Controlling LDPC Absorbing Sets via the Null Space of the Cycle Consistency Matrix

Jiadong Wang, Lara Dolecek and Richard Wesel
wjd@ee.ucla.edu, dolecek@ee.ucla.edu, wesel@ee.ucla.edu

***Abstract*— This paper focuses on controlling absorbing sets for a class of regular LDPC codes, known as separable, circulant-based (SCB) codes. For a specified circulant matrix, SCB codes all share a common mother matrix and include array-based LDPC codes and many common quasi-cyclic codes. SCB codes retain standard properties of quasi-cyclic LDPC codes such as girth, code structure, and compatibility with existing high-throughput hardware implementations. This paper uses a cycle consistency matrix (CCM) for each absorbing set of interest in an SCB LDPC code. For an absorbing set to be present in an SCB LDPC code, the associated CCM must not be full column-rank. Our approach selects rows and columns from the SCB mother matrix to *systematically* eliminate dominant absorbing sets by forcing the associated CCMs to be full column-rank. Simulation results demonstrate that the new codes have steeper error-floor slopes and provide at least one order of magnitude of improvement in the low FER region. Identifying absorbing-set-spectrum equivalence classes within the family of SCB codes with a specified circulant matrix significantly reduces the search space of possible code matrices.**

I. Introduction

High-speed systems such as data storage devices demand high-rate LDPC codes which have low FER and are compatible with high-throughput hardware architectures. Certain graphical structures can prevent good low FER performance by introducing a high error floor.

Prior work indicates that certain sub-graphs called trapping sets [1], and, in particular, a subset of trapping sets called absorbing sets [2] are a primary cause of the error floor. Absorbing sets are trapping sets that are stable under bit-flipping decoding. This paper focuses on controlling absorbing sets for a class of regular LDPC codes: separable, circulant-based (SCB) codes.

Recent papers have proposed methods to improve the absorbing set spectrum. Introducing additional check nodes [3] or increasing the girth [4] eliminates small trapping sets for some codes. The approach in [5] avoids dominant absorbing sets without compromising code properties by carefully selecting the rows of the (SCB) mother matrix. The algorithm in [6] constructs quasi-cyclic codes from Latin squares so that the Tanner graph does not contain certain trapping sets.

As in [7], this paper defines a cycle consistency matrix (CCM) for each possible absorbing set in an SCB LDPC code. For an absorbing set to be present in an SCB LDPC code, the associated CCM must not be full column-rank. Using this novel observation, a new code construction approach selects rows and columns from the SCB mother matrix to systematically eliminate the dominant absorbing sets by forcing the associated CCMs to be full column-rank. Simulation results demonstrate that the new codes have steeper error-floor slopes and provide at least one order of magnitude of improvement in the low-FER region.

Identifying absorbing-set-spectrum equivalence classes within the family of SCB codes significantly reduces the search space of possible code matrices. The CCM-based analysis can be extended to other families of quasi-cyclic codes in [8] and [9] to show that these codes can have good absorbing set spectra with the proper choice of parameters.

Section II introduces separable circulant-based (SCB) codes and the cycle consistency matrix (CCM). Section III identifies the CCMs for the dominant absorbing sets of an example family of SCB codes. Section III then selects specific rows from the SCB mother matrix to eliminate certain dominant absorbing sets by forcing the associated CCMs to be full column-rank. Section III then eliminates the remaining dominant absorbing sets by selecting specific columns from the SCB mother matrix, again forcing the associated CCMs to be full column-rank. Section IV provides simulation results demonstrating the performance improvement obtained by the new codes. Section V delivers the conclusions.

II. Definition and Preliminaries

This section introduces separable, circulant-based (SCB) codes and the cycle consistency matrix (CCM) associated with absorbing sets in SCB codes.

A. Circulant-based LDPC codes

Circulant-based LDPC codes are a family of structured regular (r, c) codes where r is the variable node degree and c is the check node degree. They are constructed as r rows and c columns of circulant matrices. They are particularly compatible with high-throughput hardware implementations [10].

The parity-check matrix of circulant-based LDPC codes has the following general structure:

$$H_{p,f}^{r,c} = \begin{bmatrix} \sigma^{f(0,0)} & \sigma^{f(0,1)} & \sigma^{f(0,2)} & \dots & \sigma^{f(0,c-1)} \\ \sigma^{f(1,0)} & \sigma^{f(1,1)} & \sigma^{f(1,2)} & \dots & \sigma^{f(1,c-1)} \\ \sigma^{f(2,0)} & \sigma^{f(2,1)} & \sigma^{f(2,2)} & \dots & \sigma^{f(2,c-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \sigma^{f(r-1,0)} & \sigma^{f(r-1,1)} & \sigma^{f(r-1,2)} & \dots & \sigma^{f(r-1,c-1)} \end{bmatrix},$$

where σ is a $p \times p$ circulant matrix.

A column (row) group is a column (row) of circulant matrices. Each variable node has a label (j, k) with $j \in \{0, ..., c-1\}$

This research was supported by a gift from Inphi Corp. and in part by grant CCF-1029030 from NSF.

that is the index of the corresponding column group and $k \in \{0, ..., p-1\}$ identifies the specific column within the group. Similarly, each check node has a label (i, l) where $i \in \{0, ..., r-1\}$ identifies the row group and $l \in \{0, ..., p-1\}$.

Circulant-based LDPC codes include, for example, the constructions in [8], [9] and [11]. The girth can be guaranteed to be at least 6 by placing a constraint on the values of the submatrix exponent $f(i,j)$ [2].

This paper focuses on separable, circulant-based (SCB) codes defined as follows:

Definition 1 (Separable, Circulant-Based (SCB) Code): An SCB code is a circulant-based LDPC code with a parity-check matrix $H^{r,c}_{p,f}$ in which $f(i,j)$ is separable, i.e., $f(i,j) = a(i) \cdot b(j)$. ■

Parity check matrices of SCB codes with the specified circulant matrix can be viewed as originating from a common SCB mother matrix $H^{p,p}_{p,f_m}$ with $f_m(i,j) = i \cdot j$. The functions $a(i)$ and $b(j)$ effectively specify which rows and columns of the mother matrix are selected for the resultant SCB matrix. The ranges of $a(i)$ and $b(j)$ are $\{0, \ldots, p-1\}$.

The SCB structure imposes certain conditions [2] on the variable and check nodes:

Bit Consistency: The neighboring check nodes of a variable node must have distinct row-group (i) labels.

Check Consistency: The neighboring variable nodes of a check node must have distinct column-group (j) labels.

Cycle consistency: As shown in [2], any length-$2t$ cycle in an SCB mother matrix, which involves t variable nodes with column-group labels j_1 through j_t and t check nodes with row-group labels i_1 through i_t, must satisfy:

$$\sum_{m=1}^{t} i_m(j_{(m+1) \mod t} - j_m) = 0 \bmod p. \quad (1)$$

After reviewing absorbing sets [2] below, (1) is used to construct a necessary matrix equation for an absorbing set to exist based on the cycles contained in that absorbing set.

B. Absorbing sets and the Cycle Consistency Matrix

An LDPC code with parity-check matrix H is often viewed as a bipartite (Tanner) graph $G_H = (V, F, E)$, where the set V represents the variable nodes, the set F represents the check nodes, and E corresponds to the edges between variable and check nodes.

For a variable node subset $V_{\text{as}} \subset V$, analogous to G_H, let $G_{\text{as}} = (V_{\text{as}}, F_{\text{as}}, E_{\text{as}})$ be the bipartite graph of the edges E_{as} between the variable nodes V_{as} and their neighboring check nodes F_{as}. Let $o(V_{\text{as}}) \subset F_{\text{as}}$ be the neighbors of V_{as} with odd degree (unsatisfied check nodes) in G_{as} and $e(V_{\text{as}}) \subset F_{\text{as}}$ be the neighbors of V_{as} with even degree in G_{as} (satisfied check nodes).

Definition 2 (Absorbing Set): An (a,b) absorbing set is a set $V_{\text{as}} \subset V$ with $|V_{\text{as}}| = a$ and $|o(V_{\text{as}})| = b$, where each node in V_{as} has strictly fewer neighbors in $o(V_{\text{as}})$ than $e(V_{\text{as}})$. ■

Suppose there are n variable nodes in the absorbing set. Let $j_1, \ldots, j_n$ be the column-group labels of these n nodes in the SCB mother matrix. Define $u_m = j_m - j_1, m = 2, ..., n$ and $\mathbf{u} = [u_2, ..., u_n]$. For each cycle in the absorbing set, by replacing the difference of j's with the difference of u's and manipulating the expression, (1) may be written as

$$\sum_{m=2}^{t} (i_{m-1} - i_m) u_m = 0 \mod p, \quad (2)$$

where $2t$ is the length of that cycle. Note that i_m will be different for different cycles reflecting the particular cycle trajectories.

Every cycle in the absorbing set satisfies an equation of the form (2). Taken together, these equations produce a matrix equation: $\mathbf{M}\mathbf{u} = 0 \mod p$, where $\mathbf{M}_{ym}$ is the coefficient of u_m in (2) for the yth cycle.

A key property of $\mathbf{M}$ is that $\mathbf{M}\mathbf{u} = 0 \mod p$ completely characterizes the requirement that every cycle in G_{as} satisfies (2). Even so, it is not necessary for $\mathbf{M}$ to include a row for every cycle in the absorbing set.

A cycle need not be included in $\mathbf{M}$ if it is a linear combination of cycles already included in $\mathbf{M}$. Thus the number of rows needed in $\mathbf{M}$ is the number of linearly independent cycles in G_{as}. Some definitions [12] from graph theory are necessary to establish the number of linearly independent cycles in G_{as} and hence how many rows are needed for $\mathbf{M}$.

Definition 3 (Incidence Matrix): For a graph with n vertices and q edges, the (unoriented) incidence matrix is an $n \times q$ matrix B with $B_{ij} = 1$ if vertex v_i and edge x_j are incident and 0 otherwise. ■

Note that since each edge is incident to exactly two vertices, each column of B has exactly two ones.

The incidence matrix of a graph is useful for identifying the cycles in the graph because every cycle has the property that the indicator vector $\mathbf{x}_c$ of the edges in the cycle satisfies $B\mathbf{x}_c = 0 \mod 2$. This is formalized in the definition below.

Definition 4 (Binary Cycle Space): The binary cycle space of a graph is the null space of its incidence matrix over $GF(2)$. ■

Any absorbing-set bipartite graph G_{as} can be transformed into a graph whose only vertices are V_{as} and where two vertices are connected iff there is a check node that connects them. We call this graph the variable-node graph of the absorbing set. If each satisfied check node in G_{as} has degree 2, then the incidence matrix B is the transpose of the submatrix of $H^{r,c}_{p,f}$ whose rows and columns correspond to F_{as} and V_{as} respectively.

The incidence matrix provides a characterization of all the cycles in an absorbing set. The number of linearly independent cycles in an absorbing set, which is the dimension of its binary cycle space, is the size of the null space of the incidence matrix B_{as}: $D_{\text{bcs}} = |E_{\text{as}}| - \text{rank}(B_{\text{as}})$.

Having established the number of rows in $\mathbf{M}$, it can be formally defined as the Cycle Consistency Matrix:

Definition 5 (Cycle Consistency Matrix): The cycle consistency matrix $\mathbf{M}$ of an absorbing-set graph G_{as} has $|V_{\text{as}}| - 1$ columns and D_{bcs} rows. The rows of $\mathbf{M}$ correspond to D_{bcs} linearly independent cycles in G_{as}. Each row has the coefficients of $\mathbf{u}$ in (2) for each one of these linearly independent cycles. ■

Note that $\mathbf{M} \cdot \mathbf{u} = 0 \mod p$ completely characterizes the requirement that every cycle in G_{as} satisfies (2).

The vector $\mathbf{u}$ cannot be an all-zero vector because an all-zero $\mathbf{u}$ indicates that all variable nodes have the same column group. This violates the Check Consistency condition, which requires that variable nodes sharing a check node have distinct column groups. Thus $\mathbf{u} \neq 0$, and a necessary condition for the existence of a given absorbing set is that its $\mathbf{M}$ does not have full column-rank in $GF(p)$.

If the variable-node graph of the absorbing set A is a subgraph of the variable-node graph of another absorbing set B with the same number of variable nodes, then we say the variable-node graph of the absorbing set A is *extensible.*

Theorem 1: Given a proposed absorbing set graph $G_{\text{as}} = (V_{\text{as}}, F_{\text{as}}, E_{\text{as}})$, where every variable node is involved in at least one cycle[1], the column group labels of the variable nodes in V_{as} in the SCB mother matrix, and the row-group labels of the check nodes in F_{as} in the SCB mother matrix, the following are necessary conditions for the proposed absorbing set to exist in each daughter SCB LDPC code (with a parity check matrix H that includes the specified row and column groups of that SCB mother matrix): (1) The CCM for G_{as} does not have full column-rank; (2) Variable nodes in V_{as} satisfy the Bit Consistency condition and can form a difference vector $\mathbf{u}$ in the null space of the CCM; and (3) Each check node in F_{as} satisfies the Check Consistency condition. Taken together, these conditions are also sufficient if the variable-node graph of this absorbing set is not extensible.

Proof: Each of the three conditions has already been shown to be a necessary condition for the existence of G_{as} in an SCB. If all of these three conditions are satisfied, all the cycles presented in the CCM exist in G_H and any linear combination of these cycles exists in G_H as well. The only issue is whether the existing graphical structures have *additional* linearly independent cycles not required by the CCM. There are only three ways for this to happen: (1) a variable node's unsatisfied check node is the same as another variable node's unsatisfied check node, or (2) a variable node's unsatisfied check node is the same as one satisfied check node in the graph, or (3) two of the satisfied check nodes are the same. In each of these cases, additional edges extend the variable-node graph. For this to be possible, the original variable-node graph must be extensible as defined above. Thus if the variable-node graph is not extensible, the above constructed solution fully describes the existence of the proposed absorbing set. This concludes that G_{as} is present in G_H. ■

III. Illustrative Case Study with $r = 4$

This section provides an example with r=4 (four row groups) that shows how to design an SCB code with a specified circulant matrix that eliminates the dominant absorbing sets by selecting rows and columns from the SCB mother matrix to force the CCMs associated with the dominant absorbing sets to be full column-rank or to preclude $\mathbf{u}$ from the null space of $\mathbf{M}$. Our example of SCB code design involves three classes of SCB codes:

Array-based codes [13] are the most elementary SCB codes in which the first r rows of the SCB mother matrix $H^{p,p}_{p,f}, f(i,j) = i \cdot j$ comprise the parity-check matrix. We will refer to this class as the elementary array-based (EAB) codes.

As shown in [5], a careful selection of the r row-groups from the overall SCB mother matrix can improve performance over the EAB codes. Thus, selected-row (SR) SCB codes are our second class of SCB codes. The parity-check matrix for these codes is $H^{r,p}_{p,f}, f(i,j) = a(i) \cdot j$ where $a(i)$ is called the row-selection function (RSF).

Removing a few column groups from an SR-SCB code provides further improvement. Hence, shortened SR (SSR) SCB codes form our third class of SCB codes. The parity-check matrix for these codes is $H^{r,c}_{p,f}, f(i,j) = a(i) \cdot b(j)$ where $b(j)$ is called the column-selection function (CSF). Note that for a $p \times p$ circulant matrix, EAB SCB and SR SCB codes have p column groups (p^2 binary columns), but SSR SCB codes have fewer column groups since $b(j)$ selects a subset of the possible column groups.

Section III-A identifies the $(6,4)$ absorbing sets as dominant for EAB SCB codes with $r = 4$. Section III-B analyzes the three possible $(6,4)$ absorbing set configurations and shows how carefully selecting four row groups from the SCB mother matrix can eliminate two of the three possible configurations. Section III-C provides an efficient provable algorithm to eliminate all $(6,4)$ absorbing sets by combining the row selection of Section III-B with column selection in which some column groups of the SCM mother matrix are removed. Section III-D identifies equivalence classes among SCB codes.

A. Identifying the dominant absorbing sets

From the previous results in [2], $(6,4)$ absorbing sets are the smallest possible structure for EAB SCB codes with r=4 for $p > 19$. Hardware simulations [10] also demonstrate that $(6,4)$ absorbing sets are the dominant cause of the error floor for example r=4 EAB codes.

Based on these results, a key goal of this example is to design an SR or SSR r=4 SCB code that avoids all $(6,4)$ absorbing sets. As a first step, the lemma and corollary below establish that SR and SSR code design approaches do not introduce absorbing sets smaller than $(6,4)$.

Lemma 1: For p large enough, careful selection of the row-selection function $a(i)$ avoids all absorbing sets smaller than $(6,4)$ in the Tanner graph corresponding to $H^{4,p}_{p,a(i)\cdot j}$.

Proof: Results in [2] imply that the smallest possible absorbing sets for an $H^{4,p}_{p,a(i)\cdot j}$ SCB code are $(4,4)$, $(5,2)$, $(5,4)$ and $(6,2)$ and that any RSF that avoids the $(4,4)$ absorbing set also avoids the $(5,2)$ absorbing set absorbing sets. The analysis in [2] proves that under the mapping $[a(0), a(1), a(2), a(3)] = [0,1,2,3]$ the Tanner graph does not contain absorbing sets smaller than $(6,4)$ for p large enough. Using a similar technique, it can be likewise shown that under

[1]It is easy to show that if the variable node degree is at least 2, then each variable node in a given absorbing set must be a part of at least onc cycle.

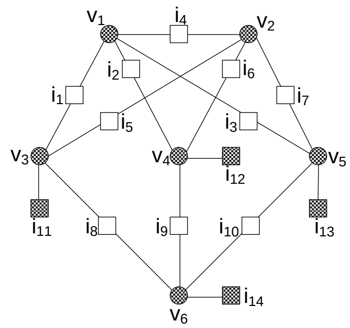

Fig. 1. Depiction of the first (6, 4) absorbing set configuration.

the mapping $[a(0), a(1), a(2), a(3)] = [0, 1, 3, 4]$, no such sets exist for p large enough. ∎

Shortening an LDPC code simply locks certain variable nodes to their correct values. Hence shortening can never introduce new absorbing sets, implying the following corollary:

Corollary 1: An SSR code contains no smaller absorbing sets than the SR code from which it is created.

Section III-B shows that SR codes always have $(6, 4)$ absorbing sets, regardless the RSF. Avoidance of all such configurations using shortening is the subject of Section III-C.

B. $(6, 4)$ absorbing sets in SR SCB codes

Fig.s 1-3 show the three distinct configurations of $(6, 4)$ absorbing sets possible for r=4 SCB codes. This section determines which configurations are possible in EAB and SR SCB codes. The first configuration exists in the EAB SCB code and in every possible SR SCB code. The second configuration exists in the EAB code but can be avoided by a proper RSF for the SR SCB code. The third configuration does not exist in either the EAB code or the SR codes.

The following lemma establishes that the EAB code and all SR codes have the first configuration (Fig. 1).

Lemma 2: In the Tanner graph corresponding to the EAB code and all SR codes with $H^{4,p}_{p,f(i,j)}$ there are $(6, 4)$ absorbing sets for any p with the configuration shown in Fig. 1.

Proof: The binary cycle space of Fig. 1 has dimension 5. Using the technique of Section II-B and the following five linearly independent cycles: $v_1{-}v_2{-}v_3$, $v_1{-}v_2{-}v_4$, $v_1{-}v_2{-}v_5$, $v_1{-}v_3{-}v_6{-}v_4$, $v_1{-}v_4{-}v_6{-}v_5$, we construct the CCM in (3):

$$\mathbf{M} = \begin{bmatrix} i_4 - i_5 & i_5 - i_1 & 0 & 0 & 0 \\ i_4 - i_6 & 0 & i_6 - i_2 & 0 & 0 \\ i_4 - i_7 & 0 & 0 & i_7 - i_3 & 0 \\ 0 & i_1 - i_8 & i_9 - i_2 & 0 & i_8 - i_9 \\ 0 & 0 & i_2 - i_9 & i_{10} - i_3 & i_9 - i_{10} \end{bmatrix}. \quad (3)$$

Note that $\det(\mathbf{M})$ is computed as $\mathbf{M}_{11}\mathbf{M}_{23}\mathbf{M}_{34}\mathbf{M}_{42}\mathbf{M}_{55} - \mathbf{M}_{12}(\mathbf{M}_{23}\mathbf{M}_{31}\mathbf{M}_{45}\mathbf{M}_{54} - \mathbf{M}_{21}\mathbf{M}_{34}(\mathbf{M}_{43}\mathbf{M}_{55} - \mathbf{M}_{45}\mathbf{M}_{53}))$, where $\mathbf{M}_{ij}$ denotes the (i, j) entry in $\mathbf{M}$.

From the proof of Lemma 8 in [2], there are only two non-isomorphic row-group labelings for the check nodes of Fig. 1: $(i_1, i_2, i_3, i_4, i_5, i_6, i_7, i_8, i_9, i_{10}) = (x, y, z, w, y, z, x, z, x, y)$ or $(x, y, z, w, y, z, x, z, w, y)$. The first labeling yields

$$\mathbf{M} = \begin{bmatrix} w - y & y - x & 0 & 0 & 0 \\ w - z & 0 & z - y & 0 & 0 \\ w - x & 0 & 0 & x - z & 0 \\ 0 & x - z & x - y & 0 & z - x \\ 0 & 0 & y - x & y - z & x - y \end{bmatrix}, \quad (4)$$

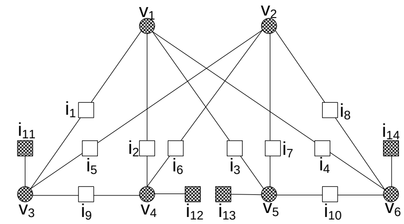

Fig. 2. Depiction of the second candidate (6, 4) absorbing set.

which has $\det(\mathbf{M}) = 0$ for any choice of w, x, y, z. Thus the first labeling always has a non-zero solution to $\mathbf{M} \cdot \mathbf{u} = 0 \mod p$. One such solution to this equation is

$$\begin{bmatrix} u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} = \begin{bmatrix} (x{-}y)(z{-}y)(x{-}z) \\ (w{-}y)(z{-}y)(x{-}z) \\ (w{-}z)(y{-}x)(x{-}z) \\ (w{-}x)(y{-}x)(z{-}y) \\ (w{-}y)(x{-}z)(z{-}y){+}(y{-}x)(w{-}z)(x{-}y) \end{bmatrix}. \quad (5)$$

For this absorbing set, the Check Consistency condition requires $u_2{\neq}0$, $u_3{\neq}0$, $u_4{\neq}0$, $u_5{\neq}0$, $u_2{\neq}u_3$, $u_2{\neq}u_4$, $u_2{\neq}u_5$, $u_3{\neq}u_6$, $u_4{\neq}u_6$, and $u_5{\neq}u_6$. These requirements as well as the Bit Consistency inequalities are met since x, y, z, w are mutually distinct. The solution in (5) satisfies the bit, check and cycle consistency constraints.

From [2], no check nodes in $(6, 4)$ absorbing sets have degree > 2 relative to the variable nodes in the absorbing set. Since $(6, 4)$ absorbing sets are thus not extensible, the existence of $(6, 4)$ absorbing sets with the configuration of Fig. 1 follows from Theorem 1 and the observation that the CCM above always has zero determinant. Any four distinct values between zero and $p - 1$ for $\{x, y, z, w\}$ identify a labeling of the first type above that induces absorbing sets with the configuration of Fig. 1 in the EAB code and every SR SCB code. ∎

Remark 1: For the second labeling above, $\det(\mathbf{M}) \neq 0$ for $\{x, y, z, w\} = \{0, 1, 2, 3\}$ and other careful choices such as $\{0, 1, 3, 4\}$. Thus $\det(\mathbf{M}) \neq 0 \mod p$ for p large enough thus precluding $(6, 4)$ configurations with the second labeling.

Fig. 2 shows the second possible configuration of a $(6, 4)$ absorbing set in an r=4 SCB code. The following lemma establishes that the EAB code has this configuration but well-designed SR codes avoid it.

Lemma 3: In the Tanner graph corresponding to $H^{4,p}_{p,f(i,j)}$, $(6, 4)$ absorbing sets of the type shown in Fig. 2 exist in the EAB code, but do not exist in certain SR codes for p large enough.

Proof: Again using the technique of Section II-B we construct the CCM for this configuration. The binary cycle space for Fig. 2 has dimension 5. We construct the following CCM by selecting the following linearly independent cycles: $v_1 - v_3 - v_4, v_1 - v_5 - v_6, v_2 - v_3 - v_4, v_2 - v_5 - v_6, v_1 - v_4 - v_2 - v_5$,

$$\mathbf{M} = \begin{bmatrix} 0 & i_1 - i_9 & i_9 - i_2 & 0 & 0 \\ 0 & 0 & 0 & i_3 - i_{10} & i_{10} - i_4 \\ i_5 - i_6 & i_9 - i_5 & i_6 - i_9 & 0 & 0 \\ i_7 - i_8 & 0 & 0 & i_{10} - i_7 & i_8 - i_{10} \\ i_6 - i_7 & 0 & i_2 - i_6 & i_7 - i_3 & 0 \end{bmatrix}. \quad (6)$$

As with the previous lemma, the proof of Lemma 8 in [2] identifies exactly two non-isomorphic labelings

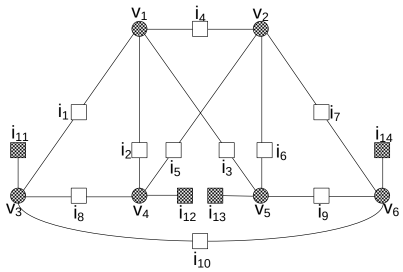

Fig. 3. Depiction of the third candidate (6, 4) absorbing set.

for $(i_1, i_2, i_3, i_4, i_5, i_6, i_7, i_8, i_9, i_{10})$ in Fig. 2. These are $(x, y, z, w, y, x, w, z, z, y)$ and $(x, y, z, w, y, z, w, x, w, y)$.

Since $\det(\mathbf{M})$=0 for $\{x, y, z, w\} = \{0, 1, 2, 3\}$ for the first labeling, the configuration in Fig. 2 exists in the EAB code.

For the second labeling, $\det(\mathbf{M}) \neq 0$ for distinct x, y, z, and w. Therefore the configuration with this labeling does not exist in EAB or SR codes for p large enough.

For both labelings, SR codes can avoid the configuration in Fig. 2 with a careful choice of the row mapping. One such choice is $\{x, y, z, w\} = \{0, 1, 3, 4\}$. Since the largest prime factor of $\det(\mathbf{M})$ is 31, this choice avoids the configuration in Fig. 2 for $p > 31$. ■

Fig. 3 shows the third configuration of a $(6, 4)$ absorbing set in an r=4 SCB code. Lemma 4 establishes that neither the EAB nor well-designed SR codes have this configuration.

Lemma 4: In the Tanner graph corresponding to $H^{4,p}_{p,f(i,j)}$ there are no $(6, 4)$ absorbing sets for p large enough with the configuration shown in Fig. 3 in either the EAB or in the SR code.

Proof: The proof technique is as in Lemmas 2 and 3, and is omitted for the lack of space. ■

C. Eliminating $(6, 4)$ absorbing sets with shortening

For a sufficiently large p, well-designed SR codes avoid the $(6, 4)$ absorbing set configurations in Fig. 2 and 3. However, as shown in Lemma 2, SR codes cannot eliminate the $(6, 4)$ absorbing set configuration in Fig. 1. We now consider shortened SR (SSR) codes that retain only certain column groups from the SCB mother matrix (thereby reducing the rate). A well-chosen column selection $b(j)$ allows the Tanner graph corresponding to $H^{4,c}_{p,a(i)\cdot b(j)}$ to avoid all $(6, 4)$ absorbing sets.

We begin with an SR code using a proper $a(i)$, for instance $[0, 1, 3, 4]$, that already avoids the $(6, 4)$ absorbing set configurations in Fig.s 2 and 3 for p large enough. We then choose a column selection $b(j)$ to also avoid the $(6, 4)$ absorbing set configurations in Fig. 1. Choosing a column selection $b(j)$ reduces to choosing a submatrix of $H^{4,p}_{p,a(i)\cdot j}$ by eliminating certain variable nodes. This operation cannot introduce smaller absorbing sets.

One solution to $\mathbf{M} \cdot \mathbf{u} = 0 \mod p$ is equation (5). The dimension of the binary cycle space is 5 and the rank of $\mathbf{M}$ in (4) is 4. Therefore this solution is a basis of the null space. Multiplying u by a constant c, for $1 \leq c \leq p-1$, also results in a solution. These $p-1$ solutions are all of the feasible solutions, i.e. the entire null space of $\mathbf{M}$.

For each of the $p-1$ choices of $u_2, ..., u_6$ in the null space, we can choose j_1 from $0, 1, ..., p-1$ and obtain $j_2, ..., j_6$. Thus there are $p(p-1)$ ways to find j_1 to j_6 for a fixed $\{x, y, z, w\}$. Since there are $4!$ ways to assign check node labels based on the set $\{x, y, z, w\}$ for a fixed row mapping, there are at most $24p(p-1)$ possible vectors $[j_1, j_2 ..., j_6]$ that can form the configuration in Fig. 1. These vectors form the set $\tilde{V}$ of vectors that completely characterizes this absorbing set configuration.

For an SSR code, each variable node label j is in a set J where $J \subset \{0, 1, ..., p-1\}$ and we can only choose $[j_1, j_2 ..., j_6]$ that $j_m \in J, m = 1, ..., 6$. There are $\binom{|J|}{6}$ possible $[j_1, j_2, ..., j_6]$ vectors and they form a set of vectors V. If $V \cap \tilde{V} = \emptyset$, the new code does not have the $(6, 4)$ configuration in Fig. 1.

We can find J so that $V \cap \tilde{V} = \emptyset$ with the greedy column-cutting procedure described in Algorithm 1.[2]

Algorithm 1 Greedy column-cutting algorithm.

$J = \{0, 1, ..., p-1\}$.
Construct the set W of all the vectors $[j_1, j_2 ..., j_6]$ that form the configuration in Fig. 1 with $j_n \in J$.
while $|W| > 0$ **do**
  Find the most frequent j in W, say j_m.
  Replace J by $J \setminus j_m$.
  Remove every $[j_1, j_2 ..., j_6]$ that involves j_m from W.
end while

Remark 2: A similar technique increases the girth [4] instead of eliminating small absorbing sets. However, increasing girth does not guarantee better performance, see [6]. ■

D. Equivalence Classes for SR codes

From the above analysis, we derive certain code equivalence conditions. Since the order of the elements in the row-mapping vector only permutes the matrix $H^{r,c}_{p,a(i)\cdot j}$ and does not change the code, we can assume that the elements of $[a_1, a_2, ..., a_r]$ monotonically increase. Consider a *difference matrix* D of the mapping vector, where $D_{ij} = a_j - a_i \mod p, 1 \leq i, j \leq r$. If $\tilde{D} = D$ or $\tilde{D}$ is D reflected on its antidiagonal, we say that $\tilde{D}$ and D are equivalent. As we now show, certain codes belong to the same absorbing-set equivalence classes.

Lemma 5: For the row mapping $(i, a(i))$, there are at least three equivalence conditions:

1) $[a_1, a_2, ..., a_r] = [\tilde{a}_1, \tilde{a}_2, ..., \tilde{a}_r] +$ constant $\mod p$
2) $[a_1, a_2, ..., a_r] = [\tilde{a}_1, \tilde{a}_2, ..., \tilde{a}_r] \times$ constant $\mod p$
3) $[a_1, a_2, ..., a_r]$ and $[\tilde{a}_1, \tilde{a}_2, ..., \tilde{a}_r]$ have equivalent difference matrices.

Proof: With the results in Section III-B and III-C, the necessary condition for the existence of certain absorbing sets depends on whether the determinant of CCM is zero. Since the determinant is only a function of the differences between the elements of the mapping vectors, if two mapping vectors $[a_1, a_2, ..., a_r]$ and $[\tilde{a}_1, \tilde{a}_2, ..., \tilde{a}_r]$ share any of the three

[2]While we omit the details, a proper column selection can also be found by progressively adding columns to an initially empty matrix.

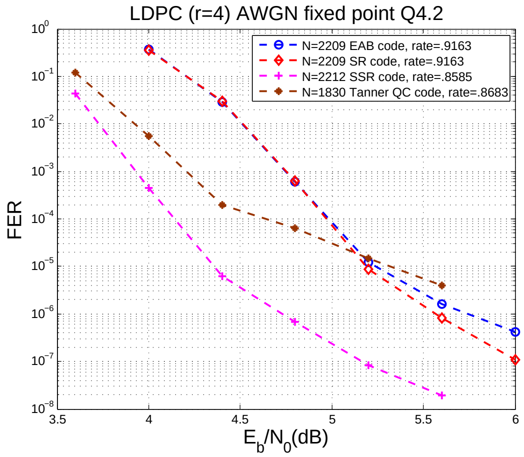

Fig. 4. Performance comparison of EAB, SR, SSR, and a code from [8].

SNR	n.r.	n.e.	(6,4)	(6,6)	(7,4)	(8,2)	(8,4)	(9,4)	(10,4)	(12,4)
5.6dB	2.0E8	322	236	2	2	27	3	1	37	1
6.0dB	8.0E8	329	329	0	0	0	0	0	0	0
5.6dB	2.0E8	167	38	3	0	40	45	3	2	0
6.0dB	8.0E8	88	4	0	0	2	48	3	0	0
5.2dB	8.0E8	98	0	6	5	21	23	1	16	5
5.6dB	1.6E9	32	0	0	0	0	11	0	0	0

TABLE I

ERROR PROFILES FOR THE EAB $(2209, 2024)$ (TOP),AN SR $(2209, 2024)$ (MIDDLE), AND AN SSR $(2212, 1899)$ (BOTTOM) CODE. THE NUMBER OF ERRORS COLLECTED (N.E.) IS DIVIDED BY THE TOTAL NUMBER OF FRAMES (N.R.) TO PRODUCE FIG. 4 DATA.

equivalent conditions, $[a_1, a_2, ..., a_r]$ has a zero determinant if and only if $[\tilde{a}_1, \tilde{a}_2, ..., \tilde{a}_r]$ has a zero determinant. ■

Corollary 2: Since the null space of CCM also only depends on the difference of the column-group indices, analogous equivalence conditions can be established column-wise.

The following is a consequence of Lemma 5.

Corollary 3: For $r = 4$, any row-mapping vector is equivalent to some $[0, 1, x, y]$ row-mapping vector.

This result enables a reduced search of structured matrices with good error- floor properties. For example, a row-mapping vector $[0, 1, 3, 4]$ is equivalent to $[1, 2, 4, 5]$, and to $[0, 2, 6, 8]$.

IV. RESULTS

In this section we experimentally demonstrate performance improvement with well-designed SR and SSR codes. In simulations, we use 200 iterations and a $Q4.2$ fixed-point quantization with 4, resp. 2, bits to represent integer, resp. fractional, values. Our decoder employs the soft-xor algorithm [14].

In Fig. 4 we compare the performance of the $(2209, 2024)$ EAB code and SR code both with check node degree = 47 and bit node degree = 4. The EAB code has $a(i)=i$ and the SR code uses $(i, a(i)) \in \{(0, 0), (1, 1), (2, 3), (3, 4)\}$. Table I shows the collected error profiles of the two codes. Consistent with the theoretical analysis, the $(6, 4)$ absorbing sets dominate the error floor of the EAB code. The reduction of the $(6, 4)$ absorbing sets is the key reason of the performance improvement of the SR code.

We also compare the performance of a high-rate quasi-cyclic (QC) code under the construction of Tanner *et al.* [8] with a similar-rate SSR code. The QC code has the following parameters: $p=61$, $f(i, j)=a^j \cdot b^i$, $a=5$, $b=11$, with multiplicative orders $o(a)=30$ and $o(b)=4$ in $GF(61)$. There are 30 column groups and 4 row groups.

Compared to the SSR code, this is a code with the same variable-node degree, a similar block length ($N = 1830$), and a similar rate (0.8683). Using the CCM based analysis, one can show that this code does not have $(6, 4)$ absorbing sets, although it does have $(4, 4)$ absorbing sets (due to an inappropriate row mapping). The SSR code has the following parameters: $p = 79$,$f(i, j) = a(i) \cdot b(j)$, $(i, a(i)) \in \{(0, 0), (1, 1), (2, 3), (3, 4)\}$ and $b(i) \in \{2, 6, 7, 14, 17, 18, 22, 26, 27, 30, 36, 37, 38, 46, 47, 49, 55, 56, 57, 58, 61, 62, 65, 66, 67, 76, 77, 78\}$.

We thus obtain a code with the same variable-node degree, a similar block length ($N = 2212$) and a similar rate (0.8585). This code provably eliminates $(6, 4)$ absorbing sets without introducing smaller absorbing sets. The profiles in Table I also support this claim. Note the significantly lower error floor achieved by the SSR code. A similar analysis can be applied to the codes in [9].

V. CONCLUSION

This paper proposes a code construction technique based on circulant matrices suitable for applications operating at low FER levels. The cycle consistency matrix description of dominant absorbing sets provides a tool for analytical rather than heuristic code design. This approach provides codes with provably better performance than some known constructions.

REFERENCES

[1] T. Richardson. Error-floors of LDPC codes. In *Proc. 41st Annual Allerton Conf.*, Monticello, IL, Oct. 2003.

[2] L. Dolecek, Z. Zhang, M. J. Wainwright, V. Anantharam, and B. Nikolic. Analysis of absorbing sets and fully absorbing sets of array-based LDPC codes. *IEEE Trans. Inform. Theory*, 56(1), Jan. 2010.

[3] S. Laendner, T. Hehn, O. Milenkovic, and J. Huber. When does one redundant parity-check equation matter? In *Proc. IEEE Global Telecomm. Conf. (GLOBECOM)*, San Francisco, CA, Nov. 2006.

[4] O. Milenkovic, N. Kashyap, and D. Leyba. Shortened array codes of large girth. *IEEE Trans. Inform. Theory*, 5(8):3707–3722, Aug. 2006.

[5] L. Dolecek, J. Wang, and Z. Zhang. Towards Improved LDPC Code Designs Using Absorbing Set Spectrum Properties. In *Proc. of 6th International symposium on turbo codes and iterative information processing*, Brest, France, Sept. 2010.

[6] D. V. Nguyen, B. Vasic, and M. Marcellin. Structured LDPC Codes from Permutation Matrices Free of Small Trapping Sets. In *Proc. IEEE Info. Theory Workshop (ITW)*, Dublin, Ireland, Sept. 2010.

[7] J. Wang, L. Dolecek, and R. D. Wesel. LDPC Absorbing Sets, the Null Space of the Cycle Consistency Matrix, and Tanners Constructions. In *IEEE Conf. on Info. Theory and its Appl.*, San Diego, CA, Feb. 2011.

[8] R. M. Tanner, D. Sridhara, A. Sridharan, T. E. Fuja, and D. J. Costello. LDPC block and convolutional codes based on circulant matrices. *IEEE Trans. Inform. Theory*, 50(15):2966 –2984, Dec. 2004.

[9] M. P. C. Fossorier. Quasi-cyclic low-density parity-check codes from circulant permutation matrices. *IEEE Trans. Inform. Theory*, 58(8):1788–1793, Aug. 2004.

[10] Z. Zhang, L. Dolecek, B. Nikolic, V. Anantharam, and M. J. Wainwright. Design of LDPC decoders for improved low error rate performance: quantization and algorithm choices. *IEEE Trans. Comm.*, pages 3258–3268, Nov. 2009.

[11] S. Lin and D. J. Costello. *Error Control Coding*, chapter 17. Prentice Hall, 2004.

[12] R. Diestel. *Graph Theory*. Springer, 2006.

[13] J. L. Fan. Array-codes as low-density parity-check codes. In *Proc. of Second Int. Symp. on Turbo Codes*, Brest, France, Sep. 2000.

[14] M.M. Mansour and N.R. Shanbhag. High-throughput LDPC decoders. *IEEE Trans. on VLSI Systems*, 1(6):976 – 996, Dec. 2003.