PAPER Special Section on Information Theory and Its Applications

Analysis of Stopping Constellation Distribution for Irregular Non-binary LDPC Code Ensemble*

Takayuki NOZAKI[†a)], ***Student Member***, **Kenta KASAI**[†b)], ***Member***, ***and*** **Kohichi SAKANIWA**[†c)], ***Fellow***

SUMMARY The fixed points of the belief propagation decoder for non-binary low-density parity-check (LDPC) codes are referred to as stopping constellations. In this paper, we give the stopping constellation distributions for the irregular non-binary LDPC code ensembles defined over the general linear group. Moreover, we derive the exponential growth rate of the average stopping constellation distributions in the limit of large code-length.

key words: *non-binary LDPC codes, stopping constellation*

1. Introduction

Gallager invented low-density parity-check (LDPC) codes [1]. Due to the sparseness of the parity check matrices, LDPC codes are efficiently decoded by the belief propagation (BP) decoder. Optimized LDPC codes can exhibit performance very close to the Shannon limit [2]. Davey and MacKay [3] found that non-binary LDPC codes can outperform binary ones.

An LDPC code is defined by a sparse parity check matrix. In this paper, we consider the non-binary LDPC codes defined over the general linear group. The general linear group over the binary field of dimension m is the set of all $m \times m$ invertible binary matrices. Each entry of parity check matrix for non-binary LDPC code defined over the general linear group is an element of the general linear group. It is known that LDPC codes defined over general linear groups outperform LDPC codes defined over finite fields in terms of the decoding performance [4].

The block and the bit erasure probabilities for binary LDPC codes over the binary erasure channels (BEC) are determined by the size of the maximal stopping set [5]. The fixed points of the BP decoder for non-binary LDPC codes are referred to as *stopping constellations* [6]. The stopping constellations for the non-binary LDPC codes correspond to the stopping sets for the binary LDPC codes. To analyze the decoding erasure probabilities of the non-binary LDPC codes over the BEC by the BP decoder, we need to analyze the stopping constellation. In this paper, as the first step, we derive the stopping constellation distribution. In this paper, we also give the exponential growth rates of the average stopping constellation distributions in the limit of large code length.

The remainder of this paper is organized as follows. In Sect. 2, we define the irregular non-binary LDPC code ensemble and the stopping constellation. In Sect. 3, we show the relationship between the stopping constellation and the stopping set. In Sect. 4, we derive the stopping constellation distributions for irregular non-binary LDPC code ensembles. In Sect. 5, we derive the exponential growth rates of the average stopping constellation distributions in the limit of large code length. In Sect. 6, we show the numerical examples for the exponential growth rates of the average stopping constellation distributions.

2. Preliminaries

In this section, we define the irregular non-binary LDPC code ensemble and the stopping constellation. We introduce some notations used throughout this paper.

2.1 Non-binary LDPC Code Ensemble

Let $\mathbb{F}_2$ be the Galois field of order 2. We denote the general linear group over $\mathbb{F}_2$ of dimension m, by $\mathrm{GL}(m, \mathbb{F}_2)$. In this paper, we consider non-binary LDPC codes defined over $\mathrm{GL}(m, \mathbb{F}_2)$. A non-binary LDPC code is defined by an $M \times N$ parity check matrix. Each entry of the parity check matrices for non-binary LDPC codes over $\mathrm{GL}(m, \mathbb{F}_2)$ is an element in $\mathrm{GL}(m, \mathbb{F}_2)$. Denote the (i, j)-th entry of parity check matrix by $H_{i,j}$. Let $[a, b]$ be the set of consecutive integers from a to b, i.e., $[a, b] = \{a, a+1, \ldots, b\}$. For $i \in [1, N]$, the i-th symbol $\boldsymbol{x}_i$ is represented by an m-tuple defined over $\mathbb{F}_2$. The code is defined as follows:

$$\left\{(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) \in (\mathbb{F}_2^m)^N \mid \sum_{j=1}^{N} H_{i,j}\boldsymbol{x}_j^T = \boldsymbol{0}^T \quad \forall i \in [1, M]\right\}.$$

The Tanner graph for a non-binary LDPC code is represented by a bipartite graph with variable nodes, check nodes and labeled edges. The parity-check matrices are represented by Tanner graphs. Let $\mathcal{L}$ and $\mathcal{R}$ be the sets of degrees of the variable nodes and the check nodes, respectively. Irregular non-binary LDPC codes are characterized with the number of variable nodes N, the dimension m of the general linear group and a pair of *degree distribution*,

Manuscript received February 18, 2011.
Manuscript revised June 10, 2011.

[†]The authors are with the Department of Communications and Integrated Systems, Tokyo Institute of Technology, Tokyo, 152-8550 Japan.

*The material in this paper was presented in part at 33rd symposium on information theory and its application (SITA2010).

a) E-mail: nozaki@comm.ss.titech.ac.jp
b) E-mail: kenta@comm.ss.titech.ac.jp
c) E-mail: sakaniwa@comm.ss.titech.ac.jp

DOI: 10.1587/transfun.E94.A.2153

$\lambda(x) = \sum_{i\in\mathcal{L}} \lambda_i x^{i-1}$ and $\rho(x) = \sum_{i\in\mathcal{R}} \rho_i x^{i-1}$, where λ_i and ρ_i are the fractions of the edges connected to the variable nodes and the check nodes of degree i, respectively.

The total number of edges in the Tanner graph is

$$\xi := N/\int_0^1 \lambda(x)\mathrm{d}x.$$

Let L_i and R_j be the fraction of the variable nodes of degree i and the check nodes of degree j, respectively, i.e.,

$$L_i := \lambda_i/\Big(i\int_0^1 \lambda(x)\mathrm{d}x\Big), \qquad R_j := \rho_j/\Big(j\int_0^1 \rho(x)\mathrm{d}x\Big).$$

Define the design rate r as follows:

$$r := 1-\Big(\int_0^1 \rho(x)\mathrm{d}x\Big)/\Big(\int_0^1 \lambda(x)\mathrm{d}x\Big).$$

Assume that we are given the number of variable nodes N, the dimension m and the degree distribution pair (λ, ρ). A non-binary LDPC code ensemble $\mathrm{EGL}(N, m, \lambda, \rho)$ is defined in the following way. There exist L_iN variable nodes of degree i and $R_jN(1-r)$ check nodes of degree j. A node of degree i has i sockets for its connected edges. Consider a permutation π on the number of edges. Join the i-th socket on the variable node side to the $\pi(i)$-th socket on the check node side. The bipartite graphs are chosen with equal probability from all the permutations on the number of edges. Each label in an edge is chosen as an element from $\mathrm{GL}(m, \mathbb{F}_2)$ with equal probability.

2.2 Stopping Constellation

Consider the transmission over the BEC. For the BEC, we are able to assume that all-zero codewords were sent without loss of generality to analyze the decoding erasure probability [7]. The indices corresponding to the nonzero entries of a message arising in the BP decoder forms a linear subspace in $\mathbb{F}_2^m$. Thus, each message in the BP decoder is represented by a linear subspace in $\mathbb{F}_2^m$ [7].

To define the fixed points of the BP decoder for the non-binary LDPC codes, which are referred to as *stopping constellations* [6], we review the *states* in the peeling decoder for non-binary LDPC codes [6].

The peeling decoder for the non-binary LDPC codes assigns a set of candidate symbols for the decoding result to each variable node. Such a set is referred to as *state* of the variable node v and denoted by E_v, where $E_\mathsf{v} \subseteq \mathbb{F}_2^m$. It is known that the states of the variable nodes are also represented by linear subspaces [6]. For any k linear subspaces $\{V_i\}_{i=1}^k$ in $\mathbb{F}_2^m$, we denote $\sum_{i=1}^k V_i := \Big\{\sum_{j=1}^k v_j \mid v_j \in V_j \text{ for } j \in [1,k]\Big\}$. For a linear subspace V in $\mathbb{F}_2^m$ and an $m \times m$ invertible binary matrix H, we denote $HV := \{Hv \mid v \in V\}$.

Definition 1: [6] Let $\mathcal{V}$ be the set of variable nodes. With some abuse of notation, we identify $\mathcal{V}$ and $[1, N]$ hereafter. We denote the set of variable nodes connecting to the check node c, by $\mathcal{N}(\mathsf{c})$. A *stopping constellation* $\{E_\mathsf{v}\}_{\mathsf{v}\in\mathcal{V}}$ is defined as an assignment of the states such that

$$E_\mathsf{v} \subseteq H_{\mathsf{c},\mathsf{v}}^{-1}\Big(\sum_{i\in\mathcal{N}(\mathsf{c})\setminus\{\mathsf{v}\}} H_{\mathsf{c},i}E_i\Big), \tag{1}$$

for any $\mathsf{v} \in \mathcal{V}$ and all the check nodes c connecting to v.

It is known that stopping constellations are fixed points of the peeling decoder and the BP decoder [6]. In this paper, for a given stopping constellation we refer to the number of states whose dimensions are not equal to 0 as the *weight* of the stopping constellation.

3. Stopping Constellation and Stopping Set

In this section, we show the relationship between the stopping constellation and the stopping set. A stopping set $\mathcal{S}$ is a set of variable nodes such that all the neighbors of $\mathcal{S}$ are connected to $\mathcal{S}$ at least twice. For a given stopping constellation $\{E_\mathsf{v}\}_{\mathsf{v}\in\mathcal{V}}$, let $\tilde{\mathcal{S}}$ be the set of variable nodes such that the dimensions of the corresponding states are not 0, i.e.,

$$\tilde{\mathcal{S}} := \{\mathsf{v} \in \mathcal{V} \mid E_v \neq \{\mathbf{0}\}\}.$$

Lemma 1: For a fixed $\mathsf{G} \in \mathrm{EGL}(N, m, \lambda, \rho)$ and a given stopping constellation, the set of variable nodes $\tilde{\mathcal{S}}$ for the stopping constellation forms a stopping set.

proof: If there exists a check node c which connects to $\tilde{\mathcal{S}}$ once, then for the variable node $\mathsf{v} \in \tilde{\mathcal{S}}$ such that a neighbor of v is c, $E_\mathsf{v} \neq \{\mathbf{0}\}$ and $H_{\mathsf{c},\mathsf{v}}^{-1}\sum_{i\in\mathcal{N}(\mathsf{c})\setminus\{\mathsf{v}\}} H_{\mathsf{c},i}E_i = \{\mathbf{0}\}$. Hence, the assignment of the states $\{E_\mathsf{v}\}_{\mathsf{v}\in\mathcal{V}}$ is not a stopping constellation if there exists a check node which connects to $\tilde{\mathcal{S}}$ once. Thus, all the neighbors of $\tilde{\mathcal{S}}$ are connected to $\tilde{\mathcal{S}}$ at least twice. Therefore, the set of variable nodes $\tilde{\mathcal{S}}$ for the stopping constellation forms a stopping set. □

Lemma 2: For a fixed $\mathsf{G} \in \mathrm{EGL}(N, m, \lambda, \rho)$ and a given stopping set $\mathcal{S}$, there exist at least one stopping constellation with the set of variable nodes $\tilde{\mathcal{S}}$ such that $\tilde{\mathcal{S}} = \mathcal{S}$.

proof: For a given stopping set $\mathcal{S}$, the assignment of state $\{E_\mathsf{v}\}_{\mathsf{v}\in\mathcal{V}}$ such that $E_\mathsf{v} = \mathbb{F}_2^m$ for all $\mathsf{v} \in \mathcal{S}$ and $E_\mathsf{v} = \{\mathbf{0}\}$ for all $\mathsf{v} \in \mathcal{V} \setminus \mathcal{S}$ is a stopping constellation. □

The stopping constellations of small weight degrade the decoding erasure rates of non-binary LDPC codes. From those lemmas, we see that in order to get a code which does not contain the stopping constellation of small weight, we need to eliminate the stopping sets of small weight.

4. Stopping Constellation Distribution for Non-binary LDPC Codes

In this section, we derive the stopping constellation distributions for irregular non-binary LDPC code ensembles. We give some lemmas to count constellations of the linear subspaces satisfying the stopping constellation constraints Eq. (1) for check nodes.

4.1 Number of Linear Subspaces

It is known that the number of distinct subspaces of dimension k of the vector space $\mathbb{F}_2^m$ is given by the Gaussian binomial coefficient [8, p. 443]. Define for a non-negative

integer m

$$[m] := \begin{cases} 1 & \text{if } m = 0, \\ \prod_{i=1}^{m}(2^i - 1) & \text{if } m \geq 1. \end{cases}$$

The Gaussian binomial coefficient $\begin{bmatrix} m \\ k \end{bmatrix}$ is given by $\begin{bmatrix} m \\ k \end{bmatrix} = \frac{[m]}{[m-k][k]}$. We denote the dimension of V_i, by $\dim V_i$. The following lemma gives the number of the sets of linear subspaces $\{V_i\}_{i=1}^k$ for a given condition for the dimension of $\{V_i\}_{i=1}^k$.

Lemma 3: Assume that two non-negative integers k, m are given. For a given set of non-negative integers $\boldsymbol{a}_k = \{a_k(S)\}_{S\subseteq[1,k]}$ such that $\sum_{S\subseteq[1,k]} a_k(S) = m$, let $B_k(\boldsymbol{a}_k)$ be the number of the sets of linear subspaces $\{V_i\}_{i=1}^k$ in $\mathbb{F}_2^m$ which satisfy

$$\dim\left(\bigcap_{i\in S} V_i\right) = \sum_{\tilde{S}\subseteq[1,k]:S\subseteq\tilde{S}} a_k(\tilde{S}), \tag{2}$$

where $\sum_{\tilde{S}\subseteq[1,k]:S\subseteq\tilde{S}} a_k(\tilde{S})$ is the sum of $a_k(\tilde{S})$ over all $\tilde{S} \subseteq [1,k]$ such that $S \subseteq \tilde{S}$. Then, we have

$$B_k(\boldsymbol{a}_k) = \frac{[m]}{\prod_{S\subseteq[1,k]}[a_k(S)]} 2^{T_k}, \tag{3}$$

where

$$T_k := \frac{1}{2} \sum_{\substack{S_1,S_2\subseteq[1,k]: \\ S_1\not\subseteq S_2, S_1\not\supseteq S_2}} a_k(S_1)a_k(S_2).$$

The proof of this lemma is in Appendix A.

Lemma 4: Assume that two non-negative integers k, m are given. Define $B_k(\boldsymbol{a}_k)$ as in Eq. (3). For a given set of non-negative integers $\boldsymbol{v} = \{v_i\}_{i=1}^k$ where $v_i \in [0, m]$ for all $i \in [1, k]$, we denote the number of the sets of linear subspaces $\{V_i\}_{i=1}^k$ such that $V_i \supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$ and $\dim V_i = v_i$ for all $i \in [1, k]$, by $\tilde{h}_k(\boldsymbol{v})$, i.e.,

$$\tilde{h}_k(\boldsymbol{v}) := \#\Big\{\{V_i\}_{i=1}^k \mid V_i \supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j \ \forall i \in [1,k], \\ \dim V_i = v_i \ \forall i \in [1,k]\Big\},$$

where $\#A$ is the cardinality of a set A. Then,

$$\tilde{h}_k(\boldsymbol{v}) = \textstyle\sum_{\boldsymbol{a}_k\in D'_k} B_k(\boldsymbol{a}_k),$$

where

$$D'_k := \Big\{\boldsymbol{a}_k \mid \textstyle\sum_{S:i\in S} a_k(S) = \dim V_i \ \ \forall i \in [1,k], \\ \textstyle\sum_{S\subseteq[1,k]} a_k(S) = m, a_k([1,k]\setminus\{i\}) = 0 \ \ \forall i \in [1,k]\Big\}.$$

The proof is in Appendix B.

Discussion 1: Assume that two non-negative integers k, m are given. For a given $\boldsymbol{d} = (d_0, \ldots, d_m)$ such that $\sum_{i=0}^m d_i = k$ and $d_i \geq 0$ for $i \in [0, m]$, we denote the number of the sets of linear subspaces $\{V_i\}_{i=1}^k$ such that $V_i \supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$ for all $i \in [1, k]$ and the number of subspaces in $\{V_j\}_{j=1}^k$ with dimension $m - i$ is d_i, by $h_k(\boldsymbol{d})$, i.e.,

$$h_k(\boldsymbol{d}) := \#\Big\{\{V_j\}_{j=1}^k \mid V_i \supseteq \textstyle\bigcap_{j\in[1,k]\setminus\{i\}} V_j \ \forall i \in [1,k], \\ \#\{i \mid \dim V_i = j\} = d_j \ \forall j \in [1,m]\Big\}.$$

There are $\binom{k}{d_0,d_1,\ldots,d_m}$ choices to the dimensions of $\{V_i\}_{i=1}^k$, where

$$\binom{k}{d_0, d_1, \ldots, d_m} := \frac{k!}{\prod_{i=0}^m d_i!}, \ \ k = \sum_{i=0}^m d_i,$$

is known as the multinomial coefficient [9]. From Lemma 4, we have $\tilde{h}_k(\{v_i\}_{i=1}^k) = \tilde{h}_k(\{v_{\pi(i)}\}_{i=1}^k)$ for any permutation π on k and $\{v_i\}_{i=1}^k$. For $j \in [1, m]$, let p_j be the smallest integer such that $j \leq \sum_{i=0}^{p_j} d_i$. Hence, we get

$$h_k(\boldsymbol{d}) = \binom{k}{d_0, d_1, \ldots, d_m} \tilde{h}_k(\{v_i\}_{i=1}^k),$$

where $v_i = m - p_i$ for all $i \in [i, k]$. Thus, we obtain

$$h_k(\boldsymbol{d}) = \binom{k}{d_0, d_1, \ldots, d_m} \sum_{\boldsymbol{a}_k\in D_k} B_k(\boldsymbol{a}_k),$$

where

$$D_k := \Big\{\boldsymbol{a}_k \mid \textstyle\sum_{S:i\in S} a_k(S) = m - p_i \ \ \forall i \in [1,k], \\ \textstyle\sum_{S\subseteq[1,k]} a_k(S) = m, a_k([1,k]\setminus\{i\}) = 0 \ \ \forall i \in [1,k]\}.$$

We denote $\boldsymbol{d} \geq \boldsymbol{0}$ if $d_i \geq 0$ for all $i \in [0, m]$. The generator function of $h_k(\boldsymbol{d})$ is written as follows:

$$f_k(\boldsymbol{u}) := \sum_{\boldsymbol{d}\geq\boldsymbol{0}:\sum_{i=0}^m d_i=k} h_k(\boldsymbol{d}) \prod_{i=1}^m u_i^{d_i}. \tag{4}$$

Since d_0 depends on $d_1, \cdots, d_n$, i.e., $d_0 = k - \sum_{i=1}^m d_i$, we drop u_0 from Eq. (4).

4.2 Stopping Constellation Distributions for Non-binary LDPC Codes

Recall that for a given stopping constellation we refer to the number of the states whose dimensions are not equal to 0 as the weight of the stopping constellation. For a given Tanner graph $\mathsf{G} \in \mathrm{EGL}(N, m, \lambda, \rho)$, we denote the number of stopping constellations of weight ℓ in G by $A^{\mathsf{G}}(\ell)$. For the ensemble $\mathrm{EGL}(N, m, \lambda, \rho)$, let $A(\ell)$ be the average stopping constellations of weight ℓ. Since each code is chosen with equal probability from $\mathrm{EGL}(N, m, \lambda, \rho)$, we get

$$A(\ell) = \sum_{\mathsf{G}\in\mathrm{EGL}(N,m,\lambda,\rho)} \frac{A^{\mathsf{G}}(\ell)}{|\mathrm{EGL}(N, m, \lambda, \rho)|}.$$

The following theorem gives the average stopping constellations for irregular non-binary LDPC code ensembles.

Theorem 1: Define $f_k(\boldsymbol{u})$ as in Eq. (4). The average stopping constellations $A(\ell)$ of weight ℓ for the non-binary LDPC code ensemble EGL(N, m, λ, ρ) is given by

$$A(\ell) = \sum_{\boldsymbol{b}\geq\mathbf{0}:\sum_{i=0}^{m} b_i=\xi} \frac{\mathrm{coef}((Q(\boldsymbol{s},t)P(\boldsymbol{u}))^N, t^\ell \prod_{i=1}^m s_i^{b_i} u_i^{b_i})}{\binom{\xi}{b_0,b_1,\ldots,b_m} \prod_{k=1}^m \begin{bmatrix} m \\ k \end{bmatrix}^{b_k}}, \tag{5}$$

$$Q(\boldsymbol{s},t) := \prod_{j\in\mathcal{L}} \Big\{1 + t \sum_{i=1}^m \begin{bmatrix} m \\ i \end{bmatrix} s_i^j\Big\}^{L_j},$$
$$P(\boldsymbol{u}) := \prod_{k\in\mathcal{R}} \{f_k(\boldsymbol{u})\}^{R_k(1-r)},$$

where $\mathrm{coef}(g(\boldsymbol{s},t,\boldsymbol{u}), t^\ell \prod_{i=1}^m s_i^{b_i} u_i^{b_i})$ is the coefficient of the term $t^\ell \prod_{i=1}^m s_i^{b_i} u_i^{b_i}$ in the polynomial $g(\boldsymbol{s},t,\boldsymbol{u})$.

proof: First, we count constellations of the linear subspaces satisfying the stopping constellation constraint Eq. (1) for all check nodes. Consider a check node c of degree k. We say that the check node c satisfies the decoding failure criterion with respect to the state assignment $\{E_\mathsf{v}\}_{\mathsf{v}\in\mathcal{V}}$ if

$$E_\mathsf{v} \subseteq H_{\mathsf{c},\mathsf{v}}^{-1}\Big(\sum_{i\in\mathcal{N}(\mathsf{c})\setminus\{\mathsf{v}\}} H_{\mathsf{c},i} E_i\Big), \quad \forall \mathsf{v} \in \mathcal{N}(\mathsf{c}).$$

Substituting $\tilde{E}_i = H_{\mathsf{c},i} E_i$ to this, we have

$$\tilde{E}_\mathsf{v} \subseteq \Big(\sum_{i\in\mathcal{N}(\mathsf{c})\setminus\{\mathsf{v}\}} \tilde{E}_i\Big), \quad \forall \mathsf{v} \in \mathcal{N}(\mathsf{c}).$$

For a linear subspace V, denote its dual subspace by $V^\perp$, i.e., $V^\perp := \{\beta \mid \langle\alpha,\beta\rangle = 0\ \forall\alpha \in V\}$, where $\langle\alpha,\beta\rangle$ denotes the inner product of α and β. Using the dual subspaces, we have

$$\tilde{E}_\mathsf{v}^\perp \supseteq \Big(\bigcap_{i\in\mathcal{N}(\mathsf{c})\setminus\{\mathsf{v}\}} \tilde{E}_i^\perp\Big), \quad \forall \mathsf{v} \in \mathcal{N}(\mathsf{c}).$$

We refer to the edges adjacent to the variable node assigned to state of dimension i as the edges of dimension i. Let d_i be the number of edges of dimension i which are adjacent to the check node c. From Discussion 1, for a given $(d_0,\ldots,d_m)$ such that $\sum_{i=0}^m d_i = k$ and $d_i \geq 0$ for all $i \in [0,m]$, the number of the constellations that satisfy the decoding failure criterion for the check node c is written as $\mathrm{coef}(f_k(\boldsymbol{u}), \prod_{i=1}^m u_i^{d_i})$. Let b_i be the total number of edges of dimension i. Since there are $R_k(1-r)N$ check nodes of degree k, the number of the constellations that satisfy the stopping constellation constraints for the $N(1-r)$ check nodes for a given $\boldsymbol{b} = (b_0,\ldots,b_m)$ such that $\sum_{i=0}^m b_i = \xi$ and $b_i \geq 0$ for all $i \in [0,m]$, is

$$\mathrm{coef}\Big(\prod_{k\in\mathcal{R}} (f_k(\boldsymbol{u}))^{R_k(1-r)N}, \prod_{i=1}^m u_i^{b_i}\Big). \tag{6}$$

Secondly, we count constellations of linear subspaces satisfying the constraints of the variable nodes. Consider a variable node v of degree k. If the variable node v is assigned to state of dimension i, the k edges adjacent to v are of dimension i. Define the parameter w as 1 if the dimension of the state of the variable node v is not 0, and otherwise 0. Denote the number of edges of dimension i adjacent to the variable node v, by d_i. For a given $w \in \{0,1\}$ and $\boldsymbol{d} = (d_0,\ldots,d_m)$ such that $\sum_{i=0}^m d_i = k$ and $d_i \geq 0$ for all $i \in [0,m]$, let $g_k(w,\boldsymbol{d})$ be the number of constellations of linear subspaces satisfying a constraint of variable node of degree k. Since the number of states of dimension i is $\begin{bmatrix} m \\ i \end{bmatrix}$, we have

$$g_k(w,\boldsymbol{d}) = \begin{cases} 1 & w = 0, d_0 = k, d_j = 0\ \forall j \in [1,m], \\ \begin{bmatrix} m \\ i \end{bmatrix} & w = 1, d_i = k, d_j = 0\ \forall j \in [0,m]\setminus\{i\}, \\ 0 & \text{otherwise.} \end{cases}$$

The generator function of $g_k(w,\boldsymbol{d})$ is written as follows:

$$\sum_{w,\boldsymbol{d}} g_k(w,\boldsymbol{d}) t^w \prod_{i=1}^m s_i^{d_i} = 1 + t\sum_{i=1}^m \begin{bmatrix} m \\ i \end{bmatrix} s_i^k.$$

Since there are $L_k N$ variable nodes of degree k, for a given ℓ and $\boldsymbol{b}$ such that $\sum_{i=0}^m b_i = \xi$ and $b_i \geq 0$ for all $i \in [0,m]$, the number of constellations of linear subspaces satisfying constraints of the N variable nodes is given by

$$\mathrm{coef}\Big(\prod_{k\in\mathcal{L}} \Big(1 + t\sum_{i=1}^m \begin{bmatrix} m \\ i \end{bmatrix} s_i^k\Big)^{L_k N}, t^\ell \prod_{i=1}^m s_i^{b_i}\Big). \tag{7}$$

Thirdly, we count the edge permutation and the edge labels which satisfy the constellation. For a given $\boldsymbol{b}$ such that $\sum_{i=0}^m b_i = \xi$ and $b_i \geq 0$ for all $i \in [0,m]$, the number of permutations of edges is given by $\prod_{i=0}^m b_i!$ and the number of edge labels is equal to $\prod_{i=0}^m ([m-i][i])^{b_i}$. Hence, for a given $\boldsymbol{b}$ such that $\sum_{i=0}^m b_i = \xi$ and $b_i \geq 0$ for all $i \in [1,m]$, the number of choices for the permutations of edges and edge labels is

$$\prod_{i=0}^m b_i!([m-i][i])^{b_i}. \tag{8}$$

Finally, the number of Tanner graphs in EGL(N,m,λ,ρ) is given by $\xi![m]^\xi$. From Eqs. (6), (7) and (8) and the number of Tanner graphs, the average stopping constellation distribution for a given ℓ and $\boldsymbol{b}$ such that $\sum_{i=0}^m b_i = \xi$ and $b_i \geq 0$ for all $i \in [0,m]$, is given by

$$A(\ell,\boldsymbol{b}) = \frac{\mathrm{coef}((Q(\boldsymbol{s},t)P(\boldsymbol{u}))^N, t^\ell \prod_{i=1}^m s_i^{b_i} u_i^{b_i})}{\binom{\xi}{b_0,b_1,\ldots,b_m} \prod_{k=1}^m \begin{bmatrix} m \\ k \end{bmatrix}^{b_k}}.$$

Since $A(\ell) = \sum_{\boldsymbol{b}\geq\mathbf{0}:\sum_{i=0}^m b_i=\xi} A(\ell,\boldsymbol{b})$, we get Theorem 1. □

5. Asymptotic Analysis

In this section, we investigate the asymptotic behavior of the average stopping constellation distributions of non-binary LDPC code ensembles in the limit of large code length. Define the *normalized* weight ω by $\omega := \ell/N$. We define

$$\gamma_m(\omega) := \lim_{N\to\infty} \frac{1}{N} \log_2 A(\omega N),$$

and refer to this as the *exponential growth rate* or simply *growth rate* of the average stopping constellation distribution. To simplify the notation, we denote logarithms to base 2 as log.

With the growth rate, we can roughly estimate the number of stopping constellations by

$$A(\omega N) \sim 2^{\gamma_m(\omega) N},$$

where $a_N \sim b_N$ means that $\lim_{N\to\infty} \frac{1}{N} \log \frac{a_N}{b_N} = 0$.

The number of the terms in Eq. (5) is at most $(\xi+1)^m$. Hence, from Eq. (5) we have

$$\max_{\boldsymbol{b}\geq\boldsymbol{0}:\sum_{i=0}^m b_i=\xi} A(\ell,\boldsymbol{b}) \leq A(\ell) \leq (\xi+1)^m \max_{\boldsymbol{b}\geq\boldsymbol{0}:\sum_{i=0}^m b_i=\xi} A(\ell,\boldsymbol{b}).$$

Therefore, we get

$$\lim_{N\to\infty} \frac{1}{N} \log A(\ell) = \lim_{N\to\infty} \frac{1}{N} \log \max_{\boldsymbol{b}\geq\boldsymbol{0}:\sum_{i=0}^m b_i=\xi} A(\ell,\boldsymbol{b}).$$

To calculate this equation, we introduce the following lemma.

Lemma 5: [10, Theorem 2] Let $\gamma > 0$ be some rational number and let $p(x_1, x_2, \ldots, x_m)$ be a function such that $p(x_1, x_2, \ldots, x_m)^\gamma$ is a multivariate polynomial with non-negative coefficients. Let $\alpha_k > 0$ be some rational numbers for $k \in [1, m]$ and let n_i be the series of all indices j such that j/γ is an integer and $\mathrm{coef}(p(x_1, \ldots, x_m)^j, x_1^{\alpha_1 j} \cdots x_m^{\alpha_m j}) \neq 0$. Then

$$\lim_{i\to\infty} \frac{1}{n_i} \log \mathrm{coef}(p(x_1,\ldots,x_m)^{n_i}, (x_1^{\alpha_1}\cdots x_m^{\alpha_m})^{n_i}) = \inf_{x_1,\ldots,x_m>0} \log \frac{p(x_1,\ldots,x_m)}{x_1^{\alpha_1}\cdots x_m^{\alpha_m}}.$$

A point $(x_1, \ldots, x_m)$ achieves the minimum of the function $p(x_1, \ldots, x_m)/(x_1^{\alpha_1} \ldots x_m^{\alpha_m})$, if and only if it satisfies the following equation for all $k \in [1, m]$:

$$x_k \frac{\partial p(x_1,\ldots,x_m)^\gamma}{\partial x_k} - \gamma\alpha_k p(x_1,\ldots,x_m)^\gamma = 0.$$

Define $\beta_i := b_i/N$ for $i \in [0, m]$ and $\epsilon := \xi/N$. We denote $\boldsymbol{s} > \boldsymbol{0}$ if $s_i > 0$ for all $i \in [1, m]$. From Theorem 1 and Lemma 5, we obtain the following theorem.

Theorem 2: The growth rate $\gamma_m(\omega)$ of the average stopping constellation distributions for the irregular non-binary LDPC code ensemble EGL(N, m, λ, ρ) is given by

$$\begin{aligned}\gamma_m(\omega) =& \sup_{\substack{\boldsymbol{\beta}>\boldsymbol{0}:\\ \sum_{i=0}^m \beta_i=\epsilon}} \inf_{\substack{\boldsymbol{s}>\boldsymbol{0},t>0,\\ \boldsymbol{u}>\boldsymbol{0}}} \Bigg\{\log Q(\boldsymbol{s},t) - \omega\log t + \log P(\boldsymbol{u}) \\ &- \sum_{i=1}^m \beta_i \log\begin{bmatrix} m \\ i \end{bmatrix} s_i u_i + \sum_{i=0}^m \beta_i \log\frac{\beta_i}{\epsilon}\Bigg\} \qquad (9)\\ =:& \sup_{\boldsymbol{\beta}>\boldsymbol{0}:\sum_{i=0}^m\beta_i=\epsilon} \inf_{\boldsymbol{s}>\boldsymbol{0},t>0,\boldsymbol{u}>\boldsymbol{0}} \hat{\gamma}_m(\omega,\boldsymbol{\beta},\boldsymbol{s},t,\boldsymbol{u})\\ =:& \sup_{\boldsymbol{\beta}>\boldsymbol{0}:\sum_{i=0}^m\beta_i=\epsilon} \tilde{\gamma}_m(\omega,\boldsymbol{\beta}).\end{aligned}$$

A point $(\boldsymbol{u}, t, \boldsymbol{s})$ which achieves the minimum of the function $\hat{\gamma}_m(\omega, \boldsymbol{\beta}, \boldsymbol{s}, t, \boldsymbol{u})$ is given in a solution of the following equations for all $i \in [1, m]$:

$$\beta_i = \frac{s_i}{Q}\frac{\partial Q}{\partial s_i} = \sum_{j\in\mathcal{L}} L_j \frac{j\begin{bmatrix} m \\ i \end{bmatrix} t s_i^j}{1 + t\sum_{k=1}^m \begin{bmatrix} m \\ k \end{bmatrix} s_k^j}, \qquad (10)$$

$$\omega = \frac{t}{Q}\frac{\partial Q}{\partial t} = \sum_{j\in\mathcal{L}} L_j \frac{t\sum_{i=1}^m \begin{bmatrix} m \\ i \end{bmatrix} s_i^j}{1 + t\sum_{k=1}^m \begin{bmatrix} m \\ k \end{bmatrix} s_k^j}, \qquad (11)$$

$$\beta_i = \frac{u_i}{P}\frac{\partial P}{\partial u_i} = \sum_{k\in\mathcal{R}} R_k(1-r)\frac{u_i}{f_k(\boldsymbol{u})}\frac{\partial f_k}{\partial u_i}, \qquad (12)$$

where

$$\frac{\partial f_k}{\partial u_i} = \sum_{\boldsymbol{d}\geq\boldsymbol{0}:\sum_j d_j=k} \binom{k}{d_0,\ldots,d_k} \frac{d_i}{u_i} \prod_{j=1}^m u_j^{d_j} \sum_{\boldsymbol{a}_k\in D_k} B_k(\boldsymbol{a}_k).$$

The point $\boldsymbol{\beta}$ which gives the maximum of $\gamma_m(\omega, \boldsymbol{\beta})$ needs to satisfy the stationary condition

$$s_k u_k \begin{bmatrix} m \\ k \end{bmatrix} \left(\epsilon - \textstyle\sum_{i=1}^m \beta_i\right) = \beta_k, \qquad (13)$$

for $k = 1, 2, \ldots, m$.

Lemma 6: For a given degree distribution pair (λ, ρ), we have $\gamma_m(\omega) \geq \gamma_1(\omega)$.

proof: We consider a fixed ω. Define β^* such that $\gamma_1(\omega) = \tilde{\gamma}_1(\omega, \beta^*)$. Note that

$$\begin{aligned}&\tilde{\gamma}_1(\omega,\beta^*)\\ &= \inf_{\substack{s_1>0,t>0,\\ u_1>0}} \Bigg[\sum_{j\in\mathcal{L}} \log\left(1+ts_1^j\right)^{L_j} + \sum_{j\in\mathcal{R}} \log\{\tilde{f}_k^{(1)}(u_1)\}^{R_j(1-r)}\\ &\quad + \epsilon\log\frac{\epsilon-\beta^*}{\epsilon} - \beta^*\log\frac{s_1u_1(\epsilon-\beta^*)}{\beta^*} - \omega\log t\Bigg], \qquad (14)\end{aligned}$$

where $\tilde{f}_k^{(1)}(u_1) = \{(1+u_1)^j - ju_1\}$. For $m > 1$, define $\boldsymbol{\beta}^{(m)}(\delta) := (\delta, \ldots, \delta, \beta^* - (m-1)\delta)$. For any $\delta > 0$, we have $\gamma_m(\omega) \geq \tilde{\gamma}_m(\omega, \boldsymbol{\beta}^{(m)}(\delta))$. Now, we consider $\tilde{\gamma}_m(\omega, \boldsymbol{\beta}^{(m)}(\delta))$ for $\delta \to 0$. For $\delta \to 0$, we have $s_i \to 0$ and $u_i \to 0$ for $i \in [1, m-1]$ from Eqs. (10) and (12). Note that $f_k(\boldsymbol{u}) = (1+u_m)^k - ku_m = f_k^{(1)}(u_m)$ for $u_i \to 0\ \forall i \in [1, m-1]$. Hence we have

$$\begin{aligned}&\lim_{\delta\to0}\tilde{\gamma}_m(\omega,\boldsymbol{\beta}^{(m)}(\delta))\\ &= \inf_{\substack{s_m>0,t>0,\\ u_m>0}} \Bigg\{\sum_{j\in\mathcal{L}} \log\left(1+ts_m^j\right)^{L_j} + \sum_{j\in\mathcal{R}} \log\{\tilde{f}^{(1)}(u_m)\}^{R_j(1-r)}\\ &\quad + \epsilon\log\frac{\epsilon-\beta^*}{\epsilon} - \beta^*\log\frac{s_mu_m(\epsilon-\beta^*)}{\beta^*} - \omega\log t\Bigg\}.\end{aligned}$$

This equation coincides with Eq. (14). Hence, we have $\gamma_m(\omega) \geq \gamma_1(\omega)$. □

Lemma 7: For t such that $t > 0$, Eqs. (10), (11) and (12) hold, we have

$$\frac{\mathrm{d}\gamma_m(\omega)}{\mathrm{d}\omega} = -\log t.$$

proof: Consider $\frac{\mathrm{d}\gamma_m(\omega)}{\mathrm{d}\omega}$. From Eq. (9), we have

$$\frac{\mathrm{d}\gamma_m(\omega)}{\mathrm{d}\omega}\ln 2 = -\ln t + \frac{1}{P}\frac{\mathrm{d}P}{\mathrm{d}\omega} - \frac{\omega}{t}\frac{\mathrm{d}t}{\mathrm{d}\omega}$$

$$+\frac{1}{Q}\frac{\mathrm{d}Q}{\mathrm{d}\omega}-\sum_{i=1}^{m}\frac{\beta_i}{s_i}\frac{\mathrm{d}s_i}{\mathrm{d}\omega}-\sum_{i=1}^{m}\frac{\beta_i}{u_i}\frac{\mathrm{d}u_i}{\mathrm{d}\omega}$$
$$-\sum_{i=1}^{m}\frac{\mathrm{d}\beta_i}{\mathrm{d}\omega}\ln s_i u_i\begin{bmatrix}m\\ i\end{bmatrix}\frac{\epsilon-\sum_{i=1}^{m}\beta_i}{\beta_i}.$$

From Eq. (13), we have $\ln\{s_i u_i\begin{bmatrix}m\\ i\end{bmatrix}(\epsilon-\sum_{i=1}^{m}\beta_i)/\beta_i\}=0$. From Eq. (12), we have $\frac{1}{P}\frac{\partial P}{\partial\omega}=\frac{1}{P}\sum_{i=1}^{m}\frac{\partial P}{\partial u_i}\frac{\mathrm{d}u_i}{\mathrm{d}\omega}=\sum_{i=1}^{m}\frac{\beta_j}{u_j}\frac{\mathrm{d}u_j}{\mathrm{d}\omega}$. Similarly, from Eqs. (10) and (11), we get $\frac{1}{Q}\frac{\partial Q}{\partial\omega}=\frac{\omega}{t}\frac{\mathrm{d}t}{\mathrm{d}\omega}+\sum_{i=1}^{m}\frac{\beta_i}{s_i}\frac{\mathrm{d}s_i}{\mathrm{d}\omega}$. Hence, we have $\frac{\mathrm{d}\gamma_m(\omega)}{\mathrm{d}\omega}=-\log t$. This concludes the proof. □

The following theorem shows the growth rate of the average stopping constellation distributions for small ω.

Theorem 3: For the irregular non-binary LDPC code ensemble EGL(N,m,λ,ρ) with $\lambda'(0)>0$, the growth rate of the average stopping constellation distributions of normalized weight ω, in the limit of large symbol code length for $\omega\to 0$, is given by

$$\gamma_m(\omega)=\log[\lambda'(0)\rho'(1)]\omega+o(\omega). \tag{15}$$

proof: From the definition of stopping constellation, we get $A(0)=1$ and $\gamma_m(0)=0$. From Lemma 7, we have for $\omega\to 0$

$$\gamma_m(\omega)=-\omega\log t(\omega)+o(\omega).$$

Recall that t satisfies Eqs.(9), (10), (11) and (12). From Eq.(11), for $\omega\to 0$, it holds that $t_i s_i^j\to 0$ for $i\in[1,m]$ and $j\in\mathcal{L}$. By using this and Eq.(10), we have $\beta_i\to 0$ for $i\in[1,m]$. Note that

$$f_k(\boldsymbol{u})=1+\textstyle\sum_{i=1}^{m}\binom{k}{2}\begin{bmatrix}m\\ i\end{bmatrix}u_i^2+o\Big(\big(\sum_{i=1}^{m}u_i\big)^2\Big). \tag{16}$$

Since $\beta_i\to 0$ for $i\in[1,m]$, from Eq.(12) we have $u_i\to 0$ for $i\in[1,m]$. From Eqs.(12) and (16) we get

$$\beta_i=\textstyle\sum_{k\in\mathcal{R}}R_k(1-r)2\binom{k}{2}\begin{bmatrix}m\\ i\end{bmatrix}u_i^2+o\Big(\big(\sum_{i=1}^{m}u_i\big)^2\Big).$$

Substituting this equation into Eq.(13), we have

$$s_i=u_i\rho'(1)+o(\textstyle\sum_{i=1}^{m}u_i). \tag{17}$$

Since $u_i\to 0$ for $i\in[1,m]$, we get $s_i\to 0$ for $i\in[1,m]$. From Eq.(10), it holds that

$$\beta_i=2L_2\begin{bmatrix}m\\ i\end{bmatrix}ts_i^2+o\Big(\big(\textstyle\sum_{i=1}^{m}s_i\big)^2\Big).$$

Substituting this equation into Eq.(13), we get

$$u_i=\lambda'(0)ts_i+o(\textstyle\sum_{i=1}^{m}s_i). \tag{18}$$

From Eqs.(17) and (18), we have for $\omega\to 0$

$$1=\lambda'(0)\rho'(1)t(\omega).$$

Hence, we obtain this theorem. □

Discussion 2: From Theorem 3, the growth rate for $\omega\to 0$ does not depend on m. The result of Theorem 3 coincides with the result for the weight distribution of non-binary LDPC code ensemble [11]. More precisely, the growth rates of the stopping constellation distributions and that of the weight distributions are the same for $\omega\to 0$. The techniques used in the proofs of Theorem 3 and Lemma 7 are originally developed in [12].

Define the *critical exponent stopping ratio* [13] as

$$\theta_m^*:=\inf\{\omega>0\mid\gamma_m(\omega)\geq 0\},\quad\text{for } m=1,2,\ldots.$$

From Lemma 6 and Theorem 3, we have the following corollary.

Corollary 1: For a given degree distribution pair (λ,ρ) which satisfies $\lambda'(0)\rho'(1)<1$, the critical exponent stopping ratio θ_1^* is larger than others, namely, $\theta_1^*\geq\theta_m^*$ for $m>1$.

Recall that the average stopping constellation of weight ωN is approximated by $A(\omega N)\sim 2^{\gamma_m(\omega)N}$. Since $\gamma_m(\omega)<0$ for $\omega\in(0,\theta_m^*)$, there are exponentially few stopping constellations of weight ωN for $\omega\in(0,\theta_m^*)$. It is known that the decoding erasure rate for the BEC with small channel erasure probability is caused by the stopping constellations of small weight. Therefore among LDPC codes with the degree distribution pair (λ,ρ) such that $\lambda'(0)\rho'(1)<1$ over the BEC, we see from Corollary 1 that the *binary* $(m=1)$ LDPC code ensemble is the best in the sense that there are exponentially few stopping constellations of weight ωN for ω within the widest range $(0,\theta_1^*)\supseteq(0,\theta_m^*)$.

6. Numerical Examples

In this section, we give some numerical examples of growth rate which illustrate the statement of Theorem 3 and Corollary 1.

Figures 1 and 2 show the growth rates of the average number of stopping constellations for the (2,4)-regular non-binary LDPC code ensembles defined over GL$(m,\mathbb{F}_2)$, where $m=1,2,3,4$. From these figures, especially from Fig. 2, we see that the growth rate for small ω does not depend on the dimension m. Moreover, we see that the gradient of the growth rate for small ω is $\log 3$. Similarly, Figs. 3 and 4 show the growth rates for the (3,6)-regular non-binary LDPC code ensembles. From these figures, especially from Fig. 4, we see that the growth rate for small ω does not depend on the dimension m even if $\lambda'(0)=0$.

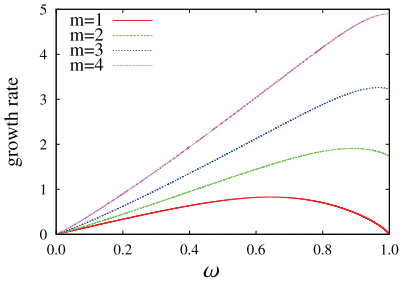

Fig. 1 The growth rates of the average stopping constellation distributions for the (2,4)-regular non-binary LDPC code ensembles defined over GL$(m,\mathbb{F}_2)$, where $m=1,2,3,4$.

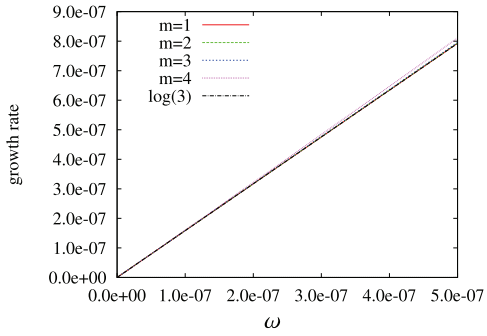

Fig. 2 The growth rates for $\omega \in [0, 5 \times 10^{-7}]$ of the average stopping constellation distributions for the (2,4)-regular non-binary LDPC code ensembles defined over $\mathrm{GL}(m, \mathbb{F}_2)$, where $m = 1, 2, 3, 4$.

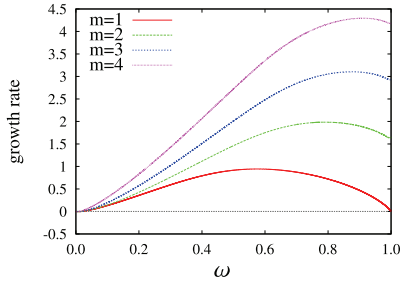

Fig. 3 The growth rates of the average stopping constellation distributions for the (3,6)-regular non-binary LDPC code ensembles defined over $\mathrm{GL}(m, \mathbb{F}_2)$, where $m = 1, 2, 3, 4$.

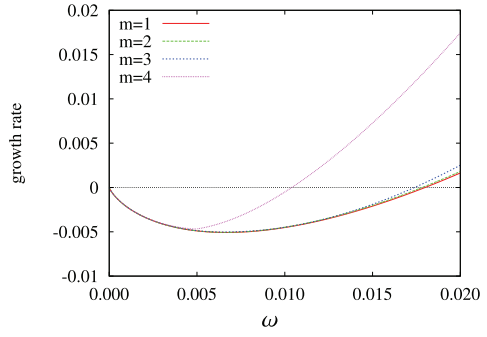

Fig. 4 The growth rates for $\omega \in [0, 0.02]$ of the average stopping constellation distributions for the (3,6)-regular non-binary LDPC code ensembles defined over $\mathrm{GL}(m, \mathbb{F}_2)$, where $m = 1, 2, 3, 4$.

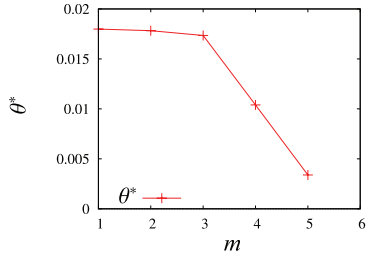

Fig. 5 The critical exponent stopping ratio of the average stopping constellation distributions for the (3,6)-regular non-binary LDPC code ensembles defined over $\mathrm{GL}(m, \mathbb{F}_2)$, where $m = 1, 2, 3, 4, 5$.

Figure 5 shows the critical exponent stopping ratio for the (3,6)-regular non-binary LDPC code ensembles defined over $\mathrm{GL}(m, \mathbb{F}_2)$, where $m = 1, 2, 3, 4, 5$. We see that the critical exponent stopping ratio monotonically decreases as the dimension m increases.

7. Conclusion and Future Work

In this paper, we derive the stopping constellation distribution and growth rate for non-binary LDPC code ensembles over general linear groups. We show that the growth rate does not depend on the dimension of the general linear group for small normalized weight. Moreover we show that the binary LDPC code ensemble is the best in terms of the critical exponent ratio for $\lambda'(0)\rho'(1) < 1$.

As a future work, we will derive the block and the symbol erasure probabilities for non-binary LDPC code ensembles.

Acknowledgment

This work was partially supported by Grant-in-Aid for JSPS Fellows.

References

[1] R.G. Gallager, Low Density Parity Check Codes, Research Monograph series, MIT Press, Cambridge, 1963.
[2] T. Richardson, M.A. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," IEEE Trans. Inf. Theory, vol.47, no.2, pp.619–637, Feb. 2001.
[3] M. Davey and D. MacKay, "Low-density parity check codes over GF(q)," IEEE Commun. Lett., vol.2, no.6, pp.165–167, June 1998.
[4] W. Chen, C. Poulliat, D. Declercq, L. Conde-Canencia, A. Al-Ghouwayel, and E. Boutillon, "Non-binary LDPC codes defined over the general linear group: Finite length design and practical implementation issues," Proc. IEEE 69th Vehicular Technology Conference, 2009, VTC Spring 2009, pp.1–5, April 2009.
[5] C. Di, D. Proietti, I. Telatar, T. Richardson, and R. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," IEEE Trans. Inf. Theory, vol.48, no.6, pp.1570–1579, June 2002.
[6] V. Rathi, "Conditional entropy of non-binary LDPC codes over the BEC," Proc. 2008 IEEE Int. Symp. Inf. Theory (ISIT), pp.945–949, July 2008.
[7] V. Rathi and R. Urbanke, "Density evolution, thresholds and the stability condition for non-binary LDPC codes," IEE Communications Proceedings, vol.152, no.6, pp.1069–1074, Dec. 2005.
[8] F.J. MacWilliams and N.J.A. Sloane, The Theory of Error-Correcting Codes, Elsevier, Amsterdam, 1977.
[9] D.E. Knuth, The Art of Computer Programming; vol.1: Fundamental Algorithms, Addison-Wesley, Reading, Massachusetts, 1973.
[10] D. Burshtein and G. Miller, "Asymptotic enumeration methods for analyzing LDPC codes," IEEE Trans. Inf. Theory, vol.50, no.6, pp.1115–1131, June 2004.
[11] K. Kasai, C. Poulliat, D. Declercq, and K. Sakaniwa, "Weight distribution of non-binary LDPC codes," IEICE Trans. Fundamentals, vol.E94-A, no.4, pp.1106–1115, April 2011.
[12] K. Kasai, T. Awano, D. Declercq, C. Poulliat, and K. Sakaniwa, "Weight distribution of multi-edge type LDPC codes," IEICE Trans. Fundamentals, vol.E93-A, no.11, pp.1942–1948, Nov. 2010.
[13] A. Orlitsky, K. Viswanathan, and J. Zhang, "Stopping set distribution of LDPC code ensembles," IEEE Trans. Inf. Theory, vol.51,

no.3, pp.929–953, March 2005.

Appendix A: Proof of Lemma 3

To prove Lemma 3, we use mathematical induction. For $k = 1$, we see that $\dim V_1 = a_1(\{1\})$ and $T_k = 0$. The number of distinct subspaces V_1 of dimension $a_1(\{1\}) = \dim V_1$ is equal to $\begin{bmatrix} m \\ a_1(\{1\}) \end{bmatrix} = \frac{[m]}{[a_1(\{1\})][a_1(\{\})]}$. Hence, Eq. (3) holds for $k = 1$.

We will show that if Eq. (3) holds for $k = k' - 1$, then Eq. (3) also holds for $k = k'$. From the induction hypothesis, we have

$$B_{k'-1}\Big(\{a_{k'}(S) + a_{k'}(S \cup \{k'\})\}_{S\subseteq[1,k'-1]}\Big) = \frac{[m]}{\prod_{S\subseteq[1,k'-1]}[a_{k'}(S) + a_{k'}(S \cup \{k'\})]} 2^{T_{k'-1}}, \quad (\mathrm{A}\cdot 1)$$

where

$$T_{k'-1} = \frac{1}{2} \sum_{\substack{S_1,S_2\subseteq[1,k'-1]: \\ S_1\not\subseteq S_2, S_1\not\supseteq S_2}} (a_{k'}(S_1) + a_{k'}(S_1 \cup \{k'\})) \times (a_{k'}(S_2) + a_{k'}(S_2 \cup \{k'\})).$$

If we fix $A_{k'}(S' \cup \{k'\})$ for all $S' \supsetneq S$, then the number of $A_{k'}(S \cup \{k'\})$ is given by

$$\begin{bmatrix} a_{k'}(S) + a_{k'}(S \cup \{k'\}) \\ a_{k'}(S \cup \{k'\}) \end{bmatrix} 2^{T_{k'}(S)}, \quad (\mathrm{A}\cdot 2)$$

where $T_{k'}(S) = a_{k'}(S \cup \{k'\}) \sum_{\tilde{S}\subseteq[1,k'-1]:\tilde{S}\supsetneq S} a_{k'}(\tilde{S})$. From Eqs. (A· 1) and (A· 2), $B_{k'}(\boldsymbol{a}_{k'})$ is given by

$$\frac{[m]}{\prod_{S\subseteq[1,k']}[a_{k'}(S)]} 2^{T_{k'-1} + \sum_{S\subseteq[1,k'-1]} T_{k'}(S)}.$$

The exponential part is written as follows:

$$\begin{aligned} &T_{k'-1} + \sum_{S\subseteq[1,k'-1]} T_{k'}(S) \\ =&T_{k'-1} + \frac{1}{2} \sum_{S\subseteq[1,k'-1]} a_{k'}(S \cup \{k'\}) \sum_{\tilde{S}:\tilde{S}\supseteq S} a_{k'}(\tilde{S}) \\ &+ \frac{1}{2} \sum_{S\subseteq[1,k'-1]} a_{k'}(S) \sum_{\tilde{S}:\tilde{S}\subseteq S} a_{k'}(\tilde{S} \cup \{k'\}) \\ =&T_{k'} \end{aligned}$$

This concludes the proof.

Appendix B: Proof of Lemma 4

From Eq. (2), we have $\sum_{S:i\in S} a_k(S) = \dim V_i$ and $\sum_{S\subseteq[1,k]} a_k(S) = m$. From Eq. (2), we have $\dim(\bigcap_{j\in[1,k]} V_j) = a([1,k])$ and $\dim(\bigcap_{j\in[1,k]\setminus\{i\}} V_j) = a([1,k]) + a([1,k]\setminus\{i\})$ for $i \in [1,k]$. From those equations, we have

$$\dim(\textstyle\bigcap_{j\in[1,k]\setminus\{i\}} V_j) = \dim(\bigcap_{j\in[1,k]} V_j) + a([1,k]\setminus\{i\}). \quad (\mathrm{A}\cdot 3)$$

Since $a([1,k]\setminus\{i\}) \geq 0$, we have

$$\dim(\textstyle\bigcap_{j\in[1,k]\setminus\{i\}} V_j) \leq \dim(\bigcap_{j\in[1,k]} V_j). \quad (\mathrm{A}\cdot 4)$$

First, we claim that for all $i \in [1,k]$, $V_i \supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$ if $a_k([1,k]\setminus\{i\}) = 0$. Since $a_k([1,k]\setminus\{i\}) = 0$, we have $\dim(\bigcap_{j\in[1,k]\setminus\{i\}} V_j) = \dim(\bigcap_{j\in[1,k]} V_j)$, from Eq. (A· 3). If $V_i \not\supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$, then $\dim(\bigcap_{j\in[1,k]\setminus\{i\}} V_j) > \dim(\bigcap_{j\in[1,k]} V_j)$. By using this equation and Eq. (A· 4), we see that $V_i \supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$ if $\dim(\bigcap_{j\in[1,k]\setminus\{i\}} V_j) = \dim(\bigcap_{j\in[1,k]} V_j)$. Thus, for all $i \in [1,k]$, $V_i \supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$ if $a_k([1,k]\setminus\{i\}) = 0$.

Next, we claim that for all $i \in [1,k]$, $a_k([1,k]\setminus\{i\}) = 0$ if $V_i \supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$. If $a_k([1,k]\setminus\{i\}) \neq 0$, we have $\dim(\bigcap_{j\in[1,k]\setminus\{i\}} V_j) > \dim(\bigcap_{j\in[1,k]} V_j)$ from Eq. (A· 3). If $V_i \supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$, then $\dim(\bigcap_{j\in[1,k]\setminus\{i\}} V_j) = \dim(\bigcap_{j\in[1,k]} V_j)$. By using this equation and Eq. (A· 4), we see that $V_i \not\supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$ if $\dim(\bigcap_{j\in[1,k]\setminus\{i\}} V_j) > \dim(\bigcap_{j\in[1,k]} V_j)$. Thus, for all $i \in [1,k]$, $V_i \not\supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$ if $a_k([1,k]\setminus\{i\}) \neq 0$. Therefore, we have for all $i \in [1,k]$, $a_k([1,k]\setminus\{i\}) = 0$ if $V_i \supseteq \bigcap_{j\in[1,k]\setminus\{i\}} V_j$. Thus, this concludes the proof.

Takayuki Nozaki received B.E. and M.E. degrees from Tokyo Institute of Technology in 2008 and 2010, respectively. He is currently pursuing the D.E. degree in Department of Communications and Integrated Systems at Tokyo Institute of Technology. His research interests are codes on graph and iterative decoding algorithm. He is a student member of IEEE.

Kenta Kasai received B.E., M.E. and Ph.D. degrees from Tokyo Institute of Technology in 2001, 2003 and 2006, respectively. Since April 2006, he has been an assistant professor in the Department of Communications and Integrated Systems, Graduate School of Science and Engineering, Tokyo Institute of Technology. His current research interests include codes on graphs and iterative decoding algorithms.

Kohichi Sakaniwa received B.E., M.E., and Ph.D. degrees all in electronic engineering from the Tokyo Institute of Technology, Tokyo Japan, in 1972, 1974 and 1977, respectively. He joined the Tokyo Institute of Technology in 1977 as a research associate and served as an associate professor from 1983 to 1991. Since 1991 he has been a professor in the Department of Electrical and Electronic Engineering, and since 2000 in the Department of Communication and Integrated Systems, Graduate School of Science and Engineering, both in the Tokyo Inst. of Tech. From November 1987 to July 1988, he stayed at the University of Southwestern Louisiana as a Visiting Professor. He received the Excellent Paper Award from the IEICE of Japan in 1982, 1990, 1992 and 1994. His research area includes Communication Theory, Error Correcting Coding, (Adaptive) Digital Signal Processing and so on. Dr. Sakaniwa is a member of IEEE, Information Processing Society of Japan and Institute of Image Information and Television Engineers of Japan.