Upper Bounds on the Rate of LDPC Codes for a Class of Finite-State Markov Channels

Pulkit Grover, *Student Member, IEEE*, and
Ajit Kumar Chaturvedi, *Senior Member, IEEE*

***Abstract*—In this correspondence, we consider the class of finite-state Markov channels (FSMCs) in which the channel behaves as a binary symmetric channel (BSC) in each state. Upper bounds on the rate of LDPC codes for reliable communication over this class of FSMCs are found. A simple upper bound for all noninverting FSMCs is first derived. Subsequently, tighter bounds are derived for the special case of Gilbert–Elliott (GE) channels. Tighter bounds are also derived over the class of FSMCs considered. The latter bounds hold *almost-surely* for any sequence of *randomly constructed* LDPC codes of given degree distributions. Since the bounds are derived for optimal maximum-likelihood decoding, they also hold for belief propagation decoding. Using the derivations of the bounds on the rate, some lower bounds on the density of parity check matrices for given performance over FSMCs are derived.**

***Index Terms*—Belief-propagation decoding, density, error probability, finite state Markov channels (FSMCs), Gilbert–Elliott (GE) channels, low-density parity-check (LDPC) codes.**

I. Introduction

In this correspondence, we say that a sequence of codes can be used for reliable communication over a given channel if the optimal (maximum-likelihood) decoding error probability of the sequence converges to zero as the block length approaches infinity [1]. In [2], Gallager derived an upper bound on the rate of regular LDPC codes for reliable communication over a binary symmetric channel (BSC). The bound was found to be $R \leq 1 - \frac{H(\eta)}{H(P_d)}$, where η is the crossover probability of the BSC, d is row weight of each row of the parity check matrix, $P_d = \frac{1+(1-2\eta)^d}{2}$, and $H(\cdot)$ is the binary entropy function. The bound was generalized by Burshtein *et al.* in [1] to memoryless binary input–output symmetric (MBIOS) channels and irregular LDPC codes.

LDPC codes have been shown to have good performance for the low-complexity iterative decoding over the memoryless channels. Recently, their performance has been analyzed over some channels with memory, for example over Gilbert–Elliott (GE) channels [3], finite-state Markov channels (FSMCs) [4] and intersymbol interference (ISI) channels [5] and found to be encouraging. It thus becomes imperative to understand the limits of LDPC codes when used over channels with memory.

In this correspondence, we generalize the bound in [2] to some channels with memory. We consider the class of FSMCs in which the channel behaves as a BSC in each state. In [6], we generalized this bound to a class of Gilbert–Elliott (GE) channels, which are two-state Markov channels with the channel behaving as a BSC in each state. Here we first derive a simple bound on the rate for all noninverting FSMCs. For FSMCs, and for GE channels in particular, we then derive tighter bounds. While the tight bounds derived here for GE channels always hold, the tight bounds for FSMCs hold only *almost-surely* for *randomly constructed* LDPC codes of given degree distributions. In the random construction, a code is selected randomly from an ensemble of codes of given degree distributions. Performance of randomly constructed LDPC codes is easy to analyze for large block lengths using the density evolution technique [7].

The derived bounds prove that for a sequence of LDPC codes to be capacity achieving, the average number of 1's in each row of the parity check matrix must approach infinity. Furthermore, these bounds lead to lower bounds on the required *density* [8] of a parity check matrix, given the code performance. Thus, we extend the results obtained by Sason and Urbanke [8] on the density of parity check matrices for MBIOS channels to the setting of FSMCs.

The rest of the correspondence is organized as follows. In Section II, we introduce some definitions and set the notation for the sequel. In Section III, we reduce the problem of finding the upper bounds on the rate to finding bounds on the entropy of syndromes. In Section IV, we derive upper bounds on the rate for GE channels as well as for FSMCs. In Section V we derive lower bounds on the density of parity check matrices of codes of given performance over FSMCs. We conclude in Section VI.

II. Notations and Definitions

A vector of length n is denoted in bold letters with superscript n. For example, the channel input vector is denoted by $\mathbf{X}^n$, and the channel output vector by $\mathbf{Y}^n$. The k^{th} element of $\mathbf{X}^n$ is denoted by X_k. Also, $lg(\cdot)$ denotes logarithm to the base 2, and $ln(\cdot)$ denotes logarithm to the base e.

We assume that the Markov chain has a unique steady state distribution, for which we require the Markov chain to be irreducible and aperiodic [9]. Throughout the correspondence, by an FSMC we mean a FSMC which behaves as a BSC in each state. C_m denotes the capacity of an FSMC, where m is the number of states in the FSMC, each of which are denoted by integers $s = 1, 2, \ldots, m$. η_i and γ_i denote the crossover probability and the steady-state probability, respectively, of the i^{th} state. The FSMC is said to be *noninverting* if $\eta_i < 0.5$ for all $i \in \{1, 2, \ldots, m\}$.

For the GE channel, the states are called "good" and "bad." We denote the "good" (respectively, "bad") state by G (resp. B), transition probability from G to B (respectively, B to G) by b (respectively, g), the corresponding crossover probability in the state G (respectively, B) by η_G (resp. η_B), where $\eta_G < \eta_B$. If $\eta_G < \eta_B < 0.5$, the channel is said to be *noninverting*, and if $g + b < 1$, the channel is said to be *nonoscillatory*. We assume that the GE channel is noninverting and nonoscillatory. The steady state probability of G (resp. B) is denoted by γ_G (resp. γ_B). Also, C_{GE} denotes the capacity of a GE channel.

For regular LDPC codes, the variable node degree is denoted by c and the check node degree is denoted by d. For irregular codes, the fraction of rows of weight d is denoted by ω_d. The left and right edge degree distributions are denoted by $\lambda(x)$ and $\rho(x)$ respectively, in accordance with the convention. That is, if $\lambda(x) = \sum_{i \geq 2} \lambda_i x^{i-1}$, then λ_i is the fraction of edges of left degree i. $\rho(x)$ is similarly defined. Note that $\rho(x)$ and row weight distribution are related by $\omega_d = \frac{\frac{\rho_d}{d}}{\sum_i \frac{\rho_i}{i}}$.

We denote a parity check matrix of a code of rate R and blocklength n by $\mathbf{H}$. The design rate of the code is denoted by R_D. Thus $R_D = 1 - \frac{c}{d}$ for regular codes and $R_D = 1 - \frac{\int_0^1 \rho(x)dx}{\int_0^1 \lambda(x)dx}$ for irregular codes. Since all the rows of $\mathbf{H}$ need not be linearly independent, $R \geq R_D$. The syndrome vector, which comprises of the results of the $n(1-R)$ parity check equations when applied on $\mathbf{Y}^n$, is denoted by $\mathbf{S}^{n(1-R)}$. The i^{th} syndrome is denoted by S_i.

Manuscript received February 24, 2005; revised April 10, 2006. The material in this correspondence was presented in part at the 2004 IEEE Information Theory Workshop, San Antonio, TX. This work was done at the Department of Electrical Engineering, IIT Kanpur, India.

P. Grover is with the Department of Electrical Engineering and Computer Science, University of California at Berkeley, Berkeley, CA 94703 USA (e-mail: pulkit@eecs.berkeley.edu).

A. K. Chaturvedi is with the Department of Electrical Engineering, Indian Institute of Technology, Kanpur 208016, India (e-mail: akc@iitk.ac.in).

Communicated by A. Ashikhmin, Associate Editor for Coding Theory.

Color versions of Fig. 1 is available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2006.887511

$C^n(c,d)$ denotes the ensemble of regular LDPC codes of length n and left and right degrees c and d respectively, generated by "socket construction" of [7]. The generalization of this construction to irregular codes is straightforward, and as per the convention, the corresponding ensemble is denoted by $C^n(\lambda,\rho)$. The average variable node degree is denoted by α, and the average check node degree is denoted by β.

Definition 1 (Density): For a binary linear code $\mathcal{C}$ with a parity check matrix $\mathbf{H}$, the density of $\mathbf{H}$, denoted by $\Delta(\mathbf{H})$, is defined as the ratio of the total number of 1's in a parity-check matrix to the code dimension (see [8]). That is, $\Delta(H) := \frac{\beta(1-R)}{R}$.

The design density, Δ_D, is defined in an analogous manner to the design rate R_D. Note that $\Delta_D \leq \Delta$.

Definition 2 (Gap): In a row of the parity check matrix $\mathbf{H}$, suppose two consecutive 1's, which are separated by a string of 0's, occur at locations n_1 and n_2. Then the *gap* between the two 1's is defined as $|n_2 - n_1|$.

Definition 3 (Random Construction of LDPC Codes): In the random construction of LDPC codes, a code is chosen randomly from $C^n(\lambda,\rho)$ with uniform probability.

Note that this random construction induces a uniform probability distribution on codes of given (λ,ρ) for a given length. This further induces a probability distribution on the product space consisting of sequences of codes of different lengths for given (λ,ρ).

III. Upper Bounds on the Rate: Reducing the Problem

Consider a linear code $\mathcal{C}_n$ of blocklength n and parity check matrix $\mathbf{H}$. If the input alphabet χ is of size M, and the input and output are denoted by $\mathbf{X}^n = \{X_1, X_2, \ldots, X_n\}$, where $X_i \in \chi$, and $\mathbf{Y}^n = \{Y_1, Y_2, \ldots, Y_n\}$, then we have the following (Fano's inequality, [10]):

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

where $\langle P_e\rangle$ is the average probability of symbol error (averaged over all symbols in the block).[1] For binary case, $M = 2$, so the equation reduces to

$$H(\langle P_e\rangle) \geq \frac{1}{n} H(\mathbf{X}^n|\mathbf{Y}^n). \tag{2}$$

Thus, if $\frac{1}{n}H(\mathbf{X}^n|\mathbf{Y}^n)$ is strictly positive, then so is $\langle P_e\rangle$, the average probability of bit error. It follows that for reliable communication $\frac{1}{n}H(\mathbf{X}^n|\mathbf{Y}^n) \to 0$. In the sequel, we prove that for use of a sequence $\{\mathcal{C}_n\}$ of LDPC codes over an FSMC, $\frac{1}{n}H(\mathbf{X}^n|\mathbf{Y}^n)$ is lower bounded by a positive constant for a rate exceeding a certain bound. For ease of exposition, we first derive the bound for regular codes.

Suppose that $\mathcal{C}_n$ is transmitted over an m-state FSMC. Suppose all rows of parity check matrices of $\mathcal{C}_n$ have a constant weight d. We derive bounds on the rate for such a sequence of codes for reliable communication. We break the derivation into two parts. In the first part, contained in this section, we reduce the problem to finding an upper bound on the entropy of the syndrome vector $\mathbf{S}^{n(1-R)}$. In the second part, which is in Section IV, we find this upper bound, and hence the bounds on the rate. The derivation in the first part is similar to that in [2]. New techniques are, however, needed in the derivation of the second part.

[1]Conventionally, in the literature of LDPC codes, the analysis is done on the bit error probability of coded symbols. For the problems addressed here, the analysis for the bit error probability of the source symbols can be performed in a similar way, and the bounds on the rate derived would be the same.

A. Reducing the Problem to Bounding the Entropy of Individual Syndromes

Using two different expressions for $I(\mathbf{X}^n;\mathbf{Y}^n)$ ([10])

$$\frac{1}{n}H(\mathbf{X}^n|\mathbf{Y}^n) = \frac{1}{n}H(\mathbf{X}^n) - \frac{1}{n}H(\mathbf{Y}^n) + \frac{1}{n}H(\mathbf{Y}^n|\mathbf{X}^n). \tag{3}$$

For any code, $\mathbf{X}^n$ is the encoded data which is in one-one mapping with the information symbols. The information symbols are the uniformly distributed over the 2^{nR} values. Thus, $\mathbf{X}^n$ takes any value from the 2^{nR} codewords with uniform probability distribution. Therefore

$$H(\mathbf{X}^n) = nR. \tag{4}$$

Denoting by $\mathbf{Z}^n := \mathbf{Y}^n + \mathbf{X}^n$ the binary error vector, observe that

$$H(\mathbf{Y}^n|\mathbf{X}^n) = H(\mathbf{Y}^n + \mathbf{X}^n|\mathbf{X}^n) = H(\mathbf{Z}^n|\mathbf{X}^n) = H(\mathbf{Z}^n). \tag{5}$$

The last equality follows from the fact that for an FSMC, the errors are independent of the input sequence.

Also, it can be inferred from [11, Th. 4.2.2] that the sequence $\{H(Z_i|\mathbf{Z}^{i-1})\}_{i=1}^{\infty}$ is monotonically decreasing in i and therefore, $\frac{1}{n}H(\mathbf{Z}^n) \geq \lim_{i\to\infty} H(Z_i|\mathbf{Z}^{i-1})$. Thus

$$\frac{1}{n}H(\mathbf{Y}^n|\mathbf{X}^n) \geq \lim_{i\to\infty} H(Z_i|\mathbf{Z}^{i-1}). \tag{6}$$

We need the following Lemma in the sequel.

Lemma 1: The capacity of an FSMC (as defined above) is given by

$$C_m = 1 - \lim_{i\to\infty} H(Z_i|\mathbf{Z}^{i-1}). \tag{7}$$

Proof: See Appendix I. □

Using (4), (6) and Lemma 1, we get

$$\frac{1}{n}H(\mathbf{X}^n|\mathbf{Y}^n) \geq R - \frac{1}{n}H(\mathbf{Y}^n) + 1 - C_m. \tag{8}$$

Since we want to lower bound $\frac{1}{n}H(\mathbf{X}^n|\mathbf{Y}^n)$, we now find an upper bound on $\frac{1}{n}H(\mathbf{Y}^n)$. As shown in [1], the received vector $\mathbf{Y}^n$ uniquely determines $\mathbf{Y_1}^{nR}$, which are the received bits at any nR linearly independent locations in the code, and $\mathbf{S}^{n(1-R)}$, the syndrome vector. Also, $\mathbf{Y_1}^{nR}$ and $\mathbf{S}^{n(1-R)}$ uniquely determine $\mathbf{Y}^n$. That is, there is a one-one mapping between the two sets.[2] Therefore,

$$\begin{aligned} H(\mathbf{Y}^n) &= H(\mathbf{Y_1}^{nR}; \mathbf{S}^{n(1-R)}) \\ &= H(\mathbf{Y_1}^{nR}) + H(\mathbf{S}^{n(1-R)}|\mathbf{Y_1}^{nR}) \\ &\leq H(\mathbf{Y_1}^{nR}) + H(\mathbf{S}^{n(1-R)}) \end{aligned} \tag{9}$$

where the last inequality follows from the fact that conditioning reduces entropy.[3] Now $\mathbf{Y_1}^{nR} = \mathbf{X_1}^{nR} + \mathbf{Z_1}^{nR}$, where $\mathbf{X_1}^{nR}$ and $\mathbf{Z_1}^{nR}$ are the vectors corresponding to characters at independent locations in the transmitted codeword and the error vector respectively. Since $\mathbf{X_1}^{nR}$ is the vector corresponding to nR independent positions in the transmitted word, it specifies a codeword uniquely, and hence the distribution of $\mathbf{X_1}^{nR}$ is uniform over its possible 2^{nR} values.

[2]Given $\mathbf{Y}^n$, both $\mathbf{Y_1}^{nR}$ and $\mathbf{S}^{n(1-R)}$ can clearly be obtained. Given $\mathbf{Y_1}^{nR}$ and $\mathbf{S}^{n(1-R)}$, finding $\mathbf{Y}^n$ is same as finding values of $\mathbf{Y}^n$ on the $n(1-R)$ locations other than those corresponding to $\mathbf{Y_1}^{nR}$. The problem reduces to solving a system of $n(1-R)$ linear equations with $n(1-R)$ variables, which has a unique solution. Note that R is the actual rate of the code, and not the design rate.

[3]It can be shown on the lines of the proof in [1] that the last inequality is, in fact, an equality. Key point is to use the uniform distribution over the codewords.

Since $\mathbf{X_1}^{nR}$ has a uniform distribution over all its possible 2^{nR} values, and $\mathbf{Z_1}^{nR}$ is independent of $\mathbf{X_1}^{nR}$, $\mathbf{Y_1}^{nR} = \mathbf{X_1}^{nR} + \mathbf{Z_1}^{nR}$ also has a uniform distribution over all its possible 2^{nR} values. Thus

$$H(\mathbf{Y_1}^{nR}) = nR. \tag{10}$$

Now it is sufficient to upper bound $H(\mathbf{S}^{n(1-R)})$, the entropy of the syndrome vector. In general, the syndromes are not independent of each other. From the chain rule, and from the fact that conditioning reduces entropy, we get

$$H(\mathbf{S}^{n(1-R)}) \leq \sum_{i=1}^{n(1-R)} H(S_i). \tag{11}$$

The problem now reduces to bounding the entropy of individual syndromes.

IV. Upper Bounds on the Rate

In this section we give the bounds on the rate for different channels by finding bounds on the entropy of the syndromes.

Consider a parity check equation, corresponding to a row of the parity check matrix $\mathbf{H}$. Let the places at which 1's occur in the equation be denoted by $n_1, n_2, \ldots n_d$, and the corresponding output random variables be denoted by $Y_{n_1}, Y_{n_2}, \ldots, Y_{n_d}$. Let $S = \sum_{i=1}^{d} Y_{n_i}$, where addition is over GF (2). The entropy of a single syndrome is given by $H(S)$.

Note that the input codeword to the channel $\mathbf{X}^n$ satisfies the parity check equations, $\sum_{i=1}^{d} X_{n_i} = 0$. Therefore,

$$S = \sum_{i=1}^{d} Y_{n_i} = \sum_{i=1}^{d} (Y_{n_i} + X_{n_i}) = \sum_{i=1}^{d} Z_{n_i}. \tag{12}$$

That is, a particular parity check is satisfied if (and only if) there are even number of errors in locations corresponding to $\{Z_{n_i}\}_{i=1}^{d}$.

Since the state space is Markov, determining $\Pr(S = 0)$ exactly is not possible without knowing the exact positions of the 1's in the row. Even if exact positions of the 1's are known, the procedure to find $\Pr(S = 0)$ would be tedious in general. So we develop some methods to bound this probability. We now proceed to derive bounds on the rate for different channels by finding bounds on $\Pr(S = 0)$.

A. A Simple Upper Bound for FSMCs

We first present a simple upper bound on the rate of LDPC codes that holds for all noninverting FSMCs. We need the following Lemma.

Lemma 2: For an m-state FSMC as defined above

$$\Pr(S=0) = \frac{1}{2} + \frac{1}{2} \sum_{r_1, r_2, \ldots, r_m} \prod_{i=1}^{m} (1-2\eta_i)^{r_i} \Pr(r_1, r_2, \ldots, r_m) \tag{13}$$

where $\Pr(r_1, r_2, \ldots, r_m)$ denotes the probability of making r_i visits to state i in d steps (with $\sum_{i=1}^{m} r_i = d$).

Proof: See Appendix II. □

Since the channel is noninverting in each state, $1 - 2\eta_i > 0$ for all i. Hence $\Pr(S = 0)$ in Lemma 2 can be lower bounded by

$$\begin{aligned}\Pr(S = 0) &\geq \frac{1}{2} + \frac{1}{2} \sum_{r_1, r_2, \ldots, r_m} (1 - 2\eta_m)^d \Pr(r_1, r_2, \ldots, r_m) \\ &= \frac{1}{2} + \frac{(1 - 2\eta_m)^d}{2}\end{aligned}$$

where, we assume without loss of generality that $\eta_m > \eta_i$ for all $i \neq m$. Define $p_{md} \triangleq \frac{1}{2} + \frac{(1-2\eta_m)^d}{2}$. Then $H(S) \leq H(p_{md})$, and from (11), $H(\mathbf{S}^{n(1-R)}) \leq n(1 - R)H(p_{md})$.

Suppose now that the rate $R = 1 - \frac{1-C_m}{H(p_{md})} + \epsilon$, for some $\epsilon > 0$. Using (2), (8) –(11), we get

$$H(\langle P_e \rangle) > \frac{1}{n} H(\mathbf{X}^n | \mathbf{Y}^n) \geq \epsilon H(p_{md}). \tag{14}$$

That is, reliable communication is not possible at this rate for any $\epsilon > 0$. Hence we get the following bound for regular codes

$$R \leq 1 - \frac{1 - C_m}{H(p_{md})}. \tag{15}$$

Furthermore, (14) also gives us an error floor for rates exceeding this bound.

For irregular codes, the expression for upper bound on $H(\mathbf{S}^{n(1-R)})$ changes to

$$H(\mathbf{S}^{n(1-R)}) \leq n(1 - R) \sum_{d} \omega_d H(p_{md}). \tag{16}$$

Hence, the bound on the rate for irregular codes can similarly be found to be

$$R \leq 1 - \frac{1 - C_m}{\sum_d \omega_d H(p_{md})}. \tag{17}$$

Similar error floors (as in (14)) can be obtained for irregular codes. Since all the bounds we give in this work are of the same form, similar error floors can be derived for all of them.

Arguments in [8] concluded, using Jensen's inequality, that for memoryless channels, the expression of upper bound on the rate gives the result that for a sequence of codes to be capacity achieving, the density $\Delta(\mathbf{H})$ must converge to infinity. Since the expression for upper bound is same here with appropriate change in the expression for the capacity of an FSMC, we can conclude that even for Markov channels, for a sequence of LDPC codes to be capacity achieving, the density must converge to infinity.

In the next subsections, we tighten these bounds for the case of GE channels, and then for FSMCs.

B. Upper Bounds for Noninverting and Nonoscillating GE Channels

From Lemma 2, it can be seen that for noninverting GE channels, $\Pr(S = 0) > 0.5$, and hence, the entropy $H(S)$ increases with decrease in $\Pr(S = 0)$. Thus, to upper bound $H(S)$, we lower bound $\Pr(S = 0)$. To that end, we need the following Lemmas.

Lemma 3: For nonoscillating and noninverting GE channels, $\Pr(S = 0)$ decreases with increase in the gap between any two 1's, keeping the gap between other 1's constant.

Proof: See Appendix IV. □

Define the average error probability of an FSMC in steady state as $q \triangleq \sum_i \eta_i \gamma_i$. Consider a row of $\mathbf{H}$. Define $k_{\min}$ as the minimum gap between any two ones in the row. Then we have the following Lemma.

Lemma 4: Given that the underlying Markov chain is irreducible and aperiodic, for an m-state FSMC, as $k_{\min} \to \infty$, $\Pr(S = 0)$ converges to $\Pr(S_{\text{memless}} = 0)$, where S_{memless} is the random variable representing result of a parity check equation for a memoryless channel with error probability same as the average error probability of the FSMC in steady state.

Proof: See Appendix V. □

From Lemma 3 and Lemma 4, $H(S) \leq H(S_{\text{memless}}) = H(\frac{1+(1-2q)^d}{2})$, where $q = \gamma_G \eta_G + \gamma_B \eta_B$ is the average probability of error in steady state of the FSMC.

Thus, similar to the derivation of (15), we get the following bound

$$R \leq 1 - \frac{1 - C_{GE}}{H(\overline{p_d})} \tag{18}$$

where C_{GE} is the capacity of a GE channel, and $\overline{p_d} \triangleq \frac{1+(1-2q)^d}{2}$

Similar to (16), the bound on the rate for reliable communication for irregular codes is

$$R \leq 1 - \frac{1 - C_{GE}}{\sum_d \omega_d H(\overline{p_d})}. \tag{19}$$

We plot this bound for typical values of the channel parameters in Fig. 1 for regular code with increasing right degree. The plot shows that the bound approaches capacity exponentially in the right degree, as would be expected from (18).

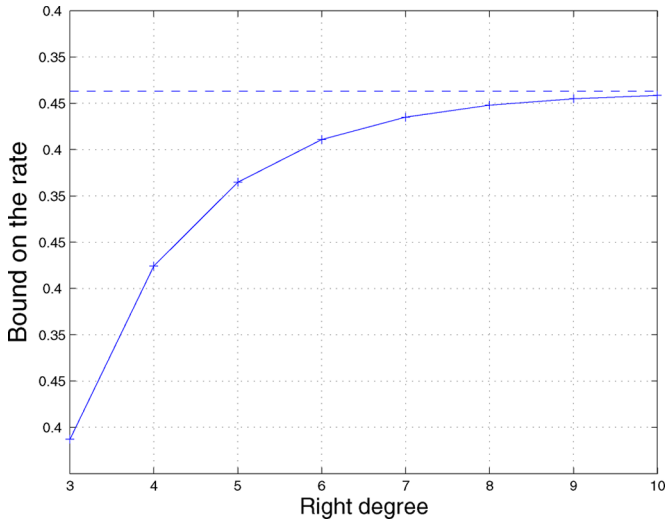

Fig. 1. The bound on the rate versus the right degree for a regular code on a GE channel with parameters $\eta_G = 0.01, \eta_B = 0.5, g = 0.6, b = 0.2$. The dashed line indicates the channel capacity.

C. Tightening the Bound for GE Channels

In this section, we tighten the upper bound on the entropy of a syndrome, which can be used to tighten the bounds on the rate for GE channels.

Suppose in a row of a parity check matrix, the *maximum* gap between any two consecutive 1's is v, that is, any two variables in that parity check equation are separated by a gap of no more than v. By Lemma 3, the entropy of the parity check increases as the gap between the 1's is increased. Hence the entropy of the given parity check will be lesser than or equal to the entropy of a parity check for which the gap between 1's is uniformly v. Notice that in the latter case, the Markov chain relating Z_{n_i}'s is homogeneous, albeit the transition probabilities have changed.

It was proved by Pedler [12] that for a homogeneous two state Markov chain, probability of visiting a particular state (say G) k times in d transitions is given by:

For $0 < k < d$

$$\begin{aligned}\Pr(N_G = k) = {} & (1-b)^k (1-g)^{d-k} F[-d+k, -k; 1; \Lambda] \\ & - \gamma_G l (1-b)^k (1-g)^{d-k-1} \\ & \quad \times F[-d+k+1, -k; 1; \Lambda] \\ & - \gamma_B l (1-b)^{k-1} (1-g)^{d-k} \\ & \quad \times F[-d+k, -k+1; 1; \Lambda]\end{aligned}$$

and

$$\begin{aligned}\Pr(N_G = 0) &= (\gamma_G b + \gamma_B (1-g))(1-g)^{d-1} \\ \Pr(N_G = d) &= (\gamma_G (1-b) + \gamma_B g)(1-b)^{d-1}\end{aligned}$$

where N_G is the random variable denoting number of times state G is visited, F is the hypergeometric function, $\gamma_G = \frac{g}{b+g}$ and $\gamma_B = \frac{b}{b+g}$ are the steady state probabilities of G and B respectively, $\Lambda = \frac{gb}{(1-g)(1-b)}$, and $l = (1-g)(1-b) - gb$.

Now, using (36) of Appendix II, the entropy of a parity check can be determined if we know $\Pr(N_G = k)$. For uniform gap v between 1's, the underlying state space for $Z'_{n_i}s$ is homogeneous Markov, and hence Pedler's result is applicable. Using v-step transition probabilities and (36), associated entropy can be calculated for different values of v. Fig. 2 shows the variation of $H(S)$ with v for typical values of d, η_G, η_B, g and b. As v increases, we see that $H(S)$ converges to the entropy in the independent case, as is expected.

Hence if v is known for each parity check, a tighter bound on entropy $H(S)$ can be obtained, and thus the bounds derived earlier for GE channels can be tightened. However, for fixed (λ, ρ), as we let $n \to \infty$, only in some specific constructions of LDPC codes would the maximum gap between $1's$ remain constant. Hence, per se, the tightening of the bound in this section is not really useful, even though it gives an insight into how the bound varies increasing gap. We revisit this point in Section V, where we bring out the utility of this result.

The tightening raises a natural question: can we conclude from this tightening that LDPC codes constructed in a manner that gap between 1's is bounded cannot achieve capacity? We show in Appendix III that the proposed tightening does *not* lead to this conclusion.

D. An Almost-Sure Bound for FSMCs

1) Upper Bound on the Rate: We now derive an *almost-sure* upper bound on the rate of a sequence of LDPC codes for reliable communication over an FSMC. For each length n, we have an ensemble of codes $C^n(\lambda, \rho)$ and a uniform probability distribution over $C^n(\lambda, \rho)$. The probability space of sequence of codes is the product space of probability spaces corresponding to each n. The probability distribution over the product space is induced by uniform probability distribution over $C^n(\lambda, \rho)$. The bound is *almost-sure* in the sense that any sequence of codes $\{\mathcal{C}_n\}$ (with $\mathcal{C}_n \in C^n(\lambda, \rho)$) has to satisfy this bound with probability 1 (in the product space) if it communicates reliably.

Consider a parity-check equation. By Lemma 4, as $k_{\min}$, the minimum gap between any two consecutive 1's in the equation, increases, $\Pr(S = 0) \to \Pr(S_{\text{memless}} = 0) = \frac{1+(1-2q)^d}{2}$, where $q = \sum_i \eta_i \gamma_i$ is the average probability of error in the steady state. Since binary entropy function is a continuous function of probability, $H(S) \to H(S_{\text{memless}})$ as $k_{\min} \to \infty$. In particular, we can choose $k_{\min}$ large enough such that $H(S) \leq H(S_{\text{memless}}) + \delta$ for any given $\delta > 0$.

Lemma 5: For any $\epsilon > 0$, any fixed k, and n large enough, the fraction of codes which have at least ϵ fraction of rows with $k_{\min} \leq k$ is upper bounded by

$$\begin{aligned}p_{bd} &= \frac{(nd(1-R))^{-n(1-R)\epsilon} (kc)^{n(1-R)\epsilon}}{(1-\epsilon)^{n(1-R)(1-\epsilon)} \epsilon^{n(1-R)\epsilon} 2\pi n(1-R)\sqrt{\epsilon(1-\epsilon)d}} \\ &= O\left(\frac{1}{ne^{n(\epsilon \ln(n) - c_2)}}\right)\end{aligned} \tag{20}$$

where c_2 is a constant.

Proof: See Appendix VI. □

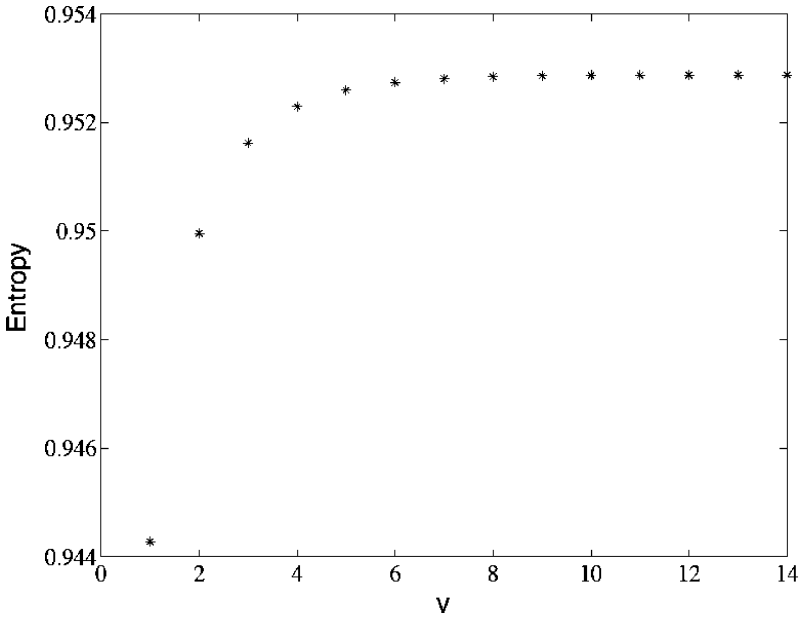

Fig. 2. $H(S)$ versus v for $d = 10, \eta_G = 0.01, \eta_B = 0.1, g = 0.2, b = 0.3$.

We note that the result also holds for irregular codes with maximum variable and check node degrees instead of c and d.

From Lemma 5, it follows that the rate of decrease of probability is at least exponential in n. Therefore, for large enough n (given k), all 1's are separated by a gap greater than k for at least $(1-\epsilon)$ fraction of rows with probability that goes to 1 at least exponentially in n. By our choice of k, the probability that $H(S) > H(S_{\text{memless}}) + \delta$ holds for more than ϵ fraction of rows converges to 0 exponentially fast in n.

From the above argument, we have an upper bound on the entropy of $(1-\epsilon)n(1-R)$ syndromes. For the entropy of rest of $\epsilon n(1-R)$ syndromes, we use the upper bound of 1, and arrive at the following *probabilistic* bound on $H(\mathbf{S}^{n(1-R)})$.

$$H(\mathbf{S}^{n(1-R)}) \leq n(1-R)(1-\epsilon)(H(S_{\text{memless}}) + \delta) + n(1-R) \quad (21)$$

which holds with probability at least $1 - p_{bd}$ with p_{bd} as in (20). Since the decrease in the probability p_{bd} is at least exponential, an application of Borel-Cantelli Lemma [13] shows that the bound holds eventually, almost surely (e.a.s.).[4] That is, with probability 1, the event $H(S) > H(S_{\text{memless}}) + \delta$ for greater than ϵ fraction of rows will happen only finitely often. We thus arrive at the following bound

$$H(\mathbf{S}^{n(1-R)}) \leq n(1-R)(1-\epsilon)(H(S_{\text{memless}}) + \delta) + n(1-R)\epsilon \ \ e.a.s. \text{ for all } \epsilon, \delta > 0. \quad (22)$$

This gives us

$$\frac{1}{n} H(\mathbf{S}^{n(1-R)}) \leq (1-R)H(S_{\text{memless}}) + \theta \ \ e.a.s. \text{ for all } \theta > 0 \quad (23)$$

where $H(S_{\text{memless}}) = H(\frac{1+(1-2q)^d}{2})$, and $\theta := (1-R)\Big((1-\epsilon)\delta + \epsilon\Big)$. For a fixed length n, θ is the slack in the probabilistic bound on entropy of the syndrome vector. For a given ϵ and k in Lemma 5, the value of θ and p_{bd} can be found using (20). Now, similar to the derivation in Section IV-A, for any sequence of regular LDPC codes that communicates reliably, the rate is bounded as follows

$$R \leq 1 - \frac{1 - C_m}{H(\overline{p_d})} + \zeta \quad \text{e.a.s. for all } \zeta > 0 \quad (24)$$

where $\overline{p_d} = \frac{1+(1-2q)^d}{2}$.

Define $R_0 \triangleq 1 - \frac{1-C_m}{H(\overline{p_d})}$. For given c, d, the design rate $R_D = 1 - \frac{c}{d}$ and R_0 are fixed, and hence if $R_D > R_0$, the code rate $R(\geq R_D)$ would always (and hence, infinitely often) exceed $R_0 + \zeta$ for $\zeta = R_D - R_0$. Thus the above bound would be violated. Hence the same bound holds for $\zeta = 0$ also[5] if we replace R by R_D. For the same reason, for fixed R_D and R_0, the term *eventually* in (24) is redundant. Thus, the bound for regular codes of given (c, d) becomes

$$R_D \leq 1 - \frac{1 - C_m}{H(\overline{p_d})} \ \text{ a.s..} \quad (25)$$

Similarly for irregular codes of given (λ, ρ), the bound is

$$R_D \leq 1 - \frac{1 - C_m}{\sum_d \omega_d H(\overline{p_d})} \ \text{ a.s..} \quad (26)$$

There can exist sequences of codes of the given (λ, ρ) which defy the bound, but the set of such sequences is of probability 0 in the product space. Note that we do not need the channel to be noninverting for the *almost-sure* bound to hold.

The utility of this bound is contingent upon the utility of the random construction of LDPC codes for FSMCs. Over memoryless symmetric channels, by the concentration theorem [7], the performance of a code chosen randomly is close to the average performance. Hence if a degree distribution pair (λ, ρ) performs well under density evolution, for large lengths, codes chosen randomly from $C^n(\lambda, \rho)$ would also perform well. Concentration theorem, therefore, validates the utility of the random construction. Hence, it would suffice to prove the concentration theorem for decoding of LDPC codes over Markov channels. The proof

[4]The term eventually, almost-surely implies that there is a set of infinite sequences of probability 1 such that for each sequence of this set, the statement holds for all $n > n_0$, where n_0 may be dependent on the sequence

[5]Consider a sequence of random variables $\{X_n\}$. Suppose $X_n < K + \epsilon$ a.s for all $\epsilon > 0$. But this does *not* guarantee that $X_n < K$ a.s. For example, take $X_n = 1 + \frac{1}{n}$ with probability 1. Then $X_n < 1 + \epsilon$ a.s. for all $\epsilon > 0$. But clearly, $X_n \not< 1$ a.s. Hence we use this alternative argument to arrive at the bound in (25) and (26).

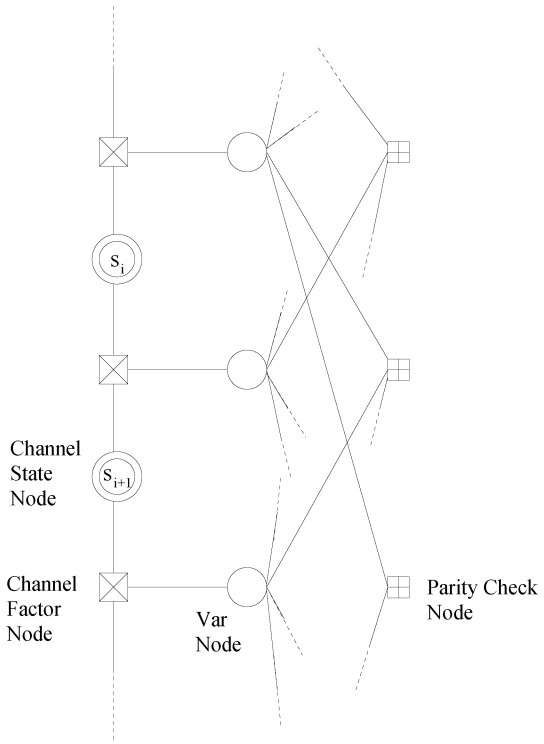

Fig. 3. Factor graph for decoding of an LDPC code over an FSMC.

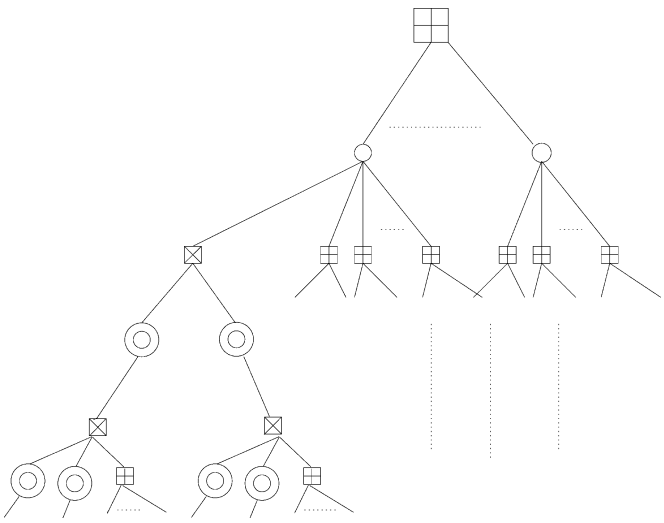

Fig. 4. Decoding neighborhood of factor graph given in Fig. 3.

is a direct extension of proof in [7], so we give only a sketch. The only difference is in finding a bound on neighborhood size after fixed l iterations, which we desire to be independent of the blocklength. Looking at the factor graph in Fig. 3, it is easy to see that for the decoding neighborhood graph in Fig. 4, the neighborhood size is indeed independent of n. Hence, the concentration theorem continues to hold.[6]

V. Lower Bounds on Parity-Check Density

Notice that the upper bounds in (15), (16), (18), (19), (25), and (26) are similar to expressions of upper bound on the rate derived in [1], with the capacity of general MBIOS channels replaced by C_m. In [8], lower bounds on parity check density of LDPC codes were derived for MBIOS channels using the same upper bounds. The key observation is that the bound on the syndrome entropy is an increasing function of the density. The bounds on the density derived in [8] are of two types.

- *Type I:* For any linear code of given finite blocklength n and given probability of error when transmitted over a channel, using the lower bound on $\frac{1}{n}H(\mathbf{X}|\mathbf{Y})$, a lower bound on $\Delta(\mathbf{H})$ can be found (see [8, Th. 2.4]).
- *Type II:* For a sequence of linear codes $\{\mathcal{C}_n\}$ which communicate reliably to achieve $1-\epsilon$ of the capacity of an MBIOS channel, asymptotic lower bound on the density of their parity check matrices can be found (see [8, Th. 2.1]).

In the sequel, we extend these bounds to the setting of FSMCs.

A. Lower Bounds on the Density for Noninverting and Nonoscillating GE Channels

Let $\Delta^{(n)}$ denote the density of parity-check matrix corresponding to $\mathcal{C}_n$. Similar to the derivation in [8], we get the following *Type I* bound on the density of a fixed code given its performance

$$\Delta \geq \frac{1-(1-\epsilon)C_{GE}}{2(1-\epsilon)C_{GE}} \cdot \frac{ln\left(\frac{1}{2ln2}\cdot\frac{1-C_{GE}+\epsilon C_{GE}}{\epsilon C_{GE}+H(\langle P_e\rangle)}\right)}{ln\left(\frac{1}{1-2q}\right)}. \tag{27}$$

Again, using methods in [8], and using the upper bound on the syndrome entropy derived in Section IV-B, we can get the following *Type II* bound on asymptotic density for reliable communication at rate $(1-\epsilon)C_{\mathrm{GE}}$ over GE channels

$$\underline{lim}_{n\to\infty}\ \Delta^{(n)} > \frac{K_1 + K_2 ln\frac{1}{\epsilon}}{1-\epsilon}. \tag{28}$$

where

$$K_1 = \frac{(1-C_{\mathrm{GE}})\ ln\left(\frac{1-C_{\mathrm{GE}}}{2\ ln2\ C_{\mathrm{GE}}}\right)}{2C_{\mathrm{GE}}\ ln\left(\frac{1}{1-2q}\right)} \text{ and } K_2 = \frac{1-C_{\mathrm{GE}}}{2C_{\mathrm{GE}}\ ln\left(\frac{1}{1-2q}\right)}$$

1) Tightening of Lower Bounds on the Density for GE Channels Using the Results in Section IV-C : For using the tight bound for GE channels of Section IV-C, we first prove the following Lemma.

Lemma 6: For a noninverting FSMC (in particular, for a noninverting GE channel), $\Pr(S=0) = \Pr(\sum_{i=1}^{d} Z_{n_i} = 0)$ decreases with increase in the row weight d.

Proof: See Appendix VII. □

Thus the syndrome entropy increases as the row weight increases. Consider regular LDPC codes. Suppose we are given a bound on the maximum gap between two consecutive 1's in the parity-check matrix of a code. Also, we are given the performance of the code over a nonoscillating and noninverting GE channel.

Since $\Pr(S=0)$ decreases with increasing d (by Lemma 6), adding an extra 1 in each row of a parity-check matrix would decrease $\Pr(S=0)$, which leads to an increase in $H(S)$. Thus we can obtain a lower bound on the number of 1's required in each row using Fano's inequality, where, in the bound on syndrome entropy, we use the tight bound on $\Pr(S=0)$ derived in Section IV-C. Therefore, similar to derivation in [8], we can use this Lemma to derive a tighter *Type I* bound for this particular case, where the bound on syndrome entropy is the one developed in Section IV-C.

Because there is no closed form expression for $P(S=0)$ in Section IV-C, we could not prove the convexity of $H(S)$ as a function of d, and hence the lower bounds on $\Delta(\mathbf{H})$ do not extend directly to irregular codes. However, in the light of Lemma 6, we observe that the bounds continue to hold for the maximum row weight, instead of the average row weight.

[6]For the neighborhood in Fig. 4, we use the scheduling proposed in [3], but it is clear that the proof works for any scheduling.

B. Lower Bounds on the Density for FSMCs

1) Deterministic Bounds: Using the simple upper bound on syndrome entropy derived in Section IV-A, we can derive the following *Type I* lower bound on the density of a fixed linear code, given its block length and performance over FSMC

$$\Delta \geq \frac{1-(1-\epsilon)C_m}{2(1-\epsilon)C_m} \cdot \frac{ln\left(\frac{1}{2ln\,2} \cdot \frac{1-C_m+\epsilon C_m}{\epsilon C_m + H(\langle P_e \rangle)}\right)}{ln\left(\frac{1}{1-2\eta_m}\right)}. \tag{29}$$

Again, using the simple upper bound on the syndrome entropy derived in Section IV-A for noninverting FSMCs, we get the following *Type II* lower bound on asymptotic density of any sequence of linear codes that achieves reliable communication at rate $(1-\epsilon)C_m$ over the FSMC

$$\underline{lim}_{n\to\infty}\ \Delta^{(n)} > \frac{K_1 + K_2 ln\frac{1}{\epsilon}}{1-\epsilon}. \tag{30}$$

where

$$K_1 = \frac{(1-C_m)\ ln\left(\frac{1-C_m}{2\ ln 2\ C_m}\right)}{2C_m\ ln\left(\frac{1}{1-2\eta_m}\right)} \text{ and } K_2 = \frac{1-C_m}{2C_m\ ln\left(\frac{1}{1-2\eta_m}\right)}.$$

2) Tighter Probabilistic Bounds: The lower bounds derived using the simple bound on the syndrome entropy for noninverting FSMCs are very loose (though they are deterministic). Tighter probabilistic *Type II* bounds for all FSMCs hold *almost-surely* for the design density Δ_D of a sequence of LDPC codes of given (λ, ρ). Similar derivation works here too, giving us the following bound:

$$\Delta_D > \frac{K_1 + K_2 ln\frac{1}{\epsilon}}{1-\epsilon} \text{ a.s.} \tag{31}$$

where

$$K_1 = \frac{(1-C_m)\ ln\left(\frac{1-C_m}{2\ ln 2\ C_m}\right)}{2C_m\ ln\left(\frac{1}{1-2q}\right)} \text{ and } K_2 = \frac{1-C_m}{2C_m\ ln\left(\frac{1}{1-2q}\right)}.$$

For finite lengths (*Type I* bound), based on the probabilistic bounds on the syndrome entropy for random construction derived in Section IV-D, we can derive the following probabilistic bounds on Δ which hold for LDPC codes of given (λ, ρ), and given performance $\langle P_e \rangle$

$$\Delta \geq \frac{1-(1-\epsilon)C_m}{2(1-\epsilon)C_m} \cdot \frac{ln\left(\frac{1}{2ln\,2} \cdot \frac{1-C_m+\epsilon C_m}{\epsilon C_m + H(\langle P_e \rangle)+\theta}\right)}{ln\left(\frac{1}{1-2q}\right)}. \tag{32}$$

Here, θ is as defined in (23). A lower bound on the probability with which this bound holds is $1-p_{bd}$, where p_{bd} is given by Lemma 5. This probability is dependent on choice of ϵ and n, and converges to 1 exponentially in n.

VI. Conclusion

In this work, we generalized the bounds on the rate of LDPC codes for reliable communication over BSC to FSMCs. We first derived a simple upper bound on the rate for reliable communication over all noninverting FSMCs. Using this bound, we proved that for a sequence of LDPC codes to be capacity achieving over an FSMC, the density $\Delta(\mathbf{H})$ must converge to infinity.

For noninverting and nonoscillating GE channels, we obtained a tighter upper bound, and showed that this can be further tightened using the knowledge of maximum gap between 1's in the rows of $\mathbf{H}$.

The bound on the syndrome entropy derived for GE channels was used to derive lower bounds on the parity check density for given performance, which is a generalization of results of Sason and Urbanke [8]. The tighter upper bound, which uses the knowledge of maximum gap between 1's in the rows of $\mathbf{H}$, leads to tighter lower bounds on the density of regular codes. For irregular codes, this leads to a necessary condition on row weights that has to be satisfied for any irregular code for given performance over a GE channel.

For FSMCs, we also proved that the tight upper bound derived holds *almost-surely* for any sequence of codes of fixed (λ, ρ).

We showed that the simple bound on the rate over FSMCs derived here can be used to derive *Type I* and *Type II* lower bounds on $\Delta(\mathbf{H})$. Using the derivation of *almost-sure* bound on the rate, the corresponding lower bound for finite length (*Type I*) holds with high probability, and lower bound for reliable communication on the design density (*Type II*) holds *almost-surely*.

Appendix I
Capacity of Markov Channels With the Channel Behaving as a BSC in Each State

We prove Lemma 1. The derivation is similar to the derivation of channel capacity for GE channels in [14]

$$\begin{aligned} C_m &= \lim_{n\to\infty} \frac{1}{n} \max_{p(\mathbf{X}^n)} I(\mathbf{X}^n; \mathbf{Y}^n) \\ &= \lim_{n\to\infty} \frac{1}{n} \max_{p(\mathbf{X}^n)} (H(\mathbf{Y}^n) - H(\mathbf{Y}^n|\mathbf{X}^n)) \\ &= 1 - \lim_{n\to\infty} \frac{1}{n} H(\mathbf{Z}^n). \end{aligned} \tag{33}$$

The last equality follows because $H(\mathbf{Y}^n)$ achieves its maximum value n when $\mathbf{X}^n$ is uniformly distributed over the possible 2^n values, and the distribution of $H(\mathbf{Z}^n)$ does not depend on distribution of $\mathbf{X}^n$. Also, as proved in [11, Th. 4.2.2]

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} H(Z_i|\mathbf{Z}^{i-1}) = \lim_{i\to\infty} H(Z_i|\mathbf{Z}^{i-1}). \tag{34}$$

Thus, from the chain rule and (33),

$$C_m = 1 - \lim_{i\to\infty} H(Z_i|\mathbf{Z}^{i-1}). \tag{35}$$

□

Appendix II
Probability of a Parity Check Equation Being Satisfied

We first prove Lemma 2 in a more general setting. Suppose we have some sequence $\{X_{n_i}\}_{i\geq 1}$ as input to an m–state channel

(not necessarily Markov)[7] which behaves as a BSC in each state. Let $\{Z_{n_i}\}_{i\geq 1}$ be the error sequence, and we want to calculate $\Pr(S=0)=\Pr(\sum_{i=1}^{d} Z_{n_i}=0)$, where the addition is modulo two. We prove the following:

$$\Pr(S=0)=\frac{1}{2}+\frac{1}{2}\sum_{r_1,\ldots,r_m}\Pi_{i=0}^{m}(1-2\eta_i)^{r_i}\Pr(r_1,\ldots,r_m) \quad (36)$$

where $\Pr(r_1,\ldots,r_m)$ denotes the probability of making r_i visits to state i in d steps, that is, $\sum_i r_i = d$.

For two-state channels (in particular for GE channels), the above result reduces to

$$\Pr(S=0)=\frac{1}{2}+\frac{1}{2}\sum_{k=0}^{d}(1-2\eta_G)^{k}(1-2\eta_B)^{d-k}\Pr(N_G=k) \quad (37)$$

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

Remark :

- Note that in (36), the probability is greater than 0.5 as long as $\eta_i < 0.5$ for all i (ensuring that the second term remains positive). This fact has been used frequently in the proofs presented here.

Proof: The proof is by induction on m. For $m=1$, the case reduces to that for a BSC. As shown in [2], the probability of even errors in d channel uses of a BSC is given by $\frac{1+(1-2\eta)^d}{2}$, where η is the crossover probability of the BSC. It is easy to verify that the expression given in (37) reduces to this.

We now assume that the result is true for $m-1$ states, and prove the result for m state systems.

Let $\text{EvErr}(r_1,r_2,\ldots,r_m)$ denote the event that there are even number of errors in $\sum_i r_i$ steps where r_i visit are made to state i. Similarly define $\text{OddErr}(\cdot)$. Note that d is the row weight in the row of the parity-check matrix under consideration. Then

$$\begin{aligned}&\Pr(\sum_{i=1}^{d} Z_{n_i}=0)\\ &=\Pr(\text{even errors in locations corresponding to } \{Z_{n_i}\})\\ &=\Pr(\text{EvErr}(r_1,r_2,\ldots,r_{m-1});\text{EvErr}(r_m))\\ &\quad+\Pr(\text{OddErr}(r_1,r_2,\ldots,r_{m-1});\text{OddErr}(r_m)). \quad (38)\end{aligned}$$

[7]In Appendix IV, we have a nonhomogeneous Markov chain, and we use the result there.

We now investigate the first term in (38) and see (39) at the bottom of the page. Doing similar analysis on second term of (38), and adding the result to (39), we arrive at (36) (the middle two terms cancel out). □

Appendix III
As $d\to\infty$, $\Pr(S=0)\to\frac{1}{2}$

Using (37)

$$\begin{aligned}\frac{1}{2}&\leq\Pr(\sum_{i=1}^{d} Z_{n_i}=0)\\ &=\frac{1}{2}+\frac{1}{2}\sum_{k=0}^{d}(1-2\eta_G)^{k}(1-2\eta_B)^{d-k}\Pr(N_G=k)\\ &\leq\frac{1}{2}+\frac{1}{2}\sum_{k=0}^{d}(1-2\eta_G)^{d}\Pr(N_G=k)\\ &=\frac{1}{2}+\frac{1}{2}(1-2\eta_G)^{d}\overset{d\to\infty}{\longrightarrow}\frac{1}{2} \qquad \square. \quad (40)\end{aligned}$$

This means that regardless of the gap between 1's, as $d\to\infty$, the bounds on the rate in Section IV-C approach capacity. Note that the same proof works for FSMCs as well.

Appendix IV
Increase in Gap Increases Syndrome Entropy for Noninverting and Nonoscillating GE Channels

First we prove that

$$\Pr(\sum_{i=1}^{d} Z_{n_i}=0|s_{n_d}=G)>\Pr(\sum_{i=1}^{d} Z_{n_i}=0|s_{n_d}=B). \quad (41)$$

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

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

Define

$$\begin{aligned}b_t&\triangleq\frac{b-b(1-g-b)^t}{g+b}\\ g_t&\triangleq\frac{g-g(1-g-b)^t}{g+b}.\end{aligned} \quad (42)$$

$$\begin{aligned}&\Pr\Big(\text{EvErr}(r_1,r_2,\ldots,r_{m-1});\text{EvErr}(r_m)\Big)\\ &=\sum_{r_m=0}^{d}\Pr\Big(\text{EvErr}(r_1,r_2,\ldots,r_{m-1});\text{EvErr}(r_m)\Big|r_m\Big)\Pr(r_m)\\ &=\sum_{r_m=0}^{d}\Pr\Big(\text{EvErr}(r_1,r_2,\ldots,r_{m-1})\Big|r_m\Big)\Pr\Big(\text{EvErr}(r_m)\Big|r_m\Big)\\ &=\sum_{r_m=0}^{d}\frac{1}{2}\Big(1+\sum_{r_1,r_2,\ldots,r_{m-1}}\Pi_{i=1}^{m-1}(1-2\eta_i)^{r_i}\Pr(r_1,r_2,\ldots,r_{m-1}|r_m)\Big)\frac{1+(1-2\eta_m)^{r_m}}{2}\Pr(r_m)\\ &=\frac{1}{4}\Big(1+(1-2\eta_m)^{r_m}+\sum_{r_1,r_2,\ldots,r_m}\Pi_{i=1}^{m-1}(1-2\eta_i)^{r_i}\Pr(r_1,\ldots,r_m)+\sum_{r_1,\ldots,r_m}\Pi_{i=1}^{m}(1-2\eta_i)^{r_i}\Pr(n_1,\ldots,r_m)\Big). \quad (39)\end{aligned}$$

Using t –step transition probability matrix for a two state Markov chain

$$P^t = \begin{bmatrix} \frac{g+b(1-g-b)^t}{g+b} & \frac{b-b(1-g-b)^t}{g+b} \\ \frac{g-g(1-g-b)^t}{g+b} & \frac{b+g(1-g-b)^t}{g+b} \end{bmatrix} = \begin{bmatrix} 1-b_t & b_t \\ g_t & 1-g_t \end{bmatrix}. \quad (43)$$

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

$$\begin{aligned} &\Pr(\sum_{i=1}^{k} Z_{n_i} = 0 | s_{n_k} = G) \\ &= \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_k} = G) \Pr(Z_{n_k} = 0 | s_{n_k} = G) \\ &\quad + \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 1 | s_{n_k} = G) \Pr(Z_{n_k} = 1 | s_{n_k} = G) \\ &= \Bigg[\Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = G)(1-b_t) \\ &\quad + \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = B) b_t \Bigg] (1-\eta_G) \\ &\quad + \Bigg[\Pr(\sum_{i=1}^{k-1} Z_{n_i} = 1 | s_{n_{k-1}} = G)(1-b_t) + \\ &\quad \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 1 | s_{n_{k-1}} = B) b_t \Bigg] \eta_G. \end{aligned} \quad (44)$$

Similarly, for conditioning on B we get

$$\begin{aligned} &\Pr(\sum_{i=1}^{k} Z_{n_i} = 0 | s_{n_k} = B) \\ &= \Bigg[\Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = B)(1-g_t) \\ &\quad + \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = G) g_t \Bigg] (1-\eta_B) \\ &\quad + \Bigg[\Pr(\sum_{i=1}^{k-1} Z_{n_i} = 1 | s_{n_{k-1}} = B)(1-g_t) \\ &\quad + \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 1 | s_{n_{k-1}} = G) g_t \Bigg] \eta_B. \end{aligned} \quad (45)$$

Now observe the terms in first square brackets in (44) and (45), which are expressions of probability of event $\sum_{i=1}^{k-1} Z_{n_i} = 0$ given s_{n_k} (G or B). Both the terms are of the form

$$\begin{aligned} f(a) = \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = G) a \\ + \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = B)(1-a) \end{aligned}$$

where $a = 1 - b_t$ in (44) and $a = g_t$ in (45). Now

$$\begin{aligned} f(a) = a(&\Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = G) \\ &- \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = B)) \\ &+ \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = B) \end{aligned}$$

which is an increasing function of a, since

$$\Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = G) > \Pr(\sum_{i=1}^{k-1} Z_{n_i} = 0 | s_{n_{k-1}} = B)$$

by induction assumption. We want to conclude that term under consideration is greater in (44). From (42), we can see that $1 - g_t > b_t$ for any nonoscillatory GE channel, and therefore, term a is greater in (44) and hence the term in first square brackets (which is under consideration here) is also greater in (44). Also, in (44) and (45), this term is greater than 0.5 (follows since the channel is noninverting, see Appendix II).

Now, the probability expressions in (44) and (45) can be written as $h(x, y) = xy + (1-x)(1-y)$, where, x takes the value of term in the first 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 (44) and hence (41) follows. □

Increase in the Gap Between 1's Increases Entropy: Let $S = \sum_{i=1}^{d} Z_{n_i} = S_1 + S_2$, where $S_1 = \sum_{i=1}^{d_1} Z_{n_i}$ and $S_2 = \sum_{i=d_1+1}^{d} Z_{n_i}$. We prove that if the gap $r \triangleq n_{d_1+1} - n_{d_1}$ increases, corresponding entropy $H(S)$ also increases. To prove this, it is sufficient to prove that $\Pr(S_1 + S_2 = 0)$ decreases as r increases (since $\Pr(S_1 + S_2 = 0) > 0.5$).

$$\Pr(S_1 + S_2 = 0) = \Pr(S_1 = 0; S_2 = 0) + \Pr(S_1 = 1; S_2 = 1). \quad (46)$$

Note that

$$\begin{aligned} &\Pr(S_1 = 1; S_2 = 1) \\ &= \Pr(S_2 = 1) - \Pr(S_1 = 0; S_2 = 1) \\ &= \Pr(S_2 = 1) - \Pr(S_1 = 0) + \Pr(S_1 = 0; S_2 = 0). \end{aligned} \quad (47)$$

Since the terms $\Pr(S_2 = 1)$ and $\Pr(S_1 = 0)$ are independent of r, to prove $\Pr(S = 0)$ decreases as r increases, it is sufficient to prove (from (46) and (47)) that $\Pr(S_1 = 0; S_2 = 0)$ decreases as r increases.

We now prove that $\Pr(S_1 = 0 | S_2 = 0)$ decreases as r increases

$$\begin{aligned} &\Pr(S_1 = 0 | S_2 = 0) \\ &= \Pr(S_1 = 0 | s_{n_{d_1+1}} = G) \Pr(s_{n_{d_1+1}} = G | S_2 = 0) \\ &\quad + \Pr(S_1 = 0 | s_{n_{d_1+1}} = B) \Pr(s_{n_{d_1+1}} = B | S_2 = 0) \\ &= \Bigg[\Pr(S_1 = 0 | s_{n_{d_1}} = G) \Pr(s_{n_{d_1}} = G | s_{n_{d_1+1}} = G) \\ &\quad + \Pr(S_1 = 0 | s_{n_{d_1}} = B) \Pr(s_{n_{d_1}} = B | s_{n_{d_1+1}} = G) \Bigg] \\ &\quad \times \Pr(s_{n_{d_1+1}} = G | S_2 = 0) \\ &\quad + \Bigg[\Pr(S_1 = 0 | s_{n_{d_1}} = G) \Pr(s_{n_{d_1}} = G | s_{n_{d_1+1}} = B) \\ &\quad + \Pr(S_1 = 0 | s_{n_{d_1}} = B) \Pr(s_{n_{d_1}} = B | s_{n_{d_1+1}} = B) \Bigg] \\ &\quad \times \Pr(s_{n_{d+1}} = B | S_2 = 0) \end{aligned}$$

Thus

$$\begin{aligned} &\Pr(S_1 = 0 | S_2 = 0) \\ &= \Bigg[\Pr(S_1 = 0 | s_{n_{d_1}} = G) \times \frac{g + b(1-g-b)^r}{g+b} \\ &\quad + \Pr(S_1 = 0 | s_{n_{d_1}} = B) \times \frac{b - b(1-g-b)^r}{g+b} \Bigg] \\ &\quad \times \Pr(s_{n_{d_1+1}} = G | S_2 = 0) \\ &\quad + \Bigg[\Pr(S_1 = 0 | s_{n_{d_1}} = G) \times \frac{g - g(1-g-b)^r}{g+b} \end{aligned}$$

$$+\Pr(S_1=0|s_{n_{d_1}}=B)\times\frac{b+g(1-g-b)^r}{g+b}\Bigg]$$
$$\times\Pr(s_{n_{d_1+1}}=B|S_2=0)$$
$$=C_1+\frac{(1-g-b)^r}{g+b}$$
$$\times\Bigg[\Pr(S_1=0|s_{n_{d_1}}=G)-\Pr(S_1=0|s_{n_{d_1}}=B)\Bigg]$$
$$\times\Bigg[b\Pr(s_{n_{d_1+1}}=G|S_2=0)-g\Pr(s_{n_{d+1}}=B|S_2=0)\Bigg]$$

where C_1 is a constant independent of r. Observe that the two terms in product with $\frac{(1-g-b)^r}{g+b}$ are positive, and since channel is nonoscillatory, $(1-g-b)>0$, so $\Pr(S_1=0|S_2=0)$ decreases with increase in r.[8] Hence $\Pr(S=0)$ decreases with increase in r. Since $\Pr(S=0)>0.5$, the entropy $H(S)$ increases as r increases.

Appendix V
Convergence of $\Pr(S=0)$ to $\Pr(S_{\text{memless}}=0)$ as $k_{\min}\to\infty$

First, we prove that as $k_{\min}\to\infty$

$$\Pr(s_{n_1}=a_1,s_{n_2}=a_2,\ldots,s_{n_d}=a_d)\to\Pi_{i=1}^{d}\gamma_{a_i}. \tag{48}$$

Notice that

$$\Pr(s_{n_1}=a_1,s_{n_2}=a_2,\ldots,s_{n_d}=a_d)$$
$$=\Pr(s_{n_1}=a_1)\Pr(s_{n_2}=a_2|s_{n_1}=a_1)$$
$$\ldots\Pr(s_{n_d}=a_d|s_{n_{d-1}}=a_{d-1}).$$

As $k_{\min}$ increases, gap between each of n_i and n_{i-1} increases. If the Markov chain is irreducible and aperiodic [9], then $\Pr(s_{n_i}=a_i|s_{n_{i-1}}=a_{i-1})\to\gamma_{a_i}$. This proves (48).

Now consider the following for b_i binary:

$$\Pr(Z_{n_1}=b_1,Z_{n_2}=b_2,\ldots,Z_{n_d}=b_d)$$
$$=\sum_{a_i=1,2,\ldots,m}\Pr(Z_{n_1}=b_1,Z_{n_2}=b_2,$$
$$\ldots,Z_{n_d}=b_d|s_{n_1}=a_1,s_{n_2}=a_2,\ldots,s_{n_d}=a_d)$$
$$\times\Pr(s_{n_1}=a_1,s_{n_2}=a_2,\ldots,s_{n_d}=a_d)$$
$$\xrightarrow{k_{\min}\to\infty}\sum_{a_i=1,2,\ldots,m;i=1,2,\ldots,d}\Pi_{i=1}^{d}\Pr(Z_{n_i}=b_i|s_{n_i}=a_i)$$
$$\times\Pr(s_{n_i}=a_i)$$
$$=\sum_{a_i=1,2,\ldots,m;i=1,2,\ldots,d}\Pi_{i=1}^{d}\Pr(Z_{n_i}=b_i|s_{n_i}=a_i)\gamma_{a_i}$$
$$=\Pi_{i=1}^{d}\sum_{j=1}^{m}\Pr(Z_{n_i}=b_i|s_{n_i}=j)\gamma_j$$
$$=(1-q)^t q^{d-t}$$

where $q=\sum_{i=1}^{m}\eta_i\gamma_i$, and t is the number of $b_i's$ which are 0.

Finally, see the equation at the bottom of the page.

[8]To see that the second term in product with $\frac{(1-g-b)^r}{g+b}$ is positive, notice that $\Pr(s_{n_{d_1+1}}=G|S_2=0)=\frac{\Pr(S_2=0|s_{n_{d_1+1}}=G)}{\Pr(S_2=0)}\times\frac{g}{g+b}$, and a similar expression holds for $\Pr(s_{n_{d_1+1}}=B|S_2=0)$. Now use (41).

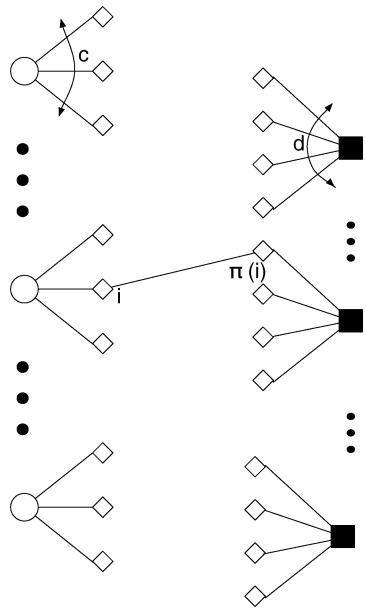

Fig. 5. Socket construction for a regular (c,d) code.

Appendix VI
Proof of Lemma 5

We prove the Lemma for regular codes.

In the following, we give detailed "socket construction" for regular LDPC codes. This construction would be used in the proof of Lemma 5.

The socket construction [7] is used to define ensembles of codes of given degree distribution. Consider (c,d) regular code. For each variable node, assign c sockets, and for each check node, assign d sockets. Total number of check node sockets is equal to the total number of variable node sockets. The ensemble of graphs is obtained by choosing a permutation π with uniform probability from the set of all possible permutations. Next, for each ith variable node socket is connected to $\pi(i)$th check node socket by an edge. If multiple edges link a pair of nodes, the nodes are considered to be connected if number of such edges is odd. The construction is shown Fig. 5.

First we look at the number of codes in the collection $C^n(c,d)$. The number of sockets in the random construction in [7] is $n(1-R)d$. Hence the number of codes is equal to the number of permutations of set of $n(1-R)d$ different elements. Thus the total number of codes in $C^n(c,d)$ is $(n(1-R)d)!$. Using Stirling's approximation, i.e., $n!\approx n^n e^{-n}\sqrt{2\pi n}$, the total number of codes can be approximated as

$$(n(1-R)d)$$
$$\approx(n(1-R)d)^{n(1-R)d}e^{-n(1-R)d}\sqrt{2\pi n(1-R)d}. \tag{49}$$

Now we find a bound on the number of codes which have at least an ϵ fraction of rows with gap between at least two 1's not exceeding k. Let τ_ϵ^k denote the set of such codes, and let $N_\epsilon^k=|\tau_\epsilon^k|$ denote the number of such codes.

Choose ϵ fraction of rows (rounding off to $\lceil\epsilon n(1-R)\rceil$ number of rows), where each row corresponds to a check node in the corresponding Tanner graph. This choice can be made in at most $\binom{n(1-R)}{\lceil\epsilon n(1-R)\rceil}$

$$\Pr(Z_{n_1}+Z_{n_2}+\cdots+Z_{n_d}=0)=\sum_{\text{even number of } b_i's \text{ are } 1}\Pr(Z_{n_1}=b_1,Z_{n_2}=b_2,\ldots,Z_{n_d}=b_d)\to\frac{1+(1-2q)^d}{2}.$$

ways. These are the rows for which at least two 1's have gap not exceeding k. For each of these rows, the first $d-1$ left sockets can be chosen in at most $(nd(1-R))^{d-1}$ ways. The last socket, which is at most a distance k from one of the 1's, can only be chosen in a constant (independent of n) number of ways, say c_1 ($c_1 \leq (k-1)c < kc$). Here we do not put any restriction for the choice of the first $d-1$ sockets. Only in choice of the last socket is the constraint on gap used.

For the rest $n(1-R) - \lceil \epsilon(1-R) \rceil$ rows, the number of ways in which variable node sockets can be chosen is at most $(nd(1-R))^d$. Thus

$$N_\epsilon^k < \binom{n(1-R)}{\lceil \epsilon n(1-R) \rceil} \left((nd(1-R))^{d-1} kc \right)^{\lceil n(1-R)\epsilon \rceil} \times \left((nd(1-R))^d \right)^{n(1-R) - \lceil \epsilon n(1-R) \rceil}. \quad (50)$$

Using Stirling's approximation

$$\binom{n(1-R)}{\lceil \epsilon n(1-R) \rceil} \approx \frac{n(1-R)^{n(1-R)} \sqrt{2\pi n(1-R)}}{\lceil \epsilon n(1-R) \rceil^{\lceil \epsilon n(1-R) \rceil}} \times \frac{1}{(n(1-R) - \lceil \epsilon n(1-R) \rceil)^{n - \lceil \epsilon n(1-R) \rceil} c_2 n} \quad (51)$$

where c_2 is a constant. Also notice that the ratio $\frac{\lceil \epsilon n(1-R) \rceil}{\epsilon n(1-R)}$ converges to 1. Using this fact and (49), (50) and (51), we can see that the probability of choosing a code from τ_ϵ^k goes to zero as

$$\frac{(nd(1-R))^{-n(1-R)\epsilon} (kc)^{n(1-R)\epsilon}}{(1-\epsilon)^{n(1-R)(1-\epsilon)} \epsilon^{n(1-R)\epsilon} 2\pi n(1-R)\sqrt{\epsilon(1-\epsilon)d}} \quad (52)$$

which falls as $O\left(\frac{1}{n e^{n(\epsilon \ln(n) - c_3)}} \right)$ for some constant c_3, and hence for n large enough, the decrease is exponential in n. □

The same proof works for irregular codes with bounded degrees, with the replacement of maximum left and right degrees in place of c, d.

Appendix VII
$\Pr(S=0)$ Decreases With Increase in d

We prove Lemma 6. The proof is by induction on d. Suppose we add another 1 to the parity check equation, which corresponds to adding another binary random variable Z to S. Now we have to consider $\Pr(S+Z=0)$.

First we note that

$$\Pr(S+Z=0) = \Pr(S=0; Z=0) + \Pr(S=1; Z=1). \quad (53)$$

Similar to the derivation of (47) in Appendix IV, we can see that

$$\Pr(S=0; Z=0) = \Pr(S=0) - \Pr(Z=1) + \Pr(S=1; Z=1). \quad (54)$$

From (53) and (54), we get

$$\begin{aligned} &\Pr(S=0) - \Pr(S+Z=0) \\ &= \Pr(Z=1) - 2\Pr(S=1; Z=1) \\ &= \Pr(Z=1) - 2\Pr(Z=1)\Pr(S=1|Z=1). \end{aligned} \quad (55)$$

Since the channel is noninverting, $\Pr(S=1|Z=1) < 0.5$ (from Appendix II). Thus the quantity in (55) is positive, and the Lemma follows. □

References

[1] D. Burshtein, M. Krivelevich, S. Litsyn, and G. Miller, "Upper bounds on the rate of LDPC codes," *IEEE Trans. Inf. Theory*, vol. 48, Sep. 2002.

[2] R. Gallager, "Low-Density Parity-Check Codes," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, 1960.

[3] A. Eckford, F. Kschischang, and S. Pasupathy, "Analysis of low-density parity-check decoding over the Gilbert-Elliott channel," *IEEE Trans. Inf. Theory*, vol. 51, Nov. 2003.

[4] A. Eckford, "Low-Density Parity-Check Codes for Gilbert-Elliott and Markov-Modulated Channels," Ph.D. dissertation, University of Toronto, Toronto, ON, Canada, 2004.

[5] A. Kavcic, X. Ma, and M. Mitzenmacher, "Binary intersymbol interference channels: Gallager codes, density evolution, and code performance bounds," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1636–1652, Jul. 2003.

[6] P. Grover and A. K. Chaturvedi, "Upper bounds on rate of LDPC codes for Gilbert-Elliott channels," in *Proc. IEEE Trans. Inf. Theory Workshop*, San Antonio, TX, USA, Oct. 24–29, 2004.

[7] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, pp. 599–618, Feb. 2001.

[8] I. Sason and R. Urbanke, "Parity-check density versus performance of binary linear block codes over memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1611–1635, Jul. 2003.

[9] S. M. Ross, *Introduction to Probability Models*, 8th, Ed. New York: Academic, 2002.

[10] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[11] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[12] P. Pedler, "Occupation times for two state Markov chains," *J. Appl. Prob.*, pp. 381–390, 1971.

[13] P. Billingsley, *Probability and Measure*, 3rd ed. New York: Wiley-Interscience, 1995.

[14] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert-Elliott channels," *IEEE Trans. Inf. Theory*, vol. 35, Nov. 1989.

Combinatorial Properties for Traceability Codes Using Error Correcting Codes

Hongxia Jin, *Member, IEEE*, and Mario Blaum, *Fellow, IEEE*

***Abstract*—In this correspondence, the combinatorial properties of traceability codes constructed from error-correcting codes are studied. Necessary and sufficient conditions for traceability codes constructed from maximum-distance separable (MDS) codes are provided. The known sufficient conditions for a traceability code are proven to be also necessary for linear MDS codes.**

***Index Terms*—Error-correcting codes, maximum-distance separable (MDS) codes, pirated copies, traceability codes, traitor tracing.**

I. Introduction

This correspondence is concerned with the traitor tracing problem for protection of copyrighted materials. The problem is typical in a broadcast encryption system [1]. When a pirated copy (decryption key) of the material is observed, a traitor tracing scheme allows to identify at least one of the real users (called traitors) who participated in the

Manuscript received August 29, 2005; revised July 18, 2006. The material in this correspondence was presented in part at the International Conference on Security and Cryptography, Setubal, Portugal, August 2006.

H. Jin is with the IBM Almaden Research Center, San Jose, CA 95120 USA (e-mail: jin@us.ibm.com).

M. Blaum is with the Hitachi Global Storage Technologies, San Jose, CA 95135 USA (e-mail: Mario.Blaum@hitachigst.com).

Communicated by E. Okamoto, Associate Editor for Complexity and Cryptography.

Digital Object Identifier 10.1109/TIT.2006.889730