RZ 3456 (# 93786) 10/28/02
Electrical Engineering 13 pages

Research Report

Capacity-Approaching Codes for the Magnetic Recording Channel

Ajay Dholakia[1], Evangelos Eleftheriou[1], Marc Fossorier[2] and Thomas Mittelholzer[1]

[1]IBM Research
Zurich Research Laboratory
8803 Rüschlikon
Switzerland

[2]Dept. Electrical Engineering
University of Hawaii at Manoa
Honolulu, HI 96822
USA

Capacity-Approaching Codes for the Magnetic Recording Channel

Ajay Dholakia[†], Evangelos Eleftheriou[†], Marc Fossorier[‡], and Thomas Mittelholzer[†]
† IBM Research, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland
‡ Dept. Electrical Engineering, University of Hawaii at Manoa, Honolulu, HI 96822, USA

Abstract

Digital signal processing and coding are increasingly being recognized as a cost-efficient approach in achieving substantial areal density gains while preserving the high reliability of disk drives, although historically advances in head and media technologies have been the main driving force behind areal density growth. The recent advances in capacity-approaching codes hold the promise to push the areal density to the ultimate limit. In this article the various configurations regarding the interplay between soft detection and soft decoding through an iterative process, as it applies to the magnetic recording channel, are presented. In particular, the state of the art in turbo and turbo-like coding, including LDPC coding is reviewed, and the serial concatenation of these coding schemes with inner generalized PR channels in a turbo-equalization structure is described.

1 Introduction

The advent of computers and the explosive growth of Internet spurred the development of vast and reliable data storage. One of the most important parameters in storage is cost per bit. In the past decade, the areal density of hard-disk drives has increased at an unprecedented rate of 60% compound annual growth. Together with the advances in VLSI technology, the introduction of the magnetoresistive (MR) recording head and the enhancements in signal processing and coding for data and servo detection have resulted in substantial reductions in the cost of storage. Nowadays it is cheaper to store information on a hard-disk drive than on paper. Note that all these areal-density improvements have been achieved while maintaining the stringent on-track error-rate requirements of 10^{-8} to 10^{-9} before outer error-correction coding.

The general similarity of the write/read process in a magnetic recording system to transmission and reception in communication systems has led to the application of advanced signal-processing and coding techniques to the magnetic recording channel. The classical communications channel perspective can be adopted not only to study equalization, detection and inner and outer coding strategies but also to estimate the ultimate information-theoretic limits for recording.

Over the past ten years, several digital signal-processing and coding advances have been incorporated into commercial hard-disk drives to increase linear density and improve the error-rate performance. Partial-response class-4 (PR4) shaping along with maximum-likelihood sequence detection was introduced in the early 1990s. This partial-response maximum-likelihood (PRML) system architecture offered a 30–50% increase in linear density over the conventional peak detection systems [1]. The quest for even higher normalized linear recording densities necessitated the departure from integer coefficient partial-response (PR) channels to generalized PR channels with noninteger coefficients. In the late 1990s, this led to the introduction of the so-called noise-predictive maximum-likelihood (NPML) system architecture [2]. Both the PRML and the NPML architecture require the use of suitable high-rate constrained or modulation inner codes to facilitate timing recovery and gain control, as well as to limit the path memory length of the sequence detectors. More recently, these constrained codes have been combined with high-rate multiparity block codes to improve the bit-error-rate performance of the inner channel even further. Currently, a 16-state NPML detector for a generalized PR channel with a first-order null at dc followed by a post-processor for soft decoding of a combined multiparity/constrained code represents the de-facto industry standard [3, 4].

Outer error-correcting codes have played an important role in achieving high data integrity in recording systems. In disk drives, interleaved byte-oriented Reed-Solomon (RS) coding is currently the standard outer coding scheme. In the future, advanced schemes based on turbo [5] or low-density parity-check (LDPC) coding [6, 7] and iterative detection/decoding promise to approach the ultimate information-theoretic limit, which is known as the capacity of the magnetic recording channel. In particular, parallel concatenated convolutional codes (PCCC), serially concatenated convolutional codes (SCCC), and LDPC codes have been shown to approach the capacity of the additive white Gaussian noise (AWGN) channel.[1] Application of such advanced coding schemes to magnetic recording systems necessitates the design of high-rate codes. It is known that for overall rates of less than 0.8, there are diminishing returns in performance owing to the more severe code-rate penalty on the magnetic recording channel than on the AWGN channel [8]. Furthermore, with data rates in excess of 1 Gbits per second (Gb/s), high-rate codes are again favorable because the internal clock speed scales inversely with the code rate. In contrast, low-rate codes, although more powerful, would require that the encoded data symbols are recorded at very high normalized densities, which in practice would render the recording process unreliable because of thermal instabilities of the media. In principle, turbo, turbo-like, and LDPC coding schemes hold the promise to push the areal density of magnetic recording systems to the ultimate limit for specific magnetic components. It is by now well understood that despite the existing sector-size constraints of hard-disk

[1] Although in the literature the term "turbo codes" is used extensively, it is more appropriate to use the qualifier "turbo" for the decoding process and refer to the codes as PCCC or SCCC.

drives, which limit the block length of the code, and the high code-rate requirements, rather simple iterative decoding schemes can improve the performance of the current state-of-the-art magnetic recording systems by more than 2 dB at medium and high normalized recording densities, and bring it to within approximately 1.5 dB of the theoretical limit [8, 9]. Clearly, the achievable gain also depends on the number of iterations performed in the detection/decoding process, which may be limited due to latency restrictions.

There are many approaches to applying the "turbo principle" to a dispersive intersymbol interference channel such as the magnetic recording channel. This article describes the various configurations regarding the interplay between soft detection and soft decoding through an iterative process. In this context, the state of the art in turbo and turbo-like coding, including LDPC coding, is reviewed, and the serial concatenation of these coding structures with inner generalized PR channels is presented.

2 Magnetic Recording System

In hard-disk drives, information is stored on the medium by means of a write head after having passed through a preprocessing circuit that transforms the binary data into a rectangular write-current waveform with two possible amplitudes. To avoid nonlinearities associated with the magnetization process, the applied magnetic field always exceeds a critical value so that the magnetization is completely polarized in one of two possible directions. This approach to storing information is known as *saturation recording*. The information can then be retrieved by using a read head and the associated data-detection circuitry.

2.1 A Communications Channel View

Write head, magnetic medium, and read head resemble the transmitter, physical channel, and receiver, respectively, in a digital communications system. Thus, the magnetic recording system can be considered as a communications channel in which binary information is stored at one point in time and retrieved at another. This allows the use of various techniques developed over the years by communications engineers and information theorists to maximize the recording and retrieval rates as well as the density of the stored information per square inch of disk space. Crucial to a successful application of signal-processing and coding techniques in hard-disk drives is an accurate understanding and characterization of the physical recording process.

The magnetic recording channel is often modeled by using the so-called Lorentzian pulse,

$$s(t) = \frac{A}{1 + \left(\frac{2t}{\mathrm{PW50}}\right)^2} , \tag{1}$$

where A is the peak amplitude and PW50 denotes the width of the pulse at its 50% amplitude level. This analytical model approximates the step response of the read head to an isolated positive transition of the write current, which corresponds to a transition in the direction of magnetization. The response of the write-head/medium/read-head assembly to a rectangular current pulse, i.e., to a positive transition followed by a negative transition separated by T seconds, where $1/T$ denotes the recording rate, is called *pulse response* and represents the effective impulse response $h(t) = s(t) - s(t-T)$ of the magnetic recording channel. Furthermore, the unitless quantity $D_c = \mathrm{PW50}/T$ is called *channel normalized linear density*. For a given PW50, the smaller the T, the closer together the magnetization transitions, which implies higher intersymbol interference (ISI). Figure 1 illustrates the frequency response characteristics $H(f) = S(f)[1 - e^{-j2\pi fT}]$ of the impulse response $h(t)$, where $S(f)$ indicates the Fourier transform of the isolated transition response.

Assuming a linear model for the magnetic recording channel, i.e., the noise-free readback signal can be constructed by linear superposition of isolated transition responses and the noise sources are Gaussian, stationary, and additive, the readback signal may be expressed in the equivalent form

[10]

$$r(t) = \sum_n a_n h(t - nT) + \eta(t) \ , \tag{2}$$

where $\{a_n\}$ denotes the two-level channel-input sequence that modulates the current source to produce the rectangular current waveform with amplitude -1 or $+1$, and $\eta(t)$ represents the compound additive Gaussian noise. Typically, $\eta(t)$ consists of a white Gaussian noise component, called electronics noise, and a colored Gaussian noise component representing stationary colored noise of the system including the dc-erase noise. At high recording densities, and depending on the type of magnetic medium, nonstationary data-dependent noise may become predominant, and the readback signal may also exhibit certain nonlinearities. Nevertheless, the linear ISI model with additive Gaussian noise is still commonly used to estimate the performance of detection and coding schemes, even at ultra high densities, and yields very accurate results.

2.2 The Magnetic Recording System Architecture

The recording system architecture used in today's commercial hard-disk drives is shown in Fig. 2. User data is organized into sectors of 512 bytes and fed into an outer RS encoder. Specifically, the sector is subdivided and encoded into a number of RS codewords, which, after byte-interleaving, are encoded by a constrained or modulation code. The modulation code is a high-rate code that imposes constraints to facilitate timing recovery and gain control, and to eliminate quasi-catastrophic error propagation [11]. Conventional modulation codes usually employ a rate-1 precoder, either of type $1/(1 \oplus D^2)$ or of type $1/(1 \oplus D)$, where $\oplus$ denotes a modulo-2 addition. The modulation encoder is followed by a parity encoder that appends typically one to four parity bits to the modulation codeword. In today's recording systems, the rate of the inner combined modulation/parity code is typically 96/102 or 96/104. For example, assuming that the 512 bytes of user data along with a four-byte cyclic-redundancy-check (CRC) code are subdivided and encoded using three RS codewords, each able to correct six byte errors, the outer coding schemes adds a total of 40 bytes of redundancy per sector. Therefore, the overall coding rate including the inner and outer coding schemes as well as the CRC code is $R = (512/552) \times (96/102) = 0.873$.

The output of the parity encoder is mapped into the bipolar symbols $\{a_n\}$ and written onto the hard disk by the write head. After low-pass filtering and sampling, the readback signal is shaped into a generalized PR signal format $g(D) = (1 - D^2)(1 + p(D))$, where $p(D)$ is the transfer function of an L-coefficient noise-prediction filter [2]. The NPML detector then performs maximum-likelihood sequence detection and provides a preliminary estimate of the binary data sequence $\{\hat{a}_n\}$. The initial estimate at the output of the NPML detector coupled with soft information extracted at the output of the shaping filter, i.e., hard and soft information, is fed into a post-processing unit. The post-processor is a sub-optimum soft-decision decoder for the parity code that corrects a specified number of most-likely error events at the output of the NPML detector by exploiting the parity information in the incoming sequence. The post-processor produces the final estimates of the recorded data sequence. This sequence is fed to the modulation decoder that delivers estimates of the modulation encoder input. Finally, after byte-deinterleaving, the RS decoder corrects any residual errors to obtain a highly reliable estimate of the original user data stored.

2.3 Information-Theoretic Limits

During the past decade, the application of advanced signal-processing and coding concepts to storage devices has been recognized as a cost-efficient means to increase areal density. The possible gains and limits of these advanced techniques are characterized by the information-theoretic capacity of the magnetic recording channel. Determining the capacity of the magnetic recording channel appears to be difficult. Owing to the binary constraint of saturation recording, even the problem of computing capacity, or developing tight bounds, for the simplified linear ISI model with stationary Gaussian noise appears to a be a considerable challenge.

Recently, a method for computing good approximations of the capacity of finite-state channel models, such as generalized PR channels, has been proposed [12, 13]. According to this method, long channel-input and channel-output sequences are used for computing the probability of a typical output sequence as well as the conditional probability of a typical output sequence given the corresponding input sequence, based on a forward sum-product algorithm on the joint source/channel trellis. The complexity of this approach depends on the number of states of the underlying trellis. By increasing the memory of the Markov input process and numerically optimizing the transition probabilities, the capacity can be closely approximated. Clearly, real magnetic recording channels are not strictly finite-state channels but can be approximated by a finite-state model, obtained either analytically or by training a parameterized model.

An alternative method to compute good approximations on the information rates for the magnetic recording channel is via the Shamai–Laroia (SL) conjectured lower bound [14]. It has been shown that the information rates obtained by the SL bound are very close to the results obtained by applying the methods in [12, 13]. The SL bound, however, has the advantage that it is simple to evaluate numerically. Figure 3 shows the signal-to-noise ratios (SNRs) for achieving the SL bound as a function of the *user* normalized linear density, $D_u = \mathrm{PW50}/T_u$, for the case of a magnetic recording channel modeled as a linear ISI channel and affected by electronics noise only, where $T_u = T/R$ is the user bit duration corresponding to the overall code rate R. The SNR is defined as the ratio of the energy E_s of the step response $s(t)$ over the two-sided power-spectral density $N_0/2$ of the AWGN. The results of these computations are shown for various code rates. It can be seen that the optimum code rate is around $R = 0.8$. Moreover, when using a coding scheme with an overall rate smaller than $R = 0.75$, the system suffers a performance penalty. These findings demonstrate the fact, well known in the industry, that lowering the code rate beyond $R = 0.8$ does not yield performance improvements in the overall magnetic recording system.

3 Iterative Detection/Decoding Schemes

Shortly after the advent of turbo coding, it was recognized that the *turbo principle* can effectively be applied to ISI channels, and the technique is now generally referred to as *turbo equalization* [15]. According to this technique, both detection and decoding are integral parts of the iterative process in which reliability information, also referred to as soft information, is exchanged between the detector and the decoder. The soft information is typically the *a posteriori* probability (APP) of each symbol. Figure 4(a) shows a generic encoder and channel, and 4(b) a turbo equalization scheme. As can be seen, turbo equalization requires interleaving to break the dependencies of neighboring bits. At each iteration, the soft information on the encoded bits at the output of the detector is fed to the decoder in the form of extrinsic information, and the decoder in turn generates new soft information about the encoded bits, which is fed back to the detector also in the form of extrinsic information. This iterative process is repeated until a suitable stopping criterion is met, and the decoder delivers the final decisions. By iterating between detector and decoder, the detection and decoding processes are combined, but the overall complexity essentially remains that of the separate detector and decoder. It is often convenient to work with *a posteriori* log-likelihood ratio (LLR) as shown in Fig. 4(b). In this case the extrinsic information, L_e^{dec}, from the decoder is subtracted from the LLR computed by the detector, and the result, after deinterleaving denoted as L_e^{det}, is passed as extrinsic information to the decoder. Similarly, the extrinsic information, L_e^{dec}, is obtained by subtracting L_e^{det} from the LLR computed by the decoder.

The performance of turbo equalization improves by increasing the length of the code and the number of iterations, at the price of larger latency. For the magnetic recording channel various turbo equalization schemes based on PCCCs, convolutional codes, and LDPC codes have been proposed (see [16, 17] and references therein). The common feature of all these schemes is the soft-input soft-output detector, which can be based on the APP algorithm, on the soft-output Viterbi algorithm, or on any simplified version of either of these two algorithms [9]. Some of the distinct aspects of the coding schemes involved are discussed below.

3.1 Turbo and Turbo-Like Coding Schemes

The encoder for a PCCC is shown in Fig. 5(a). The input data is collected into N-bit blocks, each of which is also processed in parallel by an interleaver Π_1. The interleaver performs a suitable permutation on a block-by-block basis. The two blocks are then encoded by two rate-1/2 recursive systematic convolutional (RSC) encoders. Conceptually, this generates four blocks for each input data block. However, only one of the two systematic blocks (identical up to permutation) is retained along with the two parity blocks, as shown in Fig. 5(a). The three parallel streams created in this way are then punctured and multiplexed to generate the output. As discussed above, high-rate codes are desirable in magnetic recording. Typically, the systematic bit stream is not punctured at all, and for every $2m$ systematic bits, all but one parity bit from each parity streams are punctured away. This process generates $2m+2$ encoded bits for each sub-block of $2m$ systematic bits, thereby making the overall coding rate equal to $m/m+1$. It is clear that this process allows considerable flexibility in designing a high-rate PCCC.

The turbo decoder shown in Fig. 5(b) uses two APP decoders matched to the two RSC encoders. The decoders are connected to each other through the interleaver and deinterleaver, Π_1 and Π_1^{-1}, respectively. The iterative decoding process begins by using the output of the APP detector as the initial *a priori* information about the input data bits. Each pass through the APP decoder generates *a posteriori* information, from which the extrinsic information is obtained by subtracting the corresponding *a priori* information. The extrinsic information then becomes the *a priori* information for the next round of APP decoding. Note that the interleaver Π_1 and the deinterleaver Π_1^{-1} are used in the decoder to maintain the proper ordering for each decoder. The APP information generated after each iteration can also be rearranged, via the puncturing and multiplexing step, for use in the APP detector. The delinterleaved output of the second APP decoder is used to obtain hard decisions, i.e., the highly reliable estimates of the input data.

An alternative coding scheme applicable to magnetic recording systems can be based on high-rate convolutional coding as shown in Fig. 5(c). The input data is encoded by a single rate-1/2 RSC encoder and is further punctured (and multiplexed) to obtained any desired high-rate code. The main advantage of this structure is that there is only a single APP decoder, which significantly reduces the complexity of the coding scheme. This system still delivers a performance comparable to that of its PCCC counterpart by realizing the turbo principle through the iterative processing between the APP detector and the APP decoder. A block diagram of the APP decoder, including the puncturing/depuncturing and demultiplexing/multiplexing functions, is shown in Fig. 5(d).

3.2 LDPC Codes

Although binary LDPC codes have been known for four decades, their extraordinary performance was rediscovered only in the past few years. Currently, there is a strong research activity aiming at a better understanding of the theoretical limits of these codes as well as at improved constructions [18]. As is evident from the name, a low-density parity-check code is characterized by a sparse parity-check matrix. Initially, only regular binary LDPC codes were considered, which have a constant number of "1"s per row—referred to as row weight—and a constant number of "1"s per column—referred to as column weight [5].

A parity-check matrix H can be associated with a Tanner graph. The Tanner graph is a bipartite graph with two sets of nodes, namely, symbol and check nodes, whose connectivity is determined by H [18], and is a core concept for the iterative decoding of LDPC codes. At each iteration step, soft information about the code bits are passed from the symbol nodes to the check nodes and back again along the edges of the graph. The Tanner graph concept can be extended to capture the ISI dependencies introduced by the generalized PR channel, thereby facilitating a uniform approach for the turbo equalizer as a whole (see Fig. 6). The resulting graph is generally referred to as a *factor graph*. The upper part of Fig. 6 represents the Tanner or factor graph of the LDPC code consisting of the check nodes (solid squares) and symbol nodes (circles), which correspond to the encoded bits a_n (which are the same as the encoded bits b_n in the absence of interleaving). The lower part of Fig.

6 shows the factor graph of the generalized PR channel. The states s_n of this channel correspond to state nodes (double circles), which are connected to check nodes (solid squares), representing state transitions. Moreover, the check nodes are connected to the symbol nodes representing the encoded bits b_n, which correspond to the inputs of the generalized PR channel. The check nodes are also connected to the noiseless outputs x_n of the generalized PR channel, represented by circles. The bottom row consists of function nodes (solid squares), which model the noise at the output of the generalized PR channel, expressed as conditional probability density functions $f(y_n|x_n)$, where y_n denotes the noisy generalized PR channel output. The factor graph in Fig. 6 is a refined view of the coarser block diagram shown in Fig. 4(b) for an LDPC decoder. Note that such a combined factor-graph representation does not specify the message-passing schedule between detector and decoder, allowing the investigation of efficient message-passing schedules that can lead to shorter delays.

Iterative decoding of LDPC codes, based on the sum-product algorithm, performs well provided there are no short cycles in the underlying Tanner graph. The parity-check matrix H of LDPC codes can be constructed either randomly or algebraically. The random constructions are based on selecting parts of H randomly while maintaining some constraints on the row and column weights as well as on the number of short cycles. For example, H can be constructed by appending randomly generated column vectors of length M with a specified column weight, and discarding those vectors that violate any of the constraints mentioned above [7]. Alternatively, a construction can also be based on selecting or discarding each nonzero bit in H, subject to a set of desired constraints [19].

Algebraically constructed codes having strict control over short cycles have also been proposed. The construction methods for such codes can be divided into two main categories: (1) constructions based on finite geometries (see, for example, [20, 21]) and (2) constructions based on permutation matrices (see for example [6, 22, 23, 24]). There is an intimate relationship between the two methodologies because many constructions based on finite geometries can also be equivalently represented with permutation matrices.

Constructions based on finite geometries generally exploit the intersection properties between at least two types of elements of the geometry considered, with the constraint that any two elements of one type cannot both belong to an element of the second type (which guarantees a girth of at least six in the Tanner-graph representation of the code). These codes are often cyclic or quasi-cyclic. As a result, their encoding can be realized with shift registers in linear time complexity. They also often have many orthogonal redundant check sums compared with the minimum number given by the rank of the parity-check matrix, and these additional check sums can be used in iterative decoding. This structure allows a relatively high minimum distance to be achieved, but at the price of large row and column weights, resulting in a large decoding complexity with iterative APP decoding. However, reduced-complexity algorithms for these classes of codes perform with negligible performance degradation [25].

A quasi-cyclic LDPC code with column and row weights j and k, respectively, can be constructed by appending j rows of k permutation matrices judiciously chosen. The quasi-cyclic structure allows encoding in linear time based on shift registers. In general, only few redundant rows are present in the parity-check matrix, which implies smaller row and column weights than in the first approach, but also a correspondingly smaller minimum distance.

For many of the codes obtained from these two general construction methods, good error performances have been reported up to a bit error rate of 10^{-7}. For short to medium code lengths, the performance often compares favorably with that of randomly constructed LDPC codes.

4 Performance Results

The error-rate performance of the various detection and decoding schemes described above has been studied by computer simulation. An isolated transition in the underlying magnetic recording channel is modeled by the Lorentzian pulse having energy E_s. The readback signal is first filtered by a low-pass filter and then fed to a PR shaping equalizer. The generalized PR target polynomial

is $(1 - D^2)(1 + p_1 D + p_2 D^2)$, giving rise to a 16-state detector trellis. In all computations and simulations the SNR is defined as $2E_s/N_0$.

Figure 7(a) shows the performance of two turbo-equalization schemes for the magnetic recording channel. The first scheme uses a rate-4124/4644 convolutional code. This high-rate code has been obtained by puncturing a rate-1/2, memory-4 convolutional code. Various rate-1 precoders of the form $1/(1\oplus D)$, $1/(1\oplus D^2)$, $1/(1\oplus D^3)$, $1/(1\oplus D\oplus D^3)$ etc. were also tested. As can be seen in the low-SNR region, precoding seems to offer a gain of more than 0.5 dB. However, in the high-SNR region, the error-rate curves exhibit a distinct floor, and the performance becomes worse than without any precoding. Also shown in Fig. 7(a) is the performance of the second turbo-equalization scheme based on LDPC coding. In this case, an array-code-based LDPC coding scheme of rate-4131/4652 has been adopted. Clearly, in the high-SNR region, LDPC coding offers the best performance at the two user densities $D_u = 2.6$ and $D_u = 3.6$ assumed in these simulations. Finally, note that both turbo-equalization schemes use a maximum of 10 iterations between the APP detector and the sum-product (LDPC code) or APP (convolutional code) decoder. Other approaches, such as the high-rate tail-biting convolutional codes with precoding, can delay the onset of error floors considerably [26]. Note that the convolutional-code-based turbo-equalization scheme considered in this paper is for illustration purposes only.

Figure 7(b) shows the SNR requirements of two different architectures for achieving a block- or sector-error rate of 10^{-4} as a function of the user normalized linear density D_u. In particular, curve 1 corresponds to the state-of-the-art NPML architecture, described in Subsection 2.2, which employs a rate-96/102 inner modulation/parity code and a three-way interleaved outer RS code. Curve 2 corresponds to an advanced architecture that relies on a single outer rate-4131/4652 LDPC code and turbo equalization. To avoid using an even higher-rate LDPC code, no inner run-length-limited code, although important in a practical system, has been considered. Both schemes use a block or sector size of approx. 516 bytes (including CRC). Finally, curve 3 corresponds to the achievable information rate in the form of the SL bound for $R = 0.8889$. The advanced architecture based on LDPC coding achieves a performance gain of approx. 1.7 dB at lower densities and approx. 2.4 dB at higher densities over the state-of-the-art architecture that is currently being used by disk-drive manufactures. Moreover, the gap to the SL bound of the achievable information rate is approx. 1.5 dB at low linear recording densities, whereas at high linear recording densities the gap increases to more than 3.5 dB. It is expected that by increasing the block or the sector size, the gap of the LDPC coding scheme to the achievable information rates will be reduced substantially. Similar results have also been obtained for colored stationary as well as data-dependent transition noise.

For commercial hard-disk drives, the current constraint on the block or the sector size, the requirement of an overall bit-error rate of 10^{-16}, and the necessity of a very high overall coding rate may eliminate some of the coding schemes discussed above from consideration. Note also that conventional serially concatenated schemes with a RS outer decoder can also be enhanced by employing a form of interaction between the inner and outer decoder, or by utilizing combined Chase–GMD-type RS decoding [27].

5 Concluding Remarks

Over the past decade, several advanced signal-processing and coding techniques have been introduced into hard-disk drives to improve the error-rate performance at ever increasing areal densities. The recent advances in capacity-approaching codes hold the promise to enable an areal density that is very close to the ultimate limit, for specific magnetic recording components, while maintaining the very stringent on-track error-rate requirements. By now, it is well understood that by employing iterative detection/decoding schemes, performance gains beyond those achievable by the current state-of-the-art signal-processing and coding architectures used in commercial hard-disk drives, are possible.

However, despite progress in the area of reduced-complexity detection and decoding algorithms, turbo-equalization structures with iterative detectors/decoders have not yet found their way into

digital recording systems because of the still unfavorable tradeoff between performance, implementation complexity, and latency, as well as the lack of analytical tools to predict the required error performance (in particular the occurrence of the error floor which may be caused by either the minimum distance of the code, or the suboptimality of iterative decoding). The design of high-rate, short-block-length turbo-like codes for recording systems remains an area of active research.

References

[1] R. D. Cideciyan, F. Dolivo, R. Hermann, W. Hirt, and W. Schott, "A PRML system for digital magnetic recording," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 38-56, Jan. 1992.

[2] J. D. Coker, E. Eleftheriou, R. L. Galbraith, and W. Hirt, "Noise-predictive maximum likelihood (NPML) detection," *IEEE Trans. Magn.*, vol. 34, part 1, pp. 110-117, Jan. 1998.

[3] R. D. Cideciyan, J. D. Coker, E. Eleftheriou, and R. L. Galbraith, "NPML detection combined with parity-based post-processing," *IEEE Trans. Magn.*, vol. 37 , pp. 714-720, March 2001.

[4] W. Feng, A. Vityaev, G. Burd, and N. Nazari, "On the performance of parity codes in magnetic recording systems," in *Proc. IEEE Global Telecommun. Conf., GLOBECOM '00*, San Francisco, CA, pp. 1877-1881, Nov. 2000.

[5] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int'l Conf. on Communications, ICC '93,* Geneva, Switzerland, pp. 1064-1070, May 1993.

[6] R. G. Gallager, "Low-density parity-check code," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 21-28, Jan. 1962.

[7] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, pp. 399-431, Mar. 1999.

[8] A. Dholakia, E. Eleftheriou, and T. Mittelholzer, "On iterative decoding for magnetic recording channels," in *Proc. 2nd Int'l Symposium on Turbo Codes & Related Topics*, Brest, France, pp. 219-226, Sept. 2000.

[9] T. Mittelholzer, A. Dholakia, and E. Eleftheriou, "Reduced-complexity decoding of low density parity check codes for generalized partial response channels," *IEEE Trans. Magn.*, vol. 37, pp. 721-728, Mar. 2001.

[10] J. Moon and L. R. Carley, "Performance comparison of detection methods in magnetic recording," *IEEE Trans. Magn.*, vol. 26, pp. 3155-3172, Nov. 1990.

[11] K. A. S. Immink, P. H. Siegel, and J. K. Wolf, "Codes for digital recorders," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2260-2299, Oct. 1998.

[12] D. Arnold and H.-A. Loeliger, "On the information rate of binary-input channels with memory," in *Proc. IEEE Int'l Conf. on Communications, ICC '01,* Helsinki, Finland, pp. 2692-2695, June 2001.

[13] H. D. Pfister, J. B. Soriaga, and P. H. Siegel, "On the achievable information rates of finite-state ISI channels," in *Proc. IEEE Global Telecommun. Conf., GLOBECOM '01,* San Antonio, TX, pp. 2992-2996, Nov. 2001.

[14] S. Shamai and R. Laroia, "The intersymbol interference channel: Lower bound on capacity and channel precoding loss," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1388-1404, Sept. 1996.

[15] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo-equalization," *European Trans. Telecommun.*, vol.6, no. 5, pp. 507-511, Sept.-Oct. 1995.

[16] T. Souvignier, A. Friedmannn, M. Oberg, P. H. Siegel, R. E. Swanson, and J. K. Wolf, "Turbo decoding for PR4: Parallel versus serial concatenation," in *Proc. IEEE Int'l. Conf. on Communications, ICC '99,* pp. 1638-1642, June 1999.

[17] J. Fan, A. Freedman, E. Kurtas, and S. McLaughlin, "Low density parity check codes for magnetic recording," *Proc. 37th Allerton Conf. on Communications Control, and Computing*, Urbana, IL, Sep. 1999.

[18] Special Issue on Codes and Graphs and Iterative Algorithms, *IEEE Trans. Inform. Theory*, vol. 47, Feb. 2001.

[19] X.Y. Hu, E. Eleftheriou, and D.-M. Arnold, "Progressive edge-growth Tanner graphs," in *Proc. IEEE Global Telecommun. Conf., GLOBECOM '01,* San Antonio, TX, pp. 995-1001, Nov. 2001.

[20] 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, pp. 2711-2736, Nov. 2001.

[21] P.O Vontobel and R. M. Tanner "Construction of codes based on finite generalized quadrangles for iterative decoding," *Proc. IEEE Int'l. Symp. on Information Theory, ISIT'01,* Washington, DC, p.223, June 24-29 2001.

[22] J. L. Fan, "Array codes as low-density parity-check codes," *Proc. 2nd Int'l Symposium on Turbo Codes & Related Topics*, Brest France, pp. 543-546, Sept. 2000.

[23] D. J. C. MacKay and M. C. Davey, "Evaluation of Gallager codes for short block length and high rate applications," in *Codes, Systems, and Graphical Models*, Eds. B. Marcus and J. Rosenthal, pp. 113-130, (Springer, Berlin Heidelberg, 2001).

[24] E. Eleftheriou, S. Ölçer, "Low-density parity-check codes for digital subscriber lines," in *IEEE Int'l Conf. on Communications, ICC'02,* New York, pp. 1752-1757, Apr. 2002.

[25] J. Chen and M. Fossorier, "Decoding low-density parity check codes with normalized APP-based algorithm," in *Proc. IEEE Global Telecommun. Conf., GLOBECOM '01,* San Antonio, TX, pp. 1026-1030, Nov. 2001.

[26] M. Tüchler, C. Weiß, E. Eleftheriou, A. Dholakia, and J. Hagenauer, "Application of high-rate tail-biting codes to generalized partial response channels," in *Proc. IEEE Global Telecommun. Conf., GLOBECOM '01,* San Antonio, TX, Nov. 2001, pp. 2966-2971.

[27] C. Y. Liu, H. Tang, S. Lin and M. Fossorier, "An interactive concatenated turbo coding system," *IEEE Trans. Vehicular Technol.*, vol. 51, September 2002 (in press).

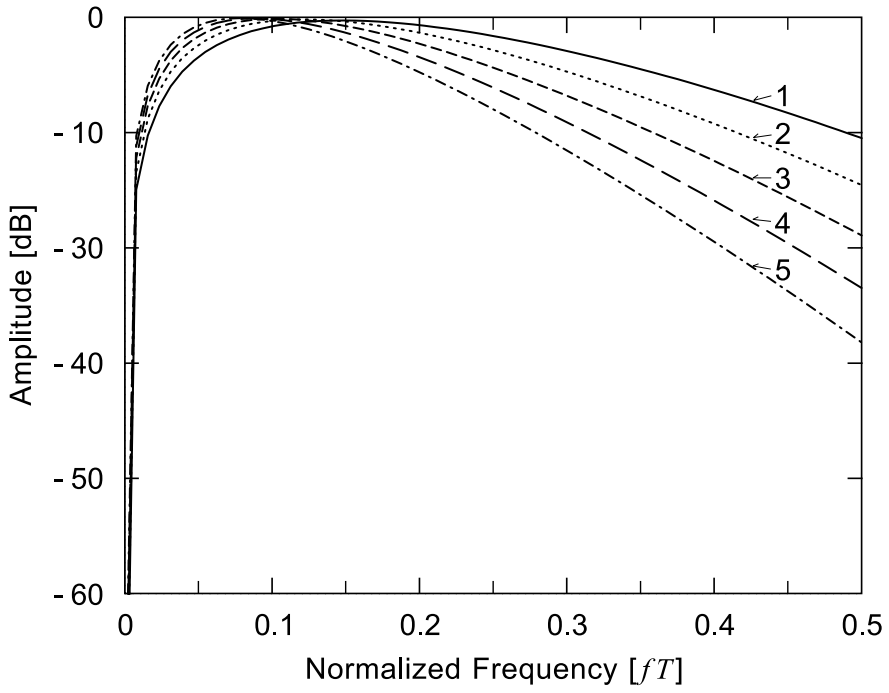

Figure 1: Frequency response to a square pulse of duration T. Curve 1: $D_c = 2.0$; curve 2: $D_c = 2.5$; curve 3: $D_c = 3.0$; curve 4: $D_c = 3.5$, and curve 5: $D_c = 4.0$.

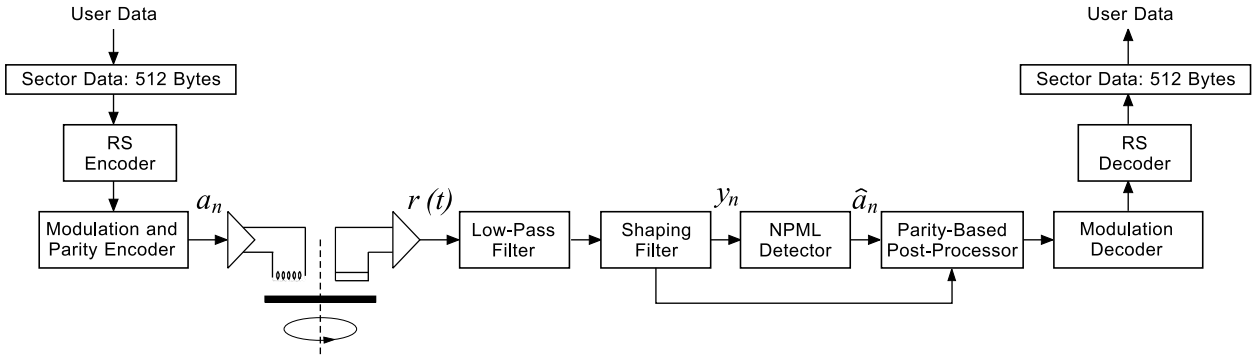

Figure 2: Recording system architecture.

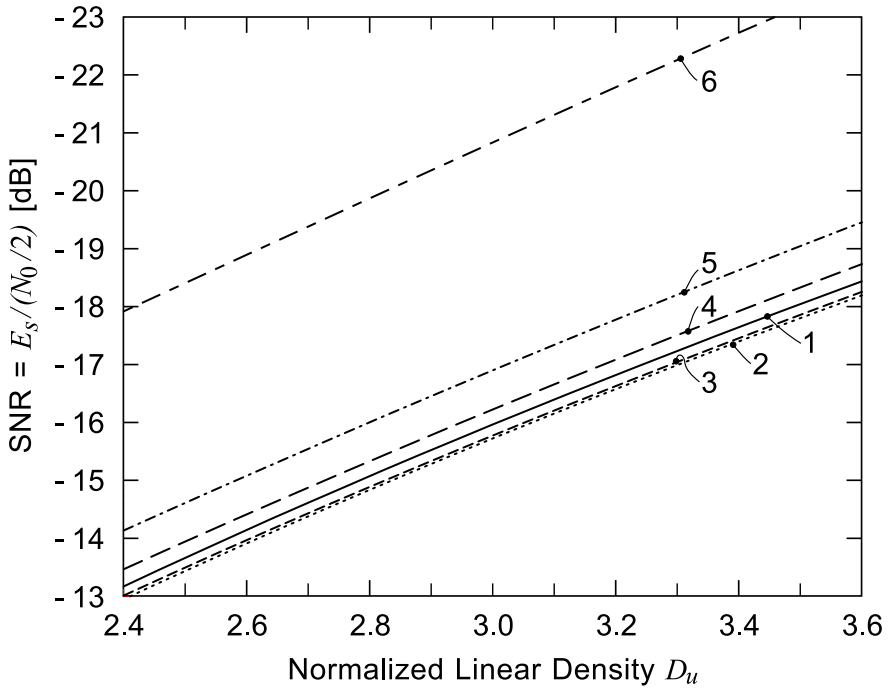

Figure 3: Capacity bounds. Curve 1: $R = 0.6$; curve 2: $R = 0.7$; curve 3: $R = 0.8$; curve 4: $R = 0.8889$; curve 5: $R = 0.9412$, and curve 6: $R = 0.999$.

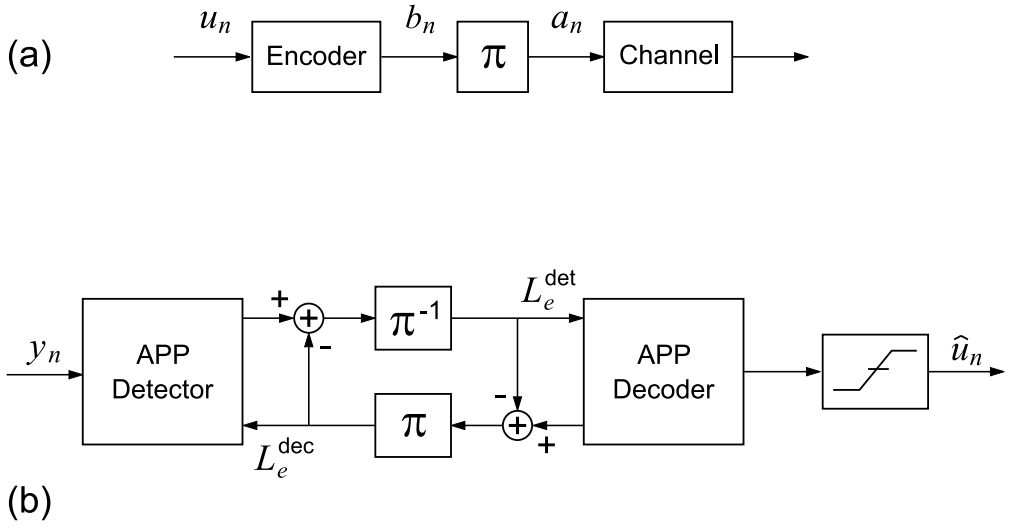

Figure 4: Iterative detection/decoding system.

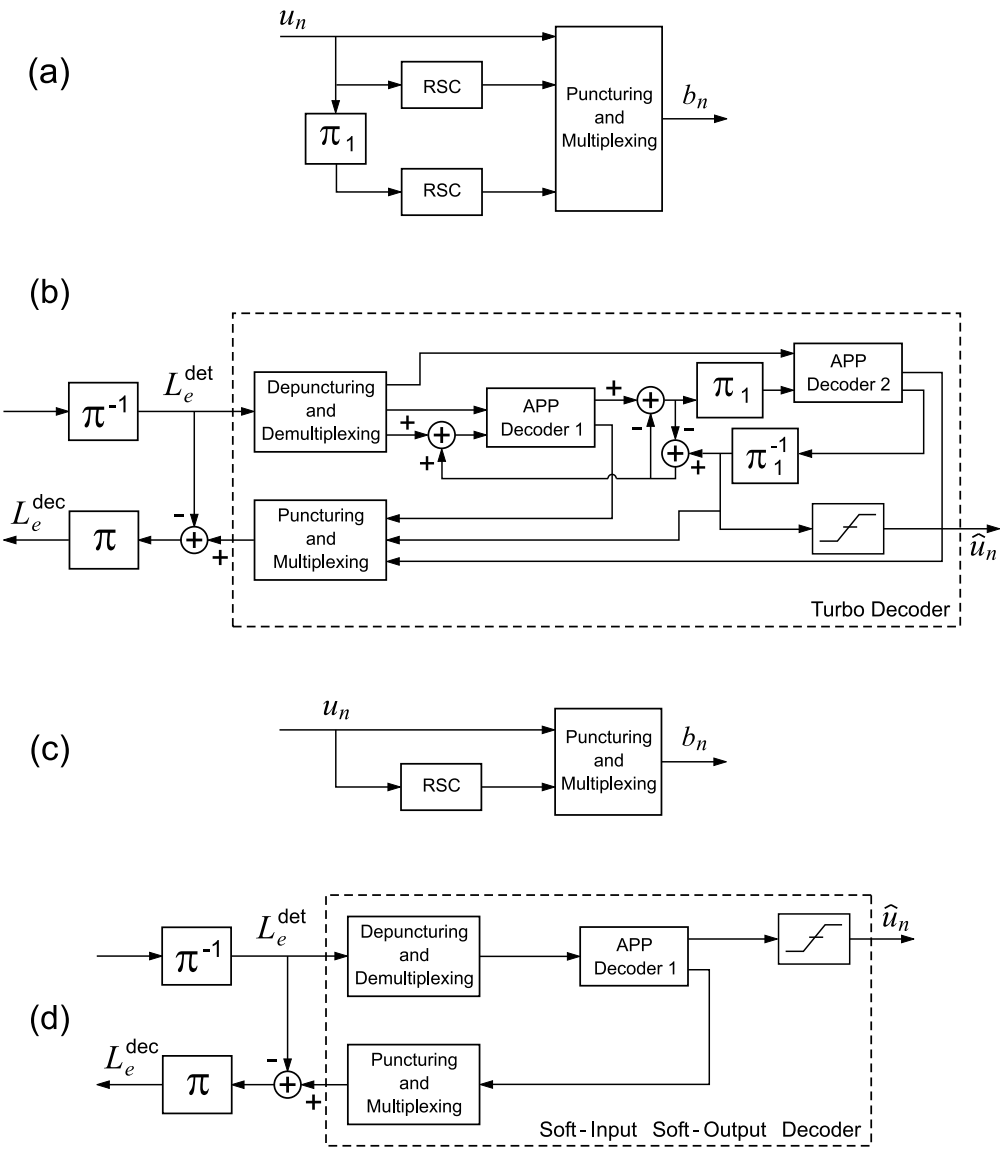

Figure 5: Turbo and turbo-like coding and decoding schemes.

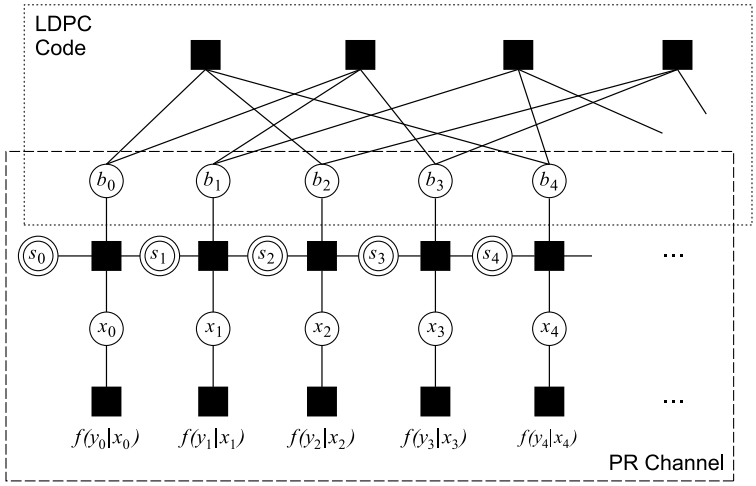

Figure 6: LDPC code and PR channel joint factor graph.

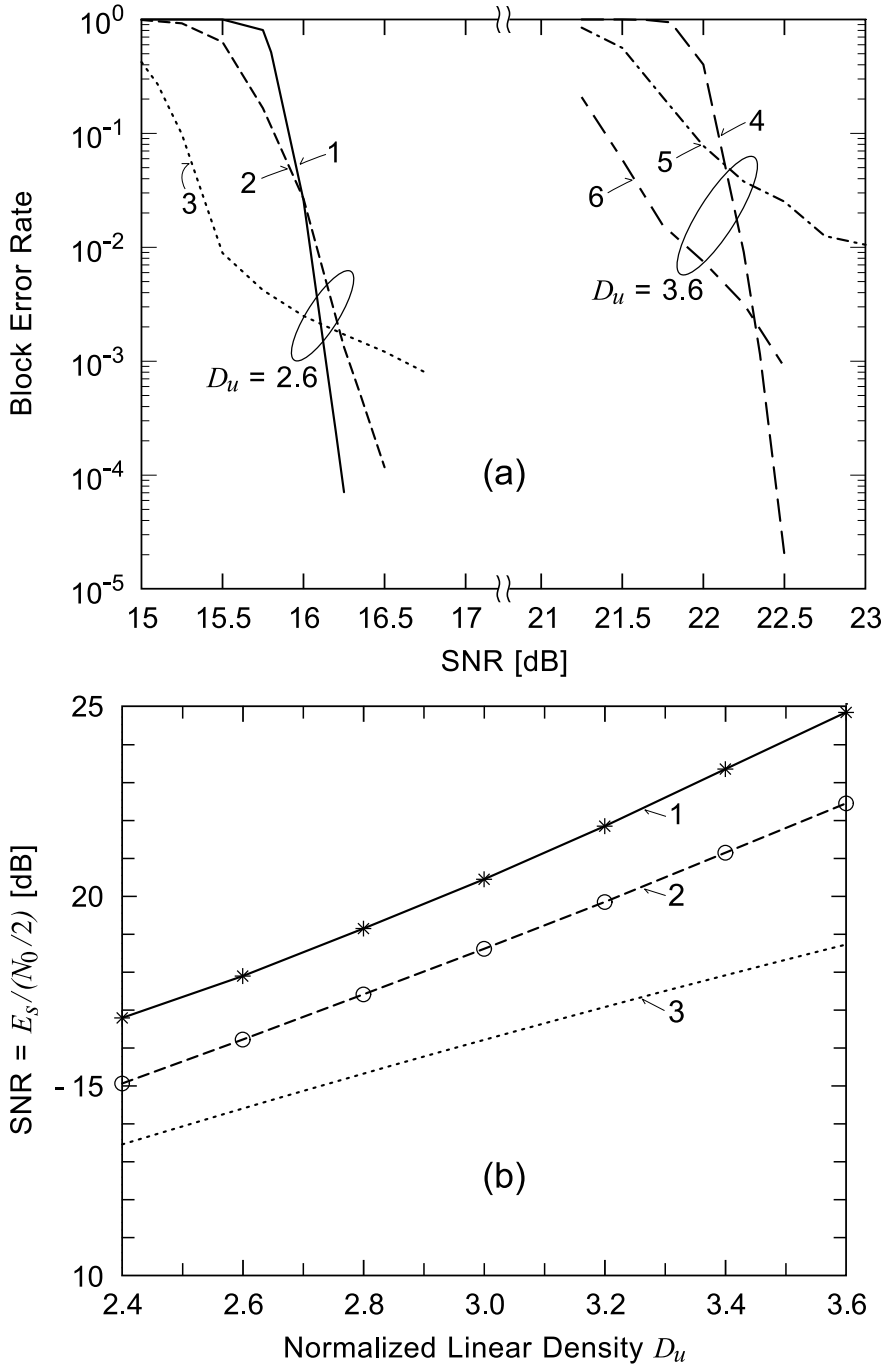

Figure 7: Performance results: (a) Block-error-rate curves showing the effect of precoding at $D_u = 2.6$ and $D_u = 3.6$. Curves 1 and 4: LDPC code; curve 2: convolutional code with $1/1 \oplus D^3$ precoder; curves 3 and 6: convolutional code with $1/1 \oplus D$ precoder, and curve 5: convolutional code, no precoding. (b) SNR requirements for a block error rate of 10^{-4}, as a function of D_c. Curve 1: NPML-RS architecture; curve 2: LDPC architecture, and curve 3: information rate.