N95-14624

TDA Progress Report 42-118

August 15, 1994

Spin Glasses and Error-Correcting Codes

M. L. Belongie
Communications Systems Research Section

In this article, we study a model for error-correcting codes that comes from spin glass theory and leads to both new codes and a new decoding technique. Using the theory of spin glasses, it has been proven that a simple construction yields a family of binary codes whose performance asymptotically approaches the Shannon bound for the Gaussian channel. The limit is approached as the number of information bits per codeword approaches infinity while the rate of the code approaches zero. Thus, the codes rapidly become impractical. We present simulation results that show the performance of a few manageable examples of these codes.

In the correspondence that exists between spin glasses and error-correcting codes, the concept of a thermal average leads to a method of decoding that differs from the standard method of finding the most likely information sequence for a given received codeword. Whereas the standard method corresponds to calculating the thermal average at temperature zero, calculating the thermal average at a certain optimum temperature results instead in the sequence of most likely information bits. Since linear block codes and convolutional codes can be viewed as examples of spin glasses, this new decoding method can be used to decode these codes in a way that minimizes the bit error rate instead of the codeword error rate. We present simulation results that show a small improvement in bit error rate by using the thermal average technique.

I. Introduction

In a 1989 article in *Nature* [2], Nicolas Sourlas claimed that by using an Ising spin glass model he could construct a family of error-correcting codes whose performance asymptotically approached the Shannon coding bound. In 1993, Pal Rujan proposed an idea in *Physical Review Letters* [6] for decoding spin glass codes with a lower resulting bit error rate than could be obtained by finding the most likely codeword, and claimed that the method could also be used for convolutional codes. In this article, we study both of these claims.

This article is organized as follows: In Section II, we introduce Ising spin glasses, which are just collections of particles with spin ± 1. We also briefly discuss the associated concepts of energy, thermal equilibrium, ground states, magnetization, phase transitions, and gauge invariance. In Section III, we explain the connection between Ising spin glasses and binary error-correcting codes on an additive white Gaussian noise channel. As an example, we will show that spin glasses with a certain type of interactions are equivalent to convolutional codes. We show in Section IV how the physical properties of spin glasses can be used to prove that a family of error-correcting codes based on spin glasses has a bit error rate

approaching zero while the rate approaches the Shannon capacity. Although the construction is simple, it will be clear that the codes rapidly grow too large to be practical. In Section V, we present simulation results that show the performance of a few manageable examples of these codes. Here, decoding is done by finding the ground state of a spin glass, which corresponds to finding the most likely codeword.

In Section VI, we consider a different decoding technique, based on the concept of a thermal average. We show that this method of decoding at an optimum, nonzero temperature minimizes bit error rate, as opposed to codeword error rate. We present simulation results that verify a small decrease in bit error rate for a few examples of spin glass codes, and also for the (8,4) Hamming code and the (24,12) Golay code, both of which can be viewed as spin glass codes. In Section VII, we demonstrate a method for decoding convolutional codes at nonzero temperatures. The method, based on the transfer matrix method of statistical mechanics, is similar to a Viterbi decoder, which it reduces to in the case of zero temperature. We present simulation results that show a small decrease in bit error rate for a convolutional code decoded at the optimum temperature. Finally, in Section VIII, we present our conclusions.

II. Properties of Spin Glasses

An Ising spin glass is a set of N particles with spin ± 1 [1]. The energy of a spin glass depends on the values of the spins and the strengths of the interactions among the particles. For instance, if there is a positive interaction between two particles, then the energy contribution from that interaction will be lower if the spins are the same and higher if they are opposite. In general, the strength of the interaction between any p particles $i_1, \cdots, i_p$ is given by the coupling coefficient $J_{i_1,\cdots,i_p}$. Let s_i be the spin of the ith particle and $S = \{s_1, \cdots, s_N\}$ a configuration of the N spins. The energy of the whole system when it has configuration S is given by the Hamiltonian

$$H(S) = -\sum_{p} \sum_{\{i_1,\cdots,i_p\} \subset \{1,\cdots,N\}} C_{i_1,\cdots,i_p} J_{i_1,\cdots,i_p} s_{i_1} \cdots s_{i_p} \tag{1}$$

where the connectivity matrix $C_{i_1,\cdots,i_p}$ is 1 if the p particles $i_1, \cdots, i_p$ interact and 0 otherwise. (Each subset is understood to appear only once in the sum.) We will assume that the coupling coefficients are given by independent identically distributed (i.i.d.) random Gaussian variables with mean J_0 and variance σ_J^2.

The physical situation that we are considering here is that the interaction strengths are fixed while the spins are free to change. Let T be the temperature of the system. Then the probability of the system being in a particular configuration S at equilibrium is given by the Gibbs distribution [1]:

$$p(S) = \frac{e^{-(1/kT)H(S)}}{Z} \tag{2}$$

where k is Boltzmann's constant and the partition function Z (a normalizing constant) is given by

$$Z = \sum_{S} e^{-(1/kT)H(S)} \tag{3}$$

As $T \to 0$, we see that $p(S) \to 0$, except for S such that $H(S)$ is a minimum. Such a minimizing configuration S is known as a ground state. Unfortunately, in many cases the ground state is degenerate, meaning that there may be multiple configurations of spins that have the same minimum energy. For instance, consider the case $p = 2$: For any configuration, the energy of the configuration obtained by reversing all the spins will be the same.

The magnetization $m(S)$ of a configuration S is simply the average spin: $m(S) = (1/N)\sum s_i$. In this article, we will need to consider limits as the number of particles $N \to \infty$. In this situation, there can occur a phase transition, which is a sort of discontinuity in some global characteristic of the system as a function of some continuous parameter(s). Specifically, for the spin glass model we have described, there is a phase transition at zero temperature at a particular critical value of J_0/σ_J: For J_0/σ_J below this cutoff, the magnetization m (of the ground state, since $T = 0$) is zero, whereas above it, $m > 0$. We will dismiss the problem of degenerate ground states by merely stating that there are ways to make the ground state unique. The existence of the phase transition also requires that the connectivity matrix not be too sparse. This condition will be satisfied in the cases for which we will invoke a phase transition.

We now define one more property of our spin glass model, called gauge invariance. Let $\{\epsilon_1, \cdots, \epsilon_N\}$ be an arbitrary configuration of Ising spins, that is, $\epsilon_i = \pm 1$. A system is gauge invariant if the configuration space is invariant under the transformation $s_i \to s_i\epsilon_i$, and if the Hamiltonian is invariant under this transformation and the simultaneous transformation $J_{i_1,\cdots,i_p} \to J_{i_1,\cdots,i_p}\epsilon_{i_1}\cdots\epsilon_{i_p}$. Clearly, our model is gauge invariant, since the result of multiplying any sequence of Ising spins by another arbitrary sequence of Ising spins is yet another sequence of Ising spins, and Eq. (1) is unchanged if both transformations are applied simultaneously.

III. Using Spin Glass Models as Error-Correcting Codes

In a 1989 article in *Nature* [2], Nicolas Sourlas suggested using an Ising spin glass model to construct error-correcting codes. In this section, we describe his proposal. For simplicity, he only considered the slightly simpler special case where interactions are restricted to a single value of p. Thus, we have the Hamiltonian

$$H(S) = -\sum_{\{i_1,\cdots,i_p\}\subset\{1,\cdots,N\}} C_{i_1,\cdots,i_p} J_{i_1,\cdots,i_p} s_{i_1}\cdots s_{i_p} \tag{4}$$

Let $\{a_1, \cdots, a_N\}, a_i = \pm 1$, be an N-bit information sequence. Let $J_{i_1,\cdots,i_p} = J_0 \prod a_{i_1}\cdots a_{i_p}$ whenever $C_{i_1,\cdots,i_p} = 1$. Then the set of spins corresponding to the data will be the ground state of the Hamiltonian with these coupling coefficients. Thus, our spin glass model yields a code, with codewords given by the computed set of coupling coefficients and decoding done by finding the ground state of the Hamiltonian specified by the coupling coefficients. Before finding the rate, we make one more simplification: Assume that the coordination number $z_i = \sum_{j_2,\cdots,j_p} C_{i,j_2,\cdots,j_p} = z$ is independent of i. (Notice that $C_{j_1,j_2,\cdots,j_p}$ is invariant under permutations of its indices, so that z_i is the number of interacting subsets that include s_i.) Then the code rate is given by $R = p/z$.

Now we consider the issue of noise. Let us assume that the transmitted codeword symbols have magnitude $\pm v$ and duration τ, and are subject to additive white Gaussian noise (AWGN) of spectral density N_0. Then the noise is included in the model via the already discussed variance σ_J^2 of the coupling coefficients, with corresponding channel $SNR = E_s/N_0 = v^2\tau/N_0 = J_0^2/2\sigma_J^2$. For $J_0 << \sigma_J$, or equivalently, $v^2\tau << N_0$, the channel thus has capacity $C = (1/\ln 2)J_0^2/2\sigma_J^2$ bits per coupling coefficient [3].

We are ultimately concerned with the probability of decoded bit error P_b, so we now consider the corresponding quantity in the spin glass model. This is where gauge invariance comes in: It allows us to assume, without loss of generality, that the spins are all $+1$. Then, P_b is just the probability that a bit is decoded as -1. Since the ground state magnetization m is given by $m = (+1)Pr\{s_i = +1\} + (-1)Pr\{s_i = -1\} = (+1)(1 - P_b) + (-1)P_b$, we have $P_b = (1 - m)/2$. In the next section, we will describe a situation where $m \to 1$, and hence $P_b \to 0$.

As a familiar example, consider a binary $(n, 1)$ convolutional code [3]. For simplicity, assume that each of its n generating polynomials $\{g_1(x), \cdots, g_n(x)\}$ has exactly p terms. It can be described by our model

by letting $N = \infty$, and $C_{i_1+k,\ldots,i_p+k} = 1$ if and only if $g_j(x) = x^{i_1} + \cdots + x^{i_p}$ for some j. For instance, if $n = p = 2$, with $g_1(x) = 1 + x$ and $g_2(x) = 1 + x^2$, then the corresponding log likelihood formula,

$$\log P \propto \sum_i J_{i,i-1} S_i S_{i-1} + J_{i,i-2} s_i s_{i-2} + \text{constant} \tag{5}$$

that is used in Viterbi decoding is equivalent to the Hamiltonian we have defined. The code can be visualized as a one-dimensional infinite spin glass with short-range, translation invariant interactions.

IV. Approaching the Shannon Bound With a Spin Glass Model

We now consider a special case, known as Derrida's random energy model [4], that is soluble. This means that the behavior of certain measurable quantities, such as the magnetization as a function of J_0/σ_J and temperature T, can be calculated. Again, we restrict interactions to a single value of p, so that our Hamiltonian is given by Eq. (4). We assume that the connectivity is extensive, i.e., $z \sim \binom{N-1}{p-1}$. Because we will be letting $N \to \infty$, we will use the scaled variables j_0 and σ_j, with $J_0 = (p!/N^{p-1})j_0$ and $\sigma_J^2 = (p!/N^{p-1})\sigma_j^2$, in order for the relevant quantities to remain finite. We then normalize by setting $\sigma_j = 1$, and so the channel $SNR = (1/2)j_0^2\, p!/N^{p-1}$. As $N \to \infty$, the asymptotic rate $R = p!/N^{p-1}$ and the capacity $C = (j_0^2/2\ln 2)p!/N^{p-1}$ bits per coupling coefficient.

In Derrida's random energy model, the number of particles $p \to \infty$ and $p/N \to 0$. He showed that in this case there is a phase transition for $T < 1/(2\sqrt{\ln 2})$ at $j_0 = \sqrt{2\ln 2}$ from a spin glass phase with $m = 0$ to a ferromagnetic phase with $m = 1$. Thus, for $j_0 > \sqrt{2\ln 2}$, the probability of error P_e can be made arbitrarily small. Therefore, we can code at a rate arbitrarily close to capacity with $P_e \to 0$. Although $R \to 0$, we note that R', the rate in bits per sec, is given by $R' = R/\tau$. Since $v^2\tau/N_0 = J_0^2/2\sigma_J^2 = (1/2)j_0^2 R = \ln 2\ R$, it follows that $R' = (1/\ln 2)v^2/N_0$. This means that the rate in bits per sec remains a constant as a function of transmitter power $P = v^2$ and noise power N_0 as $P_e \to 0$. The capacity C' in terms of bits per sec is $C' = (1/\ln 2)v^2/N_0$ [3]. The bound is approached for codeword lengths approximately equal to N^p, where both N and p approach infinity, so clearly the codes rapidly become impractical. However, we do have an explicit construction for a family of codes whose performance approaches the Shannon limit, and it is possible to simulate them, as described in the next section. Spin glass theory does give us some hint as to what we can expect: It can be shown that for large p, $m \approx 1 - (2^{-p}/\sqrt{p})$ for codes satisfying the capacity constraint.

V. Decoding by Finding the Ground State

In this section, we discuss simulations of decoding by finding the ground state for a few examples of the family of codes described in the previous section. The size of the codes grows so quickly with N and p that only fairly small values could be used. Two different methods for finding the ground state were used: exhaustive search and simulated annealing. By using efficient recursive algorithms for exhaustive search and letting simulations run for over a week in some cases, it was possible to test codes with parameters as high as $N = 20$ and $p = 5$. Results are shown in Fig. 1, where the curves are labeled as (N, p). For the four codes with $N = 8$ and 12, each point represents 2500 codeword error events. For the two $N = 16$ codes and the $(N, p) = (20, 3)$ code, each point represents 100 codeword error events, and for the $(20, 5)$ code, only 25 codeword errors were obtained. The reason for using odd values of p was to avoid the degenerate ground states that result when p is even, and no other problems from degenerate ground states appeared.

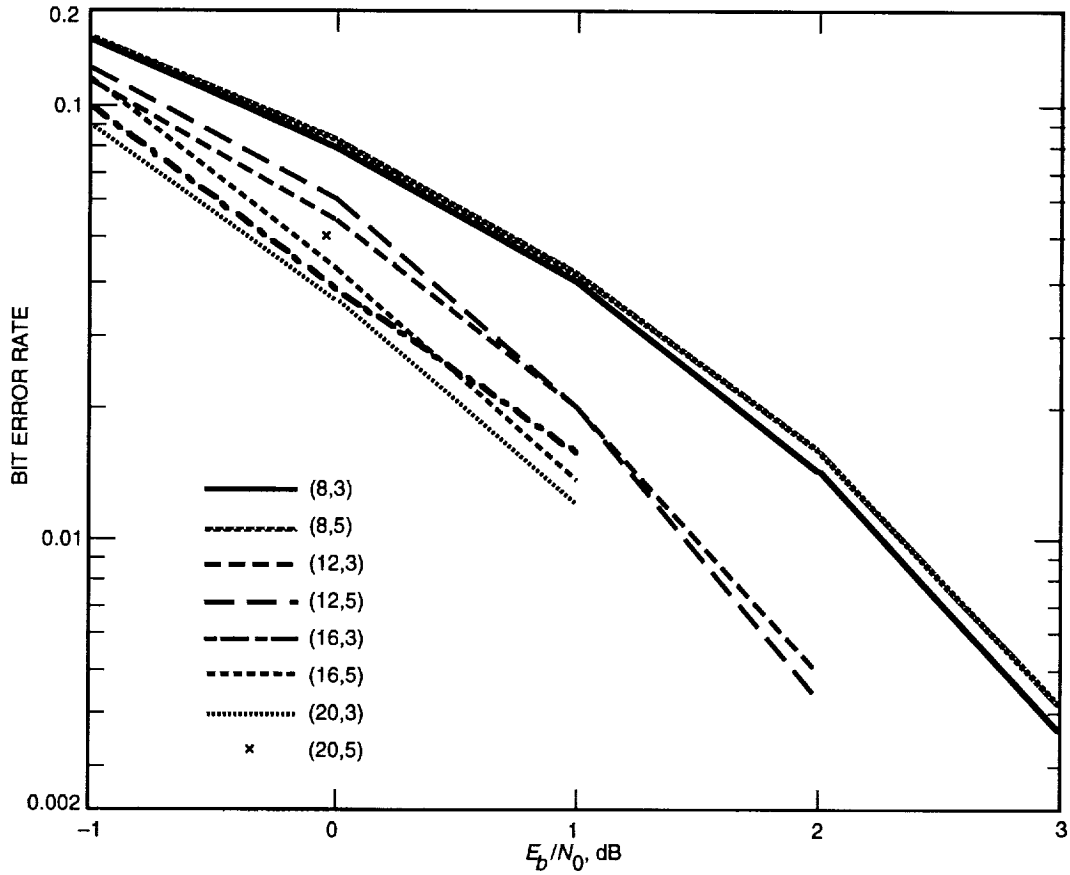

Fig. 1. Bit error rates of spin glass codes with varying parameters (N, p).

In an effort to speed up decoding, simulated annealing [5] was tried. In a physical spin glass, this corresponds to heating it up to a relatively high temperature and then slowly reducing the temperature asymptotically to 0. If the dwell time, or time spent at each temperature, is large enough for the temperature-decrementing factor used, then the system should reach equilibrium at each temperature, and thus end up in the ground state for T sufficiently close to zero. Results for $N = 12$ and $p = 3$ are shown in Fig. 2, for two different dwell times. Each point on the simulated annealing curves represents approximately 1000 codeword errors, and the exhaustive search curve is taken from Fig. 1. Unfortunately, simulated annealing took even longer than the exhaustive search. This was true even for larger parameters, such as $N = 20$ and $p = 5$, because much longer dwell times were necessary to get reasonable performance. It is possible that further customization of the simulated annealing algorithm for the particular characteristics of this problem could result in significant improvement, but the potential benefits did not seem to justify the additional effort at this time.

VI. Decoding at Nonzero Temperatures

After reading Sourlas' article, Pal Rujan proposed an idea for decoding spin glass codes with a lower resulting bit error rate than could be obtained by finding the ground state [6]. He showed that the effect of the channel was equivalent to heating up a spin glass to a particular temperature T_N (the Nishimori temperature) [7]. For the model we have described with AWGN, $T_N = \sigma_J^2/kJ_0$. This suggested decoding by computing the thermal average at T_N of the Hamiltonian given by the received codeword. The thermal average is the average over all spin configurations, weighted by the Gibbs distribution, so that the decoded

value of the ith spin, $\hat{s}_i$, is given by $(1-m_i)/2$, where m_i is the averaged magnetization of the ith particle, given by

$$m_i = \frac{\sum_S s_i e^{-\beta_N H(S)}}{\sum_S e^{-\beta_N H(S)}} \tag{6}$$

with $\beta_N = 1/(kT_N) = J_0/\sigma_J^2$. Notice that finding the ground state is equivalent to computing the thermal average at $T = 0$.

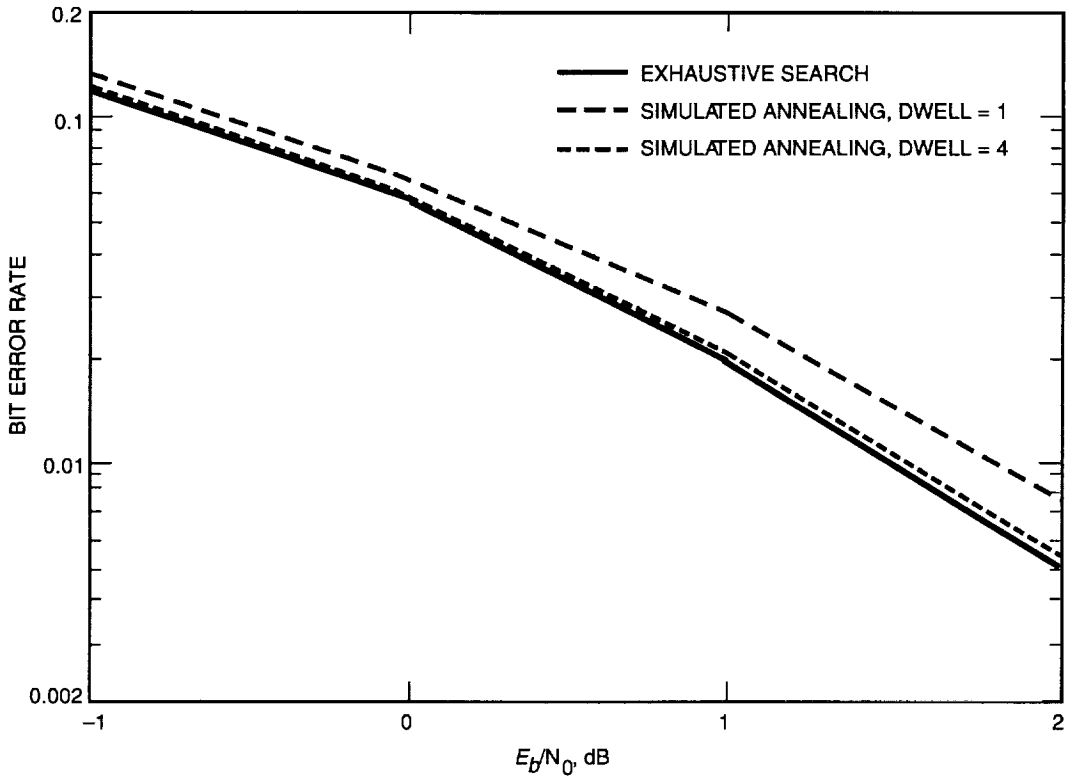

Fig. 2. Performance of simulated annealing versus exhaustive search for decoding a (12,3) spin glass code.

Rujan presented results of a simulation that showed a small decrease in bit error rate at $T = T_N$ compared to $T = 0$ [6]. Hidetoshi Nishimori responded by proving that the bit error rate was indeed lower at $T = T_N$ than at $T = 0$ [9]. In fact, he showed that this is true not only for AWGN, but for any noise with distribution $f(|J|)e^{aJ}$.

Rujan's results seemed suspicious at first glance, because Viterbi decoding is known to be optimum. The difference is that Viterbi decoding finds the most likely sequence of input bits, but not necessarily the sequence of most likely input bits. Although the formula above came from considerations of statistical mechanics, it is easy to derive using just Bayes formula. We start by expressing m_i as

$$m_i = \sum_S s_i Pr\{S|\{J_{i_1,\cdots,i_p}\}\} \tag{7}$$

(Again, we are assuming for simplicity of notation a single value of p.) Using Bayes formula, we get

$$Pr\{S|\{J_{i_1,\cdots,i_p}\}\} = \frac{Pr\{\{J_{i_1,\cdots,i_p}\}|S\}Pr\{S\}}{\sum_S Pr\{\{J_{i_1,\cdots,i_p}\}|S\}Pr\{S\}} \tag{8}$$

$Pr\{S\} = 1/2^N$ for all S, and

$$\begin{aligned} Pr\{\{J_{i_1,\cdots,i_p}\}|S\} &= \prod_{\{i_1,\cdots,i_p\}\subset\{1,\cdots,N\}} e^{-(J_{i_1,\cdots,i_p} - J_0 s_{i_1}\cdots s_{i_p})^2/2\sigma_J^2} \\ &= e^{-(1/2\sigma_J^2)\sum_{\{i_1,\cdots,i_p\}\subset\{1,\cdots,N\}}(J_{i_1,\cdots,i_p} - J_0 s_{i_1}\cdots s_{i_p})^2} \\ &= Ae^{-(1/2\sigma_J^2)\sum_{\{i_1,\cdots,i_p\}\subset\{1,\cdots,N\}} 2J_{i_1,\cdots,i_p} J_0 s_{i_1}\cdots s_{i_p}} \\ &= Ae^{-\beta_N H(S)} \end{aligned} \tag{9}$$

where A is independent of S (using $s_i^2 = 1$). Finally, substituting Eq. (9) into Eq. (8) and then Eq. (8) into Eq. (7) yields Eq. (6), the desired result.

In fact, we can now show that Rujan's method actually minimizes the bit error rate. From Eq. (7),

$$\begin{aligned} m_i &= \sum_{S:s_i=+1} s_i Pr\{S|\{J_{i_1,\cdots,i_p}\}\} + \sum_{S:s_i=-1} s_i Pr\{S|\{J_{i_1,\cdots,i_p}\}\} \\ &= (+1)Pr\{s_i = +1|\{J_{i_1,\cdots,i_p}\}\} + (-1)Pr\{s_i = -1|\{J_{i_1,\cdots,i_p}\}\} \end{aligned} \tag{10}$$

so $m_i > 0$ if and only if $Pr\{s_i = +1|\{J_{i_1,\cdots,i_p}\}\} > Pr\{s_i = -1|\{J_{i_1,\cdots,i_p}\}\}$.

Figure 3 shows the results of decoding simulations for a few of the spin glass codes described in Section IV. Each point represents 10,000 codeword error events. The bit error rate is lower for $T = T_N$ than for $T = 0$, although the difference is small. Figure 4 shows similar results for the (24,12) Golay code and the (8,4) Hamming code. Again, each point represents 10,000 codeword error events. In these simulations, calculations were done by directly computing the sum in Eq. (6), using efficient recursive algorithms. Of course, for anything but the smallest codes, using Eq. (6) directly is impractical. As mentioned previously, however, for convolutional codes one can use Rujan's transfer matrix method, which will be discussed in the next section.

Although we have shown that Eq. (6) can be derived without reference to spin glasses, the theory behind spin glasses might still be useful in dealing with error-correcting codes. One example, of course, is the use of an algorithm from statistical mechanics for decoding. The simulation results do not demonstrate a great improvement in performance, but theoretically, it would be interesting to have bounds on how much improvement is possible from minimizing the bit error rate instead of the codeword error rate. Perhaps spin glass theory could shed some light on this. Another idea that seems reasonable for decoding at the Nishimori temperature is to use simulated annealing, but only decrease the temperature to T_N instead of all the way to zero.

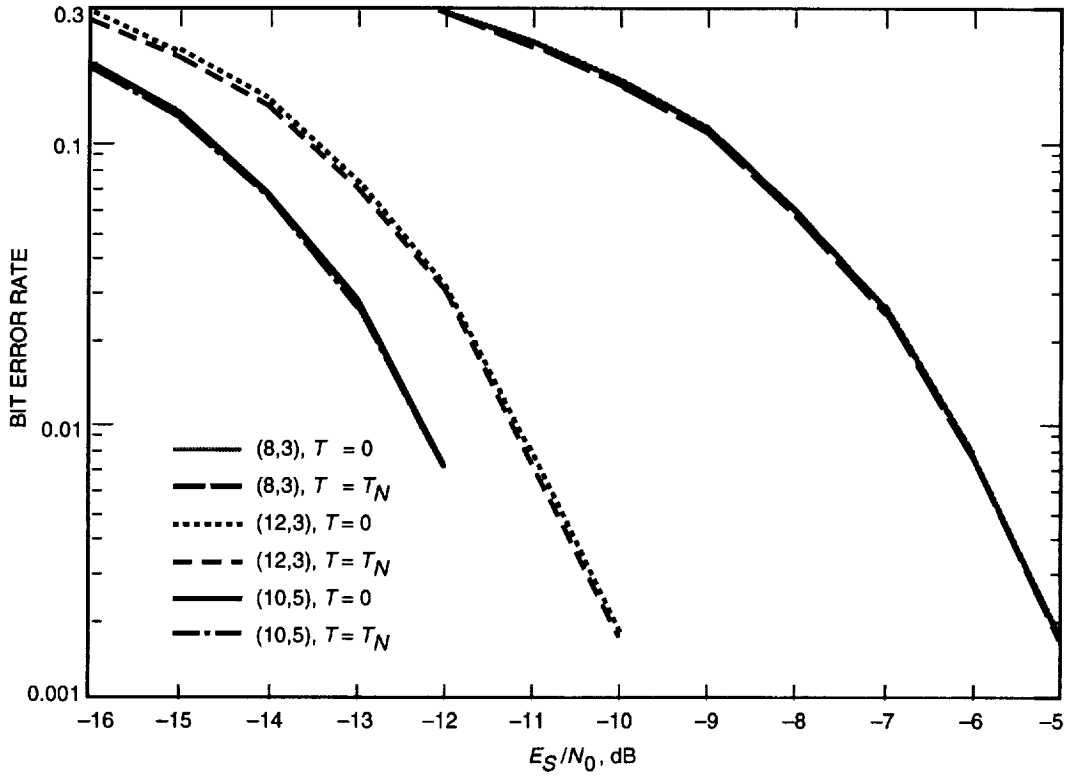

Fig. 3. Comparison of $T = 0$ and $T = T_N$ decoding for three spin glass codes.

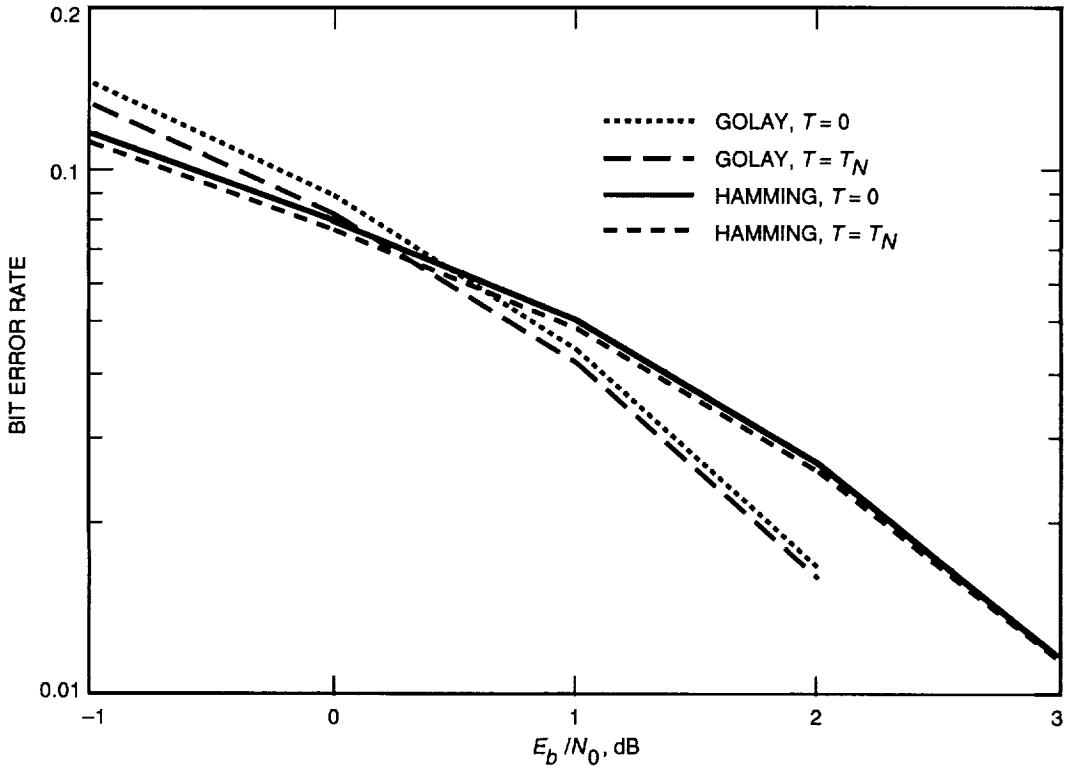

Fig. 4. Comparison of $T = 0$ and $T = T_N$ decoding for the (24,12) Golay code and the (8,4) Hamming code.

It also appears that the theory of Markov random fields [1], which is closely related to spin glass theory, might have implications for finite length codes and infinite codes (in one or more dimensions) with only short-range interactions. Codes with short-range interactions are likely to be more practical, since the rate would not decrease so rapidly as the number of information bits increases. Two-dimensional codes might be useful for encoding images, for example.

Another interesting suggestion was pointed out by Sourlas [10]: It is easy to modify Eq. (6) in order to minimize the probability of error of particular blocks of input data. For instance, if the binary input data consisted of a sequence of k-bit symbols, one could minimize the symbol error rate instead of the bit error rate or instead of the sequence error rate.

VII. Decoding Convolutional Codes at Nonzero Temperatures

Rujan [6] showed that for a one-dimensional spin glass with short-range interactions, such as a convolutional code, a variant of the transfer matrix method of statistical mechanics [8] could be used to compute the thermal average. It finds the m_i's for a given temperature T using a recursive algorithm that reduces to the Viterbi algorithm at $T = 0$, and has similar complexity. We illustrate Rujan's algorithm with a simple rate 1/2 convolutional code with generator matrix $G(x) = [1 + x^2 \quad 1 + x + x^2]$. Let $\{K_1(1), K_2(1), K_1(2), \cdots, K_2(N)\}$ be the $2N$ received coupling coefficients, $K_1(i) = a_{i-1}a_i a_{i+1} + n_1(i)$ and $K_2(i) = a_{i-1}a_{i+1} + n_2(i)$, where $n_1(i)$ and $n_2(i)$ are i.i.d. Gaussian variable of mean zero and variance σ_J^2. (Define $a_0 = a_{N+1} = 0$.) Then the energy can be written as

$$H(S) = \sum_{i=2}^{N-1} H_i(s_{i-1}, s_i, s_{i+1}) \tag{11}$$

where $H_i(s_{i-1}, s_i, s_{i+1}) = K_1(i)s_{i-1}s_i s_{i+1} + K_2(i)s_{i-1}s_{i+1}$. The value for m_i is given by

$$m_i = \frac{\sum_{s_{i-1}, s_i} \psi_{i-1}^{>} s_i \psi_{i+1}^{<}}{\sum_{s_{i-1}, s_i} \psi_{i-1}^{>} \psi_{i+1}^{<}} \tag{12}$$

where the ψ_i's are defined recursively: Let $\psi_1^{>}(s_1, s_2) = 1$ and compute

$$\psi_i^{>}(s_i, s_{i+1}) = \sum_{s_{i-1}} \psi_{i-1}^{>}(s_{i-1}, s_i) e^{-\beta H_i(s_{i-1}, s_i, s_{i+1})}. \tag{13}$$

for $i = 2, 3, \cdots, N$, $s_i = \pm 1$. The inverse temperature is $\beta = 1/(kT)$, so for $T = T_N$, $\beta = J_0/\sigma_J^2$. Then, let $\psi_N^{<}(s_{N-1}, s_N) = 1$ and compute

$$\psi_i^{<}(s_{i-1}, s_i) = \sum_{s_{i+1}} \psi_{i+1}^{<}(s_i, s_{i+1}) e^{-\beta H_i(s_{i-1}, s_i, s_{i+1})} \tag{14}$$

for $i = N - 1, \cdots, 2, 1$, $s_i = \pm 1$. Finally, the decoded sequence is obtained by setting $\hat{a}_i = \text{sgn}\ (m_i)$, $i = 1, 2, \cdots, N$, where the m_i's are computed from Eq. (12).

This algorithm can be visualized on a Viterbi decoder trellis, each column having four states, labelled $(+1, +1), (+1, -1), (-1, +1)$, and $(-1, -1)$. The branch metric from (s_{i-1}, s_i) to (s_i, s_{i+1}) is $H_i(s_{i-1}, s_i, s_{i+1})$, and $\psi_i^{>}(s_i, s_{i+1})$ is the path metric at (s_i, s_{i+1}). The difference from the Viterbi

algorithm is that instead of the path metric being solely determined by the best incoming path, it is a weighted sum of the path metrics of the incoming paths. After the $\psi_i^>$'s are computed from left to right, traceback is done by simultaneously computing $\psi_i^<$ and $\hat{a}_i$ from right to left. Unlike the Viterbi algorithm, however, instead of selecting a single best path, the algorithm computes a weighted average of all possible paths. If $T = T_N$, then this weighted average, when quantized to ± 1, yields the sequence of most likely bits, whereas the Viterbi decoder yields the most likely sequence of bits.

As T decreases, we see that the best (lowest energy) incoming path is weighted increasingly more heavily relative to the other incoming path, with the ratio of the two weights approaching infinity as $T \to 0$. Thus, the algorithm reduces to the Viterbi algorithm at $T = 0$, although renormalization is necessary to avoid all weights being infinite. The algorithm can also be modified, as the Viterbi algorithm usually is, to allow traceback to begin before reaching the end of the received symbols, at the cost of a slight loss of performance.

Figure 5 shows the results of decoding the convolutional code described above at $T = 0.1T_N$, $T = T_N$ and $T = 2T_N$. Each point represents approximately 10,000 bit errors. The bit error rate is lowest for $T = T_N$, although the difference is minuscule. The case $T = 0$ was not simulated, because the reduction to the Viterbi algorithm is not automatic, since simply using the value $T = 0$ results in dividing by zero. However, the performance at $T = 0$ should be only microscopically worse than at $T = 0.1T_N$.

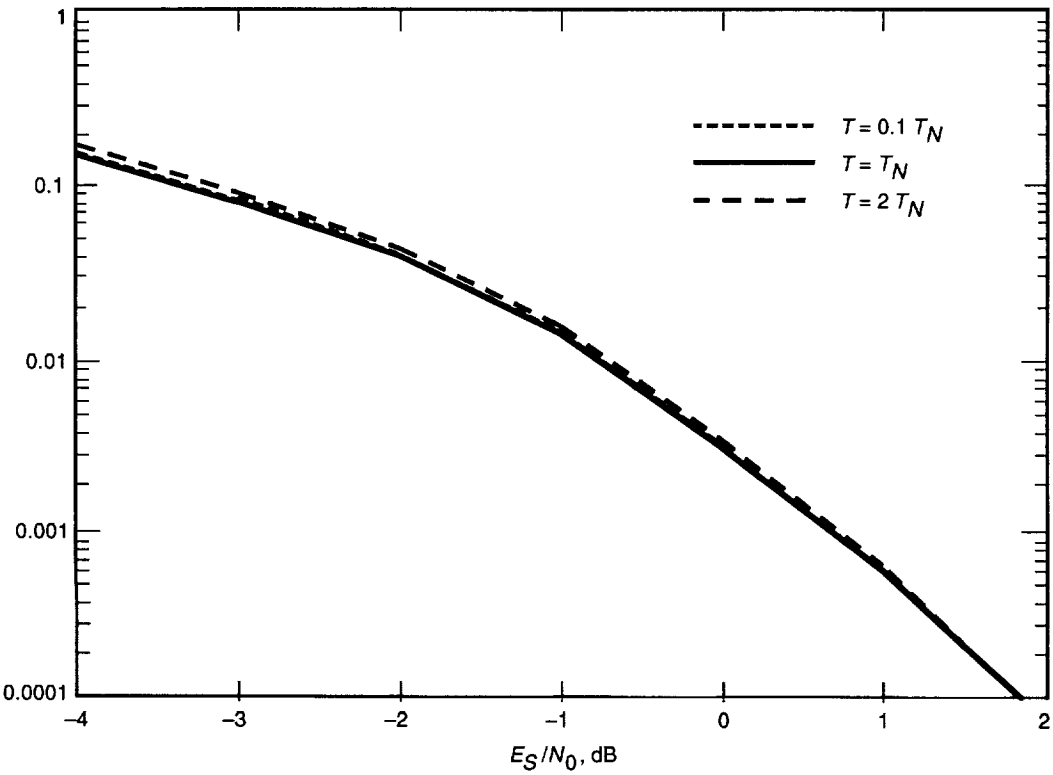

Fig. 5. Bit error rates for a rate 1/2 convolutional code decoded at three different temperatures.

VIII. Conclusion

We simulated the performance of some simple examples of a family of codes whose performance has been shown to asymptotically approach the Shannon bound. The codes have parameters N and p, with codeword length $\binom{N}{p}$ and rate $N/\binom{N}{p}$. The largest code we were able to simulate had a codeword length

of 15,504, a rate of 0.00129, and took over a week to accumulate only 25 codeword errors. Even for this code, the performance was far from capacity. Thus, these codes, although theoretically interesting, are probably not particularly useful, unless radical new decoding methods are developed. One possible such method might be to use an actual spin glass, with coupling coefficients specified by a received codeword, and let it come to equilibrium; then measure the spins to obtain the decoded data. However, this would be, at the very least, technically extremely challenging, and might even be physically impossible.

We also did simulations to measure the improvement in bit error rate achievable by using a decoding method analogous to computing the thermal average of a spin glass. The performance at the optimum temperature, which minimizes bit error rate, was compared to the performance of a standard decoder that minimizes codeword error rate. We tested some examples of the codes described in Section IV and also a (24,12) Golay code, an (8,4) Hamming code, and a simple rate 1/2 convolutional code. For the convolutional code, we used a decoding algorithm based on an algorithm from statistical mechanics for computing thermal averages in one-dimensional systems. For all the codes, there was a measurable but very small improvement in bit error rate. The improvement did not justify the increased decoding complexity, even for the convolutional code, where the decoding algorithm had complexity of the same order as the Viterbi algorithm.

References

[1] R. Kindermann and J. L. Snell, *Markov Random Fields and their Applications*, Providence, Rhode Island: American Mathematical Society, 1980.

[2] N. Sourlas, "Spin-Glass Models as Error-Correcting Codes," *Nature*, vol. 338, pp. 693–695, 1989.

[3] R. McEliece, *The Theory of Information and Coding*, Reading, Massachusetts: Addison-Wesley, 1977.

[4] B. Derrida, "Random-Energy Model: An Exactly Solvable Model of Disordered Systems," *Physical Review B*, vol. 24, pp. 2613–2626, 1981.

[5] E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines*, Chichester, England: Wiley, 1989.

[6] P. Rujan, "Finite Temperature Error-Correcting Codes," *Physical Review Letters*, vol. 70, pp. 2968–2971, 1993.

[7] H. Nishimori, "Internal Energy, Specific Heat and Correlation Function of the Bond-Random Ising Model," *Progress of Theoretical Physics*, vol. 66, pp. 1169–1181, 1981.

[8] P. Rujan, "Calculation of the Free Energy of Ising Systems by a Recursion Method," *Physica*, vol. 91A, pp. 549–562, 1978.

[9] H. Nishimori,"Optimum Decoding Temperature for Error-Correcting Codes," *Journal of the Physical Society of Japan*, vol. 62, pp. 2973–2975, 1993.

[10] N. Sourlas,"Spin Glasses, Error-Correcting Codes and Finite-Temperature Decoding," *Europhysics Letters*, vol. 25, pp. 159–164, 1994.