Upper Bounds on the Rate of LDPC Codes for Gilbert-Elliott Channels

Pulkit Grover
Indian Institute of Technology Kanpur
pulkit@iitk.ac.in

Ajit Kumar Chaturvedi
Indian Institute of Technology Kanpur
akc@iitk.ac.in

Abstract — **Recently, there has been work in use of LDPC codes over channels with memory, in particular, over Gilbert-Elliott (GE) channels. In this paper, we derive expressions for upper bounds on rate of LDPC codes for reliable communication over a large class (non-oscillatory and non-inverting) of GE channels using the methods for memoryless channels.**

I. 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 zero 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 η 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 generalised by Burshtein et al in [1] for Memoryless Binary Input Output Symmetric (MBIOS) channels and irregular LDPC codes.

We generalize these bounds to a large class of GE channels (see [5]), which have memory. For decoding of LDPC codes used over GE channel, and the associated density evolution technique, we refer the reader to [2].

As a side result, we also extend the results obtained by Sason and Urbanke [7] on density of parity check matrices for MBIOS channels to the setting of GE channels. In V, we show that similar lower bounds (as derived in [7] on density of parity check matrices) for GE channels hold here.

II. Notation and known results

For the GE channel, we denote the "good" ("bad") state by G (B), transition probability from G to B (B to G) by b (g), the corresponding probability of error in the state G (B) by η_G (η_B), where $\eta_G < \eta_B$. If $\eta_G < \eta_B < 0.5$, the channel is said to be *non-inverting*, and if $g + b < 1$, the channel is said to be *non-oscillatory*. Throughout this paper, we assume the GE channel to be non-inverting and non-oscillatory, and this is the class of GE channels on which our results are valid.

Capacity of GE channel was found in [5] to be:

$$C_{GE} = 1 - \lim_{n\to\infty} E[H(q_n)] \tag{1}$$

where $q_n = Pr[z_n = 1|\mathbf{z_{n-1}}]$ and $\mathbf{z_{n-1}}$ denotes the set $\{z_{n-1}, z_{n-2}, \ldots, z_1\}$. Note that q_n is a function of $\mathbf{z_{n-1}}$ and hence a random variable.

III. Upper bound on rate

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

$$\langle P_e\rangle log(M-1) + H(\langle P_e\rangle) \geq \frac{1}{n}H(\mathbf{X}|\mathbf{Y}) \tag{2}$$

where $\langle P_e\rangle$ is the average probability of error for 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, $\frac{1}{n}H(\mathbf{X}|\mathbf{Y}) \to 0$. In what follows, we prove that for use of a sequence $\mathcal{C}_n$ of LDPC codes of blocklength n over a GE channel, $\frac{1}{n}H(\mathbf{X}|\mathbf{Y})$ 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 a GE channel with parameters as defined above. Suppose all rows of $\mathbf{H}$ have a constant weight r. Then a necessary condition for reliable communication is:

$$R \leq 1 - \frac{\lim_{n\to\infty} E[H(q_n)]}{H(\overline{p_r})} \tag{3}$$

where

$$\overline{p_r} = \frac{1 + ((1-2\eta_G)\gamma_G + (1-2\eta_B)\gamma_B)^r}{2} \tag{4}$$

and q_n is as defined in equation 1.

Proof: 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,

$$\frac{1}{n}H(\mathbf{X}|\mathbf{Y}) = \frac{1}{n}H(\mathbf{X}) - \frac{1}{n}H(\mathbf{Y}) + \frac{1}{n}H(\mathbf{Y}|\mathbf{X}) \tag{5}$$

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 $\{z_1, \ldots z_i\}$. As shown in [5],

$$H(z_k|\mathbf{z_{k-1}}) = E[H(q_k)] \tag{6}$$

where $q_k = Pr(z_k = 1|\mathbf{z_{k-1}})$. Also, it is shown in Proposition 3 of [5] that the sequence $\{E[H(q_k)]\}_{k=0}^{\infty}$ is monotonically decreasing in k and therefore, $\frac{1}{n}H(\mathbf{Z}) = \frac{1}{n}\sum_{k=1}^{n} E[H(q_k)] \geq \lim_{n\to\infty} E[H(q_n)]$. Thus, (5) becomes:

$$\frac{1}{n}H(\mathbf{X}|\mathbf{Y}) \geq R - \frac{1}{n}H(\mathbf{Y}) + \lim_{n\to\infty} E[H(q_n)] \tag{7}$$

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,

$$H(\mathbf{Y}) = H(\mathbf{Y_1};\mathbf{P}) = H(\mathbf{Y_1}) + H(\mathbf{P}|\mathbf{Y_1}) \leq H(\mathbf{Y_1}) + H(\mathbf{P}) \tag{8}$$

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.

[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

Bounding entropy of individual parity checks

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

$$P(\zeta = 0) = P\Big(\text{even number of errors in} \{Y_{n_i}\}_{i=1}^{r}\Big) \tag{9}$$

Since the state space has Markov distribution, determining exact probability is not possible without knowing the exact positions of 1's. So we try to obtain a bound on the probability $P(\zeta = 0)$.

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.

In Appendix B, we prove that for the GE channels considered, $P(\zeta = 0)$ decreases with increase in gap between any two 1's (keeping the gap between other 1's constant). Also, in an expression of $P(\zeta = 0)$ in Appendix A, it can be seen that for non-inverting GE channels, $P(\zeta = 0) > 0.5$, and hence, the entropy $H(\zeta)$ increases with decrease in $P(\zeta = 0)$. If we keep increasing the gap between all the 1's, in the limit, Y_{n_i} become independent and state distribution assumes stationary probabilities at each time instant. Thus, the entropy $H(\zeta)$ can be upper bounded by $H(\zeta_{memless})$, where $\zeta_{memless}$ is the random variable representing result of a parity check equation for a memoryless channel with error probability the average error probability for the channel with memory under consideration.

Thus, $H(\zeta) \leq H(\zeta_{memless}) = H(\frac{1+(1-2q)^r}{2})$, where $q = \gamma_G\eta_G + \gamma_B\eta_B$ is the average probability of error in steady state of the Markov chain. This analysis obtains slightly looser bound since we do not exploit correlations between different errors for calculation of bound on entropy.

Bound on rate for regular codes

Note that $H(\mathbf{P}) \leq \sum_{i=1}^{n(1-R)} H(p_i)$. Thus, (8) becomes:

$$H(\mathbf{Y}) \leq nR + n(1-R)H(\frac{1+(1-2q)^r}{2}) \tag{10}$$

Let $\overline{p_r} \triangleq \frac{1+(1-2q)^r}{2}$. Hence, (7) becomes:

$$\frac{1}{n}H(\mathbf{X}|\mathbf{Y}) \geq R - R - (1-R)H(\overline{p_r}) + \lim_{n\to\infty} E[H(q_n)]$$
$$= \lim_{n\to\infty} E[H(q_n)] - (1-R)H(\overline{p_r})$$

Now suppose $R = 1 - \frac{\lim_{n\to\infty} E[H(q_n)]-\epsilon}{H(\overline{p_r})} > 1 - \frac{\lim_{n\to\infty} E[H(q_n)]}{H(\overline{p_r})}$, then it is easy to see that:

$$\frac{1}{n}H(\mathbf{X}|\mathbf{Y}) \geq \epsilon \tag{11}$$

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 ω_r fraction of rows of weight r. Then the bound on rate for reliable communication is:

$$R \leq 1 - \frac{\lim_{n\to\infty} E[H(q_n)]}{\sum_r \omega_r H(\overline{p_r})} \tag{12}$$

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

$$H(\mathbf{P}) \leq n(1-R)\sum_r \omega_r H(\overline{p_r}) \tag{13}$$

Remark: Notice that the bounds reduce to the known bounds for BSC [3] if we put $\eta_G = \eta_B = \eta$. Since, in that case, $q_n = Pr(z_n = 1) = \eta$ and $\overline{p_r} = P_r$, where $P_r = \frac{1+(1-2\eta)^r}{2}$.

IV. Tightening the bound

In this section, we tighten the upper bound on entropy of a parity check, which can be used to tighten the bounds in (3) and (12).

Suppose in a row of a parity check matrix, the *maximum* gap between any two 1's is v, that is, any two variables in that parity check equation are separated by a gap of no more than v. Since entropy of the parity check increases as gap between 1's is increased (as concluded in the Appendices), the entropy of the given parity check will be lesser than the entropy of a parity check for which the gap between 1's is uniformly v. In the latter case, the Markov chain relating Y_{n_r}'s is homogeneous, albeit the transition probabilities have changed.

It was proved by Pedler [6] that for a homogeneous two state Markov chain, probability of visiting a particular state (say G) k times in n transitions is given by:

For $0 < k < n$
$P(N_G = k) = (1-b)^k(1-g)^{n-k}F[-n+k,-k;1;\lambda]$
$-\pi_G d(1-b)^k(1-g)^{n-k-1}F[-n+k+1,-k;1;\lambda]$
$-\pi_B d(1-b)^{k-1}(1-g)^{n-k}F[-n+k,-k+1;1;\lambda]$
and
$P(N_G = 0) = (\pi_G b + \pi_B(1-g))(1-g)^{n-1}$
$P(N_G = n) = (\pi_G(1-b) + \pi_B g)(1-b)^{n-1}$

where N_G is the random variable denoting number of times state G is visited, F is the hypergeometric function, π_G and π_B are the steady state probabilities of G and B respectively, $\lambda = \frac{gb}{(1-g)(1-b)}$, and $d = (1-g)(1-b) - gb$. Now, as shown in Appendix A, the entropy of a parity check is given by $H(\sum_{i=0}^{r} Y_{n_i})$, and
$P(\sum_{i=0}^{r} Y_{n_i} = 0) =$
$\frac{1}{2} + \frac{1}{2}\sum_{k=0}^{r}(1-2\eta_G)^k(1-2\eta_B)^{r-k}P(N_G = k)$
For uniform gap v between 1's, the underlying state space for $Y'_{n_i}s$ is homogeneous Markov, and hence Pedler's result is applicable. Using v-step transition probabilities, associated entropy can be calculated for different values of v, and Fig. 1 shows how this variation takes place with v for typical values of r, η_G, η_B, g and b. As v increases, we see that the probability converges to the independent case, as expected. Hence if v is known for each parity check, a tighter bound on entropy $H(\zeta = 0)$ can be obtained, and thus the bounds (3) and (12) can be tightened.

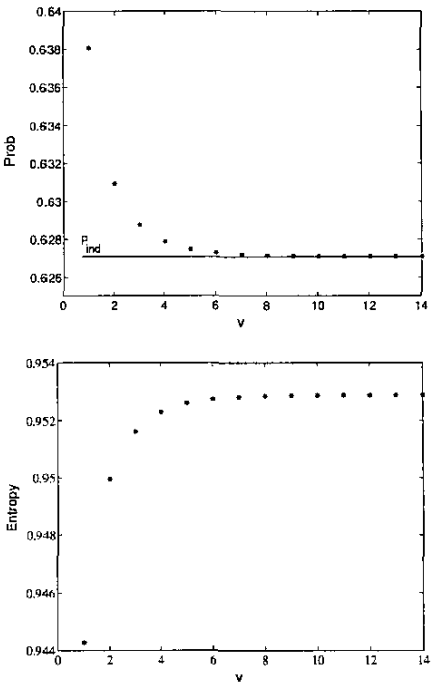

Figure 1: $P(\zeta = 0)$ and $H(\zeta)$ versus v for $n = 10$, $\eta_G = 0.01$, $\eta_B = 0.1$, $g = 0.2$, $b = 0.3$

V. Lower bounds on parity-check density

Notice that (12) can be written as:

$$R \leq 1 - \frac{1 - C_{GE}}{\sum_r \omega_r H(\overline{p_r})} \tag{14}$$

which is same as expression of upper bound on rate derived in [1], with the capacity of general MBIOS channel replaced by C_{GE}. In [7], 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 [7]) on density continue to hold here, with capacity of MBIOS channel replaced by C_{GE}. Again, these bounds can be tightened using results of section IV.

APPENDIX

A. PROBABILITY OF A PARITY CHECK EQUATION BEING SATISFIED

Suppose we have some sequence $\{X_{n_i}\}_{i\geq 1}$ as input to a two state channel which behaves as a BSC in each state. Let $\{Y_{n_i}\}_{i\geq 1}$ be the received sequence, and we want to calculate $P(\sum_{i=0}^{r} Y_{n_i} = 0)$, where the addition is modulo two. Note that here we do not assume Markov modeling of the state space. (In Appendix B, we have a non-homogeneous Markov chain, and we use this result).We prove the following result:

$$P(\sum_{i=0}^{r} Y_{n_i} = 0) = \frac{1}{2}+\frac{1}{2}\sum_{k=0}^{r}(1-2\eta_G)^k(1-2\eta_B)^{r-k}P(N_G = k) \tag{15}$$

where $P(N_G = k)$ is the probability of making k visits to the state G in the given r time instants.

Remarks:

- Note that in equation (15), the probability is greater than 0.5 as long as η_G and η_B are less than 0.5 (ensuring that the second term remains positive). We have used this fact frequently in the paper.
- In the proof below, we use the fact from [3] that probability of even errors in r channel uses of a BSC is given by $\frac{1+(1-2\epsilon)^r}{2}$, where ϵ is the crossover probability of the BSC.

Proof:

X_{n_i} are such that they satisfy the parity check equation. With the noise added due to the channel, they may be some errors. A parity check equation is satisfied as long as the number of errors in the check variables are even.

$P(\sum_{i=0}^{r} Y_{n_i} = 0) = P(\text{even number of errors in } \{Y_{n_i}\}_{i=1}^{r})$
$= \sum_{k=1}^{r} P(\text{even number of errors in } \{k \text{ good states and } r-k \text{ bad states}\} | N_G = k) \times P(N_G = k)$
$= \sum_{k=1}^{r} [P(\text{even number of errors in } k \text{ good states;even number of errors in } r-k \text{ bad states}) +$
$P(\text{odd number of errors in } k \text{ good states;odd number of errors in } r-k \text{ bad states})] \times P(N_G = k)$

$$= \sum_{k=1}^{r}\left[\left(\frac{1+(1-2\eta_G)^k}{2}\right)\left(\frac{1+(1-2\eta_G)^{r-k}}{2}\right)+\left(\frac{1-(1-2\eta_G)^k}{2}\right)\left(\frac{1-(1-2\eta_G)^{r-k}}{2}\right)\right]\times P(N_G = k)$$

Which on simplification gives:

$$\sum_{k=1}^{r}\left(\frac{1}{2}+\frac{(1-2\eta_G)^k(1-2\eta_B)^{r-k}}{2}\right)P(N_G = k) =$$

$$\frac{1}{2}+\sum_{k=1}^{r}\frac{(1-2\eta_G)^k(1-2\eta_B)^{r-k}}{2}\times P(N_G = k), \text{ as claimed.}$$

If the states s_i are i.i.d., the distribution of N_G is binomial, and simple manipulations show that the above expression reduces to $\frac{1}{2}+\frac{(1-2q)^r}{2}$, where $q = \gamma_G\eta_G + \gamma_B\eta_B$ is the average probability of error.

B. BOUND ON PROBABILITY OF EVEN NUMBER OF ERRORS IN A PARITY CHECK EQUATION

First we prove that

$$P(\sum_{i=1}^{r} Y_{n_i} = 0 | s_{n_r} = G) > P(\sum_{i=1}^{r} Y_{n_i} = 0 | s_{n_r} = B) \tag{16}$$

The proof is by induction on r. The result is trivially true for $r = 1$ (since $1-\eta_G > 1-\eta_B$).

Now, we prove the result (16) for $r = k$ assuming result is true for $r = k-1$. Let $t = n_k - n_{k-1}$ denote the gap between the k^{th} and $(k-1)^{th}$ 1's.

Using $t-$step transition probability matrix for a two state Markov chain, define b_t and g_t as:

$$P^t = \begin{bmatrix} \frac{g+b(1-g-b)^t}{g+b} & \frac{b-b(1-g-b)^t}{g+b} \\ \frac{g-g(1-g-b)^t}{g+b} & \frac{b+g(1-g-b)^t}{g+b} \end{bmatrix} =: \begin{bmatrix} 1-b_t & b_t \\ g_t & 1-g_t \end{bmatrix} \tag{17}$$

Where P is the single step transition probability matrix. Notice that given s_{n_k}, Y_{n_k} is independent of Y_{n_i} (for $i \neq k$). Thus,

$$\begin{aligned} &P(\sum_{i=1}^{k} Y_{n_i} = 0 | s_{n_k} = G) \\ &= \Big[P(\sum_{i=1}^{k-1} Y_{n_i} = 0 | s_{n_{k-1}} = G)(1-b_t) \\ &+P(\sum_{i=1}^{k-1} Y_{n_i} = 0 | s_{n_{k-1}} = B)b_t\Big](1-\eta_G) \\ &+\Big[P(\sum_{i=1}^{k-1} Y_{n_i} = 1 | s_{n_{k-1}} = G)(1-b_t)+ \\ &P(\sum_{i=1}^{k-1} Y_{n_i} = 1 | s_{n_{k-1}} = B)b_t\Big]\eta_G \end{aligned} \tag{18}$$

where, we note that the first term in square brackets is the conditional probability of the event $\sum_{i=1}^{k-1} Y_{n_i} = 0$ given $s_{n_k} = G$. Similarly, for conditioning on B we get:

$$\begin{aligned} &P(\sum_{i=1}^{k} Y_{n_i} = 0 | s_{n_k} = B) \\ &= \Big[P(\sum_{i=1}^{k-1} Y_{n_i} = 0 | s_{n_{k-1}} = B)(1-g_t) \\ &+P(\sum_{i=1}^{k-1} Y_{n_i} = 0 | s_{n_{k-1}} = G)g_t\Big](1-\eta_B) \\ &+\Big[P(\sum_{i=1}^{k-1} Y_{n_i} = 1 | s_{n_{k-1}} = B)(1-g_t) \\ &+P(\sum_{i=1}^{k-1} Y_{n_i} = 1 | s_{n_{k-1}} = G)g_t\Big]\eta_B \end{aligned} \tag{19}$$

Now observe the first terms in square brackets in (18) and (19). Both the terms are of the form $f(a) = P(\sum_{i=1}^{k-1} Y_{n_i} = 0|s_{n_{k-1}} = G)a + P(\sum_{i=1}^{k-1} Y_{n_i} = 0|s_{n_{k-1}} = B)(1-a)$, where $a = 1 - b_t$ in (18) and $a = g_t$ in (19). Now, $f(a) = a(P(\sum_{i=1}^{k-1} Y_{n_i} = 0|s_{n_{k-1}} = G) - P(\sum_{i=1}^{k-1} Y_{n_i} = 0|s_{n_{k-1}} = B)) + (1 - P(\sum_{i=1}^{k-1} Y_{n_i} = 0|s_{n_{k-1}} = B))$, which is an increasing function of a, since $P(\sum_{i=1}^{k-1} Y_{n_i} = 0|s_{n_{k-1}} = G) > P(\sum_{i=1}^{k-1} Y_{n_i} = 0|s_{n_{k-1}} = B)$. We want to conclude that term under consideration is greater in (18). From (17), it can easily be derived that $1 - g_t - b_t > 0$ for any non-oscillatory GE channel, and therefore, term a is greater in (18) and hence the first term in square brackets (which is under consideration here) is also greater in (18). Also it can be seen that in both (18) and (19), this term is greater than 0.5 (This follows since the channel is non-inverting, see Appendix A).

Now, the probability expressions in (18) and (19) can be written as $h(x, y) = xy + (1-x)(1-y)$, where, x takes the value of first term in the square brackets in each expression. Now, $h(x, y) = x(2y-1) - y + 1$, which is an increasing function of x if $y > 0.5$ and increasing function of y if $x > 0.5$. In our case, both $x > 0.5$ and $y > 0.5$, thus $h(x, y)$ is an increasing function of both x and y. It can also be seen that both x and y are larger in (18), hence $P(\sum_{i=1}^{r} Y_{n_i} = 0|s_{n_r} = G) > P(\sum_{i=1}^{r} Y_{n_i} = 0|s_{n_r} = B)$.

Increase in number of 0's between 1's increases entropy

Let $\zeta = \sum_{i=1}^{r} Y_{n_i} = \zeta_1 + \zeta_2$, where $\zeta_1 = \sum_{i=1}^{r_1} Y_{n_i}$ and $\zeta_2 = \sum_{i=r_1+1}^{r} Y_{n_i}$. We prove that if $d = n_{r_1+1} - n_{r_1}$ increases, corresponding entropy $H(\zeta)$ also increases. To prove this, it is sufficient to prove that $P(\zeta_1 + \zeta_2 = 0)$ decreases as d increases (since $P(\zeta_1 + \zeta_2 = 0) > 0.5$).

$$P(\zeta_1 + \zeta_2 = 0) = P(\zeta_1 = 0; \zeta_2 = 0) + P(\zeta_1 = 1; \zeta_2 = 1) \tag{20}$$

Note that

$$\begin{aligned} P(\zeta_1 = 1; \zeta_2 = 1) &= P(\zeta_2 = 1) - P(\zeta_1 = 0; \zeta_2 = 1) \\ &= P(\zeta_2 = 1) - P(\zeta_1 = 0) + P(\zeta_1 = 0; \zeta_2 = 0) \end{aligned} \tag{21}$$

Since the terms $P(\zeta_2 = 1)$ and $P(\zeta_1 = 0)$ are independent of d, to prove $P(\zeta = 0)$ decreases as d increases, it is sufficient to prove (from (20) and (21)) that $P(\zeta_1 = 0; \zeta_2 = 0)$ decreases as d increases.

We now prove that $P(\zeta_1 = 0|\zeta_2 = 0)$ decreases as d increases (which suffices, as $P(\zeta_2 = 0)$ is independent of d).

$$\begin{aligned} &P(\zeta_1 = 0|\zeta_2 = 0) \\ &= P(\zeta_1 = 0|s_{n_{r_1+1}} = G)P(s_{n_{r_1+1}} = G|\zeta_2 = 0) \\ &\quad + P(\zeta_1 = 0|s_{n_{r_1+1}} = B)P(s_{n_{r_1+1}} = B|\zeta_2 = 0) \\ &= \Big[P(\zeta_1 = 0|s_{n_{r_1}} = G)P(s_{n_{r_1}} = G|s_{n_{r_1+1}} = G) + \\ &\quad P(\zeta_1 = 0|s_{n_{r_1}} = B)P(s_{n_{r_1}} = B|s_{n_{r_1+1}} = G)\Big] \\ &\quad \times P(s_{n_{r_1+1}} = G|\zeta_2 = 0) \\ &\quad + \Big[P(\zeta_1 = 0|s_{n_{r_1}} = G)P(s_{n_{r_1}} = G|s_{n_{r_1+1}} = B) + \\ &\quad P(\zeta_1 = 0|s_{n_{r_1}} = B)P(s_{n_{r_1}} = B|s_{n_{r_1+1}} = B)\Big] \\ &\quad \times P(s_{n_{r_1+1}} = B|\zeta_2 = 0) \end{aligned}$$

Thus,

$$\begin{aligned} &P(\zeta_1 = 0|\zeta_2 = 0) \\ &= \Big[P(\zeta_1 = 0|s_{n_{r_1}} = G) \times \frac{g + b(1-g-b)^d}{g+b} \\ &\quad + P(\zeta_1 = 0|s_{n_{r_1}} = B) \times \frac{b - b(1-g-b)^d}{g+b}\Big] \\ &\quad \times P(s_{n_{r_1+1}} = G|\zeta_2 = 0) \\ &\quad + \Big[P(\zeta_1 = 0|s_{n_{r_1}} = G) \times \frac{g - g(1-g-b)^d}{g+b} \\ &\quad + P(\zeta_1 = 0|s_{n_{r_1}} = B) \times \frac{b + g(1-g-b)^d}{g+b}\Big] \\ &\quad \times P(s_{n_{r_1+1}} = B|\zeta_2 = 0) \\ &= C + \frac{(1-g-b)^d}{g+b} \\ &\quad \times \Big[P(\zeta_1 = 0|s_{n_{r_1}} = G) - P(\zeta_1 = 0|s_{n_{r_1}} = B)\Big] \\ &\quad \times \Big[bP(s_{n_{r_1+1}} = G|\zeta_2 = 0) - gP(s_{n_{r_1+1}} = B|\zeta_2 = 0)\Big] \end{aligned}$$

where C is a constant independent of d. It is easy to see that the two terms in product with $\frac{(1-g-b)^d}{g+b}$ are positive, and since channel is non-oscillatory, $(1-g-b) > 0$, so $P(\zeta_1 = 0|\zeta_2 = 0)$ decreases with increase in d. (To see that the second term in product with $\frac{(1-g-b)^d}{g+b}$ is positive, notice that $P(s_{n_{r_1+1}} = G|\zeta_2 = 0) = \frac{P(\zeta_2=0|s_{n_{r_1+1}}=G)}{P(\zeta_2=0)} \times \frac{g}{g+b}$, and a similar expression holds for $P(s_{n_{r_1+1}} = G|\zeta_2 = 0)$. Now use (16)) Hence $P(\zeta = 0)$ decreases with increase in d. Since $P(\zeta = 0) > 0.5$, the entropy $H(\zeta)$ increases as d increases. To find an upper bound on entropy, we can thus consider $\lim_{d\to\infty} H(\zeta)$, in which the two variables ζ_1 and ζ_2 become independent. Continuing the process inductively, maximum entropy will be obtained when all $\{Y_{n_i}\}$ are distributed independently.

ACKNOWLEDGMENTS

The authors would like to thank David J Aldous for pointing to [6].

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, FR Kschischang, and S Pasupathy. Analysis of Low-Density Parity-Check decoding over the Gilbert-Elliott channel. *submitted to IEEE Transactions on Information Theory*, 2003.

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

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

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

[6] PJ Pedler. Occupation times for two state markov chains. *Journal of Applied Probability*, pages 381–390, 1971.

[7] 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.