Minimum Pseudo-Weight and Minimum Pseudo-Codewords of LDPC Codes *

Shu-Tao Xia† and Fang-Wei Fu‡

February 1, 2008

Abstract

In this correspondence, we study the minimum pseudo-weight and minimum pseudo-codewords of low-density parity-check (LDPC) codes under linear programming (LP) decoding. First, we show that the lower bound of Kelley, Sridhara, Xu, and Rosenthal on the pseudo-weight of a non-zero pseudo-codeword of an LDPC code with girth greater than 4 is tight if and only if this pseudo-codeword is a real multiple of a codeword. Then, the lower bound of Kashyap and Vardy on the stopping distance of an LDPC code is proved to be also a lower bound on the pseudo-weight of a non-zero pseudo-codeword of an LDPC code with girth 4, and this lower bound is tight if and only if this pseudo-codeword is a real multiple of a codeword. Using these results we further obtain that for some LDPC codes, there are no other minimum pseudo-codewords except the real multiples of minimum weight codewords. This means that the LP decoding for these LDPC codes is asymptotically optimal in the sense that the ratio of the probabilities of decoding errors of LP decoding and maximum-likelihood decoding approaches 1 as the signal-to-noise ratio (SNR) tends to infinity. Finally, some LDPC codes are listed to illustrate these results.

Index Terms: LDPC codes, linear programming (LP) decoding, fundamental cone, pseudo-codewords, pseudo-weight, stopping sets.

*This research is supported in part by the National Natural Science Foundation of China under the Grants 60402031, and the DSTA research grant R-394-000-025-422.

†Shu-Tao Xia is with the Graduate School at Shenzhen of Tsinghua University, Shenzhen, Guangdong 518055, P. R. China. E-mail: xiast@sz.tsinghua.edu.cn

‡Fang-Wei Fu is with the Temasek Laboratories, National University of Singapore, 5 Sports Drive 2, Singapore 117508, Republic of Singapore (on leave from the Department of Mathematics, Nankai University, Tianjin 300071, P. R. China). E-mail: tslfufw@nus.edu.sg

I Introduction

In the study of iterative decoding of low-density parity-check (LDPC) codes, Wiberg [28] and Koetter and Vontobel [12] showed that pseudo-codewords play an important role when characterizing the performance of LDPC codes. Koetter and Vontobel [12] presented an explanation for the relevance of pseudo-codewords in iterative decoding based on graph covering and showed that the set of pseudo-codewords can be described by the so-called fundamental polytope. Recently, linear programming (LP) decoding of linear codes was introduced by Feldman, Wainwright and Karger [3][4]. The feasible region of the linear programming problem in LP decoding [3][4] agrees with the fundamental polytope. It is known that when characterizing the performance of linear codes under LP decoding, pseudo-codewords, especially the pseudo-codewords with minimum pseudo-weight (or minimum pseudo-codewords for short), also play an important role. In [2], Di *et al.* showed that the performance of an LDPC code under message passing decoding over a binary erasure channel is closely related to the stopping sets in the factor graph. Since the support of any pseudo-codeword is a stopping set [12], there are some relations between the minimum pseudo-codewords and the nonempty stopping sets of smallest size [6][20][29].

Recently, pseudo-codewords and minimum pseudo-weights of binary linear codes have been studied in [1], [3], [4], [8]-[14], [21], [25]-[27], and [29]. Chaichanavong and Siegel [1, Theorem 3] gave a lower bound on the pseudo-weight of a non-zero pseudo-codeword of an LDPC code. Xia and Fu [29] showed that the Chaichanavong-Siegel bound is tight if and only if the pseudo-codeword is a real multiple of a codeword. Using this result they further obtained that for some LDPC codes, e.g., Euclidean plane and projective plane LDPC codes [15], there are no other minimum pseudo-codewords except the real multiples of minimum weight codewords. Recently, Kelley, Sridhara, Xu, and Rosenthal [8, Theorem III.1][10, Theorem 3.1] presented a lower bound on the pseudo-weight of a non-zero pseudo-codeword of an LDPC code with girth greater than 4, which includes the Chaichanavong-Siegel bound as a special case. In [6], Kashyap and Vardy gave a lower bound on the stopping distance of an LDPC code. In this correspondence, we study the minimum pseudo-weight and minimum pseudo-codewords of LDPC codes under LP decoding. The results mentioned in the abstract are obtained. The rest of this correspondence is organized as follows. In Section II, we briefly review LP decoding and pseudo-codewords of binary linear codes.

In Section III, the main results of this correspondence are given and some LDPC codes are listed to illustrate these results. In Sections IV, the proofs of the main results are given. In Section V we end with some concluding remarks.

II Preliminaries

Let C be a binary $[n, k, d]$ linear code with length n, dimension k, and minimum distance d. The codewords with (Hamming) weight d are called *minimum codewords* of C. Let A_i be the number of codewords of weight i. Let H be an $m \times n$ parity-check matrix of C, where the rows of H may be dependent. Let $I = \{1, 2, \ldots, n\}$ and $J = \{1, 2, \ldots, m\}$ denote the sets of column indices and row indices of H, respectively. The *Tanner graph* G_H corresponding to H is a bipartite graph comprising n variable nodes labelled by the elements of I, m check nodes labelled by the elements of J, and the edge set $E \subseteq \{(i, j) : i \in I, j \in J\}$, where there is an edge $(i, j) \in E$ if and only if $h_{ji} = 1$. The *girth* g of G_H, or briefly the girth of H, is defined as the minimum length of a cycle in G_H. Note that the girth g must be an even integer not smaller than 4.

Definition 1 *A stopping set S is a subset of I such that the restriction of H to S, i.e., the $m \times |S|$ sub-matrix of H consisting of the columns indexed by S, does not contain a row of weight one. The smallest size of a nonempty stopping set, denoted by $s(H)$, is called the stopping distance of C. A stopping set with size $s(H)$ is called a smallest stopping set. The number of smallest stopping sets is denoted by $T_s(H)$.*

In other words, the stopping set S is a subset of variable nodes in G_H such that all the neighbors of S are connected to S at least twice. For more results on stopping sets and stopping distance we refer the readers to [2], [6], [19], [20], and [30].

Suppose a codeword $\mathbf{c}$ is transmitted over a binary-input memoryless channel and $\mathbf{y}$ is the output of the channel. The log-likelihood ratio vector is defined by $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ where $\lambda_i = \ln \frac{\Pr\{y_i|c_i=0\}}{\Pr\{y_i|c_i=1\}}$. Let $\mathrm{conv}(C)$ be the convex hull of C in the real space $\mathbb{R}^n$. Maximum-likelihood (ML) decoding is equivalent to the following optimization problem [3][4]: Find $\mathbf{x} \in \mathrm{conv}(C)$ that minimizes $\lambda \mathbf{x}^T$. To decrease the decoding complexity, the region $\mathrm{conv}(C)$ should be relaxed. For each row $\mathbf{h}_j$ of H, $1 \le j \le m$, let $C_j = \{\mathbf{c} \in \{0, 1\}^n : \mathbf{h}_j \mathbf{c}^T = 0 \bmod 2\}$. The *fundamental polytope* of C is defined as $P(H) = \bigcap_{j=1}^{m} \mathrm{conv}(C_j)$. LP decoding then solves the following optimization problem [3][4]: Find $\mathbf{x} \in P(H)$ that minimizes $\lambda \mathbf{x}^T$. Note that $\mathrm{conv}(C) \subseteq P(H)$.

However, usually $\mathrm{conv}(C) \subset P(H)$ which implies that the LP decoder is a sub-optimal decoder. The *support* of a real vector $\mathbf{x} \in \mathbb{R}^n$, denoted by $\mathrm{supp}(\mathbf{x})$, is defined as the set of positions of non-zero coordinates in $\mathbf{x}$, or $\mathrm{supp}(\mathbf{x}) = \{i : x_i \neq 0\}$. Assuming that the channel is a binary-input output-symmetric channel, and given that the code C is linear, we can without loss of generality assume that the all-zeros codeword was transmitted. When analyzing the LP decoder for C it is then sufficient to understand the *fundamental cone* $K(H)$ of H which is the conic hull of the fundamental polytope $P(H)$. The fundamental cone $K(H)$ can be characterized as follows [3][4][12]: it is the set of vectors of $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ such that $x_i \geq 0,\ i = 1, \ldots, n$ and

$$\forall 1 \leq j \leq m, \forall\, i \in \mathrm{supp}(\mathbf{h}_j), \sum_{l \in \mathrm{supp}(\mathbf{h}_j) \setminus \{i\}} x_l \geq x_i. \tag{1}$$

The elements of $K(H)$ are called *pseudo-codewords* of C. Hence, the question of whether the LP decoder succeeds is equivalent to whether the following optimization problem has the zero vector $\mathbf{0}$ as its optimal solution: Find $\mathbf{x} \in K(H)$ that minimizes $\sum_{i=1}^{n} x_i \lambda_i$. Two pseudo-codewords $\mathbf{x}, \mathbf{y}$ are said to be equivalent if there exists a real number $\alpha > 0$ such that $\mathbf{y} = \alpha\mathbf{x}$. Clearly, $\mathbf{x} \in K(H) \Leftrightarrow \alpha\mathbf{x} \in K(H)$. For any $\mathbf{x} \in K(H)$, let $[\mathbf{x}] = \{\alpha\mathbf{x} : \alpha > 0\}$.

Definition 2 *A pseudo-codeword* $\mathbf{x}$ *is said to be internal if there exists a real number* $\beta,\ 0 < \beta < 1$ *and* $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in K(H) \setminus [\mathbf{x}]$ *such that* $\mathbf{x} = \beta\mathbf{x}^{(1)} + (1-\beta)\mathbf{x}^{(2)}$. *If a non-zero pseudo-codeword* $\mathbf{x}$ *is not internal,* $[\mathbf{x}]$ *is called an edge of* $K(H)$. *Let* $M(H)$ *denote the set of all edges of* $K(H)$. *The pseudo-codewords on edges in* $M(H)$ *are called minimal pseudo-codewords.*

It is known from [27] and linear programming theory [1][18] that the behavior of the LP decoder is completely characterized by $M(H)$ and $|M(H)|$ must be finite for fixed C and H. From now on, we only consider the binary-input additive white Gaussian noise (AWGN) channel.

Definition 3 *The (AWGN) pseudo-weight of a non-zero real vector* $\mathbf{x} \in \mathbb{R}^n$ *is defined by* $w_P(\mathbf{x}) = \|\mathbf{x}\|_1^2 / \|\mathbf{x}\|_2^2$, *where* $\|\mathbf{x}\|_1 = |x_1| + \cdots + |x_n|$ *and* $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$. *Denote by* $d_P(H)$ *the minimum pseudo-weight of non-zero pseudo-codewords of* C. *The pseudo-codewords with pseudo-weight* $d_P(H)$ *are called minimum pseudo-codewords. Define the pseudo-weight of an edge* $[\mathbf{x}] \in M(H)$ *as the pseudo-weight of* $\mathbf{x}$. *The edges with minimum pseudo-weight are called minimum edges. The number of minimum edges is denoted by* $B_P(H)$.

It is not difficult to see from linear programming theory [18] that minimum pseudo-codewords are also minimal pseudo-codewords. Note that the minimal pseudo-codewords in the same edge have the same pseudo-weight.

Just like d and A_d of a linear code are important for characterizing the performance of ML decoding, $d_P(H)$ and $B_P(H)$ are crucial for characterizing the performance of LP decoding. In order to obtain better performance, we should try to find a desirable parity-check matrix H to maximize $d_P(H)$ and then minimize $B_P(H)$. Since the support of every codeword is a stopping set [20] and every stopping set supports a pseudo-codeword [12], it is known that $d_P(H) \leq s(H) \leq d$ regardless of the choice of H, and $B_P(H) \geq T_s(H) \geq A_d$ for any H such that $d_P(H) = s(H) = d$. It is well known that LP decoding is asymptotically optimal, in the sense that the ratio of the probabilities of decoding errors of LP decoding and ML decoding approaches 1 as the SNR tends to infinity, if and only if $d_P(H) = d$ and $B_P(H) = A_d$.

Next, we give an example to illustrate the above concepts.

Example 1 Let C be a binary $[7, 3, 4]$ cyclic simplex code. The parity-check matrix H of C is formed by a 7×7 circulant matrix, where the first row is (1 1 0 1 0 0 0). H has uniform column weight 3 and girth 6. The seven non-zero codewords of C are $(1, 0, 1, 1, 1, 0, 0)$ and its cyclic shifts, each of which is a minimum codeword. All non-empty stopping sets are $\{1, 3, 4, 5\}$, $\{2, 4, 5, 6\}$, $\{3, 5, 6, 7\}$, $\{1, 4, 6, 7\}$, $\{1, 2, 5, 7\}$, $\{1, 2, 3, 6\}$, $\{2, 3, 4, 7\}$, $\{1, 2, 3, 4, 5, 6, 7\}$, where only the first 7 ones are smallest stopping sets. We choose one minimal pseudo-codeword as a representative from each edge in $M(H)$. Then all 14 representatives are $(1, 0, 1, 1, 1, 0, 0)$ and its cyclic shifts, and $(1, 2, 1, 1, 1, 2, 2)$ and its cyclic shifts, where only the first 7 ones are minimum pseudo-codewords. Thus, there are 14 edges in $M(H)$ and 7 of which are minimum edges. Clearly, C satisfies $d_P(H) = d = 4$ and $B_P(H) = A_d = 7$, which implies that LP decoding is asymptotically optimal for C.

III Main Results

Let C be a binary $[n, k, d]$ linear code with parity-check matrix H. If the Tanner graph G_H has the girth $g \geq 6$ and H has uniform column weight γ, Tanner [22] showed that

the minimum distance d fulfills $d \geq d_L$, where

$$d_L = \begin{cases} 1+\gamma+\sum_{i=1}^{(g-6)/4}\gamma(\gamma-1)^i, & g/2 \text{ odd}, \\ 1+\gamma+\sum_{i=1}^{(g-8)/4}\gamma(\gamma-1)^i+(\gamma-1)^{(g-4)/4}, & g/2 \text{ even}. \end{cases} \tag{2}$$

Orlitsky *et al.* [19] further obtained that d_L is still a lower bound on the stopping distance, i.e., $s(H) \geq d_L$. Recently, Kelley, Sridhara, Xu, and Rosenthal [8][10] proved that the minimum pseudo-weight satisfies $d_P(H) \geq d_L$, and the bound still holds when H has non-uniform column weight with minimum column weight γ. In the next theorem, which will be proved in section IV, we give a necessary and sufficient condition for $w_P(\mathbf{x}) = d_L$ to hold for a non-zero pseudo-codeword $\mathbf{x} \in K(H)$.

Theorem 1 *Let C be a binary linear code with length n. Let H be a parity-check matrix of C with girth $g \geq 6$ and minimum column weight γ. Let $\mathbf{x} = (x_1, \ldots, x_n) \in K(H)$ be a non-zero pseudo-codeword and d_L be defined in (2). Then $w_P(\mathbf{x}) = d_L$ if and only if $d_L\mathbf{x}/\sum_{l=1}^{n} x_l$ is a codeword of C with weight d_L.*

It is easy to check that

$$d_L = \begin{cases} \beta^{(g-2)/4} + 2\left(\frac{\beta^{(g-2)/4}-1}{\beta-1}\right), & g/2 \text{ odd}, \\ 2\left(\frac{\beta^{g/4}-1}{\beta-1}\right), & g/2 \text{ even}, \end{cases} \tag{3}$$

where $\beta = \gamma - 1$. In particular,

$$d_L = \begin{cases} \beta+2, & g=6, \\ 2(\beta+1), & g=8, \\ \beta^2+2\beta+2, & g=10, \\ 2(\beta^2+\beta+1), & g=12, \\ \beta^3+2\beta^2+2\beta+2, & g=14, \\ 2(\beta^3+\beta^2+\beta+1), & g=16. \end{cases} \tag{4}$$

Clearly, for the codes satisfying the conditions of Theorem 1, $d = d_L$ will imply $d_P(H) = s(H) = d$. Furthermore, by Theorem 1, we have the next corollary.

Corollary 2 *Let C be a binary $[n, k, d]$ linear code with length n, dimension k and minimum distance d. Let H be a parity-check matrix of C with girth g ($g \geq 6$) and minimum column weight γ. If $d = d_L$, where d_L is defined in (2), then $B_P(H) = T_s(H) = A_d$, where A_d is the number of minimum codewords, $T_s(H)$ is the number of smallest stopping sets, and $B_P(H)$ is the number of minimum edges.*

Remark 1 *For a code C satisfying the conditions of Corollary 2, the minimum codewords, the nonempty stopping sets of smallest size and the minimum edges are all equivalent, which implies that LP decoding is asymptotically optimal for C.*

In Example 1, $\gamma = 3$, $g = 6$, $d = 4$, and $d_L = 1 + \gamma = 4 = d$. Hence, C satisfies the conditions of Theorem 1 and Corollary 2, and $B_P(H) = T_s(H) = A_d = 7$.

Note that [29, Theorems 1 and 2] are the special case of $g = 6$ of Theorem 1 and Corollary 2, respectively. In [29], it is shown that two classes of finite geometry LDPC codes, i.e., the projective plane LDPC codes and Euclidean plane LDPC codes [15], meet the conditions of Corollary 2. Thus, LP decoding is asymptotically optimal for finite plane LDPC codes. Below we give some more examples of LDPC codes satisfying the conditions of Corollary 2.

Example 2 A class of regular LDPC codes called $LU(3, q)$ codes were constructed in [7], where q is a prime power. $LU(3, q)$ codes have the following parameters, where n is the code length, d is the minimum distance, m is the number of rows of the parity-check matrix, ρ is the uniform row weight of the parity-check matrix, γ is the uniform column weight of the parity-check matrix, and g is the girth of the Tanner graph.

$$n = q^3,\ m = q^3,\ d = 2q,\ \rho = q,\ \gamma = q,\ g = 8.$$

This class of LDPC codes meet the conditions of Corollary 2. Thus, LP decoding is asymptotically optimal for $LU(3, q)$ codes.

Example 3 In [17] [24], regular LDPC codes were constructed from generalized polygons. In [17, Table 1], for a prime power q, LDPC codes $W(q)$, $H(3, q^2)$, $H(q)$, $T(q^3, q)$, and $\bar{O}(q)$ have the following parameters, where n is the code length, d is the minimum distance, γ is the uniform column weight of the parity-check matrix, and g is the girth of the Tanner graph.

(i) $W(q):\ n = (q+1)(q^2+1),\ d = 2(q+1),\ \gamma = q+1,\ g = 8$;

(ii) $H(3, q^2):\ n = (q^2+1)(q^3+1),\ d = 2(q+1),\ \gamma = q+1,\ g = 8$;

(iii) $H(q):\ n = (q+1)(q^4+q^2+1),\ d = 2(q^2+q+1),\ \gamma = q+1,\ g = 12$;

(iv) $T(q^3, q):\ n = (q^3+1)(q^8+q^4+1),\ d = 2(q^2+q+1),\ \gamma = q+1,\ g = 12$;

(v) $\bar{O}(q), q = 2^{2e+1}:\ n = (q^2+1)(q^3+1)(q^6+1),\ d = 2(q^3+q^2+q+1),\ \gamma = q+1,$ $g = 16.$

By (4), it is obvious that these LDPC codes meet the conditions of Corollary 2. Thus, LP decoding is asymptotically optimal for them.

Example 4 In [21], some LDPC codes with $d_P(H) = d$ were constructed by enumerating a regular tree for a fixed number l of layers and employing a connection algorithm based on mutually orthogonal Latin squares to close the tree.
(i) *Type-I A* construction [21]: It is known that if l or $g/2$ is odd, then $d = d_L$. Hence, these LDPC codes with odd $g/2$ meet the conditions of Corollary 2 and LP decoding is asymptotically optimal for them.
(ii) *Type-II* construction [21]: For the binary case and $l = 3$, the Type II construction yields exactly the projective plane LDPC codes [23][29]. For the binary case and $l = 4$, it is conjectured that $d = d_L$ in [21]. Clearly, if this conjecture is true, then these LDPC codes meet the conditions of Corollary 2 and LP decoding is asymptotically optimal for them. In particular, it is known from [21] and [24] that this is true for the $(2, 2)$-Finite-Generalized-Quadrangles-based LDPC codes.

The next theorem shows that the lower bound of Kashyap and Vardy [6] on the stopping distance of an LDPC code is also a lower bound on the pseudo-weight of a non-zero pseudo-codeword of this LDPC code, and this lower bound is tight if and only if this pseudo-codeword is a real multiple of a codeword. The proof of this theorem will be given in section IV.

Theorem 3 *Let C be a binary linear code with length n. Let H be an $m \times n$ parity-check matrix of C with minimum column weight γ. If any two distinct columns of H have at most λ common 1's and γ/λ is an integer, then $w_P(\mathbf{x}) \geq \frac{\gamma}{\lambda} + 1$ for any non-zero pseudo-codeword $\mathbf{x} = (x_1, \ldots, x_n) \in K(H)$. Moreover, equality holds if and only if $(\frac{\gamma}{\lambda} + 1)\mathbf{x}/\sum_{l=1}^{n} x_l$ is a codeword of C with weight $\frac{\gamma}{\lambda} + 1$.*

Remark 2 *If H has uniform column weight γ, Kashyap and Vardy* [6, Theorem 1] *showed that the stopping distance $s(H) \geq \frac{\gamma}{\lambda} + 1$. Since $d_P(H) \leq s(H)$, Theorem* 3 *implies the Kashyap-Vardy lower bound on the stopping distance.*

Clearly, for codes satisfying the conditions of Theorem 3, $d = \frac{\gamma}{\lambda} + 1$ will imply $d_P(H) = s(H) = d$. Furthermore, by Theorem 3, we obtain the next corollary.

Corollary 4 *Let C be a binary $[n, k, d]$ linear code with length n, dimension k and minimum distance d. Let H be a parity-check matrix of C with minimum column weight γ. If any two distinct columns of H have at most λ common 1's and γ/λ is an integer, and if $d = \frac{\gamma}{\lambda} + 1$, then $B_P(H) = T_s(H) = A_d$, where A_d is the number of minimum codewords, $T_s(H)$ is the number of smallest stopping sets, and $B_P(H)$ is the number of minimum edges.*

Remark 3 *In Theorem 3 and Corollary 4, the girth of the Tanner graph G_H is at least 6 if $\lambda = 1$ and 4 if $\lambda > 1$. The special case of $\lambda = 1$ in Theorem 3 and Corollary 4 are exactly* [29, Theorems 1 and 2] *and the special case of $g = 6$ in Theorem 1 and Corollary 2.*

Example 5 Consider the binary $[2^r-1, 2^r-r-1, 3]$ Hamming code. Let H be a parity-check matrix which consists of all the non-zero codewords of the binary $[2^r - 1, r, 2^{r-1}]$ simplex code. It is easy to see that $\gamma = 2^{r-1}$ and $\lambda = 2^{r-2}$. Hence, by Theorem 3 and Corollary 4, $d_P(H) = 3$ and any minimum pseudo-codeword must be the real multiple of some minimum codeword.

Let $q = 2^s$ and $EG(m, q)$ be the m-dimensional Euclidean Geometry over $GF(q)$. It is known from [16] and [23] that there are q^m points and $q(q^m - 1)/(q - 1)$ hyperplanes in $EG(m, q)$. By removing a point of $EG(m, q)$ together with the $(q^m - 1)/(q - 1)$ hyperplanes containing this point, we obtain a slightly modified incidence matrix H of points and hyperplanes in $EG(m, q)$. Suppose the rows of H indicate the hyperplanes. The point-hyperplane Euclidean geometry LDPC code C with the parity-check matrix H has the following parameters: length $n = q^m - 1$, uniform column weight of H $\gamma = q^{m-1}$, uniform row weight $\rho = q^{m-1}$, girth of Tanner graph $g = 4$ if $m > 2$. It is easy to see that any two distinct columns of H have at most $\lambda = q^{m-2}$ common 1's. By Theorem 3, we have that $d \geq s(H) \geq d_P \geq \gamma/\lambda + 1 = q + 1$. From [16] and [23], we know that H can be put in cyclic form and the generator polynomial $g(x)$ can be determined. By [16, p. 315, (8.33)], it is known that the dimension $k = 2^{sm} - (m+1)^s$. The following examples show that $d = q + 1$ in some cases.

Example 6 Let $m = 3$ and $s = 2$. Then C is a binary $[63, 48]$ code with generator polynomial $g(x) = 1 + x^2 + x^4 + x^{11} + x^{13} + x^{14} + x^{15}$ [16, pp. 310-311]. By Theorem 3, we know that $d \geq s(H) \geq d_P \geq q + 1 = 5$. In fact, it is easy to calculate by computer

that d does equal 5. For example, $1 + x^{23} + x^{33} + x^{36} + x^{37}$ is a weight-5 codeword. Hence, by Corollary 4, we have that $A_d = T_s(H) = B_P(H)$.

Example 7 Let $m = 3$ and $s = 3$. Then C is a binary $[511, 448]$ code [23, Example 1]. The girth of the Tanner graph is 4 and C performs very well under iterative decoding [23]. By Theorem 3, we know that $d \geq s(H) \geq d_P \geq q + 1 = 9$. In fact, it can be calculated by the method in [5] that $d = 9$. Hence, by Corollary 4, we have that $A_d = T_s(H) = B_P(H)$ and LP decoding for C is asymptotically optimal.

IV Proofs of Theorems 1 and 3

In this section, we prove Theorems 1 and 3. Chaichanavong and Siegel [1, Proposition 2] gave a lower bound on the pseudo-weight of a real vector. In [29], the necessary and sufficient condition for this bound being tight is discussed. Let u be a positive integer. Denote $\mathcal{F}_u$ the set of vectors $\mathbf{y} \in [0, 1/u]^n$ such that $\sum_{i=1}^n y_i = 1$.

Lemma 1 [29] For any $\mathbf{y} \in \mathcal{F}_u$, we have $w_P(\mathbf{y}) \geq u$. Equality holds if and only if $\mathbf{y}$ has exactly u non-zero components with value $1/u$.

A. Proof of Theorem 1

From the proof of [10, Theorem 3.1], we know that for any non-zero $\mathbf{x} = (x_1, \ldots, x_n) \in K(H)$, $d_L x_i \leq \sum_{j=1}^n x_j$ for any $i \in \mathrm{supp}(\mathbf{x})$. Let $\mathbf{y} = \mathbf{x}/\sum_{i=1}^n x_i$. Then $\mathbf{y} \in K(H)$ and $\mathbf{y} \in \mathcal{F}_{d_L}$. Hence, by Lemma 1, $w_P(\mathbf{x}) = w_P(\mathbf{y}) \geq d_L$, where equality holds if and only if $\mathbf{c} = d_L\mathbf{x}/\sum_{l=1}^n x_l$ is a binary vector with weight d_L. Now, we show that $w_P(\mathbf{x}) = d_L$ if and only if $\mathbf{c} \in C$ and $w_H(\mathbf{c}) = d_L$. If $\mathbf{c} \in C$ and $w_H(\mathbf{c}) = d_L$, then $w_P(\mathbf{x}) = w_P(\mathbf{c}) = w_H(\mathbf{c}) = d_L$. On the other hand, if $w_P(\mathbf{x}) = d_L$, then the pseudo-codeword $\mathbf{c}$ is a binary vector with weight d_L. Next, we show that $\mathbf{c} \in C$.

Clearly, $S = \mathrm{supp}(\mathbf{c})$ is a stopping set with size d_L since $\mathbf{c}$ is a pseudo-codeword. For any fixed $i \in S$, we construct a local tree of i (see Figure 1) as in the proof of [10, Theorem 3.1]. For the sake of convenience, we briefly describe the construction procedure as follows. Below we use f, e to denote check nodes and i, j to denote variable nodes of the Tanner graph G_H. Let $t = \lfloor (g-6)/4 \rfloor \geq 0$, where $\lfloor x \rfloor$ is the floor function which denotes the maximum integer not greater than x. Then $g = 4t + 6$ for odd $g/2$ and $g = 4t + 8$ for even $g/2$. In the local tree of i, i is the root of the tree. A check node f connected to i is called a child of i, and a variable node j connected to f except

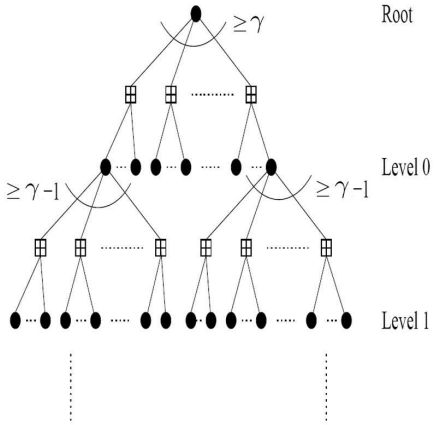

Figure 1: Local tree structure for a $\geq \gamma$-left graph

its parent i is called a child of f or a grandchild of i, and a check node e connected to j except its parent f is called a child of j, and so on. For a variable node j in the local tree, let child(j) and grch(j) denote the sets of all children and grandchildren of j respectively. Note that

$$\text{grch}(j) \quad = \quad \bigcup_{f \in \text{child}(j)} \text{child}(f). \tag{5}$$

All nodes in $L_0(i) = \text{grch}(i)$ are called *Level*-0 variable nodes. For $m = 1, 2, \ldots, t$, all nodes in

$$L_m(i) = \bigcup_{j \in L_{m-1}(i)} \text{grch}(j) \tag{6}$$

are called *Level-m* variable nodes. Fixing a check node $f^* \in \text{child}(i)$, denote $N_0(f^*) = \text{child}(f^*)$ and

$$N_m(f^*) \quad = \quad \bigcup_{j \in N_{m-1}(f^*)} \text{grch}(j), \quad m = 1, 2, \ldots, t+1. \tag{7}$$

The local tree of i has t levels if $g = 4t+6$ and $t+1$ levels if $g = 4t+8$, where $N_{t+1}(f^*)$ is the set of $(t+1)$-th level nodes. Since the Tanner graph G_H has girth $g \geq 6$, the local tree of i has the following pairwise disjoint properties: (i) all child(f) in the union of (5) are pairwise disjoint, and all grch(f) in the union of (6) are pairwise disjoint; (ii) if $g = 4t+6$, $\{i\}, L_0(i), \ldots, L_t(i)$ are pairwise disjoint; (iii) if $g = 4t+8$, all grch(f) in

the union of (7) are pairwise disjoint, and $\{i\}, L_0(i), \ldots, L_t(i)$, $N_{t+1}(f^*)$ are pairwise disjoint.

Since there are at least γ 1's in every column of H, from the construction we have that

$$|\mathrm{child}(i)| \geq \gamma \quad \text{and} \quad |\mathrm{child}(j)| \geq \gamma - 1 \tag{8}$$

for any intermediate variable node j in the local tree of i. Let j be a variable node which has some children in the local tree of i. Suppose $j \in S$. For each check node $f \in \mathrm{child}(j)$, $|\mathrm{child}(f) \cap S| \geq 1$ since S is a stopping set including j. Thus, noting that $i \in S$, by (8) and the pairwise disjoint properties, we have

$$|L_0(i) \cap S| = |\mathrm{grch}(i) \cap S| \geq \gamma; \quad |\mathrm{grch}(j) \cap S| \geq \gamma - 1 \ \text{ if } \ j \in S \setminus \{i\}. \tag{9}$$

Moreover, a necessary condition for equality in $|L_0(i) \cap S| = \gamma$ is that for each $f \in \mathrm{child}(i)$, $|\mathrm{child}(f) \cap S| = 1$. In other words, for any row $\mathbf{h}$ of H whose i-th component is 1, $w_H(\mathbf{h}_S) = 2$ where $\mathbf{h}_S$ is the restriction of $\mathbf{h}$ to S. Furthermore, by (8), (9) and the pairwise disjoint properties, for $m = 1, 2, \ldots, t$,

$$\begin{aligned} |L_m(i) \cap S| &= \Big| \bigcup_{j \in L_{m-1}(i)} \mathrm{grch}(j) \cap S \Big| \geq \Big| \bigcup_{j \in L_{m-1}(i) \cap S} \mathrm{grch}(j) \cap S \Big| \\ &\geq (\gamma - 1)|L_{m-1}(i) \cap S| \geq \cdots \geq (\gamma - 1)^m |L_0(i) \cap S| \geq (\gamma - 1)^m \gamma \end{aligned}$$

and if $g = 4t + 8$,

$$\begin{aligned} |N_{t+1}(f^*) \cap S| &= \Big| \bigcup_{j \in N_t(f^*)} \mathrm{grch}(j) \cap S \Big| \geq (\gamma - 1)|N_t(f^*) \cap S| \\ &\geq \cdots \geq (\gamma - 1)^{t+1} |\mathrm{child}(f^*) \cap S| \geq (\gamma - 1)^{t+1}. \end{aligned}$$

In other words, $|\{i\} \cap S| = 1$, $|L_0(i) \cap S| \geq \gamma$, $|L_1(i) \cap S| \geq \gamma(\gamma - 1)$, …, $|L_t(i) \cap S| \geq \gamma(\gamma - 1)^t$, and $|N_{t+1}(f^*) \cap S| \geq (\gamma - 1)^{t+1}$ if $g = 4t + 8$. Therefore, we have $|S| \geq d_L$ by adding the above inequalities and using the pairwise disjoint properties, where a necessary condition of $|S| = d_L$ is that $|L_0(i) \cap S| = \gamma$, that is, $w_H(\mathbf{h}_S) = 2$ for any row $\mathbf{h}$ of H whose i-th component is 1. This implies that $\mathbf{c}$ satisfies all the parity-check equations corresponding to the rows $\mathbf{h}$ of H whose i-th component is 1. Thus, when i varies in $S = \mathrm{supp}(\mathbf{c})$, $\mathbf{c}$ must satisfy every parity-check equation in H, i.e., $\mathbf{c}$ is a codeword.

B. Proof of Theorem 3

Let $\mathbf{y} = \mathbf{x}/\sum_{j=1}^{n} x_j$. Since $\mathbf{x} \in K(H)$, then $\mathbf{y} \in K(H)$. For fixed j, $1 \le j \le n$, let $\mathbf{h}_{q_1}, \mathbf{h}_{q_2}, \ldots, \mathbf{h}_{q_\gamma}$ be the rows of H whose j-th components are 1, i.e., $\mathbf{h}_{q_i} = (h_{q_i,1}, \ldots, h_{q_i,n})$ and $h_{q_i,j} = 1$ for $1 \le i \le \gamma$. Since $\mathbf{y} \in K(H)$, $y_j \le \sum_{l \ne j} y_l h_{q_i,l}, \quad i = 1, \ldots, \gamma$. Hence,

$$\gamma y_j \le \sum_{i=1}^{\gamma} \sum_{l \ne j} y_l h_{q_i,l} = \sum_{l \ne j} y_l \left(\sum_{i=1}^{\gamma} h_{q_i,l} \right).$$

For any $l \ne j$, since the l-th column and j-th column of H have at most λ common 1's and $h_{q_i,j} = 1$ for $1 \le i \le \gamma$, we have that $\sum_{i=1}^{\gamma} h_{q_i,l} \le \lambda$. Therefore,

$$\gamma y_j \le \lambda \sum_{l \ne j} y_l = \lambda(1 - y_j), \quad \text{i.e.,} \quad y_j \le \frac{\lambda}{\gamma + \lambda},$$

which implies that $\mathbf{y} \in \mathcal{F}_{\frac{\gamma}{\lambda}+1}$. Hence, by Lemma 1, $w_P(\mathbf{x}) = w_P(\mathbf{y}) \ge \frac{\gamma}{\lambda} + 1$, and equality holds if and only if $\mathbf{y}$ has exactly $\frac{\gamma}{\lambda} + 1$ non-zero components with value $\lambda/(\gamma + \lambda)$. In that case, $(\frac{\gamma}{\lambda} + 1)\mathbf{y} = (\frac{\gamma}{\lambda} + 1)\mathbf{x}/\sum_{j=1}^{n} x_j$ must be a binary vector, say $\mathbf{c}$, with weight $\frac{\gamma}{\lambda} + 1$.

Now, we show that $\mathbf{c}$ must be a codeword of C. Since $\mathbf{c} \in K(H)$, $\mathrm{supp}(\mathbf{c})$ is a stopping set of H, i.e., the restriction of H to $\mathrm{supp}(\mathbf{c})$, say $H(\mathbf{c})$, has no rows of weight one. Note that any two distinct columns of $H(\mathbf{c})$ have at most λ common 1's. Suppose $\mathbf{b}$ is a non-zero row of $H(\mathbf{c})$ and the j-th component of $\mathbf{b}$ is 1, where $j \in \{1, 2, \ldots, \frac{\gamma}{\lambda}+1\}$. Since the j-th column of $H(\mathbf{c})$ has at least γ 1's, there exists a $\gamma \times (\frac{\gamma}{\lambda} + 1)$ matrix, say $H(\mathbf{c}, j)$, consisting of $\mathbf{b}$ and other $\gamma - 1$ rows of $H(\mathbf{c})$ such that the j-th column of $H(\mathbf{c}, j)$ is the all-1 column. Since any two distinct columns of $H(\mathbf{c}, j)$ have at most λ common 1's, each of the columns other than the j-th column has at most λ 1's. Now we count the number of 1's in $H(\mathbf{c}, j)$, say Δ, in two ways. From the view of columns, $\Delta \le \gamma + \lambda\frac{\gamma}{\lambda} = 2\gamma$. From the view of rows, since $\mathrm{supp}(\mathbf{c})$ is a stopping set, each row of $H(\mathbf{c}, j)$ has at least two 1's, which implies $\Delta \ge 2\gamma$. Thus, $\Delta = 2\gamma$, and every row of $H(\mathbf{c}, j)$ has exactly two 1's, which implies $w_H(\mathbf{b}) = 2$. Hence, the weights of rows of $H(\mathbf{c})$ are either 0 or 2, which implies that $\mathbf{c}$ satisfies every parity-check equation in H and thus is a codeword.

V Conclusions

In this correspondence, we study the minimum pseudo-weight and minimum pseudo-codewords of LDPC codes. We characterize the pseudo-codewords of an LDPC code which attain the lower bound d_L of Kelley, Sridhara, Xu, and Rosenthal on the minimum pseudo-weight. That is, the pseudo-weight of a pseudo-codeword of an LDPC code is equal to d_L if and only if this pseudo-codeword is a real multiple of a codeword with weight d_L. Furthermore, it is shown that if the minimum distance of this LDPC code is equal to d_L, then the minimum codewords, the nonempty stopping sets of smallest size and the minimum edges are all equivalent, which implies that LP decoding is asymptotically optimal for this LDPC code. Then, we show that the lower bound of Kashyap and Vardy on the stopping distance of an LDPC code is also a lower bound on the pseudo-weight of a non-zero pseudo-codeword of an LDPC code with girth 4. The same characterization results mentioned above for the lower bound of Kelley, Sridhara, Xu, and Rosenthal are also obtained for this new lower bound on the minimum pseudo-weight. Some LDPC codes are listed to illustrate these results. Finally, we pose a further research problem: For a binary LDPC code C, construct a parity-check matrix H with minimum number of rows such that the minimum pseudo-weight of C is equal to the minimum distance of C, and the number of minimum edges is equal to the number of minimum codewords of C, i.e., LP decoding is asymptotically optimal for this LDPC code. Until now, we do not even know whether such a parity-check matrix exists for every binary linear code.

Acknowledgment

The authors would like to thank Dr. Sridhara for kindly affording the preprint [10], and Mr. X. Ge for his calculation of the minimum distance in Example 6 by computer. The authors wish to express their appreciation to the three anonymous reviewers and Associate Editor Marc Fossorier for their valuable suggestions and comments that helped to greatly improve the correspondence.

References

[1] P. Chaichanavong and P. H. Siegel, "Relaxation bounds on the minimum pseudo-weight of linear block codes," *Proc. IEEE Int. Symp. Inform. Theory*, pp. 805-809, Adelaide, Australia, Sep. 2005.

[2] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1570-1579, 2002.

[3] J. Feldman, *Decoding Error-Correcting Codes via Linear Programming*, Ph.D. Thesis, Massachusetts Institute of Technology, Sep. 2003.

[4] J. Feldman, M. J. Wainwright, and D. R. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. Inform. Theory*, vol. 51, no. 3, pp. 954-972, 2005.

[5] X.-Y. Hu, M. P. C. Fossorier, E. Evangelos, "On the computation of the minimum distance of LDPC codes," *Proc. IEEE International Conference on Communications*, vol. 2, pp. 767-771, Jun. 2004.

[6] N. Kashyap and A. Vardy, "Stopping sets in codes from designs," *Proc. IEEE Int. Symp. Inform. Theory*, p. 122, Yokohama, Japan, Jun. 29 - Jul. 4, 2003. Full version available online at http://www.mast.queensu.ca/~nkashyap/Papers/stopsets.pdf.

[7] J.-L. Kim, U. N. Peled, I. Perepelitsa, V. Pless, S. Friedland, "Explicit construction of families of LDPC codes with no 4-cycles," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2378-2388, 2004.

[8] C. Kelley, D. Sridhara, J. Xu, and J. Rosenthal, "Pseudocodeword weights and stopping sets," *Proc. IEEE Int. Symp. Inform. Theory*, p. 68, Chicago, USA, Jun. 27-Jul. 2, 2004.

[9] C. Kelley and D. Sridhara, "Structure of pseudocodewords in Tanner graphs", *Proceedings of* 2004 *International Symposium on Information Theory and Its Applications*, Parma, Italy, Oct. 2004.

[10] C. Kelley and D. Sridhara, "Pseudocodewords of Tanner graphs," preprint, 2005, available online at http://www.arxiv.org/abs/cs.IT/0504013.

[11] C. Kelley, D. Sridhara, and J. Rosenthal, "Tree-based construction of LDPC codes having good pseudocodeword weights," preprint, 2005, available online at http://www.arxiv.org/abs/cs.IT/0510009.

[12] R. Koetter and P. O. Vontobel, "Graph covers and iterative decoding of finite-length codes," *Proc. 3rd Int. Conf. Turbo Codes and Related Topics*, pp. 75-82, Brest, France, Sep. 2003.

[13] R. Koetter, W.-C. W. Li, P. O. Vontobel, and J. L. Walker, "Pseudo-codewords of cycle codes via zeta functions," *Proc. IEEE Information Theory Workshop*, pp. 7-12, San Antonio, Texas, USA, Oct. 2004.

[14] R. Koetter, W.-C. W. Li, P. O. Vontobel, and J. L. Walker, "Characterization of pseudo-codewords of LDPC codes," preprint, 2005, available online at http://www.arxiv.org/abs/cs.IT/0508049.

[15] Y. Kou, S. Lin, and M. P. C. Fossorier, "Low-density parity-check codes based on finite geometries: A rediscovery and new results," *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2711-2736, 2001.

[16] S. Lin and D. J. Costello Jr., *Error Control Coding: Fudamentals and Applications*, 2nd ed., Upper Saddle River, NJ: Prentice-Hall, 2004.

[17] Z. Liu and D. A. Pados, "LDPC codes from generalized polygons," *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3890-3898, 2005.

[18] S. G. Nash and A. Sofer, *Linear and Nonlinear Programming*, McGraw-Hill, 1996.

[19] A. Orlitsky, R. Urbanke, K. Viswanathan, and J. Zhang, "Stopping sets and the girth of Tanner graphs," *Proc. IEEE Int. Symp. Inform. Theory*, p. 2, Lausanne, Switzerland, Jun. 30 - Jul. 5, 2002.

[20] M. Schwartz and A. Vardy, "On the stopping distance and the stopping redundancy of codes," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 922-932, 2006.

[21] D. Sridhara, C. Kelley, and J. Rosenthal, "Tree-based construction of LDPC codes," *Proc. IEEE Int. Symp. Inform. Theory*, pp. 845-849, Adelaide, Australia, Sep. 2005.

[22] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory*, vol. 27, no. 5, pp. 533-547, Sep. 1981.

[23] H. Tang, J. Xu, S. Lin, and K. A. S. Abdel-Ghaffar, "Codes on finite geometries," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 572-596, 2005.

[24] P. O. Vontobel and R. M. Tanner, "Construction of codes based on finite generalized quadrangles for iterative decoding," *Proc. IEEE Int. Symp. Inform. Theory*, p. 223, Washington, DC, U.S.A., Jun. 2001.

[25] P. O. Vontobel and R. Koetter, "Lower bounds on the minimum pseudo-weight of linear codes," *Proc. IEEE Int. Symp. Inform. Theory*, p. 70, Chicago, USA, Jun. 27-Jul. 2, 2004.

[26] P. O. Vontobel and R. Smarandache, "On minimal pseudo-codewords of Tanner graphs from projective planes," *Proc. 43rd Allerton Conf. on Communications, Control, and Computing*, Allerton House, Monticello, Illinois, USA, Sep. 2005.

[27] P. O. Vontobel, R. Smarandache, N. Kiyavash, J. Teutsch, and D. Vukobratovic, "On the minimal pseudo-codewords of codes from finite geometries," *Proc. IEEE Int. Symp. Inform. Theory*, pp. 980-984, Adelaide, Australia, Sep. 2005.

[28] N. Wiberg, *Codes and Decoding on General Graphs*, Ph.D. dissertation, Linköping University, Linköping, Sweden, 1996.

[29] S.-T. Xia and F.-W. Fu, "On the minimum pseudo-codewords of LDPC codes," *IEEE Communications Letters*, vol. 10, no. 5, pp. 363-365, May 2006.

[30] S.-T. Xia and F.-W. Fu, "On the stopping distance of finite geometry LDPC Codes," *IEEE Communications Letters*, vol. 10, no. 5, pp. 381-383, May 2006.