This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the ICC 2007 proceedings.

Analysis of Absorbing Sets for Array-Based LDPC Codes

Lara Dolecek, Zhengya Zhang, Venkat Anantharam, Martin Wainwright, Borivoje Nikolić
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Email: {dolecek, zyzhang, ananth, wainwrig, bora}@eecs.berkeley.edu

***Abstract*— Low density parity check codes (LDPC) are known to perform very well under iterative decoding. However, these codes also exhibit a change in the slope of the bit error rate (BER) vs. signal to noise ratio (SNR) curve in the very low BER region. In our earlier work using hardware emulation in this deep BER regime we argue that this behavior can be attributed to specific structures within the Tanner graph associated with an LDPC code, called absorbing sets. In this paper we provide a detailed theoretical analysis of absorbing sets for array-based LDPC codes $C_{p,\gamma}$. Specifically, we identify and enumerate all the smallest absorbing sets for these array-based LDPC codes with $\gamma = 2, 3, 4$ with standard parity check matrix. Experiments carried out on the emulation platform show excellent agreement with our theoretical results.**

I. Introduction

Low density parity check (LDPC) codes are known to perform very well under iterative decoding based on message passing algorithms. However, as reported in previous work [6], [7], these codes often exhibit an error floor phenomenon, whereby the bit error rate (BER) vs. signal to noise ratio (SNR) curve shows a significant decrease in the slope in the very low BER region. For many applications it is imperative to reach the very low BER region without incurring a major increase in SNR. This region, however, is out of the reach of pure software simulations, and consequently the limitations of a given LDPC code under message-passing decoding in the very low BER region are largely unknown. In recent work we have begun to attack this bottleneck using a hardware emulation platform built from FPGAs [10].

The goal of this paper is to shed light on the behavior of array-based LDPC codes under iterative decoding in the very low BER regime. In order to explain and analyze the dominant causes of decoding failures we introduced the notion of an absorbing set in our earlier work [10]. These absorbing sets are related to (but not entirely equivalent to) previously introduced combinatorial structures, including stopping sets [1], trapping sets [7], near codewords [6] and pseudo-codewords [4]. Here we study in detail the structure of such sets for a class of high rate array-based LDPC codes. We prove the non-existence of certain candidate absorbing sets, and for the smallest sized absorbing sets we characterize their combinatorial structure and cardinalities. Furthermore, all error events captured through the simulations on a hardware emulator in the very low BER regime are attributed to absorbing sets, thus further confirming the importance of studying these objects.

The remainder of the paper is organized as follows. We begin in the following section with a brief overview of the array-based LDPC codes $C_{p,\gamma}$, and then formally introduce the concept of absorbing sets. In Section II we provide a detailed study of the absorbing sets for the column weights $\gamma = 2$, 3 and 4 for the standard parity check matrices $H_{p,\gamma}$ of such codes, and enumerate all such sets of smallest size. The experimental data is introduced in Section IV, and is put within the context of earlier theoretical results. In Section V we summarize the paper and propose future extensions of the work described here.

II. Background

A. Array-based LDPC codes

Array based LDPC codes [3] are regular LDPC codes parameterized by a pair of integers (γ, p), such that $\gamma \leq p$, p is an odd prime, with a parity check matrix $H_{p,\gamma}$ given by

$$H_{p,\gamma} = \begin{bmatrix} I & I & I & \dots & I \\ I & \sigma & \sigma^2 & \dots & \sigma^{p-1} \\ I & \sigma^2 & \sigma^4 & \dots & \sigma^{2(p-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ I & \sigma^{\gamma-1} & \sigma^{(\gamma-1)2} & \dots & \sigma^{(\gamma-1)(p-1)} \end{bmatrix} \quad (1)$$

where σ denotes a $p \times p$ permutation matrix of the form

$$\sigma = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}. \quad (2)$$

We use $C_{p,\gamma}$ to denote the binary linear code with the parity check matrix (1).

As demonstrated in [3], array-based LDPC codes have very good performance. They have been proposed for a number of applications, including digital subscriber lines [2] and magnetic recording [8].

In our earlier experimental work [10], we have observed that certain structures in the Tanner graph of the parity check matrix of the code are the limiting factor in the iterative decoding of several structured LDPC codes, including array-based codes. Motivated by the empirical findings, we formally defined these configurations, which we call *absorbing sets*. Here we study them in detail for array-based LDPC codes $C_{p,\gamma}$ for $\gamma = 2, 3, 4$, for the standard parity check matrix $H_{p,\gamma}$.

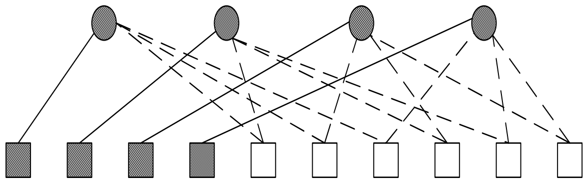

Fig. 1. An example of a (4,4) absorbing set

B. Absorbing Sets

Let $G = (V, F, E)$ be a bipartite graph with the vertex set $V \cup F$, where V and F are disjoint, and with the edge set E, such that there exists an edge $e(i, j) \in E$ iff $i \in V$ and $j \in F$. One can associate a bipartite graph $G_H = (V, F, E)$ with a parity check matrix H, such that the set V corresponds to the columns of H, the set F corresponds to the rows of H, and $E = \{e(i, j)|H(j, i) = 1\}$. Such a graph G_H is commonly referred to as the Tanner graph of the parity check matrix H of a code, [5]. Elements of V are called "bit nodes" and elements of F are called "check nodes". The Tanner graph associated with $H_{p,\gamma}$ does not have any cycles of length 4, and thus the girth is at least 6 [9]. For the subset D of V we let N_D denote the set of check nodes neighboring the elements of D.

For a subset D of V, let $\mathcal{E}(D)$ (resp. $\mathcal{O}(D)$) be the set of neighboring vertices of D in F in the graph G with even (resp. odd) degree with respect to D. Given an integer pair (a, b), an (a, b) *absorbing set* is a subset D of V of size a, with $\mathcal{O}(D)$ of size b and with the property that each element of D has strictly fewer neighbors in $\mathcal{O}(D)$ than in $F \backslash \mathcal{O}(D)$. We say that an (a, b) absorbing set D is an (a, b) *fully absorbing set*, if in addition, all bit nodes in $V \backslash D$ have strictly more neighbors in $F \backslash \mathcal{O}(D)$ than in $\mathcal{O}(D)$.

An example of an (a, b) absorbing set with $a = 4$, $b = 4$ is given in Fig. 1, where full circles constitute the set D, full squares constitute the set $\mathcal{O}(D)$, empty squares constitute the set $\mathcal{E}(D)$, $E(D, \mathcal{O}(D))$ is given by solid lines, and $E(D, \mathcal{E}(D)$ is given by dashed lines. Observe that each element in D has more even-degree than odd-degree neighbors. All check nodes not in the picture are denoted by empty squares. For this set to be a fully absorbing set, every bit node not in the figure should also have strictly more empty squares than full squares as neighbors.

Note that $D \subseteq V$ is a fully absorbing set iff for all $v \in D$, $wt(Hx_{D\Delta v}) > wt(Hx_D) = b$, where $D\Delta v$ denotes the symmetric difference between D and $\{v\}$, $wt(y)$ is the Hamming weight of a binary string y, and x_D is a binary string with support D.

We have introduced the notion of absorbing sets to qualitatively describe the convergent non-codeword state of the message passing algorithms, when the transmission channel is additive white gaussian noise (AWGN). In the asymptotic limit given by the bit flipping algorithm, the configuration described as a fully absorbing set is stable, since each bit node receives strictly more messages from the neighboring checks that reinforce its value than messages that suggest the opposite bit value.

In particular, a fully absorbing set can be viewed as a near codeword as defined in [6], though the reverse is not true, since a near codeword does not necessarily describe a stable configuration. The trapping set definition introduced in [7] also does not explicitly capture the convergent behavior since it refers to the union of all bits that are not eventually correct, and thus permits a situation in which the decoder oscillates among a finite number of states. Although stopping sets [1] also describe stable configurations, they are defined in the context of a binary erasure channel, and cannot be directly applied to an AWGN channel.

III. Theoretical Results

Our goal is to describe minimal absorbing sets and minimal fully absorbing sets (a, b) of the Tanner graph of the parity check matrix $H_{p,\gamma}$, for $\gamma = 2, 3, 4$, where the minimality refers to the smallest possible a, and where b is the smallest possible for the given a.

We use the following notation throughout the paper. For $H_{p,\gamma}$ viewed as a two-dimensional array of matrices, we let j for $0 \le j \le p - 1$ be the column-wise index in $H_{p,\gamma}$ and call it the column group j, and we let i for $0 \le i \le \gamma - 1$ be the row-wise index in $H_{p,\gamma}$ and we call it the row group (or the label) i. Let $G_{p,\gamma}$ be the Tanner graph associated with $H_{p,\gamma}$ (bit nodes and check nodes in $G_{p,\gamma}$ represent columns and rows in $H_{p,\gamma}$, respectively). Each bit node ℓ in $G_{p,\gamma}$ is uniquely indexed by (j_ℓ, k_ℓ) where j_ℓ denotes the column group of the corresponding column, and k_ℓ, $0 \le k_\ell \le p - 1$, denotes the index of that column within the column group j_ℓ it belongs to. Each check node in $G_{p,\gamma}$ receives a label i if it belongs to the row group i. Multiple check nodes can have the same label.

We note that the structure of the parity check matrix imposes the following conditions on the neighboring bit nodes and check nodes:

Vertex Consistency: For a bit node, all its incident check nodes, labelled i_{s_1} through i_{s_γ} must have distinct labels, i.e. these check nodes are in distinct row groups.

Edge Consistency: All bit nodes, say (j_{d_1}, k_{d_1}) through (j_{d_p}, k_{d_p}), participating in the same check node must have distinct j_ℓ values, i.e. they are all in distinct column groups.

Both conditions follow from the fact that the parity check matrix $H_{p,\gamma}$ of $C_{p,\gamma}$ consists of a 2-dimensional array of circulant matrices of equal size.

Our main results can be summarized as follows: Let $G_{p,\gamma}$ be the Tanner graph associated with the parity check matrix $H_{p,\gamma}$ of the array-based LDPC code $C_{p,\gamma}$.

Theorem 1: Minimality

(a) For the $G_{p,2}$ family, all minimal absorbing sets are minimal fully absorbing sets and are of size $(4, 0)$.

(b) For the $G_{p,3}$ family, the minimal absorbing sets are of size $(3, 3)$, and the minimal fully absorbing sets are of size $(4, 2)$.

(c) For the $G_{p,4}$ family, and for $p > 19$, all minimal fully absorbing sets are minimal absorbing sets, and are of size

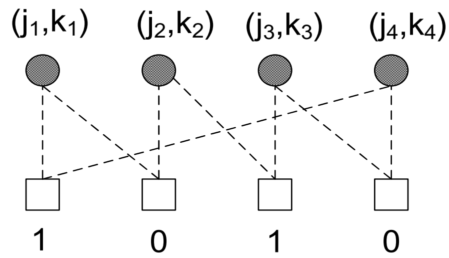

Fig. 2. (Labelled) candidate (4,0) absorbing set

$(6, 4)$. ■

Theorem 2: Scaling

(a) Suppose $\gamma = 2$ and $p > 3$. The number of minimal (fully) absorbing sets in $G_{p,\gamma}$ grows with block length n ($n = p^2$ is the number of columns in $H_{p,\gamma}$) as $O(n^2)$.

(b) Suppose that either $\gamma = 3$ and $p > 3$ or $\gamma = 4$ and $p > 19$. Then the number of minimal absorbing sets as well as the number of minimal fully absorbing sets in $H_{p,\gamma}$ grows with block length $n = p^2$ as $O(n^{3/2})$. ■

The following three subsections provide proofs of these claims, where we separately treat each of the values of γ.

A. Absorbing sets of $H_{p,2}$

The code $C_{\gamma,2}$ has uniform bit degree 2, and is thus a cycle code. Even though such codes are known to be poor, we include the analysis for the sake of completeness. We start by proving the statement in Theorem 1(a).

Let $G_{p,2} = (V, F, E)$ denote the Tanner graph of $H_{p,2}$. Let D be an (a, b) absorbing set in $G_{p,2}$. Each bit node in D has degree 2 in $G_{p,2}$ and is required to have strictly more neighbors in $\mathcal{E}(D)$ than in $\mathcal{O}(D)$. This implies that $\mathcal{O}(D)$ is empty. The absorbing set is of type $(a, 0)$. It is thus a fully absorbing set, and is in fact a codeword.

Since the matrix $H_{p,2}$ has top row consisting of identity matrices, the codewords of $C_{p,2}$ are of even weight. Moreover, $a > 2$ since no two columns of $H_{p,2}$ sum to zero. Thus $a \geq 4$.

Let (j_1, k_1), (j_2, k_2), (j_3, k_3) and (j_4, k_4) be the bit nodes participating in a $(4, 0)$ absorbing set, which must necessarily be as in Fig. 2, since there are no cycles of length 4. Consider the matrix M,

$$M = (\sigma^{0j_1})^T(\sigma^{0j_2})(\sigma^{1j_2})^T(\sigma^{1j_3})(\sigma^{0j_3})^T(\sigma^{0j_4})(\sigma^{1j_4})^T(\sigma^{1j_1}). \quad (3)$$

This matrix M has a non-zero entry on the main diagonal, and since it is itself a power of σ, it is necessary that $M = \sigma^{p\ell}$, for some ℓ, where ℓ is an integer.

Moreover, since

$$(\sigma^{\ell})^T = (\sigma^{\ell})^{-1} \quad (4)$$

it further follows that

$$p\ell = j_3 - j_2 + j_1 - j_4. \quad (5)$$

Lemma 1: There is a total of $p^2(p-1)^2$ $(4, 0)$ (fully) absorbing sets in the code described by $H_{p,2}$.

Fig. 3. Candidate (2,b) absorbing sets

Proof: It suffices to consider $\ell = 1, 0, -1$ in (5).

First, for $\ell = 1$, there are $(s + 1)(p - s - 1)$ ways of assigning values to (j_1, j_2, j_3, j_4) to make $j_2 + j_4 = s$ and $j_1 + j_3 = p + s$, for $0 \leq s \leq p - 2$. Thus, $\sum_{s=0}^{p-2}(s+1)(p - s - 1) = p(p-1)(p+1)/6$ is the total number of ways of assigning values to (j_1, j_2, j_3, j_4). By symmetry, for $\ell = -1$ there are also $p(p-1)(p+1)/6$ ways of assigning values to (j_1, j_2, j_3, j_4).

For $\ell = 0$, for each s, $0 \leq s \leq p - 1$, there are $s + 1$ ways of expressing s as a sum of an ordered pair. For s odd, each of these $s + 1$ ordered pairs can be assigned to (j_1, j_3), and for each such assignment, $s - 1$ ordered pairs can be assigned to (j_2, j_4) ($j_1 \neq j_2, j_4$ and $j_3 \neq j_2, j_4$ by the edge consistency). For s even, for s assignments out of possible $s+1$ (excluding the pair $(s/2, s/2)$) of (j_1, j_3), $s - 1$ ordered pairs can be assigned to (j_2, j_4). For the pair $(s/2, s/2)$ assigned to (j_1, j_3), there are s available assignments for (j_2, j_4). For $p - 1 \leq s \leq 2p - 2$ the number of assignments is the same as for $2p - 2 - s$. The total number of assignments for $l = 0$ is $2\sum_{i=1,i\text{odd}}^{p-2}(i + 1)(i - 1) + 2\sum_{i=2,i\text{even}}^{p-3} i^2 + (p - 1)^2 = 2p(p - 1)(p - 2)/3$.

The total number of assignments for (j_1, j_2, j_3, j_4) is then $p(p-1)^2$, and since there are in each case p ways of assigning values to (k_1, k_2, k_3, k_4), it follows that there are $p^2(p - 1)^2$ different $(4, 0)$ (fully) absorbing sets. ■

Corollary 1: The number of $(4, 0)$ (fully) absorbing sets for the code described by $H_{p,2}$ is $O(n^2)$, where n is the codeword length.

Proof: Follows immediately from Lemma 1 and $n = p^2$. ■

For $\gamma > 2$, the results are more interesting as they demonstrate the existence of minimal absorbing sets and minimal fully absorbing sets (which we have observed in our emulations to dominate the very low BER performance), for which the number of bit nodes a is strictly smaller than the minimum distance d_{min} of the code.

B. Absorbing sets of $H_{p,3}$

We now turn to the proof of Theorem 1(b).

Let $G_{p,3} = (V, F, E)$ denote the Tanner graph of $H_{p,3}$. Let D be an (a, b) absorbing set in $G_{p,3}$. Each bit node in D has degree 3 in $G_{p,3}$ and is required to have strictly more neighbors in $\mathcal{E}(D)$ than in $\mathcal{O}(D)$.

Suppose $a = 2$. In $G_{p,3}$ an even number of edges from D terminates in $\mathcal{E}(D)$. Thus either $b = 0$ or $b = 2$. These correspond to the situations in Fig. 3. In either case there would be a cycle of length 4 in $G_{p,3}$, which is false [9]. Thus $a \geq 3$.

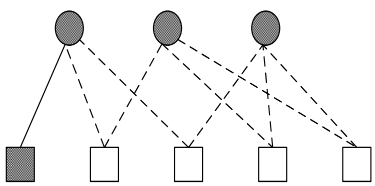

Fig. 4. Candidate (3,1) absorbing set

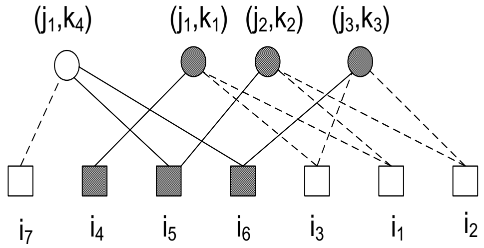

Fig. 5. Candidate (3,3) absorbing set (solid circles), with an adjacent bit node (empty circle).

Suppose $a = 3$. In $G_{p,3}$ an even number of edges from D terminates in $\mathcal{E}(D)$. Thus either $b = 1$ or $b = 3$. Suppose $b = 1$. This must correspond to the form in Fig. 4, which implies the existence of a cycle of length 4 in $G_{p,3}$, which is false [9].

Thus $b = 3$. Further, each bit node in D would then connect to exactly one check node in $\mathcal{O}(D)$ implying the unlabelled form of Fig. 5. Note that there is a cycle of length 6.

Suppose these 3 bit nodes are indexed as (j_1, k_1), (j_2, k_2) and (j_3, k_3), respectively, where j_1, j_2 and j_3 are distinct (by the edge consistency) and $0 \le j_1, j_2, j_3 \le p-1$. Without loss of generality assume that (j_1, k_1) and (j_2, k_2) share a check in the row group i_1, (j_2, k_2) and (j_3, k_3) share a check in the row group i_2, and that (j_1, k_1) and (j_3, k_3) share a check in the row group i_3, where $i_1, i_2, i_3 \in \{0, 1, 2\}$ and are distinct by the vertex consistency condition. This corresponds to the labelled representation in Fig. 5.

Consider the matrix M_1 given by

$$M_1 = (\sigma^{i_1 j_1})^T (\sigma^{i_1 j_2})(\sigma^{i_2 j_2})^T (\sigma^{i_2 j_3})(\sigma^{i_3 j_3})^T (\sigma^{i_3 j_1}). \quad (6)$$

The matrix M_1 has a non-zero entry on the main diagonal, and since it is itself a power of σ, it is necessary that $M_1 = \sigma^{p\ell}$, for some ℓ, where ℓ is an integer.

It thus follows that

$$p\ell = i_1(j_2 - j_1) + i_2(j_3 - j_2) + i_3(j_1 - j_3). \quad (7)$$

By the symmetry of the absorbing set (see Fig. 5), we may let $i_1 = 0$, $i_2 = 1$, and $i_3 = 2$. Since the column k_1 of σ^{2j_1} and column k_3 of σ^{2j_3} have a non-zero entry in the same row, it follows that

$$k_1 + 2j_1 \equiv k_3 + 2j_3 \mod p. \quad (8)$$

Likewise,

$$k_1 \equiv k_2 \mod p, \quad (9)$$
$$k_2 + j_2 \equiv k_3 + j_3 \mod p. \quad (10)$$

The existence of the solution for such a $(3, 3)$ absorbing set is given in Lemma 2.

Even though the $(3, 3)$ fully absorbing set seems plausible, care must be taken with respect to a bit node *outside* a candidate fully absorbing set when this bit node also participates in the unsatisfied checks. As we now show, such a $(3, 3)$ fully absorbing set cannot exist, though the existence of a $(3, 3)$ absorbing set implies a $(4, 2)$ fully absorbing set.

Suppose first that a $(3, 3)$ fully absorbing set exists. By definition, it is then necessary that no bit node outside of the absorbing set participates in more than one unsatisfied check described by the absorbing set. Since (j_1, k_1) and (j_3, k_3) share a check, $j_1 \neq j_3$. Hence for some k_4, the bit node labelled (j_1, k_4) connects to i_6, as in Figure 5. Since $i_3 = 2$ and $i_2 = 1$, i_6 has value 0, and $k_3 = k_4$. If the (j_1, k_3) bit node does not also participate in the check i_5 it would be necessary that $k_3 + 2j_1 \neq k_2 + 2j_2 \mod p$, which is in contradiction with (7) through (10). This eliminates a $(3, 3)$ fully absorbing set for this configuration. This also now implies a candidate $(4, 2)$ fully absorbing set with bit nodes $(j_1, k_1), (j_2, k_2)$, (j_3, k_3) and (j_1, k_3). The cases where the bit node (j_1, k_3) shares the remaining check, which we label i_7, with one of $(j_1, k_1), (j_2, k_2)$, or (j_3, k_3) can be eliminated. In the resulting $(4, 2)$ absorbing set, the unsatisfied checks are labelled i_4 and i_7. By the vertex consistency condition, $i_7 = i_4 = 1$. Since no bit node outside of this absorbing set can connect to both unsatisfied checks, this $(4, 2)$ configuration represents a fully absorbing set.

It can be shown similarly that every $(4, 2)$ fully absorbing set has the shape as the unlabelled configuration in Figure 5, and that each can be obtained from an underlying $(3, 3)$ absorbing set. Moreover for each underlying $(3, 3)$ absorbing set and for each of the three choices of the unsatisfied check pairs, there exists exactly one way of adjoining a distinct fourth bit node that neighbors these unsatisfied checks. Each resulting $(4, 2)$ fully absorbing set comes from two different $(3, 3)$ underlying configurations. Due to the space limitations, proofs of these statements are omitted.

We complete the section with the following result.

Lemma 2: The total number of $(3, 3)$ absorbing sets and $(4, 2)$ fully absorbing sets in the Tanner graph described by $H_{p,3}$ is $p^2(p-1)$, and $3p^2(p-1)/2$, respectively.

Proof: It suffices to consider $\ell = -1, 0, 1$ in (7). For $\ell = 0$ and for each value of j_1, $1 \le j_1 \le (p-1)/2$, there are $2j_1$ ways of assigning values to (j_1, j_2, j_3), and for each j_1 where $(p+1)/2 \le j_1 \le p-2$ there are $2(p-1-j_1)$ ways of assigning values to (j_1, j_2, j_3). For each assignment there are p ways to assign values to (k_1, k_2, k_3) to ensure the validity of (8)-(10). In all, there is a total of $p(p-1)^2/2$ such assignments, each describing a different $(3, 3)$ absorbing set.

Likewise, for $\ell = 1$ and $\ell = -1$ there are $p(p-1)(p+1)/4$ ways in each case to assign values to bit indices

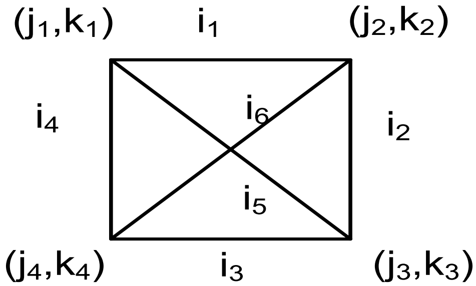

Fig. 6. Depiction of the candidate (4,4) set

$(j_1, k_1), (j_2, k_2)$ and (j_3, k_3) in the $(3,3)$ absorbing set. The total number of $(3,3)$ absorbing sets is thus $p^2(p-1)$. Depending on which two of these three bit nodes the remaining bit node (j_4, k_4) (in the $(4,2)$ fully absorbing set) shares a (satisfied) check with, we may assign (j_4, k_4) in three different ways. Note that in this way we have counted each $(4,2)$ set twice. Hence there are $3p^2(p-1)/2$ distinct $(4,2)$ fully absorbing sets. ■

Since the codeword length n is p^2, the result of Lemma 2 implies Theorem 2 for $\gamma = 3$. As a comparison, the minimum distance of the $C_{p,\gamma}$ code is 6, [9].

C. Absorbing sets of $H_{p,4}$

In order to establish that $(6,4)$ (fully) absorbing sets are minimal for $H_{p,4}$, we will first show that (a,b) absorbing sets for $a < 6$ do not exist.

Let D denote an (a,b) absorbing set in $G_{p,4} = (V, F, E)$, the Tanner graph of $H_{p,4}$. If $a = 2(3)$ then at least 6(9) edges from D in $G_{p,4}$ terminate in $\mathcal{E}(D)$, which implies the existence of a cycle of length 4 in $G_{p,4}$, which is false [9]. Thus, $a \geq 4$.

Suppose $a = 4$. The number of edges from D in $G_{p,4}$ that terminate in $\mathcal{E}(D)$ must be 12, 14, or 16, corresponding to the cases b = 4, 2, or 0, respectively. No check node in $\mathcal{E}(D)$ can connect to all four of the bit nodes in D, else there would need to be a cycle of length 4 in $G_{p,4}$, which is false [9]. Thus, each check node in $\mathcal{E}(D)$ connects to exactly two of the bit nodes in D. There are 6 pairs of nodes in D. Thus we must have $b = 4$. The following lemma establishes that such sets do not exist for large enough prime p.

Lemma 3: For $p > 7$, the Tanner graph family $G_{p,4}$ does not contain any $(4,4)$ absorbing sets.

Proof: No check node satisfied with respect to the absorbing set has degree > 2, as otherwise there would exist two bit nodes that share two distinct check nodes, which is not possible by the girth condition [9]. Similarly, all bit nodes in an absorbing set have distinct unsatisfied check nodes.

Since each bit node in the absorbing set shares exactly 3 satisfied check nodes with other bit nodes in the absorbing set, we can view the absorbing set as shown in Fig. 6 where each (labelled) vertex represents a distinct bit node and each (labelled) edge represents a check node in which the bit nodes associated with its endpoints participate.

By the edge consistency condition all j_1 through j_4 are different. Since the column k_1 of $\sigma^{i_1 j_1}$ and column k_2 of $\sigma^{i_1 j_2}$ have a non-zero entry in the same row, it follows that

$$k_1 + i_1 j_1 \equiv k_2 + i_1 j_2 \mod p$$

Likewise, for i_2 through i_6 we obtain

$$\begin{aligned}
k_2 + i_2 j_2 &\equiv k_3 + i_2 j_3 \mod p, \\
k_3 + i_3 j_3 &\equiv k_4 + i_3 j_4 \mod p, \\
k_1 + i_4 j_1 &\equiv k_4 + i_4 j_4 \mod p, \\
k_1 + i_5 j_1 &\equiv k_3 + i_5 j_3 \mod p, \text{and} \\
k_2 + i_6 j_2 &\equiv k_4 + i_6 j_4 \mod p.
\end{aligned}$$

Moreover, by imposing the vertex consistency condition and exploiting the symmetry, it suffices to consider $(i_1, i_2, i_3, i_4, i_5, i_6)$ either (x, y, x, y, z, z) or (x, y, x, y, z, w) where $x, y, z, w \in \{0, 1, 2, 3\}$ and are distinct.

For the case (x, y, x, y, z, z), we establish the following conditions based on the cycles within the graph in Fig. 6:

$$\begin{aligned}
p\ell_1 &= x(j_2 - j_1) + y(j_3 - j_2) + z(j_1 - j_3), \\
p\ell_2 &= x(j_2 - j_1) + z(j_4 - j_2) + y(j_1 - j_4), \text{and} \\
p\ell_3 &= x(j_4 - j_3) + y(j_1 - j_4) + z(j_3 - j_1).
\end{aligned}$$

for some integers ℓ_1, ℓ_2 and ℓ_3.

From this system it follows that

$$\begin{aligned}
p\ell_1' &= (y - z)(j_3 + j_4 - j_1 - j_2), \\
p\ell_2' &= (x - z)(j_2 + j_3 - j_1 - j_4), \text{and} \\
p\ell_3' &= (x - y)(j_2 + j_4 - j_1 - j_3).
\end{aligned}$$

for some integers ℓ_1', ℓ_2' and ℓ_3'. Since x, y, z are distinct, all j's would have to be the same, which contradicts the edge consistency constraint.

For the case (x, y, x, y, z, w) we have that

$$\begin{aligned}
p\ell_1 &= x(j_2 - j_1) + y(j_3 - j_2) + z(j_1 - j_3), \\
p\ell_2 &= x(j_2 - j_1) + w(j_4 - j_2) + y(j_1 - j_4), \text{and} \\
p\ell_3 &= x(j_4 - j_3) + y(j_1 - j_4) + z(j_3 - j_1).
\end{aligned}$$

for some integers ℓ_1, ℓ_2 and ℓ_3.

By manipulating above conditions one arrives at

$$(z - x)(w - y) + (z - y)(w - x) \equiv 0 \mod p. \quad (11)$$

It can be easily verified that this condition can not hold for any choice of x, y, z, w, where $x, y, z, w \in \{0, 1, 2, 3\}$ and are distinct for $p > 7$. ■

We next show that $(5, b)$ absorbing sets do not exist.

Lemma 4: For $p > 19$, the Tanner graph family $G_{p,4}$ does not contain any $(5, b)$ absorbing sets.

Proof: Since each bit node in the absorbing set has at most one neighboring unsatisfied check node, it follows that $b \leq 5$. Observe that the number of bit nodes with 3 satisfied and 1 unsatisfied check nodes is even, and thus b is even. First $b > 0$ by the minimum distance, $d_{min} \geq 8$ [9] of the code. If $b = 2$ and all satisfied check nodes had degree 2, such an absorbing set would contain a $(4,4)$ absorbing set, which by Lemma 3 does not exist. A degree-4 satisfied check node would violate the girth condition.

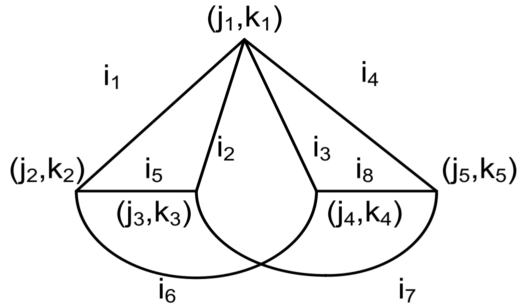

Fig. 7. Depiction of the candidate (5,4) set

We are thus left with analyzing $b = 4$ with all satisfied check nodes of degree 2. The only way that such an absorbing set could exist is if one has the configuration shown in Fig. 7, where the vertices represent bit nodes and edges represent their satisfied check nodes.

Since i_1, i_2, i_3 and i_4 are all distinct by the vertex consistency condition, we may assume that $i_1 = 0$. Then moreover, either $i_7 = 0$ or $i_8 = 0$ by the vertex consistency at (j_5, k_5).

If $i_7 = 0$ we let $x = i_2 = i_8$, $y = i_3 = i_5$ and $z = i_4 = i_6$ (by the vertex consistency condition) where $x, y, z \in \{1, 2, 3\}$ and are distinct. Note that $k_1 = k_2$ and $k_3 = k_5$. We also obtain

$$\begin{aligned} k_1 + xj_1 &\equiv k_3 + xj_3 \mod p, \\ k_1 + yj_1 &\equiv k_4 + yj_4 \mod p, \\ k_1 + zj_1 &\equiv k_5 + zj_5 \mod p, \\ k_2 + yj_2 &\equiv k_3 + yj_3 \mod p, \\ k_2 + zj_2 &\equiv k_4 + zj_4 \mod p, \text{and} \\ k_4 + xj_4 &\equiv k_5 + xj_5 \mod p. \end{aligned}$$

and likewise for $i_8 = 0$ we let $x = i_2 = i_6$, $y = i_3 = i_7$ and $z = i_4 = i_5$. Note that now $k_1 = k_2$ and $k_4 = k_5$ and

$$\begin{aligned} k_1 + xj_1 &\equiv k_3 + xj_3 \mod p, \\ k_1 + yj_1 &\equiv k_4 + yj_4 \mod p, \\ k_1 + zj_1 &\equiv k_5 + zj_5 \mod p, \\ k_2 + zj_2 &\equiv k_3 + zj_3 \mod p, \\ k_2 + xj_2 &\equiv k_4 + xj_4 \mod p, \text{and} \\ k_3 + yj_3 &\equiv k_5 + yj_5 \mod p. \end{aligned}$$

where $x, y, z \in \{1, 2, 3\}$ and distinct.

In both cases we arrive at

$$xy(z - x)(z - y) - z^2(x - y)^2 \equiv 0 \mod p, \qquad (12)$$

which does not have a solution for $p > 19$ for distinct $x, y, z \in \{1, 2, 3\}$. ■

We can now proceed with the analysis of $(6, b)$ absorbing sets. Since the number of bit nodes with 3 satisfied and 1 unsatisfied check node is even, b is even. First, $b = 0$ is not possible since $d_{min} \geq 8$ [9]. The following lemma considers $b = 2$.

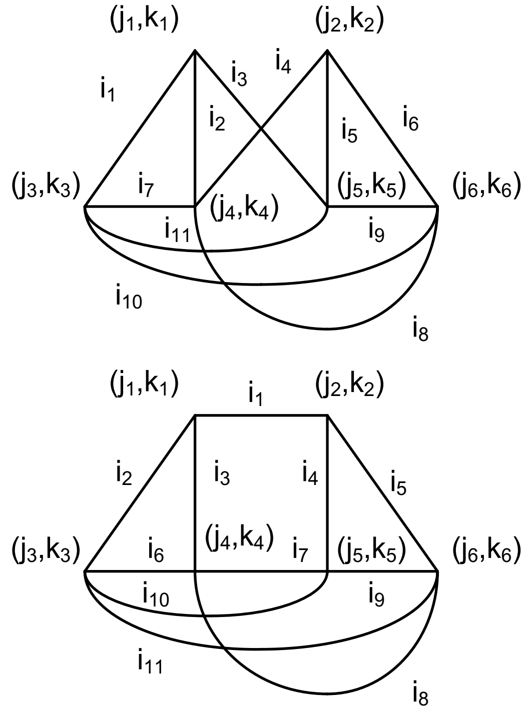

Fig. 8. Depiction of the candidate (6,2) set

Lemma 5: For $p > 19$, the Tanner graph family $G_{p,4}$ does not contain any $(6, 2)$ absorbing sets.

Proof: We first show that there is no check node of degree at least 4 with respect to the bit nodes in the absorbing set. If this were possible, there would exist at least 2 bit nodes in the absorbing set with all check nodes satisfied and with a shared check node of degree at least 4. They would necessarily share another check node, which is not possible by the girth condition [9].

We can now focus on the case where all satisfied check nodes with respect to the absorbing set have degree 2. By requiring that each vertex corresponding to a bit node in the absorbing set has either 3 or 4 outgoing edges, that there are no parallel edges and that no 3 vertices lie on the same edge, it follows that there are 2 possible configurations, as shown in Fig. 8.

Observe that the bottom configuration contains a $(4, 4)$ absorbing set which consists of (j_3, k_3), (j_4, k_4), (j_5, k_5), and (j_6, k_6). By Lemma 3 such configuration is not possible for $p > 7$.

By ensuring vertex consistency, it follows that the top configuration in Fig. 3 has 2 distinct edge labellings. Specifically, using the vertex consistency condition, we have for the top configuration

$$\begin{array}{rcll} k_1+i_1j_1 & \equiv & k_3+i_1j_3 & \mod p, \\ k_1+i_2j_1 & \equiv & k_4+i_2j_4 & \mod p, \\ k_1+i_3j_1 & \equiv & k_5+i_3j_5 & \mod p, \\ k_2+i_4j_2 & \equiv & k_4+i_4j_4 & \mod p, \\ k_2+i_5j_2 & \equiv & k_5+i_5j_5 & \mod p, \\ k_2+i_6j_2 & \equiv & k_6+i_6j_6 & \mod p, \\ k_3+i_7j_3 & \equiv & k_4+i_7j_4 & \mod p, \\ k_4+i_8j_4 & \equiv & k_6+i_8j_6 & \mod p, \\ k_5+i_9j_5 & \equiv & k_6+i_9j_6 & \mod p, \\ k_3+i_{10}j_3 & \equiv & k_6+i_{10}j_6 & \mod p, \text{and} \\ k_3+i_{11}j_3 & \equiv & k_5+i_{11}j_5 & \mod p, \end{array}$$

and either $x=i_1=i_5=i_8$, $y=i_7=i_9$, $z=i_2=i_6=i_{11}$, $w=i_3=i_4=i_{10}$ or $x=i_1=i_4=i_9$, $y=i_3=i_6=i_7$, $z=i_8=i_{11}$, $w=i_2=i_5=i_{10}$ where throughout x,y,z,w are distinct and belong to the set $\{0,1,2,3\}$.

In each case, the system of constraints reduces to one of the following constraints:

$$\begin{array}{rcll} \tilde{x} & \equiv & \tilde{y} & \mod p, \quad \text{or} \\ \tilde{x}\tilde{z}(\tilde{y}-\tilde{z})(\tilde{x}-\tilde{y}) & \equiv & \pm\tilde{y}^2(\tilde{x}-\tilde{z})^2 & \mod p, \end{array}$$

where $\{\tilde{x},\tilde{y},\tilde{z}\}=\{x,y,z\}=\{1,2,3\}$ and are distinct.

It can be verified that the above conditions cannot hold for $p>19$. ■

Lemma 6: For all $p>5$, the Tanner graph family $G_{p,4}$ has $(6,4)$ (fully) absorbing sets.

Proof: Suppose first that there exists a check node of degree 4 with respect to a $(6,4)$ absorbing set. Let t_1,t_2,t_3,t_4 be the bit nodes in the absorbing set participating in degree-4 check node, and let t_5 and t_6 be the remaining two bit nodes in the absorbing set. If at least one of t_1,t_2,t_3,t_4 had all check nodes satisfied, it would be necessary that such a bit node shares another distinct check node with some other bit node participating in the degree-4 check node, which is impossible by the girth [9]. Thus, all of t_1,t_2,t_3,t_4 are each connected to 3 satisfied and 1 unsatisfied check node. Then t_5 and t_6 are connected to 4 satisfied check nodes. Let i_j for $1\le j\le 4$ be the labels of the check nodes connecting t_j and t_5. By the vertex consistency condition at t_5, they are all different. By the vertex consistency condition at each of t_j for $1\le j\le 4$ the label of their shared degree-4 check node must be different from all i_j for $1\le j\le 4$, which is impossible as there are only 4 distinct labels available. Therefore all satisfied check nodes neighboring bit nodes in the absorbing set have degree 2.

One can show that there are 3 possible non isomorphic configurations, as shown in Fig. 9. By ensuring the vertex consistency, it further follows that for each configuration there are 8 distinct edge labellings. Let us consider the topmost configuration. The other two configurations can be analyzed similarly. As before, we establish a congruential constraint for each triplet consisting of an edge and its endpoints given in

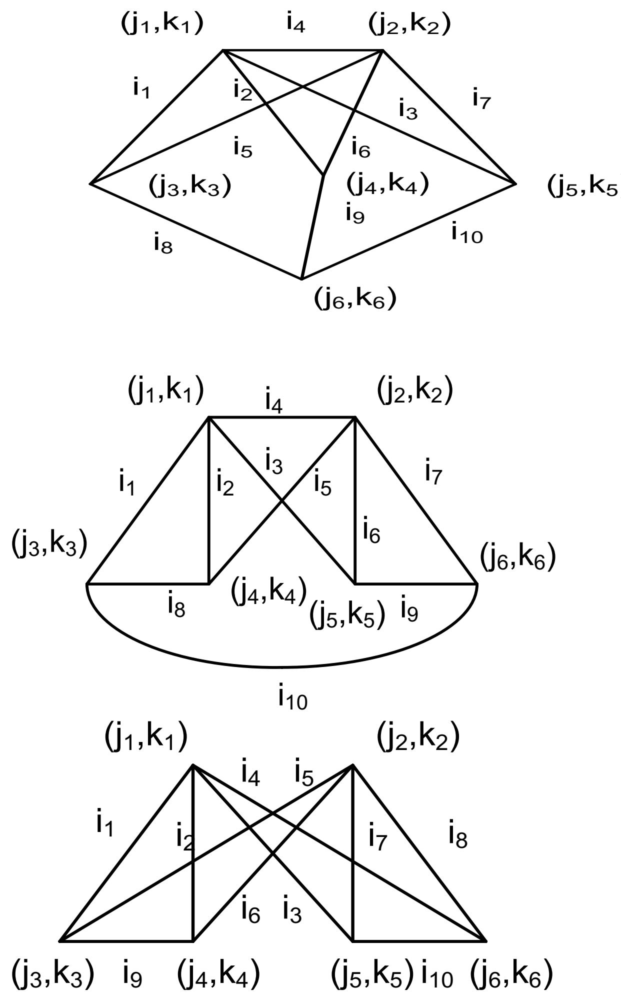

Fig. 9. Depiction of the candidate (6,4) sets

the topmost Figure 9. For example, the edge i_1 and vertices (j_1,k_1) and (j_3,k_3) satisfy

$$k_1+i_1j_1\equiv k_3+i_1j_3 \mod p.$$

The set of constraints consists of 10 such equations, one for each edge. For the topmost configuration, the valid edge labellings are by the vertex consistency as follows. Let $(i_1,i_2,i_3,i_4)=(x,y,z,w)$, for $x,y,z,w\in\{0,1,2,3\}$ and distinct. Then,

$$\begin{aligned}(i_5,i_6,i_7,i_8,i_9,i_{10})\in\{&(y,z,x,z,x,y),(z,x,y,y,z,x),\\&(y,z,x,z,w,y),(y,z,x,w,x,y),(y,z,x,z,x,w),\\&(z,x,y,y,z,w)(z,x,y,y,w,x),(z,x,y,w,z,x)\}.\end{aligned}$$

For the 6-tuple $(i_5,i_6,i_7,i_8,i_9,i_{10})$= (y,z,x,z,x,y) and $x=0$, the solution set formed from the 10 congruence constraints is given in the table in Fig. 10, where q,s are any residues $\mod p$, and t is chosen in the residue class $\mod p$ such that all resulting j values that depend on it are themselves integers. All resulting values are taken $\mod p$. Note that we have thus established the existence of $(6,4)$ absorbing sets.

j_1	j_2	j_3	j_4	j_5	j_6	k_1	k_2	k_3	k_4	k_5	k_6
q	$q+6t$	$q+4t$	$q-6t$	$q+3t$	$q-5t$	s	$s-6t$	s	$s+18t$	$s-6t$	$s+18t$
q	$q+3t$	$q+t$	$q+3/2t$	$q+6t$	$q+11/2t$	s	$s-6t$	s	$s-9/2t$	$s-6t$	$s-9/2t$
q	$q+6t$	$q+3t$	$q+12t$	$q+2t$	$q+11t$	s	$s-6t$	s	$s-24t$	$s-6t$	$s-24$
q	$q+2t$	$q-t$	$q+4t$	$q+6t$	$q+7t$	s	$s-6t$	s	$s-8t$	$s-6t$	$s-8t$
q	$q+2t$	$q-4t$	$q-2t$	$q+3t$	$q-5t$	s	$s-6t$	s	$s+2t$	$s-6t$	$s+2t$
q	$q+3t$	$q-3t$	$q+3/2t$	$q+2t$	$q-5/2t$	s	$s-6t$	s	$s-3/2t$	$s-6t$	$s-3/2t$

Fig. 10. Several solution sets for the (6,4) configuration

The (absolute) indices of columns that correspond to the bit nodes in the absorbing set are $k_i + pj_i$ for $1 \leq i \leq 6$ and the indices of rows that correspond to the unsatisfied check nodes in the absorbing set are $[k_i + j_i w]_p + wp$, for $3 \leq i \leq 6$. In particular, the solution set in row 1 holds for all $p > 5$.

Furthermore, for this 6-tuple all 4 unsatisfied checks in a candidate $(6,4)$ absorbing set belong to the same row group w. No bit node outside of such a set can be connected to more than one of these unsatisfied checks, and therefore this configuration is in fact a $(6,4)$ fully absorbing set. ■

Using these results the proof of Theorem 1(c) now follows. We complete our analysis of $\gamma = 4$ by proving the claim in Theorem 2: The number of $(6,4)$ (fully) absorbing sets scales as $O(n^{3/2})$, where n is the codeword length.

Proof: For the topmost configuration in Fig. 9 there are 8 distinct candidate edge labellings. For each labelling there are at most $4! = 24$ ways of assigning numerical values to labels. For each such assignment there are three parameters (as illustrated in example in Fig. 10) that determine all of j's and k's, and each parameter is chosen independently in at most p ways (to ensure the all j's and k's have integer values), yielding an upper bound which grows as $O(p^3)$. A lower bound on the cardinality of the $(6,4)$ fully absorbing sets is given by the solution set in Table 10, which also grows as $O(p^3)$. Likewise, the number of $(6,4)$ (fully) absorbing sets for the remaining two configurations grows as $O(p^3)$. Since $n = p^2$, the result follows. ■

IV. Experimental Results

The experiments were carried out using an LDPC code emulator described in detail in [10]. The decoder was implemented using a 4.5 (4 bits for integer and 5 bits for the fractional part) uniform quantization. The all-zero codeword was transmitted and the decoder was set to run for at most 200 iterations, halting earlier if decoding to a codeword. The frame error rate and the bit error rate for the $C_{47,4}$ code are shown in Fig. 11, along with the uncoded BER curve. In the error-floor regime all errors were found to be due to fully absorbing sets. A total of 25 errors were recorded at SNR = 6.4 dB, of which 18 errors were of smallest weight and all due to $(6,4)$ fully absorbing sets.

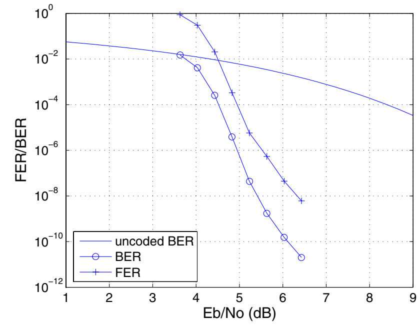

Fig. 11. Experimental Results for $C_{47,4}$

V. Conclusion and Future Work

This paper presented a detailed analysis of the dominant configurations in the error-floor regime of high rate array-based LDPC codes. We provided an explicit description of the minimal (fully) absorbing sets and showed the non-existence of certain candidate configurations. We also enumerated minimal (fully) absorbing sets and showed how their number scales with the codeword length. Experiments on an emulation platform were performed and were found to be in agreement with the theoretical description of the dominant errors. We anticipate that the techniques and analysis performed in the current work can be fruitfully extended to a larger class of structured LDPC codes.

Acknowledgment

This research was supported in part by NSF CCF grant no. 0635372, Marvell Semiconductor and Intel Corp. through UC MICRO program. NSF CNS RI grant no. 0403427 provided the computing infrastructure.

References

[1] C. Di, D. Proietti, T. Richardson, E. Telatar, and R. Urbanke, "Finite length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inform. Theory*, vol. (48), pp. 1570–1579.

[2] E. Eleftheriou and S. Ölçer, "Low density parity check codes for digital subscriber lines", *Proc. of the IEEE Int. Conf. on Commun.*, New York, NY, 2002, pp. 1752–1757.

[3] J. L. Fan, "Array-codes as low-density parity-check codes," *Second Int. Symp. on Turbo Codes*, Brest, France, pp. 543–46, Sept. 2000.

[4] J. Feldman, M. J. Wainwright and D. R. Karger, "Using linear programming to decode binary linear codes", *IEEE Trans. Inform. Theory*, vol. (51), pp. 954–972.

[5] G. D. Forney, "Codes on graphs: normal realizations", *IEEE Trans. Inform. Theory*, vol. (47), pp. 520–548, Feb. 2001.

[6] D. MacKay and M. Postol, "Weaknesses of Margulis and Ramanujan-Margulis low-density parity-check codes," *Elec. Notes in Theo. Comp. Sci.*, vol. 74, 2003.

[7] T. Richardson, "Error-floors of LDPC codes," in *Proc. of the 41st Annual Allerton Conference*, Monticello, Ill., Oct. 2003.

[8] B. Vasic and E. Kurtas, "Coding and Signal Processing for Magnetic Recording Systems", CRC press, 2005.

[9] K. Yang and T. Helleseth, "On the minimum distance of array codes as LDPC codes", *IEEE Trans. Inform. Theory*, vol. 49(12), pp. 3268–3271, Dec. 2003.

[10] Z. Zhang, L. Dolecek, B. Nikolic, V. Anantharam and M. Wainwright, "Investigation of error floors of structured low-density parity-check codes via hardware simulation", *Proc. GLOBECOM 2006*, San Francisco, CA, Nov.-Dec. 2006.