A Combinatorial Family of LDPC Codes

K. Murali Krishnan L. Sunil Chandran Priti Shankar

IISc-CSA-TR-2006-4
http://archive.csa.iisc.ernet.in/TR/2006/4/

Computer Science and Automation
Indian Institute of Science, India

May 2006

A Combinatorial Family of LDPC Codes

K. Murali Krishnan L. Sunil Chandran Priti Shankar

Abstract

An elementary combinatorial construction for a family of Low Density Parity Check (LDPC) codes is presented. The construction allows flexibility in the choice of design parameters like rate, average degree, girth and block length of the code. The construction complexity of codes of the family grows only quadratically with the block length.

1 Introduction

The fact that iterative decoding on LDPC codes performs well when the underlying Tanner graph [5] has large girth was observed right from the time of their inception [6]. Recent revival of interest in LDPC codes owing to their near capacity performance on various channel models has resulted in considerable research in the construction of LDPC code families of high rate and large girth. These constructions fall roughly into random codes (for example see [1, 11]), codes based projective and combinatorial geometries (see [12, 13, 14] and references therein), circulant matrices [2, 3, 9], algebraic constructions (see [16, 8]), expander codes [15] etc. These codes perform very well and most are fairly easy to implement although the mathematical machinery required to understand them is often complex. However, the construction complexity of the codes are often large albeit polynomial.

In this note, we present an elementary graph theoretic construction for an LDPC code family. The construction gives flexibility in fixing the parameters of the code and is an adaptation of a large girth graph construction known in the graph theory literature [7] for the problem of code design. We prove some very general bounds on code parameters achieved by the construction. The complexity of the construction grows quadratically with the block length of the code. The construction here is similar in spirit to the graph construction in [17] discovered independently. The theoretical advantage of the proposed method is that the lower bound on the girth of the resultant graph is independent of the particular run of the algorithm whereas, the bound in [17] depends on the parameter d_c — the maximum degree of a right vertex in the graph, which does not seem to be explicitly bounded although the parameter achieves good values in practice.

2 The Code Construction

Given a bipartite graph $G = (L, R, E \subseteq L \times R)$, $|L| = n$, $|R| = m$, the $m \times n$ parity check matrix $H(G) = [h_{i,j}]$ defined by $h_{i,j} = 1$ if and only if $(j, i) \in E$ specifies a binary

linear code $C(G)$. We say G is the *Tanner graph* for $C(G)$. $C(G)$ is an LDPC code if the maximum degree of any vertex in G is bounded by a constant. The length of the shortest cycle in G is called the *girth* of G denoted by $g(G)$. In the following, we describe the construction of a bipartite Tanner graph and give bounds on the parameters of the code defined by the graph.

Let m, n, p, q and d be positive integers with $p < q$, $np = mq$ and let $d < min\{(m + 4p + 1)/(5p - 1), (n + 4q + 1)/(5q - 1)\}$ be a fixed constant. We construct a bipartite graph $G = (L, R, E)$ as follows. Initially $L = \{1, 2, ..., n\}$, $R = \{1, 2, ..., m\}$ and $E = \emptyset$. We denote by $deg(x)$ the degree of a vertex $x \in L \cup R$. Define *weighted degree* of a node, $w(x) = q.deg(x)$ for $x \in L$ and $w(x) = p.deg(x)$ for $x \in R$. Denote by $\delta(x, y)$ the length of the shortest path from x to y in G. Clearly $deg(x) = w(x) = 0$ and $\delta(x, y) = \infty$ for all $x, y \in L \cup R$ initially. The algorithm operates in d phases. During each phase np edges are added to G. Each edge is formed by connecting a vertex in G of minimum weighted degree to the farthest possible node in G preserving bipartition. The steps are formalized below:

Repeat steps below for each *phase* i, $i = 1$ to $i = d$

- Repeat until np edges are added
- 1. Select $x \in L \cup R$ such that $w(x) \leq w(y)$ for all $y \in L \cup R$.
  2. If $x \in L$, Let $S = \{z \in R : \delta(x, z) > 1$ and $deg(z) < qi + q\}$. Select a $y \in S$ such that $\delta(x, y) \geq \delta(x, z)$ for all $z \in S$.
  3. If $x \in R$, Let $S = \{z \in L : \delta(x, z) > 1$ and $deg(z) < pi + p\}$. Select a $y \in S$ such that $\delta(x, y) \geq \delta(x, z)$ for all $z \in S$.
  4. Add (x, y) to E.

Theorem 2.1 *$C(G)$ is an LDPC code with rate $R \geq 1 - p/q$.*

Proof Since $H(G)$ is an $m \times n$ matrix, $R \geq 1 - m/n$. Since $m/n = p/q$ by assumption, the claim on rate follows. By construction, the left and right degrees of any node in G is bounded by $pd + p$ and $qd + q$, and hence the graph is low density. ■

The following statement proven by induction establishes the invariants maintained by the algorithm.

Lemma 2.2 *After phase i, $1 \leq i \leq d$, $pi - p \leq deg(x) \leq pi + p$ for each $x \in L$ and $qi - q \leq deg(y) \leq qi + q$ for each $y \in R$.*

Proof Initially the hypothesis holds. Assume the statement true for some i, $0 \leq i < d$ and consider the sequence of edges added during phase $i + 1$. Let e_j, $-p \leq j \leq p$ and e'_k, $-q \leq k \leq q$ denote the number of vertices of degrees $pi + j$ and $qi + k$ in L and R respectively after phase i. By induction hypothesis

$$\sum_{j=-p}^{p} e_j = n \tag{1}$$

Since each phase adds np edges to G, the sum of degrees of all vertices in L must be inp after phase i. This gives:

$$\sum_{j=-p}^{p}(ip+j)e_j = inp \tag{2}$$

substitution for n using (1) yields:

$$\sum_{j=-p}^{1} je_j = \sum_{j=1}^{p} je_j \tag{3}$$

Also,

$$\sum_{j=-p}^{p} je_j \leq \sum_{j=-p}^{p} pe_j \leq pn \tag{4}$$

where the last inequality follows from (1). Using (3) we get:

$$\sum_{j=-p}^{1} je_j = \sum_{j=1}^{p} je_j \leq np/2 \tag{5}$$

Similarly,

$$\sum_{k=-q}^{1} ke'_k = \sum_{k=1}^{q} ke'_k \leq mq/2 \tag{6}$$

The upper bound on the degree of a vertex in phase $i+1$ is explicitly guaranteed by the algorithm. Hence to show that the induction hypothesis will hold after phase $i+1$, it suffices to show that every vertex in L has degree at least ip and every vertex in R has degree at least iq after the algorithm has completed phase $i+1$.

To satisfy the above condition, during phase $i+1$, the algorithm is required to increase the degrees of vertices in L by $\sum_{j=-p}^{1} je_j$ which amounts to at most $np/2$ (by (5)) and degree of vertices in R by $\sum_{k=-q}^{1} ke'_k$ which amounts to at most $mq/2$ (by (6)). Since $np = np/2 + mq/2$ edges are added in phase $i+1$, the number of edges added in phase $i+1$ is sufficient to satisfy the induction hypothesis. Therefore, it suffices to show that the edges are always added to the deficient vertices.

Since the algorithm adds an edge to a vertex of minimum weighted degree, the algorithm will pick vertices in L and R with degree less than ip and iq respectively for adding edges before considering vertices of higher degrees. Since the number edges added to G during phase $i+1$ is sufficient to increase the degrees of all vertices in L and R to at least ip and iq respectively, it follows that when the algorithm successfully completes phase $i+1$, every vertex in L and R must have degree at least ip and iq respectively as required by the induction hypothesis.

However it remains to be shown that the algorithm will indeed complete phase $i+1$ successfully. The algorithm may fail to complete phase $i+1$ if at some stage the set S constructed by the algorithm is empty. We must also prove that the degree of the vertex of minimum weighted degree does not exceed the degree bound. The following claim rules out such cases:

Claim 2.3 *During phase $i+1$, suppose $x \in L$ (respectively $x \in R$) satisfies $w(x) \leq w(y)$ for all $y \in L \cup R$, then $deg(x) < (i+1)p$ (respectively $deg(x) < (i+1)q$) and there exists a vertex $y \in R$ (respectively $y \in L$) that satisfies $\delta(x,y) > 1$ and $deg(y) < q(i+1)+q$ (respectively $deg(y) < p(i+1)+p$).*

Proof Assume the contrary. Let $x \in L$. Then every non-neighbour of x has degree $q(i+1)+q$. Since $w(x) \leq w(y)$ for all $y \in L$, x must have minimum degree in L. As the average left degree of G after phase $i+1$ is at most $(i+1)p$, $deg(x) \leq (i+1)p - 1$ (establishing one part of the claim, the case when $x \in R$ is handled similarly). Thus x must have at least $m - (i+1)p + 1$ non-neighbours. By induction hypothesis, $deg(x) \geq ip - p$ and each neighbour of x has degree at least $iq - q$. Hence, the total degrees of all nodes in R adds up to:

$$(m-(i+1)p+1)((i+1)q+q) + (ip-p)(iq-q) \leq m(i+1)q \tag{7}$$

where, the right side of the inequality follows from the fact that the average degree of nodes in R after phase $i+1$ must be $(i+1)q$. On simplification this yields:

$$i+1 \geq (m+4p+1)/(5p-1) \tag{8}$$

A contradiction as $i+1 \leq d < (m+4p+1)/(5p-1)$ by assumption. The case when $x \in R$ is proved similarly. ■

This completes the proof of the lemma. ■

Theorem 2.4 $g(G) \geq 2\log_{st} min\{m(st-1)/2(s+1), n(st-1)/2(t+1)\} = \Omega(logn)$, *where $s = pd + p - 1$, $t = qd + q - 1$.*

Proof Assume that a smallest length cycle in G of length $g(G) = 2r$ was formed during phase i of the algorithm. Assume $x \in L$ had the least weighted degree and was connected to $y \in R$ causing the cycle. Let $T = \{z \in R : \delta(x,z) \geq g\}$. x had to be connected to y and not to any node in T because $deg(z) = qi + q$ for all $z \in T$. i Note that this implies that $|T| \leq e'_q$ during phase $i+1$. But $qe'_q \leq \sum_{k=1}^{q} ke'_k \leq mq/2$ by (6). Hence $|T| \leq e'_q \leq m/2$. This yields the lower bound $|R - T| \geq m/2$. But all nodes in $R - T$ must be at a distance at most $g - 1 = 2r - 1$ from x. Since the maximum left and right degrees of a node in G are bounded by s and t respectively, the number of such nodes is bounded above by $(s+1) + (s+1)(st) + ... + (s+1)(st)^{r-1} \leq (s+1)(st)^r/(st-1)$. Combining the lower and upper bounds,we get:

$$m/2 \leq (s+1)(st)^r/(st-1). \tag{9}$$

A similar argument for the case $x \in R$ and $y \in L$ yields the inequality:

$$n/2 \leq (t+1)(st)^r/(st-1). \tag{10}$$

The statement of the theorem follows as one of (9),(10) must hold. ■

3 Complexity

Assuming adjacency list representation for the graph, the selection of a farthest non-neighbour satisfying the degree bound necessary during each edge addition may be performed by a simple breadth first search in $O(n)$ time. Since the total number of edge additions is linear when d is fixed constant, the overall construction complexity is $O(n^2)$.

References

[1] C. Di, D. Proietti, I. E. Telatar, T. J. 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.

[2] T. Tian, C. Jones, J. D. Villasenor, R. D. Wesel, "Construction of irregular LDPC codes with low error floors", *IEEE Intl. Conf. Comm.*, 2003, pp. 3125-3129.

[3] A. Ramamoorthy, R. Wesel, "Construction of short block length irregular LDPC codes", *ICC 2004*, Paris, June 2004.

[4] A. Orlitsky, R. Urbanke, K. Viswanathan, J. Shang, "Stopping sets and girth of Tanner graphs", *ISIT 2002*, June 2002.

[5] M. Tanner, "A recursive approach to low-complexity codes", *IEEE Trans. Info. Theory*, Vol. 27, pp. 533-547, Sept 1981.

[6] R. G. Gallager, "Low density parity-check codes", MIT Press, 1963.

[7] L. Sunil Chandran, "A High girth graph construction", *SIAM J. Discrete Math.*, Vol. 16, no. 3, pp. 366-370, 2003.

[8] R. M. Tanner, D. Sridhara, T. Fuja, "A class of group structured LDPC codes", *Proc. ICSTA 2001*, Ambleside, England, 2001.

[9] R. M. Tanner, D. Sridhara, A. Sridharan, T. Fuja, D. J. Costello Jr., "LDPC block and convoluational codes based on circulant matrices", *IEEE Trans. Info. Theory*, Vol. 50, no.12, 2004. .

[10] C. Kelley, D. Sridhara, "Pseudocodewords of Tanner Graphs", *arXiv: CS. IT/0504013*, April 2005.

[11] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, D. Spielman, "Improved low density parity check codes using irregular graphs and belief propagation", *IEEE Trans. Info. Theory*, Vol 47, pp.585-588, Feb. 2001.

[12] Y. Kou, S. Lin, M. Fossorier, "Low density parity check codes based on finite geometries: A rediscovery and new results", *IEEE Trans. Info. Theory*, Vol 47, pp.2711-2736, Nov. 2001.

[13] B. Vasic, O. Milenkovic, "Combinatorial constructions of low density parity check codes for iterative decoding" *IEEE Trans. Info. Theory*, Vol 50, No. 6, June 2001.

[14] H. Tang, J. Xu, Y. Mou, S. Lin, K. Abdel-Ghaffar, "On algebraic construction of Gallager and circulant low-density parity-check codes", *IEEE. Trans. Info. Theory*, Vol. 50, No. 6, June 2004.

[15] M. Sipser, D. A. Spielman, "Expander Codes", *IEEE. Trans. Info. Theory*, Vol. 42, pp.1710-1722, Nov, 1996.

[16] J. Rosenthal, P. O. Vontobel, "constructions of regular and irregular LDPC codes using Ramanujan graphs and ideas from Margulis", *Proc. ISIT 2001*, p 4. June 2001.

[17] Xiao-Yu Hu, "Regular and irregular progressive edge-growth Tanner graphs", *IEEE. Trans. Info. Theory*, vol. 51, no. 1, Jan. 2005, pp. 386-398.