Combined Modulation and Error Correction Decoder for TDMR Using Generalized Belief Propagation

Graduate Student: Mehrdad Khatami
Advisor: Bane Vasić
Department of Electrical and Computer Engineering
University of Arizona, Tucson, AZ, 85721, USA
Email: {khatami,vasic}@ece.arizona.edu

Abstract

Constrained codes also known as modulation codes are a key component in the digital magnetic recording systems. The constrained codes forbid particular input data patterns which lead to some of the dominant error events or higher media noise. In data recording systems, a concatenated approach toward the constrained code and error-correcting code (ECC) is typically used and the decoding is done independently. In this paper, we show the improvement in combining the decoding of the constrained code and the ECC using generalized belief propagation (GBP) algorithm. We consider the performance of a combined modulation constraints and the ECC on a binary symmetric channel (BSC). We show that combining demodulation and decoding results in a superior performance compared to concatenated schemes. Furthermore, we compute the capacity of the joint ECC and modulation codes for 1-D and 2-D constraints.

Index Terms

Constrained Codes, Error-Control Codes (ECCs), Reverse Concatenation, Joint decoder/demodulator.

I. Introduction

Modulation codes are first introduced for data storage by Marcus *et al.* [1] to improve the performance of the system by matching the characteristics of the recorded signals to those of the channel. Various classes of modulation codes have been developed for different purposes in data storage systems. Run length limited (RLL) codes are one of the widely used family of constrained codes [2]. RLL codes are usually used in order to prevent the loss of clock synchronization. Maximum transition run (MTR) codes limits the maximum run of transitions written on the disk [3]. Due to the data dependent nature of the media noise, consecutive transitions leads to increasing the media noise which results in degradation of the performance.

Typically, modulation codes are followed by an error correction code (ECC) while the modulation code is designed to match the channel and the ECC is designed to correct the errors caused by the channel or even in detection. An error sequence in the input of the modulation decoder may result in error propagation since modulation codes unlike ECC codes are nonlinear in nature. Error propagation would result in overwhelming the error correcting capability of the ECC decoder. There are many works in order to provide a encoder that captures both error-correction and modulation properties. In order to solve the error propagation issue, there have been a lot of attention on reversed concatenation [4], where basically the encoders are concatenated in the reversed order. In this way, many channel errors can be corrected before reaching to the modulation decoder, and therefore decreasing the chance of error propagation.

In [5] and [6] , Wijngaarden and Immink introduced a novel concatenation scheme in which a modulation code is designed in a way so that specific positions are unconstrained such that those positions can take on any symbol without violating the constraint. In these codes, a systematic ECC is used and the the parity bits are inserted into the unconstrained positions without violating the constraint. Therefore the codewords satisfy both parity-check constraints and the modulation constraints. In a different approach, Vasic *et al.* [7] proposed a method in which deliberate bit-flipping is used to impose the k-constraint in the LDPC codewords. In this way by using strong codes with high minimum distance, it is expected that ECC would be able to correct both the deliberate errors and the channel errors.

In this paper, we propose a joint modulation/ECC decoder which performs sub-optimal decoding on the full graph using generalized belief propagation (GBP) [8]. The GBP decoder operates on an undirected graphical model. It infers the symbol *a posteriori* probabilities (APPs) using a message passing algorithm which takes into account the existence of loops in a graph. GBP decoder can be used with the modulation schemes in which the codewords generated by the scheme satisfy both the error-correction and modulation constraints in order to perform joint decoding such as the unconstrained position method [5]. Furthermore, the capacity of the low-pass constraints are calculated based on [9]. The capacity of the joint ECC/modulation code is also determined.

The paper is organized as follows: In Section II, the concatenation and the joint scheme of ECC and modulation decoders are described and the performance comparison is given. Section III defines the capacity problem in 1-D and 2-D modulation codes and provides the estimates of the capacity for local constraints. Section IV concludes the paper.

II. Joint Modulation/Decoding

In digital storage systems, the constrained code is followed by an error-control code (ECC). The inner modulation code serves the general function of matching the recorded signals to the physical channel, while the outer error-correction code is designed to remove any errors remaining after the detection and demodulation process. The constrained code and the ECC code are typically designed and decoded independently. In one type of concatenation, message bits are first encoded by ECC and then passed to the modulation codes. This type of concatenation is known as forward concatenation (FC). The problem with the forward concatenation is that the output of the constrained encoder is not necessarily an ECC codeword, and the decoding of the constrained code often results in error propagation which the outer ECC code could not able to correct. An alternative approach is provided by reverse concatenation (RC). The idea is to first encode the message bits using the modulation encoder, and then apply a systematic ECC encoder to produce parity bits. Since the parity bits do not necessarily meet the modulation constraint, they are then in turn encoded by the second modulation encoder. The block diagram of the reverse concatenation is shown in Fig. 1.

The separation of decoder and demodulator causes performance loss. Using GBP method for joint demodulation/decoding allows us to recover the performance loss due to separation of demodulation and decoding. The GBP demodulator performance is very close to the MAP demodulation. Furthermore, GBP can also be utilized to decode LDPC codes. An LDPC code

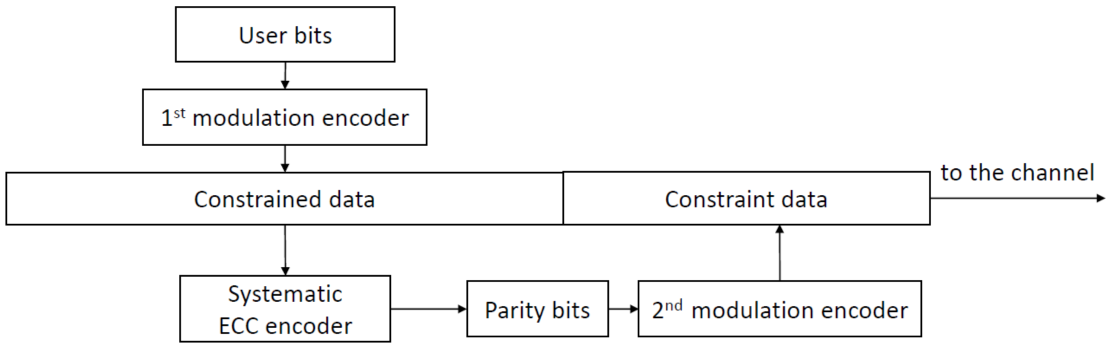

Fig. 1. Block diagram of the reverse concatenation technique

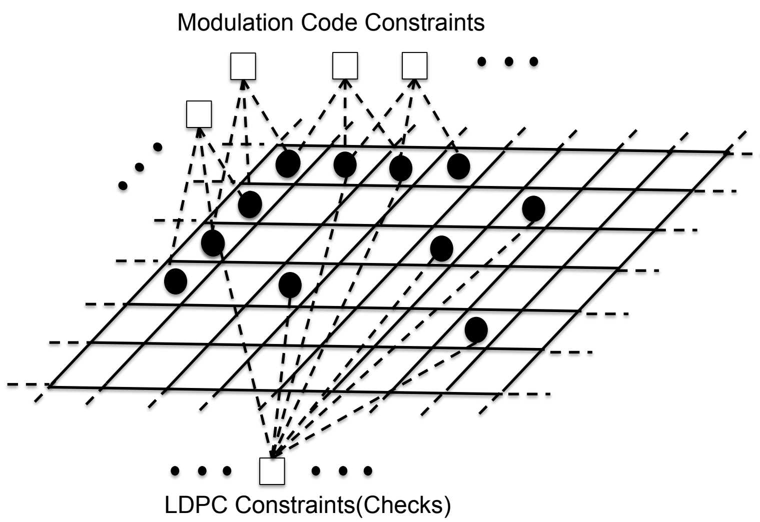

Fig. 2. Combined Tanner graph of the LDPC and the modulation code constraints for 2-D channel.

can be defined by the parity check matrix $\mathbf{H}_{m\times n}$ where there are m parity check constraints on n variable bits. These constraints can also be represented by a Tanner graph. By adjoining the modulation code constraints to the LDPC code constraints, GBP can be applied as a joint demodulator/decoder. Fig. 2 depicts a Tanner graph which is a combination of the Tanner graphs of the 2-D modulation code and the LDPC code. The upper squares represent the 2-D modulation code constraints and the lower squares represent the LDPC code constraints. Variable nodes are represented by circles inside the grid. This approach is general and it is not defined for 1-D or 2-D data and constraints. In order to get a sense of possible improvements of joint demodulator/decoder compared to concatenated demodulator and decoder, we simulated joint demodulation/decoding using GBP for 1-D case and RC is also simulated as a benchmark. The modulation code used in this simulation forbids two consecutive transitions. Let $\mathbf{C} \subset \mathbf{X}$ be the set of admissible patterns. The indicator function is defined as:

$$f(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \mathbf{C} \\ 0, & \mathbf{x} \notin \mathbf{C} \end{cases} \tag{1}$$

The constraint, which is imposed on $\mathbf{C}$, can be factored into smaller locally defined constraints which contain the three consecutive bits. The local constraints can be formulated as

$$f_a(x_i, x_{i+1}, x_{i+2}) = \begin{cases} 0, & \text{if } x_i = x_{i+2} \neq x_{i+1} \\ 1, & \text{else,} \end{cases} \tag{2}$$

where x_i, x_{i+1} and x_{i+2} are three consecutive bits and f_a is the local constraint. The indicator function for this code can be written as the product of local constraints of (2) each having some subset of $\mathbf{x}$ as arguments:

$$f(\mathbf{x}) = \prod_a f_a(\mathbf{x}_a) \tag{3}$$

The ECC code used in the simulations, is C(155,64) Tanner code of rate 0.41. The rate of the constrained code is 0.6977. In the joint scheme, we only send the codewords which satisfy our constraint. The overall rate is estimated by using GBP on the tanner graph of Fig. 2 based on [9] which is equal to 0.28. We used the same ECC code for RC method. The first constrained code is also the same. However, we used 0.5 rate modulation code for the secondary constrained code. Therefore, the overall block length of the codeword sent to the channel is $(155+91) = 246$ and the overall rate is 0.18 for the RC scheme. Fig. 3 shows the performance of these schemes. Although the block length used for RC is higher and also the rate is lower, the performance of the GBP joint demodulator /decoder is even slightly better.

III. Capacity of 2-D modulation codes

For TDMR systems, the constrained codes are viewed as 2-D codes. Therefore, we are investigating 2-D constraints to mitigate the media noise caused by grain size irregularities. Since the input transition between the neighbors is the source of media noise, we choose constraints based on 2-D low-pass constraints which forbid two consecutive transitions in both horizontal and vertical dimensions. In [10], Ashley *et al.* introduce a framework for designing encoders that transform arbitrary data sequences into two-dimensional arrays satisfying certain constraints, in particular, low-pass constraints. Since a finite state machine can not be defined for 2-D constraints, the decoding of these codes is not trivial.

In order to find the capacity of the constrained codes, the number of message patterns that satisfy the constraint must be estimated. The capacity of 1-D constrained code is given as:

$$C_{1D} = \lim_{n \to \infty} \frac{\log_2 Z(n)}{n} \tag{4}$$

where $Z(n)$, the partition function, denotes the number of 1-D message patterns of size n that satisfy the constraint.

Similar to the 1-D case, the 2-D capacity of modulation codes are as follows:

$$C_{2D} = \lim_{n,m \to \infty} \frac{\log_2 Z(m,n)}{mn} \tag{5}$$

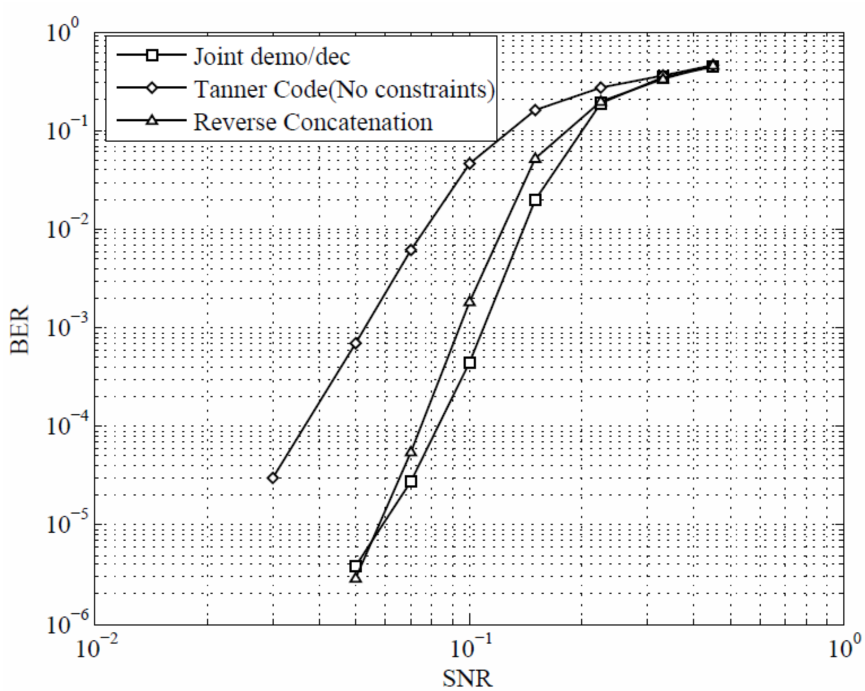

Fig. 3. Comparison of the GBP joint scheme and the RC method

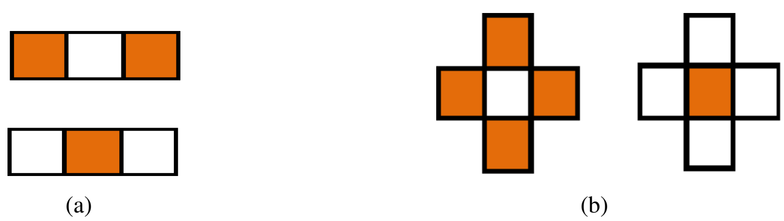

Fig. 4. (a) Forbidden pattern for the 1-D constrained code. (b) Forbidden pattern for the 2-D constrained code.

where $Z(m,n)$, the 2-D partition function, denotes the number of 2-D message patterns of size $m \times n$ that satisfy the constraint. In the general case, the partition function is defined as:

$$Z = \sum_{x \in \mathbf{X}} f(x) \tag{6}$$

where f is the indicator function of the set of admissible patterns. In statistical physics, the Helmholtz free energy is defined as:

$$F_H = -\ln(Z) \tag{7}$$

The partition function and the Helmholtz free energy are important quantities in statistical physics since they carry information about all thermodynamic properties of a system. Hence, computing the partition function is related to find the capacity of the constrained channel using (5). The Helmholtz free energy is then estimated using the region-based free energy approximation technique [9]. GBP is then used to compute the beliefs which in turn used to

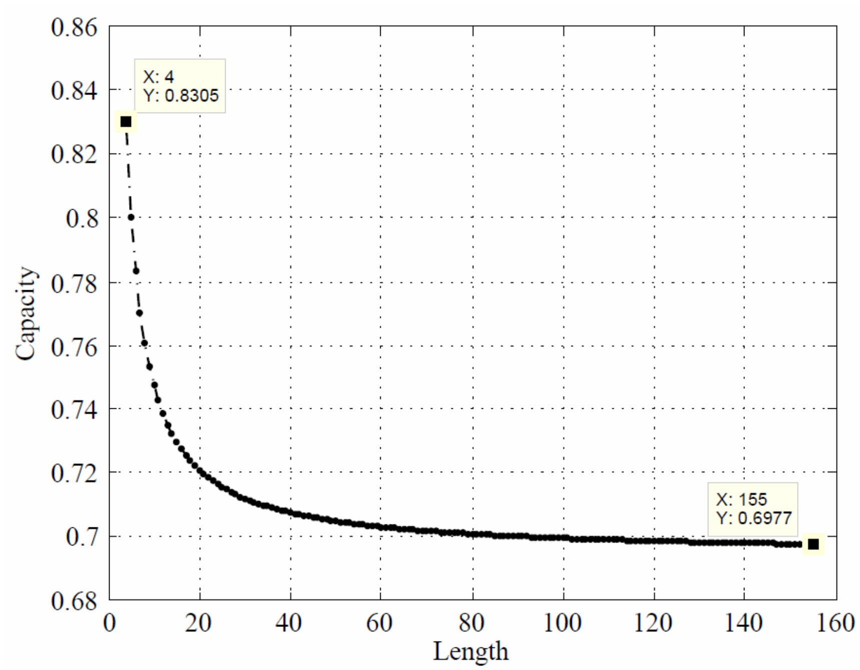

Fig. 5. Estimated capacity versus channel length for the 1-D constraints with the forbidden patterns shown in Fig. 4(a)

estimate $\hat{F}_H$. The region-based free energy $\hat{F}_H$ can then be used to estimate the partition function Z using (4). For more details, we refer the reader to [9].

Here we present the results of applying GBP to estimate the finite-sized noiseless capacity of 1-D and 2-D low-pass constraints. Fig. 4(a) shows the forbidden patterns which is defined in (2). For this constraint, Fig. 5 shows the capacity versus the channel width n over the interval $[4, 155]$.

Shown in Fig. 4(b) are the forbidden patterns of the 2-D constraint of which we estimate the capacity. For this constraint, Fig. 6 shows the capacity versus the channel width (n, n) over the interval $[4, 20]$.

IV. Conclusion

In this paper, we showed that by combining the decoding of the constrained code and the ECC significant improvement in performance can be achieved. We implemented the GBP algorithm as the joint decoder/demodulator and showed its superior performance compared to concatenated schemes such as reverse concatenation. Moreover, we compute the capacity modulation codes for local 1-D and 2-D constraints.

References

[1] B. Marcus, P. Siegel, and J. Wolf, "Finite-state modulation codes for data storage," *IEEE Journal on Selected Areas in Communications,*, vol. 10, no. 1, pp. 5–37, 1992.

[2] K. A. Schouhamer Immink, P. Siegel, and J. Wolf, "Codes for digital recorders," *IEEE Transactions on Information Theory,*, vol. 44, no. 6, pp. 2260–2299, 1998.

[3] J. Moon and B. Brickner, "Maximum transition run codes for data storage systems," *IEEE Transactions on Magnetics,*, vol. 32, no. 5, pp. 3992–3994, 1996.

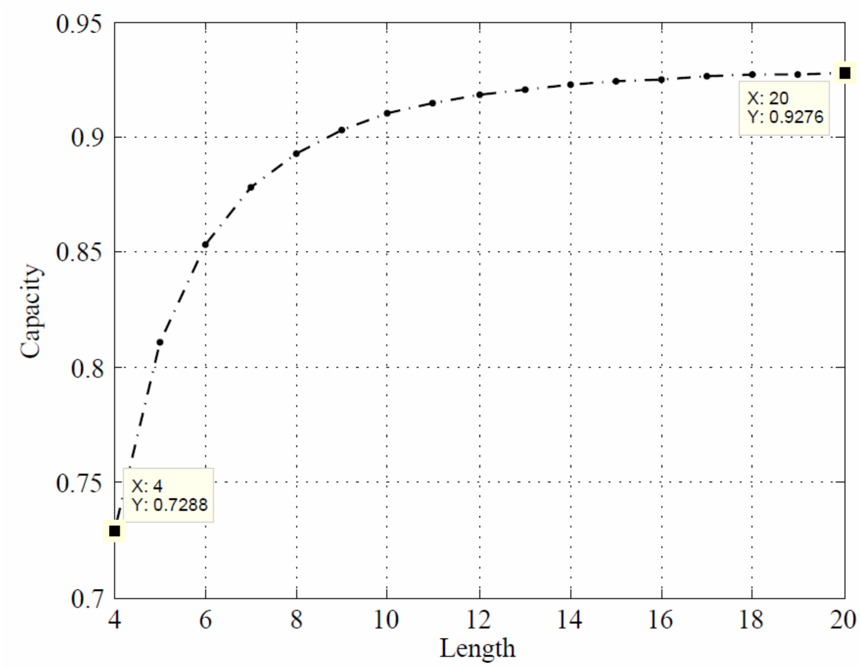

Fig. 6. Estimated capacity versus channel length for the 2-D constraints with the forbidden patterns shown in Fig. 4(b)

[4] W. G. Bliss, "Circuitry for performing error correction calculations on baseband encoded data to eliminate error propagation," *IBM Tech.*, vol. 23, p. 46334634, Mar. 1981.

[5] A. Van Wijngaarden and K. A. Schouhamer Immink, "Efficient error control schemes for modulation and synchronization codes," in *IEEE International Symposium on Information Theory,*, 1998, pp. 74–.

[6] ——, "Maximum runlength-limited codes with error control capabilities," *IEEE Journal on Selected Areas in Communications,*, vol. 19, no. 4, pp. 602–611, 2001.

[7] B. Vasic and K. Pedagani, "Run-length-limited low-density parity check codes based on deliberate error insertion," *IEEE Transactions on Magnetics,*, vol. 40, no. 3, pp. 1738–1743, 2004.

[8] J. Yedidia, W. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. on Inf. Theory*, vol. 51, no. 7, pp. 2282 – 2312, Jul. 2005.

[9] G. Sabato and M. Molkaraie, "Generalized belief propagation for the noiseless capacity and information rates of run-length limited constraints," *IEEE Transactions on Communications,*, vol. 60, no. 3, pp. 669–675, 2012.

[10] J. Ashley and B. Marcus, "Two-dimensional low-pass filtering codes," *IEEE Transactions on Communications,*, vol. 46, no. 6, pp. 724–727, 1998.