

On the Density of Parity-Check Matrices of Capacity-Approaching Binary and Linear Codes on Memoryless and Symmetric Channels

I. SASON AND R. URBANKE

Information and Communication Theory Laboratories, Swiss Federal Institute of Technology, Lausanne - EPFL

Motivation

- Low-density parity-check (LDPC) codes are well-known capacity-approaching linear codes.
- Sparse parity-check matrices \Rightarrow Low-complexity encoding and decoding algorithms. For LDPC codes, the sub-optimal iterative decoding algorithm is very efficient.
- In this work we consider the following question:
How sparse can parity-check matrices of binary linear codes be, as a function of their gap to capacity ?

Capacity-Approaching Codes

Let $\{C_m\}$ be a sequence of codes of rate R_m . It is said to achieve a fraction $1 - \epsilon$ of the channel capacity with vanishing bit (block) error probability if

- $\lim_{m \rightarrow \infty} R_m = (1 - \epsilon)C$,
- The average bit (block) error probability of the code C_m under a certain decoding algorithm tends to zero as $m \rightarrow \infty$.

Density

Let C be a binary linear code of rate R and block length n , which is represented by a parity-check matrix H . We define the *density* of H , call it $\Delta = \Delta(H)$, as the normalized number of ones in H per information bit

Information-Theoretic Lower Bound

Let $\{C_m\}$ be a sequence of binary linear codes achieving a fraction $1 - \epsilon$ of the capacity of a memoryless binary-input output-symmetric channel with vanishing bit error probability. Then,

$$\liminf_{m \rightarrow \infty} \Delta_m > \frac{K_1 + K_2 \log \frac{1}{\epsilon}}{1 - \epsilon},$$

where

$$K_1 = \frac{(1 - C) \cdot \log \left(\frac{\ln 2}{2} \cdot \frac{1 - C}{C} \right)}{2C \cdot \log \left(\frac{1}{1 - 2s} \right)}, \quad K_2 = \frac{1 - C}{2C \cdot \log \left(\frac{1}{1 - 2s} \right)}.$$

C is the channel capacity, $s = \frac{1}{2} \int_{-\infty}^{\infty} \min(f(y), f(-y)) dy$, $f(y) = p(y|x=1)$.

Lower Bound (Cont.)

For the Binary Erasure Channel (BEC), these coefficients can be improved to

$$K_1 = \frac{p \cdot \log \left(\frac{p}{1-p} \right)}{(1-p) \cdot \log \left(\frac{1}{1-p} \right)},$$

$$K_2 = \frac{p}{(1-p) \cdot \log \left(\frac{1}{1-p} \right)},$$

where p designates the probability of erasure in the BEC.

On Regular LDPC Codes used over Memoryless and Symmetric Channels

For any memoryless binary-input output-symmetric channel, there exists a sequence of ensembles of regular LDPC codes which achieves under *ML decoding* a fraction $1 - \epsilon$ of the channel capacity with vanishing *block error probability*, and

$$\lim_{n \rightarrow \infty} \Delta_n \leq \frac{K_3 + K_4 \log \frac{1}{\epsilon}}{1 - \epsilon},$$

where K_3 and K_4 only depend on the channel and

$$K_4 = \frac{(1 - C)}{(1 - a)C \cdot \log \left(\frac{1}{1 - 2h^{-1}(1 - C)} \right)},$$

On the Achievability of the Lower Bound on the Memoryless Symmetric Channels

For any memoryless binary-input output-symmetric channel, the minimal value of $\frac{K_4}{K_2}$ where K_2 and K_4 are the coefficients which are given in the lower and upper bounds on the asymptotic density satisfies

$$2 \leq \frac{K_4}{K_2} \leq \frac{2 \log \left(\frac{1}{C} \right)}{\log \left(\frac{1}{1 - 2h^{-1}(1 - C)} \right)}.$$

These lower and upper bounds on $\frac{K_4}{K_2}$ are achieved for the BSC and BEC, respectively.

Amin's LDPC Ensemble

$$\lambda_{\alpha, N}(x) = \frac{\alpha \sum_{k=1}^{N-1} \binom{\alpha}{k} (-1)^{k+1} x^k}{\alpha - N \binom{\alpha}{N} (-1)^{N+1}}$$

$$\rho_{\alpha}(x) = x^{\frac{1}{\alpha}}, \quad 0 < \alpha < 1.$$

On the Achievability of the Lower Bound on the Binary Erasure Channel

Let

$$N = \max \left(\left\lceil \frac{1 - (1-p)(1-\epsilon)c_0(p)}{\epsilon} \right\rceil, \left\lceil (1-p)^{-\frac{1}{p}} \right\rceil \right),$$

$$\alpha = \frac{\log \left(\frac{1}{1-p} \right)}{\log N}.$$

γ is Euler's constant, $c_0(p) = (1-p)^{\frac{2}{\alpha}} \cdot e^{(\frac{2}{\alpha} - \gamma)p}$ ($0 < p < 1$).

- Under *iterative message-passing decoding*, this sequence achieves asymptotically at least a fraction $1 - \epsilon$ of the channel capacity with vanishing *bit error probability*.
- The asymptotic density of its parity-check matrices satisfies the inequality

$$\lim_{n \rightarrow \infty} \Delta_n \leq \frac{K_1 + K_2 \log \frac{1}{\epsilon} + g(\epsilon, p)}{1 - \epsilon},$$

where K_1, K_2 are the coefficients in the information-theoretic lower bound. $\lim_{\epsilon \rightarrow 0^+} g(\epsilon, p)$ is upper bounded for $0 < p < 1$ by 0.5407.

Conclusion and Future Work

- The results in this work imply that for any iterative decoder based on the standard Tanner graph there is a tradeoff between performance and complexity. This tradeoff cannot be surpassed ! This begs the question if better tradeoffs can be achieved by allowing more complicated graphical models (e.g., graphs which involve state nodes, in addition to variable nodes and parity-check nodes used for representing codes by bipartite graphs).
- Is it true in general that for any sequence of codes which under an arbitrary decoding algorithm (iterative or not) achieves a certain fraction of capacity with vanishing bit or block error probability, the encoding/decoding complexity can be linked to the density of the parity-check matrices which represent these codes (or to their gap to capacity) ?