A Physically Based Approach to Information Theory for Thin Film Magnetic Recording

Joseph A. O'Sullivan
Donald G. Porter
Ronald S. Indeck
Marcel W. Muller

Department of Electrical Engineering
Washington University
St. Louis, MO 63130

ABSTRACT

Conventional approaches to the study of information theory for magnetic recording use relatively simple models for the magnetic medium which do not account for many interference sources. Typically these models account for a few types of noise sources, ignoring many others. Most models also assume Gaussian statistics for the noise sources. In this paper, we propose a more complicated model for thin film magnetic media and show preliminary analysis to demonstrate its ability to model physical properties of media. Some initial ramifications of using this model (or other advanced models) for studying the information theory of magnetic recording are explored. A major result of this paper is the development of a new methodology for computing the information capacity of magnetic media in bits per unit area. This computation is independent of the method of storing information and thus is a fundamental limit for any magnetic recording system. This represents progress in unifying physical descriptions of magnetic media with the study of the information capacity of magnetic information storage systems.

1. Introduction

This paper introduces a computationally tractable, physically based model for magnetic media intended for use in studying information theoretic properties of magnetic storage systems. The proposed model is most appropriate for thin film and particulate media. Thin film media are briefly discussed in Section 2. The model is discussed in Section 3 and is essentially a spin glass model [1,2,3]. The connection of spin glass models to neural networks [2,3] has proven worthwhile in the study of the capacity of neural networks. It is our intention to exploit these same methods to study the capacity of magnetic storage systems using our model. It is important to emphasize that the starting point of our study of capacity is to look at the capacity per unit area on the surface of the magnetic medium. This then implies fundamental limits on the capacity of magnetic storage systems. Thus the study of information theory from the input-output point of view is connected to a study of information theory at the surface of the actual storage medium.

Several models for magnetic systems are discussed in [4]. These models include the runlength-limited channel model, models which account directly for bit-shift [5], and a runlength-limited channel model with noisy transition measurements [6]. Other models account for crosstalk and intersymbol interference [7,8,9]. Several innovative ways to adaptively equalize these sources of interference have been proposed [10,11]. These are input-output models which do not describe the actual physics. They model the manner in which the channel affects the information bits and are limited. We anticipate that the model introduced here will allow for more precise computations of the capacity of magnetic systems and may lead to improved read/write techniques.

A more sophisticated model which has appeared in the literature is the French-Wolf model [12]. McLaughlin and Neuhoff have improved on the French-Wolf model [13,14] and have used their improved model to compute bounds on the capacity of magnetic information channels and to study source coding for analog data for these channels. This model accounts for media noise, electronics noise, filtering in the read process, and the variation of those quantities as the track width and the linear density of bits vary. While their model has proven very useful and some of their results are very impressive, their model still has limitations. One limitation is the assumption of Gaussian noise. A second limitation is the inability of the model to account for bit shift due to magnetic variations at the surface of the medium. A third limitation is its inability to include crosstalk.

Our model will include these other affects. To include other affects, our model must be more sophisticated. The main limitation of our model is its complexity. One area of ongoing research is the simplification of this model in special situations.

The model proposed may be thought of as a simplification of models used to study the physics of magnetic media. Models such as those discussed by Zhu and Bertram [15,16,17,18] and Mansuripur and Giles [19,20] simulate the solution of the Landau-Lifshitz-Gilbert differential equation. To our knowledge, no attempt has been made to use these very complicated models for studying signal processing properties of magnetic information systems. These models do take into account a wide variety of phenomena, however. The model presented below also takes into account these phenomena, but does not attempt to solve the Landau-Lifshitz-Gilbert equation. The model has discrete jumps in magnetic moments of individual grains. Since it is at the level of grains, it is very detailed. Some reduction in complexity is achieved by allowing only a finite number of orientations of magnetic moments of these grains.

Another major goal of this work is to bring together research in the physics of magnetic domains and research in signal processing for magnetic information systems. By developing a model for magnetic recording of use to both, interaction may be enhanced.

2. Thin Film Magnetic Media

This section presents a brief high-level description of thin film magnetic media, from the point of view our model represents. This is not meant to be a detailed physical description. Thin film media are deposited on substrates in such a manner that independent crystallites initiate growth at various sites across the surface. As the crystals grow, they merge with other crystals at irregular interfaces. Depending on the material and the orientations of the individual crystals, this interface may or may not be abrupt. All crystals are formed independently and have random orientations. At the end of the deposition process, the crystals have grown into what are usually called grains. Media are often characterized by their granularity: the typical sizes of, orientations of, and interactions between the grains. At one extreme are particulate media which are composed of physically independent grains which interact only through dipolar effects. Many media have a tendency for adjacent grains to have magnetic
moments which are aligned. The strength of this exchange interaction is a parameter in our model.

In order to get a sense for the order of magnitude of the grains under discussion, our model assumes grains are about as wide as they are thick. The media under consideration may have thicknesses of a fraction of a micron.

3. Magnetic Model

Our model assumes grains are located on sites on a rectangular two-dimensional lattice. The grains are assumed not to overlap. The magnetic moments of individual grains are assumed to lie in the plane of the lattice. Several types of interactions between grains are included in the model, and several sources of randomness are included.

Each grain has an axis of anisotropy (the easy axis), \( u_j \), in the plane of the lattice. These axes are determined by independent, identically distributed random angles, \( \theta_j \) (due to the random nature of crystal growth). To make the model computationally more tractable, we assume that the anisotropy is so strong that the magnetic moment of an individual grain lies along its easy axis. Thus, for each grain a sign, \( s_j \in \{-1, 1\} \), determines the magnetization which is given by

\[ m_j = s_j u_j , \]

where \( m \) is the magnitude of the magnetization of each grain.

The grains which are far apart interact through their magnetic fields as dipoles. This dipolar interaction depends on the distance and relative orientation of the grains and is spatially invariant. Note that the magnetic field at a given grain due to a second grain is linear in the magnetic moment of that second grain. The dipolar field at site \( j \) is equal to

\[ H_{d,j} = \sum_{i \in j} A_{j,i}(m_i) , \]

where \( A_{j,i} \) is the spatially invariant linear operator representing the dipolar interaction. Given in matrix form for the lattice being in the \( x-z \) plane and \( d_x \) and \( d_z \) being the separation of lattice sites \( j \) and \( i \) in the \( x \) and \( z \) coordinates, respectively, we have

\[ A_{j,i} = \frac{K_d}{(d_x^2 + d_z^2)^{3/2}} \begin{bmatrix} 2d_z^2 - d_x^2 & 3d_xd_z \\ 3d_xd_z & 2d_x^2 - d_z^2 \end{bmatrix} , \]

where \( K_d \) is a constant depending on the grain size and the permeability.

Adjacent grains interact additionally through an exchange coupling. This models the tendency the magnetizations of adjacent grains to be parallel. The exchange coupling is modeled by an equivalent magnetic field and is linear in the magnetizations,

\[ H_{ex,j} = \sum_{i \in N(j)} \frac{K_{ex,j,i}}{N(j)} m_i , \]

where \( K_{ex,j,i} \) is a scalar constant depending on the stiffness coefficient, the saturation magnetization, and the size of the grains. \( N(j) \) is the set of neighbors that contribute to the exchange field.

In order to determine in which direction along the easy axis a given grain should point, two steps take place. First, the magnetic field at that grain is computed. This is linear in the magnetic moments of all other grains plus the applied field. Second, this field and the previous magnetic moment of this grain are compared using the Stoner-Wohlfarth model [21]. The result of this comparison determines the direction of the magnetization of that grain.

The Stoner-Wohlfarth model is discussed in more detail in the next section.

Three comments concerning this model are appropriate:

First, other models exist. Some of the most intricate solve the Landau-Lifshitz-Gilbert equation [15,16,17,18,19,20]. Our approach approximates the solution of these nonlinear differential equations. Note that our approach tends to jump to energy minima rather than approach those minima slowly. Our approach is also much less computationally demanding.

Second, the model introduces randomness through the update strategy. A typical use of the model would be the following. A field is applied instantaneously. This means that the previous set of grain magnetizations are no longer valid. Lattice sites are chosen at random and updated by the strategy outlined above. After a large number of lattice site updates, an equilibrium (energy minimum) is declared. Since the order in which the sites are visited is random, the actual energy minimum attained is random.

Third, the model is like a neural network in that a linear combination of the outputs of other sites (their signs) plus a bias (the applied field) is passed through a nonlinear threshold function (from the Stoner-Wohlfarth model) to determine the next value of the sign of the site under consideration. Thus, the literature on the capacity of neural networks (see [22,23,24,25,26]) should prove useful in computing the capacity of our model. The main results from this literature depend on the results in the spin glass literature. While our model has many similarities to typical spin glass models, there is one large difference: the Stoner-Wohlfarth model. Usual spin glass models [1,2] do not include such a switching function to account for coercivity.

4. Stoner-Wohlfarth Model

Stoner and Wohlfarth modeled the energy function for an individual grain as being composed of two terms. The first is an anisotropy energy due to the tendency of the magnetic moment of a grain to align itself with its easy axis. The second term is the energy due to the applied field and represents the tendency of the magnetic moment to align with the applied field. For small applied fields, there are two energy minima. For large applied fields, there is only one energy minimum. A plot of the critical fields at which the transition from two minima to one minimum takes place is known as the Stoner-Wohlfarth astroid. In our model, the Stoner-Wohlfarth astroid is used to determine whether the field at a grain is sufficiently large to cause the magnetic moment to flip. That is, since we assume that the magnetic moment is aligned with the easy axis, the critical field determines when the magnetic moment must flip. For fields lower than critical (inside the astroid), no flip occurs.

5. Information Capacity

There are several ways in which the study of capacity of magnetic media may be carried out based on the model proposed here. The approach taken here is to count the relative logarithmic number of energy minima. That is, the capacity is determined by the logarithm of the number of minima divided by the logarithm of the total number of possible configurations. One limitation of using this method for computing the capacity is that the minima tend not to be isolated. This means that two or more minima may be indistinguishable from a measurement standpoint. The modeling of the measurement process should include this ambiguity.

Results on computing the capacity of media have been obtained and will be published elsewhere. This study is being carried out analytically and computationally. One goal in this study is to motivate results like the following. This result is stated in terms of energy "basins" which may include several energy minima.


On Generating Random Bits from an Arbitrary Random Source

Sridhar Venbu Sergio Verdú
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544

Abstract

We address the following question: Suppose we are given a random source and want to use it as a random number generator, how many approximately fair bits can we generate from it? We answer this question with two different measures of approximation between probability distributions: the variational distance and the normalized divergence. The results of this paper provide an operational characterization of the inf-entropy rate of a source, defined in Han and Verdú, [1]. We also relate the inf and sup entropy rates to the quantity called resolvability of channels, also defined in [1].

1 Introduction

The recent results of Han and Verdú, [1], have opened some new questions in information theory. We focus on the source coding results of [1] in this article. Specifically, we address the following question: suppose we are given a random source and we want to use it as a random number generator, how many fair bits can we generate from it? Of course, we may not be able to get too many "pure" random bits - for example, if the probability masses of the source are irrational. Instead we will focus on synthesizing an "approximately" uniform distribution from the source and address the issue of the largest asymptotic exponential rate at which this can be done. We will call this rate the Intrinsic Randomness rate (IR rate) of the source. In this paper we use two measures of approximation of probability distributions - variational distance and divergence. In both cases the answer to the question turns out to be the same: it will be the inf-entropy rate as defined in Han and Verdú, [1]. The results of this article nicely complement the source coding results of [1] where the minimum fixed length encoding rate is shown to be the sup-entropy rate for an arbitrary finite alphabet source. Thus we see (Figure 1) that intrinsic randomness plays the counterpart to the minimum source coding rate, analogously to resolvability and channel capacity.

We will also relate the inf-entropy and the sup-entropy rates to the resolvability of channels defined in [1].

*This work was supported in part by the National Science Foundation under PYI Grant ECS-8857889, and the Office of Naval Research under Grant N00014-90-J-1734. The work of the first author was also supported by an IBM Graduate Fellowship.