# Upper Bounds on the Rate of Ramdomly Constructed LDPC Codes for a Class of Markov Channels

Pulkit Grover and Ajit Kumar Chaturvedi
Department of Electrical Engineering
Indian Institute of Technology Kanpur
{pulkit, akc}@iitk.ac.in

September 20, 2004

## Abstract

We consider a class of finite-state Markov channels, in which channel behaves as a Binary Symmetric Channel (BSC) in each state. We find upper bounds on the rate of LDPC codes for reliable communication over this class of Markov channels. However, the results hold only for the construction of LDPC codes in which a code is selected randomly from a given ensemble of codes.

## 1 Introduction

We say a sequence of codes can be used for reliable communication over a given channel if the Maximum Likelihood (ML) decoding error probability of the sequence of codes converges to 0 as block length approaches infinity. In [3], Gallager derived an upper bound on rate of regular LDPC codes for use over Binary Symmetric Channel (BSC) for reliable communication. The bound was found to be $R \leq 1 - \frac{H(\eta)}{H(P_k)}$, where $\eta$ is the crossover probability of the BSC, $k$ is row weight of the parity check matrix, $P_k = \frac{1+(1-2\eta)^k}{2}$, and $H(.)$ is the binary entropy function. The bound was generalized by Burshtein et al in [1] for Memoryless Binary Input Output Symmetric (MBIOS) channels and irregular LDPC codes. In [6] we generalized the bound to a large class of Glibert-Elliott (GE) channels, which are two state Markov channels, with channel behaving as a BSC in each state. In this work, we generalize the bound further to the case of finite state Markov channels using the randomness in construction of LDPC codes. Note that the bound here holds only for *almost all* LDPC codes. There may exist a sequence of LDPC codes which does not satisfy this bound, though the fraction of such codes which defy the bound converges to 0 as $n \to \infty$. Importantly, the bounds hold for the most common construction of LDPC codes over memoryless channels, in which a code is selected randomly from ensemble of codes of given degree sequence. But in absence of concentration theorem for LDPC codes for Markov channels (as stated in [2]), we cannot say whether the same construction is a good construction for codes for Markov channels.

We assume that the Markov chain is ergodic, i.e., it has a steady state distribution.

As a side result, we also extend the results obtained by Sason and Urbanke [9] on density of parity check matrices for MBIOS channels to the setting of Markov channels (in [6], we had done this for GE channels). In section 4, we show that similar lower bounds (as derived in [9] on density of parity check matrices for memoryless channels and in [6] for GE channels) hold for Markov channels.

## 2 Notation

Throughout the paper, by Markov channel we mean a channel which behaves as a BSC in each of its $m$ states with crossover probability $\eta_i$ in $i^{th}$ state ($i = 1, \ldots, m$), and the transition from one state to another forms a Markov chain. $C_M$ denotes the capacity of Markov channel, and $\gamma_i$ denotes the steady state probability of $i^{th}$ state.

## 3 Upper bound on rate

The derivation of upper bound for Markov channels is broken into two parts. In Part 1, as in [6], we first reduce the problem to finding an upper bound on entropy of a parity check. In Part II, we find this upper bound. However, derivation of Part II is different from derivation of upper bound on entropy of a parity check derived in [6].

Consider a code of blocklength $n$. If the input alphabet (which we take to be same as output alphabet) $\chi$ is of size $M$, and the input and output are denoted by $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ and $\mathbf{Y} = \{Y_1, Y_2, \ldots, Y_n\}$ ($X_i, Y_i \in \chi$), then the following result holds ([4], Thm. 4.3.2, pp 79)

$$\langle P_e \rangle log(M-1) + H(\langle P_e \rangle) \geq \frac{1}{n} H(\mathbf{X}|\mathbf{Y}) \qquad (1)$$

where $\langle P_e \rangle$ is the average probability of error in each symbol. For binary case, $M = 2$, so the equation reduces to

$H(\langle P_e \rangle) \geq \frac{1}{n} H(\mathbf{X}|\mathbf{Y})$. Thus, if $\frac{1}{n} H(\mathbf{X}|\mathbf{Y})$ is strictly positive, then so is $\langle P_e \rangle$, the average bit error rate.

It follows that for reliable communication, $\frac{1}{n} H(\mathbf{X}|\mathbf{Y}) \rightarrow 0$. In what follows, we prove that for use of a sequence $\mathcal{C}_n$ of LDPC codes of blocklength $n$ over a Markov channel as the one described above, $\frac{1}{n} H(\mathbf{X}|\mathbf{Y})$ is lower bounded by a positive constant for a rate exceeding a certain bound. Hence for a fixed sequence of LDPC codes, we give an upper bound on rate for reliable communication.

**Regular Codes**

For ease of exposition, we first derive the bound for regular codes.

**Theorem 1:** Consider a binary linear code with parity check matrix $\mathbf{H}$ and rate $R$ over an $m$ state Markov channel as defined above. Suppose all rows of $\mathbf{H}$ have a constant weight $r$. Then a necessary condition for reliable communication is:

$$R \leq 1 - \frac{1 - C_M}{H(\overline{p_r})} \tag{2}$$

where

$$\overline{p_r} = \frac{1 + (1 - 2\sum_{i=1}^{m} \eta_i \gamma_i)^r}{2} \tag{3}$$

**Proof:**

Part I

We know that:

$$\begin{aligned} \frac{1}{n} I(\mathbf{X}; \mathbf{Y}) &= \frac{1}{n} H(\mathbf{X}) - \frac{1}{n} H(\mathbf{X}|\mathbf{Y}) \\ &= \frac{1}{n} H(\mathbf{Y}) - \frac{1}{n} H(\mathbf{Y}|\mathbf{X}) \end{aligned}$$

Thus,

$$\frac{1}{n} H(\mathbf{X}|\mathbf{Y}) = \frac{1}{n} H(\mathbf{X}) - \frac{1}{n} H(\mathbf{Y}) + \frac{1}{n} H(\mathbf{Y}|\mathbf{X}) \tag{4}$$

For any code, $\mathbf{X}$ is the encoded data which is in 1-1 mapping with the information symbols prior to encoding. The information symbols are uniformly distributed over the $2^{nR}$ values. Thus, $\mathbf{X}$ takes any value from the $2^{nR}$ codewords with uniform probability distribution. Thus $H(\mathbf{X}) = nR$. Also, $H(\mathbf{Y}|\mathbf{X}) = H(\mathbf{Y} + \mathbf{X}|\mathbf{X})$, where '+' is the binary bit-by-bit modulo two addition. But $\mathbf{Y} + \mathbf{X} = \mathbf{Z}$, where $\mathbf{Z}$ is the error vector, and errors are independent of input to the channel (because of symmetry). Channel errors are dependent only on channel state sequence. Thus $H(\mathbf{Y}|\mathbf{X}) = H(\mathbf{Z})$. By chain rule, $H(\mathbf{Z}) = \sum_{k=1}^{n} H(z_k|\mathbf{z_{k-1}})$, where $\mathbf{z_i}$ represents the vector $\{z_1, \ldots z_i\}$. It can be inferred from Lemma 4.2 of [5] that the sequence $\{H(z_k|\mathbf{z_{k-1}})\}_{k=1}^{\infty}$ is monotonically decreasing in $k$ and therefore, $\frac{1}{n} H(\mathbf{Z}) \geq \lim_{n \rightarrow \infty} H(z_k|\mathbf{z_{k-1}})$ Thus, (4) becomes:

$$\frac{1}{n} H(\mathbf{X}|\mathbf{Y}) \geq R - \frac{1}{n} H(\mathbf{Y}) + \lim_{n \rightarrow \infty} H(z_k|\mathbf{z_{k-1}}) \tag{5}$$

We prove in Appendix A that the capacity of such a Markov channel is given by $C_M = 1 - \lim_{n \rightarrow \infty} H(z_k|\mathbf{z_{k-1}})$. Hence, we get:

$$\frac{1}{n} H(\mathbf{X}|\mathbf{Y}) \geq R - \frac{1}{n} H(\mathbf{Y}) + 1 - C_M \tag{6}$$

Since we want to lower bound $\frac{1}{n} H(\mathbf{X}|\mathbf{Y})$, we now find an upper bound on $\frac{1}{n} H(\mathbf{Y})$. As shown in [3] and [1], the information content of $\mathbf{Y}$ is same as information content of $\mathbf{Y_1}$, which are the received bits at some $nR$ linearly independent locations in the code, and $\mathbf{P}$, which are the results of the parity check equations (This follows since given $\mathbf{Y_1}$ and $\mathbf{P}$, we can find $\mathbf{Y}$, and vice versa)[1]. Therefore,

$$H(\mathbf{Y}) = H(\mathbf{Y_1}; \mathbf{P}) = H(\mathbf{Y_1}) + H(\mathbf{P}|\mathbf{Y_1}) \leq H(\mathbf{Y_1}) + H(\mathbf{P}) \tag{7}$$

where the last inequality follows from the fact that conditioning reduces entropy. Now $\mathbf{Y_1} = \mathbf{X_1} + \mathbf{Z_1}$, where $\mathbf{X_1}$ and $\mathbf{Z_1}$ are the vectors corresponding to characters at independent locations in the transmitted codeword and the error vector respectively. Since $\mathbf{X_1}$ is the vector corresponding to $nR$ independent positions in the transmitted word, it specifies a codeword uniquely, and hence distribution of $\mathbf{X_1}$ is uniform over all its possible $2^{nR}$ values. Hence, $H(\mathbf{X_1}) = H(\mathbf{X}) = nR$. Since $\mathbf{X_1}$ has uniform distribution over all its possible $2^{nR}$ values, and $\mathbf{Z_1}$ is independent of $\mathbf{X_1}$, $\mathbf{Y_1} = \mathbf{X_1} + \mathbf{Z_1}$ also has a uniform distribution over all its possible $2^{nR}$ values. Thus $H(\mathbf{Y_1}) = nR$. Now its sufficient to upper bound $H(\mathbf{P})$, the entropy of the parity check vector. In general, the parity check values are not independent of each other. Thus, $H(\mathbf{P}) \leq \sum_{i=1}^{n(1-R)} H(p_i)$ where $p_i$ are the probabilities of individual parity checks being satisfied. The problem now reduces to bounding entropy of individual parity checks.

Part II

Consider the ensemble of LDPC codes of a fixed degree distribution. It is proved in the concentration theorem in [8] (for memoryless channels) that as the blocklength $n \rightarrow \infty$, probability of error for a randomly chosen code in the ensemble converges to the average probability of error of the ensemble. In constructing an LDPC code (over memoryless channels), a code is chosen randomly

[1] Given $\mathbf{Y}$, $\mathbf{Y_1}$ and $\mathbf{P}$ can obviously be obtained. Given $\mathbf{Y_1}$ and $\mathbf{P}$, finding $\mathbf{Y}$ is same as finding values of $\mathbf{Y}$ on locations other than locations of $\mathbf{Y_1}$. The problem reduces to solving a system of $n(1-R)$ linear equations with $n(1-R)$ variables, which has a unique solution by assumption of full rank of parity check matrix

from the ensemble of codes corresponding to the given degree distribution. For large $n$, it's behaviour would be the same as that of average over the ensemble (with high probability). For further details in the construction, we refer the reader to [8].
However, such a concentration theorem has not been proved for Markov channels, and is still an open problem [2]. We prove here that if we continue to use the same construction of LDPC codes as for memoryless channels (by randomly selecting a code from given ensemble), then the rate of code will be bounded by the bound given here. Note that if some different construction is used, the bound need not hold.

Lets consider a single parity check equation, which is a row of the parity check matrix $H$. Let the places at which 1's occur in the equation be denoted by $n_1, n_2, \ldots n_r$, and the corresponding output random variables be denoted by $Y_{n_1}, Y_{n_2}, \ldots, Y_{n_r}$. Let $\zeta = \sum_{i=1}^{r} Y_{n_i}$, where addition is over $GF(2)$. The entropy of a single parity check is given by $H(\zeta)$.
Since the input codeword to the channel $\mathbf{X}$ satisfies the parity check equations, for the received word $\mathbf{Y}$ a particular parity check equation will be satisfied as long as there are even number of errors in the symbols occurring in the parity check. Hence, we want to find

$$P(\zeta = 0) = P\Big(\text{even number of errors in} \{Y_{n_i}\}_{i=1}^{r}\Big) \quad (8)$$

Since the state space is Markov, determining exact probability is not possible without knowing the exact positions of 1's. However, since in the construction of LDPC codes, a code is chosen randomly from an ensemble of codes of given degree distribution, the bound can be found for large values of $n$. We prove that almost all codes in the ensemble satisfy this bound, as $n \to \infty$.

*Definition:* (gap) In a row of $H$, for any two 1's separated by a string of 0's, we define 1+ the number of 0's between the two 1's as the *gap* between the 1's.

**Lemma 1**: Consider the ensemble of regular codes of length $n$ of fixed left and right degrees. Choose an arbitrarily parity check equation from an arbitrarily chosen code from the ensemble of codes. Then, for any given integer $k$, the gap between any two $1's$ in the parity check equation becomes greater than $k$ with probability 1 as $n \to \infty$.

**Proof**: We note that for any code in the given ensemble of regular codes, all codes which are permutations of the given code (in the sense of permutation of output bits) also belong to the ensemble. Let us assume that we have $r$ 1's in each row of the corresponding parity check equation. Now note that total number of possible combinations of 1's is $\binom{n}{r}$. Whereas, if we restrict our attention to combinations in which no two 1's are at a gap $\leq k$ to each other, the number of combinations are $\geq n \times (n-2k) \times (n-4k) \ldots \times (n-2(r-1)k)/r!$. The ratio of these two values converges to 1 as $n \to \infty$ □

**Lemma 2**: For any given integer $k$, as $n \to \infty$, the gap between $1's$ is greater than $k$ for almost all rows of almost all codes in the ensemble of regular codes of length $n$ with fixed left and right degrees.

**Proof**: The proof is by contradiction. Suppose that for some $k$, at least a fraction $r$ of codes have at least a fraction $\epsilon$ of rows of at least two $1's$ with a gap $\leq k$ for infinitely many $n$. Here, $r$ and $\epsilon$ are independent of $n$. We randomly select one check node equation from one randomly selected code in the given ensemble. The probability that the selected check node equation has at least two $1's$ with gap $\leq k$ is greater than $r \times \epsilon$ for infinitely many $n$. But this is in contradiction with Lemma 1. □
Denote by $H(\zeta_{memless})$ the entropy of $\zeta$ assuming that the channel is memoryless, with the probability of error being average of probability of error in each state of Markov channel. If follows from Lemma 2 that as $n \to \infty$, the dependence between $Y'_{n_i}s$ decreases, and they become independent asymptotically. Therefore $H(\zeta) \to H(\zeta_{memless}) = H(\frac{1+(1-2q)^r}{2})$ for almost all rows of almost all codes in the given ensemble, where $q = \sum_{i=1}^{m} \gamma_i \eta_i$ is the average probability of error in steady state of the Markov chain. The upper bound now follows easily.

**Bound on rate for regular codes**
Note that $H(\mathbf{P}) \leq \sum_{i=1}^{n(1-R)} H(p_i)$. Thus, (7) becomes:

$$H(\mathbf{Y}) \leq nR + n(1-R)H(\frac{1+(1-2q)^r}{2}) \quad (9)$$

Since $\overline{p_r} = \frac{1+(1-2q)^r}{2}$ equation (6) becomes:

$$\begin{aligned}\frac{1}{n}H(\mathbf{X}|\mathbf{Y}) &\geq R - R - (1-R)H(\overline{p_r}) + 1 - C_M \\ &= 1 - C_M - (1-R)H(\overline{p_r})\end{aligned}$$

Now suppose $R = 1 - \frac{1-C_M-\epsilon}{H(\overline{p_r})} > 1 - \frac{1-C_M}{H(\overline{p_r})}$, then it is easy to see that:

$$\frac{1}{n}H(\mathbf{X}|\mathbf{Y}) \geq \epsilon \quad (10)$$

which completes the proof of Theorem 1. □

**Bound on rate for irregular codes**

**Theorem 2:** Under same notation as in Theorem 1, suppose that the irregular code has $\omega_r$ fraction of rows of weight $r$. Then the bound on rate for reliable communication is:

$$R \leq 1 - \frac{1 - C_M}{\sum_r \omega_r H(\overline{p_r})} \quad (11)$$

**Proof:** The proof follows from the observation that under the given conditions, the expression of upper bound on $H(\mathbf{P})$ changes to:

$$H(\mathbf{P}) \leq n(1-R)\sum_r \omega_r H(\overline{p_r}) \qquad \square \qquad (12)$$

**Remark:** Notice that the bounds reduce to the known bounds for BSC [3] if we put $\eta_i = \eta \; \forall i \in \{1, \ldots, m\}$, since, in that case, $C_M = C = 1 - H(\eta)$ and $\overline{p_r} = \frac{1+(1-2\eta)^r}{2}$. Though, the bounds the [3] were valid over all codes in Gallager's construction.

# 4 Lower bounds on parity-check density

Notice that (11) is same as expression of upper bound on rate derived in [1], with the capacity of general MBIOS channels replaced by $C_M$. In [9], lower bounds on parity check density of LDPC codes were derived for MBIOS channels using the same upper bound. Thus, similar lower bounds (as in [9]) on density continue to hold for the Markov channels considered here, with capacity of MBIOS channel replaced by $C_M$.

# A Capacity of Markov channels with channel behaving as a BSC in each state

The derivation is similar to the derivation of channel capacity for GE channels in [7].

$$C_M = \lim_{n\to\infty} \frac{1}{n} \max_{p(\mathbf{X})} I(\mathbf{X};\mathbf{Y}) = 1 - \lim_{n\to\infty} H(\mathbf{z_n}) \qquad (13)$$

The last equality follows because $H(\mathbf{Y})$ achieves its maximum when $\mathbf{X}$ is uniformly distributed, and $H(\mathbf{z_n})$ does not depend on distribution of $\mathbf{X}$. Also, as proved in [5]:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} H(z_i|\mathbf{z_{i-1}}) = \lim_{n\to\infty} H(z_i|\mathbf{z_{i-1}}) \qquad (14)$$

Thus, from the chain rule and (13),

$$C_M = 1 - \lim_{n\to\infty} H(z_i|\mathbf{z_{i-1}}) \qquad \square \qquad (15)$$

# References


[1] D Burshtein, M Krivelevich, S Litsyn, and G Miller. Upper bounds on the rate of LDPC codes. *IEEE Transactions on Information Theory*, 48(9), September 2002.

[2] AW Eckford. Density Evolution: A Powerful Analytical Tool for LDPC Codes. Talk, Deptt of EE, University of Notre Dame, June 2004. www.comm.toronto.edu/ eckford/pubs/ima-de-talk.pdf.

[3] RG Gallager. *Low-Density Parity Check Codes*. PhD thesis, MIT, Cambridge, 1960.

[4] RG Gallager. *Information Theory and Reliable Communication*. Wiley, 1968.

[5] AJ Goldsmith and PP Varaiya. Capacity, mutual information, and coding for finite-state markov channels. *IEEE Transactions on Information Theory*, 42(3), May 1996.

[6] Pulkit Grover and Ajit Kumar Chaturvedi. Upper bounds on rate of LDPC codes for Gilbert-Elliott channels. In *IEEE Information Theory Workshop*, San Antonio, Texas, USA, 24-29 October 2004. http://home.iitk.ac.in/student/pulkit/UpperBoundsGE.pdf.

[7] M Mushkin and I Bar-David. Capacity and coding for the gilbert-elliott channels. *IEEE Transactions on Information Theory*, 35(6), November 1989.

[8] T. J. Richardson and R. L. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Transactions on Information Theory*, 47(2):599–618, Feb 2001.

[9] Igal Sason and Rudiger Urbanke. Parity-check density versus performance of binary linear block codes over memoryless symmetric channels. *IEEE Transactions on Information Theory*, 49(7):1611–1635, July 2003.