

Analysis and Enumeration of Absorbing Sets for Non-Binary Graph-Based Codes

Behzad Amiri

Electrical Eng. Department
University of California, Los Angeles
Los Angeles, USA
Email: amiri@ucla.edu

Jörg Kliewer

Klipsch School of Electrical and Computer Eng.
New Mexico State University
Las Cruces, USA
Email: jkliewer@nmsu.edu

Lara Dolecek

Electrical Eng. Department
University of California, Los Angeles
Los Angeles, USA
Email: dolecek@ee.ucla.edu

Abstract—This work provides a generalization of absorbing sets for linear channel codes over non-binary alphabets. In a graphical representation of a non-binary channel code, an absorbing set can be described by a collection of topological and edge labeling conditions. In the non-binary case the equations relating neighboring variable and check nodes are over a non-binary field, and the edge weights are given by the non-zero elements of that non-binary field. As a consequence, it becomes more difficult for a given structure to satisfy the absorbing set constraints. This observation in part explains the superior performance of non-binary codes over their binary counterparts. We first show that, as the field order size increases, the ratio of trapping sets that satisfy the structural conditions of absorbing sets decreases. This suggests that a trapping set-only performance estimation of non-binary codes may not be as accurate in the error floor/high reliability regime. By using both insights from graph theory and combinatorial techniques, we establish the asymptotic distribution of non-binary elementary absorbing sets for regular code ensembles. Finally, we provide design guidelines for finite-length non-binary codes free of small absorbing sets.

I. INTRODUCTION

It is well known that non-binary low density parity check (NB-LDPC) codes offer performance improvement over their binary counterparts. Recent results, e.g., [1], [2] are examples of such constructions. It is also well-known that under low-complexity but suboptimal message passing algorithms, certain non-codewords are vying with codewords to be the output of the decoder. The presence of these non-codeword objects can significantly undermine the performance of iteratively decoded graph based codes and may even result in an undesirable error floor. Given the significance of the error floor behavior for the finite block-length performance of coding schemes, extensive recent work has been devoted to understanding and eliminating this phenomenon for the binary case.

In contrast, for the non-binary case much less is known about how non-codeword objects and specific substructures in the Tanner graph affect the error floor performance. Important recent results include stopping constellation distributions for irregular NB-LDPC code ensembles [3], as well as the analysis of zig-zag cycles and their relationship with the error floor [4], [5]. Further, ensemble enumerators for stopping set and

trapping set sizes are found in [6] for protograph-based ensembles of NB-LDPC codes. However, apart from the recent result in [7], where only a simplified absorbing set (AS) definition compared to the one in [8] without further analysis is given, no other results are known for non-binary absorbing sets (NB-AS) to the best of our knowledge.

The main goals of this paper are multifold: 1) to define and analyze NB-ASs for codes over $\text{GF}(q)$, $q > 2$, 2) to highlight the difference between binary ASs and NB-ASs, 3) to compare the distribution of NB-ASs with the one for non-binary trapping sets, and 4) to propose an efficient code design based on our classification of ASs.

II. NON-BINARY ABSORBING SETS

We quickly summarize the well-known definitions of (elementary) trapping and absorbing sets in the binary regime before motivating our new definition of NB-AS. Consider a subgraph of the Tanner graph of a (binary) code induced by a variable nodes, given by the node set \mathcal{V} . Set all variable nodes in \mathcal{V} to 1 and set all other variable nodes to 0. Let $\mathcal{O}(\mathcal{E})$ be the set of check nodes connected to the set \mathcal{V} an odd (even) number of times. Clearly, the set \mathcal{E} represents the satisfied checks, and the set \mathcal{O} the unsatisfied checks, respectively.

Definition 1. [9] *The set \mathcal{V} is an (a, b) trapping set (TS) if $|\mathcal{O}| = b$.*

Definition 2. [8] *The set \mathcal{V} is an (a, b) AS if $|\mathcal{O}| = b$ and if each of the a variable nodes has strictly more neighbors in \mathcal{E} than in \mathcal{O} .*

Definition 3. [9], [10] *An elementary AS (TS) is an AS (TS) with each of its neighboring satisfied checks having two edges connected to the AS (TS), and each of its neighboring unsatisfied checks having exactly one edge connected to the AS (TS).*

Figure 1(a) shows a binary $(4, 4)$ elementary AS (EAS). We tacitly assume that the edge weights (edge labels) are $w_i = 1$, $i = \{1, 2, \dots, 16\}$.

In contrast to binary codes, each edge in the Tanner graph of a non-binary code admits a label taken as a non-zero element of the underlying non-binary field. Consequently, the edges of the subgraphs which correspond to the fixed points of NB-LDPC decoders need not only be topologically connected in

This work was supported in part by NSF grants CCF-0830666, CCF-1161774, and CCF-1161798.

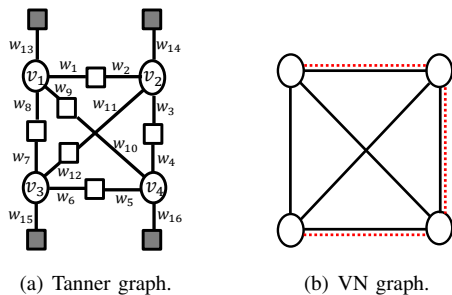


Fig. 1. (4, 4) AS and its VN graph.

specific ways, but also the labels on these edges must satisfy certain conditions. In other words, suppose we consider a topology (with no edge labeling) that satisfies the conditions of the (binary) AS, as given by Definition 2. Once the edge labels, taken over some field $\text{GF}(q)$ are chosen, the resulting object, when interpreted over $\text{GF}(q)$, may or may not cause a decoding error.

Example 1 illustrates the difference between binary AS and NB-AS and gives a motivation for the definition of NB-AS.

Example 1: Consider the graphical structure in Figure 1(a). If there exists a set of non-zero inputs for all variable nodes that makes all degree-two check nodes satisfied, the resulting configuration would have 4 unsatisfied checks and each variable node would have strictly more satisfied than unsatisfied neighboring checks (3 vs. 1). Mathematically, the inputs v_1, v_2, v_3, v_4 and weights of the edges w_1, \dots, w_{12} would then satisfy:

$$\begin{aligned} v_1 w_1 &= v_2 w_2, v_2 w_3 = v_4 w_4, v_4 w_5 = v_3 w_6, \\ v_3 w_7 &= v_1 w_8, v_2 w_{11} = v_3 w_{12}, v_1 w_9 = v_4 w_{10}, \end{aligned}$$

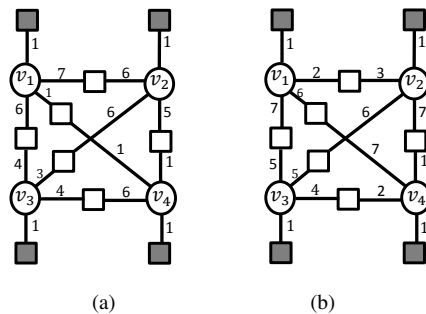
which leads to the following conditions:

$$\begin{aligned} w_1 w_7 w_{11} &= w_2 w_8 w_{12}, & w_3 w_5 w_{12} &= w_4 w_6 w_{11}, \\ w_2 w_4 w_9 &= w_1 w_3 w_{10}, \end{aligned} \quad (1)$$

where the calculations are over $\text{GF}(q)$, and all w and v are non-zero elements of $\text{GF}(q)$.

For example, for $q = 8$, Figure 2(a) shows a choice of weights satisfying the conditions in (1). With these weights, there are $q - 1$ choices (out of $(q - 1)^4$) for the set (v_1, v_2, v_3, v_4) such that each variable node has exactly 3 satisfied and 1 unsatisfied checks. One example is $(1, 2, 4, 1)$. In contrast, the weights in Figure 2(b) do not satisfy the conditions in (1) and result in a configuration that has 4 variable nodes and z unsatisfied checks. Here, $5 \leq z \leq 10$, and the value of z depends on the input values v . For example, the same input $(1, 2, 4, 1)$ results in $z = 10$. Clearly, this configuration with all 10 neighboring checks being unsatisfied is not expected to be problematic for belief propagation decoding. ■

Example 1 provides a motivation for studying NB-AS. Consider a NB-LDPC code with an $m \times n$ parity check matrix H defined over $\text{GF}(q)$. The corresponding Tanner graph has n variable nodes and m check nodes. Definition 4 states the conditions for a subset of variable nodes with cardinality a , $\mathcal{V} = \{v_i, v_j, \dots, v_k\}$ to be an (a, b) NB-AS.


 Fig. 2. (a) Non-binary elementary (4, 4) AS over $\text{GF}(8)$, (b) $(4, z)$ TSs, weights do not satisfy AS conditions over $\text{GF}(8)$.

We form an $\ell \times a$ matrix A , a submatrix of matrix H , consisting of the columns of matrix H that correspond to variable nodes in \mathcal{V} and ℓ check nodes connected to \mathcal{V} .

Definition 4. The configuration \mathcal{V} is an (a, b) AS over $\text{GF}(q)$ if there exists an $(\ell - b) \times a$ submatrix B of rank r_B , with elements $b_{j,i}, 1 \leq j \leq \ell - b, 1 \leq i \leq a$, in matrix A that satisfies the following conditions.

- 1) Let $N(B)$ be the null-space of matrix B and let $\mathbf{d}_i, 1 \leq i \leq b$ be the i th row of matrix D , where D is formed by excluding the matrix B from A . Then,

$$\exists \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_a \end{bmatrix} \in N(B) \text{ s.t. } x_i \neq 0 \text{ for } \forall i \in \{1, \dots, a\} \text{ and } \nexists i, \mathbf{d}_i \mathbf{x} = 0. \quad (2)$$

- 2) Let $d_{j,i}, 1 \leq j \leq b, 1 \leq i \leq a$, be the elements of the matrix D . Then,

$$\forall i \in \{1, 2, \dots, a\} : \left(\sum_{j=1}^{\ell-b} S(b_{j,i}) \right) > \left(\sum_{j=1}^b S(d_{j,i}) \right), \quad (3)$$

where the function S is

$$S(x) = \begin{cases} 1 & \text{when } x > 0 \\ 0 & \text{when } x \leq 0. \end{cases} \quad (4)$$

Condition 1 in Definition 4 requires that there exists a vector in the null-space of the matrix B all of whose elements are non-zero. Therefore, there exists a solution to $B\mathbf{x} = 0 \pmod{q}$ such that all components of the solution are non-zero. A consequence of this condition is that $r_B < a$. Also, condition 1 guarantees that for vector \mathbf{x} in the null-space of matrix B , none of the check nodes associated with the rows of the matrix D are satisfied (otherwise, input \mathbf{x} results in an $(a, \tilde{b}), \tilde{b} < b$, AS). Condition 2 ensures that for each variable node, the number of connected satisfied checks is larger than the number of connected unsatisfied checks.

Observe that the proposed definition is in agreement with the existing definition of a binary AS [8]. In particular, for $q = 2$, the condition 1 is automatically satisfied since the variable nodes input vector must be an all-ones vector to satisfy the checks. Also, the condition that each variable node in the AS has strictly more neighbors in \mathcal{E} than in \mathcal{O} corresponds to condition 2.

Remark 1. Our definition of a NB-AS is different from the definition proposed in [7] where the authors defined a non-binary (primitive) AS as an object which has more satisfied

checks than unsatisfied checks, taken collectively over all variable nodes in this object. In contrast, our definition, similar to the original definition of binary ASs [8], requires each variable node to be connected to more satisfied checks than unsatisfied checks.

Remark 2. Note that whether an NB-AS as defined in Definition 4 results in a decoding error depends on the choice of the variable nodes inputs. The set of all \mathbf{x} 's which satisfy (2) is the set of all variable nodes inputs that result in an (a, b) NB-AS. Other choices of variable nodes inputs results in other (a, \tilde{b}) , $\tilde{b} \neq b$ TSs.

As in the binary case, we say that a NB-AS \mathcal{V} is **elementary** if all neighboring satisfied checks have degree 2 and all neighboring unsatisfied checks have degree 1 with respect to \mathcal{V} . An elementary NB-AS is necessarily an elementary binary AS where all non-zero edge labels and non-zero variable node values are converted to 1 and all operations are taken modulo 2. Clearly, the converse is not true as the choice of non-binary labels may violate the AS constraints.

In the case of elementary ASs, the AS conditions can be simplified, as the following lemma shows. Consider a code \mathcal{C} with a parity check matrix H over $\text{GF}(q)$. Let \mathcal{G} be its Tanner graph. Let C_p be an arbitrary cycle involving p distinct variable nodes and p distinct neighboring check nodes in the graph induced by an elementary (a, b) NB-AS in \mathcal{G} , $p \leq a$. We write C_p as the oriented traversal $c_1 - v_1 - c_2 - v_2 - \dots - c_p - v_p - c_1$, where v and c denote the spanned variable and check nodes, respectively. Let w_{2i+1} denote the label on the $c_i - v_i$ edge, and let w_{2i} denote the label on the $v_i - c_{i+1}$ edge.

Lemma 1. In the case of elementary ASs over $\text{GF}(q)$, the conditions in Definition 4 can be simplified to:

- 1) Unlabeling of all edges (converting all edge weights to one) of the non-binary subgraph results in an elementary binary AS.
- 2) For every cycle C_p (as introduced above), the weights of the edges w_i , $i \in \{1, 2, \dots, 2p\}$, satisfy the following relation:

$$\prod_{k=1}^p w_{2k-1} = \prod_{k=1}^p w_{2k} \pmod{q}. \quad (5)$$

Proof: Condition 1 is immediate. For the cycle C_p (of length $2p$) we form a $p \times p$ submatrix B_{C_p} of H , corresponding to the p variable nodes and p check nodes in C_p . Since the check nodes in B_{C_p} are satisfied, there exists a non-zero solution x to $B_{C_p}x = 0 \pmod{q}$. Therefore, for the square matrix B_{C_p} , we have $\det(B_{C_p}) = 0 \pmod{q}$. Thus,

$$\det(B_{C_p}) = \det \begin{bmatrix} w_1 & w_2 & 0 & \dots & 0 \\ 0 & w_3 & w_4 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & w_{2p-3} & w_{2p-2} \\ w_{2p} & 0 & \dots & 0 & w_{2p-1} \end{bmatrix} = 0 \pmod{q}. \quad (6)$$

Now, in order to satisfy (6) the condition in (5) must hold. Since every permutation of the columns of B_{C_p} results in same determinant, the chosen ordering was performed without loss of generality. ■

Definition 5. A set of cycles \mathcal{F} is a cycle span of a graph if all the cycles of the graph can be constructed as a combination of the cycles in \mathcal{F} , and this set has the fewest cycles. The cycles in the cycle span are called fundamental cycles.

It can be shown that if the cycle span satisfies (5), this also holds for all other cycles in the graph. Thus, a subgraph of the Tanner graph is an elementary NB-AS if it satisfies the topological conditions, i.e., the unlabeled subgraph satisfies the conditions of an elementary binary ASs, and all of its fundamental cycles satisfy (5).

Remark 3. Conditions similar to Lemma 1 are presented in [4] but only for regular codes of column weight 2. In the case of column weight 2, the smallest ASs correspond to the minimum weight codewords, and moreover contain only one (fundamental) cycle. The results presented in this paper apply to general column weights wherein ASs may be spanned by more than one fundamental cycle.

Example 2: Let us interpret the configuration in Figure 1(a) as a $(4, 4)$ NB-AS. The conditions on the cycles in the cycle span (containing 3 cycles) in (5) are precisely the one given in (1). If, however, (5) is violated for one of the three fundamental cycles, the resulting configuration will no longer be an AS; it will instead become a $(4, z)$, $5 \leq z \leq 12$, trapping set (where the value of z depends on the input and on the number of unsatisfied checks). ■

The following result is easy to prove.

Lemma 2. Based on the non-binary labels, an unlabeled (a, b) (i.e., binary) elementary AS becomes an (a, b^+) AS/TS in the non-binary case (b^+ denotes any integer greater than or equal to b).

III. ASYMPTOTIC DISTRIBUTION OF NB-AS IN REGULAR CODE ENSEMBLES

In this section, we find the asymptotic distribution of non-binary elementary ASs in regular code ensembles by using techniques from graph theory. We show that in the non-binary regime, as the alphabet size q gets larger, it is harder to satisfy edge labeling conditions of ASs.

Definition 6. [11] For a bipartite graph \mathcal{G} corresponding to an elementary AS, a variable node (VN) graph is constructed by defining variable nodes of \mathcal{G} as its vertices and those checks that connect the variable nodes in \mathcal{G} as its edges.

Figure 1(b) shows the VN graph for the $(4, 4)$ AS displayed in Figure 1(a). The edges in Figure 1(b) marked in dashed red form a spanning tree of the VN graph. Note that a spanning tree is not unique. Also, from graph theory we know that for an undirected connected graph with a vertices the spanning tree has $a - 1$ edges, and that adding an edge to the spanning tree creates a fundamental cycle. Therefore, there is a one-to-one correspondence between the fundamental cycles and the edges outside the spanning tree. As a result, the corresponding VN graph of \mathcal{G} (which is a subgraph of the Tanner graph) with a vertices (variable nodes) and e edges (satisfied checks) has $e - a + 1$ fundamental cycles.

Theorem 1. Given an (a, b) unlabeled (i.e., binary) elementary AS with e satisfied checks, a fraction of $(q - 1)^{a-e-1}$ of

all possible edge label assignments taken from $GF(q)$ results in (a, b) elementary NB-ASs.

Proof: The VN graph of the given unlabeled elementary AS includes a vertices and e edges. Each edge in the spanning tree of the VN graph corresponds to two edges in the Tanner graph. The labels of the edges in the Tanner graph that are represented in the spanning tree can be chosen arbitrarily. Therefore, we have $(q-1)^2$ choices for the weights for each edge. Thus, there are $(q-1)^{2(a-1)}$ weight assignments for the $a-1$ edges in the spanning tree.

For each of the remaining $e-(a-1)$ edges in the VN graph that are not in the spanning tree, one of the two edges in the Tanner graph can again be chosen arbitrarily but the other edge is uniquely determined according to (5). Thus, there are $(q-1)^{e-a+1}$ weight assignments for edges not in the spanning tree. Hence, the total number of the weight assignments resulting in elementary NB-ASs is given as

$$(q-1)^{2(a-1)} \cdot (q-1)^{e-a+1} = (q-1)^{a+e-1}.$$

Since, we have a total of $(q-1)^{2e}$ possible weight assignments, the resulting fraction follows. ■

Based on these results we have the following corollary. Note that to distinguish unlabeled ASs and NB-ASs, we refer to the weight assignments obeying the conditions in Definition 4 as *problematic*.

Corollary 1. For regular code ensembles with column weight c and a given unlabeled (a, b) elementary AS, the fraction of problematic weight assignments is given as $(q-1)^{\frac{b-(c-2)a-2}{2}}$.

We now consider the asymptotic distribution of the NB-ASs. Let $\mathcal{G}_{n,m,q}^{c,r}$ denote the regular code ensemble with parity check matrices from the set $\Lambda_{n,m,q}^{c,r}$ which consists of all regular matrices of column-weight c and row weight r with elements from $GF(q)$. The asymptotic distribution of NB-ASs in the regular code ensemble is defined as follows:

$$e^{c,r}(\theta, \lambda) \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{z_{a,b,n}^{c,r}}{|\Lambda_{n,m,q}^{c,r}|}, \quad (7)$$

where $\theta = \frac{a}{n}$, $\lambda = \frac{b}{n}$, and $z_{a,b,n}^{c,r}$ is the number of (a, b) ASs over all matrices in $\Lambda_{n,m,q}^{c,r}$. In (7) and in the following, \log is taken to the base e .

The following lemma specifies the normalized asymptotic distribution for the $c=3$ case.

Lemma 3. The normalized logarithmic asymptotic distribution of $(a = \theta n, b = \lambda n)$ elementary ASs in $\mathcal{G}_{n,m}^{3,r}$ regular code ensemble over $GF(q)$ is given as

$$\begin{aligned} e^{3,r}(\theta, \lambda) &= -2H_b(\theta, 1-\theta) - H_b(\zeta, 1-\zeta) + \lambda \log 3r \\ &+ H_b\left(1-\zeta, \lambda, \frac{3\theta-\lambda}{2}, \frac{2\zeta-3\theta-\lambda}{2}\right) \\ &+ \frac{3\theta-\lambda}{2} \log \binom{r}{2} + \theta H_b\left(\frac{\lambda}{\theta}, 1-\frac{\lambda}{\theta}\right) \\ &- 3\theta H_b\left(\frac{\lambda}{3\theta}, 1-\frac{\lambda}{3\theta}\right) - \frac{\theta-\lambda}{2} \log(q-1), \end{aligned} \quad (8)$$

where $\zeta = \frac{m}{n}$, and $H_b(p_1, \dots, p_N) = -\sum_{i=1}^N p_i \log p_i$ with $\sum_{i=1}^N p_i = 1$ denotes the entropy function.

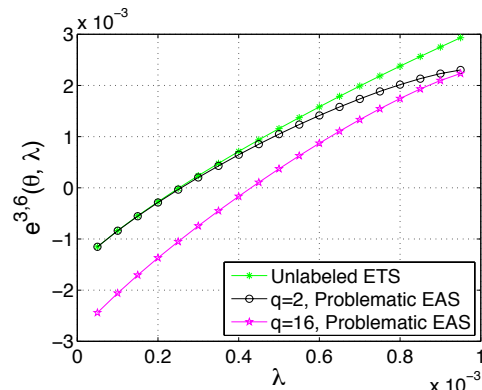


Fig. 3. Normalized logarithmic asymptotic distributions of elementary TSs (ETS) and ASs (EAS) for fixed $\theta = 0.001$ for $G_{n,m}^{3,6}$, $q = 2, 16$.

Proof: Using Corollary 1, from all unlabeled (binary) $(\theta n, \lambda n)$ ASs in the Tanner graphs of regular code ensembles (enumerated in [10]), a fraction of $(q-1)^{\frac{(\lambda-(c-2)\theta)n-2}{2}}$ of all possible weight assignments results in $(\theta n, \lambda n)$ NB-ASs. Details are omitted due to space constraints. ■

Note that (8) indicates that the normalized asymptotic number of ASs in the ensemble is a *decreasing* function in the field size q for fixed θ and λ .

For a $(3, 6)$ regular ensemble, Figure 3 shows the asymptotic distribution of unlabeled $(\theta n, \lambda n)$ elementary (binary) TSs using the results in [9], as well as $(\theta n, \lambda n)$ problematic elementary ASs for $q = 2, 16$ and fixed $\theta = 0.001$. We can observe that in particular for larger λ and for $q = 16$ the normalized asymptotic number of elementary ASs is smaller than the one for elementary TSs. The reason for this behavior is that after the weight assignment in the non-binary case a smaller fraction of unlabeled TSs leads to problematic ASs. Therefore, the resulting NB-AS enumerators provide a better assessment of the error floor compared to TS enumerators. Also, as the value of λ gets smaller, both curves for unlabeled elementary TSs and binary elementary AS converge, which means that most of the unlabeled elementary TSs result in problematic elementary ASs in this case.

IV. FINITE LENGTH ANALYSIS

In this section, we propose an algorithm to improve the performance of non-binary codes in the error floor region. We exploit the fact that non-binary edge weights enable us to reduce the number of ASs by just changing the weights of edges in the Tanner graph without changing its structure. The method is stated in Algorithm 1. Note that AS A_y is a child of AS A_z if A_z is a subgraph of A_y .

Remark 4. The work in [4] focuses on the $c=2$ case and presents an approach to cancel all cycles of length l , $g \leq l \leq l_{max}$, where g is the girth, for this column weight choice. As stated in [4], it is impossible to cancel all cycles for all lengths l . In contrast, for codes with column weights $c \geq 2$, our approach only seeks to cancel a selected number of cycles, i.e., one fundamental cycle per AS of interest. As a result, our approach allows for further flexibility in canceling cycles of various lengths for column weights $c \geq 2$.

Figure 4 shows the simulation results for random regular codes (denoted by ‘Random’) over $GF(q)$, $q = 2, 4, 8, 16$

Algorithm 1 Reduction in the number of ASs in the Tanner graph of a non-binary code

- 1: **Input:** Tanner graph T with edge weights over $\text{GF}(q)$.
- 2: Choose W , the set of NB-AS to be eliminated.
- 3: Let X be the set of NB-AS which can not be eliminated.
- 4: Let $X = \emptyset$.
- 5: Let set $A = \emptyset$.
- 6: For every edge $j \in T$, C_j is the set of canceled cycles which include j .
- 7: $\forall j \in T$, $C_j = \emptyset$.
- 8: Find (a, b) , the smallest NB-AS in $W \setminus A$.
- 9: If this AS is a child of another AS in A , go to 22.
- 10: Find U , the set of all (a, b) ASs in unlabeled Tanner graph T (e.g., using techniques in [12]).
- 11: **for** $\forall u \in U$ **do**
- 12: Find F_u , a set of fundamental cycles of u .
- 13: Let E_u be the set of all edges in u .
- 14: For an edge $k \in E_u$, let M_k be the set of cycles in F_u which include k .
- 15: **if** (5) is satisfied for all cycles in F_u **then**
- 16: If $E_u = \emptyset$, $X \leftarrow X \cup U$ and go to 22.
- 17: Find edge $i \in E_u$ with minimum $|C_i|$.
- 18: $w_i \leftarrow w_i + v$ with $v \neq w_i$ and $v \neq 0$ such that all cycles in C_i do not satisfy (5). If no v exists, $E_u \leftarrow E_u \setminus i$, and go to 17.
- 19: For every edge e in the cycles of M_i , update $C_e \leftarrow C_e \cup M_e$.
- 20: **end if**
- 21: **end for**
- 22: Add (a, b) AS to the set A .
- 23: If $A \neq W$, go to 8.
- 24: If $X = \emptyset$, all ASs of interest are eliminated. Otherwise, it is not possible to eliminate ASs in X .

TABLE I
ERROR PROFILE, $E_b/N_0 = 5.3$ dB, AND $\text{GF}(4)$.

Error Type	(4, 4)	(5, 0)	(5, 2)	(6, 2)	(6, 4)	(6, 6)	(7, 4)	(8, 2)	other
Random	85	16	24	12	16	32	16	36	45
P-method	0	0	0	0	0	15	5	4	19
A-method	0	0	0	0	0	0	0	0	15

with block length $N \approx 2800$ bits, rate $R \approx 0.87$, column weight $c = 4$ and girth $g = 6$. The figure shows the results obtained by using the approach as in [4] (denoted by P-method), which tries to cancel all cycles of length l , $g \leq l \leq l_{max}$. The figure also shows the results obtained by our code modification specified in Algorithm 1 (denoted by A-method). All codes are decoded using Fast Fourier Transform-based q -ary SPA (FFT-QSPA) decoder [13]. Both P-method and A-method approaches enjoy the improved performance relative to random code construction. Performance comparison for different values of q reveals that the improvement is more pronounced for smaller values of q . As shown in Corollary 1, for random code construction, under higher field orders, there are fewer elementary NB-AS to begin with. As a result there are fewer NB-ASs to cancel using either P-method or A-method.

Table I includes the error profile for all three codes over $\text{GF}(4)$. Both P-method and A-method eliminate all (4, 4) ASs

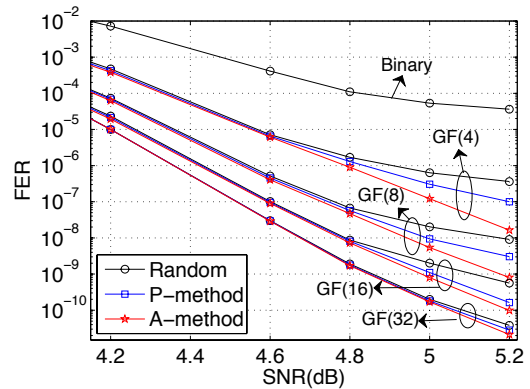


Fig. 4. Performance comparison of the original random non-binary code with P-method modified and A-method modified codes.

as well as (5, 0), (5, 2), (6, 2) and (6, 4) ASs that are children of (4, 4) ASs. Additionally, A-method successfully eliminates all (6, 6), (7, 4) and (8, 2) ASs since in this approach, we selectively cancel some (but not necessary all) length 6 cycles followed by cancellation of some of length 8 cycles. As q increases, there are more degrees of freedom to change the edge weights to cancel cycles and as a result, the cycle-only approach appears to be sufficient for larger values of q .

V. CONCLUSION

We presented a generalized absorbing set definition for non-binary graph based codes over $\text{GF}(q)$. We showed that as q gets larger, it is harder to satisfy edge labeling conditions of absorbing sets and as a result the number of absorbing sets decreases for larger field sizes. Further, we also proposed an algorithm to decrease the number of absorbing sets in the Tanner graph by changing carefully chosen edge weights.

REFERENCES

- [1] J. Kang *et al.*, "Quasi-cyclic LDPC codes: an algebraic construction," *IEEE Trans. on Commun.*, vol. 58, no. 5, pp. 1383-1396, 2010.
- [2] J. Huang *et al.*, "Large-girth nonbinary QC-codes of various lengths," *IEEE Trans. on Commun.*, vol. 58, no. 12, pp. 3436-3447, 2010.
- [3] T. Nozaki *et al.*, "Analysis of stopping constellation distribution for irregular non-binary LDPC code ensemble" in *Proc. IEEE ISIT*, 2011.
- [4] C. Poulliat, M. Fossorier, and D. Declercq, "Design of regular $(2, d_c)$ -LDPC codes over $\text{GF}(q)$ using their binary images" *IEEE Trans. on Commun.*, vol. 56, no. 10, pp. 1626-1635, 2008.
- [5] T. Nozaki, K. Kasai, and K. Sakaniwa, "Analysis of error floors of generalized non-binary LDPC codes over q -ary memoryless symmetric channels" in *Proc. IEEE ISIT*, 2012.
- [6] D. Divsalar and L. Dolecek, "Enumerators for protograph-based ensembles of nonbinary LDPC codes," in *Proc. IEEE ISIT*, 2011.
- [7] A. Poloni, S. Valle, and S. Vincenti, "NB-LDPC: absorbing set and importance sampling," in *Proc. IEEE ISTC*, 2012.
- [8] L. Dolecek *et al.*, "Analysis of absorbing sets and fully absorbing sets of array-based LDPC codes," *IEEE Trans. on Inf. Theory*, vol. 56, no. 1, pp. 181-201, 2010.
- [9] O. Milenkovic, E. Soljanin, and P. Whiting, "Asymptotic spectra of trapping sets in regular and irregular LDPC code ensembles," *IEEE Trans. on Inf. Theory*, vol. 53, no. 1, pp. 39-55, 2007.
- [10] B. Amiri, C. W. Lin, and L. Dolecek, "Asymptotic distribution of absorbing sets and fully absorbing sets for regular sparse code ensembles," *IEEE Trans. on Commun.*, to appear, 2013.
- [11] J. Wang, L. Dolecek and R. Wesel, "The cycle consistency matrix approach to LDPC absorbing sets in separable circulant-based codes," *IEEE Trans. on Inf. Theory*, vol. 59, no. 4, pp. 2293-2314, 2013.
- [12] M. Karimi and A. H. Banihashemi, "An efficient algorithm for finding dominant trapping sets of LDPC codes," *IEEE Trans. on Inf. Theory*, vol. 58, pp. 6942-6958, Nov. 2012.
- [13] D. Declercq and M. Fossorier, "Decoding algorithms for nonbinary LDPC codes over $\text{GF}(q)$," *IEEE Trans. on Commun.*, vol. 55, no. 4, pp. 633-643, 2007.