

THE UNIVERSALITY OF LDPC CODES ON WIRELESS CHANNELS

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

ABSTRACT

Root and Variaya proved the existence of codes that can communicate reliably over any linear Gaussian channel for which the mutual information exceeds the information rate of the code. In this paper we demonstrate that properly designed Low-Density Parity-Check (LDPC) codes are such codes and that their performance lies in close proximity to the Root and Variaya capacity for the linear Gaussian vector channels (a.k.a. space-time channels). We also demonstrate the robustness of the codes on the partial-band jamming channel and in fast Rayleigh fading.

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. Root and Variaya's compound channel theorem [1] applied to the linear Gaussian vector channel,

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k \quad \mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, N_0 \mathbf{I}_n), \quad (1)$$

indicates that for a given rate R and input distribution there exists a *single* code that can achieve reliable information transmission at rate R on every channel \mathbf{H} for which the input distribution induces a mutual information (MI) higher than R . The immediate implication of this result is that good error performance on one particular channel does not have to come at the expense of significant performance degradation on others. Codes that have consistently good proximity to capacity (to the extent their blocklength and decoding complexity permit) over a class of channels will be referred to as *universal* codes in this paper. Since the linear Gaussian vector channels are commonly called space-time channels today, we refer to such codes as universal space-time codes.

The capacity promise of multiple-input multiple-output (MIMO) systems in rich scattering environments [2] makes the existence and use of universal codes of practical interest. In [3] Wesel et. al constructed universal trellis codes for periodic erasure channels. The universal property of LDPC codes in the context of periodic fading channels was described by Jones et. al in [4]. In [5][6] Köse and Wesel found by exhaustive search universal space-time trellis codes for the 2×2 linear Gaussian vector channel.

Low-density parity-check (LDPC) codes were proposed by Gallager in the early 1960s [7]. The structure of Gallager's codes (uniform column and row weight) led them to be called regular LDPC codes. Gallager provided simulation results for codes with block lengths on the order of hundreds of bits.

Funded by NSF Grant CCR0209110, Texas Instruments, and the State of California through UC MICRO Grant 03-093

However, these codes were too short for the sphere packing bound to approach Shannon capacity, and the computational resources for longer random codes were decades away from being broadly accessible.

Following the groundbreaking demonstration by Berrou *et al.* [8] of the impressive capacity-approaching capability of long random linear (turbo) codes, MacKay [9] re-established interest in LDPC codes during the mid to late 1990s. Luby *et al.* [10] formally showed that properly constructed irregular LDPC codes can approach capacity more closely than regular ones. Richardson, Shokrollahi and Urbanke [11] created a systematic method called density evolution to analyze and synthesize the *degree distribution* in asymptotically large random bipartite graphs under a wide range of channel realizations.

In this paper we demonstrate that LDPC codes are universal over a class of matrix channels by showing that bit-multiplexed LDPC coding on the MIMO channel yields essentially constant per dimension excess mutual information performance (around 0.1 bits per real dimension) for all the channels that we examine. For a matrix channel, the excess mutual information per dimension is defined as the capacity margin between the operational channel MI per transmit antenna and the information transmission rate per transmit antenna.

The rest of the paper is organized as follows. Section II describes the partial-band jamming channel and demonstrates the performance of LDPC in this channel environment. Typical ρ , E_b/N_f measures are provided, but an additional technique will be introduced specifically for the purpose of comparison to mutual information. Section III presents LDPC performance results across a uniform sampling of 2×2 MIMO channels to illustrate the universal property of the codes. The performance of the codes on any specific 2×2 quasi-static distribution (for instance Rayleigh) can be inferred from the provided results. Though fast Rayleigh fading does not fit in the Root and Variaya framework, we provide LDPC results for 1×1 through 4×4 MIMO systems with fast Rayleigh fading in Section IV. Section V summarizes our conclusions.

II. LDPC PERFORMANCE ON THE PARTIAL-BAND JAMMING CHANNEL

A special case of (1) is the partial-band jamming channel. The channel model used in the results that follow is that same as the one previously described in [12] and [13]. We limit our discussion to the case of coherently detected BPSK modulation under a frequency hopped scenario in which a fraction ρ of the available channels are jammed. All of the channels experience additive thermal noise due to the receiver front end. The SNR of this noise is fixed to $E_b/N_o = 20$ dB so as to be consistent with results in [12]. Channels that are jammed, however, also

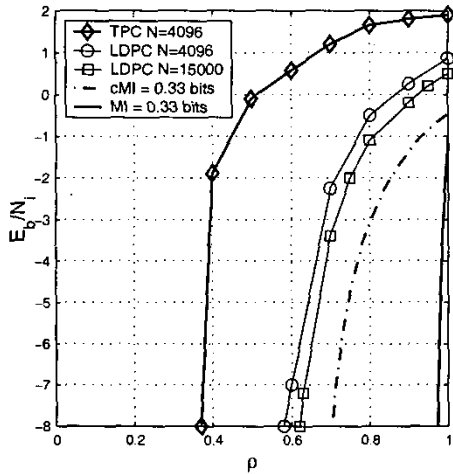


Fig. 1. Performance of Rate 1/3 LDPC codes with length 4096 and 15000 on the partial-band jamming channel compared to turbo product code performance [12]. Gaussian signaling capacity (MI) and BPSK constrained capacity (cMI) at 1/3 of a bit are also displayed. FER = 10^{-3}

incur the addition of band-limited white Gaussian noise with power spectral density $\rho^{-1}N_I$. Scaling by ρ ensures that the total jamming noise power $\rho(\rho^{-1}N_I) + (1 - \rho)0$ is equal to N_I , but is independent of ρ . Bit energy to interference ratio, E_b/N_I is the most common measure of performance on this channel. Perfect channel state information has been assumed for the LDPC results that will be presented. This implies that very low values of E_b/N_I tend to make jammed channels look like erasures as the log-likelihood ratios computed from channel observations are inversely scaled by the noise variance in a given subchannel. On the other hand, as ρ is increased to unity (all subchannels are jammed), the channel begins to appear much like a standard AWGN channel.

Fig. 1 provides simulation results for two rate 1/3 LDPC codes. Both are realized from the degree sequence of the (1,0) optimized code described in Section III. The first code has length 4096 and the second length 15000. For sake of comparison with another coding technique, the performance of a length 4096 turbo product code with comparable rate [12] is also provided. An important parameter for code performance on the partial-band jamming channel is the so-called dwell interval. This quantity describes the number of successive code symbols that will be transmitted on a given sub-channel before the modulation is hopped to another sub-channel. For sake of comparison with results in [12] we have fixed the dwell interval to 32 for the length 4096 code and to 30 for the length 15000 code. We have also made the assumption that channels are ‘framed’ around single code words. This implies that for the length 4096 code there are 128 subchannels and $\lceil \rho 128 \rceil$ of these will be jammed. There are 500 subchannels per frame for the length 15000 code. The distribution of jammed subchannels is realized uniformly and independently from one codeword transmission to the next. This technique is meant to yield an ‘ergodic’ jamming result for a given code

across a parameterization of ρ and E_b/N_o .

Mutual information (MI) computations for the partial-band jamming channel are also included in Fig. 1. To compute these measures consider the MI level in partial band jamming,

$$MI = \rho f(SNR_J) + (1 - \rho)f(SNR_{NJ}), \quad (2)$$

where SNR_J defines the symbol signal to noise ratio in the jammed subchannels and SNR_{NJ} defines the symbol signal to noise ratios in the non-jammed subchannels. In the case of Gaussian signaling, $f(\cdot) = \log_2(\cdot)$, and for the BPSK constrained case $f(\cdot)$ is evaluated via numerical integration. In the partial-band jamming simulations performed for this paper, SNR_{NJ} is held fixed at a level which corresponds to $E_b/N_o = 20$ dB. In the unconstrained case the term $\log_2(1 + SNR_{NJ})$ is therefore a constant (η) which can be determined via solution to the equation $\eta = \log_2(1 + \eta E_b/N_o)$, e.g. $\eta = 9.96$ bits. In the BPSK constrained case $MI_{BPSK}(SNR_{NJ})$ saturates to $\eta = 1$ bit at this high SNR.

We are interested in values of (ρ, SNR_J) that yield constant levels of mutual information. We therefore fix the MI to some constant level, say 1/3 of a bit. If we also fix ρ , it is possible to uniquely determine SNR_J (analytically for unconstrained and via table lookup for the BPSK constrained case). The resulting SNR_J can then be converted to E_b/N_I via the following relations,

$$\begin{aligned} SNR_J &= \frac{E_b}{\frac{N_I}{\rho} + N_o} = \frac{1}{2\sigma_{N_I}^2 + 2\sigma_{N_o}^2} \\ &= \frac{1}{\frac{1}{\rho R \frac{E_b}{N_I}} + \frac{1}{R \frac{E_b}{N_o}}} \\ \frac{E_b}{N_I} &= \frac{SNR_J}{\rho R \left(1 - \frac{SNR_J}{R \frac{E_b}{N_o}}\right)}. \end{aligned} \quad (3)$$

A large discrepancy can be observed between the BPSK constrained and Gaussian signaling mutual information curves in Fig. 1. This is due primarily to the fact that the non-jammed subchannels provide far more mutual information (9.96 bits) than can be provided by BPSK modulation, which in turn implies that with Gaussian signalling just a small fraction of the subchannels need to be non-jammed for the expected mutual information in the channel to reach 1/3 of a bit. We note that a system that achieves an average spectral efficiency of 1/3 bits, and that approaches the unconstrained Gaussian capacity, can be achieved by simultaneously increasing modulation cardinality and decreasing code rate. For instance a rate 1/6 code driving QPSK can be expected to perform better than the rate 1/3 BPSK system.

All of the curves in Fig. 1 represent systems with rate that is close to 1/3 of a bit of mutual information. In order to demonstrate the consistency of the mutual information performance of the two LDPC codes across SNR and ρ we plot the abscissa in Fig. 2 in terms of SNR_J rather than in terms of E_b/N_I . This avoids comparing E_b/N_I terms that differ only because of different rate (since $\frac{E_b}{N_I} \approx SNR_J - 10 \log_{10}(\rho R)$).

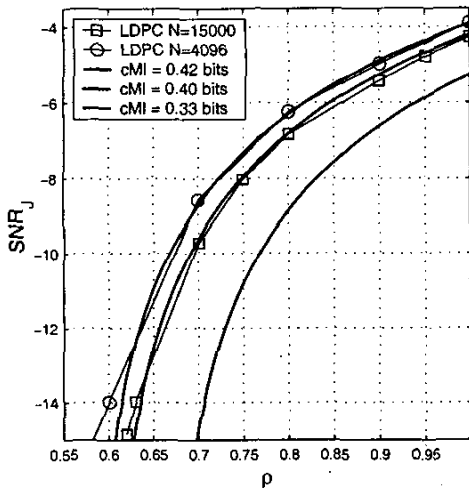


Fig. 2. SNR, ρ performance of length 4096 and 15000 LDPC codes compared to SNR, ρ levels required to achieve 0.42, 0.4, and 1/3 of a bit of mutual information. FER = 10^{-3}

When the partial-band jamming channel achieves 0.4 bits of mutual information, the length 15000 LDPC code operates with an FER = 10^{-3} at all but the lowest values of ρ . The same can be stated for the length 4096 code when the channel supports 0.42 bits of mutual information. The length 15000 results are consistent with the results that will be given in Section III, i.e. the code provides reliable communication when the excess MI in the channel is $\Delta MI = 0.4 - 1/3$, or roughly 0.065 bits. Furthermore, the performance on a parameterization of the channel that corresponds to an ‘erased’ case is slightly inferior to the performance on an AWGN ($\rho = 1$) parameterization.

III. MUTUAL INFORMATION PERFORMANCE OF LDPC CODES ON LINEAR GAUSSIAN VECTOR CHANNELS

In this section we consider the full Gaussian vector channel of (1). We note that mutual information identifies the fundamental information-carrying potential of a channel for a specific input distribution. Thus a universal code should provide performance that is consistent in terms of required excess mutual information. The common way to plot BER performance is versus channel signal-to-noise ratio (SNR). Since MI on the additive white Gaussian noise (AWGN) channel is a monotonic (and almost linear) function of SNR in dB, $MI_{\text{Gauss}} = \log_2(1 + \text{SNR})$, this representation is essentially equivalent to plotting BER against MI. For a fixed transmission rate R , SNR gap is defined as the difference between the SNR required to achieve the desired BER and the SNR at which the channel capacity in (4) is equal to R bits per channel use. However, when assessing the performance of a code over a variety of linear Gaussian channels, considering SNR performance or SNR gap is problematic because the monotonic relationship between SNR in dB and MI is different for different channel eigenvalue skews.

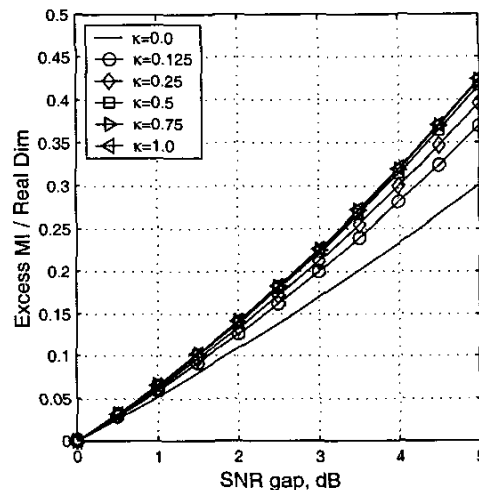


Fig. 3. Excess MI per real dimension vs. SNR gap for 2×2 matrix channels, $R = 4/3$ bits/channel use, $\lambda_1 = 1$, for eigenvalue skews (top to bottom) $\kappa = 1, 0.75, 0.5, 0.25, 0.125, 0$.

To better understand the previous statement, consider a 2×2 linear Gaussian vector channel \mathbf{H} and its MI under the assumption of uniform power distribution across antennas,

$$\begin{aligned} MI(\mathbf{H}) &= \log_2 \left(1 + \frac{\text{SNR}}{2} \lambda_1 \right) \left(1 + \frac{\text{SNR}}{2} \lambda_2 \right) \\ &= \log_2 \left(1 + \frac{\text{SNR}}{2} \lambda_1 \right) \left(1 + \frac{\text{SNR}}{2} \kappa \lambda_1 \right). \end{aligned} \quad (4)$$

Where λ_1 and λ_2 are the eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$, $\kappa = \lambda_2/\lambda_1$ is the eigenvalue skew and ρ is the average SNR per receive antenna.

Fig. 3 illustrates the excess MI per real dimension as a function of the SNR gap in dB for $R = 4/3$ bits/channel use and different eigenvalue skews κ . Note that the excess MI curves are approximately linear functions of the SNR gap in dB, however the slope depends on the eigenvalue spread (eigens skew) of the channel. In other words, a constant level of excess MI is achieved by differing excess SNR levels (depending on the eigens skew of the channel). The MI available in the channel is the absolute measure for performance, while an excess SNR measure depends both on the MI level and the specific channel realization (eigens skew).

In the remainder of this section the performance of two different length-15,000 bit-multiplexed LDPC codes on parameterized 2×2 MIMO channels will be described. The first set of channels is described via an explicit sampling of the set of all 2×2 channels. Each of these channels is characterized by several parameters, but of greatest importance (as it directly affects capacity) is eigens skew. The point of such an assessment is that ‘worst’ and ‘best’ case channels are easily found due to the absence of channel matrix averaging that occurs in both fast and quasi-static Rayleigh fading experiments. Furthermore, ‘flatness’ of the excess mutual information measure versus channel skew becomes a design criteria for creating

LDPC codes that exhibit the highest possible level of robustness. Again, in our terminology, a code is universal if reliable communication (for instance $\text{BER} = 10^{-5}$) occurs at the same small excess mutual information level across all channels. Of course, the absolute SNR required to achieve a given mutual information level on an eigenskew = 0 channel will be different than on an eigenskew = 1 channel.

Consider a sampling of the continuum of 2×2 channels. A simple description of the channel sampling problem is given as follows,

$$\mathbf{H} = \sqrt{\lambda_1} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\kappa} \end{bmatrix} \begin{bmatrix} \cos(\phi) & \sin(\phi)e^{j\theta} \\ -\sin(\phi)e^{-j\theta} & \cos(\phi) \end{bmatrix} \quad (5)$$

where $\kappa = \lambda_2/\lambda_1$ is the eigenvalue skew of the Hermitian matrix $\mathbf{H}\mathbf{H}^\dagger$, $\phi \in [0, \pi/2]$, and $\theta \in [0, 2\pi]$. In this paper we sample the above matrix via the parameters $\kappa = \{0, 0.125, 0.25, 0.5, 0.75, 1\}$, $\phi = \{0, \pi/4, \pi/2\}$ and $\theta = \{\pi/4, 3\pi/4\}$. Performance on each channel in this sampling is measured by the mutual information in excess of the transmission rate of 4/3 bits/channel-use required by a rate 1/3 code (modulating QPSK onto 2 antennas) to achieve a BER of 10^{-5} .

Before discussing the performance across these channels we draw an analogy to periodic fading SISO channels. Note that a period-2 SISO fading channel, $y_k = a_{(k \bmod 2)}x_k + n_k$, with fading vector $\mathbf{a} = (\sqrt{\lambda_1}, \sqrt{\lambda_2})$ is equivalent to a diagonal $\phi = 0$ 2×2 matrix fading channel (but requires two channel uses to relay the same information). Expanding the work in [4], we are able to optimize LDPC degree distributions for period- p fading via an adaptation of the Gaussian approximation to density evolution. In specific, at iteration l , degree i variable nodes have their mean values updated in correspondence to the periodic initial means given by $m_{a_i} = \frac{2a_i^2}{\sigma^2}$ (where a_j are known fading gains) and the means of messages arriving from check nodes (m_u),

$$m_{v,i}^{(l)}(j) = m_{a_i} + (i-1)m_u^{(l-1)}, j = \{0, \dots, p-1\}. \quad (6)$$

Randomly selected edges emanating from variable nodes adhere to the following Gaussian mixture density,

$$f_v^{(l)} = \sum_{j=0}^{p-1} \sum_{i=2}^{d_i} \frac{\lambda_i}{p} \mathcal{N}(m_{v,i}^{(l)}(j), 2m_{v,i}^{(l)}(j)) \quad (7)$$

Using this kernel, codes optimized for (1,1) and (1,0) period-2 fading channels were designed and tested across the parameterization of 2×2 channels. Both codes have $(n, k) = (15000, 5000)$. The code optimized for (1,1) has variable node degree distribution $\lambda_{(1,1)}(x) = 0.27603x + 0.11195x^2 + 0.17229x^3 + 0.01712x^4 + 0.42261x^{14}$ and constrained node degree distribution $\rho_{(1,1)}(x) = x^5$. The second code (optimized for (1,0)) has degree distributions $\lambda_{(1,0)}(x) = 0.3317x + 0.1247x^2 + 0.0417x^3 + 0.1780x^4 + 0.3239x^{14}$ and $\rho_{(1,0)}(x) = 0.5x^4 + 0.5x^5$.

Points plotted as circles in Fig. 4 illustrate the excess MI per real dimension for the (1,1) optimized code. Note that the results across ϕ and θ at each plotted κ are tightly clustered.

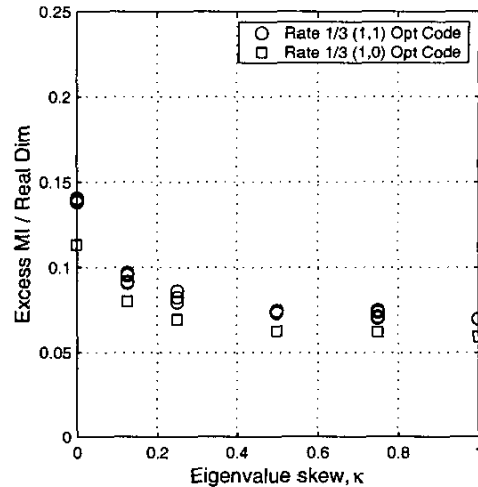


Fig. 4. Simulation results showing excess MI vs. eigenvalue skew κ at $\text{BER} = 10^{-5}$. Circles indicate results for a length 15,000 rate 1/3 code optimized for a (1,1) period-2 SISO channel. Squares indicate results for a length 15,000 rate 1/3 code optimized for a (1,0) period-2 SISO channel.

For this code, the best case performance is on the unitary ($\kappa = 1$) channel where the excess MI per real dimension required to achieved $\text{BER} = 10^{-5}$ is 0.07 bits (which corresponds to 1.05 dB excess SNR on the eigenskew = 1 channel). The worst performance is attained on the singular channels ($\kappa = 0$) where the excess MI per real dimension is 0.14 bits (2.5 dB excess SNR on the eigenskew = 0 channel).

Significant improvement in the overall robustness is found in the (1,0) optimized code, results for which are plotted as squares in Fig. 4 (for $\phi = \pi/4$ and $\theta = \pi/4$). The performance on the unitary channel with this code, 0.06 bits excess MI (0.91 dB excess SNR), is comparable to the performance with the (1,1) code. However, the performance on the singular channels improves and the excess MI requirement is reduced to 0.11 bits (2.0 dB excess SNR). Code design work to further flatten the profile of excess MI versus skew is underway.

In Fig. 5 λ_1 and λ_2 are related through $\text{MI} = \log_2(1 + \lambda_1)(1 + \lambda_2)$. Contours of MI corresponding to 1.33 and 1.61 total bits per channel use (or 4 times the number per real dimension) are shown. The difference between these two level curves corresponds to an excess mutual information of 0.28 bits (or 0.07 bits per real dimension). The adherence of the (1,1) optimized (circles) and (1,0) optimized (squares) codes to the 1.61 bit curve is tightest near eigenskew = 1.0 (the $\lambda_1 = \lambda_2$ line in the plot) and is looser at eigenskew = 0.

IV. MI PERFORMANCE IN FAST RAYLEIGH FADING

Fast Rayleigh fading, where the channel matrix is realized independently from one transmission time to the next, does not fit in the Root and Varaiya paradigm where performance is measured across a set of static channels. Nevertheless, this measure is frequently of interest to the community. Therefore, we examine the performance of the previous rate 1/3 (1,1) optimized code and a rate 1/2 (1,1) optimized code each

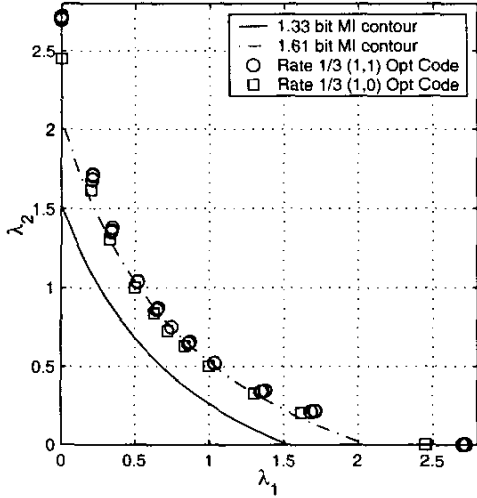


Fig. 5. Simulation results showing (λ_1, λ_2) operating points at $\text{BER} = 10^{-5}$ for length 15,000 codes optimized for (1,1) and (1,0) period-2 SISO channels. Also shown are λ_2 vs. λ_1 level curves for $\text{MI} = 1.33$ and 1.61 bits/channel use.

bit-multiplexed onto 1×1 , 2×2 , 3×3 , and 4×4 symbol-by-symbol (fast) fading Rayleigh channels. Fig. 6 describes code performance at $\text{BER} = 10^{-5}$ against constrained ergodic Rayleigh capacity for (1×1) BPSK and $(N \times N)$ QPSK (for $N = 1, 2, 3, 4$). Code robustness in terms of excess mutual information per real dimension required to achieve $\text{BER} = 10^{-5}$ is provided in Fig. 7. QPSK 1×1 Rayleigh faded performance is the same as the BPSK 1×1 performance on a per dimension basis. Since QPSK is an orthogonal extension of BPSK (when the real and imaginary additive noise components are independent) this result is expected. The excess mutual information performance per real dimension for increasing dimensionality (2×2 , 3×3 , 4×4 , etc.) remains essentially constant, around 0.1 bits, indicating the robust and scalable operation of the codes.

V. CONCLUSION

This paper has taken a mutual information, rather than a signal-to-noise ratio, approach to measuring code performance over linear Gaussian and partial-band jamming channels. Root and Varaiya showed that ‘universal’ codes must exist 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. Our results show that properly designed LDPCs can be thought of as *universal* codes for partial-band jamming, quasi-static fading, and fast Rayleigh fading channels.

APPENDIX - MAXIMUM A-POSTERIORI DETECTION FOR THE LINEAR GAUSSIAN VECTOR CHANNEL

We are interested in finding the probability of each transmit hypothesis given the channel and the received observation (we drop time index k to simplify notation),

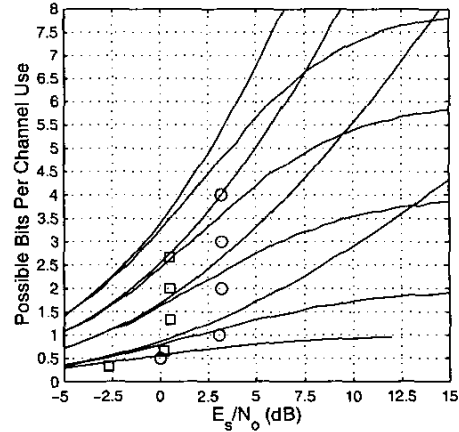


Fig. 6. Modulation constrained and unconstrained ergodic $N \times N$ Rayleigh capacities. Top to bottom, 4×4 QPSK, 3×3 QPSK, 2×2 QPSK, 1×1 QPSK, 1×1 BPSK. Rate 1/3 LDPC performance (square), Rate 1/2 LDPC performance (circle), both at $\text{BER} = 10^{-5}$. The SNR Gap for each case is less than 1.5dB, the MI gap is less than 0.11 bits/real-dim (see Fig. 7.)

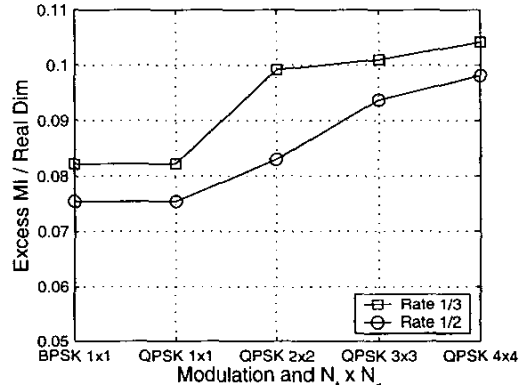


Fig. 7. Simulations results show that excess mutual information per real dimension in fast Rayleigh fading remains essentially constant across $N \times N$ channels for $N = 1$ to 4.

$$p(\mathbf{x}^{(\ell)} | \mathbf{y}, \mathbf{H}) = \frac{p(\mathbf{y} | \mathbf{x}^{(\ell)}, \mathbf{H}) p(\mathbf{x}^{(\ell)})}{p(\mathbf{y})}, \quad (8)$$

where $\mathbf{x}^{(\ell)}$ is the ℓ th transmit hypothesis (of $2^{N_s M_c}$ total possible hypotheses). Since likelihood ratios are ultimately computed, the quantity $p(\mathbf{y})$ is normalized away and the quantities of interest are $p(\mathbf{y} | \mathbf{x}^{(\ell)}, \mathbf{H})$ and $p(\mathbf{x}^{(\ell)})$. The latter is initialized as a uniform distribution, but will be updated via the recursions of the message passing decoder. This leaves the key measure to be extracted from the channel,

$$p(\mathbf{y} | \mathbf{x}^{(\ell)}, \mathbf{H}) = \frac{1}{(\pi \sigma_n^2)^{N_r}} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}^{(\ell)}\|^2}{\sigma_n^2}\right). \quad (9)$$

The joint in (8) must be marginalized before it can be used to express variable node reliabilities in the message passing

decoder. We express this marginalization in terms of log-likelihood ratios as follows [14],

$$\begin{aligned} L_{\text{Det}}(\mathbf{x}(b_j)) &= \ln \frac{p(\mathbf{x}(b_j=0)|\mathbf{y}, \mathbf{H})}{p(\mathbf{x}(b_j=1)|\mathbf{y}, \mathbf{H})} \\ &= \ln \frac{\sum_{\ell \in \{\ell|\mathbf{x}^{(\ell)}(b_j)=0\}} p(\mathbf{y}|\mathbf{x}^{(\ell)}, \mathbf{H}) p(\mathbf{x}^{(\ell)})}{\sum_{\ell \in \{\ell|\mathbf{x}^{(\ell)}(b_j)=1\}} p(\mathbf{y}|\mathbf{x}^{(\ell)}, \mathbf{H}) p(\mathbf{x}^{(\ell)})} \end{aligned} \quad (10)$$

where $\mathbf{x}^{(\ell)}(b_j)$ specifies the j th bit of transmit hypothesis ℓ .

There are $2^{N_t M_c - 1}$ terms in each summation and a total of $N_t M_c$ log-likelihood-ratios (LLRs) must be calculated for each received vector \mathbf{y} . Reliability updates from the decoder are incorporated in the detector via the approximation,

$$p(\mathbf{x}) \simeq \prod_{j=0}^{N_t M_c - 1} p(\mathbf{x}(b_j)), \quad (11)$$

where $\mathbf{x} = [b_0 b_1 \dots b_{N_t M_c - 1}]$. This product can be used in its entirety to update (10) and then the *intrinsic* reliability

$$L_{\text{Graph}}(\mathbf{x}(b_j)) = \ln \left(\frac{p(\mathbf{x}(b_j=0))}{p(\mathbf{x}(b_j=1))} \right) \quad (12)$$

should be backed out to form updated observation values for the next decoder iteration [15],

$$L_{\text{Ext}}(\mathbf{x}(b_j)) = L_{\text{Det}}(\mathbf{x}(b_j)) - L_{\text{Graph}}(\mathbf{x}(b_j)). \quad (13)$$

Iterative processing in the above manner approximates an optimal joint detection/decoding algorithm. If the feedback path to the detector is removed then the operational capacity of the system is given by the sum of the independent bit plane capacities. This result is stated in the following theorem.

Theorem 1: The mutual information available to an iterative decoder for the case of a generalized Gaussian channel without feedback between the detector and decoder is given by

$$\tilde{I}(\mathbf{X}; \mathbf{Y}|\mathbf{H}) = \sum_{j=0}^{N_t M_c - 1} I(\mathbf{X}(b_j); \mathbf{Y}|\mathbf{H}), \quad (14)$$

where $\mathbf{X}(b_j)$ calls out the j th bit in the transmit vector \mathbf{X} which carries a total of $N_t M_c$ bits, and \mathbf{X}, \mathbf{Y} are RVs and \mathbf{H} is the given channel matrix.

Proof: Consider the decoding scenario where the *instantaneous* mutual information entering the decoder due to each received vector \mathbf{y}_k is given by \hat{I}_k (here we re-introduce time index k),

$$\begin{aligned} \hat{I}_k &= \sum_{j=0}^{N_t M_c - 1} \hat{I}_k^{b_j}, \quad \text{where} \\ \hat{I}_k^{b_j} &= h(p(\mathbf{x}_k(b_j))) - h(p(\mathbf{x}_k(b_j)|\mathbf{y}_k, \mathbf{H}_k)). \end{aligned} \quad (15)$$

Observe that \hat{I}_k is the sum of the instantaneous marginalized bit reliability mutual informations. The marginalization step reduces the information entering the decoder at each receive time k from the instantaneous vector-wise joint mutual information, $I(\mathbf{x}_k; \mathbf{y}_k|\mathbf{H}_k)$, to the sum of the instantaneous bit-wise

mutual informations, or \hat{I}_k . Explicitly, \hat{I}_k is defined as (for uniform $p(\mathbf{x}_k(b_j))$),

$$\hat{I}_k = \sum_{j=0}^{N_t M_c - 1} \left(\sum_{b_j=\{0,1\}} p(\mathbf{x}_k(b_j)|\mathbf{y}_k, \mathbf{H}_k) \log_2(p(\mathbf{x}_k(b_j)|\mathbf{y}_k, \mathbf{H}_k)) - \log_2(1/2) \right) \quad (16)$$

To complete the proof, observe that the expectation of the instantaneous mutual information converges to the ergodic mutual information of the sum of the individual bit planes in the limit of infinite sample size n ,

$$\frac{1}{n} \sum_{k=0}^{n-1} \hat{I}_k \xrightarrow[n \rightarrow \infty]{} \sum_{j=0}^{N_t M_c - 1} I(\mathbf{X}(b_j); \mathbf{Y}|\mathbf{H}). \quad (17)$$

An exchange in the order of the two outermost summations in (17) left-hand-side and (16) respectively yields this asymptotic result. Hence the average given by (14) is the information available to the *closed* iterative decoder. ■

The results generated in this paper for MIMO systems with dimension greater than 1×1 have each full graph iteration (messages propagate from variable to constraint nodes and back) followed by a detector update.

REFERENCES

- [1] 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.
- [2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.
- [3] R.D. Wesel, X. Liu, and W. Shi, "Trellis codes for periodic erasures," *IEEE Trans. on Comm.*, vol. 48, pp. 938–947, June 2000.
- [4] C. Jones, R. Tian, A. Matache, R. Wesel, and J. Villaseñor, "Robustness of LDPC Codes on Periodic Fading Channels," in *Proceedings of GlobeCom*, Nov 2002.
- [5] C. Köse and R. Wesel, "Universal space-time trellis codes," in *Proceedings of GlobeCom*, Nov 2002.
- [6] C. Köse and R. Wesel, "Universal space-time trellis codes," *To Appear - IEEE Transactions on Information Theory, Special Issue on Space-Time Transmission, Reception, Coding and Signal Design*, Oct 2003.
- [7] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 21–28, Jan. 1962.
- [8] B. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. IEEE Int. Conf. Commun.*, Geneva, Switzerland, pp. 1064–1070, May 1993.
- [9] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, pp. 399–431, Mar. 1999.
- [10] M. Luby, M. Mitzenmacher, A. Shokrollahi, and D. Spielman, "Improved low-density parity-check codes using irregular graphs," *IEEE Trans. Inform. Theory*, vol. 47, pp. 585–598, Feb. 2001.
- [11] T. Richardson, M. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 638–656, Feb. 2001.
- [12] M. Pursely and J. Skinner, "Decoding strategies for turbo product codes in frequency-hop wireless communications," in *Proceedings of ICC*, May 2003.
- [13] J.H. Kang and W.E. Stark, "Turbo codes for coherent fh-ss with partial band interference," in *Proceedings of Military Communications Conference*, Nov 1997.
- [14] A. van Zelst and G. Atwater, "Turbo-blast and its performance," in *Proceedings of VTC*, May 2001.
- [15] A.M. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," in *Proceedings of VTC*, Sept 2000, pp. 473–478.