Data-dependent Write Channel Model for Magnetic Recording

Aravind R. Iyengar, Paul H. Siegel, and Jack K. Wolf
University of California, San Diego
9500 Gilman Dr. La Jolla CA 92093
email:{aravind,psiegel,jwolf}@ucsd.edu

***Abstract*—We propose a new channel model for the write channel in Magnetic Recording with Bit-Patterned Media. We study information theoretic propoerties of this channel and suggest a simplistic rate-$1/2$ coding scheme that achieves zero error. Based on this channel model, we propose a channel with insertion and deletion errors.**

I. Introduction

MAGNETIC Recording channels are typically studied as binary-input intersymbol interference (ISI) channels. However, the write channel–the transformation of information to be written to the data stored on a magnetic medium– is itself non-ideal. We aim to model the write channel and study the effects of its noisy nature. To this end, we proposed a simplistic concatenated channel model for the Bit-Patterned Media (BPM) channel in [1]. Here, we focus on the write channel and characterize the data dependent nature of the noise.

This paper is organized as follows. Section II explains the channel model considered and its relevance with the BPM write channel. We then characterize the information theoretic properties of the channel with substitution errors. Based on this channel model, we propose a channel that considers insertion and deletion errors in Section III and give its information theoretic properties. We summarize our findings in Section IV.

II. Write channel model with substitution errors

Let the input process $\mathscr{X}$ be a binary process, i.e. $X_i \in \{0,1\}\ \forall\ i = 1,2,\cdots,n$. The channel model can be written as

$$Y_i = X_i \oplus (X_i \oplus X_{i-1}) \otimes W_i, \qquad (1)$$

where $W_i \sim \mathcal{B}(\mathsf{p})$ is i.i.d. and independent of the input process $\mathscr{X}$, $\oplus$ and $\otimes$ represent addition and multiplication over $GF(2)$ respectively. We can write this as $Y_i = X_i \oplus N_i$, where the *noise process* $\mathscr{N}$ is given as $N_i = (X_i \oplus X_{i-1}) \otimes W_i$. This is a two-state channel where the present state $S_i = X_{i-1}$ with $S_1 = X_0 = 0$ by convention. The channel moves from one state to another based on the present input, and the state transition is accompanied by a probabilistic output – when the channel state changes (from 0 to 1, or vice versa) the output is flipped (with respect to the input) with a probability p, whereas when the channel state remains the same (i.e. state transition from 0 to 0, or 1 to 1) the output is the same as the input. Thus, the channel introduces errors only when the channel state changes. In other words, when the channel state changes ($X_i \oplus S_i = 1$), the output Y_i is equal to the input X_i with probability $1-\mathsf{p}$ (when $W_i = 0$) and the state S_i with probability p (when $W_i = 1$), whereas when the channel state remains the same ($X_i \oplus S_i = 0$), the output Y_i is equal to the input X_i with probability 1.

Relevance with the BPM write channel

The channel model in (1) depicts a BPM write channel. During the write process, the write head moves over the magnetic islands constituting a *track* on the disk. To write on a particular magnetic island, the write-head is positioned to fall within the *writing window zone* [2] of the island to be written. However, the write-head field potential affects, along with the magnetic state of the island presently being written, the magnetic states a few other islands that are yet to be written. Let us call the number of islands other than the present island influenced by the head as the *writing depth* W . When there is an error during the write process, it is assumed that the magnetic state of the present state is the only wrong state, i.e. the head is not off the island on which data is to be written by more than an island. The following example clarifies this.

Suppose $W = 3$ and the present information to be written is a 1. Let the state of the present island correspond to a 0 (so that it is to be flipped). Therefore, before writing, the states of the islands are $\times$ *0* $\times$ $\times$ $\times$ $\times$, where the state in italics is the present island and the $\times$s could be either 0 or 1. According to the assumed model, a correct writing would correspond to the next states of the islands being $\times$ *1* 1 1 1 $\times$, whereas an erroneous writing would result in $\times$ *0* 1 1 1 $\times$. Thus, when there is an error, only the present state is written erroneously while the rest of the islands within the writing depth are written correctly with the data corresponding to the present island (before being overwritten with their own data). Further, if the data on the present island before writing (which is same as X_{i-1}) was the same as the data to be written X_i, then there is no error even if the head misses writing on the present island. When the erroneous writing happens independently with probability p, (1) depicts this channel mathematically. Note that the channel model is independent of the writing depth W as long as $W \geq 1$.

*The work of A. R. Iyengar is supported by the National Science Foundation under the Grant CCF-0829865.

Information theoretic properties

Let us consider the capacity of the channel in (1). Since the channel has memory, by definition [3, Chap. 8], the infinite-letter characterization of capacity is

$$C = \lim_{n\to\infty} \sup_{\mathbb{P}\{X_1^n\}} \frac{1}{n} I\left(X_1^n; Y_1^n\right). \tag{2}$$

We will assume that the entropies and mutual informations are calculated in bits, so that the capacity is always measured in bits per channel use.

Consider

$$\begin{aligned}\frac{1}{n} I\left(X_1^n; Y_1^n\right) &= \frac{1}{n} H(Y_1^n) - \frac{1}{n} H(Y_1^n|X_1^n) \\ &\overset{(a)}{=} \frac{1}{n}\sum_{i=1}^{n} H(Y_i|Y_1^{i-1}) - \frac{1}{n}\sum_{i=1}^{n} H(Y_i|X_{i-1}^i) \\ &= \frac{1}{n}\sum_{i=1}^{n} H(Y_i|Y_1^{i-1}) \\ &\quad - \frac{h_2(\mathsf{p})}{n}\sum_{i=1}^{n} \mathbb{P}\{X_i \neq X_{i-1}\}, \end{aligned} \tag{3}$$

where the equality labelled (a) follows from the definition of Y_i in (1). Since the input distribution that maximizes the above quantity is unknown, we can examine the rate of information transfer under different assumptions of the input process.

A. I.I.D. Input process

We will first assume that the input process $\mathscr{X}$ is an i.i.d. $\mathcal{B}(\alpha)$ process. With this assumption, the maximum achievable information rate, called the *i.i.d. capacity*, denoted $C_{iid}(\alpha)$, gives a lower bound to the capacity. Thus, we have from (3)

$$\begin{aligned} C \geq C_{iid}(\alpha) &= \lim_{n\to\infty} \frac{1}{n} I(X_1^n; Y_1^n)\Big|_{\mathscr{X}\sim\mathcal{B}(\alpha)} \\ &= \mathcal{H}(\mathscr{Y})\Big|_{\mathscr{X}\sim\mathcal{B}(\alpha)} - 2\alpha(1-\alpha)h_2(\mathsf{p}). \end{aligned} \tag{4}$$

We can lower bound the capacity by discarding the memory in the channel, and combining the two states into a single state. This gives a channel equivalent to a BSC$(2\alpha\bar{\alpha}\mathsf{p})$ so that

$$C_{iid}(\alpha) \geq h_2(\alpha) - h_2\left(2\alpha\bar{\alpha}\mathsf{p}\right) = L_0^{iid}, \tag{5}$$

where $h_2(\cdot)$ is the *binary entropy function*. We can obtain a tighter lower bound for the i.i.d. capacity as follows. We will drop the description of the input process whenever it is clear from the context, for convenience.

$$\begin{aligned} C_{iid}(\alpha) &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} H(Y_i|Y_1^{i-1}) - 2\alpha(1-\alpha)h_2(\mathsf{p}) \\ &\geq \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} H(Y_i|Y_1^{i-1}, X_{i-1}) \\ &\quad - 2\alpha(1-\alpha)h_2(\mathsf{p}) \\ &= \bar{\alpha}h_2(\alpha\bar{\mathsf{p}}) + \alpha h_2(\bar{\alpha}\bar{\mathsf{p}}) - 2\alpha\bar{\alpha}h_2(\mathsf{p}) \\ &= L_1^{iid}, \end{aligned} \tag{6}$$

where we have used the fact that Y_i depends only on X_{i-1} and X_i, and given X_{i-1}, $Y_i \perp\!\!\!\perp Y_1^{i-1}$, and $\bar{a} = 1 - a \ \forall \ a \in [0, 1]$.

We can obtain a tighter lower bound

$$\begin{aligned} C_{iid}(\alpha) &\geq \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} H(Y_i|Y_1^{i-1}, X_{i-2}) - 2\alpha\bar{\alpha}h_2(\mathsf{p}) \\ &= (\bar{\alpha}^2 + \alpha\bar{\alpha}\mathsf{p})h_2\left(\frac{\alpha(\bar{\alpha}\bar{\mathsf{p}} + \alpha\mathsf{p}) + \alpha\bar{\alpha}\mathsf{p}^2}{\bar{\alpha} + \alpha\mathsf{p}}\right) + \\ &\quad (\alpha^2 + \alpha\bar{\alpha}\mathsf{p})h_2\left(\frac{\bar{\alpha}(\alpha\bar{\mathsf{p}} + \bar{\alpha}\mathsf{p}) + \alpha\bar{\alpha}\mathsf{p}^2}{\alpha + \bar{\alpha}\mathsf{p}}\right) + \\ &\quad \alpha\bar{\alpha}\bar{\mathsf{p}}\left(h_2(\alpha\bar{\mathsf{p}}) + h_2(\bar{\alpha}\bar{\mathsf{p}})\right) - 2\alpha\bar{\alpha}h_2(\mathsf{p}) \\ &= L_2^{iid}. \end{aligned} \tag{7}$$

A straightforward upper bound to the i.i.d. capacity, from (4), is that

$$C_{iid}(\alpha) \leq 1 - 2\alpha\bar{\alpha}h_2(\mathsf{p}) = U_0^{iid}, \tag{8}$$

which follows because the entropy rate for a binary process $\mathcal{H}(\mathscr{Y}) \leq 1$. Note that this bound is achieved when $\mathscr{Y}$ is the iid $\mathcal{B}(1/2)$ process. Again, from (4),

$$\begin{aligned} C_{iid}(\alpha) &\leq \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} H(Y_i|Y_{i-1}) - 2\alpha\bar{\alpha}h_2(\mathsf{p}) \\ &= \bar{\alpha}h_2\left(\alpha(\bar{\mathsf{p}} + \mathsf{p}^2)\right) + \alpha h_2\left(\bar{\alpha}(\bar{\mathsf{p}} + \mathsf{p}^2)\right) \\ &\qquad - 2\alpha\bar{\alpha}h_2(\mathsf{p}) \\ &= U_1^{iid}. \end{aligned} \tag{9}$$

We can find a tighter upper bound for the entropy rate $\mathcal{H}(\mathscr{Y})$ as follows

$$\begin{aligned} C_{iid}(\alpha) &\leq \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} H(Y_i|Y_{i-2}^{i-1}) - 2\alpha\bar{\alpha}h_2(\mathsf{p}) \\ &= \left(\bar{\alpha}^4 + \bar{\alpha}^3\alpha + \bar{\alpha}^2\alpha + \bar{\alpha}\alpha\mathsf{p}\bar{\mathsf{p}}\right) \\ &\quad h_2\left(\frac{\bar{\alpha}^2\alpha(1-\mathsf{p}\bar{\mathsf{p}}) + \bar{\alpha}\alpha^2 + \alpha^3\mathsf{p}\bar{\mathsf{p}}}{\bar{\alpha}^3 + \bar{\alpha}^2\alpha + \bar{\alpha}\alpha + \alpha\mathsf{p}\bar{\mathsf{p}}}\right) + \\ &\quad \bar{\alpha}\alpha(1-\mathsf{p}\bar{\mathsf{p}})h_2\left(\frac{\bar{\alpha}\alpha + \bar{\alpha}^2\mathsf{p}\bar{\mathsf{p}} + \alpha^2(1-\mathsf{p}\bar{\mathsf{p}})}{1-\mathsf{p}\bar{\mathsf{p}}}\right) + \\ &\quad \bar{\alpha}\alpha(1-\mathsf{p}\bar{\mathsf{p}})h_2\left(\frac{\bar{\alpha}\alpha + \bar{\alpha}^2(1-\mathsf{p}\bar{\mathsf{p}}) + \alpha^2\mathsf{p}\bar{\mathsf{p}}}{1-\mathsf{p}\bar{\mathsf{p}}}\right) + \\ &\quad \left(\bar{\alpha}^3\alpha\mathsf{p}\bar{\mathsf{p}} + \bar{\alpha}^2\alpha^2(1+2\mathsf{p}\bar{\mathsf{p}}) + \bar{\alpha}\alpha^3(2+\mathsf{p}\bar{\mathsf{p}}) + \alpha^4\right) \\ &\quad h_2\left(\frac{\bar{\alpha}^3\mathsf{p}\bar{\mathsf{p}} + \bar{\alpha}^2\alpha + \bar{\alpha}\alpha^2(1-\mathsf{p}\bar{\mathsf{p}})}{\bar{\alpha}^3\mathsf{p}\bar{\mathsf{p}} + \bar{\alpha}^2\alpha(1+2\mathsf{p}\bar{\mathsf{p}}) + \bar{\alpha}\alpha^2(2+\mathsf{p}\bar{\mathsf{p}}) + \alpha^3}\right) \\ &\quad - 2\alpha\bar{\alpha}h_2(\mathsf{p}) \\ &= U_2^{iid}. \end{aligned} \tag{10}$$

Figure 1 shows the contours of the bounds discussed here. Also shown, in dotted lines, is the α that maximizes the corresponding bound for a given p. As is clear from the figure, and as can be expected from the symmetry in the channel, the information rate maximizing input process is the process with equally likely inputs for the entire range of p values.

Symmetric Information Rate: From the observation made in the last section, it is clear that the maximum information rate is achieved for an i.i.d. input process when the inputs are equally likely. In this section, we give the bounds obtained in the previous section for the special case of $\alpha = \frac{1}{2}$, wherein

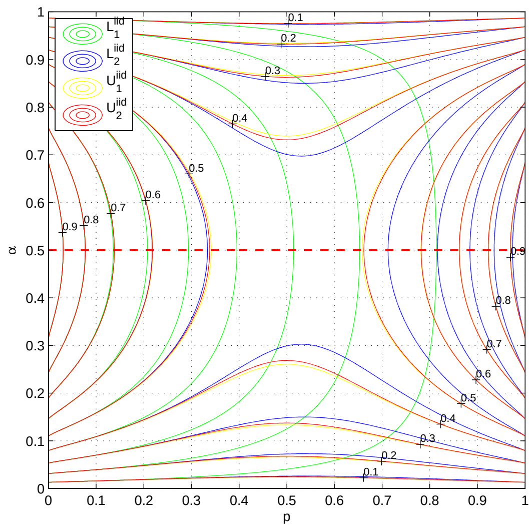

Fig. 1. Lower and upper bounds for the I.I.D. Capacity given in (6), (7), (9) and (10).

the achievable rate is referred to as the *Symmteric Information Rate* (SIR), denoted[1] C_{iud}.

$$C \geq \max_{\alpha} C_{iid}(\alpha) = C_{iid}(0.5) = C_{iud}$$
$$= \lim_{n\to\infty} \frac{1}{n} I(X_1^n; Y_1^n)\Big|_{\mathbb{P}\{X_1^n\}=2^{-n}}$$
$$= \mathcal{H}(\mathscr{Y}) - \frac{h_2(\mathsf{p})}{2}. \tag{11}$$

The SIR is of special interest because it can be achieved using linear coset codes [4, §6.2]. We now give the bounds obtained in the previous section for the SIR. We start with the lower bound $C_{iud} \geq 1 - h_2\left(\frac{\mathsf{p}}{2}\right) = L_0^{iud}$. The lower bound L_1^{iid} becomes

$$C_{iud} \geq h_2\left(\frac{1-\mathsf{p}}{2}\right) - \frac{h_2(\mathsf{p})}{2} = L_1^{iud}, \tag{12}$$

and similarly,

$$C_{iud} \geq \left(\frac{1+\mathsf{p}}{2}\right) h_2\left(\frac{1+\mathsf{p}^2}{2(1+\mathsf{p})}\right) + \left(\frac{1-\mathsf{p}}{2}\right) h_2\left(\frac{1-\mathsf{p}}{2}\right) - \frac{h_2(\mathsf{p})}{2} = L_2^{iud}. \tag{13}$$

The corresponding upper bounds are

$$C_{iud} \leq 1 - \frac{h_2(\mathsf{p})}{2} = U_0^{iud}, \tag{14}$$

$$C_{iud} \leq h_2\left(\frac{1-\mathsf{p}+\mathsf{p}^2}{2}\right) - \frac{h_2(\mathsf{p})}{2} = U_1^{iud}, \tag{15}$$

and

$$C_{iud} \leq \frac{1-\mathsf{p}+\mathsf{p}^2}{2} h_2\left(\frac{1}{2(1-\mathsf{p}+\mathsf{p}^2)}\right) + \frac{1+\mathsf{p}-\mathsf{p}^2}{2} h_2\left(\frac{1}{2(1+\mathsf{p}-\mathsf{p}^2)}\right) - \frac{h_2(\mathsf{p})}{2} = U_2^{iud}. \tag{16}$$

[1]"iud" for independent, uniformly distributed input

Figure 2 plots the lower bounds to capacity (and SIR) and the upper bounds to SIR discussed here. A couple of

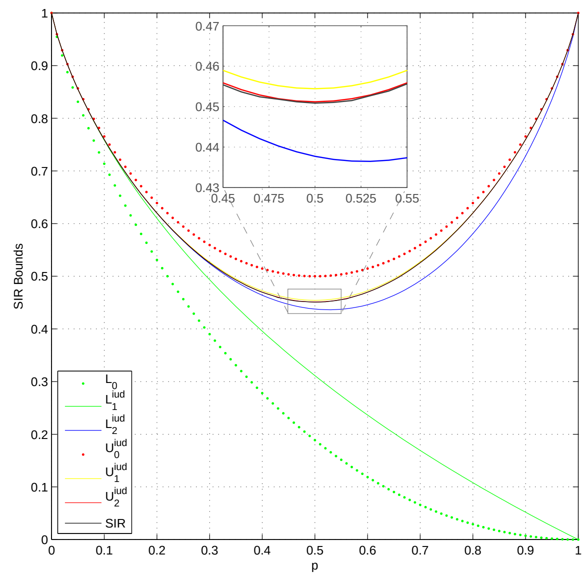

Fig. 2. Lower and upper bounds for SIR given in (5), (12), (13), (14), (15) and (16). The SIR from simulations is also shown.

observations are in order. First, note that the symmetry in p for the upper bounds suggests a symmetric SIR (and capacity). This can be seen by noting that when $\mathsf{p} = 1$, the output process $\mathscr{Y}$ is the same as the input process $\mathscr{X}$ delayed by one time instant. This suggests that when the channel parameter $\mathsf{p} > 0.5$, we can decode the output to the channel state rather than the channel input. The symmetry of the bounds in p can be seen from the following. The channel model can be restated as

$$\begin{aligned} Y_i &= X_i \oplus (X_i \oplus X_{i-1}) \otimes W_i \\ &= X_i \otimes (1 \oplus W_i) \oplus X_{i-1} \otimes W_i \\ &= X_i \otimes \bar{W}_i \oplus X_{i-1} \otimes W_i \\ &= X_i \otimes \bar{W}_i \oplus S_i \otimes W_i, \end{aligned}$$

so that when $\mathsf{p} < 0.5$, the output process $\mathscr{Y}$ can be decoded to the input process $\mathscr{X}$ and when $\mathsf{p} > 0.5$, the output can be decoded to the state process $\mathscr{S}$.

The second observation is that the tighter lower and upper bounds (L_2^{iud} in (13) and U_2^{iud} in (16)) almost meet for $\mathsf{p} \leq 0.3$ (and from symmetry, for $\mathsf{p} \geq 0.7$), i.e. when the bounds themselves are > 0.5, so that the SIR is fairly accurately given by these bounds in this range. It was suggested in [5] with reference to the binary ISI channel that the SIR in this range (> 0.8) is close to the best known lower bounds for capacity of the binary ISI channel. Given the similarity in the memory involved in the channel in (1) to that in the binary ISI channel, it would not be unreasonable to expect the same here.

The process $\mathscr{Y}$ is ergodic (as all channel states can be reached within a finite number of steps at any time [6, §5.3]), and converges to a stationary process. As a result, from the Asymptotic Equipartition Property [3, §4.2], we have the entropy rate

$$\mathcal{H}(\mathscr{Y}) = \lim_{n\to\infty} \frac{1}{n} H(Y_1^n) = -\lim_{n\to\infty} \frac{1}{n} \log_2 \mathbb{P}\{Y_1^n\},$$

which can be numerically evaluated by simulations [6, §5.3]. By simulating using long enough sequences X_1^n and Y_1^n, the SIR can be obtained fairly accurately. This is shown as the "SIR" in Figure 2, from which we conclude that the upper bound U_2 in (16) is a good approximation for the SIR.

B. *Markov*-1 *Source*

To explore the loss in terms of the maximum achievable rate due to the usage of linear coset codes, we consider a source with memory. The simplest source with memory is when the process $\mathscr{X}$ is a First-order Markov process. Since the channel has no bias for 0's and 1's, we consider a symmetric Markov process described as $\mathscr{X} \in \{0,1\}^*$ with $\mathbb{P}\{X_i \neq X_{i-1}\} = \beta$. We write $\mathscr{X} \sim \mathcal{M}_1^{(2)}(\beta)$ to mean that $\mathscr{X}$ is a binary Markov source with memory 1 and transition parameter β.

Starting from Equation (3), we can arrive at a lower bound similar to L_1 in (12) for the *Markov*-1 *Capacity* C_{M1}, which we define as the maximum rate of information transfer when $\mathscr{X} \sim \mathcal{M}_1^{(2)}(\beta)$, as

$$\begin{aligned} C_{M1} &= \mathcal{H}(\mathscr{Y}) - \beta h_2(\mathsf{p}) \\ &\geq \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} H(Y_i|Y_1^{i-1}, X_{i-1}) - \beta h_2(\mathsf{p}) \\ &= h_2(\beta\bar{\mathsf{p}}) - \beta h_2(\mathsf{p}) = L_1^{M1}. \end{aligned} \tag{17}$$

A tighter lower bound follows

$$\begin{aligned} C_{M1} &\geq \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} H(Y_i|Y_1^{i-1}, X_{i-2}) - \beta h_2(\mathsf{p}) \\ &= (1-\beta\bar{\mathsf{p}})\, h_2\left(\frac{\beta\left(1-\beta(1-\mathsf{p}^2)\right)}{1-\beta\bar{\mathsf{p}}}\right) \\ &\quad + \beta\bar{\mathsf{p}} h_2(\beta\bar{\mathsf{p}}) - \beta h_2(\mathsf{p}) = L_2^{M1}. \end{aligned} \tag{18}$$

The trivial upper bound is

$$C_{M1} \leq 1 - \beta h_2(\mathsf{p}) = U_0^{M1}. \tag{19}$$

An upper bound similar to U_1^{iid} in Equation (9) is

$$C_{M1} \leq h_2\left(1-\beta+2\beta^2\mathsf{p}\bar{\mathsf{p}}\right) - \beta h_2(\mathsf{p}) = U_1^{M1}. \tag{20}$$

The upper bound analogous to U_2^{iid} in (10) is

$$\begin{aligned} C_{M1} \leq &\left\{\beta^2\bar{\beta}(1-\mathsf{p}\bar{\mathsf{p}}) + \beta^2(3-\beta)\mathsf{p}\bar{\mathsf{p}} + (1+\beta)\bar{\beta}^2\right\} \\ &h_2\left(\frac{\beta^2\bar{\beta}(1-\mathsf{p}\bar{\mathsf{p}}) + \beta^3\mathsf{p}\bar{\mathsf{p}} + \beta\bar{\beta}^2}{\beta^2\bar{\beta}(1-\mathsf{p}\bar{\mathsf{p}}) + \beta^2(3-\beta)\mathsf{p}\bar{\mathsf{p}} + (1+\beta)\bar{\beta}^2}\right) + \\ &\left\{2\beta^2\bar{\beta}(1-\mathsf{p}\bar{\mathsf{p}}) + \beta^3(1-2\mathsf{p}\bar{\mathsf{p}}) + \beta\bar{\beta}^2\right\} \\ &h_2\left(\frac{\beta^2\bar{\beta}(1-\mathsf{p}\bar{\mathsf{p}}) + \beta^3\mathsf{p}\bar{\mathsf{p}} + \beta\bar{\beta}^2}{2\beta^2\bar{\beta}(1-\mathsf{p}\bar{\mathsf{p}}) + \beta^3(1-2\mathsf{p}\bar{\mathsf{p}}) + \beta\bar{\beta}^2}\right) - \beta h_2(\mathsf{p}) \\ = &\, U_2^{M1}. \end{aligned} \tag{21}$$

Figure 3 shows the contours of the bounds for C_{M1} in (17), (18), (20) and (21). Also shown (in dotted lines) is the β that maximizes the corresponding bound for a given p. Denoting

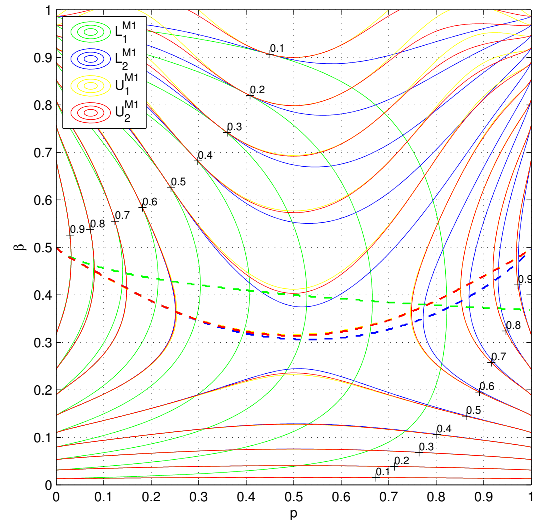

Fig. 3. Lower and upper bounds for Markov-1 Capacity given in (17), (18), (20) and (21).

the Markov-1 capacity maximizing source parameter by $\beta^*(\mathsf{p})$, we can see that $\beta^*(\mathsf{p})$ is monotonically decreasing in p in the interval $[0, 0.5]$ with $\beta^*(0) = 0.5$ and $\beta^*(0.5) \approx 0.31$.

Figure 4 compares the SIR (Solid line representing L_2^{iud} in (13), and dashed line representing U_2^{iud} in (16)) and the Markov-1 Capacity (Solid line – $L_2^{M1}(\beta_{L_2}^*)$ in (18), dashed line – $U_2^{M1}(\beta_{U_2}^*)$ in (21)) over the range of p values. It is

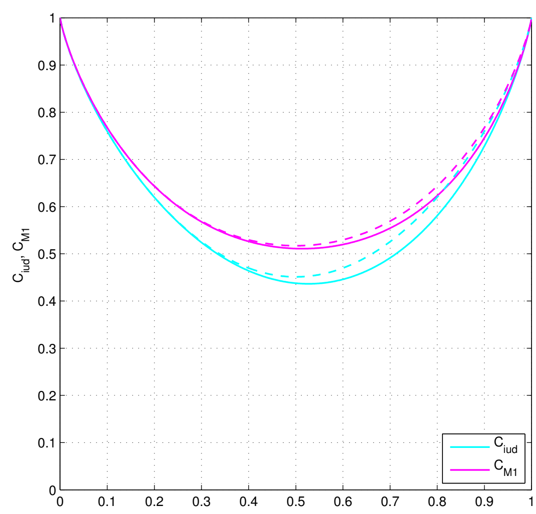

Fig. 4. Comparison between SIR and Markov-1 Capacity.

clear that considerable gains in reliable information transfer rate is possible by using an input with memory.

Zero-error achieving code: Consider a rate-0.5 convolutional code with the message bit at instant i M_i and two coded bits $C_i^1 = C_i^2 = M_i$, so that the corresponding message process $\mathscr{M} = (M_1, M_2, \cdots, M_i, \cdots)$. The channel input process $\mathscr{X}$ is then given as $X_{2i-1} = C_i^1 = M_i, X_{2i} = C_i^2 = M_i$ for $i = 1, 2, \cdots$. Note that since $X_{2i-1} = X_{2i}$, $Y_{2i} = X_{2i} = M_i$ so that discarding the Y_{2i-1}s gives us $\mathscr{M}$ exactly, thereby achieving zero error.

III. Write channel model with insertion-deletion errors

We now consider a channel with insertions and deletions that is built on the channel model in Section II. We first note that, in (1), when $\mathsf{p} = 0$, $Y_i = X_i$ so that the output reproduces the input exactly, and when $\mathsf{p} = 1$, $Y_i = X_{i-1}$ so that the output reproduces the input exactly, but with a delay of one time instant (The channel therefore has memory). As expected, and as is clear from the capacity bounds obtained in the previous section, the capacities of both these channels (with $\mathsf{p} = 0$ and $\mathsf{p} = 1$) are 1 bit per channel use. Let us call these cases the *identity channel* ($\mathsf{p} = 0$) and the *delayed channel* ($\mathsf{p} = 1$).

Let us now consider a two-state channel, i.e., the state space is $\Sigma = \{0, 1\}$. In state 0, the channel is the identity channel and in state 1, the channel is the delayed channel. The channel moves between the states according to a Markov process $\mathscr{S} \sim \mathcal{M}_1^{(2)}(\gamma)$, i.e., $\mathbb{P}\{\Sigma_i = 0 | \Sigma_{i-1} = 1\} = \mathbb{P}\{\Sigma_i = 1 | \Sigma_{i-1} = 0\} = \gamma$. This model is an insertion-deletion channel, as insertion of a bit accompanies the transition of the channel from state 0 to 1, and a bit deletion accompanies the transition from state 1 to 0. This is illustrated in Table I.

Σ_i	0	0	0	1	1	1	1	0	0	0
X_i	1	1	1	0	0	0	1	0	1	1
Y_i	1	1	1	1	0	0	0	0	1	1

TABLE I
An insertion-deletion example.

The example shows the inserted bit in $\mathscr{Y}$ in green when the channel state changes from 0 to 1, and the deleted bit from $\mathscr{X}$ in red when the channel state changes from 1 to 0. We will call this channel the *Symmetric* 1 *insertion-deletion channel*, Symmetric 1 InsDel Channel for short; symmetric because insertions and deletions are equally likely (depending on the channel state), and 1 insertion-deletion channel because there can be a maximum of 1 consecutive insertions or deletions.

Figure 5 shows the SIR obtained for the Symmetric 1-InsDel Channel. The extremes of the SIR when $\gamma = 0$ and $\gamma = 1$ are easily explained. When $\gamma = 0$, the channel is nothing but the identity channel and every bit goes through uncorrupted. When $\gamma = 1$, the channel deterministically flip-flops between the identity and the delayed channel so that every odd bit is repeated twice, and every even bit is lost, and the maximum achievable information transfer rate is 0.5 bits per channel use.

Asymmetric **1** *insertion-deletion Channel:* We now consider the case where $\mathbb{P}\{\Sigma_i = 1 | \Sigma_{i-1} = 0\} = \mathsf{p}_i$ and $\mathbb{P}\{\Sigma_i = 0 | \Sigma_{i-1} = 1\} = \mathsf{p}_d$ are not necessarily equal. In this case, the SIR can be found as in the symmetric case, and is given in Figure 6.

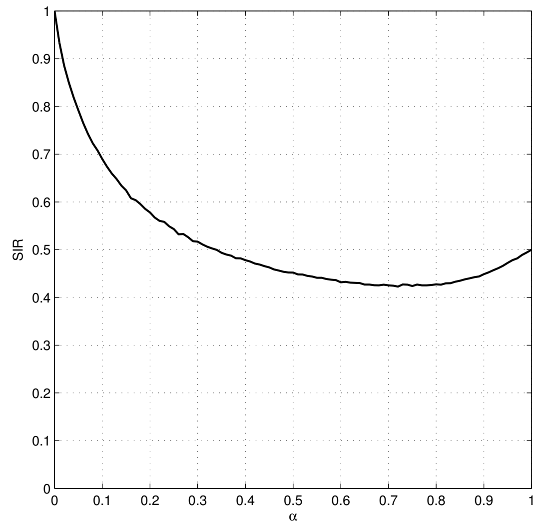

Fig. 5. SIR for the Symmetric 1 InsDel channel.

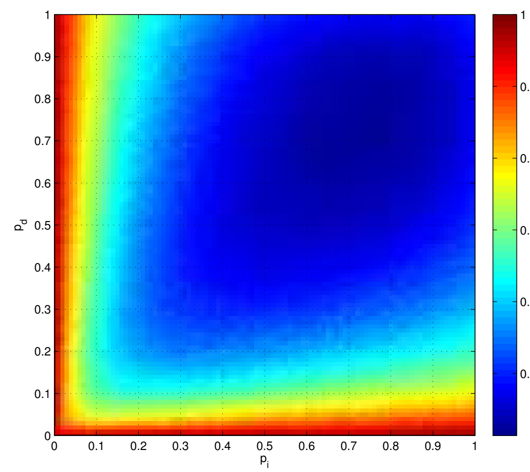

Fig. 6. SIR for the Asymmetric 1 InsDel Channel.

IV. Conclusions

We proposed a new channel model for the write channel in Bit-Patterned Media that models substitution errors or insertion-deletion errors. We studied its information theoretic properties by giving bounds to the capacity, i.i.d. capacity, symmetric information rate and the Markov-1 rates; we gave a rate-0.5 code construction that achieves zero error over this channel. Using this channel model in the study of Magnetic recording channels and exploring good code constructions constitutes further work in this direction.

References

[1] A. R. Iyengar, P. H. Siegel and J. K. Wolf, "LDPC Codes for the Cascaded BSC-BAWGN channel," *Proceedings of the 47th Annual Allerton Conference on Communications, Control and Computing*, Sep. 2009.

[2] B. Livshitz, A. Inomata, H. N. Bertram, V. Lomakin, "Semi-analytical approach for analysis of BER in conventional and staggered bit patterned media," *IEEE Trans. Magn.*, 2009.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley India, 2008.

[4] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley and Sons, 1968.

[5] J. B. Soriaga, H. D. Pfister, and P. H. Siegel, "Determining and approaching achievable rates of binary intersymbol interference channels using multistage decoding," *IEEE Trans. Inform. Theory*, vol. 53, no. 4, pp. 1416–1429, Apr. 2007.

[6] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.