Coding for High-Density Storage Systems*

Joseph A. O'Sullivan
Dept. of Electrical Engineering
Washington University in St Louis
St Louis, MO 63130
jao@ee.wustl.edu

Lihao Xu
Dept. of Computer Science
Washington University in St Louis
St Louis, MO 63130
lihao@cs.wustl.edu

Naveen Singla
Dept. of Electrical Engineering
Washington University in St Louis
St Louis, MO 63130
singla@essrl.wustl.edu

Abstract

We study the issues relating to coding for a proposed high-density storage system consisting of arrays of storage devices. The storage devices store data in a two-dimensional (2-D) manner. The data on each storage device is coded using an inner code and coding is also done across the devices using an outer code. Array codes are shown to be a good choice for the outer codes and low-density parity-check codes for the inner codes. The read-back data from the storage devices has 2-D intersymbol interference. Joint equalization and decoding methods are used for detection and decoding on each storage device. The issue of the trade off between decoding performance and passing soft or hard information from the inner code to the outer code is also discussed.

1 Introduction

As traditional information storage devices approach predicted limits, new and improved technologies may be required to keep up with the trends in data storage. One alternative technology could be the use of two-dimensional (2-D) storage. On magnetic media this may correspond to a bit aspect ratio of one and no inter-track spacing, leading to a large increase in the storage capacity. Optical storage systems may be two or three-dimensional (holographic). While we cannot yet implement 2-D recording and reading, progress is being made to achieve this technology in the future. Multi-track optical transducers or magnetic recording sensors [1], [2] could be used for read and write for such storage media. The "millipede," an array of atomic force microscopy tips, developed by IBM Research in Zurich [3] is designed for polymer substrates employing mechanical storage. The "millipede" could be modified for read and write on 2-D magnetic recording media.

In this paper we envision an array of high-density storage devices, where each storage device stores data in a 2-D manner. The demands of such a system necessitate coding with

---

*This research is supported in part by NSF Grants ECS-0000434, and ECS-9900159.

good error correction capability and low data read back latency. An important aspect of the storage devices is intersymbol interference (ISI) which, due to the 2-D nature of data storage, is also 2-D. For conventional magnetic recording, one-dimensional (1-D) ISI has been modeled and partial response maximum likelihood decoding is used to combat its effects. On the other hand, combating ISI in two dimensions is much harder. The reason for this being that the BCJR and the Viterbi algorithms, which are commonly employed for 1-D ISI, have no direct generalization to two dimensions. Also, in two dimensions every bit has significantly more interfering neighbours than in one dimension. Several equalization methods used to reduce 2-D ISI have been proposed in the literature [4, 5, 6]. However, relatively little work has been done on joint equalization and decoding for 2-D ISI. Some methods for joint equalization and decoding proposed in [7], [8] are used to mitigate the effects of the 2-D ISI on the storage devices in the high-density storage system.

The coding scheme proposed here for the high-density storage system uses an outer code which spreads data over the array of storage devices. In conjunction with the outer code, an inner code is used on each individual storage device. The inner code also includes equalization for reducing the 2-D ISI. A block diagram of the envisioned system is shown in Fig. 1. The channel (not shown) is assumed to be an additive white Gaussian noise (AWGN) channel. We show that array codes, a special class of 2-D codes, are a good choice for the high-rate outer codes due to their simple algebraic structure and good performance when decoded using message-passing algorithms. Each row (or column) of the array code is mapped to a disk. We also show that low-density parity-check (LDPC) codes are a good choice for the low-rate inner codes. The issue of the trade off between decoding performance and passing soft or hard information from the inner code to the outer code is also studied.

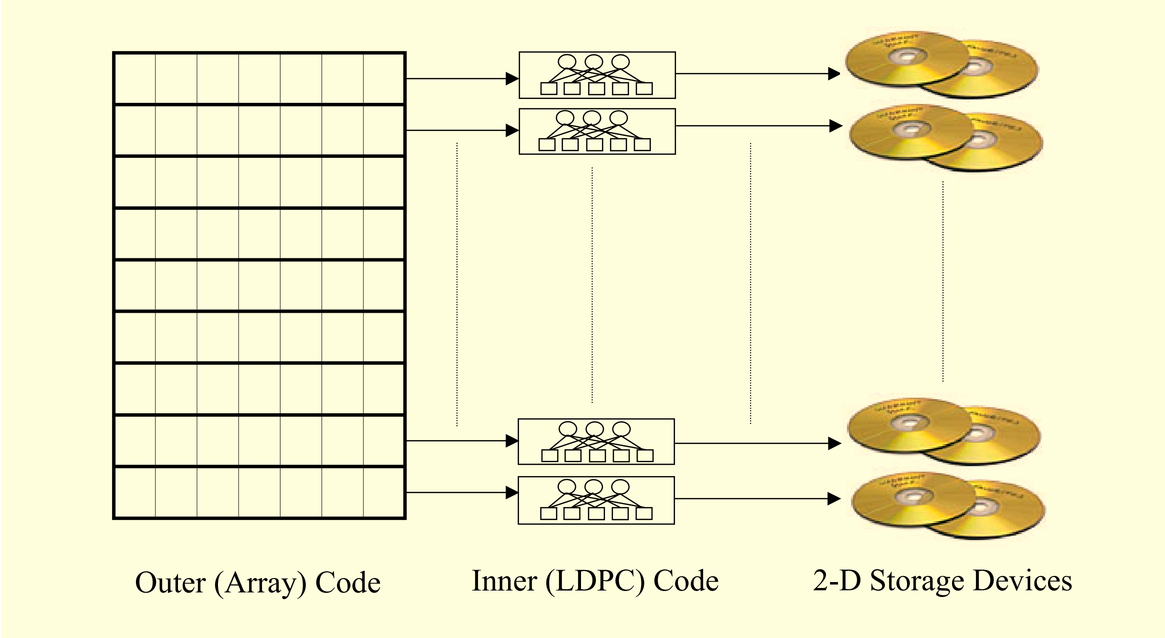


Figure 1: High-Density Storage System. Each row of the array code is mapped to a disk. The inner (LDPC) code block also performs equalization to mitigate the effects of the 2-D ISI. Each disk stores data in a 2-D pattern.

The paper is organized as follows: Section 2 briefly describes the joint equalization and decoding schemes developed for combating 2-D ISI. Section 3 discusses in detail the paradigm of the proposed high-density storage system and the issues pertaining to coding. Conclusions are provided in Section 4.

2 Joint Equalization and Decoding

This section briefly describes four iterative equalization and decoding schemes used for systems that have 2-D ISI during read-back. The channel is modeled as a discrete 2-D ISI channel characterized by its impulse response, $\mathbf{h}$. The system can be modeled by a discrete-time communication system as shown in Fig. 2. For our simulations we use

$$h = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 0.25 \end{pmatrix}.$$

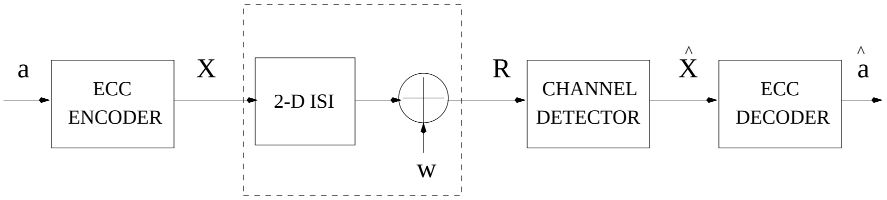


Figure 2: The equivalent discrete-time communication model for a system with 2-D ISI.

The observed data is $\mathbf{R} = \mathbf{h} * *\mathbf{X} + \mathbf{W}$, where ** denotes 2-D convolution. $\mathbf{X}$ is a matrix, typically rectangular, obtained by taking the output of the LDPC code encoder over a 2-D index set and $\mathbf{W}$ is AWGN. Two of the joint equalization and decoding schemes are designed for a general channel response while the other two are designed for a separable channel response. A 2-D channel response is separable if it can be written as a product of two 1-D vectors. The considered channel response is separable into $[1\ \ 0.5]^T[1\ \ 0.5]$. For a detailed description of these schemes refer to [7], [8].

2.1 MMSE Equalization and Decoding

The first scheme performs equalization based on minimum mean squared error (MMSE) criterion followed by decoding using the LDPC codes. 2-D MMSE equalization has been shown to be very effective for detection on 2-D ISI channels [6]. The Wiener filter is applied iteratively with soft information being passed from the LDPC decoder to the Wiener filter. The soft information is the estimated mean of the codeword calculated by the LDPC decoder using the posterior probabilities of the codeword bits.

2.2 Joint Equalization and Decoding using Message-Passing

message-passing algorithms that take advantage of the 2-D dependence are less well-developed than their 1-D counterparts. The second scheme is a pure *a-posteriori* probability (APP) based algorithm that is used for both equalization and decoding. This algorithm computes approximate APPs of the codeword bits, given the observed data, by performing message-passing on a three-level graph of the LDPC code and the channel ISI. The upper two levels in this “full graph” represent the LDPC code bipartite graph which shows the connection of the codeword bits to the parity-check bits through the parity-check matrix. The lower two levels represent the channel ISI graph showing how the ISI connects the codeword bits to the observed data.

2.3 Joint Equalization and Decoding for Separable Channels

The following two schemes are developed for a special case of 2-D ISI channels where the channel response is separable. In this case the channel is treated as a 1-D ISI acting along the row dimension, followed by a 1-D ISI acting along the column dimension.

The third scheme, named "Equalization-MAP," uses MMSE equalization to reduce the ISI in the column dimension followed by a MAP detector that performs a row-by-row detection. This is followed by the LDPC decoder which iterates with the MAP detector. A block diagram of the decoding scheme is shown in Fig. 3.

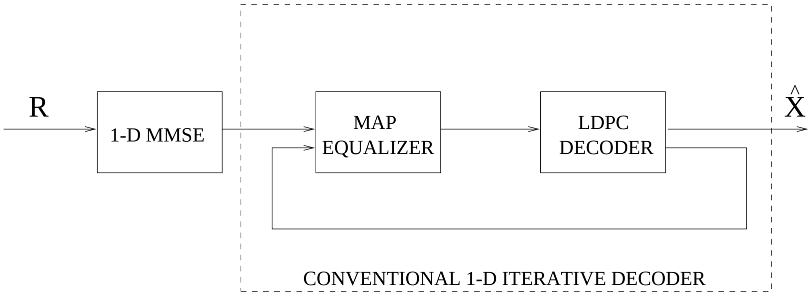


Figure 3: Decoder for the equalization-MAP scheme. **R** is the received data and $\hat{\mathbf{X}}$ is the estimate of the user data. The MAP equalizer iterates with the LDPC decoder.

The second scheme, named "Row-and-Column," employs modified MAP algorithms for the ISI in each dimension. The modifications have to be made due to the fact that the trellis of the column MAP detector is nonbinary, thus requiring the use of a MAP for nonbinary inputs. Also, no direct information from the channel is available for the row MAP detector necessitating modifications in it [8]. The LDPC decoder only iterates with the row detector. A block diagram of this scheme is shown in Fig. 4.

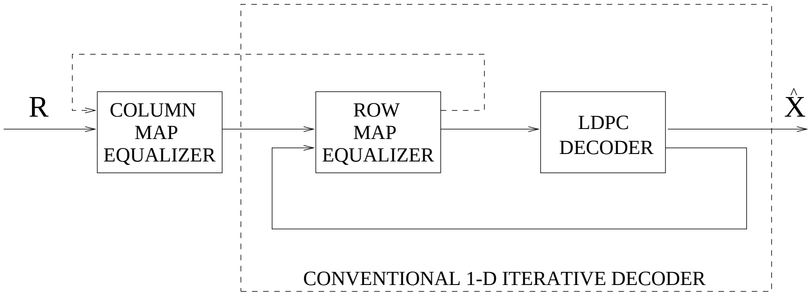


Figure 4: Decoder for the row-and-column scheme. **R** is the received data and $\hat{\mathbf{X}}$ is the estimate of the user data. The row MAP equalizer iterates with the LDPC decoder. The decoder can be modified to allow the MAP detectors to iterate between themselves also.

The performance curves for all the schemes are shown in Fig. 5. We compare the performance of the schemes to the performance of the LDPC code on the channel without ISI. The best performance is obtained using the Row-and-Column decoding algorithm. For a general channel response the full graph algorithm gives the best performance. The performance of the full graph algorithm is degraded due to the presence of short cycles in the channel ISI graph [7].

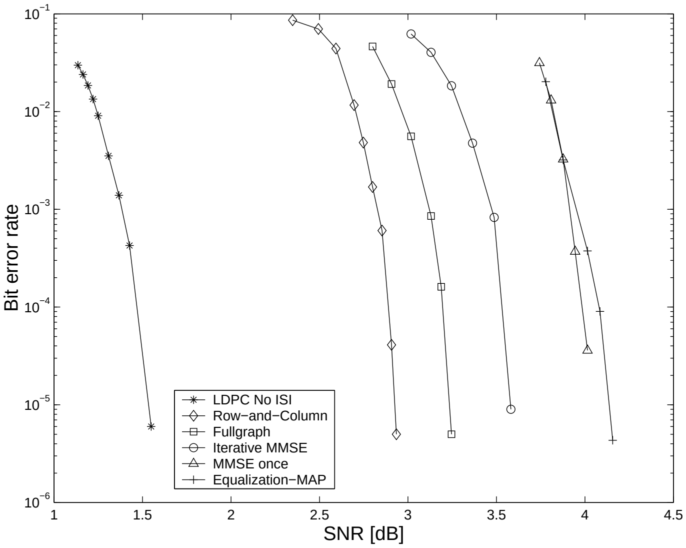


Figure 5: Performance curves for the proposed schemes using a [10000,5000] regular LDPC code. The curves from left to right are, the performance of the LDPC code on an AWGN channel with no ISI, the Row-and-Column decoding scheme, the full graph message-passing scheme, iterative 2-D MMSE equalization and decoding, 2-D MMSE equalization performed once followed by decoding, and the equalization-MAP decoding scheme.

3 High-Density Storage System

As the capacity of an individual storage device approaches its physical limit, a natural extension is to use multiple storage devices connected together through a high speed network. This array of disks can not only effectively increase storage capacity, but also bring fault tolerance to the storage system by using redundant disks to store parity-check data. This idea has been applied in *RAID* (Redundant array of independent disks) systems [9]. From a coding point of view for such a system of disk arrays, high-rate codes can be used as outer codes to spread data over a number of high-density storage devices. Each high-density storage device is assumed to store data in a 2-D manner. An inner code is used on each high-density storage device.

3.1 Array codes versus LDPC codes

Various structured constructions of LDPC codes have been proposed [10, 11, 12]. Irregular LDPC codes, when optimized, can get very close to capacity on an AWGN channel for very large block lengths (of the order of Mbits). Practical data storage applications, however, require much shorter blocklengths (of the order of kbits) to enable low read-back latency and low memory space needed for data read-write control device. Random LDPC codes usually do not perform well for short blocklengths. Thus it is important to seek more structured code constructions with short codeword length, which allow for easier encoding, guaranteed distance properties, less implementation memory, and more importantly low decoding complexity resulting in low read-back latency.

Array codes, a special class of 2-D codes, have been studied in the context of data storage applications and burst error communication channels [13, 14, 15]. For an array code, data is placed in a 2-D array rather than a 1-D vector. A common property of these codes is that the encoding and decoding computations use only simple *XOR* (exclusive OR) operations, which can be implemented easily in hardware or software or both; thus these codes are more efficient than Reed-Solomon codes in terms of computational complexity. Algebraically, array codes can be defined over an Abelian group $G(q)$. Furthermore many binary array codes have low-density in their parity-check matrices [16] and thus are suitable for decoding using the message-passing algorithm. In this context, it is worthwhile to consider several families of array codes as LDPC codes. In paritcular, it is beneficial to study the performance of array codes as LDPC codes for 2-D data storage devices and data storage arrays. Also, almost all current RAID systems use the simple parity-check code to distribute redundant data. Thus these RAID systems usually tolerate only one disk failure. Array codes on the other hand are capable of tolerating more than one disk failures when each row (or column) of the array code is mapped to a disk.

Here we compare the performance of several families of array codes with LDPC codes for use as the high-rate outer code or low-rate inner code. Array codes and LDPC codes are compared for low-rate and high-rate scenarios. The array codes considered are the EVENODD code [14], the X-code [17], the BR-93 code [13], and the BR-99 code [18]. These classes of array codes are all high-rate and maximum distance separable (MDS) codes capable of tolerating two disk failures. Fig. 6 shows this performance comparison. As expected, the array codes perform poorly in the low-rate regime. For high-rates the performance of the array codes is very close to those of the LDPC codes on an AWGN channel with no ISI. Again the use of practical blocklengths is emphasized here. The results suggest using the array codes as the high-rate outer codes and LDPC codes as the low-rate inner codes for the proposed storage system.

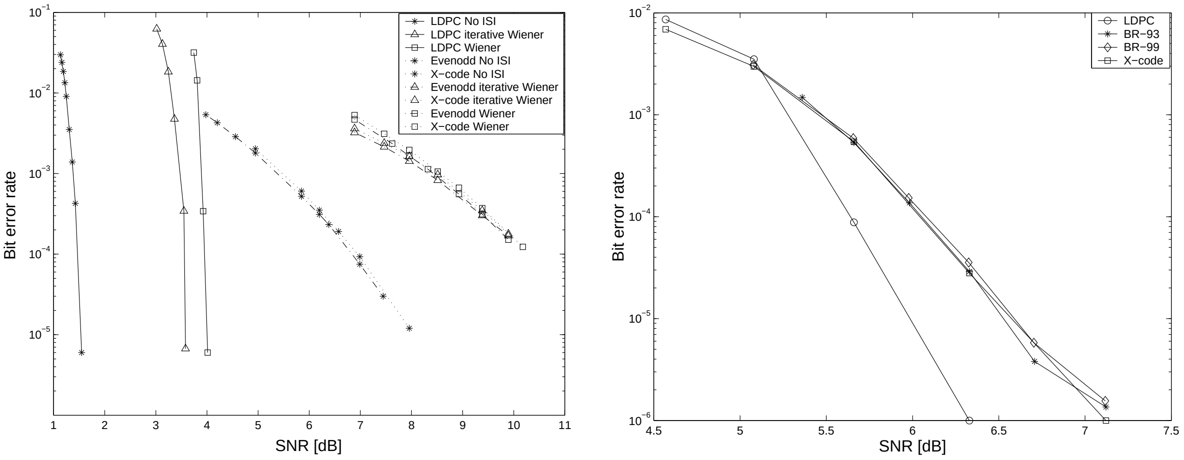
3.2 Hard versus Soft Information

Current RAID systems only employ hard algebraic decoding for error recovery i.e., the data output from each individual disk in the system is hard (including erasure when a disk failure is detected). Inherently each storage device maybe able to output soft real-valued data signal stored on it. Fig. 7 shows the performance of an array code using soft decoding on an AWGN channel, which corresponds to passing soft information from the inner code to the outer code. The figure also shows the performance of the array code on a binary symmetric channel using the bit-flipping algorithm which would correspond to hard information passed to it from the inner code. We can see a significant improvement in performance from this curve when soft-decoding is used.

Thus, upon data retrieval from the system, iterative decoding is first applied to correct errors on each individual disk. When a disk has too many errors, it can output soft information instead of hard information. Together with the decoding outputs from other disks, the outer array code can be used to correct more errors using soft decoding. Of course soft information is more expensive than hard information so there is a trade off between performance and information passed from the inner code to the outer code.

4 Conclusions

We have proposed a new high-density storage system using an array of high-density storage devices that store data in a 2-D manner. Issues pertaining to coding for such a system are discussed. The coding is done using a high-rate outer code that spreads



(a) (b)

Figure 6: (a) Comparison of LDPC codes and array codes for a [10000,5000] code. The curve on the left is the performance of the LDPC code on an AWGN channel with no ISI. The next two curves from the left are the performance using iterative MMSE equalization and decoding for the 2-D ISI channel and applying the Wiener filter once followed by decoding. The next two curves show the performance of the EVENODD code and the X-code on an AWGN channel with no ISI. The last four curves show the performance of the EVENODD and X-code for MMSE equalization and decoding scheme on the 2-D ISI channel. (b) Comparison of LDPC code and array codes for a [4489,4355] code. The curves from left to right are the performance of the LDPC, BR-93, BR-99, X-code on an AWGN channel with no ISI.

data across the devices followed by a low-rate inner code for error correction on each disk. Array codes are shown to be a good choice for the outer codes due to their simple algebraic structure and good decoding performance when decoded using message-passing. Joint equalization and decoding is used to retrieve data reliably from individual storage devices that have 2-D ISI. The benefit of passing soft information from the inner code to the outer code vis-a-vis the performance gain is also shown. Future work includes quantifying the complexity versus performance trade off between using the LDPC codes and array codes as outer codes. The trade off between the cost of passing more information from the inner code to the outer code and performance gain also has to be quantified.

References

[1] E. Soljanin and C. N. Georghiades, "Multihead detection for multi-track recording channels," *IEEE Trans. Inform. Theory*, vol. 44, No. 7, pp. 2988-2997, Nov. 1998.

[2] A. Kikawada, N. Honda, and K. Ouchi, "A multi-track coding scheme designed for perpendicular recording using double layer medium and MR head," *Journal of the Magnetics Society of Japan*, vol. 21, Supp S2, pp. 415-418, Oct. 1997.

[3] P. Vettiger *et al.*, "The 'Millipede'-more than one thousand tips for future AFM data storage," *IBM J. Res. Develop.* vol. 44, No. 3, pp. 323-340, May 2000.

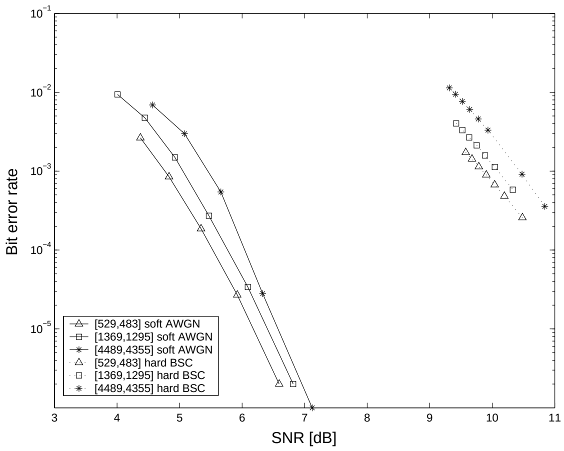


Figure 7: The performance comparison of X-codes of varying blocklengths for soft vs hard decoding. The three curves on the left are the performance of the X-code for different blocklengths on an AWGN channel using soft decoding. The three curves on the right are the performance of the same X-codes on a binary symmetric channel using the bit-flipping algorithm.

[4] W. Weeks IV, “Full surface data storage,” Ph.D. thesis, University of Illinois at Urbana Champaign, 2000.

[5] R. Krishnamoorthi, “Two-dimensional Viterbi like algorithms,” M.S. thesis, University of Illinois at Urbana Champaign, 1998.

[6] X. Chen, K. M. Chugg, and M. A. Neifeld, “Near-optimal parallel distributed data detection for page-oriented optical memories,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 4, No. 5, pp. 866-879, Sept./Oct. 1998.

[7] N. Singla, J. A. O’Sullivan, R. S. Indeck, and Y. Wu, “Iterative decoding and equalization for 2-D recording channels,” *IEEE Trans. on Magn.*, Sept. 2002.

[8] Y. Wu, J. A. O’Sullivan, R. S. Indeck, and N. Singla, “Iterative detection and decoding for separable two-Dimensional intersymbol interference,” submitted to *IEEE Trans. on Magn.*, Jun. 2002.

[9] D. A. Patterson, G. A. Gibson, and R. H. Katz, “A case for redundant array of inexpensive disks,” *Proc. SIGMOD Int. Conf. Data Management*, pp. 109-116, Chicago, 1998.

[10] S. Johnson and S. Weller, “Combinatorial constructions of LDPC codes,” *Proc. Information Theory Workshop*, Cairns, Australia, Sep. 2001.

[11] Y. Kou, S. Lin, and M. P. C. Fossier, “Low-density parity-check codes based on finite geometries: a rediscovery and new results,” *IEEE Trans. on Information Theory*, vol. 47, pp. 2711-2736, Nov. 2001.

[12] R. Michael Tanner, Deepak Sridhara, and Tom Fuja, “”A Class of Group-Structured LDPC Codes,” *Proc. ICSTA*, Ambleside, England, Jul. 2001.

[13] M. Blaum and R. M. Roth, “New array codes for multiple phased burst correction,” *IEEE Trans. on Information Theory*, vol. 39, No. 1, pp. 66-77, Jan. 1993.

[14] M. Blaum, J. Brady, J. Bruck, and J. Menon, "EVENODD: an efficient scheme for tolerating double disk failures in RAID architectures," *IEEE Trans. on Computers*, vol. 44, No. 2, pp. 192-202, Feb. 1995.

[15] R. M. Goodman, R. J. McEliece, and M. Sayano, "Phased burst error correcting array codes," *IEEE Trans. on Information Theory*, vol. 39, pp. 684-693, 1993.

[16] J. L. Fan, "Array codes as low-density parity-check codes," *Proc. Int'l Symp. on Turbo Codes*, Brest, France, pp. 543-546, Sept. 2000.

[17] L. Xu and J. Bruck, "X-code: MDS array codes with optimal encoding," *IEEE Trans. on Information theory*, vol. 45, No. 1, pp. 272-276, Jan. 1999.

[18] M. Blaum and R. M. Roth, "On lowest-density MDS codes," *IEEE Trans. on Information theory*, vol. 45, No. 1, pp. 46-59, Jan. 1999.