Robustness of LDPC Codes on Periodic Fading Channels

Chris Jones, Tao Tian, Adina Matache, Richard Wesel, and John Villasenor
Electrical Engineering Department, University of California, Los Angeles, CA, 90095, USA
{christop, ttian, villa}@icsl.ucla.edu {matache, wesel}@ee.ucla.edu

Abstract— **Root and Variya proved the existence of codes that can communicate reliably over any member of a set of linear Gaussian channels where each member exceeds a given amount of mutual information. In this paper we show that LDPC codes are such codes and that their performance lies within 0.1 bits of the Root and Variya capacity for a large family of periodic Gaussian channels. Specifically, the robustness of LDPC codes to periodic fading is demonstrated through the consistency of their mutual information performance across period-2 and period-256 fading profiles. The latter case implies that these codes are ideal candidates for coding in OFDM.**

I. INTRODUCTION

A compound channel occurs when the actual channel is unknown to both transmitter and receiver but belongs to a set of possible channels known to both. In this paper we consider periodic fading channels that have at time i input x_i and output $y_i = \tilde{a}_{(i \bmod p)} x_i + n_i$, where n_i is additive white Gaussian noise (AWGN) with variance $N_o/2$ per dimension. The p-element vector $\tilde{a} = [a_0, \ldots, a_{p-1}]$ consists of complex scalars, which could be the subchannel gains of an OFDM system with p subcarriers. We are interested in codes that perform well over a family of periodic fading channels and denote the set of vectors $\tilde{a}$ that comprise such a family with $\tilde{A}$. This set defines a compound Gaussian channel.

Previous work on the design and analysis of trellis codes for the compound Gaussian channel was done by Wesel *et al.* in [1]. The authors of [1] showed that robust communication using a rate-k/n trellis code with v memory elements over a periodic channel can be achieved provided that the channel period p is less than the maximum *effective length* of the code $\lfloor v/k \rfloor + 1$. For puncturing periods longer than this, an $\tilde{a}$ can always be found that causes zero-output weight for non-zero input weight. Thus trellis codes with 1024 or fewer states cannot be expected to communicate reliably on channel sets with periods beyond 11 unless special constraints (for instance, strong adjacent subcarrier correlation) are assumed. As we will see, no such correlation or periodic length constraints are necessary when using Low-Density Parity-Check (LDPC) codes.

LDPC codes were first proposed by Gallager [2] in 1960. They describe a class of linear random codes whose encoding and decoding complexity follow $O(n^2)$ and $O(n)$ respectively. Though, it has been shown that the coefficient in front of n^2 will be small if care is taken in constructing the generator matrix [3]. Section II motivates the robustness property that we seek by making information theoretic statements about a set of channels that are collectively described as the *compound* channel. Background on LDPCs will be provided in Section III. Section IV constructs the metric computations needed to incorporate channel state information (CSI) in the LDPC iterative decoding algorithm. The robust operation of an LDPC code on period-2 and period-256 fading channels is demonstrated in Section V.

II. MUTUAL INFORMATION FOR PERIODIC CHANNELS

The mutual information of the channel $y = ax + n$ can be expressed as,

$$I(X;Y,A) = I(X;A) + I(X;Y|A) = I(X;Y|A), \quad (1)$$

where X and A are independent RVs. The expectation $E_A[I(X;Y|A)]$ defines the average mutual information of this channel if A is known (estimated) at the receiver. If, however, A is a deterministic constant (a) then the mutual information can be computed directly,

$$I(X;aX+N) = I(X;X+\frac{N}{a}) = h(X+\frac{N}{a}) - h(\frac{N}{a}). \quad (2)$$

The extension of this result to periodic fading follows for a particular instance of the p-element vector $\tilde{a}$,

$$I(\tilde{a}) = \frac{1}{p}\sum_{i=0}^{p-1} I(X_i; X_i + \frac{N}{a_i}), \quad (3)$$

which can also be used to define the capacity of a frequency selective fading ISI channel. If both X_i and N are Gaussian and each X_i has the same average power, $E[\mathrm{X_i}^2] = E_x$, then the average mutual information over period p is,

$$I(\tilde{a}) = \frac{1}{p}\sum_{i=0}^{p-1} \log_2(1+\tfrac{|a_i|^2 E_x}{2\sigma^2}). \quad (4)$$

The constant power constraint causes the mutual information in (4) to be less than the water-filling capacity that can be achieved if the transmitter knows $\tilde{a}$. Nevertheless, Shannon's basic noisy coding theorem ensures that for each $\tilde{a}$ there is a code with fixed symbol power E_x and rate R that achieves reliable communication with R arbitrarily close to $I(\tilde{a})$. For example, p parallel Gaussian alphabet codes could be designed with the ith code assigned rate $R_i = \log_2(1 + \frac{|a_i|^2 E_x}{2\sigma^2})$. This solution is unattractive as it requires transmitter and receiver to coordinate code

Funded by NSF Career Award CCR9733089 and the Office of Naval Research

0-7803-7632-3/02/$17.00 ©2002 IEEE

selection depending on $\tilde{a}$ and of course has tremendous complexity for large p .

In this paper we turn to the result of Root and Variaya [4] who proved that a single code exists that can communicate reliably at rates arbitrarily close to the compound channel capacity given by,

$$C(\tilde{A}) = \inf_{\tilde{a} \in \tilde{A}} (I(\tilde{a})). \tag{5}$$

For a given desired rate R (say for an OFDM broadcast), one might choose the set $\tilde{A} = \{\tilde{a} | I(\tilde{a}) > R\}$, i.e. every periodic channel $\tilde{a}$ such that the mutual information is above the transmitted rate. In this way, Root and Varaiya's theorem says that "universal" or "robust" codes exist that support rate R over every channel where *any* code exists that supports that rate. We conjecture that this robustness can be achieved within the LDPC paradigm.

Simulation and density evolution results will show that that a single LDPC code performs with less than 0.1 bits of *excess* mutual information for compound channels where $|\tilde{A}|$ is large. Note that excess mutual information is defined as the difference between the capacity of the channel, at the SNR for which the code operates at some arbitrarily low BER, and the rate of the code.

III. LOW-DENSITY PARITY-CHECK CODES

LDPC codes received little attention in the coding community until the introduction of turbo codes in the early 1990's. These codes both belong to the class of *iterative* codes. The demonstration of capacity approaching performance in turbo codes stimulated interest in the improvement of Gallager's original LDPC codes to the extent that in various measures they are now considered superior to turbo codes. For instance, for a given rate and number of coded bits, LDPC decoding complexity is lower than that of turbo codes [5]. Also, since they do not use convolutional constituent codes, they avoid catastrophic behaviour under random puncturing.

A major breakthrough in LDPC code design came with the invention of ***density evolution*** by Richardson, Shokrollahi, and Urbanke [6]. The authors showed that it is possible to predict a noise threshold below which a code realized from a given ensemble can be expected to converge to zero errors with high probability. A code ensemble is most often described via a pair of polynomials,

$$\lambda(x) = \sum_{\mathrm{i=1}}^{\mathrm{d_l max}} \lambda_i x^{i-1}, \rho(x) = \sum_{\mathrm{i=1}}^{\mathrm{d_r max}} \rho_i x^{i-1}. \tag{6}$$

The coefficients of $\lambda(x)(\rho(x))$ represent the fraction of *edges* emanating from variables (constraints) of various degree (in the bi-partite graph describing the code) as indicated by the powers of the place holding variables x^{deg-1}. For instance, the (3,6) regular code has $(\lambda(x), \rho(x)) = (x^2, x^5)$. Furthermore, since $\lambda(x)$, $\rho(x)$ are distributions $\lambda(1) = \rho(1) = 1$ always holds. Conversion between edge and node perspective is useful for defining the rate of the code in terms of a particular $(\lambda(x), \rho(x))$. Let E be the total number of edges in the graph, then $\frac{E\lambda_i}{i}$ equals the number of variable nodes with degree i. The rate of the code follows from the well known definition,

$$R = 1 - \frac{\sum \frac{\rho_i}{i}}{\sum \frac{\lambda_i}{i}} = 1 - \frac{\int \rho(x)}{\int \lambda(x)}. \tag{7}$$

The parity check matrix H of a linear binary $(n,\ k)$ systematic code has dimension $(n-k) \times n$. The rows of H comprise the null space of the rows of the code's $k \times n$ generator matrix G. H can be written as,

$$H = \left[\begin{array}{cc} H_1 & H_2 \end{array}\right], \tag{8}$$

where H_1 is an $(n-k) \times k$ matrix and H_2 is an $(n-k) \times (n-k)$ matrix. H_2 is constructed to be invertible, so by row transformation through left multiplication with H_2^{-1}, we obtain a systematic parity check matrix H_{sys} that is range equivalent to H,

$$H_{sys} = H_2^{-1} H = \left[\begin{array}{cc} H_2^{-1} H_1 & I_{n-k} \end{array}\right]. \tag{9}$$

The left hand portion of which can be used to define a null basis for the rows of H.

$$G_{sys} = \left[\begin{array}{cc} I_k & \left(H_2^{-1} H_1\right)^T \end{array}\right]. \tag{10}$$

As an example, the matrix and graph descriptions of an $(n = 9, k = 3)$ code are shown in Fig. 1. Because a vertex cannot connect to another vertex from the same (left or right) node set in a bipartite graph, the length of any cycle in the graph is an even number. Fig. 1 also shows a length-6 cycle (solid line) involving variable nodes v_0, v_4 and v_6.

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \tag{11}$$

Although the relationship of graph topology to code performance in the case of a specific code is not fully understood, work has been done to investigate the effects of graph structures such as cycles [7][8], stopping sets, linear dependencies, and expanders. The code used in this paper was designed to preclude small stopping sets and short cycles.

IV. CHANNEL INFORMATION IN THE DECODER

From here forward the periodic fading sequence $\tilde{a}$ will be deterministic and time invariant for a given channel. Hence it is known (or can be estimated) at the receiver and should be incorporated in the decoding if the system hopes to achieve capacity approaching performance. Consider as an example antipodal signaling where (χ^0, χ^1) represent transmission of (0,1) respectively. We define the

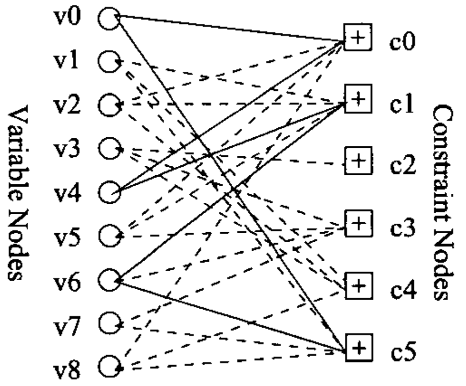

Fig. 1. Matrix and graph descriptions of a (9, 3) code.

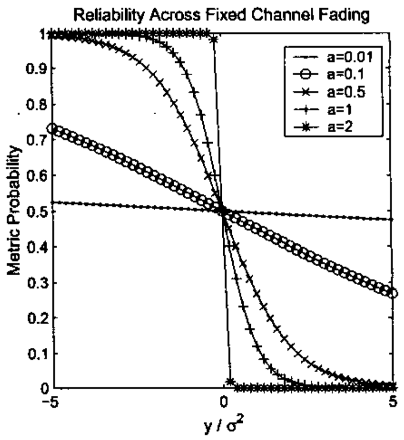

Fig. 2. Probability $p(v = 0)$ for several fading coefficients.

normalized probability that the ith message node (v_i) was transmitted with value χ^0, $p(v_i = \chi^0)$, via the progression of equalities,

$$\frac{p(y_i|x_i = \chi^0, |a_{(i \bmod p)}|)}{p(y_i|x_i = \chi^0, |a_{(i \bmod p)}|) + p(y_i|x_i = \chi^1, |a_{(i \bmod p)}|)} \tag{12}$$

$$= \frac{1}{1 + \frac{\phi(y_i - |a_{(i \bmod p)}|\chi^1)}{\phi(y_i - |a_{(i \bmod p)}|\chi^0)}} \tag{13}$$

$$= \frac{1}{1 + e^{\sigma^{-2}|a_{(i \bmod p)}|(\chi^1 - \chi^0)y_i}}. \tag{14}$$

where $\phi(\gamma)$ is the normal distribution $\mathcal{N}(0, \sigma^2)$. A plot of (14) for several values of fixed $|a|$ and $(\chi^0, \chi^1) = (-0.5, 0.5)$ is provided in Fig. 2. The domain of $|a|$ is $[0, \infty)$ and the figure shows that as $|a| \rightarrow 0$ equal probability is assigned to the likelihood of message observations, regardless of the observation value or noise power. This limiting case represents the occurrence of a symbol erasure. Intermediate values of $|a| \in (0, 1)$ have monotonically increasing levels of reliability assigned for a given y/σ^2. The case $|a| = 1$ represents symbol transmission across the AWGN channel and larger amplitude cases $|a| \in [1, \infty)$ display an ever greater slew rate across the intermediate probabilities.

A few comments concerning the role of reliability in a belief propagation (BP) message passing decoder are in order. The BP message passing decoder, whose message alphabet $M \in \Re$, can be thought of as a superset of a large family of smaller alphabet decoders. For instance, if $|a|$ were to take values from the two element set $\{0, +\infty\}$, indicating a deterministic BEC channel, then the BP decoder performance on this channel would no better than that of the much simpler BEC message passing decoder with message alphabet $M = \{-\infty, 0, +\infty\}$. In fact, the BP variable and constraint node update rules degenerate to BEC update rules under these conditions. The belief propagation decoder is thus a superset of the BEC decoder and many other decoders that operate with finite message alphabets. To summarize, the BP decoder, with metric information supplied as described, incorporates a continuum of reliability measures concerning received symbols. In one limiting case, it is capable of implementing erasure decoding where total confidence or total indecision is initially assigned to any given message node.

V. LDPC PERFORMANCE OVER PERIODIC CHANNELS

The LDPC code used in the following simulations has (n, k)=(15000, 5000) and was realized from the ensemble of codes with distribution: $\lambda(x) = 0.27603x + 0.11195x^2 + 0.17229x^3 + 0.01712x^4 + 0.42261x^{14}$, $\rho(x) = x^5$. The density evolution threshold for this code in AWGN is -0.147815 dB E_b/N_o, or -4.919 dB E_x/N_o. This degree sequence was found via a linear program that sought the highest rate ensemble under a given threshold and maximum left and right node degree constraints [9]. Note that we use E_x/N_o to describe the signal to noise ratio before channel scaling by $\tilde{a}$. As a benchmark, rate-1/3 BPSK constrained capacity is -5.3 dB (E_x/N_o) and rate-1/3 QPSK constrained capacity is 3 dB higher, or -2.3 dB (E_x/N_o).

For a first example consider the period-2 channel with $\tilde{a} = (1, a)$ and BPSK modulation. Fig. 3(a) clearly shows that a decrease in a requires an increase in E_x/N_o to maintain constant BER. The plot versus SNR, however, does not provide a clear view of the code's performance versus $I(\tilde{a})$ in (3) for the set of channels $\tilde{A} = \{(1, a)|0 \leq a \leq 1\}$. Instead, for each of the five channels in Fig. 3(b) we compute the BER versus mutual information using (3) in the context of BPSK modulation. The SNR of each plotted point in Fig. 3(a) has a corresponding mutual information under the BPSK and periodic channel constraints, which yields the plot of BER versus mutual information given in Fig. 3(b). The robustness of the code to period-2 fading

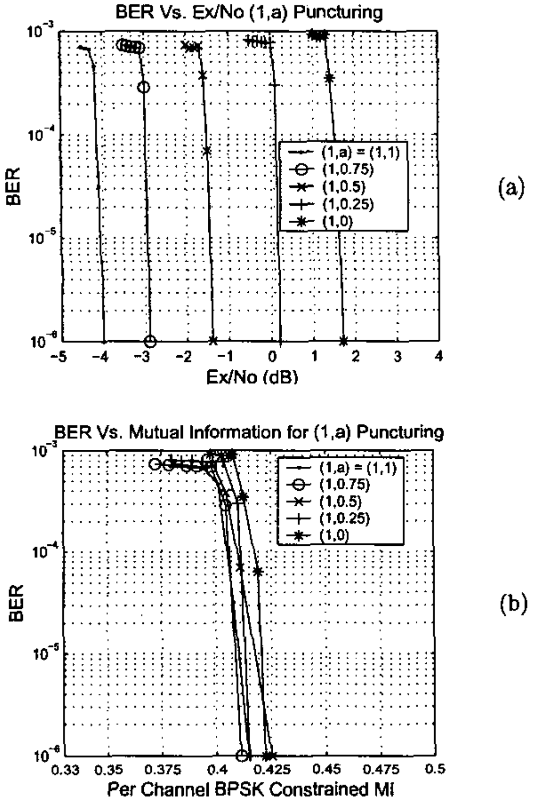

Fig. 3. (a) Code performance on the $(1,a)$ fading channel. (b) Code performance on the $(1,a)$ fading channel in terms of MI.

is now apparent. When each of the five channels provides a mutual information of at least 0.4125 bits and no more than 0.425 bits, the code communicates at or below BER=10^{-6}.

Density evolution using a Gaussian approximation as described in [9], but adapted for a periodic variation in the channel observations, was performed on all five of these channels. The lowest threshold, in terms of mutual information, was associated with the $\tilde{a} = (1,1)$ channel, $T(1,1) = 0.3495$ bits. The highest threshold, $T(1,0) = 0.3645$ bits, was found for the $\tilde{a} = (1,0)$ channel. Fig. 4 shows density evolution results for the code at initializing SNRs where the mutual information level on both channels equals 0.3645 bits. It is clear that the $\tilde{a} = (1,0)$ channel is "pinched off" while the AWGN channel exhibits an "open tunnel."

In general, the threshold $T(1,a)$ decreased monotonically as a increased. Thus for this particular code, the least favorable channel (the one requiring the most excess mutual information) is the $(1,0)$ erasure channel and the most favorable is the AWGN channel with $(1,1)$. We emphasize that the narrowness of the simulated mutual information

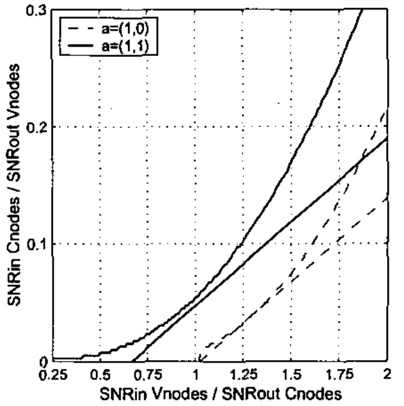

Fig. 4. Density evolution at pinch-off for a=(1,0) (E_x/N_0 = 0.1dB → 0.3645 bits MI on a=(1,0)). For a=(1,1) at the same mutual information an open tunnel is observed (0.3645 bits MI → E_x/N_0 = -4.760dB on a=(1,1)).

range, $\Delta T_{sim} = 0.425 - 0.4125 = 0.0125$ bits, shown Fig. 3(b) is in close agreement with the range predicted by density evolution, $\Delta T_{devo} = 0.3645 - 0.3495 = 0.015$ bits. However, the ordering in Fig.3(b) is not monotonic in a. We believe this is an artifact of not being able to experimentally resolve performance with an accuracy of less than one hundreth of a bit using reasonably sized Monte Carlo simulations.

To demonstrate that the code's robust operation is not limited to channels with small p, consider the four period-256 channels in Fig. 5. These fading profiles were generated by realizing channels with 4, 8, and 16 multi-path components in the time domain. The time channels were randomly generated with each tap magnitude and phase being a realization of a Rayleigh random variable. Exponential interarrival times between taps were assumed and an exponentially decaying envelope was imposed on the randomly realized taps. The 256 point FFT of each of these channels was taken and the magnitude of the resulting FFT coefficients (OFDM subcarrier gains) are shown for each channel in the plot. Channel (d), is identical to channel (c), with the exception of the erasure of an arbitrarily selected block of 125 consecutive subcarriers.

The performance of the Rate 1/3 LDPC code on these channels using QPSK modulation, where even(odd) code bits are mapped to I(Q) components, is given in Fig. 6(a). As a manifestation of the law of large numbers on the random subcarrier mutual informations, the QPSK constrained mutual information for these three channels is very similar. This explains the proximity of the three BER vs. SNR curves. Because of severe erasure distortion, channel (d) requires much more SNR for a given BER than the other channels to achieve the same level of mutual information. However, Fig. 6(b) shows that from the mutual information point of view the code works virtually as well on channel (d) as on channels (a,b,c).

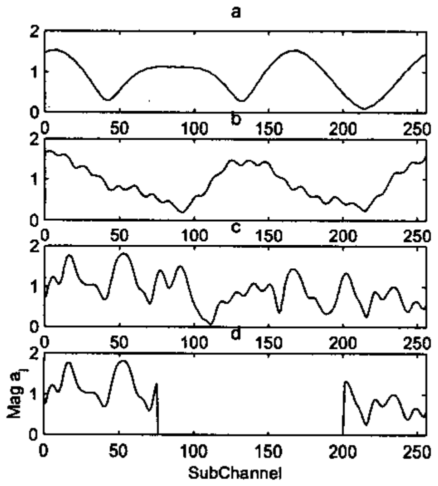

Fig. 5. Four period-256 Fading Channels.

At first glance, it may seem surprising that the code can communicate with 125 of the 256 subcarriers completely erased. However, the supremum of erasure rates for the code on the BEC channel,

$$\varepsilon^* = \sup(\varepsilon = x_0 | x_l = x_0\lambda(1-\rho(1-x_{l-1})) \rightarrow 0, l \rightarrow \infty)$$

has $\varepsilon^* = 0.613$. Note that ε^* is an asymptotic measure that can only be achieved in the limit of infinite block length. For the finite length code used throughout this paper, $\varepsilon^* = 0.59$ was found via simulation. Thus the minumum capacity of the QPSK BEC channel on which this code can be expected to communicate reliably is given by $C_{BEC} = 2(1-\varepsilon^*) = 0.82$. The high SNR (erasure) capacity of channel (d) is equal to $2(1-(125/256)) = 1.02$ bits. Therefore, it is reasonable to expect that the code can operate on this channel when E_x/N_o is large. However, we emphasize that the more remarkable result is that the difference in mutual information required for the code to operate on each of these four period-256 channels is less than 0.025 bits.

VI. Conclusion

This paper has taken a mutual information, rather than a signal-to-noise ratio, approach to measuring code performance over periodic Gaussian channels. Root and Variya showed that a single code exists that can communicate reliably on all of the channels in a given set provided that the rate of the code is greater than the smallest mutual information of all channels in the set. It has been shown for a quantized spread of all period-2 channels and for several arbitrarily selected period-256 channels that LDPC codes are an incarnation of Root and Variya's promise of 'universal' codes. Future work involves extending these arguments and demonstrations for general (matrix) linear Gaussian channels.

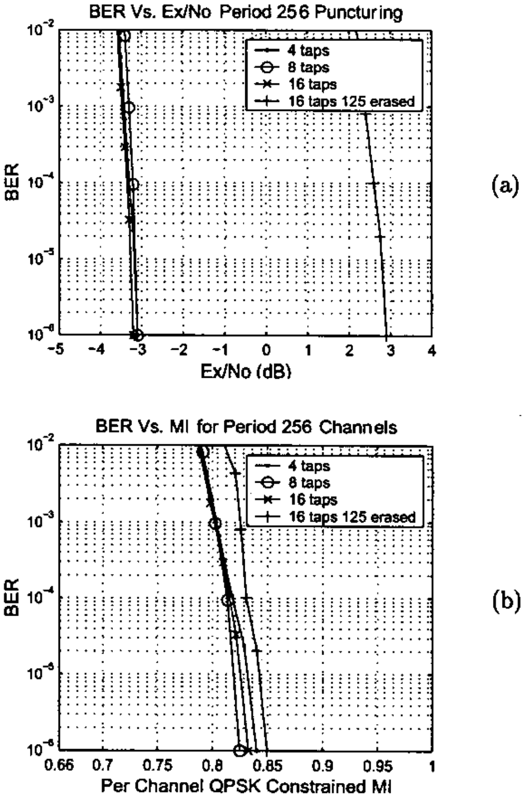

Fig. 6. (a) Code performance on four period-256 fading channels. (b) Code performance on four period-256 fading channels in terms of MI.

References

[1] R.D. Wesel, X. Liu, and W. Shi, "Trellis codes for periodic erasures," *IEEE Trans. on Comm.*, vol. 48, pp. 938–947, June 2000.

[2] R.G. Gallager, *Low-Density Parity-Check Codes*, Ph.D. thesis, Mass. Inst. Tech., Cambridge, Sept. 1960.

[3] T. Richardson and R. Urbanke, "Efficient encoding of low-density parity-check codes," *IEEE Trans. on Inform. Theory*, vol. 47, pp. 638–656, Feb 2001.

[4] W.L. Root and P.P Varaiya, "Capacity of classes of Gaussian channels," *SIAM J. Appl. Math*, vol. 16, no. 6, pp. 1350–1393, Nov. 1968.

[5] D.J.C. Mackay, "Good error correcting codes bases on very sparse matrices," *IEEE Trans. on Inform. Theory*, vol. 45, pp. 399–431, Mar 1999.

[6] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of capacity approaching irregular low density parity check codes," *IEEE Trans. on Inform. Theory*, vol. 47, pp. 618–637, Feb 2001.

[7] T. Tian, C. Jones, J. Villasenor, and R. D. Wesel, "Characterization and selective avoidance of cycles in irregular LDPC code construction," Submitted to *IEEE Trans. on Comm.*

[8] C. Di, D. Proietti, E. Telatar, T. Richardson, and R. Urbanke, "Finite length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1570–1579, June 2002.

[9] S.Y. Chung, T. Richardson, and R. Urbanke, "Analysis of Sum-Product Decoding of Low-Denisty Parity-Check Codes Using a Gaussian Approximation ," *IEEE Trans. on Inform. Theory*, vol. 47, pp. 657–670, Feb 2001.