On Universal Properties of Capacity-Approaching LDPC Code Ensembles

Igal Sason

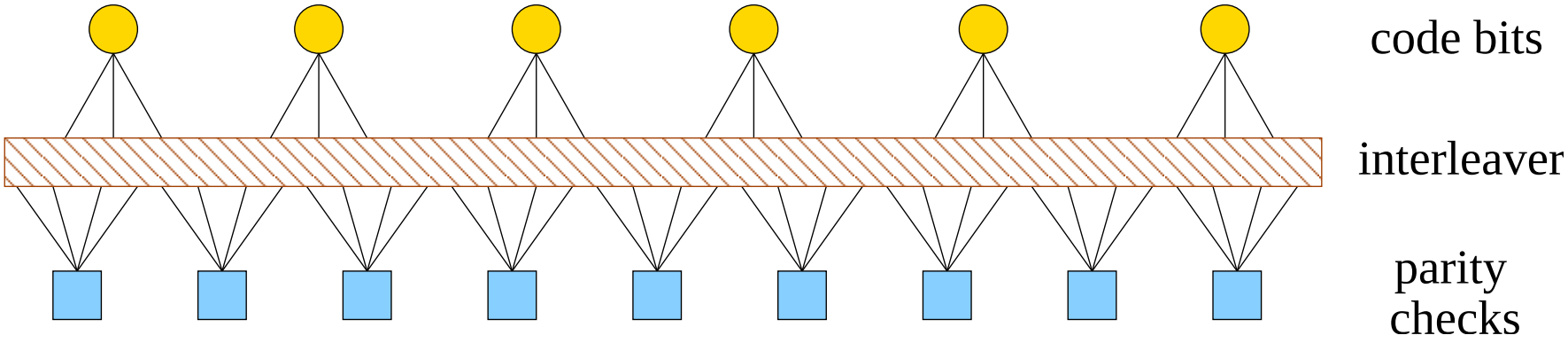
Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel

November 11, 2010

This work was supported by the Israel Science Foundation (ISF)
(Grant 1070/07)

Outline

Low-Density Parity-Check Codes



- LDPC codes are well-known capacity-approaching linear codes which are characterized by sparse parity-check matrices.
- Sparse parity-check matrices $\Rightarrow$ Capacity-approaching codes with low-complexity encoding and decoding algorithms.

Capacity-Approaching LDPC Code Ensembles

Capacity-approaching LDPC ensembles were introduced around ’00.

- M. G. Luby, M. Mitzenmacher, A. Shokrollahi and D. Spielman, “Efficient erasure correcting codes,” *IEEE Trans. on Information Theory*, vol. 47, pp. 569–583, February 2001.
- T. J. Richardson and R. Urbanke, “The capacity of low-density parity-check codes under message-passing decoding,” *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 599–618, February 2001.
- T. J. Richardson, A. Shokrollahi and R. Urbanke, “Design of capacity-approaching irregular low-density parity-check codes,” *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.
- S. Y. Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, “On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit,” *IEEE Communications Letters*, vol. 5, no. 2, pp. 58–60, February 2001.

Representing a Linear Block Code by a Bipartite Graph

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

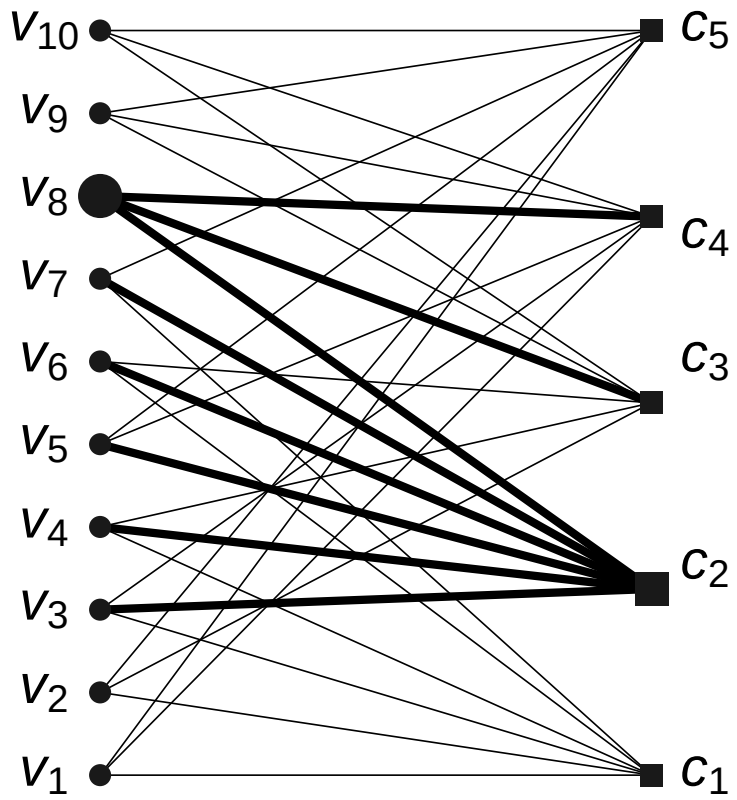


Figure: A parity-check matrix H and the corresponding bipartite graph. Column 8 & row 2 of H and the corresponding nodes & edges are bolded.

Graphical Complexity versus Performance

Question

Consider the representation of a finite-length binary linear block code by an arbitrary bipartite graph. How simple can such a graphical representation be as a function of the channel model, target block error probability, and code rate ?

Answer $\Rightarrow$

- An information-theoretic measure of the inherent graphical complexity of finite-length LDPC codes as a function of their gap to capacity.
- Provides a measure of the sub-optimality of explicit constructions of LDPC codes by comparing to some lower bounds.

Tools Used for Considering this Issue

- Sphere-packing bounds and some recent improvements for finite-length codes $\Rightarrow$ A lower bound on the required block length given the channel model, target block error probability and code rate.
- An improved (information-theoretic) lower bound on the average degree of the variable nodes $\Rightarrow$ A lower bound on the number of edges normalized per code bit.

The graphical complexity is defined as the number of edges in the bipartite graph $\Rightarrow$ The graphical complexity is lower bounded by the product of the two lower bounds above.

Sphere-Packing Bounds

- Lower bounds on the decoding error probability of optimal block codes, given in terms of
  1. block length
  2. rate
  3. communication channel
- Based on geometrical properties of the decoding regions.
- Decay to zero exponentially with the block length.

The 1967 sphere-packing (SP67) bound (Shannon et al.)

- Applies to codes transmitted over DMCs.
- Valid under optimal ML decoding or even under list decoding.
- Error exponent is exact between the critical rate and channel capacity.

Theorem (The 1967 Sphere-Packing Bound)

- Let $\mathcal{C}$ be a block code consisting of M codewords each of length N.
- Assume communication over a DMC, and let $P(j|k)$ designate the transition probabilities where $k \in \{1,\ldots,K\}$ and $j \in \{1,\ldots,J\}$ are the channel input and output alphabets, respectively.
- Assume a list decoder where the size of the list is limited to L.
- Define

$$R \triangleq \frac{\ln\left(\frac{M}{L}\right)}{N} \quad - \text{ code rate in nats per channel use}$$

$$P_{\min} \quad - \text{ smallest non-zero transition probability of the DMC.}$$

- Then, the *average decoding error probability* is lower bounded by

$$P_{\mathrm{e}}(N,M,L) \geq \exp\left\{-N\left[E_{\mathrm{sp}}\left(R - O_1\left(\frac{\ln N}{N}\right)\right) + O_2\left(\frac{1}{\sqrt{N}}\right)\right]\right\}$$

Sphere-Packing Bounds (Cont.)

- The original focus in the derivation of the SP67 bound was on asymptotic analysis.
- The aim was to make the derivation as simple as possible, as long as there is no loss in the asymptotic behavior.
- Problem: The SP67 bound is in general very loose for codes of short to moderate block lengths.
- Goal: Improve the tightness of the sphere-packing bound for finite-length codes.

This direction was studied by Valembios and Fossorier (IEEE Trans. on Information Theory, December 2004), followed by Wiechman and Sason (IEEE Trans. on Information Theory, May 2008).

Theorem (ISP Bound, IEEE Trans. on IT, May 2008)

- Let $\mathcal{C}$ be an arbitrary block code consisting of M codewords, each of length N.
- Assume communication over a symmetric memoryless channel specified by the transition probabilities (or densities) $P(j|k)$.
- Assume a list decoder where the size of the list is limited to L.
- Then, the *average decoding error probability* is lower bounded by

$$P_e(N,M,L) \geq \exp\Big\{-NE_{\text{ISP}}(R,N)\Big\}$$

where

$$E_{ISP}(R,N) \triangleq \inf_{x>\frac{\sqrt{2}}{2}} \left\{E_0(\rho_x) - \rho_x\left(R - O_1\left(\frac{1}{N},x\right)\right) + O_2\left(\frac{1}{\sqrt{N}},x,\rho_x\right)\right\}$$

On the average degree of the parity-check nodes

- In this work, we introduce an information-theoretic lower bound which is related to the average right degree of binary linear block codes which are represented by an arbitrary bipartite graph.
- This forms an improvement to some bounds previously reported by Sason and Urbanke (July 2003) and Wiechman and Sason (Feb. 2007).
- The proof of the new bound refers to
  - Finite-length code (as opposed to the asymptotic case of infinite block length in the bounds above).
  - A rigorous adaptation of the theorem to LDPC code ensembles.
- An improvement of the new bound over previously reported bounds will be exemplified.

Theorem (On the average degree of the parity-check nodes, IEEE Trans. on IT, July 2009)

- Let $\mathcal{C}$ be a binary linear block code of block length n.
- Assume the transmission takes place over a memoryless binary-input output-symmetric (MBIOS) channel.
- Let $\mathcal{G}$ be a bipartite graph which corresponds to a full-rank parity-check matrix of $\mathcal{C}$.
- Let C be the channel capacity, in bits per channel use, and L be the RV which designates the log-likelihood ratio (LLR) at the channel output, given that the binary input is 0).
- Assume that the code rate is (at least) $(1-\varepsilon)C$ (where $0 < \varepsilon < 1$), and the code achieves a block error probability P_{B} or a bit error probability P_{b} under some decoding algorithm.

Theorem (Cont.)

Then, the average right degree of the bipartite graph (i.e., the average degree of the parity-check nodes in $\mathcal{G}$) satisfies

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

where $g_1 \triangleq \mathbb{E}[\tanh^2(L/2)]$, and

$$\delta \triangleq \begin{cases} P_{\mathrm{B}} + \frac{h_2(P_{\mathrm{B}})}{n} & \text{for a block error probability } P_{\mathrm{B}} \\ h_2(P_{\mathrm{b}}) & \text{for a bit error probability } P_{\mathrm{b}} \end{cases} . \tag{2}$$

Theorem (Cont.)

For LDPC code ensembles, where the parity-check matrices are not necessarily full-rank (i.e., there is some linear dependence between the parity-check equations), the theorem still holds for **for every code from the ensemble** when the rate of a code is replaced by the design rate of the ensemble.

A proof of this theorem (IEEE Trans. on Information Theory, July 2009) follows from

- A 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 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.

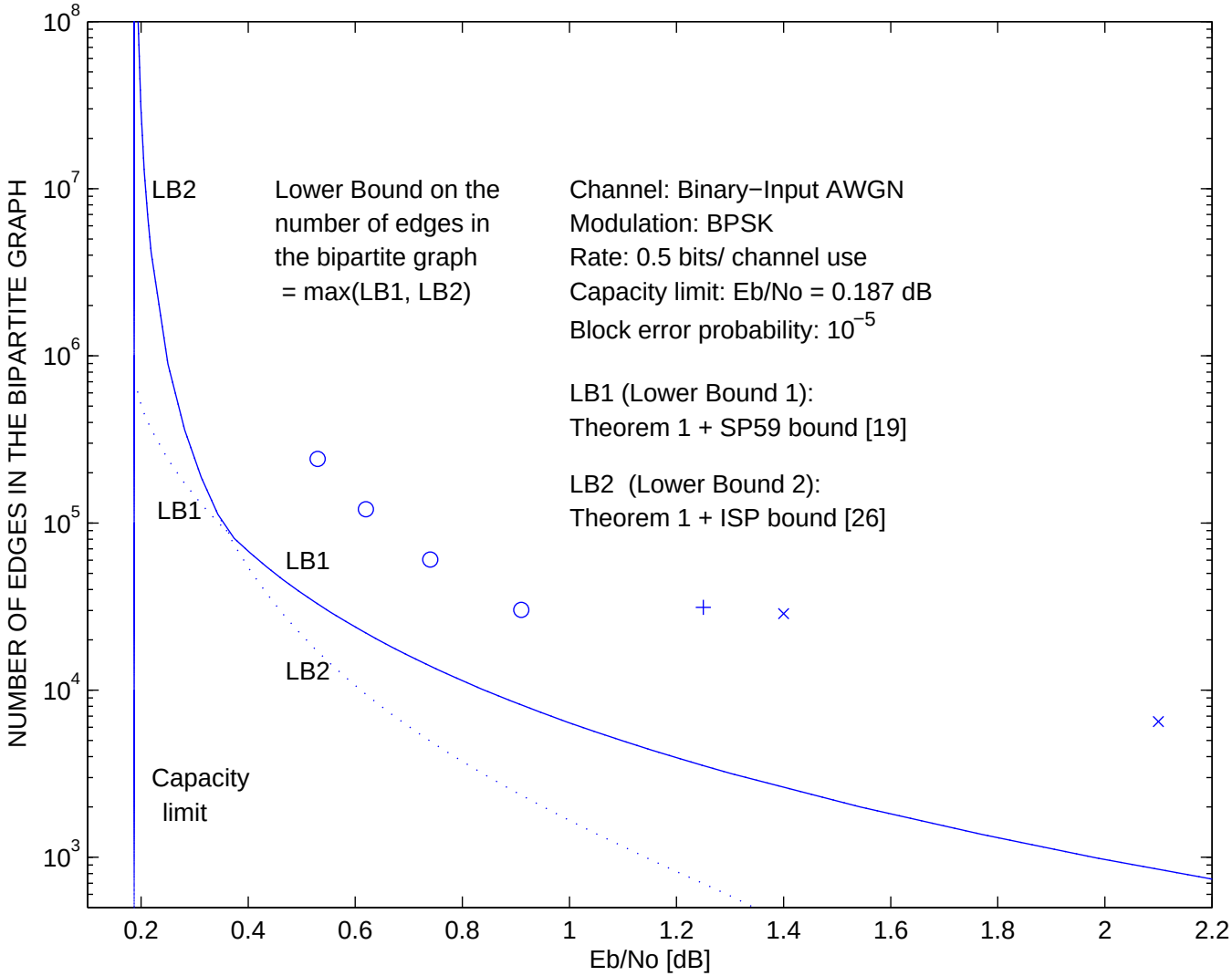


Figure: A comparison between the graphical complexity of various efficient LDPC code ensembles and an information-theoretic lower bound.

Degree Distributions of LDPC Code Ensembles

Question

How do the degree distributions of LDPC code ensembles behave as a function of their achievable gap (in rate) to capacity ?

Degree Distributions of LDPC Code Ensembles

In this work

- *Communication over a memoryless binary-input output-symmetric (MBIOS) channel is assumed.*
- *Linear programming (LP) upper bounds on the degree distributions of LDPC code ensembles are expressed in terms of their gap (in rate) to capacity and the bit/ block error probability.*
- *Analytical solutions of these bounds are obtained via Lagrange duality, and these bounds are easy to calculate.*
- *These information-theoretic bounds give an indication on the behavior of degree distributions of capacity-approaching LDPC code ensembles.*

Degree Distributions of LDPC Code Ensembles

Consider the case where transmission takes place over a memoryless, binary-input output-symmetric (MBIOS) channel.

- From the symmetry property
- Consider LDPC code ensembles whose design rate forms a fraction $1 - \varepsilon$ of the channel capacity with a target bit error probability P_{b}.

Question

What can be said about the degree distributions of the LDPC code ensembles in this setting ?

An Outline of the Derivation of the LP Bounds

- A 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

$$g_p \triangleq \int_0^{\infty} a(I)(1+e^{-I}) \tanh^{2p}\left(\frac{I}{2}\right) dI, \quad p \in \mathbb{N}.$$

and $\Gamma(x) \triangleq \sum_k \Gamma_k x^k$ forms the degree distribution of the parity-check nodes, from the node perspective, of an arbitrary representation of the code by a **full-rank** parity-check matrix (Wiechman & Sason, IEEE Trans. on IT, Feb. 2007).

Question

But what if the parity-check matrix has some linearly dependent parity-check equations ? (e.g., parity-check matrices of some LDPC code ensembles).

Theorem (Sason, '09)

For (regular and irregular) LDPC code ensembles of binary LDPC codes, the above lower bound on the conditional entropy stays valid for every code from the ensemble when

- *The rate R of the code is replaced by the design rate of the ensemble.*
- *The sequence $\{\Gamma_k\}$ denotes the degree distribution of the parity-check nodes of the ensemble (where the representation of a code by a parity-check matrix, with the given degree distribution, possibly includes some linearly dependent rows).*

Some additional things needed for the derivation of the LP bounds in this work:

- Fano inequality.
- The inequality

$$g_p \geq (g_1)^p$$

holds for every $p \in \mathbb{N}$, with equality for the BSC.

- The derivation of the LP bounds finally relies on the equality

$$\frac{1}{2 \ln 2} \sum_{k=1}^{\infty} \frac{u^k}{k(2k-1)} = 1 - h_2\left(\frac{1-\sqrt{u}}{2}\right), \ \forall u \in [0,1].$$

where h_2 designates the binary entropy function on base 2.

LP1 Bound for the Degree Distribution of the Parity-Check Nodes for LDPC Code Ensembles

$$\text{maximize} \quad \sum_{i=1}^{k} \rho_i, \quad k = 1, 2, \ldots$$

$$\text{subject to}$$

$$\begin{cases} \displaystyle\sum_{i=1}^{\infty} \left\{ \left[1 - h_2\left(\frac{1-g_1^{\frac{i}{2}}}{2} \right) \right] \frac{\rho_i}{i} \right\} \leq \frac{\varepsilon C + h_2(P_b)}{1-(1-\varepsilon)C} \displaystyle\sum_{i=1}^{\infty} \frac{\rho_i}{i} \\ \displaystyle\sum_{i=1}^{\infty} \rho_i = 1 \\ \rho_i \geq 0, \quad i = 1, 2, \ldots \end{cases}$$

where the optimization variables are $\{\rho_i\}_{i \geq 1}$. The quantity g_1 above depends on the channel statistics only.

Closed-Form Solution of the LP1 Bound

- Since strong duality holds for a feasible LP problem, then the LP1 problem can be solved via Lagrange duality.
- The dual problem gets the form

$$\begin{aligned}
&\text{minimize } \lambda_2 \\
&\text{subject to} \\
&\begin{cases}
-1 + \lambda_1 d_i + \lambda_2 - \theta_i = 0, \quad i = 1, 2, \ldots, k \\
\lambda_1 d_i + \lambda_2 - \theta_i = 0, \quad i = k+1, k+2, \ldots \\
\lambda_1, \lambda_2 \geq 0 \\
\theta_i \geq 0, \quad i = 1, 2, \ldots
\end{cases}
\end{aligned}$$

where $d_i \triangleq \frac{1}{i}\left[1 - h_2\left(\frac{1-g_1^{\frac{i}{2}}}{2}\right) - \frac{\varepsilon C + h_2(P_b)}{1-(1-\varepsilon)C}\right]$ for $i \geq 1$.

Closed-Form Solution of the LP1 Bound (Cont.)

- The sequence $\{d_i\}$ is non-negative if and only if $i \leq k_0$ where

$$k_0 \triangleq \alpha \ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C-h_2(P_b)}{1-(1-\varepsilon)C}\right)}\right), \quad \alpha \triangleq \frac{2}{\ln\left(\frac{1}{g_1}\right)}$$

- For $k \leq k_0$, the sequence $\{d_i\}_{i=1}^{k}$ is non-negative and monotonic decreasing: $d_1 > d_2 > \ldots, > d_k > 0, \quad \forall\, k \leq k_0$.
- $d_i < 0$ for $i > k_0$, and $\lim_{i\to\infty} d_i = 0$. Let $d^* \triangleq \min_{i\geq 1} d_i$ where the minimum of the sequence $\{d_i\}$ is attained for some index $i > k_0$, and $d^* < 0$.
- The optimal value of the dual LP is equal to $-\frac{d^*}{d_k - d^*}$ (which is indeed bounded between 0 and 1) for $k \leq k_0$, and 1 for $k > k_0$.

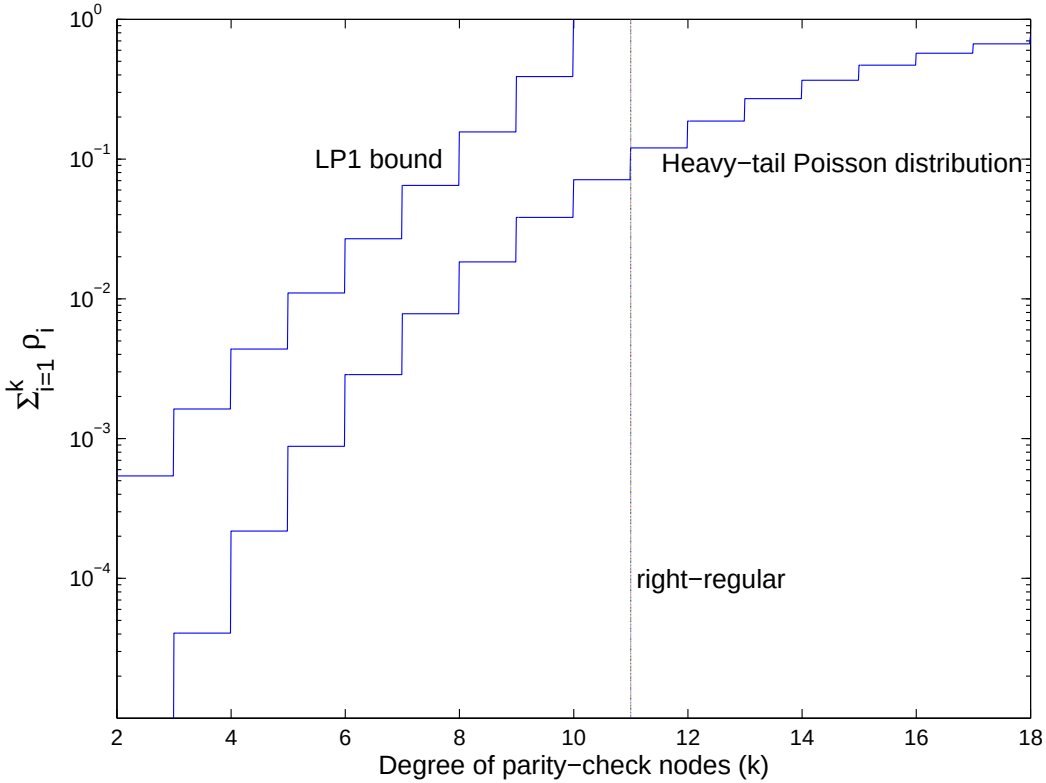


Figure: The LP1 bound for a BEC is compared to the heavy-tail Poisson distribution, and the right-regular LDPC ensemble. This figure refers to the fraction of edges which are attached to parity-check nodes of degree $\leq k$ for an integer $k \geq 2$. We refer to a BEC of capacity $\frac{1}{2}$ bit per channel use, and consider the case where 99.9% of capacity is achieved under iterative message-passing decoding with vanishing bit erasure probability.

Closed-Form Solution of the LP Upper Bound on the Degree Distribution of Variable Nodes

$$\sum_{i=1}^{k} \lambda_i \leq \min \left\{ 1, \frac{k \ln\left(\frac{1}{g_1}\right)}{2(1-C)\left(1+\frac{\varepsilon C}{1-C}\right) \ln\left(\frac{1}{1-2h_2^{-1}\left(\frac{1-C-h_2(P_b)}{1-(1-\varepsilon)C}\right)}\right)} \right\}$$

and for the universal bound for equi-capacity MBIOS channels, the parameter g_1 is replaced by the capacity (C).

Example (the Degree Distribution of the Variable Nodes)

- Consider LDPC code ensembles whose design rate is $\frac{1}{2}$ bits per channel use.
- Transmission over a Binary-input AWGN channel.
- A target bit error probability of $P_{\mathrm{b}} = 10^{-10}$ at $\frac{E_{\mathrm{b}}}{N_0} = 0.188$ dB.

From our analytical LP bound, we get

$$\lambda_2 \leq 0.2683$$
$$\lambda_2 + \lambda_3 \leq 0.4025$$
$$\lambda_2 + \lambda_3 + \lambda_4 \leq 0.5367$$
$$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 \leq 0.6709$$
$$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 \leq 0.8051$$
$$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 \leq 0.9392$$
$$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 \leq 1.0000.$$

Asymptotic Behavior of the Degree Distributions

Corollary

If the asymptotic bit error/ erasure probability vanishes, then the following properties hold for an arbitrary finite degree i

$$L_i = O(1)\,, \qquad R_i = O(\varepsilon)\,,$$
$$\lambda_i = O\left(\frac{1}{\ln\frac{1}{\varepsilon}}\right)\,, \quad \rho_i = O\left(\frac{\varepsilon}{\ln\frac{1}{\varepsilon}}\right)\,.$$

where $\{L_i\}$ and $\{R_i\}$ are the degree distributions of the variable and parity-check nodes, respectively, and $\{\lambda_i\}$ and $\{\rho_i\}$ are the corresponding degree distributions from the edge perspective.

- These bounds hold under ML decoding (or any other algorithm).
- The upper bounds on the left degree distribution look at first glance looser than those for the right degree distribution (due to the additional factor ε in the latter case).
- However, it is not an artifact of the bounding technique, as it indeed reflects reality, e.g.:
  - For various capacity-achieving degree distributions on the BEC with iterative message-passing decoding, the fraction of degree-2 variable nodes tends to $\frac{1}{2}$.
  - The upper bound on the fraction of edges connected to degree-2 variable nodes (λ_2) is shown in this work to be obtained for the right-regular LDPC code ensemble of Shokrollahi which achieves capacity on the BEC under iterative decoding.

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 ?

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:

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

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

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*

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

where a_{R} can be replaced by a lower bound.

As was already presented, a lower bound on a_{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 is the gap (in rate) to capacity.

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 this work, 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 some simple arguments from graph theory.

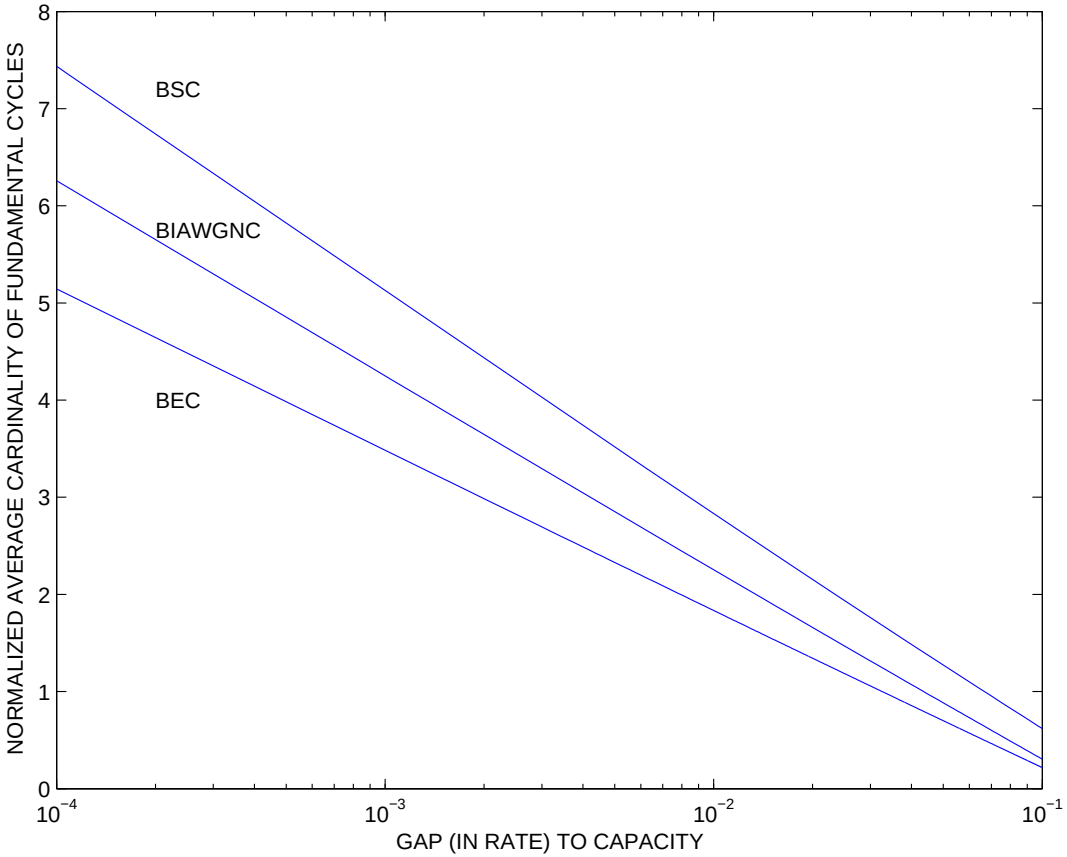


Figure: Asymptotic lower bounds on the average cardinality of the fundamental system (see Theorem 1). The bounds refer to the BSC, BIAWGNC and BEC where the design rate is $\frac{1}{2}$ bit per channel use.

Summary

- This talk presented some universal properties of capacity-approaching LDPC code ensembles whose transmission takes place over memoryless binary-input output-symmetric channels.
- These properties refer to the
  1. Graphical complexity
  2. Degree distributions
  3. Cardinality of fundamental system of cycles

  of LDPC code ensembles as a function of the fractional gap between the achievable rate and the channel capacity.
- The theoretical results rely on information-theoretic bounds.

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.