On the Fundamental System of Cycles in the Bipartite Graphs of LDPC Code Ensembles

Igal Sason

Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel

2009 IEEE International Symposium on Information Theory
Seoul, Korea

June 28– July 3, 2009.

Some Background on Cycles in Graphs

Definition

A **tree** is a connected graph that has no cycles.

$\Rightarrow$ A removal of any edge from a tree makes the graph disconnected. Any two vertices are connected by a single path.

Definition

A **spanning tree** of a connected graph $\mathcal{G}$ is a tree which spans all the vertices of $\mathcal{G}$.

$\Rightarrow$ By repeatedly removing edges which originally create cycles in the graph, it follows that every connected graph has a spanning tree.

Some Background on Cycles in Graphs (Cont.)

Definition

The **number of components** of a (possibly disconnected) graph $\mathcal{G}$ is the minimal number of its connected subgraphs whose union forms the graph $\mathcal{G}$.

Definition

The **cycle rank** of $\mathcal{G}$, denoted by $\beta(\mathcal{G})$, is defined as the maximal number of edges which can be removed from the graph without increasing its number of components.

Each component forms a spanning tree after removing these edges.

Some Background on Cycles in Graphs (Cont.)

- The cycle rank of a graph is a measure of the edge redundancy with respect to the connectedness of this graph.
- Let $\mathcal{G}$ be an un-directed graph with $|V_{\mathcal{G}}|$ vertices, $|E_{\mathcal{G}}|$ edges and $C(\mathcal{G})$ components. Then, the cycle rank satisfies the equality

$$\beta(\mathcal{G}) = |E_{\mathcal{G}}| - |V_{\mathcal{G}}| + C(\mathcal{G}).$$

Definition

Let $\mathcal{G}$ be an un-directed graph. A **full spanning forest** $\mathcal{F}$ of the graph $\mathcal{G}$ is the subgraph of $\mathcal{G}$ that results from removing the $\beta(\mathcal{G})$ edges from $\mathcal{G}$ without increasing its number of components.

$\Rightarrow$ The number of components of $\mathcal{F}$ and $\mathcal{G}$ is the same.
A graph may have a multiplicity of full spanning forests.

Some Background on Cycles in Graphs (Cont.)

Definition

Let $\mathcal{F}$ be a full spanning forest of an un-directed graph $\mathcal{G}$, and let e be an edge in the relative complement of $\mathcal{F}$. The cycle of the subgraph $\mathcal{F} \cup \{e\}$ is called a **fundamental cycle** of $\mathcal{G}$ that is associated with $\mathcal{F}$.

Each of the edges in the relative complement of a full spanning forest $\mathcal{F}$ gives rise to a *different* fundamental cycle of the graph $\mathcal{G}$.

The cardinality of the fundamental system of cycles of $\mathcal{G}$ associated with a full spanning forest of this graph is equal to the cycle rank $\beta(\mathcal{G})$.

Fundamental System of Cycles: An Example

$$H := \begin{array}{c} \\ 1 \\ \mathbf{2} \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & \mathbf{8} & 9 & 10 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & \mathbf{0} & 0 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & \mathbf{1} & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & \mathbf{1} & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & \mathbf{0} & 1 & 1 \end{pmatrix} \end{array}$$

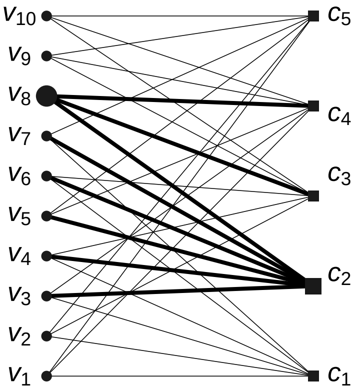

Figure: A parity-check matrix H and the corresponding bipartite graph.

- In this graph $|V(\mathcal{G})| = 15$, $|E(\mathcal{G})| = 30$, $C(\mathcal{G}) = 1$ form the number of vertices, edges, and components respectively.
- The cycle rank of $\mathcal{G}$ is therefore $\beta(\mathcal{G}) = 16$.
- To get a spanning tree of this graph, we remove repeatedly 16 edges which create cycles while preserving connectivity.

$$\widetilde{H} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & \mathbf{0} & 0 & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & \mathbf{0} & 1 & 1 & 1 & 0 & 0 \\ 0 & \mathbf{0} & 0 & \mathbf{0} & 0 & \mathbf{0} & 0 & \mathbf{0} & 1 & \mathbf{0} \\ 1 & 0 & \mathbf{0} & 0 & 1 & 0 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 1 & \mathbf{0} & 0 & 0 & \mathbf{0} & 0 & \mathbf{0} & 0 & 1 & 1 \end{pmatrix} \end{array}$$

Figure: A parity-check matrix which corresponds to a spanning tree of the bipartite graph in Fig. 1. The new parity-check matrix $\widetilde{H}$ is obtained by changing the values of the bolded 16 entries in H from 1 to 0.

Fundamental System of Cycles in Bipartite graphs

Question

How does the average cardinality of the fundamental system of cycles of bipartite graphs behave as a function of the achievable gap to capacity of the underlying LDPC code ensemble ?

Answer to this question $\Rightarrow$
Quantitative measure to the statement that bipartite graphs of good LDPC codes should have cycles (even under ML decoding).

Cardinality of the Fundamental System of Cycles of Good LDPC Code Ensembles

- Binary Linear block codes which are represented by cycle-free bipartite graphs are not good even under ML decoding.
- A theoretical treatment of cycle-free codes was provided by T. Etzion, A. Trachtenberg and A. Vardy, "Which codes have cycle-free Tanner graphs ?," *IEEE Trans. on Information Theory*, vol. 45, no. 6, pp. 2173–2181, September 1999.

Question

What can be said about the cardinality of the fundamental system of cycles of LDPC code ensembles as a function of the achievable gap (in rate) to capacity ?

Cardinality of the Fundamental System of Cycles of Good LDPC Code Ensembles (Cont.)

A lower bound on the average cardinality of the fundamental system of cycles of LDPC code ensembles was previously derived in the paper:

I. Sason and R. Urbanke, “Parity-check density versus performance of binary linear block codes over memoryless symmetric channels,” *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1611–1635, July 2003.

In the following, we present an improved information-theoretic lower bound.

Theorem

Let $\{(n, \lambda, \rho)\}$ *be a sequence of LDPC code ensembles transmitted over an MBIOS channel. Suppose that the design rate is a fraction* $1 - \varepsilon$ *of the channel capacity* C*, and the average bit error probability of this sequence vanishes under some decoding algorithm as* $n \to \infty$*. Consider the average cardinality of the fundamental system of cycles,* $\beta_n(\mathcal{G})$*, where the graphs* $\mathcal{G}$ *are chosen uniformly at random from the LDPC code ensemble* (n, λ, ρ)*. Then, the following result holds:*

$$\liminf_{n\to\infty} \frac{\mathbb{E}[\beta_n(\mathcal{G})]}{n} \geq \frac{(1-C)\ \ln\left(g_1\ \left[1 - 2h_2^{-1}\left(\frac{1-C}{1-(1-\varepsilon)C}\right)\right]^{-2}\right)}{\ln\left(\frac{1}{g_1}\right)} - 1$$

where $g_1 \triangleq \mathbb{E}\left[\tanh^2\left(\frac{L}{2}\right)\right]$ *and* L *forms the LLR at the channel output.*

Cardinality of the Fundamental System of Cycles of Good LDPC Code Ensembles (Cont.)

Lemma

[Cardinality of the fundamental system of cycles] *Under the assumptions of the theorem, the cardinality of the fundamental system of cycles of a bipartite graph $\mathcal{G}$, associated with a full spanning forest of $\mathcal{G}$, is larger than*

$$n\Big[(1-R)(a_{\mathrm{R}}-1)-1\Big]$$

where a_{R} designates the average right degree (i.e., the average degree of the parity-check nodes).

Lower Bound on the average right degree a_R

- A lower bound is derived in this work via a previously reported lower bound on the conditional entropy for binary linear block codes transmitted over MBIOS channels:

$$\frac{H(\mathbf{X}|\mathbf{Y})}{n} \geq R - C + \frac{1-R}{2\ln 2} \sum_{p=1}^{\infty} \frac{\Gamma(g_p)}{p(2p-1)}$$

where $g_p \triangleq \mathbb{E}[\tanh^{2p}(L/2)]$ for $p \in \mathbb{N}$, and $\Gamma(x) \triangleq \sum_k \Gamma_k x^k$ forms the degree distribution of the parity-check nodes for an arbitrary representation of the code by a **full-rank** parity-check matrix (Wiechman & Sason, IEEE Trans. on IT, Feb. 2007).

- This inequality was extended in this work for the case where the parity-check matrix is not necessarily full-rank.
- The bound on the average right degree was then tightened as compared to its previous form in the IEEE Trans. on IT, Feb. 2007.

Lower Bound on the average right degree a_{R} (Cont.)

In the limit where the bit error probability vanishes as we let the block length tend to infinity, the resulting lower bound on the average right degree ($a_{\mathcal{R}}$) admits the form

$$a_{\mathrm{R}} \geq \frac{2 \ln \left(\frac{1}{1 - 2h_2^{-1}\left(\frac{1-C}{1-(1-\varepsilon)C}\right)} \right)}{\ln\left(\frac{1}{g_1}\right)}$$

where ε forms an achievable gap (in rate) to capacity (under ML decoding or any sub-optimal decoding algorithm).

Lemma

[Extreme values of g_1 among all MBIOS channels with a given capacity] *Among all the MBIOS channels with a given capacity C, the value of g_1 satisfies*

$$C \leq g_1 \leq \left(1 - 2h_2^{-1}(1 - C)\right)^2$$

and these upper and lower bounds on g_1 are attained for a BSC and BEC, respectively.

This lemma is equivalent to Theorem 1 of the paper

Y. Jiang, A. Ashikhmin, R. Koetter and A. C. Singer, “Extremal problems of information combining,” *IEEE Trans. on Information Theory*, vol. 54, no. 1, pp. 51–71, January 2008.

In the paper, we present an alternative (more elementary) proof.

Corollary

The average cardinality of the fundamental system of cycles grows at least like $\log \frac{1}{\varepsilon}$ *where the achievable design rate forms a fraction* $1 - \varepsilon$ *of the channel capacity.*

$\Rightarrow$ The fundamental system of cycles becomes unbounded as the achievable gap to capacity vanishes (even under ML decoding).

Essence of the proof of this theorem: A combination of an improved lower bound on the average right degree (which behaves like $\log \frac{1}{\varepsilon}$), which follows from the lower bound on the conditional entropy, with the lemma on slide 14.

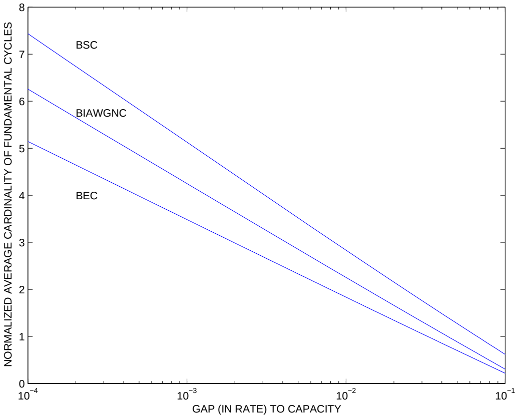

Figure: Asymptotic lower bounds on the average cardinality of the fundamental system (see Theorem 1). The bounds refer to the BSC, BIAWGNC and BEC where the design rate is $\frac{1}{2}$ bit per channel use.

Full Paper Version

I. Sason, “On Universal Properties of Capacity-Approaching LDPC Code Ensembles,” *IEEE Trans. on Information Theory*, vol. 55, no. 7, pp. 2956–2990, July 2009.