective.

The peeling algorithm is best understood in terms of the *Tanner graph*, $\mathcal{G}$, representing the parity-check matrix $\mathbf{H} = [h_{I,J}]_{0 \leq I < M, 0 \leq J < N}$ [8],[13],[14]. This is a bipartite graph in which the set of vertices is partitioned into a set of *variable nodes* indexed by the columns of $\mathbf{H}$ and a set of *check nodes* indexed by the rows of $\mathbf{H}$. Edges connect only variable nodes to check nodes. In particular, there is an edge connecting the variable node corresponding to the J-th column to the check node corresponding to the I-th row if and only if $h_{I,J} = 1$. Since the code is the null space of $\mathbf{H}$, if the variable nodes assume the bit values of a codeword, then the sum of the values of the variable nodes adjacent to each check node is even. The peeling algorithm looks for a check node which is adjacent to only one erased variable node and determines its value as the sum over GF(2), i.e., XOR, of the values of all other variable nodes adjacent to the check node. The number of erasures is then reduced by one and the process is repeated until all erased bits are recovered, in which case decoding is successful, or there is no check node that checks exactly one erased variable node, in which case decoding fails as the remaining variable nodes form a stopping set. The success of the peeling algorithm depends on the parity-check matrix used and its associated Tanner graph. We say that a parity-check matrix is *ML peeling-decodable* if every recoverable

set of erasures by an ML decoder can also be recovered by the peeling algorithm.

Constructions of ML peeling-decodable parity-check matrices for an (N, K) linear code are presented in [15]–[17] where the number of rows of the constructed matrices is exponential in $N-K$. For such matrices, the peeling algorithm may cease to be appealing if $N-K$ is large. As a motivation of our investigation of codes with parity-check matrices of column weight two we give the following result, the proof of which is presented in Appendix A.

Theorem 1. *Let* $\mathbf{H}$ *be a parity-check matrix of a linear code in which each column has weight at most two. Then,* $\mathbf{H}$ *is ML peeling-decodable.*

Let $\mathbf{H} = [h_{I,J}]_{0\le I<M, 0\le J<N}$ be a binary matrix. We say that $\mathbf{H}$ satisfies the *row-column (RC) constraint* [8] if there are no four 1's in the positions specified by a pair of distinct rows and a pair of distinct columns, i.e., for any $0 \le I_0 < I_1 < M$ and $0 \le J_0 < J_1 < N$, at least one of the elements $h_{I_0,J_0}, h_{I_0,J_1}, h_{I_1,J_0}, h_{I_1,J_1}$ is zero. In this case, the girth of the Tanner graph $\mathcal{G}$ representing $\mathbf{H}$, which is the shortest length of a cycle in $\mathcal{G}$, is at least 6. However, there is a more important consequence to the RC-constraint. Suppose that the code with $\mathbf{H}$ as a parity-check matrix is used over a channel causing erasures. If the number of erasures, e, is at most equal to the minimum weight $w_{\min}$ of a column in $\mathbf{H}$, then not only the code can recover the erasures but it can do so by applying the computationally simple peeling algorithm to $\mathbf{H}$. Indeed, an erasure is checked by at least $w_{\min}$ parity equations and, because of the RC-constraint, each of the other $e-1 < w_{\min}$ erasures is checked by at most one of these parity equations. Hence, there is a parity equation that checks only that erasure and no other from which the erasure can be recovered. The procedure is repeated until all erasures are recovered. In particular, if $\mathbf{H}$ satisfies the RC-constraint, then it is a parity-check matrix of a code with minimum Hamming distance at least $w_{\min}+1$.

B. QC-LDPC Codes and Their Parity-Check Matrices

Throughout this paper, we use $(x)_t$ for an integer x and a positive integer t to denote the least nonnegative integer congruent to x modulo t, i.e., $(x)_t = x - \lfloor x/t \rfloor t$. All indices of vectors and of rows and columns of matrices are numbered starting with 0.

By an $m\times n$ array $\mathbf{H} = [\mathbf{H}_{i,j}]_{0\le i<m, 0\le j<n}$ of $t\times t$ matrices $\mathbf{H}_{i,j}$ we mean the $mt \times nt$ matrix in which the (I, J) entry in $\mathbf{H}$, $0 \le I < mt$, $0 \le J < nt$, is the (i', j') entry in $\mathbf{H}_{i,j}$ where $i' = (I)_t$, $j' = (J)_t$, $i = \lfloor I/t \rfloor$, and $j = \lfloor J/t \rfloor$. In general, we use (I, J), $0 \le I < mt, 0 \le J < nt$, to denote indices of entries in the $mt \times nt$ matrix $\mathbf{H}$, (i', j'), $0 \le i', j' < t$, to denote indices of entries in a $t\times t$ submatrix, and (i, j), $0 \le i < m, 0 \le j < n$, to denote the indices of the submatrix within the array $\mathbf{H}$. For $0 \le i < m$, the $t \times nt$ submatrix $[\mathbf{H}_{i,0}, \mathbf{H}_{i,1}, \mathbf{H}_{i,1}, \ldots, \mathbf{H}_{i,n-1}]$ is called the i-th *row block* of $\mathbf{H}$ and for $0 \le j < n$, the $mt \times n$ matrix $[\mathbf{H}_{0,j}^{\mathsf{T}}, \mathbf{H}_{1,j}^{\mathsf{T}}, \ldots, \mathbf{H}_{m-1,j}^{\mathsf{T}}]^{\mathsf{T}}$ is called the j-th *column block* of $\mathbf{H}$. For $0 \le i < n$ and $0 \le i' < t$, a row in $\mathbf{H}$ is indexed by $(i; i')$ if it is the i'-th row in the i-th row block. Thus, a row in $\mathbf{H}$ can be indexed by I for some I, $0 \le I < mt$, or by the pair $(i; i')$, $0 \le i < m$, $0 \le i' < t$, where $i' = (I)_t$, $i = \lfloor I/t \rfloor$, and $I = it + i'$. Similarly, a column in $\mathbf{H}$ can be indexed by J for some J, $0 \le J < nt$, or by the pair $(j; j')$, $0 \le j < n$, $0 \le j' < t$, where $j' = (J)_t$, $j = \lfloor J/t \rfloor$, and $J = jt + j'$, indicating the j'-th column in the j-th column block.

A *circulant* is a square matrix in which every row other than the top row is the cyclic shift of the row above it by one position to the right. It follows that the top row is also the cyclic shift of the bottom row. Hence, a circulant is completely characterized by its top row. In particular, the square zero matrix (ZM) is a circulant. A binary $t \times t$ matrix is called a *circulant permutation matrix* (CPM) if its top row has weight one. A CPM in which the single 1 in its top row is in position p, $0 \le p < t$, is denoted by $\mathbf{CPM}_t(p)$[1]. Notice that all the entries in $\mathbf{CPM}_t(p)$ are zeros except those in positions $(i', (i'+p)_t)$ for $0 \le i' < t$, i.e., positions $((j'-p)_t, j')$ for $0 \le j' < t$. Suppose that $\mathbf{H}$ is an array of $m\times n$ of $t\times t$ CPM's, i.e., $\mathbf{H} = [\mathbf{CPM}_t(p_{i,j})]_{0\le i<m, 0\le j<n}$. Then each column in $\mathbf{H}$ has weight m and each row has weight n. To capture the parameters of $\mathbf{H}$ we denote it by $\mathbf{H}_{m,n,t}$ and reserve this notation for arrays composed exclusively of CPMs without any ZMs. A necessary and sufficient condition for $\mathbf{H}$ to satisfy the RC-constraint is given in the following proposition which follows as a special case of [18, Theorem 2.1].

Proposition 1. *The matrix* $\mathbf{H}_{m,n,t} = [\mathbf{CPM}_t(p_{i,j})]_{0\le i<m, 0\le j<n}$ *satisfies the RC-constraint if and only if* $p_{i_1,j_1} - p_{i_0,j_1} - p_{i_1,j_0} + p_{i_0,j_0}$ *is not divisible by* t *for* $0 \le i_0 < i_1 < m$, $0 \le j_0 < j_1 < n$.

A code is *quasi-cyclic* (QC) [8],[14] if it is the null space of an array of circulants of equal size. In particular, if $\mathbf{H}_{m,n,t} = [\mathbf{CPM}_t(p_{i,j})]_{0\le i<m, 0\le j<n}$, then it is a parity-check matrix of a QC code, $C_{m,n,t}$, of length nt and dimension $nt - \mathrm{rank}(\mathbf{H}_{m,n,t})$. Assuming that t is not small, then $\mathbf{H}_{m,n,t}$ is sparse and the code $C_{m,n,t}$ is a QC-LDPC code.

The composition of the parity-check matrix $\mathbf{H}_{m,n,t}$ as an array of circulants naturally defines a sectionalized structure for codewords. A binary sequence $\mathbf{v} = (v_0, v_1, \ldots, v_{nt-1})$ composed of nt bits can be written as $\mathbf{v} = (\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{n-1})$, where $\mathbf{v}_i = (v_{it}, v_{it+1}, \ldots, v_{it+t-1})$ forms a section. Erasures affecting only one section of the transmitted codeword form a phased burst. Thus, a phased burst may contain up to t erasures. If the number of erasures in a phased burst is t, then we say that the phased burst is *solid*. We say that two phased bursts affecting two sections are *mutually semi-solid* if the total number of erasures is $2t-1$, i.e., one phased burst is solid and the other contains $t-1$ erasures.

For the code $C_{m,n,t}$, let $e(r)$ be the maximum number such that any $e(r)$ erasures confined to any r sections are correctable and $e_{\mathrm{adj}}(r)$ be the maximum number such that any $e_{\mathrm{adj}}(r)$ erasures confined to any r adjacent sections, i.e, sections $j, j+1, \ldots, j+r-1$ for some integer j, $0 \le j \le n-r$, are correctable. Clearly, $e(r) \le e_{\mathrm{adj}}(r)$ for

[1]More commonly denoted by $I(p)$ or P^p, see, e.g., [18] and [19].

$1 \le r \le n$. Furthermore, $e(1) = e_{\rm adj}(1) = t$ as the columns in any column block are linearly independent. We also have $e(n) = e_{\rm adj}(n) = d - 1$, where d is the minimum Hamming distance of the code.

By circularly shifting the columns in each column block and the rows in each row block of $\mathbf{H}_{m,n,t} = [\mathbf{CPM}_t(p_{i,j})]_{0\le i<m, 0\le j<n}$, we can put $\mathbf{H}_{m,n,t}$ in a form of an $m \times n$ array of CPMs in which the 0-th row block and the 0-th column block consist only of $t \times t$ identity matrices $\mathbf{CPM}_t(0)$. These shifting operations do not change the rank of the matrix $\mathbf{H}_{m,n,t}$ and, being confined to columns in the same column block, do not change the capability of $C_{m,n,t}$ to correct phased bursts using ML decoding or the peeling algorithm. Therefore, from now on, we only consider matrices $\mathbf{H}_{m,n,t}$ in this form.

As each column block is composed of CPMs, the columns in any two column blocks are linearly dependent as their sum is the all-zero vector. This implies that $e(r) \le e_{\rm adj}(r) \le 2t-1$ for all $r \ge 2$. For $m = 1$, $e_{\rm adj}(2) = 1$ as $\mathbf{H}_{1,n,t}$ is just a row of CPMs and, therefore, there are two identical columns in any two distinct column blocks. To have $e_{\rm adj}(2) > 1$, m should be at least two. We will show that with proper choice of the CPMs, the upper bound $2t-1$ on the number of erasures that can be corrected in a pair of phased bursts can be attained for $m = 2$. Since the dimension of code may decrease by increasing m, it is interesting to consider the case $m = 2$ which is treated in the next section.

III. QC Codes with Parity-Check Matrices of Column Weight Two

A. Correcting Pairs of Semi-Solid Phased Bursts of Erasures

With $m = 2$, we consider a parity-check matrix, $\mathbf{H}_{2,n,t}$, in the form of

$$\begin{bmatrix} \mathbf{CPM}_t(0) & \mathbf{CPM}_t(0) & \cdots & \mathbf{CPM}_t(0) \\ \mathbf{CPM}_t(p_0) & \mathbf{CPM}_t(p_1) & \cdots & \mathbf{CPM}_t(p_{n-1}) \end{bmatrix}, \quad (1)$$

where $p_0 = 0$ and $n \ge 2$. For convenience, we call the two row blocks in $\mathbf{H}_{2,n,t}$ the top row block and the bottom row block. Then $\mathbf{H}_{2,n,t}$ is a parity-check matrix of a QC-LDPC code, $C_{2,n,t}$, of length nt and dimension $nt-{\rm rank}(\mathbf{H}_{2,n,t})$. The rank of $\mathbf{H}_{2,n,t}$, which equals the redundancy of $C_{2,n,t}$, is given in the following theorem in which GCD stands for the greatest common divisor. The proof is presented in Appendix B.

Theorem 2. *For the matrix $\mathbf{H}_{2,n,t}$ in (1),* ${\rm rank}(\mathbf{H}_{2,n,t}) = 2t - {\rm GCD}(p_1, p_2, \ldots, p_{n-1}, t)$.

The following theorem, the proof of which is given in Appendix C, specifies the phased-burst erasure correcting capabilities of the code $C_{2,n,t}$ with the parity-check matrix $\mathbf{H}_{2,n,t}$. First we say that a collection of integers $p_0, p_1, \ldots, p_{n-1}$ forms a *t-modular Golomb ruler* [20, Section 19.3] if $(p_i - p_j)_t$ are nonzero and distinct for distinct ordered pairs (i, j), $0 \le i \ne j < n$. This means that for every positive integer less than t, there is at most one pair of i and j such that $(p_i - p_j)_t$ equals this integer[2]. The integers $p_0, p_1, \ldots, p_{n-1}$ are called the *markers* of the ruler.

We say that the parity-check matrix $\mathbf{H}_{2,n,t}$ in (1) has the *distinct property* if p_j, $0 \le j < n$, are distinct. We also say that $\mathbf{H}_{2,n,t}$ has the *modular Golomb ruler property* if the numbers p_j, $0 \le j < n$, form a t-modular Golomb ruler. In particular, if $\mathbf{H}_{2,n,t}$ has the modular Golomb ruler property, then it also has the distinct property. For example, the parity-check matrix $\mathbf{H}_{2,3,7}$ for which $p_0 = 0$, $p_1 = 1$, and $p_2 = 2$ has the distinct property since p_0, p_1, and p_2 are distinct, but not the modular Golomb ruler property since $(p_1 - p_0)_7 = (p_2 - p_1)_7$ as both equal 1. On the other hand, the parity-check matrix $\mathbf{H}_{2,3,7}$ for which $p_0 = 0, p_1 = 1$, and $p_2 = 3$ has the modular Golomb ruler property since $(p_1 - p_0)_7 = 1$, $(p_0 - p_1)_7 = 6$, $(p_2 - p_0)_7 = 3$, $(p_0 - p_2)_7 = 4$, $(p_2 - p_1)_7 = 2$, and $(p_1 - p_2)_7 = 5$, i.e., $(p_i - p_j)_7$ are nonzero and distinct for distinct ordered pairs (i, j), $0 \le i \ne j < 3$. Clearly, $t \ge n$ is a necessary condition for $\mathbf{H}_{2,n,t}$ to have the distinct property. Also, $t \ge n^2 - n + 1$ is a necessary condition for $\mathbf{H}_{2,n,t}$ to have the modular Golomb ruler property. For $1 \le r \le n$, we say that $\mathbf{H}_{2,n,t}$ has the *r-adjacent distinct property* or the *r-adjacent modular Golomb ruler property* if the corresponding property holds for any submatrix of $\mathbf{H}_{2,n,t}$ composed of r consecutive column blocks.

Theorem 3. *For the code $C_{2,n,t}$ with the parity-check matrix $\mathbf{H}_{2,n,t}$ in (1), we have $e(1) = e_{\rm adj}(1) = t$,*

$$e(2) = \frac{2t}{\max\limits_{0\le j_0<j_1<n} {\rm GCD}(p_{j_1} - p_{j_0}, t)} - 1$$

$$e_{\rm adj}(2) = \frac{2t}{\max\limits_{0\le j_0<n-1} {\rm GCD}(p_{j_0+1} - p_{j_0}, t)} - 1$$

$$e(r) = \begin{cases} 1, & \text{if } \mathbf{H}_{2,n,t} \text{ does not have the distinct property} \\ 3, & \text{if } \mathbf{H}_{2,n,t} \text{ has the distinct property but not the modular Golomb ruler property} \\ 5, & \text{if } \mathbf{H}_{2,n,t} \text{ has the modular Golomb ruler property,} \end{cases}$$

$$e_{\rm adj}(r) = \begin{cases} 1, & \text{if } \mathbf{H}_{2,n,t} \text{ does not have the } r\text{-adjacent distinct property} \\ 3, & \text{if } \mathbf{H}_{2,n,t} \text{ has the } r\text{-adjacent distinct property but not the } r\text{-adjacent modular Golomb ruler property} \\ 5, & \text{if } \mathbf{H}_{2,n,t} \text{ has the } r\text{-adjacent modular Golomb ruler property,} \end{cases}$$

for $r \ge 3$. In particular, if t is a prime and $\mathbf{H}_{2,n,t}$ has the distinct property, then $C_{2,n,t}$ can correct any two mutually semi-solid phased bursts of erasures regardless of whether or not they are adjacent.

Since the minimum Hamming distance of the code is $d = e(n) + 1$, we have the following corollary to Theorem 3.

[2]In case $(p_i - p_j)_t$ is replaced by $(p_i + p_j)_t$, the sequence is a modular Sidon sequence [22] while if the difference sign is kept but "at most" is replaced by "exactly", the modular Golomb ruler is a perfect difference set [20, Section 19.3]. These combinatorial objects and variations thereof were used in numerous papers, e.g., [19],[23]–[26], to construct LDPC codes with Tanner graphs of large girths.

Corollary 1. *The minimum Hamming distance, d, of the code $C_{2,n,t}$ with the parity-check matrix $\mathbf{H}_{2,n,t}$ in (1), where $n \geq 3$, is*

$$d=\begin{cases} 2, & \text{if } \mathbf{H}_{2,n,t} \text{ does not have the distinct property} \\ 4, & \text{if } \mathbf{H}_{2,n,t} \text{ has the distinct property but not the} \\ & \quad \text{modular Golomb ruler property} \\ 6, & \text{if } \mathbf{H}_{2,n,t} \text{ has the modular Golomb ruler property.} \end{cases}$$

It is worth mentioning that Gallager [27, Theorem 2.5] has shown that the minimum Hamming distances of codes, with parity-check matrices in which each column has weight two, grow at most logarithmically with the code length. A result by MacKay and Davey [28, Theorem 2] implies, as a special case, that the minimum Hamming distance of QC-LDPC codes with parity-check matrices of the form $\mathbf{H}_{2,n,t}$ in (1) is at most 6. Corollary 1 specifies exactly the minimum Hamming distances for such codes. In spite of the poor minimum distance, Theorem 1 implies that $\mathbf{H}_{2,n,t}$ is ML peeling-decodable. In particular, all erasures recoverable by the ML decoder, and not only those limited in number by the minimum Hamming distance, are also correctable by the peeling algorithm. We also notice from the proofs in Appendix C that the girth of the Tanner graph representing $\mathbf{H}_{2,n,t}$ is twice the minimum Hamming distance, i.e., it is 4, 8, or 12 as observed earlier by Fossorier [18, Corollary 2.1]. We should also mention here that Chen, Bai, and Wang have shown that the girth is 12 if and only if $\mathbf{H}_{2,n,t}$ has the modular Golomb ruler property [23].

Although $C_{2,n,t}$ has poor correcting capability if the erasures are in three or more sections, it may correct large number of erasures confined to two sections. As mentioned earlier, a linear (N, K) code is optimal for correcting some erasures if these erasures are correctable by the code and the redundancy, $N - K$, equals the number of erasures. By combining Theorems 2 and 3, it follows that if $e(2) = 2t-1$, then $\mathrm{rank}(\mathbf{H}_{2,n,t}) = 2t-1$ and the code $C_{2,n,t}$ is optimal for correcting two mutually semi-solid phased bursts of erasures.

In the following, we give two constructions of general classes of codes with parity-check matrices as given in (1) by specifying the parameters $p_0 = 0, p_1, \ldots, p_{n-1}$. The two classes of codes are denoted by $C^{\mathrm{RS}}_{2,n,t}$ and $C^{\mathrm{Gabidulin}}_{2,n,t}$. The superscripts RS and $\mathrm{Gabidulin}$ refer to Reed-Solomon and Gabidulin codes, respectively. The parameters $p_0 = 0, p_1, \ldots, p_{n-1}$ are the exponents, modulo t, of elements in a finite field used to define parity-check matrices of these codes. This will be elaborated upon later after generalizing the constructions in Examples 3 and 4.

Example 1. Let $\mathbf{H}^{\mathrm{RS}}_{2,n,t}$ be the parity-check matrix given in (1) in which $t \geq n \geq 3$ and $p_j = j$ for $0 \leq j < n$. From Theorem 2, we have $\mathrm{rank}(\mathbf{H}^{\mathrm{RS}}_{2,n,t}) = 2t - \mathrm{GCD}(1, 2, \ldots, n-1, t) = 2t - 1$. Notice that $\mathbf{H}^{\mathrm{RS}}_{2,n,t}$ has the distinct property but not the modular Golomb ruler property or the r-adjacent modular Golomb ruler property for any $r \geq 3$. Indeed, for the pairs $(p_0, p_1) = (0, 1)$ and $(p_1, p_2) = (1, 2)$, we have $(p_1 - p_0)_t = (p_2 - p_1)_t$. Hence, from Theorem 3, we have $e(1) = e_{\mathrm{adj}}(1) = t$,

$$\begin{aligned} e(2) &= \frac{2t}{\max\limits_{0 \leq j_0 < j_1 < n} \mathrm{GCD}(j_1 - j_0, t)} - 1 = \frac{2t}{t_n} - 1 \\ e_{\mathrm{adj}}(2) &= \frac{2t}{\max\limits_{0 \leq j_0 < n-1} \mathrm{GCD}((j_0+1) - j_0, t)} - 1 = 2t - 1, \end{aligned}$$

and $e(r) = e_{\mathrm{adj}}(r) = 3$ for $r \geq 3$ where t_n is the largest factor of t less than n. The null space of $\mathbf{H}^{\mathrm{RS}}_{2,n,t}$ is a QC-LDPC code which we denote by $C^{\mathrm{RS}}_{2,n,t}$. This code has minimum Hamming distance of four. It can correct any pair of adjacent mutually semi-solid phased bursts of erasures and it is optimal for correcting these erasures. If t is a prime, then $t_n = 1$ and the code can also correct any pair of mutually semi-solid phased bursts of erasures and, in this case, it is also optimal for correcting these erasures. ■

Example 2. Let $\mathbf{H}^{\mathrm{Gabidulin}}_{2,n,t}$ be the parity-check matrix given in (1) in which $t \geq n \geq 3$ and $p_j = (q^j - 1)_t$ for $0 \leq j < n$, $q \geq 2$ is an integer, and t is relatively prime to q and $q - 1$. From Theorem 2, we have $\mathrm{rank}(\mathbf{H}_{2,n,t}) = 2t - \mathrm{GCD}(q - 1, q^2 - 1, \ldots, q^{n-1} - 1, t) = 2t - 1$ and from Theorem 3, we have $e(1) = e_{\mathrm{adj}}(1) = t$,

$$\begin{aligned} e(2) &= \frac{2t}{\max\limits_{0 \leq j_0 < j_1 < n} \mathrm{GCD}(q^{j_1} - q^{j_0}, t)} - 1 \\ &= \frac{2t}{\max\limits_{1 \leq j < n} \mathrm{GCD}(q^j - 1, t)} - 1 \\ e_{\mathrm{adj}}(2) &= \frac{2t}{\max\limits_{0 \leq j_0 < n-1} \mathrm{GCD}(q^{j_0+1} - q^{j_0}, t)} - 1 \\ &= 2t - 1. \end{aligned}$$

For $\mathbf{H}^{\mathrm{Gabidulin}}_{2,n,t}$ to have the distinct property, t should be chosen such that $q^j - 1$ is not divisible by t for every j, $1 \leq j < n$. To have the modular Golomb ruler property, in addition to the distinct property, $q^{j_0} - q^{j_1} - q^{j_2} + q^{j_3}$ should not be divisible by t for all j_0, j_1, j_2, j_3, $0 \leq j_0 \neq j_1 < n, 0 \leq j_2 \neq j_3 < n$, $(j_0, j_1) \neq (j_2, j_3)$. For such t, we have $e(r) = e_{\mathrm{adj}}(r) = 5$ for $r \geq 3$ and the minimum Hamming distance of the code is six. The null space of $\mathbf{H}^{\mathrm{Gabidulin}}_{2,n,t}$ is a QC-LDPC code which we denote by $C^{\mathrm{Gabidulin}}_{2,n,t}$. It can correct any pair of adjacent mutually semi-solid phased bursts of erasures and it is optimal for correcting these erasures.

As a special case, we can take $q = 2$ and $t = 2^\tau - 1$ where $\tau \geq n$. With this choice, $\mathbf{H}_{2,n,t}$ has the distinct property. It also has the modular Golomb ruler property. Indeed, suppose that $2^{j_0} - 2^{j_1} - 2^{j_2} + 2^{j_3}$ is divisible by t for $0 \leq j_0 \neq j_1 < n, 0 \leq j_2 \neq j_3 < n$, $(j_0, j_1) \neq (j_2, j_3)$. Since $-2^n + 2 \leq 2^{j_0} - 2^{j_1} - 2^{j_2} + 2^{j_3} \leq 2^n - 2$ and $t \geq 2^n - 1$, it follows that $2^{j_0} - 2^{j_1} - 2^{j_2} + 2^{j_3} = 0$. Without loss of generality, assume that $j_3 \geq j_0, j_1, j_2$. Since $2^{j_3} > 2^{j_3-1} + 2^{j_3-2} + \cdots + 1$, we conclude that $j_2 = j_3$ and $j_1 = j_0$ or $j_1 = j_3$ and $j_2 = j_0$. Both cases contradict the conditions imposed on the two pairs. Therefore, $\mathbf{H}_{2,n,t}$ has the modular Golomb ruler property. With this choice of q and t, $e(r) = e_{\mathrm{adj}}(r) = 5$ for $r \geq 3$ and $e(2) = 2t/(2^{\tau_n} - 1) - 1$, where τ_n is the largest factor of τ less than n. The drawback of this construction is that the value of t is exponential in n. In Table I we list

in the second and third columns, respectively, the smallest t, denoted by $t_{q\le 5,\min}^{\text{Gabidulin}}(n)$, minimized over $q = 2, 3, 4, 5$, such that $\mathbf{H}_{2,n,t}^{\text{Gabidulin}}$ has the modular Golomb property and a value of $q \le 5$ that yields this minimum. ■

We can construct t-modular Golomb rulers with markers $p_0, p_1, \ldots, p_{n-1}$ with values of t that are substantially less than those obtained above by not restricting p_j to be $(q^j - 1)_t$ as in Example 2. Let $t_{\min}(n)$ be the minimum value of t such that there are n nonnegative integers $p_0, p_1, \ldots, p_{n-1}$ less than t that form a t-modular Golomb ruler. This function has been studied extensively, see e.g., [20]. It is stated in [29] that $n^2 - n+1 \le t_{\min}(n) \le n^2+O(n^{36/23})$, which shows that quadratic growth in n is necessary and sufficient. Constructions of t-modular Golomb rulers with t equal or close to the lower bound for some values of n are due to Singer [30], Bose [31], and Ruzsa [32]. The last two columns in Table I extracted from [33] give for each n, $2 \le n \le 14$, the value of $t_{\min}(n)$ and the n markers of a $t_{\min}(n)$-modular Golomb ruler. The ruler is optimal in the sense that there is no ruler of the same size which is a t-modular Golomb ruler for any $t < t_{\min}(n)$. If the difference between any two consecutive markers is relatively prime to $t_{\min}(n)$, one can use the ruler to construct a parity-check matrix for a code that can correct adjacent phased bursts of erasures which are mutually semi-solid. We succeeded in ordering the markers of each ruler to satisfy this condition except in the case $n = 7$. It should be noted, however, that for any given n there is an infinite number of t-modular Golomb rulers satisfying the condition as shown in Example 2.

B. Correcting Solid Phased Bursts of Erasures

Any code with dimension at most $2t - 1$, such as $C_{2,n,t}$, cannot correct two solid phased bursts of $2t$ erasures. If $C_{2,n,t}$ can correct two adjacent mutually semi-solid phased bursts then its redundancy is $2t - 1$. In this case a subcode of $C_{2,n,t}$ can correct any two adjacent solid phased bursts. The parity-check matrix of this subcode is obtained by augmenting $\mathbf{H}_{2,n,t}$ with an additional row that contains 1 in column $(j; 0)$ whenever j is even and 0's everywhere else. This gives an extra parity equation that can be used to recover one of the erased bits if the channel causes two adjacent solid phased bursts of erasures. The remaining erasures form two adjacent mutually semi-solid phased bursts which are within the correcting capability of $C_{2,n,t}$. In particular, the peeling algorithm applied to the augmented parity-check matrix can correct any two adjacent solid phased bursts of erasures. The dimension of the subcode is $2t$ and, hence, is optimal for correcting such erasures.

If $C_{2,n,t}$ can correct any two mutually semi-solid bursts of erasures, not necessarily adjacent, then it is easy to come up with a subcode, C, of $C_{2,n,t}$ that can correct any two solid phased bursts of erasures. Since any vector of weight $2t$ in which all its 1's are confined to two sections is in the null space of $\mathbf{H}_{2,n,t}$, a parity-check matrix, $\mathbf{H}$, of C can be obtained by augmenting $\mathbf{H}_{2,n,t}$ with a matrix that does not have any such vector in its null space. Hence, the sums of the columns in each column block in the augmenting matrix should be distinct. Therefore, the number of rows in the augmenting matrix is at least $\lceil \log_2(n) \rceil$. A possible choice for such matrix with that many rows is to have the $(j; 0)$ column to be the binary representation of j, $0 \le j < n$, and all other columns to be all-zero columns. The $(j; 0)$ columns in this matrix are all distinct. Hence, if the channel causes two solid phased bursts of erasures, then there is a parity equation that can be used to recover one of the erased bits. Again, the remaining erasures form two mutually semi-solid phased bursts which are within the correcting capability of $C_{2,n,t}$. In particular, the peeling algorithm applied to the augmented parity-check matrix can correct any two solid phased bursts of erasures. Unlike the case for adjacent solid phased bursts, the code is not optimal for correcting any two solid phased bursts of erasures. Indeed, if $n \le 2^t$, then a (possibly shortened or lengthened) Reed-Solomon code of length n and dimension $n - 2$ over GF(2^t) in which each symbol is represented by a binary vector of length t is optimal for correcting pairs of solid phased bursts of erasures.

C. Globally Coupled QC-LDPC Codes for Correcting Multiple Phased Bursts of Erasures

The matrix $\mathbf{H}_{2,n,t}$ given in (1) can be used as a building block to construct long QC-LDPC codes to correct multiple phased bursts of erasures. Here we present an approach in which a number of matrices $\mathbf{H}_{2,n,t}$ are connected globally [34, Chapter 10]. For an integer $l \ge 2$, we define the following $(2l + n)t \times ntl$ matrix

$$\mathbf{H}_{2,n,t,l}^{\text{Global}} = \left[\begin{array}{cccc} \mathbf{H}_{2,n,t} & & & \\ & \mathbf{H}_{2,n,t} & & \\ & & \ddots & \\ & & & \mathbf{H}_{2,n,t} \\ \hline \mathbf{CPM}_{nt}(0) & \mathbf{CPM}_{nt}(0) & \cdots & \mathbf{CPM}_{nt}(0) \end{array}\right]. \tag{2}$$

This matrix consists of two submatrices. The upper submatrix is an $l \times l$ diagonal array of $2t \times nt$ matrices with copies of $\mathbf{H}_{2,n,t}$ on the diagonal. The lower submatrix, which we call the *global coupling matrix*, is a $1 \times l$ array of $nt \times nt$ matrices $\mathbf{CPM}_{nt}(0)$, the $nt \times nt$ identity matrix. The matrix $\mathbf{H}_{2,n,t,l}^{\text{Global}}$ has column weight three and two row weights n and l, for the upper and lower submatrices, respectively. It can also be viewed as a $(2l+n) \times nl$ array of $t \times t$ matrices, each is either a CPM or a ZM. Hence, its null space is a QC code denoted by $C_{2,n,t,l}^{\text{Global}}$. The code $C_{2,n,t,l}^{\text{Global}}$ is a *globally coupled* QC-LDPC code of length ntl and dimension $ntl - \text{rank}(\mathbf{H}_{2,n,t,l}^{\text{Global}})$. The rank of $\mathbf{H}_{2,n,t,l}^{\text{Global}}$ is given in the following theorem, the proof of which appears in Appendix D.

Theorem 4.

$$\text{rank}(\mathbf{H}_{2,n,t,l}^{\text{Global}}) = (l - 1)\text{rank}(\mathbf{H}_{2,n,t}) + nt.$$

It follows from Theorem 4 that the dimension of $C_{2,n,t,l}^{\text{Global}}$ is $(l-1)(nt-\text{rank}(\mathbf{H}_{2,n,t}))$. This product suggests that $C_{2,n,t,l}^{\text{Global}}$ is a product code, which is indeed the case. It is the product of the code $C_{2,n,t}$ and the single parity-check (SPC) code composed of all even-weight words of length l. The minimum Hamming distance of $C_{2,n,t,l}^{\text{Global}}$ is twice that of $C_{2,n,t}$.

TABLE I
$t^{\text{Gabidulin}}_{q\leq 5,\min}(n)$ AND OPTIMAL $t_{\min}(n)$-MODULAR GOLOMB RULERS OF SIZE n FOR $2 \leq n \leq 14$

n	$t^{\text{Gabidulin}}_{q\leq 5,\min}(n)$	q	$t_{\min}(n)$	Optimal Modular Golomb Rulers' Markers
3	7	2	7	$0, 1, 3$
4	15	2	13	$0, 1, 3, 9$
5	25	2	21	$0, 17, 18, 10, 12$
6	41	2	31	$0, 1, 3, 8, 12, 18$
7	69	5	48	$0, 1, 3, 15, 20, 38, 42$
8	73	2	57	$0, 23, 19, 45, 47, 30, 44, 39$
9	73	2	73	$0, 1, 3, 7, 15, 31, 36, 54, 63$
10	191	3	91	$0, 1, 56, 27, 9, 3, 49, 81, 77, 61$
11	197	4	120	$0, 13, 110, 69, 76, 95, 78, 29, 106, 75, 8$
12	239	5	133	$0, 1, 3, 12, 20, 38, 34, 81, 94, 88, 104, 109$
13	295	2	168	$0, 167, 120, 107, 102, 29, 10, 161, 82, 45, 2, 33, 146$
14	295	2	183	$0, 1, 3, 16, 23, 28, 42, 76, 82, 86, 119, 137, 154, 175$

From (2), we see that $\mathbf{H}^{\text{Global}}_{2,n,t,l}$ has a *local structure* represented by each matrix $\mathbf{H}_{2,n,t}$ on the main diagonal of the upper submatrix. These matrices are connected together by the global coupling matrix endowing the matrix $\mathbf{H}^{\text{Global}}_{2,n,t,l}$ with a *global structure* as well. This allows for a two-phase decoding procedure. Let $\mathbf{u} = (\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{l-1})$ be a codeword in $C^{\text{Global}}_{2,n,t,l}$ where each $\mathbf{u}_i$, $0 \leq i < l$, is a sequence of length nt. Then, each such sequence is a codeword in $C_{2,n,t}$, which we call a *local codeword*. If erasures occurring in each local codeword are within the erasure correcting capability of $C_{2,n,t}$, then they can be recovered by applying the peeling algorithm to the parity-check matrix $\mathbf{H}_{2,n,t}$ of $C_{2,n,t}$. If an entire local codeword is erased, and the remaining codewords suffer from erasures that can be corrected by $C_{2,n,t}$, then after recovering them the erased local codeword can be recovered as each bit is checked by a row in the global coupling matrix that checks that bit and no other in the erased local codeword. In particular, if t is a prime and $\mathbf{H}_{2,n,t}$ has the distinct property, then from Theorem 3, $C^{\text{Global}}_{2,n,t,l}$ can recover any local codeword which is entirely erased, in addition to correcting two mutually semi-solid phased bursts in each of the other $l-1$ local codewords. The number of erasures is then $(l-1)(2t-1)+nt$. This is precisely equal to the rank of the matrix $\mathbf{H}^{\text{Global}}_{2,n,t,l}$ given in Theorem 4 where $\text{rank}(\mathbf{H}_{2,n,t}) = 2t-1$ as t is a prime, see Theorem 2. Hence, $C^{\text{Global}}_{2,n,t,l}$ is optimal for correcting these erasures.

IV. QC-LDPC CODES WITH PARITY-CHECK MATRICS OF COLUMN WEIGHT THREE OR MORE

The capability of the codes in Section III with parity-check matrices of column weight two to correct a pair of mutually semi-solid bursts of erasures may not be surpassed by codes with parity-check matrices with higher column weights. However, most noisy channels affect the transmitted data adversely in many ways besides causing bursts of erasures. As stated in Theorem 3, QC-LDPC codes with parity-check matrices of column weight two have minimum Hamming distance of at most six which renders them ineffective in combating random noise. By having $m \geq 3$, the code $C_{m,n,t}$ with parity-check matrix $\mathbf{H}_{m,n,t} = [\mathbf{CPM}_t(p_{i,j})]_{0\leq i<m, 0\leq j<n}$ can be made more effective compared to $C_{2,n,t}$. For this purpose, we choose the parity-check matrix, $\mathbf{H}_{m,n,t}$, to satisfy the RC-constraint. Based on the discussion following Theorem 1, any code with parity-check matrix of column weight m, such as $C_{m,n,t}$, has minimum Hamming distance at least $m+1$ if the matrix satisfies the RC-constraint. We give two explicit examples for the parameters $p_{i,j}$, $0 \leq i < m, 0 \leq j < n$, $m \geq 3$, such that $\mathbf{H}_{m,n,t}$ satisfies the RC-constraint. In both examples, the top two row blocks of the parity-check matrix $\mathbf{H}_{m,n,t}$ of $C_{m,n,t}$ constitute the parity-check matrix $\mathbf{H}_{2,n,t}$ of $C_{2,n,t}$. Hence, $C_{m,n,t}$ can correct all phased bursts of erasures that can be corrected by $C_{2,n,t}$ using the same peeling algorithm.

Example 3. Let $n \geq m \geq 3$ and choose t such that none of the products $i \times j$, where $1 \leq i < m, 0 \leq j < n$, is divisible by t. This is satisfied, for example, if t is greater than $(m-1)(n-1)$ or its largest prime factor is at least equal to both m and n. Let $\mathbf{H}^{\text{RS}}_{m,n,t} = [\mathbf{CPM}_t((ij)_t)]_{0\leq i<m, 0\leq j<n}$, i.e., $p_{i,j} = (ij)_t$. From Proposition 1, $\mathbf{H}^{\text{RS}}_{2,n,t}$ satisfies the RC-constraint since

$$\begin{aligned} &p_{i_1,j_1} - p_{i_0,j_1} - p_{i_1,j_0} + p_{i_0,j_0} \\ &\quad = i_1 j_1 - i_0 j_1 - i_1 j_0 + i_0 j_0 = (i_1 - i_0)(j_1 - j_0) \end{aligned}$$

is not divisible by t. Hence, $\mathbf{H}^{\text{RS}}_{m,n,t}$ is a parity-check matrix of a QC-LDPC code, $C^{\text{RS}}_{m,n,t}$, of minimum Hamming distance at least $m+1$. Notice that the top two row blocks of $\mathbf{H}^{\text{RS}}_{m,n,t}$ constitute the matrix $\mathbf{H}^{\text{RS}}_{2,n,t}$ in Example 1. ■

Example 4. Let $n \geq m \geq 3$ and choose t to be relatively prime to $q \geq 2$ such that none of the products $i \times (q^j - 1)$, where $1 \leq i < m, 0 \leq j < n$, is divisible by t. This is satisfied, for example, if t is greater than $(m-1)(q^{n-1}-1)$ or its largest prime factor is at least equal to both m and q^{n-1}. Let $\mathbf{H}^{\text{Gabidulin}}_{m,n,t} = [\mathbf{CPM}_t((i(q^j-1))_t)]_{0\leq i<m, 0\leq j<n}$, i.e., $p_{i,j} = (i(q^j-1))_t$. From Proposition 1, $\mathbf{H}^{\text{Gabidulin}}_{2,n,t}$ satisfies the RC-constraint since

$$\begin{aligned} &p_{i_1,j_1} - p_{i_0,j_1} - p_{i_1,j_0} + p_{i_0,j_0} \\ &\quad = i_1(q^{j_1}-1) - i_0(q^{j_1}-1) - i_1(q^{j_0}-1) + i_0(q^{j_0}-1) \\ &\quad = (i_1 - i_0)(q^{j_1} - q^{j_0}) \end{aligned}$$

is not divisible by t for $0 \leq i_0 < i_1 < m, 0 \leq j_0 < j_1 < n$. Hence, $\mathbf{H}^{\text{Gabidulin}}_{m,n,t}$ is a parity-check matrix of a QC-LDPC code, $C^{\text{Gabidulin}}_{m,n,t}$, of minimum Hamming distance at least $m+1$. Notice that the top two row blocks of $\mathbf{H}^{\text{Gabidulin}}_{m,n,t}$ constitute the matrix $\mathbf{H}^{\text{Gabidulin}}_{2,n,t}$ in Example 2. ■

The superscript RS in Examples 1 and 3 refers to Reed-Solomon (RS) code [8] with a parity-check matrix of the form

$[\beta^{ij}]_{0\le i<m,0\le j<n}$ while superscript $\mathrm{Gabidulin}$ in Examples 2 and 4 refers to Gabidulin code [9] with a parity-check matrix of the form $[\beta^{i(q^j-1)}]_{0\le i<m,0\le j<n}$ where q is a power of a prime and β is an element in an extension field of GF(q) satisfying certain properties. In both cases, the exponents of β reduced modulo t are precisely the values of $p_{i,j}$ defining the CPMs in the constructions. The RS and the Gabidulin codes are nonbinary codes that can be considered as base codes for constructing QC-LDPC codes, see [35] where techniques to determine the dimensions of the QC-LDPC codes are also presented. The use of RS codes to construct QC-LDPC codes for correcting erasures is explored in [36],[37].

We simulated the performances of the codes $C^{\mathrm{RS}}_{2,12,239}$, $C^{\mathrm{Gabidulin}}_{2,12,239}$, $C^{\mathrm{RS}}_{6,12,239}$, and $C^{\mathrm{Gabidulin}}_{6,12,239}$ over the AWGN channel using a scaled min-sum decoder. The four codes are of length 2868 and their parity-check matrices are composed of 12 column blocks of 239×239 CPMs. The codes $C^{\mathrm{RS}}_{2,12,239}$ and $C^{\mathrm{Gabidulin}}_{2,12,239}$ have parity-check matrices $\mathbf{H}^{\mathrm{RS}}_{2,12,239} = [\mathbf{CPM}_{239}((ij)_{239})]_{0\le i<2,0\le j<12}$ and $\mathbf{H}^{\mathrm{Gabidulin}}_{2,12,239} = [\mathbf{CPM}_{239}((i(5^j-1))_{239})]_{0\le i<2,0\le j<12}$, respectively. Both codes have dimension 2391 and rate 0.8337. The codes $C^{\mathrm{RS}}_{6,12,239}$ and $C^{\mathrm{Gabidulin}}_{6,12,239}$ have parity-check matrices $\mathbf{H}^{\mathrm{RS}}_{6,12,239} = [\mathbf{CPM}_{239}((ij)_{239})]_{0\le i<6,0\le j<12}$ and $\mathbf{H}^{\mathrm{Gabidulin}}_{6,12,239} = [\mathbf{CPM}_{239}((i(5^j-1))_{239})]_{0\le i<6,0\le j<12}$, respectively. Both codes have dimension 1439 and rate 0.5017. The constructions of $C^{\mathrm{Gabidulin}}_{2,12,239}$ and $C^{\mathrm{Gabidulin}}_{6,12,239}$ are based on $q = 5$, see the entry in Table I for $n = 12$. The bit error rate (BER) and the frame error rate (FER) of the four codes are shown in Fig. 1 where α is the scaling factor used in min-sum decoding. As shown, $C^{\mathrm{Gabidulin}}_{2,12,239}$ performs better than $C^{\mathrm{RS}}_{2,12,239}$ as their minimum Hamming distances are six and four, respectively, see Examples 1 and 2.

V. Conclusion

In this paper, we considered the use of QC-LDPC codes with parity-check matrices of the form $\mathbf{H}_{m,n,t} = [\mathbf{CPM}_t(p_{i,j})]_{0\le i<m,0\le j<n}$ for correcting phased bursts of erasures. Such codes cannot correct erasures forming two solid phased bursts regardless of the column weight m of their parity-check matrices. Since, in general, codes with parity-check matrices of column weight $m = 1$ cannot correct a single error, we first focused on codes with parity-check matrices of column weight $m = 2$ and determined their abilities to correct phased bursts of erasures. Using these parity-check matrices, we showed how to modify them to correct two solid phased bursts of erasures and how to globally couple them to correct more erasures. To improve performance over the AWGN channel, we considered QC-LDPC codes with parity-check matrices $\mathbf{H}_{m,n,t}$ with $m \ge 3$ which include as submatrices well designed $\mathbf{H}_{2,n,t}$.

In the proof of Theorem 3 in Appendix C we related the linear dependence of columns in the parity-check matrix to cycles in the associated Tanner graphs. The girths of the constructions in this paper are 4, 8, or 12. It should be noted that girths of up to 24 can be achieved with parity-check matrices of column weight two which are arrays not only of CPMs but also ZMs, see, e.g., [38] and references therein. Codes with Tanner graphs having such high girths should have better erasure correcting capabilities than those specified in Theorem 3 for the code $C_{2,n,t}$, including the correction of multiple solid phased bursts of erasures. The price to be paid for such better capabilities is a reduction in code rate.

Appendix A
Proof of Theorem 1

Suppose $\mathbf{H}$ is not ML peeling-decodable. Then, there is a nonempty set, $\mathcal{J}$, of variable nodes in $\mathcal{G}$ that form a stopping set such that the columns in $\mathbf{H}$ indexed by $\mathcal{J}$ are linearly independent. Let $\mathcal{I}$ be the set of check nodes in the subgraph $\mathcal{G}(\mathcal{J})$ induced by $\mathcal{J}$. This is the subgraph of $\mathcal{G}$ consisting of the variable nodes in $\mathcal{J}$, all edges incident on these variable nodes, and all check nodes adjacent to these nodes. As every column in $\mathbf{H}$ has weight at most two, the number of edges incident on $\mathcal{J}$ is at most $2|\mathcal{J}|$. Since $\mathcal{J}$ forms a stopping set, every check node in $\mathcal{I}$ is adjacent to at least two variable nodes in $\mathcal{J}$. Hence, the number of edges incident on these check nodes is at least $2|\mathcal{I}|$. Since the columns in $\mathbf{H}$ indexed by $\mathcal{J}$ are linearly independent, we have $|\mathcal{J}| \le |\mathcal{I}|$. As the edges in $\mathcal{G}(\mathcal{J})$ incident on $\mathcal{I}$ are the same as those incident on $\mathcal{J}$, we conclude that $|\mathcal{I}| = |\mathcal{J}|$ and every node in $\mathcal{I}$ or $\mathcal{J}$ is incident on exactly two edges. This is to say that every row in the submatrix of $\mathbf{H}$ composed of the columns indexed by $\mathcal{J}$ has weight two, contradicting the assumption that the columns indexed by $\mathcal{J}$ are linearly independent. ■

Appendix B
Proof of Theorem 2

We start with the following lemma which gives the rank of an array of circulants that are not necessarily CPMs. The proof is based on Bézout's identity which states that given polynomials $a_0(x), a_1(x), \ldots, a_n(x)$ over some field with greatest common divisor (GCD) $f(x)$, there exists polynomials $q_0(x), q_1(x), \ldots, q_n(x)$ such that $f(x) = q_0(x)a_0(x) + q_1(x)a_1(x) + \cdots + q_n(x)a_n(x)$, see e.g., [39, Corollary 1.37].

Lemma 1. [3] *Let $\mathbf{A}_0, \mathbf{A}_1, \ldots, \mathbf{A}_{n-1}$ be $t \times t$ circulants over some field and $\mathbf{A} = [\mathbf{A}_0, \mathbf{A}_1, \ldots, \mathbf{A}_{n-1}]$. For $0 \le j < n$, let $\mathbf{a}_j = (a_{0,j}, a_{1,j}, \ldots, a_{t-1,j})$ be the top row of $\mathbf{A}_j$ and $a_j(x) = a_{0,j} + a_{1,j}x + \cdots + a_{n-1,j}x^{n-1}$. Then, $\mathrm{rank}(\mathbf{A}) = t - \deg(f(x))$ where $f(x) = \mathrm{GCD}(a_0(x), a_1(x), \ldots, a_{n-1}(x), x^t - 1)$ and $\deg(f(x))$ is the degree of the polynomial $f(x)$.*

Proof: Let $f(x) = \sum_{j=0}^{t-1} f_j x^j$ and define the sequence $\mathbf{f} = (f_0, f_1, \ldots, f_{t-1})$. Then, with $a_n(x) = x^t - 1$, Bézout's identity yields

$$f(x) \equiv q_0(x)a_0(x) + q_1(x)a_1(x) + \cdots + q_{n-1}(x)a_{n-1}(x) \pmod{x^t - 1}$$

[3] In the special case in which $n = 1$, Lemma 1 gives the rank of a circulant. In this special case, if the circulant is over GF(q) and $t = q - 1$, the result is known as the König-Rados Theorem [40]. More generally, if t is not divisible by the characteristic of the field, Newman [41] provided a proof based on a similarity transformation of the circulant. Although the result for general t in case of a single circulant may be a folk theorem to coding theorists, we did not find it explicitly stated in the coding literature. We believe that the generalization to an array of $n > 1$ circulants is new.

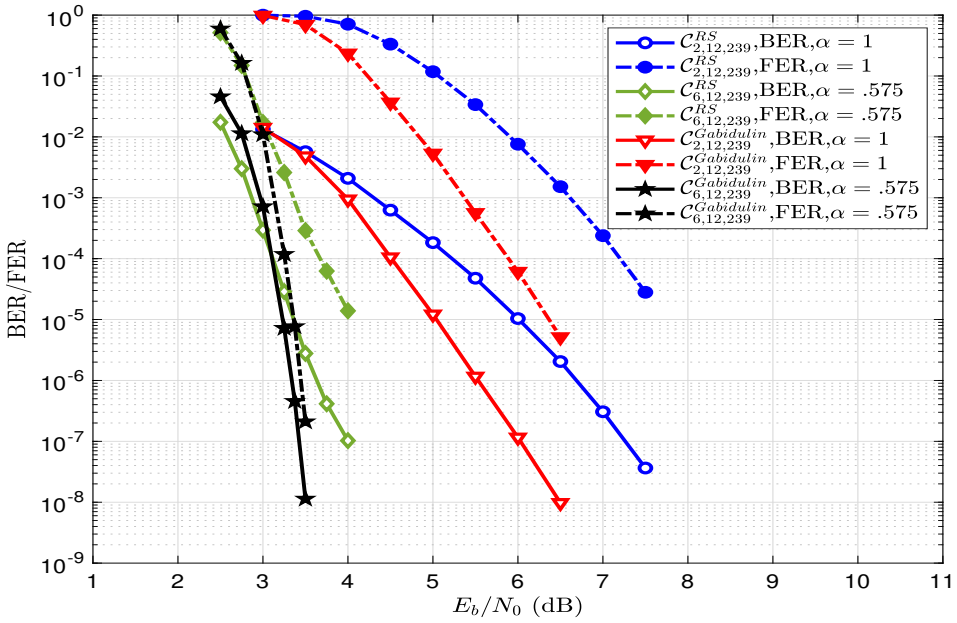

Fig. 1. The BER and BLER performances of $C^{\mathrm{RS}}_{2,12,239}$, $C^{\mathrm{Gabidulin}}_{2,12,239}$, $C^{\mathrm{RS}}_{6,12,239}$, and $C^{\mathrm{Gabidulin}}_{6,12,239}$ over the AWGN channel decoded with a scaled min-sum decoder.

for some polynomials $q_0(x), q_1(x), \ldots, q_{n-1}(x)$. This implies that $\mathbf{f}$ is a linear combination of $\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{n-1}$ and their cyclic shifts. Since the columns of a circulant are the the same as its rows read in reverse, then $\overleftarrow{\mathbf{f}} = (f_{t-1}, f_{t-2}, \ldots, f_0)$ is in the column space of the matrix $\mathbf{A}$. Notice that $\overleftarrow{\mathbf{f}}$ starts with exactly $t - \deg(f(x)) - 1$ zeros. Hence, $\overleftarrow{\mathbf{f}}$ and its $t-\deg(f(x))-1$ cyclic shifts are linearly independent. Since $\overleftarrow{\mathbf{f}}$ is in the column space of $\mathbf{A}$ which is a row of circulants, all cyclic shifts of $\overleftarrow{\mathbf{f}}$ are also in the same column space. Thus, $\mathbf{A}$ has rank at least $t - \deg(f(x))$ as it contains that many linearly independent vectors. To show that the rank of $\mathbf{A}$ does not exceed $t - \deg(f(x))$, we argue that every vector $\mathbf{s} = (s_0, s_1, \ldots, s_{t-1})$ in the column space of $\mathbf{A}$ is a linear combination of these $t-\deg(f(x))$ linearly independent vectors composed of $\overleftarrow{\mathbf{f}}$ and its cyclic shifts. Indeed, as $\mathbf{s}$ is in the column space of $\mathbf{A}$, $\overleftarrow{\mathbf{s}} = (s_{t-1}, s_{t-2}, \ldots, s_0)$ is a linear combination of $\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{n-1}$ and their $t-1$ cyclic shifts. In particular, for some polynomials $u_0(x), u_1(x), \ldots, u_{n-1}(x)$, we can write

$$\overleftarrow{s}(x) \equiv u_0(x)a_0(x) + u_1(x)a_1(x) + \cdots + u_{n-1}(x)a_{n-1}(x) \pmod{x^t - 1},$$

where $\overleftarrow{s}(x) = \sum_{i=0}^{t-1} s_{t-1-i}x^i$. Since $f(x) = \mathrm{GCD}(a_0(x), a_1(x), \ldots, a_{n-1}(x), x^t - 1)$, it follows that $f(x)$ divides $\overleftarrow{s}(x)$, i.e., $\overleftarrow{s}(x) = q(x)f(x)$ for some polynomial $q(x)$ of degree less than $t - \deg(f(x))$. This is equivalent to saying that $\overleftarrow{\mathbf{s}}$ is a linear combination of $\mathbf{f}$ and its $t-\deg(f(x))-1$ cyclic shifts which is the same as saying that $\mathbf{s}$ is a linear combination of $\overleftarrow{\mathbf{f}}$ and its $t-\deg(f(x))-1$ cyclic shifts. ∎

We continue with the proof of Theorem 2. First, we subtract the top row block of the matrix $\mathbf{H}_{2,n,t}$ given in (1) where $p_0 = 0$ from the bottom row block to obtain the matrix

$$\mathbf{H}' = \begin{bmatrix} \mathbf{CPM}_t(0) & \mathbf{CPM}_t(0) & \cdots & \mathbf{CPM}_t(0) \\ \mathbf{0} & \mathbf{A}_1 & \cdots & \mathbf{A}_{n-1} \end{bmatrix},$$

where $\mathbf{0}$ is the $t \times t$ all-zero matrix and $\mathbf{A}_j = \mathbf{CPM}_t(p_j) - \mathbf{CPM}_t(0)$. For $1 \le j < n$, the matrix $\mathbf{A}_j$ is a circulant in which its top row is either the all-zero vector or has exactly two 1's at positions p_j and 0. Since $\mathbf{H}'$ is obtained from $\mathbf{H}_{2,n,t}$ by elementary row operations, they have the same rank. Furthermore, as $\mathbf{CPM}_t(0)$, being an identity matrix, has rank t, we have

$$\mathrm{rank}(\mathbf{H}_{2,n,t}) = \mathrm{rank}(\mathbf{H}') = t + \mathrm{rank}(\mathbf{A}), \tag{3}$$

where $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_{n-1}]$ is composed of $n-1$ circulants. We invoke Lemma 1 to find the rank of this matrix. For this purpose, let $a_j(x) = x^{p_j} - 1$ for $1 \le j < n$. Then,

$$\begin{aligned} f(x) &= \mathrm{GCD}(a_1(x), \ldots, a_{n-1}(x), x^t - 1) \\ &= \mathrm{GCD}(x^{p_1} - 1, \ldots, x^{p_{n-1}} - 1, x^t - 1) \\ &= x^{\mathrm{GCD}(p_1, \ldots, p_{n-1}, t)} - 1, \end{aligned}$$

where we used the well-known fact that $\mathrm{GCD}(x^a - 1, x^b - 1) = x^{\mathrm{GCD}(a,b)} - 1$ for nonnegative integers a and b. The result now follows directly from Lemma 1 and (3). ∎

Appendix C
Proof of Theorem 3

Since the t columns in any column block are linearly independent, $e(1) = t$. As any r column blocks, where $2 \le r \le n$, have linearly dependent columns, $e(r)$ is one less than the minimum number of linearly dependent columns confined to r column blocks. Since each column in $\mathbf{H}_{2,n,t}$ has a single 1 in the top row block and a single 1 in the bottom

row block, only an even number of columns in $\mathbf{H}_{2,n,t}$ can sum up to the all-zero vector and, therefore, $e(r)$ is odd for $2 \leq r \leq n$.

Our approach is based on relating linear dependence of columns in $\mathbf{H}_{2,n,t}$ to cycles in the Tanner graph $\mathcal{G}$ representing $\mathbf{H}_{2,n,t}$. Indeed, the columns of $\mathbf{H}_{2,n,t}$ indexed by a nonempty set $\mathcal{J}$ of indices are linearly dependent if and only if there is a cycle in the subgraph, $\mathcal{G}(\mathcal{J})$, of $\mathcal{G}$ induced by $\mathcal{J}$. Notice that "if" uses the fact that every column in $\mathbf{H}_{2,n,t}$ has exactly two 1's.

Without loss of generality, we can assume that a cycle in $\mathcal{G}$ starts with the variable node $(j_0; j_0')$ followed by a check node in the top row block followed by the variable node $(j_1; j_0')$ followed by a check node in the bottom row block and so on until it reaches a variable node $(j_{z-1}; j_{z-1}')$ followed by a check node in the bottom row block and finally ends at the variable node $(j_z; j_z') = (j_0; j_0')$ we started with for some positive even integer z. Based on this, the cycle can be completely specified by the sequence $(j_0; j_0'), (j_1; j_1'), \ldots, (j_{z-1}; j_{z-1}')$ of variable nodes without listing the check nodes or the ending variable node which is the same as the starting node. The length of the cycle is $2z$. For such a sequence to form a cycle it is necessary that

1) $j_\ell \neq j_{\ell+1}$ for $0 \leq \ell < z$ where $j_z = j_0$ as no check node is adjacent to two variable nodes in the same column block;
2) If ℓ is even, then $j_\ell' = j_{\ell+1}'$ for the variables nodes $(j_\ell; j_\ell')$ and $(j_{\ell+1}; j_{\ell+1}')$ to be adjacent to a check node in the top row block;
3) If ℓ is odd, then $(j_\ell' - p_{j_\ell})_t = (j_{\ell+1}' - p_{j_{\ell+1}})_t$, where $(j_z; j_z') = (j_0; j_0')$, for the variables nodes $(j_\ell; j_\ell')$ and $(j_{\ell+1}; j_{\ell+1}')$ to be adjacent to a check node in the bottom row block.

Combined with the condition that $(j_0; j_0'), (j_1; j_1'), \ldots, (j_{z-1}; j_{z-1}')$ are distinct gives a necessary and sufficient condition for the sequence to form a cycle. If this extra condition is not met, then the sequence represents a closed walk that contains a cycle of length less than $2z$.

To determine $e(2)$, we consider the minimum number of linearly dependent columns confined to the column blocks j_0 and j_1, where $0 \leq j_0 \neq j_1 < n$. There are z such columns only if there is a sequence $(j_0; j_0'), (j_1; j_1'), \ldots, (j_{z-1}; j_{z-1}')$ of variable nodes satisfying conditions 1), 2), and 3). Then, for even ℓ we have $j_\ell = j_0$ and $j_\ell' = j_{\ell+1}'$ while for odd ℓ we have $j_\ell = j_1$ and $(j_\ell' - p_{j_\ell})_t = (j_{\ell+1}' - p_{j_{\ell+1}})_t$. Summing over $\ell = 0, 1, \ldots, z-1$, we get $\frac{1}{2}(p_{j_1} - p_{j_0})z \equiv 0 \pmod t$. The minimum value of z for this congruency to hold is $2t/\mathrm{GCD}(p_{j_1} - p_{j_0}, t)$. Hence, there is no cycle of length less than $2z$ with $z = 2t/\mathrm{GCD}(p_{j_1} - p_{j_0}, t)$ involving only variable nodes confined to the column blocks j_0 and j_1. For such z, we can find a closed walk of length $2z$. Indeed, let $(j_\ell; j_\ell') = (j_0; ((p_{j_0} - p_{j_1})\frac{\ell}{2})_t)$ if $\ell = 0, 2, \ldots, z$ and $(j_\ell; j_\ell') = (j_1; ((p_{j_0} - p_{j_1})\frac{\ell-1}{2})_t)$ if $\ell = 1, 3, \ldots, z-1$. Then, $(j_z; j_z') = (j_0; j_0')$ and the three conditions 1), 2), 3) hold. We conclude that the length of a shortest cycle of variable nodes confined to the column blocks j_0 and j_1 is $2z$ and z is the minimum number of linearly dependent columns confined to these column blocks. From this, the expression of $e(2)$ follows.

Next, we consider $e(r)$ in case $r \geq 3$. If $\mathbf{H}_{2,n,t}$ does not have the distinct property, then $p_{j_0} = p_{j_1}$ for some $j_0 \neq j_1$. In this case, the j'-th column, $0 \leq j' < t$, in the j_0-th column block is the same as the j'-th column in the j_1-th column block. Hence, $e(r) = 1$. If $\mathbf{H}_{2,n,t}$ has the distinct property, then no two columns in $\mathbf{H}_{2,n,t}$ are identical and $e(r) > 1$ which implies that $e(r) \geq 3$. If $\mathbf{H}_{2,n,t}$ does not have the modular Golomb ruler property, then $(p_{j_1} - p_{j_0})_t = (p_{j_2} - p_{j_3})_t$ for distinct pairs (j_0, j_1) and (j_3, j_2) such that $0 \leq j_0, j_1, j_2, j_3 < n$, $j_0 \neq j_1$, and $j_2 \neq j_3$. Then $j_1 \neq j_2$ otherwise $p_{j_0} = p_{j_3}$ which implies that $j_0 = j_3$ as $p_0, p_1, \ldots, p_{n-1}$ are distinct. Similarly, $j_3 \neq j_0$. The sequence of variable nodes $(j_0; 0), (j_1; 0), (j_2; (p_{j_2} - p_{j_1})_t), (j_3; (p_{j_2} - p_{j_1})_t)$ satisfies conditions 1), 2), and 3) and forms a cycle of length 8. Therefore, the columns of $\mathbf{H}_{2,n,t}$ indexed by these four variable nodes are linearly dependent and $e(r) = 3$. If $\mathbf{H}_{2,n,t}$ has the modular Golomb ruler property, then there is no such sequence of variable nodes and $e(r) > 3$, which implies that $e(r) \geq 5$. However, consider the sequence of the six variable nodes $(j_0; 0), (j_1; 0), (j_2; (p_{j_2} - p_{j_1})_t), (j_0; (p_{j_2} - p_{j_1})_t), (j_1; (p_{j_2} - p_{j_0})_t), (j_2; (p_{j_2} - p_{j_0})_t)$, where $0 \leq j_0 < j_1 < j_2 < n$. This sequence satisfies conditions 1), 2), and 3) and forms a cycle of length 12. Therefore, the columns of $\mathbf{H}_{2,n,t}$ indexed by these six variable nodes which are confined to the three column blocks j_0, j_1, j_2 are linearly dependent. This proves that $e(r) \leq 5$ for all $3 \leq r \leq n$.

The results for $e_{\mathrm{adj}}(r)$ follow by confining the erasures to r consecutive column blocks. ■

APPENDIX D
PROOF OF THEOREM 4

Clearly, the upper and lower submatrices have row spaces of dimensions $l\mathrm{rank}(\mathbf{H}_{2,n,t})$ and nt, respectively. Because of the structure of the lower submatrix, a linear combination of the rows of the upper submatrix belongs to the row space of the lower submatrix if and only if it involves the same linear combination of the rows of each matrix $\mathbf{H}_{2,n,t}$ on the diagonal. Hence, the dimension of the intersection of the row spaces of the two submatrices is $\mathrm{rank}(\mathbf{H}_{2,n,t})$. The dimension of the row space of $\mathbf{H}_{2,n,t,l}^{\mathrm{Global}}$ is the sum of the dimensions of the row spaces of the upper and lower submatrices excluding the dimension of their intersection. We conclude that the rank of $\mathbf{H}_{2,n,t,l}^{\mathrm{Global}}$, which is the dimension of its row space, is as given in the theorem. ■

REFERENCES

[1] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Efficient erasure correcting codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 569–584, Feb. 2001.
[2] M. Yang and W. E. Ryan, "Performance of efficiently encodable low-density parity-check codes in noise bursts on the EPR4 channel," *IEEE Trans. Magn.*, vol. 40, no. 2, pp. 507–512, Mar. 2004.
[3] G. Hosoya, H. Yagi, T. Matsushima, and S. Hirasawa, "A modification method for constructing low-density parity-check codes for burst erasures," *ICICE Trans. Fundamentals*, vol. E89-A, no. 10, pp. 2501–2509, Oct. 2006.

[4] Y. Y. Tai, L. Lan, L. Zeng, S. Lin, and K. A. S. Abdel-Ghaffar, "Algebraic construction of quasi-cyclic LDPC codes for the AWGN and erasure channels," *IEEE Trans. Commun.*, vol. 54, no. 10, pp. 1765–1774, Oct. 2006.

[5] S. J. Johnson, "Burst erasure correcting LDPC codes," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 641–652, Mar. 2009.

[6] K. Li, A. Kavčić, and M. F. Erden, "Construction of burst-erasure efficient LDPC codes for use with belief propagation decoding," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Cape Town, South Africa, May 23–27, 2010, pp. 1–5.

[7] X. Ge and S. -T. Xia, "Structured non-binary LDPC codes with large girth," *Electron. Lett.*, vol. 43, no. 22, pp. 1220–1221, Oct. 2007.

[8] S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2004.

[9] E. M. Gabidulin, "Theory of codes with maximum rank distance," *Problems Inf. Transmiss.*, vol. 21, no. 1, pp. 3–16, Jul. 1985.

[10] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1570–1579, Jun. 2002.

[11] H. Pishro-Nik and F. Fekri, "On decoding of low-density parity-check codes over the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 439–454, Mar. 2004.

[12] V. Savin, "LDPC decoders," in *Channel Coding: Theory, Algorithms, and Applications*, D. Declerq, M. Fossorier, and E. Biglieri Eds., Oxford, UK: Academic Press, 2014.

[13] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 533 – 547, Sep. 1981.

[14] W. E. Ryan and S. Lin, *Channel Codes: Classical and Modern*. New York, NY: Cambridge University Press, 2009.

[15] H. D. L. Hollmann and L. M. G. M. Tolhuizen, "Generic erasure correcting sets: bounds and constructions," *J. Combin. Theory, Ser. A*, vol. 113, no. 8, pp. 1746–1759, 2006.

[16] H. D. L. Hollmann and L. M. G. M. Tolhuizen, "On parity check collections for iterative erasure decoding that correct all correctable erasure patterns of a given size," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 823–828, Feb. 2007.

[17] J. H. Weber and K. A. S. Abdel-Ghaffar, "Results on parity-check matrices with optimal stopping and/or dead-end set enumerators," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1368–1374, Mar. 2008.

[18] M. P. C. Fossorier, "Quasi-cyclic low-density parity-check codes from circulant permutation matrices," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1788–1793, Aug. 2004.

[19] B. Vasić and O. Milenkovic, "Combinatorial constructions of low-density parity-check codes for iterative decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1156–1176, Jun. 2004.

[20] C. J. Colbourn and J. H. Dinitz, *Handbook of Combinatorial Designs*, 2nd ed. CRC Press: Boca Raton, FL, 2007.

[21] A. Sălăgean, D. Gardner, and R. Phan, "Index tables of finite fields and modular Golomb rulers," in *Sequences and Their Applications* (Lecture Notes in Computer Science, vol. 7280), T. Helleseth and J. Jedwab, Eds. Berlin, Germany: Springer, 2012, pp. 136–147.

[22] K. O'Bryant, "A complete annotated bibliography of work related to Sidon sequences," *Electron. J. Combin.*, vol. DS11, pp. 1–39, Jul. 2004.

[23] C. Chen, B. Bai, and X. Wang, "Construction of nonbinary quasi-cyclic LDPC cycle codes based on Singer perfect difference set," *IEEE Commun. Lett.*, vol. 14, no. 2, pp. 181–183, Feb. 2010.

[24] M. Esmaeili and M. Javedankherad, "4-cycle free LDPC codes based on difference sets," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3579–3586, Dec. 2012.

[25] G. Zhang, R. Sun, and X. Wang, "New quasi-cyclic LDPC codes with girth at least eight based on Sidon sequences," in *Proc. Int. Symp. Turbo Codes and Iterative Information Processing (ISTC)*, Gothenburg, Sweden, Aug. 27–31, 2012, pp. 31–35.

[26] H. Park, S. Hong, J.-S. No, and D.-J. Shin, "Construction of high-rate regular quasi-cyclic LDPC codes based on cyclic difference families," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3108–3113, Aug. 2013.

[27] R. G. Gallager, *Low Density Parity Check Codes*. Cambridge, MA: MIT Press, 1963.

[28] D. J. C. MacKay and M. C. Davey, "Evaluation of Gallager codes for short block length and high rate applications," in Codes Syst. Graphical Models, New York: Springer-Verlag, vol. 123, pp. 113–130, 2001.

[29] R. L. Grahams and N. J. A. Sloane, "On additive bases and harmonious graphs," *SIAM J. Alg. Disc. Meth.*, vol. 1, no. 4, Dec. 1980.

[30] J. Singer, "A theorem in finite projective geometry and some applications to number theory," *Trans. Amer. Math. Soc.*, vol. 43, pp. 377–385, 1938.

[31] R. C. Bose, "An affine analogue of Singer's theorem," *J. Indian Math. Soc.*, vol. 6, pp. 1–15, 1942.

[32] I. Ruzsa, "Solving a linear equation in a set of integers I," *Acta Arith.*, vol. 65, no. 3, 259–282, 1993.

[33] H. Haanpää, A. Huima, and P. Östergård, "Sets in $\mathbb{Z}_n$ with distinct sums of pairs," *Disc. Appl. Math.*, vol. 138, no. 1, pp. 99–106, 2004.

[34] J. Li, S. Lin, K. Abdel-Ghaffar, W. E. Ryan, and D. J. Costello, Jr., *LDPC Code Designs, Constructions, and Unification*. Cambridge, UK: Cambridge University Press, 2017.

[35] Q. Diao, Q. Huang, S. Lin, and K. Abdel-Ghaffar, "A matrix-theoretic approach for analyzing quasi-cyclic low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 4030–4048, Jun. 2012.

[36] J. Li, K. Liu, S. Lin, and K. Abdel-Ghaffar, "Reed-Solomon based nonbinary globally coupled LDPC codes: Correction of random errors and bursts of erasures," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Aachen, Germany, Jun. 25 – 30, 2017, pp. 381–385.

[37] X. Xiao, W. E. Ryan, B. Vasić, S. Lin, and K. Abdel-Ghaffar, "Reed-Solomon-based quasi-cyclic LDPC codes: Designs, cycle structure and erasure correction," in *Proc. Inform. Theory Applic. Workshop (ITA)*, San Diago, CA, Feb. 11–16, 2018, pp. 1–10.

[38] M. Gholami and M. Samadieh, "Design of binary and nonbinary codes from lifting of girth-8 cycle codes with minimum lengths," *IEEE Commun. Lett.*, vol. 17, no. 4, pp. 777–780, Apr. 2013.

[39] P. A. Fuhrmann, *A Polynomial Approach to Linear Algebra*, 2nd ed. New York, NY: Springer, 2012.

[40] R. Lidl and H. Niederreiter, *Finite Fields*, second edition. Cambridge, UK: Cambridge University Press, 1997.

[41] M. Newman, "Circulants and difference sets," *Proc. Amer. Math. Soc.*, vol. 88, no. 1, pp. 184–188, May 1983.

PLACE PHOTO HERE

Xin Xiao received the B.S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2012. Under a dual degree program, she received the M. E. degree in system LSI from the Graduate School of Information, Production and System, Waseda University, Kitakyushu, Fukuoka, Japan, in 2013, and the M. S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2015. She is currently pursuing the Ph.D. degree in electrical and computer engineering at the University of Arizona, Tucson, AZ, United States.

From 2016, she is a Research Assistant with the Electrical and Computer Engineering Department at the University of Arizona, Tucson, AZ, United States. Her current research interests are in the general area of error correction coding for communication and storage systems, and deep learning and optimization in channel coding.

PLACE PHOTO HERE

Bane Vasić is a Professor of Electrical and Computer Engineering and Mathematics at the University of Arizona and a Director of the Error Correction Laboratory. He is also a member of the Center for Quantum Networks funded by the NSF and the Superconducting Materials and Systems Center funded by the Department of Energy. He is an inventor of the soft error event decoding algorithm, and the key architect of a detector/decoder for Bell Labs data storage read channel chips which were regarded as the best in industry. His pioneering work on structured low density parity check (LDPC) error correcting codes and invention of codes has enabled low-complexity iterative decoder implementations. Structured LDPC codes are today adopted in a number of communications standards and data storage systems. Dr. Vasić is known for his theoretical work in error correction coding theory and codes on graphs which has led to characterization of the hard decision iterative decoders of LDPC codes, and design of decoders with best error-floor performance known today. He is a co-founder of Codelucida, a startup company developing advanced error correction solutions for communications and data storage. He is an IEEE Fellow, Fulbright Scholar, da Vinci Fellow, and a past Chair of IEEE Data Storage Technical Committee.

PLACE PHOTO HERE

Shu Lin (S'62-M'65-SM'78-F'80-LF'00) received the B.S.E.E. degree from the National Taiwan University, Taipei, Taiwan, in 1959, and the M.S. and Ph.D. degrees in electrical engineering from Rice University, Houston, TX, in 1964 and 1965, respectively. In 1965, he joined the Faculty of the University of Hawaii, Honolulu, as an Assistant Professor of Electrical Engineering. He became an Associate Professor in 1969 and a Professor in 1973. In 1986, he joined Texas A&M University, College Station, as the Irma Runyon Chair Professor of Electrical Engineering. In 1987, he returned to the University of Hawaii. From 1978 to 1979, he was a Visiting Scientist at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, where he worked on error control protocols for data communication systems. He spent the academic year of 1996-1997 as a Visiting Chair Professor at the Technical University of Munich, Munich, Germany.

He retired from University of Hawaii in 1999 and he is currently an Adjunct Professor at University of California, Davis, California. He has published at least 800 technical papers in prestigious refereed technical journals and international conference proceedings. He is the author of the book, An Introduction to Error-Correcting Codes (Englewood Cliff, NJ: Prentice-Hall, 1970). He also co-authored (with D. J. Costello) the book, Error Control Coding: Fundamentals and Applications (Upper Saddle River, NJ: Prentice-Hall, 1st edition, 1982, 2nd edition, 2004), the book (with T. Kasami, T. Fujiwara, and M. Fossorier), Trellises and Trellis-Based Decoding Algorithms, (Boston, MA: Kluwer Academic, 1998), and the book, Channel Codes: Classical and Modern (Cambridge University Press 2009).

Dr. Lin was elected to IEEE (Institute of Electrical and Electronic Engineering) Fellow in 1980 and Life Fellow in 2000. In 1996, he was a recipient of the Alexander von Humboldt Research Prize for U.S. Senior Scientists and a recipient of the IEEE Third-Millennium Medal, 2000. In 2007, he was a recipient of The Communications Society Stephen O. Rice Prize in the Field of Communications Theory. In 2014, he was awarded the NASA Exceptional Public Achievement Medal. He is the recipient of the IEEE 2020 IEEE Leon K. Kirchmayer Graduate Teaching Award.

PLACE PHOTO HERE

Juane Li received the B.E. degree from Harbin Institute of Technology, Harbin, Heilongjiang, China, in 2010, in electrical and computer engineering. She received her Ph.D degree in electrical and computer engineering in the University of California, Davis, in 2016. Her research interests are in the general area of channel coding for communication and storage systems, and the hardware implementation of encoder and decoder of LDPC codes. She is currently a staff systems architect at Micron Technology Inc., San Jose, CA 95131.

PLACE PHOTO HERE

Khaled Abdel-Ghaffar received the B.Sc. degree from Alexandria University, Alexandria, Egypt, in 1980, and the M.S. and Ph.D. degrees from the California Institute of Technology, Pasadena, CA, USA, in 1983 and 1986, respectively, all in electrical engineering. He is currently a Professor of electrical and computer engineering with the University of California, Davis. His main interest is coding theory.

Dr. Abdel-Ghaffar was a co-recipient of the IEEE Communications Society 2007 Stephen O. Rice Prize Paper Award. He served as an Associate Editor for Coding Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY from 2002 to 2005 and as an Associate Editor for Algebraic and LDPC Codes for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2012 to 2017.