The Universal Operation of LDPC Codes Over Scalar Fading Channels

Christopher R. Jones
Jet Propulsion Laboratory
Pasadena, CA, USA
crjones@jpl.nasa.gov

Tao Tian
QualComm Corp., San Diego, CA, USA
ttian@qualcomm.com

John Villasenor and Richard Wesel
Electrical Engineering Department, University of California, Los Angeles, CA, USA
{villa, wesel}@ee.ucla.edu

Abstract— Root and Varaiya proved the existence of a code that can communicate reliably over any linear Gaussian channel for which the channel mutual information level exceeds the transmitted rate. This paper provides several examples of scalar (single-input single-output) fading channels and shows that on these channels the performance of Low-Density Parity-Check (LDPC) codes lies in close proximity to the performance limits identified by Root and Varaiya. Specifically, we consider periodic fading channels and partial-band jamming channels. A special case of periodic fading is the variation of signal-to-noise ratio across orthogonal frequency division modulation subchannels. The robustness of LDPC codes to periodic fading and partial-band jamming across parameterizations of these different channels is demonstrated through the consistency of the required mutual information to provide a specified bit error rate. For the periodic fading case, the Gaussian approximation to density evolution has been adapted such that asymptotic threshold measures can be compared to simulated code performance in various periodic fading scenarios.

Index Terms – low-density parity-check code design, orthogonal frequency division modulation, partial-band jamming, robust channel coding, universal codes.

I. INTRODUCTION

Channel coding techniques that approach capacity for a large set of channel realizations, without specializing the transmission to the channel, are desirable from complexity and system-usability points of view. In the discussion that follows, a single code that can communicate reliably near the capacity of many different channels will be called "universal,"

a term introduced in [1][2]. A proof of the existence of codes that exhibit this property was provided by Root and Varaiya in [3]. Root and Varaiya's proof considered the compound channel that occurs when the actual channel is unknown to both transmitter and receiver but belongs to a set of possible channels known to both. Specifically, they proved that a single code exists that can communicate reliably over all channels $\mathbf{a}$ in the set of channels $\mathbf{A}$ at rates arbitrarily close to the compound channel capacity given by,

$$C(\mathbf{A}) = \inf_{\mathbf{a} \in \mathbf{A}} (I(\mathbf{a})), \qquad (1)$$

where $I(\mathbf{a})$ is the mutual information (MI) induced by the transmitted power spectrum on the channel $\mathbf{a}$. For a given desired rate R, reliable transmission is theoretically possible on the set of channels $\mathbf{A} = \{\mathbf{a} : I(\mathbf{a}) > R\}$, such that the MI of every channel in the set is above the transmitted rate. In this way, Root and Varaiya's theorem says that a "universal" code exists that supports rate R over every channel where at least one code (with the specified transmit power spectrum) exists that supports rate-R. That is, if some code works at rate R on a channel, then the rate-R universal code will also work on that channel.

This paper examines the degree to which Low-Density Parity-Check (LDPC) codes designed for the AWGN channel or designed for the AWGN channel with 50% erasures realize the promise of universal operation over channels with periodic fading and partial-band jamming. Gallager proposed LDPC codes in the early 1960s [4]. 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. These codes were too short for the sphere-packing bound to approach Shannon capacity, and the computational resources for longer codes were decades away from being broadly accessible. While Gallager included proofs of asymptotic capacity-approaching performance, the infeasability of long-block-length simulations and system design delayed their practical adoption.

A revival of interest in LDPC codes followed the groundbreaking demonstration by Berrou *et al.* [5] of the capacity-approaching capability of random linear (turbo) codes. Specifically, MacKay [6] re-established interest in LDPC codes during the mid-to-late 1990s. Luby *et al.* [7] formally showed that properly constructed irregular LDPC codes can approach capacity more closely than regular ones. Richardson, Shokrollahi, and Urbanke [8] 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 statistics. This paper characterizes the performance of LDPC codes under two distinct types of channels, periodic fading channels and partial-band jamming channels. Each of these two channel types displays the distinct robustness properties of LDPC codes.

Period-p fading channels have at time i complex input x_i and output $y_i = a_{(i \bmod p)}x_i + n_i$, where n_i is additive white Gaussian noise (AWGN) with variance $N_o/2$ per dimension. The p-element vector $\mathbf{a} = [a_0 \; a_1 \; \ldots \; a_{p-1}]$ consists of complex scalars, which could be the subchannel gains of an Orthogonal Frequency Division Modulation (OFDM) system with p subcarriers. Nearly complete characterization of the performance of LDPC codes on channels with small p can be carried out experimentally and analytically through exhaustive parameterization of the fading vector $\mathbf{a}$. Analytic characterization is provided via the periodic channel extension of Chung's Gaussian approximation [9] to density evolution. Robust operation in channels with longer periods is more difficult to completely characterize, though we do provide several specific examples. To further demonstrate the robustness with greater generality, we turn to the partial-band jamming channel.

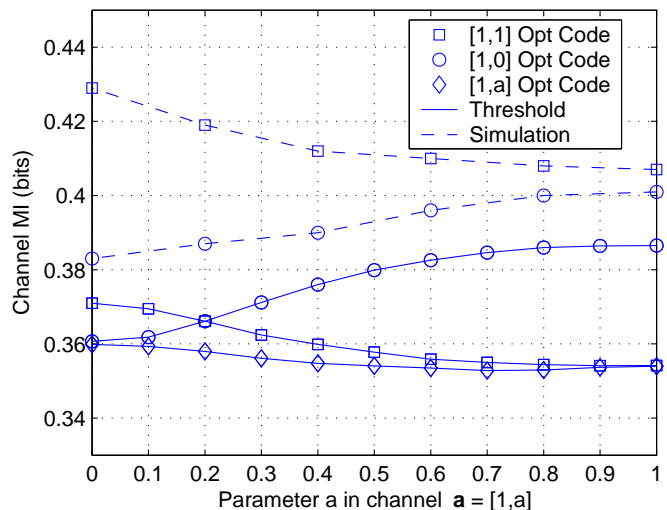Partial-band jamming (PBJ) [10][11] occurs when a fraction of transmitted code symbols has a rela-

Fig. 1. MI thresholds of $\mathbf{a} = [1, 1]$, $\mathbf{a} = [1, 0]$, and $\mathbf{a} = [1, a]$ optimized codes across $\mathbf{a} = [1, a]$ fading (solid lines). All codes are rate-1/3. Simulation results are for BPSK at BER $= 10^{-5}$ for length 15,000 codes realized from the corresponding degree distributions.

tively poor signal-to-noise ratio (SNR) at the receiver (the jammed fraction) and the remaining code symbols experience a relatively good received SNR. It is usually the case that the selection of received signals that incur jamming versus those that do not is random (in adherence to the pre-determined proportions). Simulations performed such that jamming locations are varied from one codeword to the next provide a means for testing a large set of long-period fading channels and of measuring the average fading performance of a code on this set. Good average performance over the set of channels is a necessary, but not sufficient, condition for the code to perform well on every channel individually as the Root and Varaiya result would predict. However, complete characterization of every partial-band jamming channel is beyond our computational ability.

Work on the design and characterization of universal channel codes has been conducted by Wesel *et al.* across several code paradigms for numerous channels. For periodic fading channels, trellis codes were developed in [12], universal serially concatenated turbo codes were developed in [13], and universal LDPC codes (the first appearance of the present work) were developed in [2]. Universal LDPC codes for partial-band jamming and more generally for Gaussian (multi-input multi-output) channels were studied in [14]. In [15] Tse discusses the theory of universal codes for periodic channels (referred to as parallel channels) as well as for

MIMO channels. In particular, universal codes are presented as a solution to the tradeoff between diversity and multiplexing. Furthermore, they note that parallel universal codes may be transformed to MIMO universal codes via DBLAST as was demonstrated by Matache in [16].

We also note the work of Ha *et al.* [17] who have hybridized Gaussian approximated density evolution to include channel erasures. Ha determines optimal variable degree puncturing proportions in order to achieve rate compatibility. In contrast, the codes in this paper employ fixed rates, and our main motivation is to emphasize the robustness of these codes under varying channel conditions. Block fading with block interleaving yields periodic fading. Related work that addresses the general problem of coded performance in block fading channels is given in [18] and [19]. In particular the specifics of LDPC-coded performance under block fading are discussed in [20] and [21]. As part of the universal discussion we present results for several instances of OFDM-like channels. Prior work applying turbo coding to OFDM can be found in [22] and [23]. The authors of [24] considered LDPC-coded OFDM systems with M-PSK signaling.

The next section of the paper provides MI definitions and a design methodology for LDPC codes in the context of periodic scalar fading channels. Section III discusses the design and operation of LDPC codes for the period-2 channel in detail. To demonstrate that robust performance is not limited to the period-2 channel, Section IV provides performance results for an LDPC code on four period-256 channels. A test of average performance on long periodic channels is made in Section V using the partial-band jamming channel. Finally, conclusions from this work are drawn in Section VI.

II. PERIODIC SCALAR FADING CHANNELS

A. *Mutual Information for Periodic Scalar Fading Channels*

The MI of the channel $y = ax + n$, where scale factor a is known at the receiver, can be expressed as,

$$I(X; Y, A) = I(X; A) + I(X; Y|A)$$
$$= I(X; Y|A), \quad (2)$$

where X and A are independent (scalar) RVs. $I(X; Y|A) = E_A[I(X; Y|A = a)]$ is the expectation

that defines the average MI of this channel if A is varying at the receiver. If, however, A is a deterministic constant ($A = a$) then the MI can be computed directly,

$$I(X; aX + N) = I(X; X + \frac{N}{a})$$
$$= h(X + \frac{N}{a}) - h(\frac{N}{a}). \quad (3)$$

The extension of this result to periodic fading follows for a particular instance of the p-element vector **a**,

$$I(\mathbf{a}) = \frac{1}{p} \sum_{i=0}^{p-1} I(X_i; X_i + \frac{N}{a_i}), \quad (4)$$

which can also be used to define the capacity of a frequency selective fading channel in the context of OFDM modulation where i indexes the subcarriers. If both X_i and N are complex Gaussian random variables and each X_i has the same average power, $E[X_i^2] = E_x$, then the average MI per symbol is,

$$I(\mathbf{a}) = \frac{1}{p} \sum_{i=0}^{p-1} \log_2\left(1 + \frac{|a_i|^2 E_x}{2\sigma^2}\right), \quad (5)$$

where σ^2 is the per-dimension variance of the AWGN channel and **a** is a vector of complex scalars. The constant power constraint causes the MI in (5) to be less than the water-filling capacity that can be achieved if the transmitter knows **a**.

Nevertheless, Shannon's basic noisy coding theorem ensures that for each **a** there is a code with fixed symbol power E_x and rate R that achieves reliable communication with R arbitrarily close to $I(\mathbf{a})$. For example, p parallel Gaussian-alphabet codes could be designed with the ith code assigned rate $R_i = \log_2\left(1 + \frac{|a_i|^2 E_x}{2\sigma^2}\right)$. A solution that uses a separate code for each subcarrier is, of course, unattractive as it requires transmitter and receiver to coordinate code selection depending on **a** and, of course, has tremendous complexity for large p. Practically speaking, systems that do have knowledge of **a** at the transmitter can apply bit filling by using symbols from a single code and varying the modulation cardinality on each subcarrier. Asynchronous digital subscriber line (ADSL) systems employ this technique [25].

In this paper we turn to the broader result of Root and Varaiya [3] who proved that a *single*
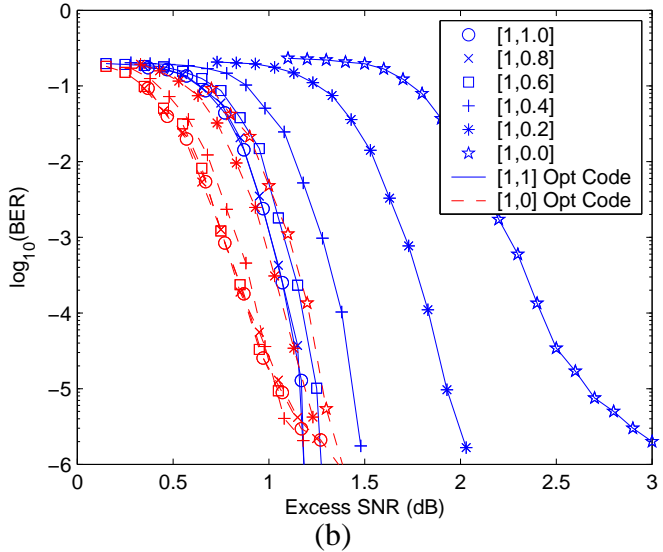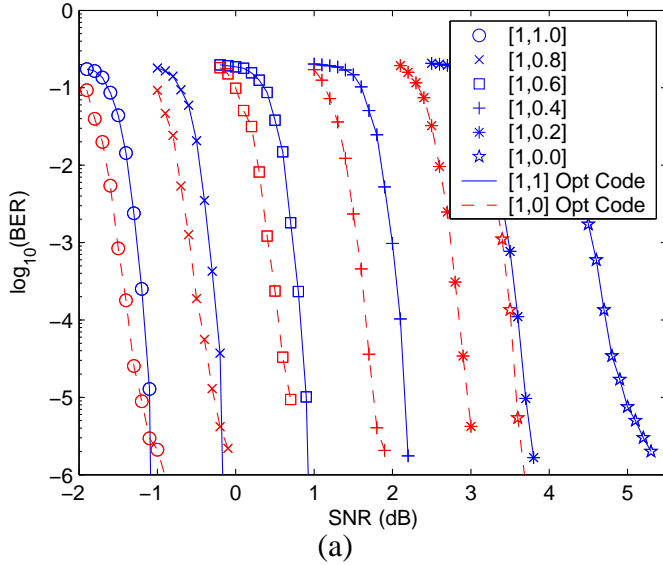
Fig. 2. (a) Code performance on the $\mathbf{a} = [1, a]$ fading channel in terms of SNR. (b) Code performance on the $\mathbf{a} = [1, a]$ fading channel in terms of SNR in excess of channel capacity for each corresponding channel. Dashed lines indicate operation of a code optimized for the $\mathbf{a} = [1, 0]$ channel, solid lines indicate operation of a code optimized for the $\mathbf{a} = [1, 1]$ channel.
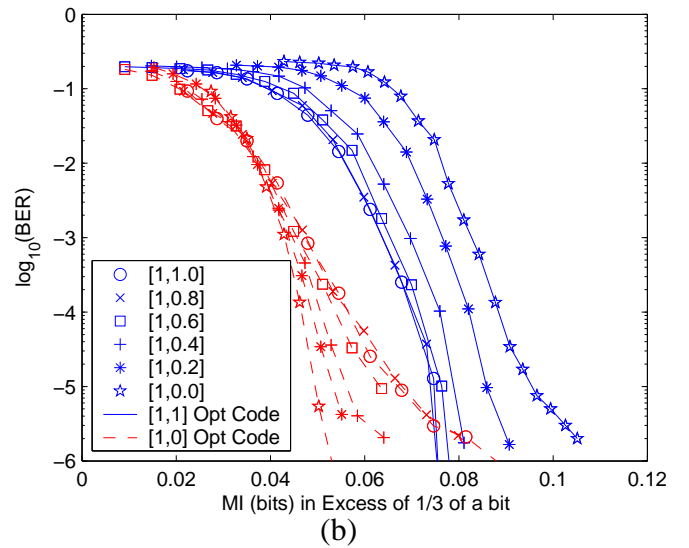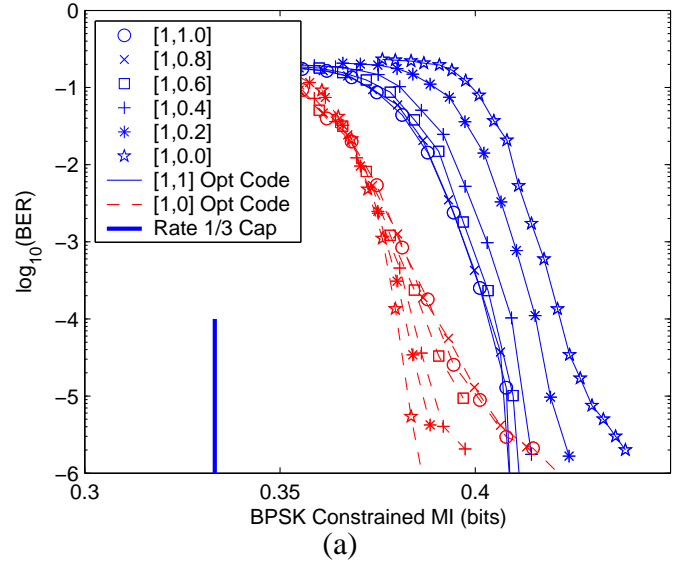
Fig. 3. (a) Code performance on the $\mathbf{a} = [1, a]$ fading channel in terms of MI. (b) Code performance on the $\mathbf{a} = [1, a]$ fading channel in terms of MI in excess of code rate (1/3). Dashed lines indicate operation of a code optimized for the $\mathbf{a} = [1, 0]$ channel, solid lines indicate operation of a code optimized for the $\mathbf{a} = [1, 1]$ channel.

code exists that can communicate reliably at rates arbitrarily close to the compound channel capacity given by (1). While Shannon stated that for each channel there exists a code that provides reliable communication for that channel, Root and Varaiya showed that for a given *set* of channels (collectively this set forms the compound channel) there is *a* code that provides reliable communication on *all* channels within this set. In the succeeding sections,

simulation and density evolution results will show that a single LDPC code can perform with less than 0.1 bits of *excess* MI (per real signaling dimension) for compound channels where the cardinality of the channel set is large. Excess MI is defined as the difference between the channel MI where the desired error probability is achieved and the information transmission rate R. We use excess MI as a performance measure throughout the remainder of the paper. Section III includes and explanation of

why we prefer excess MI rather than excess SNR.

B. LDPC Code Design for period-p Fading Channels

Following the assumption in [9] that the output of an individual variable or constraint node has a Gaussian distribution we extend the derivation of [9] to include a known periodic fading vector **a**. We attempt to preserve the notation originally proposed in [9], however, we introduce and simplify notation at points and also begin the derivation from the constraint side of the graph in order to arrive at a set of recursion relations that are linear in λ (λ (ρ) is the distribution of variable (constraint) node degrees from an edge-wise perspective). This is done to facilitate the use of a linear program for selection of a rate-maximizing λ vector. We begin by stating that the log-likelihood messages entering a degree-k constraint node, denoted v_i, are related to an outgoing message, denoted u, as follows:

$$\tanh\left(\frac{u}{2}\right) = \prod_i \tanh\left(\frac{v_i}{2}\right), \quad (6)$$

if we assume that messages v_i and u are drawn from independent and identically distributed (i.i.d.) Gaussian random variables $v_i \sim \mathcal{N}(m_v, 2m_v)$ and $u \sim \mathcal{N}(m_u, 2m_u)$ (where the subscript is dropped from v's because they are identically distributed) then the expectations of the left and right hand sides of this update relation can be expressed as,

$$E\left[\tanh\frac{u}{2}\right] = E\left[\tanh\frac{v}{2}\right]^{k-1}. \quad (7)$$

This is the expected value of LLR messages departing degree-k constraint nodes (an additional subscript is added to m_u to denote the degree, i.e. $m_{u,k}$). Chung introduced the function $\phi(x) = 1 - E\left[\tanh\frac{n}{2}\right]$ for $n \sim N(x, 2x)$ which permits (7) to be described by,

$$m_{u,k} = \phi^{-1}\left(1 - E\left[\tanh\frac{v}{2}\right]^{k-1}\right). \quad (8)$$

The introduction of parameter $r = 1 - E\left[\tanh\frac{v}{2}\right]$ further simplifies (8) to $m_{u,k} = \phi^{-1}\left(1 - [1 - r]^{k-1}\right)$. The expected value of any given constraint-to-variable message is obtained by averaging over the constraint degree distribution of the code,

$$m_u = \sum_{k=2}^{d_r} \rho_k m_{u,k} = \sum_{k=2}^{d_r} \rho_k \phi^{-1}\left(1 - [1 - r]^{k-1}\right), \quad (9)$$

where ρ_k denotes the fraction of edges connected to degree-k constraint nodes. Degree-i variable nodes have their mean values, $m_{v,i}$, updated in correspondence to the periodic initial means given by $m_{a_j} = \frac{2a_j^2}{\sigma^2}$ and the means of messages arriving from constraint nodes (m_u),

$$m_{v,i}(j) = m_{a_j} + (i-1)m_u, \quad j \in \{0, \ldots, p-1\}. \quad (10)$$

According to the Gaussian approximation, randomly selected edges emanating from variable nodes adhere to the following Gaussian mixture density,

$$f_v = \sum_{j=0}^{p-1} \sum_{i=2}^{d_l} \frac{\lambda_i}{p} \mathcal{N}\left(m_{v,i}(j), 2m_{v,i}(j)\right). \quad (11)$$

Where the outer summation mixes over the periodic fading vector and the inner summation mixes over the variable node edge-wise degree distribution. Literal evaluation of the expectation $E\left[\tanh\frac{v}{2}\right]$ yields,

$$\begin{aligned} E[\tanh\frac{v}{2}] &= \int_{\mathcal{R}} \left(\tanh\frac{x}{2}\right) f_v dx \\ &= 1 - \frac{1}{p}\sum_{j=0}^{p-1}\sum_{i=2}^{d_l} \lambda_i \phi(m_{v,i}(j)). \end{aligned} \quad (12)$$

Parameter $r = 1 - E\left[\tanh\frac{v}{2}\right]$ is expressed in terms of the right hand side of (12); while (10) is substituted with an additional index ℓ to denote iteration.

$$\begin{aligned} r &= \frac{1}{p}\sum_{j=0}^{p-1}\sum_{i=2}^{d_l} \lambda_i \phi\left(m_{v,i}(j)\right) \\ r^{(\ell)} &= \sum_{i=2}^{d_l} \lambda_i \sum_{j=0}^{p-1} \frac{1}{p} h_i\left(m_{a_j}, r^{(\ell-1)}\right), \end{aligned} \quad (13)$$

The above recursion is linear in λ, and uses,

$$h_i\left(m_{a_j}, r^{(\ell-1)}\right) = $$
$$\phi\left(m_{a_j} + (i-1)\sum_{k=2}^{d_r} \rho_k \phi^{-1}\left(1 - \left(1 - r^{(\ell-1)}\right)^{k-1}\right)\right)$$

note that if $m_u^{(\ell)} \to \infty$ (which implies that an asymptotically long code will converge for the given set of initial means $m_{a_0}, \ldots, m_{a_{p-1}}$) iff $r^{(\ell)} \to 0$.

The parameter $r^{(\ell)}$ is initialized such that $r^{(0)} = \frac{1}{p} \sum_{j=0}^{p-1} \phi\left(m_{a_j}\right)$. A linear program (LP) which seeks a rate maximizing λ given a set of initial means $m_{a_0}, \ldots, m_{a_{p-1}}$ and a fixed right degree distribution ρ follows:

$$\text{Min} \left(- \sum \lambda_i / i \right)$$

$$\text{s.t.}$$
$$1^T \lambda = 1$$
$$\lambda_i > 0, \forall i$$
$$r > \sum_{i=2}^{d_l} \lambda_i \sum_{j=0}^{p-1} \frac{1}{p} h_i\left(m_{a_j}, r\right) \tag{14}$$
$$0 < r < \frac{1}{p} \sum_{j=0}^{p-1} \phi\left(m_{a_j}\right).$$

To obtain a code of a given rate (say 1/3) for a particular channel (say $\mathbf{a} = [1, a]$ period-2), the above LP operates as the kernel of a doubly nested loop. The outer loop monotonically increases channel SNR in small increments such that the inner loop is able to find the first (λ, ρ) pair that achieves the rate constraint. For each step of the outer loop, the inner loop sweeps a set of concentrated ρ values (concentrated right degree sequences were proven to be nearly optimal in [9]) and performs the above LP at each step to find the rate maximizing λ values for each specified ρ. The ρ concentration starts at relatively high value (degree 15 for instance) and is decremented in steps of size 0.25 until the achievable rate peaks for the given SNR constraint. Manipulations of the above procedure suggest empirically that the problem of optimizing rate for a given channel SNR constraint is convex in the parameterization of concentrated ρ's (meaning a unique rate-maximizing concentrated ρ distribution exists). The columns in Table I labeled '[1,0]' are the result of the optimization procedure when it is applied to the period-2 $\mathbf{a} = [1, 0]$ channel. Optimization results for the $\mathbf{a} = [1, 1]$ channel are also provided in the table.

Figure 1 provides asymptotic threshold and simulation results for the $\mathbf{a} = [1, 1]$ and $\mathbf{a} = [1, 0]$ optimized codes across $\mathbf{a} = [1, a]$ period-2 fading.

i	λ_i [1,1]	λ_i [1,0]	ρ_i [1,1]	ρ_i [1,0]
2	0.3010	0.3317	-	-
3	0.1474	0.1247	-	-
4	-	0.0416	-	-
5	0.2075	0.1780	0.3000	0.5000
6	0.0097	-	0.7000	0.5000
15	0.3344	0.3239	-	-

TABLE I

DEGREE DISTRIBUTIONS OPTIMIZED USING GAUSSIAN APPROXIMATION TO DENSITY EVOLUTION ADAPTED TO PERIODIC FADING. COLUMNS LABELED $\mathbf{a} = [1, 0]$ INDICATE THE DISTRIBUTION RESULTING FROM OPTIMIZATION FOR THE PERIOD-2 CHANNEL WHERE HALF OF ALL RECEIVED SYMBOLS ARE ERASED. COLUMNS LABELED $\mathbf{a} = [1, 1]$ INDICATE A PERIOD-2 CODE OPTIMIZED FOR AWGN.

As expected, the thresholds of each of these codes is best on the channel for which it was designed and worst on the opposite channel. A third curve (diamond) that provides thresholds for channel-specific designs on each of eleven $\mathbf{a} = [1, a]$ channels is also plotted. Note that for $0 < a < 1$ the channel-specific $[1, a]$ designs do have lower thresholds than the $[1, 1]$ and $[1, 0]$ designs, but these channel specific designs provide a relatively small improvement. Also note that the excess MI required by the simulations is always less that 0.1 bits.

III. LDPC PERFORMANCE ON PERIOD-2 FADING CHANNELS

A. Demonstration of Universal Performance

Figures 2 and 3 present simulation results for $(n, k) = (15000, 5000)$ LDPC codes realized from degree distributions in Table I. As described in Section II, these degree distributions were found by constraining periodic density evolution for the $\mathbf{a} = [1, 1]$ (i.e. Gaussian) and $\mathbf{a} = [1, 0]$ (i.e. Gaussian with every other symbol erased) channels and using a linear program (LP) solver to find the respective minimum threshold rate-1/3 codes with maximum left degree 15.

The codes were conditioned using the Approximate Cycle EMD (ACE) technique developed in [26], where EMD stands for extrinsic message degree. This graph construction technique is particularly attractive for use in conjunction with density evolution as it places no constraints on the underlying degree distribution of the code. The ACE

technique sets out to maximize the multiplicity of "extrinsic" edges connected to short cycles in the graph (where extrinsic edges are edges that do not participate in the cycle). Such a construction improves the stopping set distribution in the graph by increasing the mean stopping set size [27]. Stopping sets can be shown to be formed by closed clusters of cycles [26] (e.g. cycles that are completely interconnected). Ensuring that short cycles have at least a minimum number of extrinsic connections increases the average number of nodes required to form a stopping set. For the codes used in this paper every cycle of length 24 or less has at four or more extrinsic connections.

Figures 2 and 3 describe the performance of these codes in a period-2 channel with $\mathbf{a} = [a_0 \ a_1]$, where $a_0 = 1$ and $a_1 = a$ (signaling is via BPSK modulation). Fig. 2(a) clearly shows that a decrease in a requires an increase in SNR to maintain constant BER. The plot of BER versus absolute SNR, however, does not provide an adequate view of the respective performance on each of the channels. Meaning, we might gather that performance on the $\mathbf{a} = [1,1]$ channel is about 1.2 dB away (at BER $= 10^{-5}$) from binary-input (BI) additive white Gaussian noise (BI-AWGN) capacity (-2.27 dB), however we can less easily determine if the code is performing as well on say the $\mathbf{a} = [1, 0.5]$ channel. To gain a better view from this perspective, Fig. 2(b) plots BER performance vs. *excess* SNR. This plot is constructed by subtracting the SNR required to achieve a MI level of 1/3 of a bit on each respective $\mathbf{a} = [1, a]$ channel from the absolute operating SNRs for each curve in Fig. 2(a). On the plot we see that the performance of the $[1, 0]$ optimized code is clustered in a 0.5 dB range near 1 dB of excess SNR while that of the $[1, 1]$ optimized code varies over roughly a 1.5 dB range. Also note that on this plot the $[1, 0]$ optimized code has better excess SNR performance on the $\mathbf{a} = [1, 1]$ channel than on the $\mathbf{a} = [1, 0]$ channel. These are points that we will return to shortly.

Fig. 3(a) plots BER versus MI using (4) in the context of BPSK constrained signaling. This plot is created by noting that the SNR of each plotted point in Fig. 2(a) has a corresponding MI (under the BPSK and periodic channel constraints). MI in excess of 1/3 of a bit (the code-rate) is given in Fig. 3(b). From Figs. 3(a) and 3(b) we can see that the $[1, 0]$ code performs better at low SNR but
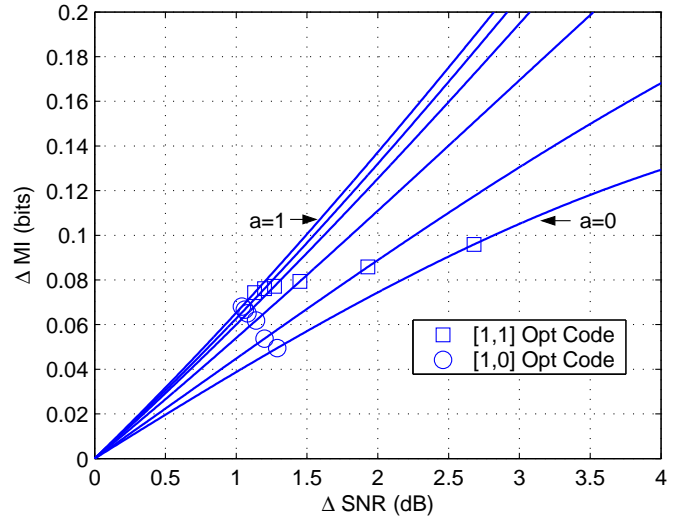

Fig. 4. MI and SNR in excess of that required for 0.33 bits per channel use under BPSK modulation in $\mathbf{a} = [1, a]$ period-2 fading. The operating points of two LDPC codes at BER=10^{-5} are plotted. Curves from left to right indicate excess MI and excess SNR for a = $\{1.0, 0.8, 0.6, 0.4, 0.2, 0.0\}$

has a more severe error floor (i.e. flattening of the BER curve at high SNR) than the $[1, 1]$ code. The trade-off between low-SNR performance and error floor is well-known, but it is not the focus of this paper. Instead, we focus on the universal behavior displayed by both codes.

Note the tightening of the performance of both codes (the $[1, 1]$ code at all BERs and the $[1, 0]$ code in the high BER regime before its error floors appear) when measured in terms of excess MI in Fig. 3(b) as compared to the measurement in terms of excess SNR Fig. 2(b). This tightening across channel variation can be seen as a consequence of removing the bias that is inherent to using SNR as a comparative performance measure across different channel realizations. We emphasize this point in Fig. 4 which provides curves describing the excess SNR required to achieve MI in excess of 1/3 of a bit under BPSK signaling for each of the considered $\mathbf{a} = [1, a]$ channels (six total).

Each of the curves in Fig. 4 describes the relationship between excess MI (on the vertical axis) and excess SNR (on the horizontal axis) for a particular $\mathbf{a} = [1, a]$ channel. The origin on the vertical axis represents 1/3 of a bit of MI and the origin on the horizontal axis represents the SNR level required to produce 1/3 bit of MI on each of the given channels. Note that each channel requires a different absolute SNR to achieve 1/3 bit of MI, but that on this plot

these differing SNRs are the single point of zero excess SNR. Points on the plot away from the origin measure MIs and SNRs in *excess* of 0.33 bits per symbol (bps) capacity-achieving levels.

The differing slopes for each of these channels reveals that there is a choice to be made between excess MI and excess SNR as a metric for comparing performance across channels. For a fixed value of one metric, the other varies with channel selection. For example, a constant excess MI of 0.1 bits could correspond to an excess SNR anywhere from 1.5 dB to 2.75 dB. Conversely, a constant excess SNR of 2 dB could correspond to an excess MI anywhere from 0.0675 bits to 0.14 bits. Given this choice, we select excess MI because it appears in the exponent of the information theoretic proofs of the channel coding theorem [28] while excess SNR does not. We also note that excess MI may be applied to channels where excess SNR has no meaning, such as erasure channels and discrete alphabet channels (such as the common binary symmetric channel).

Excess SNR and MI operating points from Figures 2(b) and 3(b) at BER = 10^{-5} have been superimposed on Fig. 4. From this plot we see that at BER = 10^{-5} the $[1,0]$ code performs a little better than the $[1,1]$ code. However, that performance difference is not the focus of this paper. In fact, both codes are good universal codes. In terms of excess MI, each code performs best on the channel for which it was designed but requires only about 0.02 bits more MI to maintain a BER = 10^{-5} on the opposite channel.

In terms of excess SNR, the $[1,0]$ code performs best on the $[1,1]$ channel and worst on the $[1,0]$ channel. This illustrates the problems inherent with using excess SNR to compare performance across different channels. Also in terms of excess SNR, the $[1,1]$ code performs best on the $\mathbf{a} = [1,1]$ channel, where it requires about 1.1 dB of excess SNR to achieve BER = 10^{-5}, and worst on the $\mathbf{a} = [1,0]$ channel, where it requires about 2.7 dB of excess SNR to achieve BER = 10^{-5}.

This difference of more than 1.5 dB in required excess SNR would not indicate an especially robust code if one considered excess SNR rather than excess MI as the measure of interest. Again note that the total MI variation from worst to best channel is approximately 0.02 bits for both codes. The large excess SNR difference can be explained by observing that the slope of excess MI versus excess

SNR is smaller for the $\mathbf{a} = [1,0]$ channel than for the $\mathbf{a} = [1,1]$ channel. This is because the mutual information of the $\mathbf{a} = [1,0]$ channel is entering saturation. Specifically, the MI of a BPSK, $\mathbf{a} = [1,0]$ channel saturates at 0.5 bits of MI (or 0.167 = 0.5 - 0.33 bits of excess MI).

Fig. 1 uses dashed curves to show the simulated performance of the $[1,1]$ and $[1,0]$ optimized codes at BER = 10^{-5} from data plotted in Fig. 3(b), across the channel parameterization. At this BER the $[1,0]$ optimized code slightly outperforms the $[1,1]$ optimized code on the $\mathbf{a} = [1,1]$ channel. However, the $[1,0]$ optimized code has already entered a flooring region on this channel and if the measurement was taken at BER=10^{-6} (see Fig. 3(b)) then the $[1,1]$ code would exhibit better performance than the $[1,0]$ code for some channels. Error flooring of the $[1,1]$ code on the $\mathbf{a} = [1,0]$ channel is also observed. In general we state that both the $[1,0]$ and $[1,1]$ codes exhibit flooring on their respective opposing channels.

The MI gap between threshold and simulation in Fig. 1 remains nearly constant for each of the two codes as the channel is varied the from $\mathbf{a} = [1,1]$ (Gaussian) to $\mathbf{a} = [1,0]$ (Gaussian with every other symbol erased). In the case of the $[1,1]$ optimized code this gap is approximately $\Delta_{[1,1]} = 0.06$ bits and in the case of the $[1,0]$ code the gap is $\Delta_{[1,0]} = 0.02$ bits. We believe that the outperformance in terms of absolute gap to threshold of the $[1,0]$ versus the $[1,1]$ optimized code is related to the particular code that was generated for each of these cases and also the difference in the achievable matrix conditioning (the approximate cycle conditioning technique of [26] was used to construction the final parity matrix for each code) given the separate input degree distributions. The absolute gap, however, is less consequential than the observation that a given code realization tracks threshold prediction with a fixed offset across channel realizations.

In summary, this section has empirically described the robustness of the designed codes to period-2 fading. When each of the six channels provides MI of at least 0.43 bits (i.e. approximately 0.1 bits above the transmitted rate of 1/3) the codes communicate at or below BER=10^{-5}. Furthermore, the performance variation of each code, on all channels, is approximately 0.02 bits of MI. We say that these two codes are universal for period-2 fading since their MI requirement is essentially
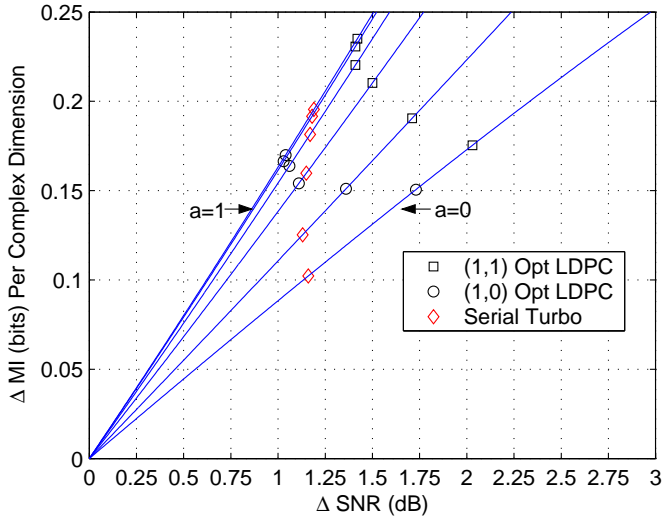
Fig. 5. MI and SNR in excess of that required for 1.0 bit per channel use on 8-PSK in $\mathbf{a} = [1, a]$ period-2 fading. The operating points of two LDPC codes and a serial turbo code are plotted. Each code modulates 10,000 8-PSK symbols per block at BER $= 10^{-5}$. Curves from left to right indicate excess MI and excess SNR for a = {1.0,0.8,0.6,0.4,0.2,0.0}.

constant (within 6% variation from best to worst case excess MI) against channel variation. In the next section we compare the performance of the designed LDPC codes to that of a serially concatenated convolutional code.

B. Comparison with Serially Concatenated Convolutional Codes

LDPC codes represent one of several well known realizations of random linear codes with manageable decoding complexity. Parallel and serially concatenated convolutional (turbo) codes also exhibit capacity approaching performance under AWGN channel conditions. The original work to describe the performance of serial turbo codes in AWGN was performed in [29]. Work that considers serial turbo codes as universal codes has been conducted in [13]. The authors in [13] present an 8-PSK serially concatenated turbo code (SCTC) for period-2 periodic fading. We compare this to $\mathbf{a} = [1, 0]$ and $\mathbf{a} = [1, 1]$ LDPC codes combined with 8-PSK modulation.

Figure 5 simultaneously plots excess MI and excess SNR for six different channel parameterizations under an 8-PSK modulation constraint. Each of the six curves (one for each of the six channels) represents the relationship between excess MI (on the vertical axis) and excess SNR (on the horizontal

axis), as in Fig. 4. In this plot, however, the origin on the vertical axis represents 1.0 bit of MI and the origin on the horizontal axis represents the SNR level required to produce 1.0 bit of MI on each of the given channels.

Also plotted in Fig. 5 are the BER$=10^{-5}$ operating points of $\mathbf{a} = [1, 1]$ and $\mathbf{a} = [1, 0]$ optimized (from Table I) length 30,000, rate-1/3, LDPC codes on these six channels using Gray-labeled 8-PSK modulation (10,000 total channel symbols). Note that no interleaving is used between the code and the 8-PSK mapper, instead the columns of the parity matrix that describes the code are uniformly permuted. The BER$=10^{-5}$ operating points of a rate-1/3 length-10,000 serial turbo code optimized for period-2 fading [13] are also provided. This code was constructed from a 4-state rate-1/2 outer code and a 4-state rate-2/3 linear recursive systematic inner code ([13] provides a description of code generator polynomials).

It is important not to neglect scale in these plots. For instance the difference in the MI performance of the $\mathbf{a} = [1, 0]$ optimized LDPC code across the channels is less than 0.05 bits. The serial turbo code exhibits a consistent excess SNR requirement that is striking, but has a variation in excess MI of 0.1 bits. The wider variation in excess MI of the serial turbo code comes from its exceptionally good performance on channels with low a values. Although the worst case performance is quite similar, the serial turbo code is strictly better than the $\mathbf{a} = [1, 1]$ optimized LDPC code and exhibits performance that is comparable to that of the $\mathbf{a} = [1, 0]$ optimized LDPC code.

IV. LDPC PERFORMANCE ON PERIOD-p CHANNELS

So far we have designed LDPC codes for the two extremes of period-2 fading channels and found that each code did well across the entire spectrum of period-2 fading channels. For the rest of the paper we will see how LDPC codes designed for period-2 fading work on channels with fading period much larger than 2. The only justification for such an unmotivated application is that the codes turn out to do surprisingly well. The point of this exercise is to illustrate that even without careful design for universality, LDPC codes often turn out to approximate universal behavior.
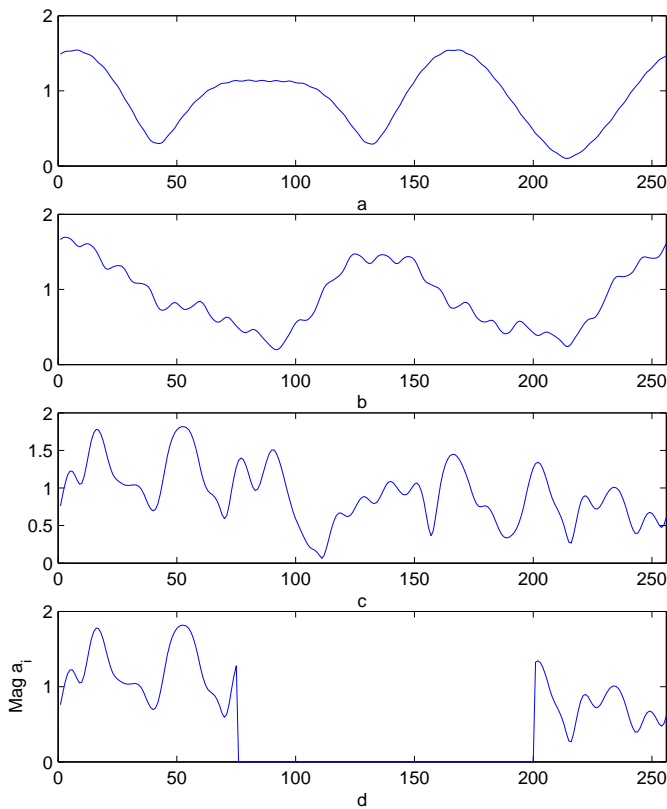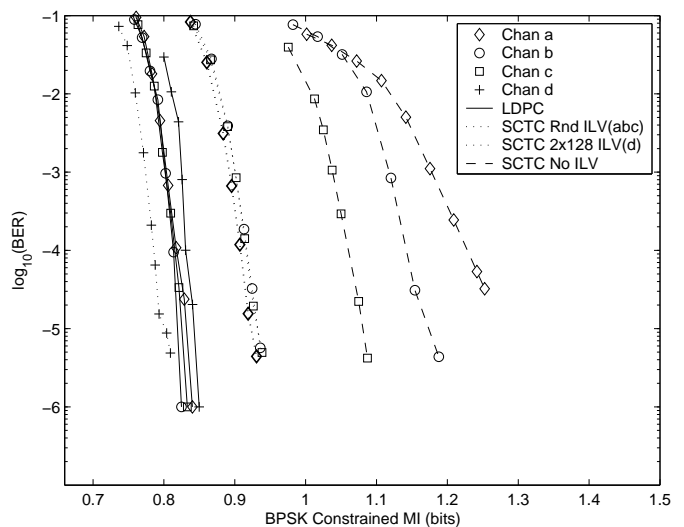
Fig. 6.   Four period-256 fading channels.

Fig. 7.   Code performance on four period-256 fading channels in terms of MI. The SCTC code converges on channel d only for the case 2x128 block interleaving.

Consider the four period-256 channels in Fig. 6. These fading profiles are generated by realizing channels with 4, 8, and 16 multi-path components in the time domain. The time channels are randomly generated with each tap magnitude drawn from a Rayleigh distribution and each tap phase is drawn uniformly. Exponential interarrival times between taps are assumed and an exponentially decaying envelope is imposed on the randomly realized taps. The 256-point fast-fourier-transform (FFT) of each of these channels is 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 MI performance of the rate-1/3, blocklength-15,000 $\mathbf{a} = [1, 1]$ optimized LDPC code on these channels using QPSK modulation, where even(odd) code bits are mapped to I(Q) components, is given in Fig. 7. Half of the subcarriers in channel (d) provide no MI. Thus, for the same total MI, the non-erased subcarriers of channel (d) must provide twice as much MI, on the average, as the subcarriers

of channels (a), (b), and (c). As a result, channel (d) requires a significantly higher SNR for a given level of performance than channels (a), (b), and (c). However, Fig. 7 shows that from the MI point of view the LDPC code works virtually as well on channel (d) as on channels (a,b,c).

It may seem surprising that the code can communicate with 125 of the 256 subcarriers completely erased. However, the supremum of erasure rates for this code on the BEC channel is

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

which has $\varepsilon^* = 0.613$ (equivalent to 157 of 256 subcarriers being erased). Note that ε^* is an asymptotic measure that can only be achieved in the limit of infinite block length. For the length 15,000 code used in this simulation (the $\mathbf{a} = [1, 1]$ optimized code), $\varepsilon^* = 0.59$ (151 of 256 subcarriers) was found via simulation. Thus the minimum 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 SNR is large. However, we emphasize the more remarkable result that the difference in MI required for the code to operate on each of these

four very different period-256 channels is less than 0.025 bits.

As described in [30], an SCTC can perform well on channels (a),(b), and (c) through the use of a random channel interleaver. Fig. 7 shows this performance and the performance on these three channels without channel interleaving for an SCTC with equivalent rate and block length as the presented LDPC code. The SCTC without random channel interleaving performs better on channels with more rapid fading characteristics as the dwell time within a given fade is comparatively shorter on these channels.

Note that the SCTC fails to provide reliable communication at any SNR on channel (d) (the 50% erasure channel) unless the interleaver is "matched" to the channel. The best SCTC performance on channel (d) is given by a 2-by-128 block interleaver resulting in a channel with every other symbol erased which then becomes similar to a period-2 channel. This particular SCTC is designed to have optimal performance under the period-2 $\mathbf{a} = [1, 0]$ channel. Therefore, it is not surprising that the SCTC slightly outperforms the ($\mathbf{a} = [1, 1]$ optimized) LDPC code by 0.04 bits of excess MI. However, such channel and interleaver matching schemes do not follow the theme of universality in a Root and Varaiya sense and LDPC codes seem to be a more reasonable choice than SCTC to achieve universality. In fact, to date we know of no coding methodology, other than LDPC, that can communicate so close to the theorectical limits without the augmentation of a matched channel interleaver on channels such as (d).

V. LDPC PERFORMANCE ON THE PARTIAL-BAND JAMMING CHANNEL

Now we consider the partial-band jamming channel. The PBJ model used in this paper is the same as the one previously described in [10] and [11]. 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 to be consistent with the results in [10]. Channels that are jammed also incur the addition of band-limited white Gaussian noise with power spectral
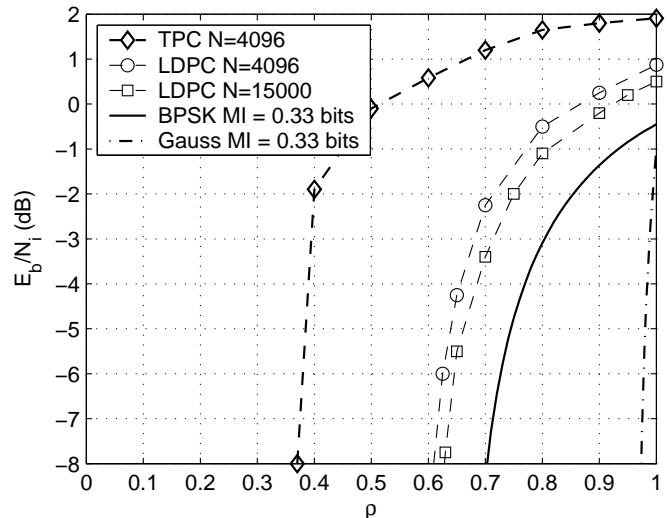

Fig. 8. Performance of rate-1/3 LDPC codes with blocklength 4096 and 15,000 on the partial-band jamming channel compared to a blocklength 4096 turbo product code. E_b/N_i vs. ρ curves that maintain a constant Gaussian signaling capacity (Gauss MI) and BPSK constrained capacity (BPSK MI) of 1/3 of a bit are also displayed. FER = 10^{-3} for the three simulated curves.

density $\rho^{-1}N_I$ over a fraction ρ of the band. The total jamming noise power $\rho(\rho^{-1}N_I) + (1 - \rho)0$ is equal to N_I, and 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 (where all subchannels are jammed), the channel begins to appear much like a standard AWGN channel.

Fig. 8 provides simulation results for two rate-1/3 LDPC codes. Both are realized from the degree sequence of the $\mathbf{a} = [1, 0]$ optimized code described in Table I. The first code has length 4096 (k=1368, n=4096, rate = 0.334) and the second length 15,000. The performance of a length 4096 turbo product code with comparable rate [10] 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
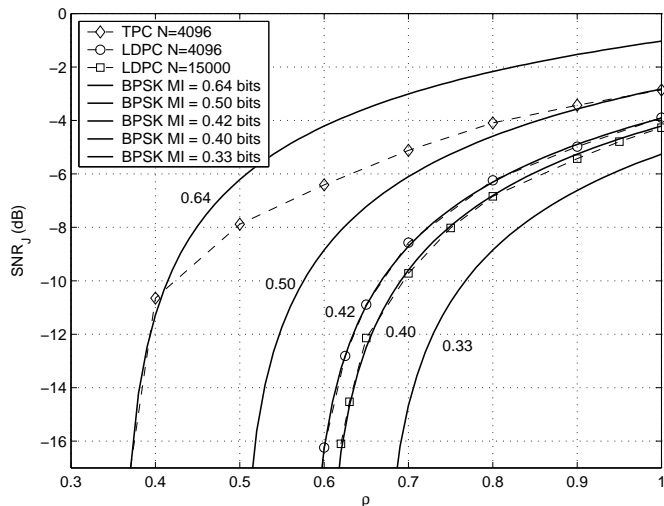
Fig. 9. SNR vs. ρ operating points for FER=10^{-3} of rate-1/3 length 4096 and 15,000 LDPC codes and a rate-1/3 length 4096 TPC code compared to BPSK-constrained constant-MI curves of 0.33, 0.40, 0.42, 0.5, and 0.64 bits.

Fig. 10. Operating points for FER=10^{-3} of a rate-1/2 LDPC code with blocklength 4096 on the partial-band jamming channel. E_b/N_i vs. ρ curves that maintain a constant Gaussian signaling capacity (Gauss MI) of 0.5 bits and BPSK constrained capacities (BPSK MI) of 0.5 and 0.64 bits are also displayed.

results in [10] we have fixed the dwell interval to 32 for the length 4096 code and to 30 for the length 15,000 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 15,000 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 average jamming result for a given code across a parameterization of ρ and E_b/N_o.

Constant MI curves for the partial-band jamming channel are also included in Fig. 8. To compute these curves consider the MI level in partial-band jamming,

$$MI = \rho f\left(SNR_J\right) + (1-\rho)f\left(SNR_{NJ}\right), \quad (15)$$

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 complex Gaussian signaling, $f(x) = \log_2(1+x)$, and for the BPSK constrained case $f(x)$ 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 $\log_2\left(1 + SNR_{NJ}\right)$ is
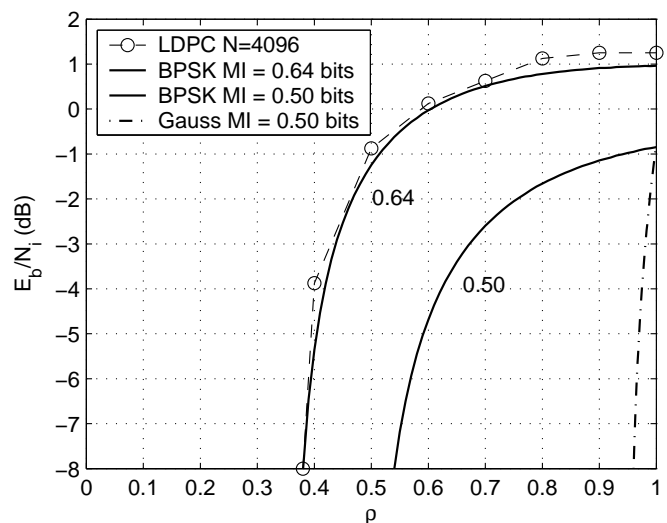
therefore a constant (η) which can be determined via solution to the equation $\eta = \log_2\left(1 + \eta E_b/N_o\right)$, which is $\eta = 9.96$ bits (for $E_b/N_o = 20$ dB). In the BPSK constrained case $MI_{BPSK}\left(SNR_{NJ}\right)$ saturates to $\eta = 1$ bit at this high SNR.

We are interested in values of (ρ, SNR_J) that yield constant levels of MI. 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:

$$SNR_J = \frac{E_s}{\frac{N_I}{\rho} + N_o} = \frac{1}{\frac{1}{\rho R \frac{E_b}{N_I}} + \frac{1}{R \frac{E_b}{N_o}}}$$
$$\frac{E_b}{N_I} = \frac{SNR_J}{\rho R \left(1 - \frac{SNR_J}{R \frac{E_b}{N_o}}\right)}. \quad (16)$$

A large discrepancy can be observed between the BPSK-constrained and Gaussian-signaling MI curves in Fig. 8. This is due primarily to the fact that the non-jammed subchannels provide far more MI (9.96 bits) than the 1 bit maximum provided by BPSK modulation, which in turn implies that with Gaussian signaling, just a small fraction of the subchannels need to be non-jammed for the

expected MI in the channel to reach $1/3$ of a bit. We note that a capacity-approaching system that achieves an average spectral efficiency of $1/3$ of a bit 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 in an absolute E_b/N_I sense than the rate-1/3 BPSK system.

The curves in Fig. 8 represent the FER $= 10^{-3}$ operating points of rate-1/3 systems or the contours of constant MI of $1/3$ of a bit. Fig. 9 plots differen constant MI curves in terms of an absolute SNR_J ordinate. This avoids rate-dependent compression of the ordinate (since $\frac{E_b}{N_I} \simeq SNR_J - 10 \log_{10}(\rho R)$) and allows for distinguishable contours of constant MI. In the plot, each constant contour is labeled with its represented level of MI.

When the partial-band jamming channel provides 0.4 bits of MI, the length 15,000 LDPC code (the $\mathbf{a} = [1, 0]$ optimized code) operates with an FER $= 10^{-3}$ at all values of ρ. The same can be stated for the length 4096 code when the channel supports 0.42 bits of MI. Restating this result, the length 15,000 code provides reliable communication when the excess MI in the channel is $\Delta MI = 0.4 - 1/3$, or roughly 0.067 bits and similarly the length-4096 code requires 0.087 bits of excess MI. The closeness with which the simulated performance tracks constant contours of MI in the figure provides clear empirical evidence of code robustness across an extremely broad range of channels.

To demonstrate that not all codes exhibit this property, we have also plotted the performance of the turbo product code in Fig. 9. When $\rho = 1$ this code requires 0.5 bits of MI to achieve FER 10^{-3}. At lower ρ values, however, the code requires up to 0.64 bits of MI to maintain constant FER. We note that the MI level (e.g. 0.4 bits) required for the LDPC codes to achieve reliable communication in this channel is comparable to the levels required by the period-2 and period-256 fading channels. Finally, since the majority of channels in this paper have operating modes that provide less than 0.5 bits of MI (the $[1, 0]$ channel and OFDM channel d), we have been restricted (in the context of robust coding) to the use of codes with rate less than 1/2. Since the PBJ channel provides no such obstacle, Fig. 10 plots the performance of a rate-1/2 code. The degree distribution for the selected code was designed using Gaussian approximation for the $\mathbf{a} = [1, 1]$ channel

with maximum left and right degrees less than 16 and was built with an information length $k = 2048$. The plot shows that when the PBJ channel provides roughly 0.64 bits of MI that the code operates reliably at FER $= 10^{-3}$.

VI. CONCLUSION

In this paper, we have taken an MI, rather than an SNR, approach to measuring code performance over periodic Gaussian and partial-band jamming channels. Root and Varaiya 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 less than the smallest MI of all channels in the set. It has been shown for a quantized spread of all period-2 channels, under BPSK and 8-PSK modulation, and for several arbitrarily selected period-256 channels, under QPSK modulation, that LDPC codes provide a practical example of Root and Varaiya's promise of "universal" codes. We have described and used periodic density evolution to design codes matched to channels and to determine the thresholds of existing codes across parameterizations of the $\mathbf{a} = [1, a]$ channel. Root and Varaiya's theorem applies to any particular instance of the partial-band jamming channel. However, we have averaged the performance of a given code across many thousands of instances of the PBJ channel in order to test the universality of the codes across a large sampling of channels. While it is true that the performance of the codes on some particular PBJ channel may have been poor (and such an event would go undetected due to the averaging process), we have nevertheless shown that the average excess MI requirements of the codes on this channel are very similar to those of the codes on the periodic fading channels.

REFERENCES

[1] C. Köse and R. Wesel, "Universal space-time trellis codes," in *Proceedings of GlobeCom*, Nov 2002.
[2] C. Jones, T. Tian, A. Matache, R. D. Wesel, and J. Villasenor, "Robustness of LDPC codes on periodic fading channels," in *Proceedings of GlobeCom*, Nov 2002.
[3] W. Root and P. Varaiya, "Capacity of classes of Gaussian channels," *SIAM J. Appl. Math*, vol. 16, no. 6, pp. 1350–1393, Nov. 1968.
[4] R. Gallager, "Low-density parity-check codes," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 21–28, Jan 1962.
[5] C. 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.

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

[7] M. Luby, M. Mitzenmacher, M. Shokrollahi, and D. Spielman, "Improved low-density parity-check codes using irregular graphs," *IEEE Trans. on Inform. Theory*, vol. 47, pp. 569–584, Feb 2001.

[8] T. Richardson, M. 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.

[9] S. Chung, T. Richardson, and R. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a gaussian approximation," *IEEE Trans. on Inform. Theory*, vol. 47, pp. 657–670, Feb 2001.

[10] M. Pursely and J. Skinner, "Decoding strategies for turbo product codes in frequency-hop wireless communications," in *Proceedings of ICC*, May 2003.

[11] J. Kang and W. Stark, "Turbo codes for coherent FH-SS with partial band interference," in *Proceedings of Military Communications Conference*, Nov. 1997.

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

[13] C. Kose, W. Weng, and R. Wesel, "Serially concatenated trellis coded modulation for the compound periodic erasures channel," in *Proceedings of ICC*, May 2003.

[14] C. Jones, A. Matache, T. Tian, J. Villasenor, and R. D. Wesel, "The universality of LDPC codes on wireless channels," in *Proceedings of MilCom*, Oct 2003.

[15] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[16] A. Matache and R. D. Wesel, "Universal trellis codes for diagonally layered space-time systems," *IEEE Transactions on Signal Processing, Special Issue on MIMO Wireless*, vol. 51, no. 11, pp. 2773–2783, Nov 2003.

[17] J. Ha, K. Jaehong, and S. W. McLaughlin, "Rate-compatible puncturing of low-density parity-check codes," *IEEE Trans. on Comm.*, vol. 50, pp. 2824–2836, Nov 2004.

[18] R. Knopp and P. Humblet, "On coding for block fading channels," *IEEE Trans. on Inform. Theory*, vol. 46, no. 1, pp. 189–205, Jan. 2000.

[19] M. Chiani, A. Conti, and V. Tralli, "On the design of convolutional codes over block fading channels," in *Proceedings of International Symposium on Information Theory*, June 2000, pp. 25–30.

[20] M. Chiani, A. Conti, and A. Ventura, "Evaluation of low-density parity-check codes over block fading channels," in *Proceedings of International Conference on Communications*, June 2000, pp. 18–22.

[21] J. Xiaowei, A. Eckford, and T. Fuja, "Analysis of LDPC decoding for correlated and uncorrelated block fading channels," in *Proceedings of International Symposium on Information Theory*, June 2004, p. 375.

[22] Y. Li and S. Sun, "Design of bit-interleaved turbo-coded modulation in OFDM-based wireless LAN systems," in *4th International Workshop on Mobile and Wireless Communications Networks*, Sept. 2002, pp. 549–553.

[23] J. Kim, "Turbo-coded OFDM system for a mobile satellite broadcasting system," in *Proceedings of Vehicular Technology Conference*, May 2001, pp. 2945–2949.

[24] H. Futaki and T. Ohtsuki, "Low-density parity-check (LDPC) coded OFDM systems with M-PSK," in *Proceedings of Vehicular Technology Conference*, Oct. 2002, pp. 1035–1039.

[25] P. Chow, J. Cioffi, and J. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. on Comm.*, vol. 43, no. 2, pp. 773–775, April 1995.

[26] T. Tian, C. Jones, J. Villasenor, and R. D. Wesel, "Characterization and selective avoidance of cycles in irregular LDPC code construction," *IEEE Trans. on Comm.*, vol. 52, pp. 1242–1247, Aug. 2004.

[27] A. Ramamoorthy and R. D. Wesel, "Expansion properties of generalized ACE codes," in *Proceedings of 42nd Allerton Conference on Communication, Control and Computing 2004*, 2004.

[28] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley and Sons, 1991.

[29] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: performance analysis, design, and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 44, pp. 909–926, May 1998.

[30] W. Weng, C. Kose, and R. Wesel, "Serially concatenated trellis coded modulation for the compound periodic erasures channel," *In Submission to the IEEE Transactions on Communications*.