Reed-Solomon Based Quasi-Cyclic LDPC Codes: Designs, Girth, Cycle Structure, and Reduction of Short Cycles

Xin Xiao, Bane Vasić, *Fellow, IEEE,* Shu Lin, *Life Fellow, IEEE,* Khaled Abdel-Ghaffar, *Senior Member, IEEE,* and William E. Ryan, *Fellow, IEEE*

***Abstract*—Designs and constructions of quasi-cyclic (QC) LDPC codes for the AWGN channel are presented. The codes are constructed based on *the conventional parity-check matrices of Reed-Solomon* (RS) codes and are referred to as *RS-QC-LDPC codes*. Several classes of RS-QC-LDPC codes are given. Cycle structural properties of the Tanner graphs of codes in these classes are analyzed and specific methods for constructing codes with girth at least eight and reducing their short cycles are presented. The designed codes perform well in both waterfall and low error-rate regions.**

***Index Terms*—Cycles, Dispersion, Girth, LDPC code, Masking, Reed-Solomon code, Tanner graph.**

I. Introduction

Some recent research [2]–[5] combines two powerful categories of codes, namely Reed-Solomon (RS) [6], and LDPC codes [7], to form powerful classes of hybrid codes which, not only perform well, but are also practically implementable [8], [9]. Starting with the *conventional* parity-check matrix of an RS code under *certain constraints*, each element in the matrix, which belongs to the finite field over which the RS is designed, is replaced by a circulant permutation matrix (CPM) that uniquely represents that element. This replacement procedure, called *dispersion* [2], results in a parity-check matrix of a quasi-cyclic (QC) LDPC code which is referred to as an RS-QC-LDPC code.

It is widely recognized by researchers that short cycles in the Tanner graphs of LDPC codes upon which iterative decoding is performed can degrade performance. There are numerous constructions of LDPC codes with no cycles of length less than six including properly designed RS-QC-LDPC codes, see e.g., [10] for a survey of many such constructions. The literature is also rich in techniques to reduce or eliminate cycles of lengths six or more. These techniques can be broadly classified into two approaches. In the first approach, techniques are proposed to eliminate short cycles from the Tanner graph of a given LDPC code, in particular, a QC-LDPC code, see e.g., [11]–[15]. In the second approach, QC-LDPC codes are constructed subject to certain constraints to ensure that their Tanner graphs have no short cycles, see e.g., [16]–[24].

This work was partially supported by the NSF grants ECCS-1500170 and SaTC-1813401, and gift grants from AppoTech and Nufront. This paper was presented in part in [1] at the Information Theory and Applications Workshop, La Jolla, CA, Feb. 11-16, 2018. 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). 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). W. E. Ryan is with the Zeta Associates Inc. USA (e-mail: bill.ryan.work@gmail.com).

In this paper, we investigate several aspects of RS-QC-LDPC codes, including new designs and constructions, girth, cycle structure, reduction of short cycles in their Tanner graphs, and error performance over the AWGN channel (AWGNC). Our technique for eliminating short cycles belongs to the second aforementioned approach in which constraints are imposed on the design of the codes so that the resulting codes do not have short cycles. However, it differs from most of the previous work in developing and applying the design constraints specifically to the powerful class of RS-QC-LDPC codes. Our approach is similar to that presented in [25] in which the cycle structure was analyzed and a computer search algorithm was proposed to obtain generalized RS-QC-LDPC codes of girth 8. We use masking [26] to further reduce the number of short cycles and to adjust the column and row weights of the parity-check matrices of the constructed codes for performance improvement. Our goal is to propose systematic and flexible constructions of LDPC codes that *balance the requirements of good waterfall performance and low error-floor* while maintaining a structure that facilitates implementation.

The rest of the paper is organized as follows. Section II presents the basic construction of RS-QC-LDPC codes. Section III analyzes cycle structure of the Tanner graph of an RS-QC-LDPC code and presents methods for constructing codes with girth at least 8. Section IV gives two special classes of cycle-8 RS-QC-LDPC codes constructed based on the conventional parity-check matrices of general RS codes. Section V concludes the paper with some comments. Simulation results are presented to demonstrate the error performance of the designed codes. Throughout the paper, performance is measured for BPSK signaling over the AWGNC. Decoding is done using scaled min-sum algorithm (MSA) [10], [27] and simulations are performed on each code, in terms of bit error rate (BER) and block error rate (BLER), to obtain a scaling factor that yields good performance. We use the algorithm in [28] to count cycles.

II. Construction of RS-Based QC-LDPC Codes

In this section, we present the basic construction of binary RS-QC-LDPC codes. The construction of such a code begins with the *conventional* parity-check matrix $\mathbf{B}_{RS}$ of a chosen RS code $\mathcal{C}_{RS}$ over a finite field GF(2^s) which satisfies the following constraint: any 2×2 submatrix of $\mathbf{B}_{RS}$ is *non-singular* (NS). We call such a constraint the 2×2 *submatrix nonsingular* (SNS) constraint, denoted by 2×2 *SNS-constraint*.

Such an RS parity-check matrix $\mathbf{B}_{RS}$ is referred to as a 2×2 SNS-constrained RS parity-check matrix. Once a 2×2 SNS-constrained RS parity-check matrix is constructed, we represent each of its entries by a *circulant permutation matrix* (CPM) of a fixed size, say $n \times n$, where n is a factor of $2^s - 1$. This results in an array $\mathbf{H}_{RS,n}$ of CPMs of size $n \times n$. It is a matrix over GF(2) that has the following structure: *any two rows (or two columns) do not have more than one position in which both have 1-entries*. Such a structure is referred to as *row-column (RC) constraint* [10]. The null space over GF(2) of $\mathbf{H}_{RS,n}$ gives an RS-QC-LDPC code, denoted by $\mathcal{C}_{RS,n,ldpc}$. The RC-constraint on $\mathbf{H}_{RS,n}$ ensures that the Tanner graph of $\mathcal{C}_{RS,n,ldpc}$ has girth at least 6.

A. Construction of 2×2 SNS-Constrained RS Matrices

Let β be an element of order n in GF(2^s) where n is a factor of $2^s - 1$. The set $\mathbf{S}_n = \{1, \beta, \ldots, \beta^{n-1}\}$ forms a cyclic subgroup of GF(2^s). Let d be a positive integer such that $1 \leq d \leq n$. Form the following $d \times n$ matrix over GF(2^s):

$$\mathbf{B}_{RS,n}(d,n) = \left[(\beta^i)^j\right]_{1 \leq i \leq d, 0 \leq j < n}. \quad (1)$$

The null space over GF(2^s) of $\mathbf{B}_{RS,n}(d,n)$ gives a 2^s-ary $(n, n-d, d+1)$ RS code, denoted by $\mathcal{C}_{RS,n}(d,n)$ of length n, dimension $n-d$, rate $(n-d)/n$ and minimum distance $d+1$ [29]. The matrix $\mathbf{B}_{RS,n}(d,n)$ is the parity-check matrix of the RS code $\mathcal{C}_{RS,n}(d,n)$ in the conventional form. Let b be a positive integer such that $d < b < n$. Suppose we delete $n-b$ columns from $\mathbf{B}_{RS,n}(d,n)$. We obtain a $d \times b$ matrix $\mathbf{B}_{RS,n}(d,b)$ over GF(2^s) whose null space gives a $(b, b-d, d+1)$ shortened RS code. We call both $\mathbf{B}_{RS,n}(d,n)$ and $\mathbf{B}_{RS,n}(d,b)$ RS matrices.

In general, the RS matrix $\mathbf{B}_{RS,n}(d,n)$ does not satisfy the 2×2 SNS-constraint. In the following, we give three types of RS matrices that satisfy the 2×2 SNS-constraint.

For the first type, type-1, $n = mk$ is a product of two proper factors m and $k \leq m$. For an integer d such that $1 \leq d \leq k+1$, we form the $d \times m$ matrix over GF(2^s):

$$\mathbf{B}_{RS,n}(d,m) = \left[(\beta^i)^j\right]_{1 \leq i \leq d, 0 \leq j < m}. \quad (2)$$

The null space over GF(2^s) of $\mathbf{B}_{RS,n}(d,m)$ gives an $(m, m-d, d+1)$ shortened code $\mathcal{C}_{RS,n}(d,m)$ of the RS code $\mathcal{C}_{RS,n}(d,n)$. We label the columns of $\mathbf{B}_{RS,n}(d,m)$ from 0 to $m-1$ and the rows from 1 to d.

Theorem 1: The RS matrix $\mathbf{B}_{RS,n}(d,m)$ given by (2) satisfies the 2×2 SNS-constraint.

Proof: Consider a 2×2 submatrix in $\mathbf{B}_{RS,n}(d,m)$ with four entries $(\beta^i)^s, (\beta^i)^t, (\beta^j)^s, (\beta^j)^t$ at the locations $(i,s), (i,t), (j,s)$ and (j,t) with $1 \leq i < j \leq d$ and $0 \leq s < t < m$. If this matrix is singular, then $\beta^{(j-i)(t-s)} = 1$. Since $0 < j - i < d \leq k+1$ and $0 < t - s < m$, the product $(j-i)(t-s)$ is less than $n = km$ and nonzero. Since the order of β is n, $\beta^{(j-i)(t-s)} \neq 1$ and the above 2×2 submatrix of $\mathbf{B}_{RS,n}(d,m)$ is nonsingular. We conclude that $\mathbf{B}_{RS,n}(d,m)$ satisfies the 2×2 SNS-constraint. ∎

For $1 \leq d \leq k+1$, the RS matrix $\mathbf{B}_{RS,n}(d,m)$ given by (2) is a submatrix of $\mathbf{B}_{RS,n}(d,n)$ given by (1). It simply consists of the first m columns of $\mathbf{B}_{RS,n}(d,n)$. In fact, any m *consecutive* columns of $\mathbf{B}_{RS,n}(d,n)$ form a 2×2 SNS-constrained $d \times m$ RS matrix. This can be proved in a similar manner as that given in Theorem 1.

Now we consider the second type, type-2, for which the RS matrix $\mathbf{B}_{RS,n}(d,n)$ in the form given by (1) satisfies the 2×2 SNS-constraint. Suppose n is not a prime. Let p_s be the *smallest prime* factor of n and let d be a positive integer such that $1 \leq d \leq p_s$. Then, the RS matrix $\mathbf{B}_{RS,n}(d,n)$ in the form of (1) satisfies the 2×2 SNS-constraint. The third type, type-3, is that n is a *prime factor* of $2^s - 1$. Let d be a positive integer such that $1 \leq d \leq n$. Then, the RS matrix $\mathbf{B}_{RS,n}(d,n)$ given by (1) satisfies the 2×2 SNS-constraint. The proofs of 2×2 SNS-constraint structure of type-2 and type-3 are similar to that given in Theorem 1.

It is clear that any submatrix of a 2×2 SNS-constrained RS matrix also satisfies the 2×2 SNS-constraint. Besides the above three types, other 2×2 SNS-constrained RS matrices can be constructed by selecting columns from the general RS matrix $\mathbf{B}_{RS,n}(d,n)$ given by (1) under *certain constraints*. This will be presented in Section IV.

B. CPM-Dispersions of Elements of a Finite Field

Consider the cyclic subgroup $\mathbf{S}_n = \{1, \beta, \ldots, \beta^{n-1}\}$ of GF(2^s). For $0 \leq i < n$, we represent the element β^i by a CPM, denoted by $CPM_n(\beta^i)$, over GF(2) of size $n \times n$ (with rows and columns labeled from 0 to $n-1$), whose *generator* (the top row) has the unit-element "1" of GF(2^s) as its *single nonzero component* at position i. The representation of the element β^i by $CPM_n(\beta^i)$ is *unique* and the mapping between β^i and $CPM_n(\beta^i)$ is *one-to-one*. This matrix representation of β^i is referred to as the $n \times n$ *CPM-dispersion* of β^i with respect to the cyclic subgroup $\mathbf{S}_n$ [4]. We represent the 0-element of GF(2^s) by a *zero matrix* (ZM) of size $n \times n$.

Since n is a factor of $2^s - 1$, there is some l for which ln divides $2^s - 1$. Let δ be an element in GF(2^s) of order ln, i.e., $\delta^{ln} = 1$. Then, β can be expressed as the l-th power of δ, i.e., $\beta = \delta^l$. Let $\mathbf{W}_{ln}$ be the cyclic subgroup of GF(2^s) of order ln generated by the powers of δ. Then, $\mathbf{W}_{ln}$ contains $\mathbf{S}_n$ as a subgroup. If we disperse each element in $\mathbf{W}_{ln}$ by a CPM of size $ln \times ln$ as described above, then the element $\beta^i = \delta^{il}$, $0 \leq i < n$, as an element in $\mathbf{W}_{ln}$, is dispersed into a CPM of size $ln \times ln$, denoted by $CPM_{ln}(\beta^i)$, whose generator has its single 1-component at position il. In this case, every element β^i in $\mathbf{S}_n$ is uniquely dispersed into a CPM of size $ln \times ln$ which is referred to as the $ln \times ln$ CPM-dispersion of β^i with respect to the super group $\mathbf{W}_{ln}$ of $\mathbf{S}_n$. Clearly, this CPM-dispersion of each element β^i, $0 \leq i < n$, in $\mathbf{S}_n$ with respect to the super group $\mathbf{W}_{ln}$ is unique and the mapping between β^i and $CPM_{ln}(\beta^i)$ is one-to-one with respect to $\mathbf{W}_{ln}$. The number ln is called the *dispersion factor*. Therefore, each element in $\mathbf{S}_n$ can be one-to-one dispersed into a CPM of a size equal to the order of a cyclic subgroup of GF(2^s) which contains $\mathbf{S}_n$ as a subgroup (including $\mathbf{S}_n$ itself).

C. Construction of RS-QC-LDPC Codes

Any of the three types of 2×2 SNS-constrained RS matrices given in Section II.A can be used to construct RS-QC-LDPC codes.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2916605, IEEE Transactions on Communications

Consider the 2×2 SNS-constrained $d \times m$ RS matrix $\mathbf{B}_{RS,n}(d,m)$ over GF(2^s) in the form of (2) constructed based on a cyclic subgroup $\mathbf{S}_n = \{1, \beta, \ldots, \beta^{n-1}\}$ of GF(2^s). We form an $lnd \times lnm$ matrix $\mathbf{H}_{RS,ln}(d,m)$ over GF(2) by dispersing each entry in $\mathbf{B}_{RS,n}(d,m)$ into a binary CPM of size $ln \times ln$ with respect to a cyclic subgroup $\mathbf{W}_{ln}$ of GF(2^s) of order ln that contains $\mathbf{S}_n$ as a subgroup. The matrix $\mathbf{H}_{RS,ln}(d,m)$ is a $d \times m$ array of CPMs of size $ln \times ln$. This array $\mathbf{H}_{RS,ln}(d,m)$ is called the $ln \times ln$ *CPM-dispersion* of $\mathbf{B}_{RS,n}(d,m)$, denoted by $CPM_{ln}(\mathbf{B}_{RS,n}(d,m))$. Typically, $ln >> 1$, and $\mathbf{H}_{RS,ln}(d,m)$ is a sparse matrix with column and row weights d and m, respectively.

The array $\mathbf{H}_{RS,ln}(d,m)$ consists of d (m) *row* (*column*)*-blocks* of CPMs. Each CPM row (column)-block of $\mathbf{H}_{RS,ln}(d,m)$ consists of ln rows (columns). Each row (column) in a CPM row (column)-block contains m (d) 1-entries which reside in m (d) *separate* CPMs in the CPM row (column)-block, one in each. The m CPMs in each CPM row-block of $\mathbf{H}_{RS,ln}(d,m)$ are distinct, and the d CPMs in each CPM column-block of $\mathbf{H}_{RS,ln}(d,m)$, except the 0-th column-block, are distinct.

The null space over GF(2) of $\mathbf{H}_{RS,ln}(d,m)$ gives a (d,m)-regular RS-QC-LDPC code, denoted by $\mathcal{C}_{RS,ln,ldpc}(d,m)$, of length lnm. The rate of $\mathcal{C}_{RS,ln,ldpc}(d,m)$ is *at least* $(m-d)/m$ which is the rate of the RS code $\mathcal{C}_{RS,n}(d,m)$. Since the RS-QC-LDPC code $\mathcal{C}_{RS,ln,ldpc}(d,m)$ is constructed by CPM-dispersion of the RS matrix $\mathbf{B}_{RS,n}(d,m)$, we call $\mathbf{B}_{RS,n}(d,m)$ the *base matrix*.

The following necessary and sufficient condition on a base matrix $\mathbf{B}$ over GF(2^s) for which the CPM-dispersion of $\mathbf{B}$ satisfies the RC-constraint, i.e., the Tanner graph associated with the CPM-dispersion of $\mathbf{B}$ to have a girth of at least 6, is proved in [30, Corollary 1].

Theorem 2: Let $\mathbf{B}$ be a matrix over GF(2^s) and $\mathcal{C}$ be the QC-LDPC code given by the null space over GF(2) of the CPM-dispersion of $\mathbf{B}$. A necessary and sufficient condition for the Tanner graph of $\mathcal{C}$ to have girth of at least 6 is that every 2×2 submatrix of $\mathbf{B}$ either contains *at least one zero entry* or is nonsingular.

We refer to the necessary and sufficient condition in Theorem 2 as the 2×2 *submatrix (SM) constraint*. Notice that it includes the 2×2 SNS-constraint as a special case. Since the RS matrix $\mathbf{B}_{RS,n}(d,m)$ given by (2) satisfies the 2×2 SNS-constraint, it satisfies the 2×2 SM-constraint. As a result, the Tanner graph of the RS-QC-LDPC code $\mathcal{C}_{RS,ln,ldpc}(d,m)$ has girth of at least 6.

In [30, Corollary 2], the following necessary and sufficient condition on a base matrix $\mathbf{B}$ for which the Tanner graph associated with the its CPM-dispersion has girth at least 8, is also proved.

Theorem 3: Let $\mathbf{B}$ be a matrix over GF(2^s) and $\mathcal{C}$ be the QC-LDPC code given by the null space over GF(2) of the CPM-dispersion of $\mathbf{B}$. A necessary and sufficient condition for the Tanner graph of $\mathcal{C}$ to have girth of at least 8 is that no 2×2 or 3×3 submatrix of $\mathbf{B}$ has two identical non-zero terms in its determinant expansion.

For convenience, we refer to the necessary and sufficient condition given in Theorem 3 as the $2 \times 2/3 \times 3$ SM-constraint.

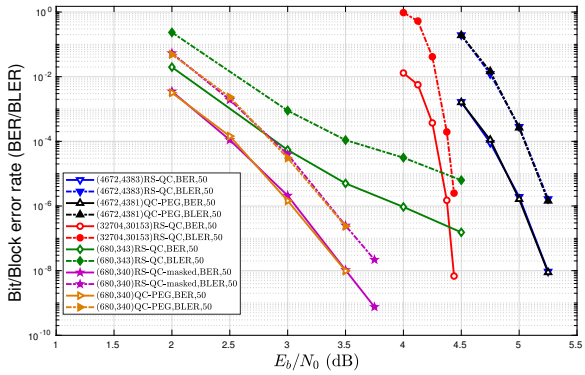

Fig. 1. The BER and BLER performances of the four RS-QC-LDPC codes given in Examples 1 and 2.

By imposing this constraint on base matrices, RS-QC-LDPC codes whose Tanner graphs have girth at least 8 will be presented in Sections III and IV.

Construction of RS-QC-LDPC codes whose Tanner graphs have girth at least 6 using RS matrices of type-2 and type-3 as base matrices is exactly the same as that using the RS matrix $\mathbf{B}_{RS,n}(d,m)$ in the form of (2) of type-1 as the base matrix.

Example 1: Suppose we use the field GF(2^9) for code construction. Note that $2^9 - 1 = 511$ can be factored as the product of two primes, 7 and 73. Setting $n = 73$ and $d = 4$, we construct a 4×73 RS matrix $\mathbf{B}_{RS,73}(4,73)$ as given in (1) where β is an element in GF(2^9) of order 73. The RS matrix $\mathbf{B}_{RS,73}(4,73)$ satisfies the 2×2 SNS-constraint from the type-3 construction. Deleting the last 9 columns from $\mathbf{B}_{RS,73}(4,73)$, we obtain a 4×64 RS matrix $\mathbf{B}_{RS,73}(4,64)$. The 73×73 CPM-dispersion of $\mathbf{B}_{RS,73}(4,64)$ results in a 4×64 array $\mathbf{H}_{RS,73}(4,64)$ of CPMs of size 73×73 which is a 292×4672 matrix of rank 289 with column and row weights 4 and 64, respectively. The null space over GF(2) of $\mathbf{H}_{RS,73}(4,64)$ gives a $(4,64)$-regular $(4672,4383)$ RS-QC-LDPC code $\mathcal{C}_{RS,73,ldpc}(4,64)$ of rate 0.938. Its associated Tanner graph is of girth 6 with 1,022,876 and 167,500,398 cycles of lengths 6 and 8, respectively. The BER and BLER performances of the code $\mathcal{C}_{RS,73,ldpc}(4,64)$ decoded with 50 iterations of MSA scaled by a factor of 0.70 are shown in Fig. 1. The code achieves a BER of 10^{-8} at an SNR of 5.25 dB without a visible error-floor and at the BER of 10^{-8}, the code performs 1.35 dB from the Shannon limit (3.906 dB) and 0.911 dB from its threshold (4.339 dB).

For comparison, a QC $(4672,4381)$ LDPC code is constructed using the *PEG-algorithm* for QC codes [33]. The parity-check matrix of the QC-PEG code has constant column weight 4 but three different row weights, 63, 64, and 65. Its Tanner graph has girth 6 with 1,136,464 cycles of length 6 and 162,030,946 cycles of length 8. The BER and BLER performances of the $(4672,4381)$ QC-PEG code decoded with 50 iterations of MSA scaled by a factor of 0.70 are also shown in Fig. 1. We see that the performances of the $(4672,4383)$ RS-QC-LDPC code and the $(4672,4381)$ QC-PEG code overlap with each other in the range of simulation.

Setting $n = 511, k = 7$ and $m = 73$, we construct a

2×2 SNS-constrained 8×73 RS matrix $\mathbf{B}_{RS,511}(8,73)$ in the form of (2) using type-1 construction. Any submatrix of $\mathbf{B}_{RS,511}(8,73)$ can be used as a base matrix for constructing an RS-QC-LDPC code. In this case, the dispersion factor must be 511. Suppose we set $d = 5, b = 64$ and take a 5×64 submatrix $\mathbf{B}_{RS,511}(5,64)$ from $\mathbf{B}_{RS,511}(8,73)$ by deleting the last 9 columns and last 3 rows of $\mathbf{B}_{RS,511}(8,73)$. The null space over GF(2) of the 511×511 CPM-dispersion of $\mathbf{B}_{RS,511}(5,64)$ gives a $(5,64)$-regular $(32704, 30153)$ RS-QC-LDPC code $\mathcal{C}_{RS,511,ldpc}(5,64)$ of rate 0.922. The BER and BLER performances of the code $\mathcal{C}_{RS,511,ldpc}(5,64)$ decoded with 50 iterations of the MSA scaled by a factor of 0.70 are also shown in Fig. 1. At the BER of 10^{-8}, the code performs about 0.84 dB from the Shannon limit (3.6 dB) and about 0.27 dB from its threshold (4.171 dB).

D. Masking

Masking was first proposed in [26] and further investigated in many papers, see, e.g., [5] and the references cited therein. It is a technique for removing short cycles and/or enlarging the girth of the Tanner graph of a QC-LDPC code constructed by CPM-dispersion of a base matrix. It is also useful in adjusting the column weights in a parity check matrix of a QC-LDPC codes composed of CPMs to improve performance. If properly applied, masking may increase the minimum distance of the code. It also reduces the number of wires in hardware implementation of the decoder.

Masking an RS base matrix $\mathbf{B}_{RS,n}(d,m) = [(\beta^i)^j]_{1\le i\le d, 0\le j<m}$ can be modeled mathematically as follows. Let $\mathbf{Z}(d,m) = [z_{i,j}]_{1\le i\le d, 0\le j<m}$ be a $d \times m$ matrix with the zero and unit elements of GF(2^s) as entries. Define the following product of $\mathbf{Z}(d,m)$ and $\mathbf{B}_{RS,n}(d,m)$: $\mathbf{B}_{RS,n,mask}(d,m) = \mathbf{Z}(d,m) \otimes \mathbf{B}_{RS,n}(d,m) = [z_{i,j}\beta^{ij}]_{1\le i\le d, 0\le j<m}$, where $z_{i,j}\beta^{ij} = \beta^{ij}$ if $z_{i,j} = 1$, and $z_{i,j}\beta^{ij} = 0$ if $z_{i,j} = 0$. In this matrix product, the nonzero entries in $\mathbf{B}_{RS,n}(d,m)$ at the locations corresponding to zero entries in $\mathbf{Z}(d,m)$ are replaced (or *masked*) by zeros. The binary $ln \times ln$ CPM-dispersion of $\mathbf{B}_{RS,n,mask}(d,m)$ gives a $d \times m$ masked array, denoted by $\mathbf{H}_{RS,ln,mask}(d,m)$, of CPMs and ZMs of size $ln \times ln$. We call $\mathbf{Z}(d,m)$ and $\mathbf{B}_{RS,n,mask}(d,m)$ the *masking matrix* and the *masked base matrix*, respectively. The null space of $\mathbf{H}_{RS,ln,mask}(d,m)$ gives a (masked) RS-QC-LDPC code, denoted by $\mathcal{C}_{RS,ln,mask,ldpc}(d,m)$.

Example 2: In this example, we construct an RS-QC-LDPC code of rate 1/2 using the field GF(2^8). First, we factor $2^8 - 1$ as the product of three primes 3, 5 and 17. Set $n = 5 \times 17 = 85$ and $d = 4$. (Note that 5 is the smallest prime factor of $n = 85$.) Using type-2 construction, we first construct a 2×2 SNS-constrained 4×85 RS matrix $\mathbf{B}_{RS,85}(4,85)$ as given in (1) where β is an element in GF(2^8) of order 85. We can take any submatrix of $\mathbf{B}_{RS,85}(4,85)$ as a base matrix and mask it to produce a masked RS base matrix for constructing an RS-QC-LDPC code whose Tanner graph may have girth larger than 6 and much smaller number of short cycles. In addition, the resulting RS-QC-LDPC code constructed from the masked base matrix has larger minimum distance and exhibits better performance compared to the RS-QC-LDPC code constructed from the unmasked base matrix.

Label the columns of $\mathbf{B}_{RS,85}(4,85)$ from 0 to 84. Suppose we choose the columns labeled by 2, 5, 7, 13, 20, 31, 48 and 54 from the mother matrix $\mathbf{B}_{RS,85}(4,85)$ and form a 4×8 RS submatrix $\mathbf{B}_{RS,85}(4,8)$. We find that $\mathbf{B}_{RS,85}(4,8)$ satisfies the necessary and sufficient condition given in Theorem 3, i.e., the $2 \times 2/3 \times 3$ SM-constraint. (How to choose these columns will be discussed in Section IV.) Choose 85 as the dispersion factor. The 85×85 CPM-dispersion of $\mathbf{B}_{RS,85}(4,8)$ gives a 4×8 array $\mathbf{H}_{RS,85}(4,8)$ of CPMs of size 85×85. It is a 340×680 matrix over GF(2) with column and row weights 4 and 8, respectively. The rank of $\mathbf{H}_{RS,85}(4,8)$ is 337 and hence, $\mathbf{H}_{RS,85}(4,8)$ has 3 redundant rows. Its null space over GF(2) gives a $(4,8)$-regular $(680, 343)$ RS-QC-LDPC code $\mathcal{C}_{RS,85,ldpc}(4,8)$ of rate 0.5044. Since its RS base matrix $\mathbf{B}_{RS,85}(4,8)$ satisfies the $2 \times 2/3 \times 3$ SM-constraint, the Tanner graph $\mathcal{G}_{RS,85,ldpc}(4,8)$ of $\mathcal{C}_{RS,85,ldpc}(4,8)$ has girth at least 8. The numbers of cycles of lengths 8, 10, 12 and 14 in $\mathcal{G}_{RS,85,ldpc}(4,8)$ are 32,810, 386,240, 7,256,535, 128,090,240, respectively. We see that $\mathcal{G}_{RS,85,ldpc}(4,8)$ has girth 8 and in total 135,765,825 (short) cycles of lengths 8, 10, 12, and 14.

The BER and BLER performances of the $(680, 343)$ RS-QC-LDPC code $\mathcal{C}_{RS,85,ldpc}(4,8)$ decoded with 50 iterations of the MSA with a scaling factor of 0.7 are shown in Fig. 1. We see that the code performs poorly and starts to have an error-floor below the BLER of 10^{-4} even though its Tanner graph has girth 8. This poor error performance is caused by the large number of short cycles and a small minimum distance of 10 computed using the improved impulse method [31].

Suppose we mask the RS matrix $\mathbf{B}_{RS,85}(4,8)$ with the following 4×8 masking matrix over GF(2) with column and row weights 3 and 6, respectively, which is proposed in [5], [32]:

$$\mathbf{Z}(4,8) = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}. \quad (3)$$

We obtain a 4×8 masked RS matrix $\mathbf{B}_{RS,85,mask}(4,8)$ with column and row weights 3 and 6, respectively. Dispersing each non-zero entry in $\mathbf{B}_{RS,85,mask}(4,8)$ into a CPM of size 85×85 and a zero entry into a ZM of size 85×85, we obtain a 4×8 masked array $\mathbf{H}_{RS,85,mask}(4,8)$ of CPMs and ZMs of size 85×85. The array $\mathbf{H}_{RS,85,mask}(4,8)$ is a 340×680 matrix with column and row weights 3 and 6, respectively. It is a full-rank matrix with rank 340. Note that masking removes the redundant rows of the unmasked array $\mathbf{H}_{RS,85}(4,8)$. The null space over GF(2) of $\mathbf{H}_{RS,85,mask}(4,8)$ gives a $(3,6)$-regular $(680, 340)$ RS-QC-LDPC code $\mathcal{C}_{RS,85,mask,ldpc}(4,8)$ of rate 0.5. The numbers of cycles of lengths 8, 10, 12, and 14 in the Tanner graph $\mathcal{G}_{RS,85,mask,ldpc}(4,8)$ of the masked code $\mathcal{C}_{RS,85,mask,ldpc}(4,8)$ are 1,020, 9,945, 85,170, 720,970, respectively. The masked Tanner graph $\mathcal{G}_{RS,85,mask,ldpc}(4,8)$ also has girth 8 but the number of cycles of length 8 is only 1,020 which is much smaller than the 32,810 cycles of length 8 in the unmasked Tanner graph $\mathcal{G}_{RS,85,ldpc}(4,8)$. The total number of cycles of lengths 8, 10, 12 and 14 in

$\mathcal{G}_{RS,85,mask,ldpc}(4,8)$ is 817,105. Comparing the short cycle distributions of the unmasked Tanner graph $\mathcal{G}_{RS,85,ldpc}(4,8)$ and the masked Tanner graph $\mathcal{G}_{RS,85,mask,ldpc}(4,8)$, we find that there is a very large reduction in short cycles from 135,765,825 to 817,105, *a reduction by a factor of almost* 166.

The BER and BLER performances of the $(680,340)$ $\mathcal{C}_{RS,85,mask,ldpc}(4,8)$ decoded with 50 iterations of the MSA with a scaling factor of 0.75 are also shown in Fig. 1. We see that the $(680,340)$ masked code $\mathcal{C}_{RS,85,mask,ldpc}(4,8)$ performs much better than the $(680,343)$ unmasked code $\mathcal{C}_{RS,85,ldpc}(4,8)$. The masked code achieves a BER of 10^{-9} without a visible error-floor. At the BER of 10^{-9}, it performs 2.6 dB from the threshold (1.1 dB) for rate 0.5. From [31], we determine that the minimum distance of the masked code $\mathcal{C}_{RS,85,mask,ldpc}(4,8)$ is exactly 34. Thus, in this example, *masking more than triples the minimum distance of the code*.

For comparison, a QC $(680,340)$ LDPC code is constructed using the PEG-algorithm for QC codes. Its parity-check matrix has regular column and row weights 3 and 6, respectively. Its Tanner graph has girth 8 with 595, 10,455, 90,865, and 715,020 cycles of lengths 8, 10, 12, and 14, respectively. Using the impulse method in [31], it follows that the minimum distance of the QC-PEG code is at most 30. The BER and BLER performances of the $(680,340)$ QC-PEG code decoded with 50 iterations of MSA scaled by a factor of 0.75 are also shown in Fig. 1. We see that the performances of the $(680,340)$ masked RS-QC-LDPC code $\mathcal{C}_{RS,85,mask,ldpc}(4,8)$ and the PEG code overlap with each other down to the BER of 10^{-8} (BLER of 10^{-7}).

Example 3: In this example, we construct a high rate LDPC code using the field GF(2^{11}). First, we factor $2^{11}-1$ as the product of two primes 89 and 23. We construct a 2×2 SNS-constrained 8×89 matrix, $\mathbf{B}_{RS,89}(8,89)$, using type-3 construction with $n=89$ and $d=8$. Then we use the 64 columns in $\mathbf{B}_{RS,89}(8,89)$ with labels 1, 2, 4, 5, 7, 8, 10, 12, 13, 14, 15, 17, 18, 19, 21, 22, 24, 25, 26, 27, 29, 30, 32, 33, 34, 35, 37, 38, 39, 40, 41, 43, 44, 46, 47, 48, 50, 51, 52, 54, 55, 57, 58, 60, 61, 62, 64, 65, 66, 68, 69, 71, 72, 74, 75, 76, 79, 80, 81, 83, 84, 86, 87, 88 to form an 8×64 matrix $\mathbf{B}_{RS,89}(8,64)$ as a submatrix of $\mathbf{B}_{RS,89}(8,89)$. In addition, we design an 8×64 masking matrix $\mathbf{Z}$, consisting of eight 8×8 circulant matrices $\mathbf{Z}_i$ in order with generators $\mathbf{g}_i$, $0\le i<8$, given by

$$\begin{aligned}
\mathbf{g}_0&=[1,0,1,0,1,1,0,0], \quad \mathbf{g}_1=[0,1,0,1,0,0,1,1],\\
\mathbf{g}_2&=[0,0,1,0,1,1,0,1], \quad \mathbf{g}_3=[0,1,0,1,1,1,0,0],\\
\mathbf{g}_4&=[0,1,1,0,1,1,0,0], \quad \mathbf{g}_5=[0,0,0,1,1,1,0,1],\\
\mathbf{g}_6&=[0,0,0,1,1,1,1,0], \quad \mathbf{g}_7=[0,1,1,0,0,1,1,0].
\end{aligned}$$

The masked submatrix $\mathbf{B}_{RS,89,mask}(8,64)$ has regular column and row weights of 4 and 32, respectively. The CPM dispersion of $\mathbf{B}_{RS,89,mask}(8,64)$ with dispersion factor of 89 gives a 712×5696 parity check matrix $\mathbf{H}_{RS,89,mask}(8,64)$ of rank 711 and regular column and row weights of 4 and 32, respectively. Its null space over GF(2) gives a $(4,32)$-regular $(5696,4985)$ RS-QC-LDPC code $\mathcal{C}_{RS,89,mask,ldpc}(8,64)$ of rate 0.875. Fig. 2 shows the BER and the BLER performances of the code decoded with 50 iterations of SPA. For comparison, the figure also includes the BER performance of the RJA $(5792,5068)$ LDPC code of the same rate 0.875 presented in [34]. The $\mathcal{C}_{RS,89,mask,ldpc}(8,64)$ code has a lower error-floor than the RJA code.

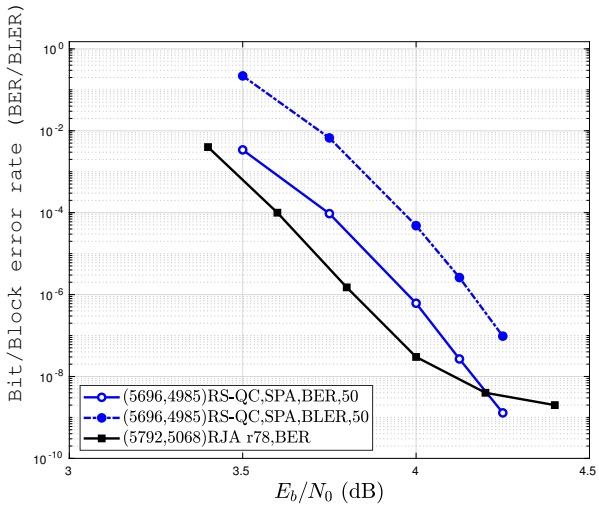

Fig. 2. The BER and BLER performances of the RS-QC-LDPC code given in Example 3

III. RS-Based QC-LDPC Codes with Girth at Least Eight

Theorem 2 implies that the 2×2 SNS-constraint on an RS matrix over GF(2^s) only ensures that the Tanner graph of the RS-QC-LDPC code constructed based on the CPM-dispersion of the RS matrix has girth at least 6. Theorem 3 gives a necessary and sufficient condition on a base matrix whose CPM-dispersion gives a QC-LDPC code with girth at least 8 but it does not provide *a specific method* for constructing such a code. Testing all the 2×2 and 3×3 submatrices of a base RS matrix $\mathbf{B}_{RS,n}(d,n)$ for large d and n is either practically impossible or requires a very large number of computations over GF(2^s). In this section, we develop conditions on the selection of columns from a 2×2 SNS-constrained RS matrix to form an RS submatrix to meet the $2\times 2/3\times 3$ SM-constraint given in Theorem 3. Hence, the null space of its CPM-dispersion gives an RS-QC-LDPC code whose Tanner graph has girth at least 8.

To develop the girth-8 conditions on the columns of an RS matrix, we start with the 2×2 SNS-constrained $d\times m$ RS matrix $\mathbf{B}_{RS,n}(d,m)$ over GF(2^s) given by (2), the type-1 construction. All the developments and results also apply to type-2 and type-3 constructions of 2×2 SNS-constrained RS base matrices.

Label the columns of $\mathbf{B}_{RS,n}(d,m)$ from 0 to $m-1$. Let $d\le t\le m$. Suppose we take a set $\Lambda_t=\{l_1,l_2,\ldots,l_t\}$ of labels of t columns in $\mathbf{B}_{RS,n}(d,m)$ with $0\le l_1<l_2<\cdots<l_t<m\le n$ and form the following $d\times t$ submatrix of $\mathbf{B}_{RS,n}(d,m)$:

$$\mathbf{B}_{RS,n,\Lambda_t}(d,t)=\left[\left(\beta^i\right)^{l_j}\right]_{1\le i\le d,1\le j\le t}. \tag{4}$$

The $n\times n$ CPM-dispersion of $\mathbf{B}_{RS,n,\Lambda_t}(d,t)$ gives a $d\times t$ array $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ of CPMs of size $n\times n$ which is a $dn\times tn$

matrix over GF(2) with column and row weights d and t, respectively. The null space of $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ gives a (d,t)-regular RS-QC-LDPC code $\mathcal{C}_{RS,n,\Lambda_t,ldpc}(d,t)$ of length tn. Since $\mathbf{B}_{RS,n,\Lambda_t}(d,t)$ satisfies the 2×2 SNS-constraint, the array $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$, as a matrix, satisfies the RC-constraint and its associated Tanner graph has girth at least 6.

The array $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ is composed of t CPM column-blocks, each consisting of d CPMs of size $n\times n$. Each CPM column-block consists of n columns (labeled from 0 to $n-1$), each containing d 1-entries residing in d *separate* CPMs. For a column-label $l_j, 1\le j\le t$, in Λ_t, let $Col(l_j)$ denote the CPM column-block of $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ which is the CPM-dispersion of the column in $\mathbf{B}_{RS,n,\Lambda_t}(d,t)$ labeled by l_j. Then, $Col(l_j)=\Big[CPM_n\big(\beta^{l_j}\big)^T, CPM_n\big(\beta^{2l_j}\big)^T,\ldots, CPM_n\big(\beta^{dl_j}\big)^T\Big]^T$ and $\mathbf{H}_{RS,n,\Lambda_t}(d,t)=[Col(l_1), Col(l_2),\ldots, Col(l_t)]$.

Consider the k-th column $\mathbf{c}_k$ of $Col(l_j)$ with $0\le k<n$. For $1\le i\le d$, the i-th 1-entry in $\mathbf{c}_k$ is located at the position $(i-1)_n+(k-il_j)_n$ where $(x)_n$ denotes the least nonnegative integer equal to x modulo n. The d-tuple

$$\begin{aligned}\mathbf{Loc}(l_j,k)=((k-l_j)_n, n+(k-2l_j)_n,\\ 2n+(k-3l_j)_n,\ldots,\\ (d-1)n+(k-dl_j)_n)\end{aligned}$$

specifies the locations of the d 1-entries in the k-th column $\mathbf{c}_k$ of $Col(l_j)$. We call this d-tuple, $\mathbf{Loc}(l_j,k)$, the 1-entry *location-vector* of the k-th column $\mathbf{c}_k$ of $Col(l_j)$.

Let $\mathcal{G}_{RS,n,\Lambda_t,ldpc}(d,t)$ be the Tanner graph associated with the parity-check matrix $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$. Notice that no two VNs corresponding to two columns in the same CPM column-block are connected to the same CN and, similarly, no two CNs corresponding to two rows in the same CPM row-block are connected to the same VN. Since $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ satisfies the RC-constraint, $\mathcal{G}_{RS,n,\Lambda_t,ldpc}(d,t)$ has girth at least 6. In the following, we first investigate the scenarios under which cycles of length 6, called 6-cycles, exist in $\mathcal{G}_{RS,n,\Lambda_t,ldpc}(d,t)$ and then develop the conditions on the labels of Λ_t under which $\mathcal{G}_{RS,n,\Lambda_t,ldpc}(d,t)$ does not contain 6-cycles so that the girth of $\mathcal{G}_{RS,n,\Lambda_t,ldpc}(d,t)$ is at least 8.

Suppose $\mathcal{G}_{RS,n,\Lambda_t,ldpc}(d,t)$ contains a 6-cycle, denoted by C_6. This cycle consists of three VNs, denoted by v_1,v_2,v_3 and three CNs, denoted by c_1,c_2,c_3. Then, v_1,v_2,v_3 correspond to three columns in three distinct CPM column-blocks of $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$, say, $Col(l_{i_1}), Col(l_{i_2}), Col(l_{i_3})$, respectively, where without loss of generality, we assume $1\le i_1<i_2<i_3\le t$, which implies that $l_{i_1}<l_{i_1}<l_{i_3}$. Suppose v_1,v_2,v_3 correspond to columns k_1,k_2,k_3 in the three CPM column-blocks $Col(l_{i_1}), Col(l_{i_2}), Col(l_{i_3})$, respectively, where $0\le k_1,k_2,k_3<n$. For $j=1,2,3$, the 1-entry location-vector of column k_j is

$$\begin{aligned}\mathbf{Loc}(l_{i_j},k_j)=((k_j-l_{i_j})_n, n+(k_j-2l_{i_j})_n,\\ 2n+(k_j-3l_{i_j})_n,\ldots,\\ (d-1)n+(k_j-dl_{i_j})_n).\end{aligned}$$

Suppose the three CNs c_1,c_2,c_3 on C_6 correspond to the three rows, numbered by r_1,r_2,r_3, in $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$. Assume that $0\le r_1<r_2<r_3<dn$. For $1\le i\le 3$, let $r_i=x_in+(r_i)_n$ where $x_i=\lfloor\frac{r_i}{n}\rfloor<d$ and $0\le(r_i)_n<n$. For $1\le i\le 3$, $(r_i)_n$ and x_i simply indicate that the row r_i is the row labeled by $(r_i)_n$ in the x_i-th CPM row-block of $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$. Since the rows r_1,r_2,r_3 are in different CPM row-blocks and $r_1<r_2<r_3$, it follows that $x_1<x_2<x_3$.

The 6-cycle C_6 has six possible configurations, denoted by $C_{6,\tau}, 1\le\tau\le 6$, which are in the following form:

$$\begin{aligned}C_{6,1}&=(v_1\to c_1\to v_2\to c_2\to v_3\to c_3\to v_1),\\ C_{6,2}&=(v_1\to c_1\to v_2\to c_3\to v_3\to c_2\to v_1),\\ C_{6,3}&=(v_1\to c_1\to v_3\to c_3\to v_2\to c_2\to v_1),\\ C_{6,4}&=(v_1\to c_2\to v_2\to c_1\to v_3\to c_3\to v_1),\\ C_{6,5}&=(v_1\to c_2\to v_3\to c_1\to v_2\to c_3\to v_1),\\ C_{6,6}&=(v_1\to c_1\to v_3\to c_2\to v_2\to c_3\to v_1),\end{aligned}$$

where $\to$ represents an edge that connects a VN (or CN) to a CN (or VN).

For $1\le\tau\le 6$, each possible configuration $C_{6,\tau}$ of C_6 corresponds to a *unique sequence* $L_{6,\tau}$ of 6 different locations of 1-entries which reside in 6 separate CPMs in $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ with the *ending* location the same as the *starting* location. This sequence $L_{6,\tau}$ is called the *location-sequence* (LS) of 1-entries of the configuration $C_{6,\tau}$. The six 1-entry location-sequences of the 6 possible configurations of C_6 are given below:

$$\begin{aligned}L_{6,1} &= (((k_1-x_1l_{i_1})_n,k_1),((k_2-x_1l_{i_2})_n,k_2),\\ &\quad((k_2-x_2l_{i_2})_n,k_2),((k_3-x_2l_{i_3})_n,k_3),\\ &\quad((k_3-x_3l_{i_3})_n,k_3),((k_1-x_3l_{i_1})_n,k_1),\\ &\quad((k_1-x_1l_{i_1})_n,k_1)),\\ L_{6,2} &= (((k_1-x_1l_{i_1})_n,k_1),((k_2-x_1l_{i_2})_n,k_2),\\ &\quad((k_2-x_3l_{i_2})_n,k_2),((k_3-x_3l_{i_3})_n,k_3),\\ &\quad((k_3-x_2l_{i_3})_n,k_3),((k_1-x_2l_{i_1})_n,k_1),\\ &\quad((k_1-x_1l_{i_1})_n,k_1)),\\ L_{6,3} &= (((k_1-x_1l_{i_1})_n,k_1),((k_3-x_1l_{i_3})_n,k_3),\\ &\quad((k_3-x_3l_{i_3})_n,k_3),((k_2-x_3l_{i_2})_n,k_2),\\ &\quad((k_2-x_2l_{i_2})_n,k_2),((k_1-x_2l_{i_1})_n,k_1),\\ &\quad((k_1-x_1l_{i_1})_n,k_1)),\\ L_{6,4} &= (((k_1-x_2l_{i_1})_n,k_1),((k_2-x_2l_{i_2})_n,k_2),\\ &\quad((k_2-x_1l_{i_2})_n,k_2),((k_3-x_1l_{i_3})_n,k_3),\\ &\quad((k_3-x_3l_{i_3})_n,k_3),((k_1-x_3l_{i_1})_n,k_1),\\ &\quad((k_1-x_2l_{i_1})_n,k_1)),\\ L_{6,5} &= (((k_1-x_2l_{i_1})_n,k_1),((k_3-x_2l_{i_3})_n,k_3),\\ &\quad((k_3-x_1l_{i_3})_n,k_3),((k_2-x_1l_{i_2})_n,k_2),\\ &\quad((k_2-x_3l_{i_2})_n,k_2),((k_1-x_3l_{i_1})_n,k_1),\\ &\quad((k_1-x_2l_{i_1})_n,k_1)),\\ L_{6,6} &= (((k_1-x_1l_{i_1})_n,k_1),((k_3-x_1l_{i_3})_n,k_3),\\ &\quad((k_3-x_2l_{i_3})_n,k_3),((k_2-x_2l_{i_2})_n,k_2),\\ &\quad((k_2-x_3l_{i_2})_n,k_2),((k_1-x_3l_{i_1})_n,k_1),\\ &\quad((k_1-x_1l_{i_1})_n,k_1)).\end{aligned}$$

Consider the 1-LS $L_{6,\tau}, 1\le\tau\le 6$, for the possible configuration $C_{6,\tau}$ of C_6. This 1-LS consists of three pairs of 1-entry locations,$(((k_e-x_fl_{i_e})_n,k_e),((k_{e'}-x_fl_{i_{e'}})_n,k_{e'})), 1\le e,e'\le 3, e\ne e'$ and $1\le f\le 3$. The two 1-entry

locations, $((k_e - x_f l_{i_e})_n, k_e)$ and $((k_{e'} - x_f l_{i_{e'}})_n, k_{e'})$, in each pair are in the same row (the $(r_f)_n$-th row) of the x_f-th CPM row-block of $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ and hence their first coordinates must be equal. This results in three equalities: $(k_e - x_f l_{i_e})_n = (k_{e'} - x_f l_{i_{e'}})_n$ for $f = 1, 2, 3$. For example, for the configuration $C_{6,1}$, the three equalities are: $(k_1 - x_1 l_{i_1})_n = (k_2 - x_1 l_{i_2})_n$, $(k_2 - x_2 l_{i_2})_n = (k_3 - x_2 l_{i_3})_n$, and $(k_3 - x_3 l_{i_3})_n = (k_1 - x_3 l_{i_1})_n$. Adding these three equalities and with some algebraic manipulations, we get:

$$((x_2 - x_1)(l_{i_2} - l_{i_1}) + (x_3 - x_2)(l_{i_3} - l_{i_1}))_n = 0. \quad (5)$$

Then, the equality given by (5) is the necessary and sufficient condition for the existence of C_6 in configuration $C_{6,1}$.

Similarly, for the other five possible configurations of C_6, we can derive the necessary and sufficient conditions for their existence. They are:

$$\begin{aligned}
&C_{6,2} : ((x_2 - x_1)(l_{i_2} - l_{i_1}) + (x_3 - x_2)(l_{i_2} - l_{i_3}))_n = 0,\\
&C_{6,3} : ((x_2 - x_1)(l_{i_3} - l_{i_1}) + (x_3 - x_2)(l_{i_3} - l_{i_2}))_n = 0,\\
&C_{6,4} : ((x_2 - x_1)(l_{i_3} - l_{i_2}) + (x_3 - x_2)(l_{i_3} - l_{i_1}))_n = 0,\\
&C_{6,5} : ((x_2 - x_1)(l_{i_2} - l_{i_3}) + (x_3 - x_2)(l_{i_2} - l_{i_1}))_n = 0,\\
&C_{6,6} : ((x_2 - x_1)(l_{i_3} - l_{i_1}) + (x_3 - x_2)(l_{i_2} - l_{i_1}))_n = 0.
\end{aligned} \quad (6)$$

Hence, the Tanner graph associated with the parity-check matrix $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ given by the $n \times n$ CPM-dispersion of the RS matrix $\mathbf{B}_{RS,n,\Lambda_t}(d,t)$ in (4) contains cycles of length 6 if and only if there are six integers $i_1, i_2, i_3, x_1, x_2, x_3$, where $1 \le i_1 < i_2 < i_3 \le t$ and $0 \le x_1 < x_2 < x_3 < d$, such that at least one of the 6 equalities given by (5) and (6) holds. From this we deduce the following theorem.

Theorem 4: The Tanner graph associated with the parity-check matrix $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ has girth at least 8 if and only if for any six integers $i_1, i_2, i_3, x_1, x_2, x_3$, where $1 \le i_1 < i_2 < i_3 \le t$ and $0 \le x_1 < x_2 < x_3 < d$,

$$\begin{aligned}
&((x_2 - x_1)(l_{i_2} - l_{i_1}) + (x_3 - x_2)(l_{i_3} - l_{i_1}))_n \neq 0,\\
&((x_2 - x_1)(l_{i_2} - l_{i_1}) + (x_3 - x_2)(l_{i_2} - l_{i_3}))_n \neq 0,\\
&((x_2 - x_1)(l_{i_3} - l_{i_1}) + (x_3 - x_2)(l_{i_3} - l_{i_2}))_n \neq 0,\\
&((x_2 - x_1)(l_{i_3} - l_{i_2}) + (x_3 - x_2)(l_{i_3} - l_{i_1}))_n \neq 0,\\
&((x_2 - x_1)(l_{i_2} - l_{i_3}) + (x_3 - x_2)(l_{i_2} - l_{i_1}))_n \neq 0,\\
&((x_2 - x_1)(l_{i_3} - l_{i_1}) + (x_3 - x_2)(l_{i_2} - l_{i_1}))_n \neq 0.
\end{aligned} \quad (7)$$

From the six inequalities given by (7), we readily see that if the labels of a set $\Lambda_t = \{l_1, l_2, ..., l_t\}$ satisfy the six inequalities, then the labels of the set $\Lambda_t^* = \{(r + l_1)_n, (r + l_2)_n, ..., (r + l_t)_n\}$ obtained by adding an integer $r, 0 \le r < n$, to each label in Λ_t (modulo-n addition) also satisfy the six inequalities given by (7). Using the label set Λ_t^* , we can construct another $d \times t$ RS matrix $\mathbf{B}_{RS,n,\Lambda_t^*}(d,t)$ which satisfies the $2 \times 2/3 \times 3$ SM-constraint. For convenience, we call Λ_t a $2 \times 2/3 \times 3$ *column-label set*.

Theorem 4 tells us how to choose columns from the 2×2 SNS-constrained $d \times m$ RS matrix $\mathbf{B}_{RS,n}(d,m)$ given by (2) to form a $d \times t$ RS matrix $\mathbf{B}_{RS,n,\Lambda_t}(d,t)$ which satisfies the $2 \times 2/3 \times 3$ SM-constraint. Consequently, the Tanner graph associated with the CPM-dispersion $\mathbf{H}_{RS,n,\Lambda_t}(d,t)$ of $\mathbf{B}_{RS,n,\Lambda_t}(d,t)$ has girth at least 8. All the six inequalities given by (7) are expressed in terms of the labels of the CPM column-blocks and row-blocks. Computations required are relatively simple. There are $\binom{n}{t}$ sets of t columns among the n columns in the matrix $\mathbf{B}_{RS,n}(d,n)$. This is the number of column-label sets to be examined until a valid one is found. To check if a set $\{l_1, l_2, ..., l_t\}$ is valid, we need to verify that the six conditions stated in Theorem 4 are satisfied by all tuples of six integers $l_{i_1}, l_{i_2}, l_{i_3}, x_1, x_2, x_3$, where $1 \le i_1 < i_2 < i_3 \le t$ and $0 \le x_1 < x_2 < x_3 < d$. Examining a tuple requires 5 integer subtractions, 6 integer additions, 6 integer multiplications and 6 comparisons. If X is the computational complexity of examining a tuple, then, in the worst case when all tuples are examined before finding a valid one, the total computational complexity is $\mathrm{O}\Big(\binom{n}{t}\binom{t}{3}\binom{d}{3}\Big) X)$.

All the results developed above for 2×2 SNS-constrained RS matrix $\mathbf{B}_{RS,n}(d,m)$ of type-1 apply to 2×2 SNS-constrained RS matrices of type-2 and type-3.

Example 4: In this example, we use the same field GF(2^9) as used in Example 1 to construct a rate-1/2 RS-QC-LDPC code whose Tanner graph has girth 8. First, we factor $2^9 - 1 = 511$ as the product of 7 and 73. Set $n = 511, k = 7$ and $m = 73$. Choose $d = 4$. Using the construction of type-1, we form a 2×2 SNS-constrained 4×73 RS matrix $\mathbf{B}_{RS,511}(4,73)$ as in (2). Label the columns of $\mathbf{B}_{RS,511}(4,73)$ from 0 to 72. From the 73 column labels, we find a set Λ_8 of eight column labels 2, 5, 9, 15, 26, 42, 64, and 72 which satisfy the six inequalities given by (7) in Theorem 4. Hence, the set $\Lambda_8 = \{2, 5, 9, 15, 26, 42, 64, 72\}$ forms a $2 \times 2/3 \times 3$ constrained column-label set. Choose the eight columns from $\mathbf{B}_{RS,511}(4,73)$ which are labeled by 2, 5, 9, 15, 26, 42, 64, and 72 and form a 4×8 RS submatrix $\mathbf{B}_{RS,511,\Lambda_8}(4,8)$ of $\mathbf{B}_{RS,511}(4,73)$. Then, the matrix $\mathbf{B}_{RS,511,\Lambda_8}(4,8)$ satisfies the $2 \times 2/3 \times 3$ SM-constraint.

The 511×511 CPM-dispersion of $\mathbf{B}_{RS,511,\Lambda_8}(4,8)$ gives a 4×8 array $\mathbf{H}_{RS,511,\Lambda_8}(4,8)$ of CPMs of size 511×511 which is a 2044×4088 matrix over GF(2) with column and row weights 4 and 8, respectively. The null space over GF(2) of $\mathbf{H}_{RS,511,\Lambda_8}(4,8)$ gives a (4, 8)-regular (4088, 2047) RS-QC-LDPC code $\mathcal{C}_{RS,511,\Lambda_8,ldpc}(4,8)$ of rate 0.501, slightly higher than 1/2. The Tanner graph $\mathcal{G}_{RS,511,\Lambda_8,ldpc}(4,8)$ of the code has girth 8. The numbers of short cycles of lengths 8, 10, 12, and 14 in $\mathcal{G}_{RS,511,\Lambda_8,ldpc}(4,8)$ are 87,892, 623,420, 12,511,835, and 192,430,366, respectively. The total number of such short cycles in $\mathcal{G}_{RS,511,\Lambda_8,ldpc}(4,8)$ is 205,653,483. We also constructed RS-QC-LDPC codes using other label sets generated randomly. Compared to the code generated using Λ_8, the BER of these codes show higher error-floors. For example, the RS-QC-LDPC codes constructed by the label sets $\{43, 118, 135, 231, 308, 335, 353, 383\}$, $\{50, 65, 143, 280, 324, 417, 463, 467\}$, and $\{19, 216, 336, 405, 434, 468, 478, 491\}$ have similar waterfall performance but show an error-floor of over an order of magnitude higher at SNR of 3dB.

Suppose we mask the RS matrix $\mathbf{B}_{RS,511,\Lambda_8}(4,8)$ with the 4×8 masking matrix given by (3). We obtain a 4×8 masked RS matrix $\mathbf{B}_{RS,511,\Lambda_8,mask}(4,8)$. The 511×511 CPM-dispersion of $\mathbf{B}_{RS,511,\Lambda_8,mask}(4,8)$ gives a 4×8 masked array $\mathbf{H}_{RS,511,\Lambda_8,mask}(4,8)$ of CPMs and ZMs of size 511×511 which is a 2044×4088 full-rank matrix with

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2916605, IEEE Transactions on Communications

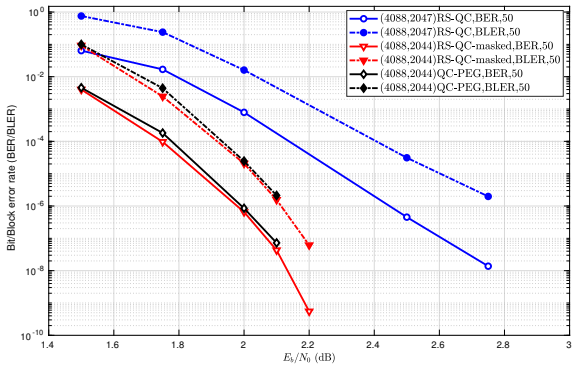

Fig. 3. The BER and BLER performances of the RS-QC-LDPC codes given in Example 4.

column and row weights 3 and 6, respectively. The null space over GF(2) of $\mathbf{H}_{RS,511,\Lambda_8,mask}(4,8)$ gives a $(3,6)$-regular $(4088, 2044)$ RS-QC-LDPC code $\mathcal{C}_{RS,511,\Lambda_8,mask,ldpc}(4,8)$ of rate exactly 1/2. The Tanner graph $\mathcal{G}_{RS,511,\Lambda_8,mask,ldpc}(4,8)$ of the masked code $\mathcal{C}_{RS,511,\Lambda_8,mask,ldpc}(4,8)$ also has girth 8. The numbers of short cycles of lengths 8, 10, 12, and 14 in $\mathcal{G}_{RS,511,\Lambda_8,mask,ldpc}(4,8)$ are 1,022, 14,308, 141,547, and 1,016,890, respectively. The total number of short cycles in $\mathcal{G}_{RS,511,\Lambda_8,mask,ldpc}(4,8)$ is 1,173,767.

Comparing the short cycle distributions of the masked Tanner graph $\mathcal{G}_{RS,511,\Lambda_8,mask,ldpc}(4,8)$ and the unmasked Tanner graph, $\mathcal{G}_{RS,511,\Lambda_8,ldpc}(4,8)$, we see that masking results in an enormous reduction in short cycles. The reduction of the number of cycle-8 is from 87,892 to 1,022, and the reduction of the total number of short cycles is from 205,653,483 to 1,173,767, a reduction by a factor of more than 175.

The BER and BLER performances of $\mathcal{C}_{RS,511,\Lambda_8,ldpc}(4,8)$ and $\mathcal{C}_{RS,511,\Lambda_8,mask,ldpc}(4,8)$, decoded with 50 iterations of the MSA scaled by factors of 0.7 and 0.75, respectively, are shown in Fig. 3. We see that masking results in a significant performance improvement. The masked code achieves a BER of 10^{-9} at an SNR of 2.18 dB and shows no error-floor. At the BER of 10^{-9}, it performs about 1.9 dB from the Shannon limit (0.188 dB) and 1 dB from its threshold (1.1 dB). This performance improvement of the masked code is mainly caused by the large reduction of short cycles and the change of degree distribution from $(4,8)$-regular distribution to $(3,6)$-regular distribution due to masking.

The $(3,6)$-regular masked $(4088, 2044)$ code $\mathcal{C}_{RS,511,\Lambda_8,mask,ldpc}(4,8)$ performs slightly better than the $(4080, 2040)$ QC-LDPC code given in [10, Fig. 11.9, p. 503] and the Euclidean geometry code C_2 of girth 8 in [10, Fig. 12.9, p. 541] decoded with 50 iterations of the sum product algorithm (SPA).

For comparison, a QC $(4088, 2044)$ LDPC code is constructed using the PEG-algorithm for QC codes. Its parity-check matrix has regular column and row weights 3 and 6, respectively. Its Tanner graph has girth 8 with 1,022, 7,665, 89,425, and 760,879 cycles of lengths 8, 10, 12, and 14, respectively. The BER and BLER performances of the $(4088, 2044)$ QC-PEG code decoded with 50 iterations of MSA scaled by a factor of 0.75 are also shown in Fig. 3. The masked $(4088, 2044)$ RS-QC-LDPC code $\mathcal{C}_{RS,511,\Lambda_8,mask,ldpc}(4,8)$ performs slightly better than the QC-PEG code.

Example 5: In this example, we intend to design a rate-2/3 RS-QC-LDPC code using the field GF(2^9). Set $d = 5$. Since d is less than the smallest prime factor 7 of 511, we can use type-2 construction to form a 5×511 RS matrix $\mathbf{B}_{RS,511}(5,511)$ which satisfies the 2×2 SNS-constraint. Label the columns of $\mathbf{B}_{RS,511}(5,511)$ from 0 to 510. Using the constraints given in Theorem 4, we find a set $\Lambda_{15} = \{2, 5, 9, 18, 38, 77, 165, 172, 255, 283, 299, 314, 360, 379, 460\}$ of 15 labels that satisfies the 6 constraints. Next, we form a 5×15 RS matrix $\mathbf{B}_{RS,511,\Lambda_{15}}(5,15)$ using the columns in $\mathbf{B}_{RS,511}(5,511)$ labeled by the numbers in Λ_{15}. Then, $\mathbf{B}_{RS,511,\Lambda_{15}}(5,15)$ satisfies the $2 \times 2/3 \times 3$ SM-constraint.

Form a 5×15 masking matrix $\mathbf{Z}(5,15)$ which consists of three 5×5 circulants whose generators are (1 0 1 0 1), (1 0 0 1 1), and (0 1 1 0 1) in order. Masking $\mathbf{B}_{RS,511,\Lambda_{15}}(5,15)$ with $\mathbf{Z}(5,15)$, we obtain a 5×15 masked RS matrix $\mathbf{B}_{RS,511,\Lambda_{15},mask}(5,15)$ with column and row weights 3 and 9 respectively. Using $\mathbf{B}_{RS,511,\Lambda_{15}}(5,15)$ and $\mathbf{B}_{RS,511,\Lambda_{15},mask}(5,15)$ as base matrices and 511 as the dispersion factor, we can construct a $(5,15)$-regular $(7665, 5114)$ RS-QC-LDPC code $\mathcal{C}_{RS,511,\Lambda_{15},ldpc}(5,15)$ of rate 0.667 and a $(3,9)$-regular $(7665, 5110)$ masked RS-QC-LDPC code $\mathcal{C}_{RS,511,\Lambda_{15},mask,ldpc}(5,15)$ of rate 2/3. The Tanner graphs of both codes have girth 8. The Tanner graph $\mathcal{G}_{RS,511,\Lambda_{15},ldpc}(5,15)$ of the unmasked code $\mathcal{C}_{RS,511,\Lambda_{15},ldpc}(5,15)$ contains 1,635,200 cycles of length 8 and 53,696,902 cycles of length 10, however the Tanner graph $\mathcal{G}_{RS,511,\Lambda_{15},mask,ldpc}(5,15)$ of the masked code $\mathcal{C}_{RS,511,\Lambda_{15},mask,ldpc}(5,15)$ contains 6,132 cycles of length 8 and 107,821 cycles of length 10. We see that masking reduces short cycles of lengths 8 and 10 drastically.

The BER and BLER performances of $\mathcal{C}_{RS,511,\Lambda_{15},ldpc}(5,15)$ and $\mathcal{C}_{RS,511,\Lambda_{15},mask,ldpc}(5,15)$, decoded with 50 iterations of the MSA scaled by 0.7 and 0.75, respectively, are shown in Fig. 4. We see that the masked code outperforms the corresponding unmasked code. At the BER of 10^{-8}, the $(3,9)$-regular $(7665, 5110)$ masked code $\mathcal{C}_{RS,511,\Lambda_{15},mask,ldpc}(5,15)$ performs 1.42 dB from the Shannon limit (1.08 dB) and 0.713 dB from its threshold (1.787 dB).

IV. Two Special Constructions of RS-QC-LDPC Codes with Girth at Least 8

In this section, we first present a special case of the construction given in Section III. Next, we give a new construction of 2×2 SNS-constrained RS matrices based on which we construct $2 \times 2/3 \times 3$ SM-constrained base matrices for RS-QC-LDPC codes whose Tanner graphs have girth at least eight.

Let $d = 4$. Consider a $4 \times t$ RS matrix $\mathbf{B}_{RS,n,\Lambda_t}(4,t)$ in the form of (4) constructed from a mother 2×2 constrained RS matrix $\mathbf{B}_{RS,n}(4,n)$ based on a chosen column-label set Λ_t. Since $0 \leq x_1 < x_2 < x_3 < 4$, there are 4 possibilities for (x_1, x_2, x_3), which are (0, 1, 2), (0, 1, 3), (0, 2, 3), (1, 2, 3). Substituting these four possibilities of (x_1, x_2, x_3) into

the six inequalities given by (7) in Theorem 4, we obtain the following corollary.

Corollary 1: Let $\mathbf{B}_{RS,n,\Lambda_t}(4,t)$ be the $4\times t$ matrix formed by the columns of the RS matrix $\mathbf{B}_{RS,n}(4,n)$ labeled by the numbers in a column-label set Λ_t. Let $\mathbf{H}_{RS,n,\Lambda_t}(4,t)$ be the $n\times n$ CPM-dispersion of $\mathbf{B}_{RS,n,\Lambda_t}(4,t)$. The Tanner graph associated with $\mathbf{H}_{RS,n,\Lambda_t}(4,t)$ has girth at least 8 if and only if the following 9 conditions on the column-labels in Λ_t hold:

$$\begin{array}{ll}(\text{p1})\ n\nmid l_{i_3}-2l_{i_2}+l_{i_1}, & (\text{p2})\ n\nmid l_{i_3}-3l_{i_2}+2l_{i_1},\\ (\text{p3})\ n\nmid 2l_{i_3}-3l_{i_2}+l_{i_1}, & (\text{p4})\ n\nmid l_{i_3}+l_{i_2}-2l_{i_1},\\ (\text{p5})\ n\nmid 2l_{i_3}+l_{i_2}-3l_{i_1}, & (\text{p6})\ n\nmid l_{i_3}+2l_{i_2}-3l_{i_1},\\ (\text{p7})\ n\nmid 2l_{i_3}-l_{i_2}-l_{i_1}, & (\text{p8})\ n\nmid 3l_{i_3}-l_{i_2}-2l_{i_1},\\ (\text{p9})\ n\nmid 3l_{i_3}-2l_{i_2}-l_{i_1}.\end{array} \tag{8}$$

We note that the columns of the 4×8 RS base matrix $\mathbf{B}_{RS,85}(4,8)$ over GF(2^8) used for constructing the rate-1/2 $(680,340)$ RS-QC-LDPC codes in Example 2 are chosen from the mother matrix $\mathbf{B}_{RS,85}(4,85)$ over GF(2^8) with column labels satisfying the nine conditions given by (8). In this example, with a 8-core CPU, g++ compiler, for $d=4, t=8, n=85$, it needs only 1.67ms to find a valid column-label set.

Next, we consider the RS matrix $\mathbf{B}_{RS,n}(d,n)$ over GF(2^s) given in the general form of (1) which is constructed by using the cyclic subgroup $\mathbf{S}_n=\{1,\beta,...,\beta^{n-1}\}$ of GF(2^s) generated by an element β of order n with $d<n$. This RS matrix, in general, does not satisfy the 2×2 SNS-constraint, except for the type-2 and type-3 cases. In the following, we consider a special case for which we can construct 2×2 SNS-constrained matrices by choosing columns from $\mathbf{B}_{RS,n}(d,n)$ under certain conditions. From these 2×2 SNS-constrained matrices, we can construct $2\times 2/3\times 3$ SM-constrained matrices based on Corollary 1.

Let $d=4$. Suppose 3 is a factor of n, i.e., $3|n$. Label the columns of the $4\times n$ RS matrix $\mathbf{B}_{RS,n}(4,n)$ from 0 to $n-1$. Partition the column-labels of $\mathbf{B}_{RS,n}(4,n)$ into $n/3$ disjoint triplets, $(0,n/3,2n/3),(1,1+n/3,1+2n/3),\ldots,(n/3-1,2n/3-1,n-1)$. From each triplet $(i,i+n/3,i+2n/3), 0\le i<n/3$, we take *any one* label j_i and form a set $\Lambda_{n/3}=\{j_0,j_1,\ldots,j_{n/3-1}\}$ with $n/3$ column labels. Next, we select the $n/3$ columns, labeled by $j_0,j_1,\ldots,j_{n/3-1}$, from $\mathbf{B}_{RS,n}(4,n)$ and form a $4\times n/3$ sub-matrix $\mathbf{B}_{RS,n,\Lambda_{n/3}}(4,n/3)$. Following the proof of Theorem 1, it can be shown that $\mathbf{B}_{RS,n,\Lambda_{n/3}}(4,n/3)$ satisfies the 2×2 SNS-constraint. This gives another construction of 2×2 SM-constrained RS matrices.

Using $\mathbf{B}_{RS,n,\Lambda_{n/3}}(4,n/3)$ as the mother matrix and choosing a set of t columns with labels satisfying the nine conditions given by (8), we can construct a $2\times 2/3\times 3$ SM-constrained RS matrix $\mathbf{B}_{RS,n,\Lambda_t}(4,t)$. Then, the null space of the $n\times n$ CPM-dispersion $\mathbf{H}_{RS,n,\Lambda_t}(4,t)$ of $\mathbf{B}_{RS,n,\Lambda_t}(4,t)$ gives an RS-QC-LDPC code with girth at least 8.

In Examples 2 and 4, we showed that the 4×8 masking matrix $\mathbf{Z}(4,8)$ given by (3) is very effective in reducing short cycles of the Tanner graph of an RS-QC-LDPC code. This masking matrix can be used as a *building block* to construct larger masking matrices for larger RS base matrices. This masking matrix has a simple structure. The second pair of columns is a repetition of the first pair of columns and the fourth pair of columns is a repetition of third pair of columns. A simple expansion of this masking matrix is to repeat the first pair and the third pair of columns t times [5], [32]. This expansion results in a $4\times 4t$ masking matrix $\mathbf{Z}(4,4t)$.

If we permute the columns of the $\mathbf{Z}(4,8)$, we can put it into the following form:

$$\mathbf{Z}_c(4,8)=\begin{bmatrix}1&1&1&0&1&1&0&1\\0&1&1&1&1&1&1&0\\1&0&1&1&0&1&1&1\\1&1&0&1&1&0&1&1\end{bmatrix}. \tag{9}$$

The matrix $\mathbf{Z}_c(4,8)$ given by (9) consists of two 4×4 circulants. The second circulant is simply obtained by shifting the rows (or columns) of the first circulant downward (or to the left) one position cyclically. We can use the first circulant as a building block to construct a $4\times 4t$ masking matrix $\mathbf{Z}_c(4,4t)$ by repeating it t times and/or using its downward (or left) cyclic-shifts. The matrix $\mathbf{Z}_c(4,4t)$ has circulant structure and has column and row weights 3 and $3t$, respectively. The subscript "c" in $\mathbf{Z}_c(4,16)$ stands for "circulant". Suppose we set $t=4$ and form the following 4×16 masking matrix with column and row weights 3 and 12, respectively:

$$\mathbf{Z}_c(4,16)=\begin{bmatrix}1&1&1&0&1&1&1&0&1&1&1&0&1&1&0&1\\0&1&1&1&0&1&1&1&0&1&1&1&1&1&1&0\\1&0&1&1&1&0&1&1&1&0&1&1&0&1&1&1\\1&1&0&1&1&1&0&1&1&1&0&1&1&0&1&1\end{bmatrix}. \tag{10}$$

In forming the matrix of $\mathbf{Z}_c(4,16)$, we use the first circulant of $\mathbf{Z}_c(4,8)$ three times, and its first downward cyclic-shift once.

Example 6: In this example, we use the field GF(2^8) to construct an RS-QC-LDPC code of rate 3/4. First, we construct a 4×255 RS matrix $\mathbf{B}_{RS,255}(4,255)$ over GF(2^8) in the form of (1) which does not satisfy the 2×2 SNS-constraint. Label the columns of $\mathbf{B}_{RS,255}(4,255)$ from 0 to 254. Since 3 is a factor of 255, we can partition the column numbers of $\mathbf{B}_{RS,255}(4,255)$ into 85 disjoint triplets, $(i,i+85,i+170), 0\le i<85$. For $4\le t\le 85$, suppose we choose a set of t triplets. From each triplet of this set, we take one number. This results in a set $\Lambda_t=\{l_1,l_2,...,l_t\}$ of labels of t columns of $\mathbf{B}_{RS,255}(4,255)$. Using the t columns in $\mathbf{B}_{RS,255}(4,255)$ labeled by the numbers in Λ_t, we form a $4\times t$ RS matrix $\mathbf{B}_{RS,255,\Lambda_t}(4,t)$. This matrix satisfies the 2×2 SNS-constraint. If the labels in Λ_t satisfy all of the nine conditions given by (8), then the $4\times t$ matrix $\mathbf{B}_{RS,255,\Lambda_t}(4,t)$ satisfies the $2\times 2/3\times 3$ SM-constraint. The null space of the 255×255 CPM-dispersion of $\mathbf{B}_{RS,255,\Lambda_t}(4,t)$ gives an RS-QC-LDPC code of length $255t$ whose Tanner graph has girth at least 8.

Set $t=16$. Suppose we choose the following set of column labels of $\mathbf{B}_{RS,255}(4,255)$: $\Lambda_{16}=\{1,3,6,13,21,32,44,59,64,73,77,83,111,212,226,239\}$. The column labels in Λ_{16} satisfy the 9 conditions given by (8). Using the column labels in Λ_{16}, we form a 4×16 RS matrix $\mathbf{B}_{RS,255,\Lambda_{16}}(4,16)$ which satisfies the $2\times 2/3\times 3$ SM-constraint. The null space over GF(2) of the 255×255 CPM-dispersions of $\mathbf{B}_{RS,255,\Lambda_{16}}(4,16)$ gives a $(4,16)$-regular

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2019.2916605, IEEE Transactions on Communications

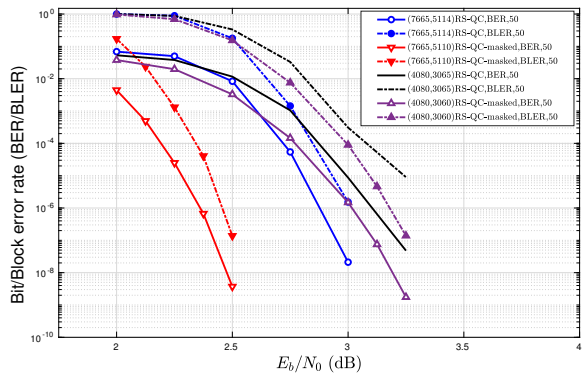

Fig. 4. The BER and BLER performances of the four RS-QC-LDPC codes given in Examples 5 and 6.

$(4080, 3065)$ RS-QC-LDPC code $\mathcal{C}_{RS,255,\Lambda_{16},ldpc}(4, 16)$ whose Tanner graph $\mathcal{G}_{RS,255,\Lambda_{16},ldpc}(4, 16)$ has girth 8. The numbers of cycles of lengths 8, 10, and 12 in $\mathcal{G}_{RS,255,\Lambda_{16},ldpc}(4, 16)$ are 688,500, 17,485,860, and 703,291,020, respectively, for a total of 721,465,380, which is a quite large number of short cycles.

In order to reduce the number of short cycles in $\mathcal{G}_{RS,255,\Lambda_{16},ldpc}(4, 16)$, we can mask the RS matrix $\mathbf{B}_{RS,255,\Lambda_{16}}(4, 16)$ with the masking matrix given by (10). Masking results in a masked RS matrix $\mathbf{B}_{RS,255,\Lambda_{16},mask}(4, 16)$ with column weights 3 and 12, respectively. The null space of the 255×255 CPM-dispersion $\mathbf{H}_{RS,255,\Lambda_{16},mask}(4, 16)$ of $\mathbf{B}_{RS,255,\Lambda_{16},mask}(4, 16)$ gives a $(3, 12)$-regular $(4080, 3060)$ masked RS-QC-LDPC code $\mathcal{C}_{RS,255,\Lambda_{16},mask,ldpc}(4, 16)$ of rate 3/4. The numbers of short cycles of lengths 8, 10, and 12 in the Tanner graph $\mathcal{G}_{RS,255,\Lambda_{16},mask,ldpc}(4, 16)$ of $\mathcal{C}_{RS,255,\Lambda_{16},mask,ldpc}(4, 16)$ are 32,640, 495,210, and 9,570,915, respectively, for a total of 10,098,765. We see that masking reduces the total number of short cycles of lengths 8, 10, and 12 in the unmasked Tanner graph by a factor of more than 71.

The BER and BLER performances of $\mathcal{C}_{RS,255,\Lambda_{16},ldpc}(4, 16)$ and $\mathcal{C}_{RS,255,\Lambda_{16},mask,ldpc}(4, 16)$, decoded with 50 iterations of the MSA are also shown in Fig. 4. The scaling factor for the unmasked and the masked codes are 0.70 and 0.75, respectively. We see that masking improves the error performance of the unmasked code. The $(3, 12)$-regular $(4080, 3060)$ masked RS-QC-LDPC code $\mathcal{C}_{RS,255,\Lambda_{16},mask,ldpc}(4, 16)$ achieves a BER of 10^{-9} without a visible error-floor and performs less than 1 dB from its threshold (2.2564 dB) and 1.62 dB from the Shannon limit (1.628 dB). The masked code performs well in both waterfall and low error rate regions. It outperforms the $(4080, 3093)$ QC-LDPC code of rate 0.758 presented in [10, Fig. 11.10, p. 505] and the $(4096, 3073)$ finite-geometry code of rate 0.750 presented in [29, Fig. 17.31, p. 910], both decoded wth SPA.

V. Conclusion

In this paper, we presented designs and constructions of QC-LDPC codes for the AWGN channel based on the conventional parity-check matrices of Reed-Solomon codes, called RS-QC-LDPC codes. Four classes of RS-QC-LDPC codes whose Tanner graphs have girth at least 6 were given. Cycle structural properties of the Tanner graphs of codes in these classes are analyzed and specific methods for constructing codes with girth at least 8 and reducing their short cycles are presented. The designed codes perform well in both waterfall and low error-rate regions. The methods presented in this paper for constructing binary RS-QC-LDPC codes can be generalized for constructing nonbinary RS-QC-LDPC codes.

References

[1] 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)*, La Jolla, CA, Feb. 2018, pp. 1 – 10.

[2] L. Lan, L. -Q. Zheng, Y. Y. Tai, L. Chen, S. Lin, and K. Abdel-Ghaffar, "Construction of quasi-cyclic LDPC codes for AWGN and binary erasure channels: A finite field approach," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2429–2458, Jul. 2007.

[3] H. Liu, Q. Huang, G. Deng, and J. Chen, "Quasi-cyclic representation and vector representation of RS-LDPC codes," *IEEE Trans. Commun.*, vol. 63, no. 4, pp.1033–1042, Apr. 2015.

[4] 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," *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 381–385.

[5] 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.

[6] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *J. Soc. Indust. Appl. Math.*, vol. 8, no. 2, pp. 300–304, Jun. 1960.

[7] R. G. Gallager, "Low density parity-check codes," *IRE Trans. Inf. Theory*, vol. IT-8, no. 1, pp. 21 – 28, Jan. 1962.

[8] M. Zhang, L. Tang, Q. Huang, and Z. Wang, "Low complexity encoding algorithm of RS-based QC-LDPC codes," in *Proc. Inf. Theory and Applic. Workshop (ITA)*, San Diego, CA, Feb. 2014, pp. 1 – 4.

[9] S. H. Kuo, Z. U Liu, and J. Yang, "On practical LDPC code construction for NAND flash applications," in *Proc. IEEE Inform. Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017, pp. 191 – 195.

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

[11] L. Lan, Y. Y. Tai, L. Chen, S. Lin, and K. Abdel-Ghaffar, "A trellis-based method for removing cycles from bipartite graphs and construction of low density parity check codes," *IEEE Commun. Lett.*, vol. 8, no. 7, pp. 443–445, Jul. 2004.

[12] F. C. M. Lau and W. M. Tam, "A fast searching method for the construction of QC-LDPC codes with large girth," *IEEE Symp. Comp. & Commun. (ISCC)*, Cappadocia, Turkey, Jul. 2012, pp. 125 – 128, Jun. 2012.

[13] A. Tasdighi, A. H. Banihashemi, and M.-R. Sadeghi, "Efficient search of girth-optimal QC-LDPC codes," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1552–1564, Apr. 2016.

[14] H. Xu, D. Feng, R. Luo, and B. Bai, "Construction of quasi-cyclic LDPC codes via masking with successive cycle elimination," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2370–2373, Dec. 2016.

[15] X. He, L. Zhou, and J. Du, "PEG-like design of binary QC-LDPC codes based on detecting and avoiding generating small cycles," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 1845-1858, May 2018.

[16] H. Zhang and J. M. F. Moura, "The design of structured regular LDPC codes with large girth," in *Proc. IEEE Global Telecom. Conf. (Globcom)*, San Francisco, CA, Dec. 2003, pp. 4022–4027.

[17] J. M. F. Moura, J. Lu, and H. Zhang, "Structured low-density parity-check codes," *IEEE Sig. Proc. Mag.*, vol. 21, no. 1, pp. 42 – 55, Jan. 2004.

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

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

[20] X. Jiang and M. H. Lee, "Large girth quasi-cyclic LDPC codes based on the Chinese remainder theorem," *IEEE Commun. Lett.* vol. 13, no. 5, pp. 342 – 344, May 2009.
[21] M. Esmaeili, M. H. Tadayon, and T. A. Gulliver, "Low-complexity girth-8 high-rate moderate length QC LDPC codes," *AEU - Int. J. Electron. & Commun.*, vol. 64, no. 4, pp. 360 – 365, Apr. 2010.
[22] J.-F. Huang, C.-M. Huang, and C.-C. Yang, "Construction of one-coincidence sequence quasi-cyclic LDPC Codes of large girth," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1825–1836, Mar. 2012.
[23] 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 Inform. (ISTC)*, Gothenburg, Sweden, Aug. 2012, pp. 31 – 35.
[24] A. Tasdighi, A. H. Banihashemi, and M.-R. Sadeghi, "Symmetrical constructions for regular girth-8 QC-LDPC codes," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 14–22, Jan. 2017.
[25] R. Sun, Y. Tian, and J. Liu, "Construction of QC-LDPC codes based on generalized RS codes with girth larger than 6," in *Proc. IEEE Int. Conf. Commun. Syst. (ICCS)*, Shenzhen, China, Dec. 2016. pp. 1–6.
[26] J. Xu, L. Chen, I. Djurdjevic, S. Lin, and K. Abdel-Ghaffar, "Construction of regular and irregular LDPC codes: Geometry decomposition and masking," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 121–134, Jan. 2007.
[27] J. Chen and M. P. C. Fossorier, "Near optimum universal belief propagation based decoding of low-density parity check codes," *IEEE Trans. Commun.*, vol. 50, no.3, p. 406 – 414, Mar. 2002.
[28] J. Li, S. Lin, and K. Abdel-Ghaffar, "Improved message-passing algorithm for counting short cycles in bipartite graphs," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Hong Kong, China, Jun. 2015, pp. 416–420.
[29] S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, 2nd Ed. Upper Saddle River, NJ: Prentice Hall, 2004.
[30] Q. Diao, Q. Huang, S. Lin, and K. Abdel-Ghaffar, "A transform approach for analyzing and constructing quasi-cyclic low-density parity-check codes," in *Proc. Inform. Theory Applic. Workshop (ITA)*, La Jolla, CA, Feb. 2011, pp. 1 – 8.
[31] D. Declercq and M. Fossorier, "Improved impulse method to evaluate the low weight profile of sparse binary linear codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Toronto, Canada, Jul. 2008, pp. 1963–1967.
[32] J. Li, K. Liu, S. Lin, and K. Abdel-Ghaffar, "Algebraic LDPC codes: Construction, low error-floor, large girth and a reduced-complexity decoding scheme," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2626–2637, Aug. 2015.
[33] Z. Li and B. V. K.V. Kumar, A class of good quasi-cyclic low-density parity check codes based on progressive edge growth graph," in *Proc. 38th Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, Nov. 2004, pp. 1990–1994.
[34] D. Divsalar, S. Dolinar, C. R. Jones, and K. Andrews, "Capacity-approaching protograph codes," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 6, pp. 876–888, Aug. 2009.

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, Japan, in 2013, and the M. S. degree in Electrical Engineering from Shanghai Jiao Tong University, in 2015. She is currently pursuing the Ph.D. degree with the University of Arizona, Tucson, United States.

Her current research interests are in the general area of error correction coding for communication and storage systems, and applications of neural networks in channel coding.

PLACE PHOTO HERE

Bane Vasić is a Professor of Electrical and Computer Engineering and Mathematics at the University of Arizona. He is affiliated with BIO5, the Institute for Collaborative Beoresearch, and is a Director of the Error Correction Laboratory.

Dr. Vasić is an inventor of the soft error-event decoding algorithm, and the key architect of a detector/decoder for Bell Labs data storage chips which were regarded as the best in industry.

Dr. Vasić is a Chair of the IEEE Data Storage Technical Committee, Editor of the IEEE JSAC Special Issues on Data Storage Channels and Member of the Editorial Board of the IEEE Transactions on Magnetics. He is da Vinci Circle and Fulbright Scholar and founder and Chief Scientific Officer of Codelucida.

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 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. He has published numerous technical papers in IEEE Transactions and other refereed journals. 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), (with T. Kasami, T. Fujiwara, and M. Fossorier) the book, Trellises and Trellis-Based Decoding Algorithms, (Boston, MA: Kluwer Academic, 1998), (with W. E. Ryan) the book, Channel Codes: Classical and Modern (Cambridge University Press, 2009), and (with J. Li, K. Abdel-Ghaffar, W. E. Ryan, and D. J. Costello, Jr.) the book, LDPC Code Designs, Constructions, and Unification (Cambridge University Press, 2017). His current research areas include algebraic coding theory, coded modulation, error control systems, and satellite communications. He has served as the Principle Investigator on 32 research grants.

Dr. Lin is a Member of the IEEE Information Theory Society and the Communication Society. He served as the Associate Editor for Algebraic Coding Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY from 1976 to 1978, and as the Program Co-Chairman of the IEEE International Symposium of Information Theory held in Kobe, Japan, in June 1988. He was the President of the IEEE Information Theory Society in 1991. 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.

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, in 1983 and 1986, respectively, all in electrical engineering.

Currently, he is a Professor of Electrical and Computer Engineering at the University of California, Davis. His main interest is coding theory.

Dr. Abdel-Ghaffar 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. He is a co-recipient of the IEEE Communications Society 2007 Stephen O. Rice Prize paper award.

PLACE PHOTO HERE

William E. Ryan (S'83-M'84-SM'01-F'11) received the Ph.D. degree in electrical engineering from the University of Virginia in 1988 after receiving the B.S. and M.S. degrees from Case Western Reserve University and the University of Virginia, respectively, in 1981 and 1984.

Prior to his academic positions, Dr. Ryan held positions in industry for five years, first at The Analytic Sciences Corporation, then at Ampex Corporation, and finally at Applied Signal Technology. From 1993 to 1998, he was on the faculty in the Department of Electrical and Computer Engineering at New Mexico State University. From 1998 to 2011, he was a faculty member in the Department of Electrical and Computer Engineering at the University of Arizona, first as an associate professor and then as full professor. He is now a Senior Associate with Zeta Associates, headquartered in Fairfax, VA.

Dr. Ryan has published approximately 150 publications in the leading conferences and journals in the area of communication theory and channel coding and he is co-author (with Shu Lin) of Channel Codes: Classical and Modern (Cambridge University Press, 2009). He is also a contributor to several other books. His research interests are in coding and signal processing with applications to data storage and wireless data communications. He was an associate editor for the IEEE Transactions on Communications from 1998 through 2005.