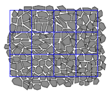
Codes for High Density Magnetic Memory

Alexander Barg (joint work with Arya Mazumdar, Navin Kashyap)

ISR@25 — The Institute for Systems Research — A. James Clark School of Engineering — University of Maryland

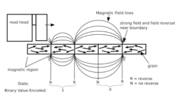
Granularity in magnetic medium

- The recording medium on a hard disk drive (HDD) is a thin film of magnetic material.
- The magnetic medium is physically composed of fundamental magnetizable units, called grains, of irregular shapes and sizes.



Grains comprising a portion of the magnetic surface

- The granular magnetic medium is conceptually divided into many sub-micrometer-sized bit cells, each of which records a single bit of data. In current HDD technology, each bit cell contains several hundred grains.
- The write head records data on the medium by magnetizing each bit cell directionally, to represent either a 0 or a 1; all grains within a bit cell acquire the same magnetic polarity.



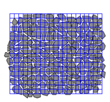
Cross-section of magnetic material (figure courtesy Allan Haldane @ Wikipedia)

- The read head reads the data back by detecting the magnetic polarity of each bit cell.

Towards Ultra-High-Density Magnetic Recording

- Areal density is the number of bits of data stored per square inch of medium.
- Currently, HDDs with areal density of 300-400 Gbits/sq.in. are commercially available.

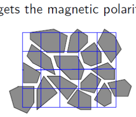
  The push is on towards 1 Terabit/sq.in.
- It is believed that 10 Tbits/sq.in. is achievable.
- As areal density increases, the bit cells must get smaller.
- Eventually, bit cells must become as small as grains.
- Since each grain can store only one bit of data, there is a fundamental one-bit-per-grain limit to the storage capacity of the magnetic medium.



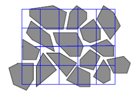
Bottleneck ...

..in achieving the one-bit-per-grain limit

- The read/write head is typically unaware of the shapes and positions of the grains in the medium.
- Bits get written into the bit cells of the medium in some prescribed order (say, raster scan).
- At each step of the write process, if a grain has significant (say, > 30%) overlap with the bit cell being currently written, then that grain gets the magnetic polarity of that bit cell.



The Effect of Not Knowing Grain Boundaries



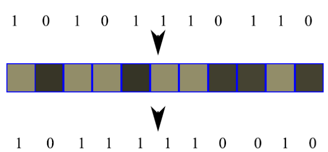
The read head detects the polarity value of a small region around the center of each bit cell.

The net effect is that some bit cells may get overwritten by the polarity values of subsequent bit cells;

A Simplified 1-D Model

Grey-shaded regions are grains; blue lines demarcate bit cells

- 1-dimensional medium consisting of n bit cells
- grain boundaries fall on bit cell boundaries
- bits are written from right to left
- the last bit to be written within a grain overwrites all bits previously written within the same grain
- all grains are 1 or 2 bit cells long
- the above assumption means that the sizes of the grains does not vary beyond a certain limit.



The "Grains Channel"

We consider a probabilistic model of a "grains channel":

- define $\gamma_j = \begin{cases} 1 & \text{if a length-2 grain ends at position } j; \\ 0 & \text{otherwise.} \end{cases}$
- $P(\gamma_j = 1 \mid \gamma_{j-1} = 0) = p$

  $P(\gamma_j = 1 \mid \gamma_{j-1} = 1) = 0$

Channel Input: $\mathbf{x} = (x_j)_{j\in\mathbb{N}}, \quad x_j \in \{0,1\}$

Channel Output: $\mathbf{y} = (y_j)_{j\in\mathbb{N}}, \quad y_j \in \{0,1\}$

Channel State: $\mathbf{s} = (s_j)_{j\in\mathbb{N}}, \quad s_j = (\gamma_j, x_j \oplus x_{j-1})$

Input-Output Transition: $y_j \neq x_j$ iff $s_j = (1,1)$

- What is the maximum possible rate of information storage i such a medium?
- Is that rate achievable?
- Can we give an explicit answer as a function of p?

The grain channel is a finite-state channel, there are two channel capacities, the *lower* (or *pessimistic*) *capacity* $\underline{C} = \lim_{n\to\infty} \underline{C}_n$, and *upper* (or *optimistic*) *capacity* $\overline{C} = \lim_{n\to\infty} \overline{C}_n$, where

$$\underline{C}_n = n^{-1} \max_{Q^n(\mathbf{x}^n)} \min_{s_0} I(\mathbf{x}^n; \mathbf{y}^n \mid s_0)$$

$$\overline{C}_n = n^{-1} \max_{Q^n(\mathbf{x}^n)} \max_{s_0} I(\mathbf{x}^n; \mathbf{y}^n \mid s_0).$$

with:

- $\mathbf{x}^n = (x_1, \ldots, x_n)$: length-$n$ input.
- $\mathbf{y}^n = (y_1, \ldots, y_n)$: length-$n$ output.
- s_0: the initial state
- $Q^n(\mathbf{x}^n)$: set of probability distributions on the input $\mathbf{x}^n$.

Capacity

- The grains channel is indecomposable: the effect of initial state dies away with time.
- $\overline{C} = \underline{C}$.
- However, It is difficult to explicitly compute the capacity of the grain channel.

Lower Bound: Achievable rate with uniform random codes

- $\mathbf{x}$ is an i.i.d. Bernoulli(1/2) random sequence.
- $I(\mathbf{x}^n; \mathbf{y}^n) = H(\mathbf{y}^n) - H(\mathbf{y}^n|\mathbf{x}^n)$, the terms can be evaluated with some effort.

$$H(\mathbf{y}^n|\mathbf{x}^n) = \frac{1+p/2}{1+p} \sum_{j=2}^{\infty} \left(\frac{1}{2}\right)^j \mathrm{h}\left(\frac{1-(-p)^j}{1+p}\right).$$

$$H(\mathbf{y}^n) = \frac{1}{2(1+p)} \sum_{j=2}^{\infty} \mathrm{h}(\beta_j) \prod_{k=2}^{j-1} (1-\beta_k),$$

where

$$\beta_j := \Pr[y_{j+1} = 1 \mid y_j = y_{j-1} = \cdots = y_2 = 0, y_1 = 1]$$

is given by the following recursion: $\beta_2 = \frac{1}{2}(1-p)$, *and for* $j \geq 3$,

$$\beta_j = \frac{1}{2}\left(\frac{1-(1+p)\beta_{j-1}}{1-\beta_{j-1}}\right).$$

Lower Bound: Zero error capacity

- Zero-error capacity:

$$C_{\text{zero-error}} = \lim_{n\to\infty} R_n.$$

where R_n is the maximum achievable zero-error rate.

It is not difficult to show that for $n \geq 1$,

$$R_n = \begin{cases} \frac{1}{2} & \text{if } n \text{ is even} \\ \frac{1}{2} + \frac{1}{2n} & \text{if } n \text{ is odd.} \end{cases}$$

Hence, $C_{\text{zero-error}} = 1/2$.

The Repeat-Each-Bit Code

Consider $\mathcal{R}_n \subset \{0,1\}^n$ consisting of all words of the form

$$(c_1, c_1, c_2, c_2, c_3, c_3, \ldots)$$

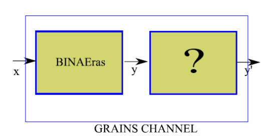
Note that the second bit in each (c_i, c_i) pair cannot be changed by the action of any length-1 or length-2 grain.

So, the decoder can correctly recover any codeword written on the medium.

- $|\mathcal{R}_n| = 2^{\lceil n/2 \rceil}$
- $\mathcal{R}_n$ is t-grain-correcting for *any* t

Upper Bounding by a Better Channel

The grain channel is a degraded version of the "no-adjacent-erasures" channel:



Input-Output Transition of '?': $y'_j = \begin{cases} y_j & \text{if } y_j \neq \varepsilon \\ y_{j-1} & \text{if } y_j = \varepsilon \end{cases}$

- define $\gamma_j = \begin{cases} 1 & \text{if an erasure occurs at position } j; \\ 0 & \text{otherwise.} \end{cases}$
- $P(\gamma_j = 1 \mid \gamma_{j-1} = 0) = p$

  $P(\gamma_j = 1 \mid \gamma_{j-1} = 1) = 0$

Channel Input: $\mathbf{x} = (x_j)_{j\in\mathbb{N}}, \quad x_j \in \{0,1\}$

Channel Output: $\mathbf{y} = (y_j)_{j\in\mathbb{N}}, \quad y_j \in \{0,1,\varepsilon\}$

Channel State: $\mathbf{s} = (s_j)_{j\in\mathbb{N}}, \quad s_j = \gamma_j$

Input-Output Transition: $y_j = \begin{cases} x_j & \text{if } s_j = 0 \\ \varepsilon & \text{if } s_j = 1 \end{cases}$

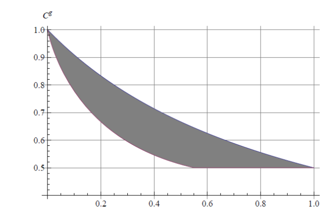
We have

$$C_{\text{grain}} \leq C_{\text{no-adj-erase}}$$

The capacity of the no-adjacent erasure channel is computed to be

$$C_{\text{no-adj-erase}} = 1 - \frac{p}{1+p} = \frac{1}{1+p}$$

Plot of Capacity Bounds



t-Grain-Correcting Codes

The effect of a given grain pattern is denoted by an operator ϕ that acts upon binary sequences $\mathbf{x} \in \{0,1\}^n$:

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \xrightarrow{\phi} \mathbf{y} = \phi(\mathbf{x}) = (y_1, y_2, \ldots, y_n)$$

For $n, t \in \mathbb{Z}^+$, let $\Phi_{n,t}$ denote the set of all operators ϕ corresponding to grain patterns comprising n bit cells, with at most t grains of length 2. [Note: $t \leq n/2$]

For $\mathbf{x} \in \{0,1\}^n$, define $\Phi_{n,t}(\mathbf{x}) = \{\phi(\mathbf{x}) : \mathbf{x} \in \Phi_{n,t}\}$.

$\mathcal{C} \subset \{0,1\}^n$ is called a t-grain-correcting code if for all $\mathbf{x}, \mathbf{x}' \in \mathcal{C}$ with $\mathbf{x} \neq \mathbf{x}'$, we have

$$\Phi_{n,t}(\mathbf{x}) \cap \Phi_{n,t}(\mathbf{x}') = \emptyset.$$

Constructions and Bounds

- $M(n,t)$:= max. size of any length-n t-grain-correcting code.
- If n is a power of 2, then it is easy to construct a t-grain-correcting code with $\geq 2^n/n^t$ codewords.
- Simply observe that a t-error-correcting code is a t-grain-correcting code and there can never be an error in the first position.

$$M(n,t) \leq \frac{2^{n+t}\, t!}{n^t}(1+o(1)) \text{ for any constant } t.$$

Idea of Proof: In a t-grain-correcting code $\mathcal{C}$, the sets $\Phi_{n,t}(\mathbf{x})$, $\mathbf{x} \in \mathcal{C}$, must be disjoint. Hence,

$$2^n \geq \Big|\bigcup_{\mathbf{x}\in\mathcal{C}} \Phi_{n,t}(\mathbf{x})\Big| = \sum_{\mathbf{x}\in\mathcal{C}} |\Phi_{n,t}(\mathbf{x})|$$

The bound follows from an estimate of $|\Phi_{n,t}(\mathbf{x})|$ for a "typical" $\mathbf{x} \in \mathcal{C}$.

Cardinality/Rate Lower Bounds

When $t = \tau n$, for $0 \leq \tau \leq 1/2$, define:

$$\underline{R}(\tau) := \liminf_{n\to\infty} \frac{\log_2 M(n, n\tau)}{n}, \quad \overline{R}(\tau) := \limsup_{n\to\infty} \frac{\log_2 M(n, n\tau)}{n}.$$

- $M(n,t) \geq 2^n / \sum_{j=0}^{2t} \binom{n}{j}$ [Gilbert-Varshamov bound]

  Hence, $\underline{R}(\tau) \geq 1 - h(2\tau)$.
- $M(n,t) \geq 2^{\lceil n/2 \rceil}$ for any n, t [Repeat-each-bit code]

  Hence, $\underline{R}(\tau) \geq 0.5$.

Clique-Partition Upper Bound

- Define the confusability graph $G(n,t)$ as follows:
  - Vertex set $= \{0,1\}^n$
  - $\mathbf{x}, \mathbf{x}'$ are joined by an edge iff $\Phi_{n,t}(\mathbf{x}) \cap \Phi_{n,t}(\mathbf{x}') \neq \emptyset$

Let $\overline{\chi}_{n,t}$:= smallest size of a clique partition of $G(n,t)$

For m, n, s, t such that $t/n \leq s/m$, we have

$$M(n,t) \leq (\bar{\chi}_{m,s})^{\lfloor t/s \rfloor} 2^{n - m\lfloor t/s \rfloor}.$$

Hence, for $\tau \leq s/m$,

$$\overline{R}(\tau) \leq 1 - \tau\left(\frac{m}{s} - \frac{1}{s}\log_2 \bar{\chi}_{m,s}\right).$$

Cardinality/Rate Upper Bounds

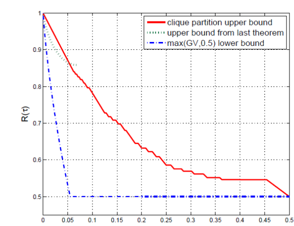
Let $\alpha_0 \equiv \alpha_0(\tau)$ be the smallest positive solution of the following equation:

$$\mathrm{h}\left(\frac{1-x}{2}\right) + \frac{1-x}{4}\mathrm{h}\left(\frac{4\tau}{1-x}\right) = 1,$$

where $\mathrm{h}(\cdot)$ denotes the binary entropy function. For all τ such that $\alpha_0 \leq 1 - 8\tau$, the following bound holds true:

$$\overline{R}(\tau) \leq \mathrm{h}\left(\frac{1-\alpha_0}{2}\right).$$

The bound follows from an estimate of $|\Phi_{n,t}(\mathbf{x})|$ for a "typical" $\mathbf{x} \in \mathcal{C}$ and count of the the "atypical" $\mathbf{x}$.



Constructions

- For small lengths can be found by computer search.

  $(0000, 0011, 0110, 1000, 1011)$ is a 1-grain-correcting code of length 4. The best 1-error-correcting code of length 4 can have at most 2 codewords.
- Concatenated construction for the large lengths (inner codes are small grain-correcting codes that are better than error-correcting codes)
- Work in progress.

- Find better codes, cardinality bounds, and capacity bounds!
- How to handle 2-D granular media?