A Class of High-Rate, Low-Complexity, Well-Structured LDPC Codes from Combinatorial Designs and their Applications on ISI Channels

Jing Li (Tiffany)*
Electrical Engineering Dept, Texas A&M University
College Station, TX 77843-3128
email: JingLi@ee.tamu.edu

Erozan Kurtas
Seagate Technology, Seagate Research, 1251 Waterfront Place
Pittsburgh, PA 15222-4215
email: Erozan.M.Kurtas@seagate.com

ABSTRACT

We present a systematic construction of a class of high-rate, well-structured low density parity check (LDPC) codes based on combinatorial designs. We show that the proposed $(2\rho, \rho^2, \rho, 2, \{0,1\})$-design results in a class of $(2,\rho)$-regular LDPC codes, which are systematic, quasi-cyclic, free of length-4 and length-6 cycles, linear-time encodable and decodable, and which have high code rates of $R=(1-\frac{1}{\rho})^2$. Analysis from the maximum likelihood perspective shows that the distance spectrum of the proposed LDPC codes are better than that of the Gallager ensemble codes for the same code length and rate. The proposed codes are then applied to several inter-symbol interference channels, where 2 high code rates and 3 block sizes from short to mediumn are evaluated. For best performance gain, the i.i.d. capacity is computed to choose the best precoder and iterative decoding and equalization is performed The proposed LDPC codes demonstrate performance that is slightly (but noticeably) better than an average random LDPC code of column weight 3. Unlike random codes, well-structured LDPC codes can lend themselves to a very low-complexity implementation for high-speed applications.

KEY WORDS

low density parity check (LDPC) codes, combinatorial design, distance spectrum, inter-symbol interference (ISI), iterative decoding and equalization (IDE), turbo equalization, binary precoding

1 Introduction

Fifty years after Claude Shannon computed the capacity limit of memoryless channels, we have finally constructed practical coding schemes that perform close to the capacity limit. The breakthrough of turbo and low density parity check (LDPC) codes [1] [2] have revitalized the coding research by introducing new concepts and techniques like code graph, random interleaving and iterative decoding. Numerous simulations have demonstrated their remarkable performance: within a fraction of a dB from the Shannon limit on additive white Gaussian noise (AWGN) channels for fairly large block sizes. However, it is fair to say that for full deployment in commercial systems, complexity and implementation issues remain to be addressed.

The major criticism of turbo codes have been its high decoding complexity where each decoding iteration of turbo codes involves two *a posteriori* probability (APP) decoding units matched to the component convolutional codes. An APP decoder usually implements the BCJR algorithm, which is about 4 times as complex as the Viterbi algorithm for the same trellis code.

For LDPC codes, although both the encoding and the decoding can be made linear (or near-linear) in the block length [3], hardware implementation has not been easy due to their random structure. In his original proposal, Gallager defined LDPC codes as a class of block codes which use random, sparse parity check matrices H to denote the relations between code bits and parity-check sums [1]. Although research work has indicated that (properly constrained) randomness is important for near capacity performance, codes with structure and regularity are far from being abandoned. In addition to the random construction of LDPC codes (i.e. to randomly allocate bits in the parity check matrix subject to certain constraints), like bit filling and/or optimization of bit/check degree profiles using density evolution, systematic constructions are also being proposed, which include the approaches from combinatorial designs [5] [6] [8], finite geometries [7], Ramanujan graph [9] and lattice designs [8]. It has been shown that in some cases (especially for short block sizes and/or high code rates, like those used for digital recording systems), structured LDPC codes might be more advantageous than random LDPC codes with comparable performance and with more implementable structure [6] [5] [7].

In general, an LDPC code is represented using either a parity check matrix H or its corresponding Tanner graph which is a bipartite graph using *bit* nodes and *check* nodes to represent the columns and the rows of the H matrix and using inter-connecting *edges* to represent the relations between bits and checks. Major parameters for an LDPC code include the column weight (or the bit node degree) γ, the

*Jing Li is currently with the Electrical and Computer Engineering Dept. in Lehigh Unversity, Bethlehem, PA 18015.

row weight (or the check node degree) ρ [1], and the *girth* (defined as the length of the shortest cycle in the Tanner graph). An LDPC code is said to be (γ, ρ)-*regular* if all columns have weight γ and all rows have weight ρ. The girth is important to LDPC codes because the existing decoder is an iterative message-passing decoder whose efficiency is adversely affected by the existence of short cycles.

In this work, we introduce a class of high-rate, well-structured LDPC codes from $(2\rho, \rho^2, \rho, 2, \{0,1\})$ combinatorial designs. A key merit of considering combinatorial designs for LDPC codes, in addition to their regular and thus easily-implementable structure, is that they are free of length-4 cycles [4] [6] [5] [8]. In fact, for the specific class studied in this work, length-6 cycles are also systematically avoided. Furthermore, the proposed LDPC codes are quasi-cyclic and have low uniform column weight of only 2, which considerably simplifies the encoding/decoding process (recall that the decoding complexity per bit is proportional to the average column weight). The proposed codes are rich in high rate (and short block size) region. We discuss the properties of this class of codes not only from the viewpoints of LDPC codes, but also from that of the irregular repeat accumulate (IRA) codes [14] and turbo product codes (TPC) [15] to facilitate the understanding. Further, we examine the distance spectrum of the proposed LDPC codes and compare it to the Gallager's original ensemble. We show that the proposed structured LDPC codes are better than the ensemble average of random codes from the maximum likelihood (ML) perspective (i.e. assuming an optimal decoder is used).

We then move on to evaluate the performance of this class of regular LDPC codes on inter-symbol interference (ISI) channels. Inter-symbol interference channels are an important channel model, common in both wireless communications and digital data recording systems. There is currently great deal of interest in using the concatenation of an LDPC code with an ISI channel (see for example [10]-[13] and the references therein). A detector/decoder of such systems typically comprises of two parts: a detector/decoder matched to the (inner) ISI channel, and a decoder matched to the (outer) LDPC code. Inspired by the ideas and practice of serial turbo codes, current research trends have focused on joint detection/decoding of the two devices in an iterative fashion to achieve additional coding gains. This is known as turbo equalization or iterative decoding and equalization (IDE). To maximize the coding gains, we also explore binary precoding in conjunction with IDE in this study. We use partial response class IV (PR4) family channels as an example, and demonstrate how i.i.d. capacity can be computed to facilitate the choice of the precoder. Simulation results show that 4-5 dB gains can be achieved over uncoded ISI systems (using a maximum likelihood sequence detector (MLSD) such as the Viterbi detector). We observe that ISI systems using our combinatorially designed LDPC codes outperform (slightly but noticeably) those using randomly constructed LDPC codes (which have column weight 3 and which are manually removed of length-4 cycles). Unlike random codes, the proposed codes are well-structured and, hence, lend a low-complexity implementation for high code rate and high speed applications.

The rest of the paper is organized as follows. Section II presents the preliminary of combinatorial designs followed by the discussion and analysis of the proposed $(2\rho, \rho^2, \rho, 2, \{0,1\})$-designed LDPC codes. Section III discusses their application on (binary precoded) ISI channels. Section IV reports simulation results and Section V concludes the paper.

[1] For irregular LDPC codes which do not constrain uniform row or column weights, degree profiles, $\rho(i)$ and $\gamma(i)$, are usually used to describe the distributions of row weights and column weights.

2 Combinatorial Designs for LDPC Codes

2.1 Preliminary

Borrowing terms from [6], we present here the definitions and some properties of combinatorial design which are used in the work.

Definition : (Combinatorial design)

[1] A *combinatorial design* is an arrangement of a set of m *points* into n subsets, called *blocks*, which satisfy certain regularity constraints.

[2] The *incidence* matrix of a combinatorial design gives the (0,1)-matrix (of dimensionality $n \times m$) which has a row for each point v and a column for each block B, and $(v, B) = 1$ iff point v is incident with block B.

[3] The *covalency* λ_{v_1,v_2} of two points v_1 and v_2 is the number of blocks that contain both of them.

[4] A design is said to be *regular* if the number of points contained in each block (denoted as γ) is the same for every block and the number of blocks each point is incident with (denoted as ρ) is the same for every point.

[5] A design is said to be *balanced* if the covalency λ_{v_1,v_2} of the point pair (v_1, v_2) is the same for all pairs. A regular and balanced design can be denoted as a $(m, n, \rho, \gamma, \lambda)$-design, where $m\rho = n\gamma$.

It follows from the above definitions that a combinatorial design with favorable constraints can define a binary LDPC code, where the transpose of the incidence matrix can serve as the parity check matrix H. Apparently, a *point* in a combinatorial design corresponds to a *check* node in the Tanner graph or a row in H, a *block* corresponds to a *bit* node or a column in H. The H matrix has m rows/checks, n columns/bits (the codeword length), with row weight ρ and column weight γ. Covalency $\lambda < 2$ guarantees the absence of length-4 cycles in the Tanner graph. Further note that all the rows/checks of the resulting H may not be independent and, hence, the actual rate of the LDPC code depends on the rank of the H matrix

$$R = 1 - \frac{rank(H)}{n} \leq 1 - \frac{m}{n}. \quad (1)$$

An example is shown in Fig. 1, where $m = 8$ points are grouped in $n = 16$ blocks with each point incident with 4 blocks and each block containing 2 points, such that $B_1 = (v_1, v_2)$, $B_2 = (v_1, v_4)$, $\cdots$, $B_{16} = (v_7, v_8)$. Fig. 1(a) shows the combinatorial design (where a line connecting 2 points is used to denote a group containing two points), (b) shows the parity check matrix H of the resulting LDPC code, and (c) shows the corresponding Tanner graph. It can be easily verified that this combinatorial graph has covalency $\lambda = \{0, 1\}$ for all pairs of points and, hence, is free of length-4 cycles. In fact, as can be seen from the Tanner graph (Fig. 1(c)), this combinatorial design has also eliminated length-6 cycles.

Some popular classes of combinatorial designs that have already been studied for generating LDPC codes are *Steiner systems* or $(m, n, \rho, \gamma, 1)$-designs [4] and *Kirkman triple systems* (KTS) or $(m, n, \rho, 3, \{0, 1\})$-designs which are resolvable Steiner triple systems (STS) [6] [5] [8]. Other designs from lattice [8] and Ramanujan graphs [9] are also proposed. These systematically-designed LDPC codes share the same desirable properties of simplicity in construction and regularity in code structure. Some of these codes have also been evaluated for use on ISI channels, where simulations have shown that 3 dB gains over uncoded systems are generally achievable. Below we present a new class of LDPC codes from combinatorial design which is structurally much simpler and more implementable than a random LDPC code, and whose performance on ISI channels are (slightly) better than an average random LDPC code.

2.2 LDPC Codes from $(m, n, \rho, 2, \{0, 1\})$-Design

Recall that the previous combinatorial designs of LDPC codes have primarily focused on triple systems, like STS and KTS. This is probably because that in his original work, Gallager proved that the average minimum distance of the ensemble of (γ, ρ)-regular LDPC code will increase linearly in the block size, so long as column weight $\gamma \geq 3$. In this work, however, we present a design with $\gamma = 2$ and evaluate its performance on ISI channels. To ensure the absence of length-4 cycles, we have constrained $\lambda = \{0, 1\}$. The motivation of choosing $\gamma = 2$ is two-folded. First, the decoding complexity is proportional to the (average) column weight and, hence, low column weight leads to low decoding complexity (reducing column weight from 3 to 2 saves 1/3 of the decoding effort!). Second, while most of the LDPC codes studied have column weight around 3 or 4, it does not follow that performance will be inferior beyond this region, see, for example, linear time encodable LDPC codes (whose average column weight < 3) [16] and EG- and PG-LDPC codes (whose average column weight $>> 4$) [7]. The point here is that the relation between the weight and the performance of an LDPC code is multi-faceted, and needs to be judged case by case. In the proposed construction, by choosing a low column weight of only 2, we have reduced the possibility of creating (short) cycles which are undesirable for iterative probabilistic decoding (and are more so for high-rate codes). On the other hand, this raises a concern for insufficient constraints, since each bit participates in only 2 checks. We note that although the performance of the proposed LDPC codes are slightly worse than those proposed by MacKay [2] on AWGN channels, the LDPC codes proposed in this work show advantage both in performance and in complexity for ISI channels (with proper binary precoding). This is primarily because, as we mentioned above, through the use of iterative decoding and equalization, the overall performance is not shaped by the outer code (i.e. LDPC) alone, but rather by each and every component in the system which hopefully interacts with and complements each other in harmony. The term "in harmony" is used vaguely here, and its sufficient conditions are hard to define, but at least one necessary condition, that the outer code have a minimum distance ≥ 3 (in order to assure interleaving gain), is generously satisfied in the proposed design (since the girth of the proposed codes is at least 8).

The construction we present is as follows. For a set of points containing even number of points, denoted as $V = \{v_1, v_2, \cdots, v_{2\rho-1}, v_{2\rho}\}$ (ρ is an integer), a block B will contain a pair of points, (v_i, v_j), from V where

$$i = (j + k) \bmod 2\rho, \;\; \forall \mathrm{k} = 1, 3, 5, \cdots 2\lceil\frac{\rho}{2}\rceil - 1. \quad (2)$$

The example of $\rho = 4$ is shown in Fig. 1. As mentioned before, this $(m, n, \rho, 2, \{0, 1\})$-design results in a $(2, \rho)$-regular LDPC code. We have:

Lemma 1: The $(2, \rho)$-regular LDPC codes based on the proposed $(m, n, \rho, 2, \{0, 1\})$-design (see (2)) have the following properties:

1. It exists for all even integer m, such that the codeword size is $n = (m/2)^2$.
2. The H matrix derived from the design presents a quasi-cyclic LDPC code, that is, shifting a valid codeword leftward or rightward by $m/2$ bits produces another valid codeword. However, they are not M-sequences since the codeword length n is a multiple of $m/2$.
3. The resulting LDPC codes are linear time encodable and linear time decodable.
4. The resulting LDPC codes are systematic codes.
5. The girth of the Tanner graph is 8.
6. The rank of the H matrix is $(m - 1)$ and, hence, the codeword rate is $R = (1 - 2/m)^2 = (1 - 1/\sqrt{n})^2$.

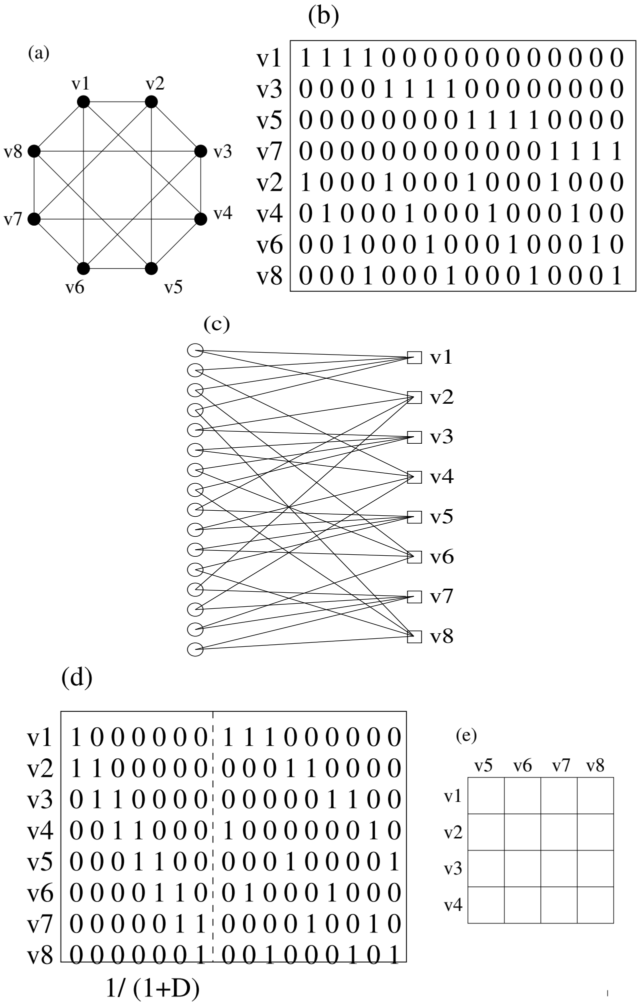

Figure 1. (a) $(m, n, \rho, 2, \{0, 1\})$ Combinatorial design, where $\rho = 4$, $m = 2\rho = 8$, and $n = \rho^2 = 16$; (b) H matrix of the resulting LDPC code; (c) Corresponding Tanner graph; (d) A form of linear time encodable LDPC codes; (e) A form of turbo product codes.

7. This class of codes can be specified using only one parameter ρ, that is, the design is a $(2\rho, \rho^2, \rho, 2, \{0, 1\})$-design, and the LDPC code is $(2, \rho)$-regular with code length $n = \rho^2$ and rate $R = (1 - 1/\rho)^2$.

The above properties can be conveniently verified. Here are some comments.

Property 1 assures the richness of this design as compared to some other combinatorial designs. Specifically, we have demonstrated the available code choices fit for magnetic recording applications (rate > 0.85 and block size < 4096 bits, a sector in a hard disk driver) in Fig. 2.

Property 2 eliminates the necessity to store the generator matrix, since the encoding can be implemented with a linear shift register with feedback connections based on its generator polynomial.

Whereas LDPC codes are generally linear time decodable, i.e., the decoding complexity per bit is only proportional to the average column weight (and the number of iterations), linear-time encodability does not come for free for most of the random constructions [3]. For the proposed construction, linear time encodability of Property 3 can be directly inferred from Property 2, and we will get back to this later as we interpret this class of LDPC codes in the form of irregular repeat accumulate codes [14].

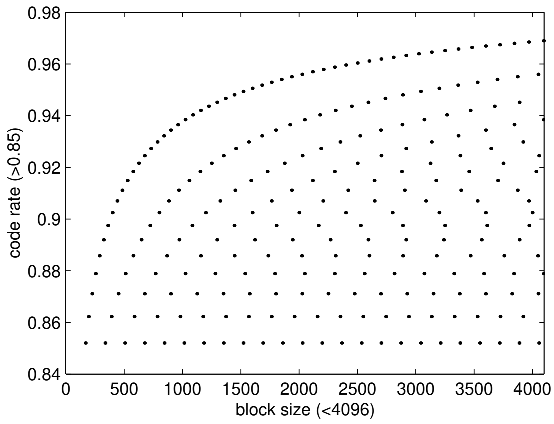

Figure 2. Possible choices of codes for use in magnetic recording systems. ($R > 0.85$ and $N < 4096$)

Table 1. The code rate and code length of short block and high rate LDPC codes from $(2\rho, \rho^2, \rho, 2, \{0, 1\})$-design

ρ	13	14	15	$\cdots$	62	63	64
R	.852	.862	.871	$\cdots$	.968	.969	.969
n	169	196	225	$\cdots$	3844	3969	4096

Although almost all LDPC codes are systematic, Property 4 will become more intuitive as we later interpret it in the form of turbo product codes.

Property 5 guarantees a minimum distance of at least 4, which, as mentioned above, ensures the interleaving gain when serially concatenated with a recursive inner code or a (precoded) ISI channel.

Property 6 shows that the proposed codes tend to be a class of high rate codes, which are intrinsically fitful for applications like high-density digital data recording systems. Tab. 1 presents a list of rates and lengths of some high rate and short block LDPC codes resulting from the above design.

2.3 Alternative Viewpoints

As mentioned before, in addition to the perspectives of regular LDPC codes, the proposed $(2\rho, \rho^2, \rho, 2, \{0, 1\})$-design can also be interpreted as a form of irregular repeat accumulate codes [14] and turbo product codes [15]. Before we provide alternative viewpoints, we note that even disregarding the quasi-cyclic property, this class of codes are linear time encodable. In fact, we have:

Lemma 2: For an LDPC code specified by an $m \times n$ parity check matrix H, if there are at least (m-1) weight-2 columns in H which do not complete a cycle among them, then encoding can be performed in linear time using the parity check matrix H.

The proof is quite simple. In fact, if there are at least $(m-1)$ weight-2 columns in H among which there is no cycle, then we can reorder the rows and columns of the H matrix such that the submatrix formed by these $(m-1)$ columns are lower triangular. An example is shown in Fig. 1(d) which is a reorder of the same H matrix in Fig. 1(b). Apparently this becomes a form of linear time encodable LDPC codes which were initially presented in [16]. Furthermore, a closer observation of Fig. 1(d) reveals that it is also a form of irregular repeat accumulate codes proposed in [14]. The lower triangular part (left) of the H matrix does nothing but plays the role of an accumulator $1/(1 \oplus D)$, and the other part (right) functions to repeat and form checks of the data bits. In other words, this H matrix defines an encoding procedure where bits are first repeated and formed checks (in an irregular fashion), and then the parity sequence is passed through an accumulator. This is exactly what a systematic IRA code does, and this (again) confirms the linear time encodability of the code.

It is quite interesting that this same class of LDPC codes can also be viewed as a special type of turbo product codes (or block turbo codes (BTC)) [15]. Turbo product codes are typically described as arrays of codewords from systematic linear block codes concatenated in a multi-dimensional form. As can be seen from the example in Fig. 1(e), the parity check matrix H has an equivalent form of a 2-dimensional turbo product code where each row and each column satisfy a single-parity check. This reflects the regularity in the structure of the proposed combinatorial design.

2.4 Distance Spectrum Analysis

In his original construction [1], Gallager specified a class of (γ, ρ)-regular LDPC codes whose $m \times n$ parity check matrix H can be split into γ submatrices of dimensionality $\frac{m}{\gamma} \times n$ each, with each submatrix of column weight 1 and row weight ρ (denote such a submatrix as $H_{(1,\rho)}$). Gallager used the ensemble of this class of codes to derive many useful analytical results. It can be seen from Fig. 1(b) that the quasi-cyclic $(2, \rho)$-regular LDPC codes constructed from the proposed $(2\rho, \rho^2, \rho, 2, \{0, 1\})$-design actually falls into this special class of Gallager codes (call it "Gallager ensemble") for the case of $\gamma = 2$.

For Gallager ensemble codes with random construction, the expectation (or the average) of the output weight enumerator function (OWEF) can be derived fairly easily [1]. Considering Gallager ensemble of (γ, ρ)-regular codes with code length n, the parity check matrix, $H_{(\gamma,\rho)}$, constitutes of γ submatrices, $H_{(1,\rho)}$, each of which has output weight enumerator function

$$A_{(1,\rho)}(w) = \underbrace{B(w) * B(w) * \cdots * B(w)}_{\rho}, \tag{3}$$

where $*$ denotes convolution operation and

$$B(w) = \begin{cases} \binom{\rho}{w}, & w \text{ even}, \\ 0, & w \text{ odd}. \end{cases} \tag{4}$$

The average OWEF of (γ, ρ)-regular Gallager ensemble is thus given by

$$A^{gallager}_{(\gamma,\rho)}(w) = A_{(1,\rho)} \cdot \left(\frac{A_{(1,\rho)}}{\binom{n}{w}} \right)^{\gamma-1}. \tag{5}$$

If we substitute $\gamma = 2$, we could compute the average distance spectrum of $(2, \rho)$-regular Gallager ensemble (random) LDPC codes from the coefficients of (5). We mentioned that the proposed LDPC codes from combinatorial design form a subset of Gallager ensemble. To be more precise, the proposed $(2\rho, \rho^2, \rho, 2, \{0, 1\})$-design leads to a deterministic $(2, \rho)$-regular LDPC code which is an instance in Gallager ensemble if the relevant order of the bits in the codeword is ignored. In fact, the exact output weight enumerator function of the proposed $(2\rho, \rho^2, \rho, 2, \{0, 1\})$-design can be derived from the perspective of turbo product codes [17]

$$A^{proposed}_{(2,\rho)}(w) = \frac{1}{2^\rho} \sum_{\alpha=0}^{\rho} \binom{\rho}{\alpha} \left(\sum_{\substack{\beta=0, \\ \beta\ even}}^{\rho} P(\beta, \alpha, \rho) w^\beta \right)^\rho, \tag{6}$$

where

$$P(\beta, \alpha, \rho) = \sum_{k=0}^{\beta} (-1)^k \binom{\alpha}{k} \binom{\rho - \alpha}{m - k}. \tag{7}$$

In Tab. 2, we compare the output weight enumerators, A_w, of the Gallager ensemble $(2, \rho)$-regular (random) LDPC codes (see (5)) and the proposed combinatorial designed $(2, \rho)$-regular (structured) LDPC codes (see (6)). For the purpose of clarity, we represent the numbers in logarithmic scale. It is obvious that the proposed structured LDPC codes are better than the ensemble average of random codes, with fewer codewords at the low weight end of the distance spectrum. Put another way, even with this rigidly constrained structure, the proposed LDPC codes are above average from the maximum likelihood perspective (with optimal decoding).

3 Application on ISI Channels

3.1 System Description

We evaluate the proposed LDPC codes on inter-symbol interference channels. A block diagram of the LDPC-coded

Table 2. Comparing the Output Weight Enumerator of Gallager Ensemble (Random) LDPC Codes and the Proposed (Structured) Combinatorial Designed LDPC Codes (ρ=16, n=256, Logarithm Scale)

Output weight w	Gallager $\log_{10}(A_w)$	Proposed $\log_{10}(A_w)$
2	2.0529	-
4	4.2471	4.1584
6	6.4509	6.2745
8	8.6383	8.5254
10	10.7988	10.7074
12	12.9265	12.8570
14	15.0175	14.9651
16	17.0689	17.0300
18	19.0781	19.0500
20	21.0431	21.0232
22	22.9620	22.9483
24	24.8333	24.8242
26	26.6558	26.6499
28	28.4286	28.4249
30	30.1510	30.1488

ISI channel and a matching iterative decoder/equalizer is shown in Fig. 3. The ISI channel in the presence of additive white Gaussian noise is interpreted as a rate-1, nonlinear trellis code with binary input and real-value output. We insert a random interleaver in-between the LDPC code and the PR channel, where the interleaver size is an integer multiple of the LDPC codeword length. As shown in the block diagram, data sequence is first encoded by the outer LDPC code, then passed into a random interleaver followed by a precoder (if it exists), and then BPSK modulated ($1 \rightarrow +1, 0 \rightarrow -1$) before finally being put onto the ISI channel. For the inner code (i.e. the (precoded) ISI channel), a MAP decoder implementing the BCJR algorithm is used, and for the outer LDPC code, the message-passing algorithm is used. Overall, a soft-in soft-out (SISO) iterative decoding/equalization is exploited to jointly detect and decode the system. The turbo principle, which is an iterative *a posteriori* probability estimation/detection with successively refined *a priori* information, is strictly followed to help the iterative process to approximate the optimal solution. During each iteration of the message flow, out-bound information (i.e. the extrinsic information) from a local processor (either the inner decoder or the outer decoder) is constrained to have the least correlation with the in-bound information (i.e. the *a priori* information) to this processor.

3.2 Binary Precoding and i.i.d. Capacity

That a binary precoder has a direct impact on the performance and the convergence of the iterative process of coded ISI channels is no longer news. Several analytical tools

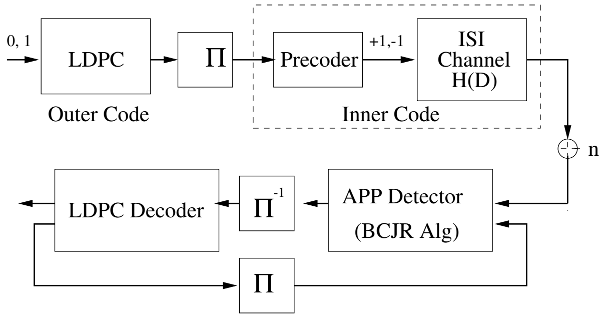

Figure 3. System model for LDPC-coded PRML channels.

have been used to facilitate the choice of the best precoder, like the threshold computed using density evolution [12] and the extrinsic information transfer (EXIT) chart [18]. In this work, we compute i.i.d. capacity to facilitate the choice of precoder for coded ISI channels.

I.i.d. capacity is a useful tool to evaluate the asymptotic performance given an infinite block/interleaver size and perfect interleaving. For a given code rate R, the proposed LDPC code has fixed code length of $n = 1/(1-\sqrt{R})^2$. The perfect interleaver which is of infinite length will scramble infinite number of blocks of LDPC codewords and, hence, although each codeword always has girth 8, the overall scrambled coded bits (from infinite blocks of codewords) will appear cycle-free or mutually independent. This validates the application of i.i.d. capacity to the aforementioned system.

To compute i.i.d. capacity, we need to evaluate log-likelihood ratio (LLR) information as it evolves through the decoder/detector. Each (sub-)decoder transforms input LLRs to output LLRs where LLRs are assumed to follow a Gaussian distribution that is governed by a single parameter, μ, the mean value of the LLRs:

$$f_d^(l) = \frac{1}{\sqrt{4\pi\mu}} exp\left(-\frac{(l-d\mu)^2}{4\mu}\right) = \mathcal{N}(\mu, 2\mu), \qquad (8)$$

where $d = \pm 1$ is BPSK modulated signal put onto the channel. In the initial iteration, the inner channel decoder/detector takes the LLRs from the ISI channel with mean value $\mu_{ch} = 4R/N_0$, that is, $f_d^{(ch)} \sim \mathcal{N}(4dR/N_0, 8R/N_0)$, where R is the code rate and N_0 is the one-sided power spectral density of the Gaussian noise. In subsequent runs, the inner decoder takes the LLRs from the outer decoder and vice versa.

For (inner) ISI channels, APP decoder/detector based on the BCJR algorithm is used to examine the message flow, and for (outer) LDPC codes, the massage-passing decoder is used. The i.i.d. capacity of the system is the maximum of the mutual information between input and output LLRs:

$$I = \frac{1}{2}\sum_{d=\pm 1}\int_{-\infty}^{\infty} f_d^{(ldpc)}(\eta)\log_2 \frac{2f_d^{(ldpc)}(\eta)}{f_{+1}^{ch}(\eta)+f_{-1}^{(ch)}(\eta)} d\eta, \qquad (9)$$

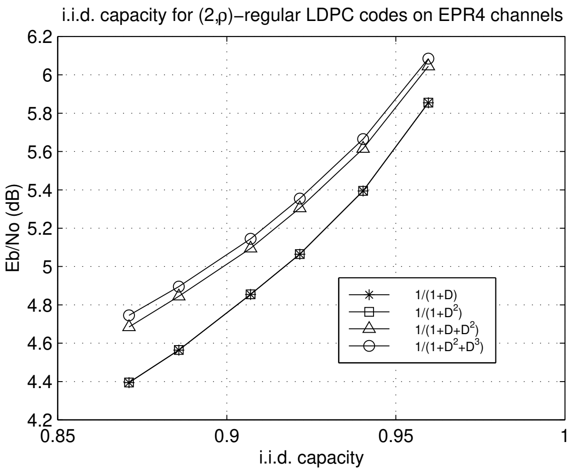

Figure 4. I.i.d. capacity of $(2\rho, \rho^2, \rho, 2, \{0, 1\})$-designed LDPC codes on PRML channels with different precoding.

$$= \quad 1 - \int_{\infty}^{\infty} f_{+1}^{(ldpc)}(\eta) \cdot \log_2(1 + e^{-\eta}) d\eta, \qquad (10)$$

where $d = \pm 1$, $f_d^{(ch)}(\eta)$ and $f_d^{(ldpc)}(\eta)$ are the pdf's of the input LLRs (from the ISI channel) to the system and the output LLRs from the LDPC code after joint decoding/detection, respectively. The second equation holds because both the symmetry condition [19], $f_{-1}(\eta) = f_{+1}(-\eta)$, and the consistency condition, $f_d(\eta) = f_d(-\eta)e^{d\eta}$, are satisfied for LLRs (from the channel and LDPC code) in message passing.

As an example, we compute the i.i.d. capacity of the proposed $(2, \rho)$-LDPC codes on EPR4 channels (whose channel response is $H(D) = 0.5+0.5D-0.5D^2-0.5D^3$) to evaluate several binary precoders. It is obvious from Fig. 4 that $1/(1 \oplus D \oplus D^2)$ and $1/(1 \oplus D^2 \oplus D^3)$ are worse precoders than $1/(1 \oplus D)$ and $1/(1 \oplus D^2)$. We note that although the latter two demonstrate almost identical i.i.d. capacity which implies the same asymptotic performances, simulations with short block sizes and finite complexity reveal a marginal better performance of $1/(1 \oplus D^2)$ over $1/(1 \oplus D)$. It is nevertheless fair to say that i.i.d capacity provides a convenient means to evaluate a precoder, and that similar treatment can be used for an arbitrary ISI channel to evaluate an arbitrary outer code and/or an arbitrary precoder.

4 Results

We report in this section the computer simulated performance of the proposed regular LDPC codes from $(2\rho, \rho^2, \rho, 2, \{0, 1\})$-designs on ISI channels. Two high code rates of $R = 0.88$ and 0.94, and three interleaver sizes of 1024, 2048 and 4096 bits are considered, respectively. The channels we investigate have responses $H(D) = 0.7071 - 0.7071D$ (known as PR4 channel), $H(D) = 0.5 + 0.5D - 0.5D^2 - 0.5D^3$ (EPR4 channel)

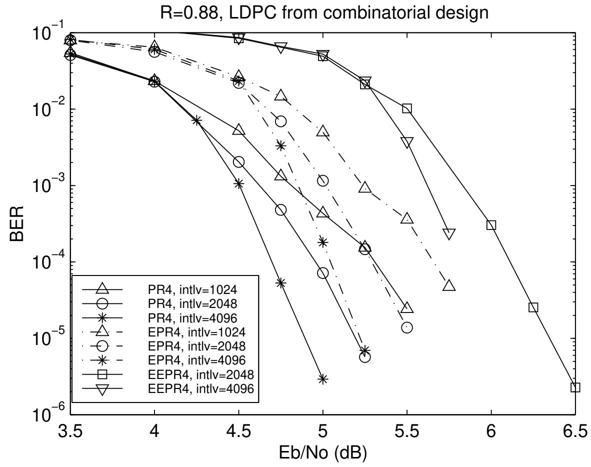

Figure 5. BER performance of rate 0.88 LDPC codes from combinatorial designs on ISI channels. PR4, EPR4 and E^{2}PR4 channel model. Interleaver size 1024, 2048 and 4096 bits.

and $H(D) = 0.3162 + 0.6325D - 0.6325D^3 - 0.3162D^4$ (E^{2}PR4 channel), respectively. Unless otherwise stated, all curves shown are after 8 turbo iterations.

Fig. 5 plots the bit error rate (BER) curves of rate 0.88 $(32, 256, 16, 2, \{0, 1\})$-designed LDPC codes on PR4, EPR4, and E^{2}PR4 channels with precoder $1/(1 \oplus D^2)$. Although not shown here, for uncoded system with MLSD to reach BER of 10^{-5}, it requires about 10.25 dB for PR4, 10.5 dB for EPR4, and 10.8 dB for E^{2}PR4 channels, respectively. Hence, 4-5 dB gains are achieved with these high-rate, low-complexity LDPC codes of relatively short block sizes. Further, interleaving gain phenomenon is also observed. By increasing the interleaver size from 1K to 4K, additional 0.5 dB gain is obtained.

Fig. 6 shows the BER performance of a rate 0.88 and a rate 0.94 LDPC codes from $(32, 256, 16, 2, \{0, 1\})$-design and $(64, 1024, 32, 2, \{0, 1\})$-design, respectively (solid lines). For comparison purpose, the performance of a typical LDPC code from the random construction which has column weight 3 and which is removed of length-4 cycles is also plotted (dashed lines)[2]. We see that the proposed structured LDPC codes actually outperform (slightly but noticeably) the random LDPC code, in addition to simpler and more implementable structure!

5 Conclusion

We present in this work a class of high-rate, regular LDPC codes from combinatorial $(2\rho, \rho^2, \rho, 2, \{0, 1\})$-design. As opposed to the prevalent practice of random(-like) con-

[2]The PR channel is not precoded for random LDPC codes. This is because the random LDPC codes investigated here have pretty large minimum distances which perform better without precoding than precoded (about 0.5 dB difference). Hence, the comparison here is fair, for it compares the best cases in both codes.

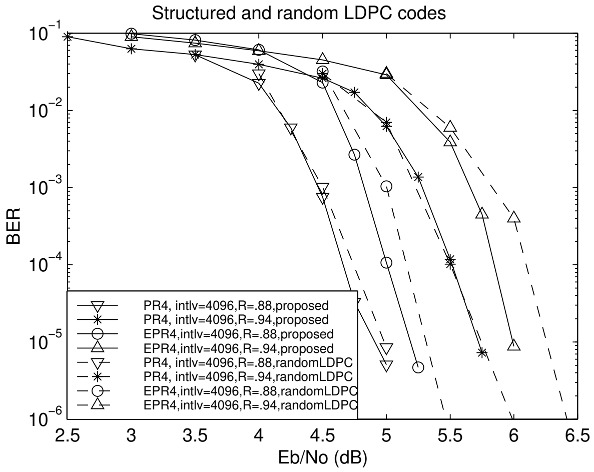

Figure 6. Performance comparison of combinatorially designed LDPC codes and random LDPC codes on ISI channels. Proposed LDPC codes are $(2, 16)$-regular with rate 0.88 and $(2, 32)$-regular with rate 0.94. Random LDPC codes have uniform column weight of 3 and concentrated row weight, and have no length-4 cycles.

struction, our systematic approach results in codes that are very well-defined and balanced in structure and hence, unlike random codes, can lend themselves to very low-complexity implementation for high speed applications. Analysis of code distance spectrum and evaluation of the performance on ISI channels reveal encouraging evidence for deployment of the proposed codes in future high-density digital data recording devices and high-speed wireless communication systems.

References

[1] R. G. Gallager, *Low-density parity-check codes*, MIT press, Cambridge, MA, 1963.

[2] `http://www.inference.phy.cam.ac.uk/mackay/CodesFiles.html`

[3] T. Richardson, and R. Urbanke, "Efficient encoding of low-density parity-check codes," *IEEE Trans. Inform. Theory*, Feb. 2001,

[4] D. J. MacKay and M. C. Davey, "Evaluation of Gallager codes for short block length and high rate applications," *Proc. of the IMA Workshop on Codes, System and Graphical Models*, 1999.

[5] B. Vasic, "Structured iteratively decodable codes based on Steiner systems and their application in Magnetic recording," *Proc GLOBECOM*, San Antonio, TX, Nov. 2001, 770-974.

[6] S. J. Johnson, and S. R. Weller, "Construction of low-density parity-check codes from Kirkman Triple Systems," *Proc BLOBECOM*, San Antonio, Nov. 2001.

[7] Y. Kou, S. Lin, and M. P. C. Fossorier, "Low-density parity-check codes based on finite geometries: a rediscovery and new results," *IEEE Trans. Inform. Theory*, Vol 47, Nov 2001, 2711-2736.

[8] Erozan M. Kurtas, Bane Vasic, and Alexander V. Kuznetsov, " Design and Analysis of Low Density Parity Check Codes for Applications to Perpendicular Recording Channels," **invited chapter** for *The Wiley Encyclopedia of Telecommunications*, 2002.

[9] I. J. Rosenthal, and P. Vontobel, "Construction of LDPC codes using Ramanujan graphs and ideas from Margulis," *Proc Intl. Symp. on Inform. Theory*, Washington D.C., 2001.

[10] J. Fan, A. Friedmann, E. Kurtas, and S. W. McLaughlin, "Low density parity check codes for partial response channels," *Allerton Conf. on Commun., Control and computing*, Urbanan, IL, Oct 1999.

[11] T. Oenning, and J. Moon, "A low-density generator matrix interpretation of parallel concatenated single-bit parity codes," *Proc The Magnetic Recording Conf.*, Santa Clara, CA, Aug. 2000.

[12] J. Li, K. R. Narayanan, E. Kurtas, and C. N. Georghiades, "On the performance of high-rate TPC/SPC codes and LDPC codes over partial response channels," *IEEE Trans. Commun*, May, 2002, 723-734.

[13] H. Song, Richard, M. Todd, and J. R. Cruz, "Low density parity check codes for magnetic recording channels," *Digest of Intl. Magnetic Conf.*, 2000, GD-07.

[14] H. Jin, A. Khandekar and R. McEliece, "Irregular repeat-accumulate codes," *2nd Intl. Symp. on Turbo Codes and Related Topics*, Brest, France, Sept 2000.

[15] P. Elias, "Error-free coding," *IRE Trans. Inform. Theory*, vol. IT-4, Sept., 1954, 29-37.

[16] Ping Li, W. K. Leung, and Nam Phamdo, "Low density parity check codes with semi-random parity check matrix," *Electronics Letters,* vol. 35, Jan. 1999, 38-39.

[17] G. Caire, and C. Taricco, "Weight distribution and performance of the iterated product of single-parity-check codes," *Proc. GLOBECOM Conf.*, 1994, 206-211.

[18] M. Tuchler, C. Weib, E. Eleftheriou, A. Dholakia, and J. Hagenauer, "Application of high-rate tail-biting codes to generalized partial response channels," *Proc. GLOBECOM*, San Antonio, TX, Nov. 2001, 2966-2971.

[19] S.-Y. Chung, R. Urbanke and T. J. Richardson, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," *IEEE Trans. Inform. Theory*, Feb. 2001, 657-670.