arXiv:cond-mat/0104079 v1 4 Apr 2001

The glassy phase of Gallager codes

Andrea Montanari

Laboratoire de Physique Théorique de l'Ecole Normale Supérieure[*]
24, rue Lhomond, 75231 Paris CEDEX 05, FRANCE
Internet: `Andrea.Montanari@lpt.ens.fr`

October 1, 2006

Abstract

Gallager codes are the best error-correcting codes to-date. In this paper we study them by using the tools of statistical mechanics. The corresponding statistical mechanics model is a spin model on a sparse random graph. The model can be solved by elementary methods (i.e. without replicas) in a large connectivity limit. For low enough temperatures it presents a completely frozen glassy phase ($q_{EA} = 1$). The same scenario is shown to hold for finite connectivities. In this case we adopt the replica approach and exhibit a one-step replica symmetry breaking order parameter. We argue that our ansatz yields the exact solution of the model. This allows us to determine the whole phase diagram and to understand the performances of Gallager codes.

LPTENS 01/19

[*]UMR 8549, Unité Mixte de Recherche du Centre National de la Recherche Scientifique et de l' Ecole Normale Supérieure.

1 Introduction

Information theory [1, 2] deals with the problem of reliable communication through an imperfect (noisy) communication channel. This can be done by properly encoding the information message in such a way to increase its redundancy. If a transmission error occurs due to the noise, the correct message can be restored by exploiting this redundancy.

The price to pay for error-correction to be possible is to increase the length of the transmitted message, i.e. to decrease the information rate through the channel. In 1948 C. E. Shannon [3] computed the maximal achievable rate at which information can be transmitted through a given communication channel (the so-called *capacity* of the channel). Since then a lot of work has been spent for constructing practical error-correcting codes that could realize Shannon prediction, i.e. that could saturate the channel capacity.

In the past few years it has become progressively clear that such an objective is not unreachable. It has become possible to construct error-correcting codes which remain effective extremely near to the Shannon capacity [4]. The reasons of this revolution have been the invention of "turbo codes" [5] and the re-invention of "low-density parity check codes" (LDPCC) [6]. The last ones [7] were proposed for the first time by R. Gallager in 1962, but were soon forgotten afterwards, probably because of the lack of computational resources at that time.

As it has been shown by N. Sourlas [8–10], error-correcting codes can be mapped onto disordered spin models. This mapping allows to employ statistical mechanics techniques to investigate the behavior of the former. Both turbo codes [11, 12] and LDPCC [13–19] have been already studied using this approach. However all previous studies were restricted to particular regions of the phase diagram. The principal technical reason was the difficulty of implementing replica symmetry breaking in finite connectivity systems.

In this work we focus on regular Gallager codes (a particular family of LDPCC), and we address the fundamental problem of determining the corresponding phase diagram. There are two type of motivations for such a task to be undertaken. First, the spin model corresponding to Gallager codes is a disordered spin model on a diluted graph. The study of such systems has greatly improved our understanding of glassy systems over the last few years. Second, it is of great practical importance to have a complete quantitative picture of the behavior of Gallager codes. For instance, the existence of a glassy phase can have important effects on the decoding algorithms, and the knowledge of the phase diagram can be used to improve them.

The model is presented Sec. 2. In Sec. 3 we prove some exact properties which hold at inverse temperature $\beta = 1$. The line $\beta = 1$ can be regarded as the Nishimori line [20] of the phase diagram. In Sec. 4 we solve the model in the large connectivity limit. We show that it becomes identical to a simplified model which we call the *random codeword model* (RCM). The RCM is shown to have a freezing phase transition analogous to the one of the random energy model (REM) [21]. In Sec. 5 we adopt the replica approach [22] and prove that the same scenario applies for finite connectivities. In particular we construct a replica symmetry breaking solution of the saddle point equations. The proposed solution is much simpler than the generic one-step replica symmetry breaking solution. Rather than being parametrized by a functional over a probability space [23], it depends simply upon the probability distribution of a local field. Such a probability distribution can be easily computed numerically. It can be also obtained from a large connectivity expansion, see Sec. 6. In Sec. 7 we compute the finite-size corrections of the free energy for the RCM, and compare the result with exact enumerations. Finally in Sec. 8 we discuss the validity of our replica symmetry breaking ansatz.

2 The model

Let us suppose we want to transmit an information message consisting of L bits. There are 2^L such messages. Each of them is encoded in a string of $N > L$ bits (*codewords*).

This motivates the following model. There are 2^L possible configurations of the system (the codewords), each one corresponding to a distinct sequence of $N > L$ bits. We shall denote the codewords as $\underline{x}^{(\alpha)} = (x_1^{(\alpha)}, \ldots, x_N^{(\alpha)})$, with $\alpha = 1, \ldots, 2^L$. The set of codewords $\mathcal{C}$ is a linear space. This means that $\underline{0} \equiv (0, \ldots, 0) \in \mathcal{C}$, and that, if $\underline{x}^{(\alpha)}, \underline{x}^{(\beta)} \in \mathcal{C}$, then $\underline{x}^{(\alpha)} + \underline{x}^{(\beta)} \in \mathcal{C}$ (where the sum has to be carried modulo 2).

Like any linear space, the set of codewords $\mathcal{C}$ can be specified as the kernel of a linear operator. In other words, we can find an M by N matrix $\mathbb{C} = \{C_{ij}\}_{i=1\ldots M,\, j=1\ldots N}$, with $C_{ij} = 0, 1$, and $M = N - L$, such that

$$\mathcal{C} = \{\underline{x}^{(\alpha)} : \alpha = 1, \ldots, 2^L\} = \{\underline{x} \in \{0,1\}^N : \mathbb{C}\underline{x} = \underline{0} \,(\mathrm{mod}\, 2)\}\,. \tag{2.1}$$

The condition $\mathbb{C}\underline{x} = \underline{0}\,(\mathrm{mod}\, 2)$ can be regarded as a set of M linear equations (called *constraints* or *parity checks*) of the form:

$$C_{i1}x_1 + C_{i2}x_2 + \ldots + C_{iN}x_N = 0 \;(\mathrm{mod}\, 2)\,, \tag{2.2}$$

with $i = 1, \ldots, M$.

To each bit x_i, $i = 1, \ldots, N$, we assign an *a priori* probability distribution $p_i(x_i)$. In the information-theory context, the *a priori* distributions $p_i(x_i)$ are induced by the observation of the channel output, and by the knowledge of the statistical properties of the channel. We are interested in studying the induced probability distribution over the codewords $\underline{x}^{(\alpha)}$. In other words we want to consider the following probability distribution over the strings $\underline{x}$ of N bits:

$$P(\underline{x}) = \frac{1}{Z}\,\delta[\mathbb{C}\underline{x}] \prod_{i=1}^{N} p_i(x_i)\,, \tag{2.3}$$

where Z is a normalization constant; $\delta[\underline{z}] = 1$ if $\underline{z} = \underline{0}\,(\mathrm{mod}\, 2)$, and $\delta[\underline{z}] = 0$ otherwise.

There are several graphical representations of the above model. The most used in the coding theory community makes use of the so-called Tanner graph [24], cf. Fig. 1. This is a bipartite graph which is constructed as follows. A node on the left is associated to each binary variable x_j, and a node on the right to each constraint, i.e. to each linear equation (2.2) with $i = 1, \ldots, M$. There are therefore N left nodes (*variable* nodes), and M right nodes (*check* nodes). A given check i is connected to the variables x_j which appear with nonzero coefficient in the corresponding equation (2.2).

The model (2.3) has a spin-wise formulation [13–19] which we shall employ hereafter. We replace any bit sequence $\underline{x} = (x_1, \ldots, x_N)$, with a spin configuration $\underline{\sigma} = (\sigma_1, \ldots, \sigma_N)$, where $\sigma_i = (-1)^{x_i}$. The constraints (2.2) on the sums of bits x_i, get translated into constraints on the product of spins σ_i. These have the form

$$\sigma^{\omega_i} \equiv \prod_{j \in \omega_i} \sigma_j = +1\,, \tag{2.4}$$

where $\omega_i = \{j \in \{1, \ldots, N\} : C_{ij} = 1\}$. The other ingredient of the model are the *a priori* probability distributions $p_i(x_i)$. They can be encoded into properly chosen magnetic fields: $p_i(x_i) = e^{\beta h_i \sigma_i}/(2\cosh \beta h_i)$, with $2\beta h_i = \log(p_i(0)/p_i(1))$, where we introduced the inverse

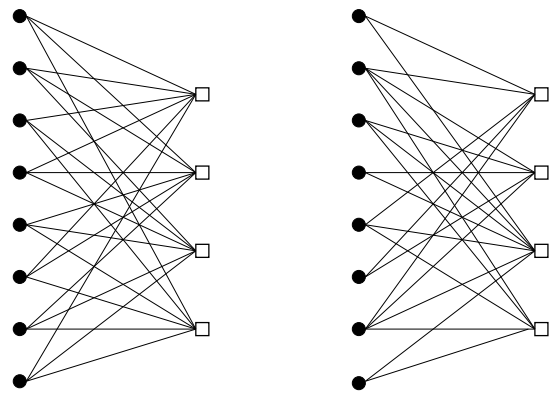

Figure 1: Two Tanner graphs: a regular one with $(k, l) = (6, 3)$ on the left, and an irregular one on the right. In both cases $N = 8$, $M = 4$ (and therefore the rate is $R = 1/2$).

temperature β for later convenience. With these building blocks, we can write down the spin model equivalent of Eq. (2.3):

$$P(\underline{\sigma}) = \frac{1}{Z(\beta)} \prod_{j=1}^{M} \delta[\sigma^{\omega_j}, +1] \exp\left(\beta \sum_{i=1}^{N} h_i \sigma_i\right) , \tag{2.5}$$

where $\delta[a, b]$ is the Kronecker delta function. This can be regarded as a spin model with infinite strength multi-spin interactions (which enforce $\sigma^{\omega_j} = +1$) and a random magnetic field.

Instead of insisting on the motivations for the probabilistic model (2.5) coming from coding theory, we shall remark that, as it stands, it is remarkably general. Any spin-model hamiltonian $H(\underline{\sigma}) = -\sum_{i_1 \ldots i_p} J_{i_i \ldots i_p} \sigma_{i_1} \ldots \sigma_{i_p}$ can be written in the form (2.5). This can be done by introducing the auxiliary spin variables $\sigma_{i_1 \ldots i_p}$. The Kronecker delta functions in Eq. (2.5) can be used to enforce $\sigma_{i_1 \ldots i_p} = \sigma_{i_1} \ldots \sigma_{i_p}$. The couplings $J_{i_i \ldots i_p}$ become magnetic fields acting on the variables $\sigma_{i_1 \ldots i_p}$.

Untill now we have been pretty generic in the presentation of the model. In order to be more precise, we have to choose the constraint matrix $\mathbb{C}$, and the magnetic fields $\{h_i\}_{i=1,\ldots,N}$.

Following Gallager [7], we shall take $\mathbb{C}$ to be *random* and *sparse*. More precisely $\mathbb{C}$ will be constrained to have k non-zero elements for each row and l non-zero elements for each column (with $l < k$), and not to have two identical rows[1]. This choice corresponds to taking the Tanner graph (cf. Fig. 1) as a random bipartite graph, with variable (*left*) nodes of fixed degree l, and check (*right*) nodes of degree k. We shall choose among the matrices of this *ensemble* with flat probability distribution. We shall use the pair (k, l) to denote the spin model (or the error-correcting code) defined by this *ensemble* of matrices. An important characteristic of the code is its *rate* $R = 1 - l/k$, which measures the redundancy of the encoded message (infact $R = L/N$).

The magnetic fields h_i will be random i.i.d. variables with probability distribution $p_h(h_i)$. We consider $p_h(h_i)$ to be biased towards positive values of h_i (i.e. $\int dh_i \, p_h(h_i) h_i > 0$). We

[1]Remark that, with this choice, some of the parity check equations (2.2) may be linearly dependent. However, such an event is *rare* for $k > l$ [7].

shall refer often to two simple examples: the two-peak distribution

$$p_h(h_i) = (1-p)\delta(h_i - h_0) + p\delta(h_i + h_0)\,, \tag{2.6}$$

with $p < 1/2$ and $h_0 > 0$, and the gaussian distribution

$$p_h(h_i) = \frac{1}{\sqrt{2\pi\tilde{h}^2}} \exp\left\{-\frac{(h_i - h_0)^2}{2\tilde{h}^2}\right\}\,, \tag{2.7}$$

with $h_0 > 0$. It can be shown that, if the model describe communication through a noisy "symmetric" channel, the condition

$$p_h(-h_i) = e^{-2h_i} p_h(h_i) \tag{2.8}$$

follows. This implies $h_0 = (1/2)\log(1-p)/p$ for the example (2.6) (which corresponds to a binary symmetric channel), and $h_0 = \tilde{h}^2$ for the example (2.7) (corresponding to a gaussian channel). Hereafter we shall denote with $\langle\cdot\rangle_h$ and $\langle\cdot\rangle_{\mathbb{C}}$ the averages with respect to the magnetic fields $\{h_i\}$, and the *ensemble* of matrices $\mathbb{C}$.

More details on the model introduced in this Section, and on analogous examples can be found in Refs. [11–19]

3 The Nishimori line

Nishimori [20, 25] showed that the physics of disordered spin models simplifies considerably on a particular line in the phase diagram. In particular, it has been recently shown [26] that replica symmetry breaking is absent on this line. The Nishimori line plays a distinguished role in the correspondence between error-correcting codes and disordered spin models. As shown in Refs. [27, 28], *maximum a posteriori symbol probability* (MAP) decoding for a given error-correcting code is equivalent to computing expectation values on the Nishimori line of the corresponding spin model.

In this Section we extend the results concerning the Nishimori line to the model (2.5). We shall consider a generic magnetic field distribution $p_h(h_i)$ satisfying Eq. (2.8). In this case the Nishimori line is simply given by $\beta = 1$. Although the proofs are very similar to the ones of Refs. [25, 26], we present them for sake of completeness. Some consequences of the exact results of this Section will be outlined in Sec. 5.

Let us start with some convention. Notice that there are two sources of disorder in our model (2.3): the magnetic field h_i (which is determined by the channel output), and the check matrix $\mathbb{C}$. Different $\mathbb{C}$ correspond to different error-correcting codes. In this Section we keep the parity check matrix $\mathbb{C}$ fixed, and average uniquely over the random magnetic fields $\{h_i\}$, with distribution $p_h(h_i)$. Our results will remain valid after averaging with respect to any *ensemble* of check matrices $\mathbb{C}$ (i.e. to any *ensemble* of codes). It is convenient to introduce the notation $\delta_{\mathbb{C}}[\underline{\sigma}]$ to denote the product of Kronecker delta functions in Eq. (2.5). In other words $\delta_{\mathbb{C}}[\underline{\sigma}] = 1$, if and only if $\underline{\sigma}$ satisfies all the parity checks encoded in $\mathbb{C}$, i.e. if the corresponding string of bits $\underline{x}$ is a codeword. We assume that the parity check matrix $\mathbb{C}$ selects $2^L = 2^{NR}$ codewords. This means that there are 2^L distinct configurations $\underline{\sigma}$, such that $\delta_{\mathbb{C}}[\underline{\sigma}] = 1$. Finally we shall take the distribution of the random fields to satisfy the identity (2.8).

We start by writing down the definition of the (field averaged) free energy density $f_{\mathbb{C}}(\beta)$ for a given parity check matrix $\mathbb{C}$:

$$-\beta N f_{\mathbb{C}}(\beta) = \int_{-\infty}^{+\infty} \prod_{i=1}^{N} dh_i\, p_h(h_i)\, \log\left\{\sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}]\, e^{\beta\sum_i h_i\sigma_i}\right\}\,. \tag{3.1}$$

Then we notice, following Ref. [25], that the integral over the field h_i can be decomposed into an integral over its absolute value and a sum over its sign. Using Eq. (2.8), we get, for any function $\mathcal{O}(h_i)$

$$\int_{-\infty}^{+\infty} dh_i\, p_h(h_i)\mathcal{O}(h_i) = \int_0^{+\infty} dh_i\, \rho(h_i) \sum_{\tau_i} e^{h_i\tau_i}\mathcal{O}(h_i\tau_i)\,, \tag{3.2}$$

where $\rho(h_i)$ is given by

$$\rho(h_i) = \frac{p_h(h_i) + p_h(-h_i)}{2\cosh h_i}\,. \tag{3.3}$$

By using the decomposition (3.2) into the definition (3.1), we get

$$-\beta N f_{\mathbb{C}}(\beta) = \int_0^{+\infty} \prod_{i=1}^N dh_i\, \rho(h_i) \sum_{\underline{\tau}} e^{\sum_i h_i\tau_i} \log\left\{\sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}]\, e^{\beta\sum_i h_i\tau_i\sigma_i}\right\}\,. \tag{3.4}$$

To be more compact, we shall use hereafter the shorthand $\langle\cdot\rangle_\rho \equiv \int_0^{+\infty}\prod_{i=1}^N dh_i\,\rho(h_i)\,(\cdot)$ for the average over the absolute values of the fields $\{h_i\}$.

The next step consists in performing a gauge transformation $\tau_i \to \sigma_i'\tau_i$, $\sigma_i \to \sigma_i'\sigma_i$. Because of the constraint term $\delta_{\mathbb{C}}[\underline{\sigma}]$, the free energy (3.4) is not invariant with respect to such a transformation for a generic choice of $\{\sigma_i'\}$. However, if $\delta_{\mathbb{C}}[\underline{\sigma}'] = 1$, i.e. if $\underline{\sigma}'$ is a codeword, then the gauge transformation leaves invariant the free energy. We can sum over all such "allowed" transformations, and divide by their number, namely 2^{NR}, obtaining

$$-\beta N f_{\mathbb{C}}(\beta) = \left\langle \frac{1}{2^{NR}} \sum_{\underline{\tau}} \sum_{\underline{\sigma}'} \delta_{\mathbb{C}}[\underline{\sigma}'] e^{\sum_i h_i\tau_i\sigma_i'} \log\left\{\sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}]\, e^{\beta\sum_i h_i\tau_i\sigma_i}\right\}\right\rangle_\rho\,, \tag{3.5}$$

where the constraint $\delta_{\mathbb{C}}[\underline{\sigma}']$ force the gauge transformation $\underline{\sigma}'$ to be an allowed one.

In Eq. (3.5) we wrote the sums over quenched and dynamical variables in a symmetric form. This allows to derive several exact identities for $\beta = 1$, where the symmetry is complete. In particular, let us consider the internal energy per spin $\epsilon_{\mathbb{C}}(\beta) = \partial_\beta(\beta f_{\mathbb{C}}(\beta))$. From Eq. (3.5) we get

$$\epsilon_{\mathbb{C}}(\beta = 1) = -\left\langle \frac{1}{2^{NR}} \sum_{\underline{\tau}} \sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}] \left(\frac{1}{N}\sum_{i=1}^N h_i\tau_i\sigma_i\right) e^{\sum_i h_i\tau_i\sigma_i}\right\rangle_\rho\,. \tag{3.6}$$

We can now perform a second gauge transformation $\tau_i \to \tau_i\sigma_i$, sum over the $\{\sigma_i\}$ using the constraint, and finally sum over the τ_i. We obtain $\epsilon_{\mathbb{C}}(\beta = 1) = -\langle h \tanh h\rangle_h$. Analogously to Ref. [25], we can further simplify this result, obtaining

$$\epsilon_{\mathbb{C}}(\beta = 1) = -\langle h\rangle_h\,, \tag{3.7}$$

which is the first important result of this Section.

We want now to prove the absence of replica symmetry breaking on the Nishimori line of our model (2.3), i.e. for $\beta = 1$. As in Ref. [26], we consider the magnetization distribution

$$P^{(1)}_{\beta,\mathbb{C}}(m) \equiv \int_{-\infty}^{+\infty} \prod_{i=1}^N dh_i\, p_h(h_i) \frac{\sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}]\, e^{\beta\sum_i h_i\sigma_i}\, \delta(m - N^{-1}\sum_i \sigma_i)}{\sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}]\, e^{\beta\sum_i h_i\sigma_i}}\,, \tag{3.8}$$

and the overlap distribution

$$P^{(2)}_{\beta,\mathbb{C}}(q) \equiv \int_{-\infty}^{+\infty} \prod_{i=1}^{N} dh_i\, p_h(h_i) \frac{\sum_{\underline{\sigma},\underline{\sigma}'} \delta_{\mathbb{C}}[\underline{\sigma}]\, \delta_{\mathbb{C}}[\underline{\sigma}']\, e^{\beta \sum_i h_i \sigma_i + \beta \sum_i h_i \sigma'_i}\, \delta(q - N^{-1} \sum_i \sigma_i \sigma'_i)}{\sum_{\underline{\sigma},\underline{\sigma}'} \delta_{\mathbb{C}}[\underline{\sigma}]\, \delta_{\mathbb{C}}[\underline{\sigma}']\, e^{\beta \sum_i h_i \sigma_i + \beta \sum_i h_i \sigma'_i}} \,. \tag{3.9}$$

As before, we keep the parity check matrix $\mathbb{C}$ fixed. We shall prove that the two probability distributions defined above are indeed identical on the Nishimori line $\beta = 1$, i.e. $P^{(1)}_{1,\mathbb{C}}(x) = P^{(2)}_{1,\mathbb{C}}(x)$. Since the probability distribution of the magnetization is expected to be a single delta function[2] [22], this implies the absence of replica symmetry breaking for $\beta = 1$.

We begin by using the decomposition (3.2) in Eq. (3.8). This yields:

$$P^{(1)}_{\beta,\mathbb{C}}(m) = \left\langle \sum_{\underline{\tau}} e^{\sum_i h_i \tau_i} \frac{\sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}]\, e^{\beta \sum_i h_i \tau_i \sigma_i}\, \delta(m - N^{-1} \sum_i \sigma_i)}{\sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}]\, e^{\beta \sum_i h_i \tau_i \sigma_i}} \right\rangle_{\rho} \,. \tag{3.10}$$

Then we notice that the above distribution is invariant under an "allowed" gauge transformation $\tau_i \to \sigma'_i \tau_i$, $\sigma_i \to \sigma'_i \sigma_i$. As before, "allowed" means that $\delta_{\mathbb{C}}[\underline{\sigma}'] = 1$. We can therefore average over these transformations, obtaining

$$P^{(1)}_{\beta,\mathbb{C}}(m) = \left\langle \sum_{\underline{\tau},\underline{\sigma}'} \delta_{\mathbb{C}}[\underline{\sigma}'] e^{\sum_i h_i \tau_i \sigma'_i} \frac{\sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}]\, e^{\beta \sum_i h_i \tau_i \sigma_i}\, \delta(m - N^{-1} \sum_i \sigma_i \sigma'_i)}{2^{NR} \sum_{\underline{\sigma}} \delta_{\mathbb{C}}[\underline{\sigma}]\, e^{\beta \sum_i h_i \tau_i \sigma_i}} \right\rangle_{\rho} \,. \tag{3.11}$$

We then insert $1 = (\sum_{\underline{\widehat{\sigma}}} \delta_{\mathbb{C}}[\underline{\widehat{\sigma}}] e^{\sum_i h_i \tau_i \widehat{\sigma}_i}) / (\sum_{\underline{\sigma}'} \delta_{\mathbb{C}}[\underline{\sigma}'] e^{\sum_i h_i \tau_i \sigma'_i})$, perform a second gauge transformation $\tau_i \to \widehat{\sigma}_i \tau_i$, $\sigma_i \to \widehat{\sigma}_i \sigma_i$, $\sigma'_i \to \widehat{\sigma}_i \sigma'_i$, and sum over $\underline{\widehat{\sigma}}$. Finally we set $\beta = 1$, obtaining $P^{(1)}_{1,\mathbb{C}}(m) = P^{(2)}_{1,\mathbb{C}}(m)$, as anticipated above.

4 The random codeword limit

The limiting case $k, l \to \infty$, with $l/k = 1 - R$ fixed, plays an important role. We shall call it the random codeword limit for reasons which will be clear later. It is a non-trivial limit since the redundancy of the error-correcting code is kept fixed. From a theoretical point of view, it allows a simple solution of the model without changing its qualitative features. Our methods will be similar to the ones used by Derrida to solve the REM [21]. Finally, we will show that the corrections for finite values of k and l are exponentially small in k. Therefore this limit is interesting also from a quantitative point of view.

4.1 The limit $k, l \to \infty$

Let us consider the probability for a given sequence of bits $\underline{x} = (x_1, \ldots, x_N)$ to be a codeword with respect to the *ensemble* of parity check matrices $\mathbb{C}$. This coincides with the probability $P_{\underline{\sigma}}$ for a given spin configuration $\underline{\sigma}$ to satisfy the constraints (2.4). In other words:

$$P_{\underline{\sigma}} \equiv \frac{1}{\mathcal{N}_{\mathbb{C}}} \sum_{\mathbb{C}} \prod_{j=1}^{M} \delta[\sigma^{\omega_j}, +1] \,, \tag{4.1}$$

[2]Notice that our model (2.3) has no spin-reversal symmetry.

where the sum over $\mathbb{C}$ runs over all the matrices of the (k,l)-*ensemble* , and $\mathcal{N}_{\mathbb{C}}$ is their number.

Clearly $P_{\underline{\sigma}}$ depend upon $\underline{\sigma}$ uniquely through the magnetization $m_\sigma \equiv (1/N)\sum_i \sigma_i$. In general it has the form

$$P_{\underline{\sigma}} \sim \exp\left[N\Sigma_1^{(k,l)}(m_\sigma)\right] . \tag{4.2}$$

The function $\Sigma_1^{(k,l)}(m)$ is computed in Appendix A for general values of k and l, and is not particularly illuminating. However, in the limit $k,l \to \infty$, $l/k = 1-R$ fixed, we have

$$\Sigma_{(k,l)}(m) \to -(1-R)\log 2 , \tag{4.3}$$

for any $-1 < m < 1$. In other words any spin configuration $\underline{\sigma}$ has the same probability $P_{\underline{\sigma}} \sim 2^{-(1-R)N}$ of being a codeword. In addition we must keep track of the completely ordered configurations $\sigma_i = +1$ for $i = 1,\dots,N$, and $\sigma_i = -1$ for $i = 1,\dots,N$. The positive one satisfies the all constraints for any k and l, and for any matrix $\mathbb{C}$ (this configuration is quite important for the thermodynamics of the model). The negative one satisfies the constraints for k even, but it is irrelevant for the thermodynamics.

Let us now turn to a slightly more complicated quantity. We consider the joint probability $P_{\underline{\sigma},\underline{\tau}}$ for two different spin configurations $\underline{\tau}$ and $\underline{\sigma}$ to satisfy the same set of constraints (2.4), corresponding to some matrix $\mathbb{C}$ taken from the (k,l)-*ensemble*. In formulae:

$$P_{\underline{\sigma},\underline{\tau}} = \frac{1}{\mathcal{N}_{\mathbb{C}}}\sum_{\mathbb{C}}\prod_{j=1}^{M}\delta[\sigma^{\omega_j},+1]\delta[\tau^{\omega_j},+1] . \tag{4.4}$$

As before we can argue that $P_{\underline{\sigma},\underline{\tau}}$ depends upon $\underline{\sigma}$ and $\underline{\tau}$ only through their magnetizations m_σ, m_τ, and their overlap $q \equiv (1/N)\sum_i \sigma_i\tau_i$. The form of $P_{\underline{\sigma},\underline{\tau}}$ in the thermodynamic limit is

$$P_{\underline{\sigma},\underline{\tau}} \sim \exp[N\Sigma_2^{(k,l)}(m_\sigma, m_\tau, q)] . \tag{4.5}$$

The function $\Sigma_2^{(k,l)}(m_1,m_2,q)$ is computed in Appendix A. Again, we shall not report here the result, but we remark that in the $k,l \to \infty$ limit

$$\Sigma_2^{(k,l)}(m_1,m_2,q) \to -2(1-R)\log 2 , \tag{4.6}$$

for any $-1 < m_1, m_2, q < 1$. In other words, the probability for two configurations $\underline{\sigma}$, and $\underline{\tau}$ to satisfy the same set of constraints is $P_{\underline{\sigma},\underline{\tau}} \sim P_{\underline{\sigma}}P_{\underline{\tau}} \sim 2^{-2(1-R)N}$: the two configurations can be regarded as independent ones.

4.2 The random codeword model

The previous considerations allow us to replace (in the $k,l \to \infty$ limit) the original model (2.5), with the following *random codeword model* (RCM). The model has 2^{NR} possible states which we shall index with the letter $\alpha = 1,\dots,2^{NR}$. To each of these states we associate a *random* spin configuration $\underline{\sigma}^{(\alpha)} = (\sigma_1^{(\alpha)},\dots,\sigma_N^{(\alpha)})$. By *random* we mean that each spin $\sigma_i^{(\alpha)}$ is chosen independently from the others, and that $\sigma_i^{(\alpha)} = +1$ or -1 with equal probability. Let us underline that, in the random codeword model, the $\sigma_i^{(\alpha)}$ are quenched variables, the dynamical one being the index α. There is a second set of quenched variables: the magnetic fields h_i,

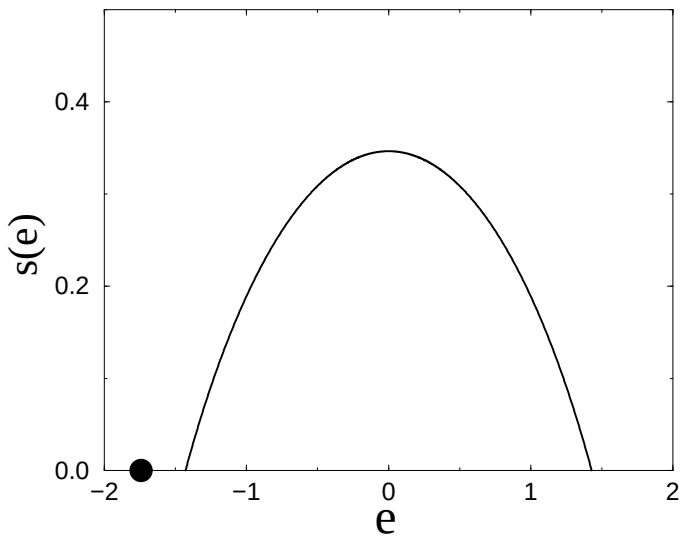

Figure 2: The microcanonical entropy density of the RCM with binary field distribution, cf. Eq. (2.6). Here we set $R = 1/2$, $p = 0.025$, $h_0 = \text{arctanh}(1 - 2p)$. Notice the continuous contribution coming from the random configurations (solid line), and the isolated ordered configuration (filled circle).

with $i = 1, \ldots, N$. As in the original model we take them to be random i.i.d. variables with distribution $p_h(h_i)$. The energy of the state α reads

$$E^{(\alpha)} = -\sum_{i=1}^{N} h_i \sigma_i^{(\alpha)} \,. \tag{4.7}$$

To the 2^{NR} "disordered" states described above we add the ordered state $\alpha = 0$, and the corresponding spin configuration $\underline{\sigma}^{(0)}$, with $\sigma_i^{(0)} = +1$ for $i = 1, \ldots, N$. This corresponds to the "all zeros" codeword $\underline{0}$. Its energy is obviously $E^{(0)} = -\sum_i h_i$.

The random codeword model can be solved through elementary methods. Here we shall solve it for the $\pm h_0$ distribution of fields, see Eq. (2.6). At the end of this Section, we shall quote the result for a general distribution $p_h(h_i)$. For sake of clarity we shall report the calculation for this case, which is slightly less straightforward, in the Appendix B.

We begin by taking into account the "random" states $\alpha = 1, \ldots, 2^{NR}$. Later we shall consider the contribution coming from the ordered state $\alpha = 0$. Let us consider a fixed configuration of the magnetic fields $\{h_i\}$. Since the probability distribution of the $\sigma_i^{(\alpha)}$ is flat, $P(\{\sigma_i^{(\alpha)}\}) = 2^{-N^2 R}$, we can apply a gauge transformation $\sigma_i^{(\alpha)} \to \varepsilon_i \sigma_i^{(\alpha)}$, with $\varepsilon_i = \pm 1$, without changing their statistical properties. If we choose $\varepsilon_i = \text{sign}(h_i)$, the energy (4.7) becomes $E^{(\alpha)} = -h_0 \sum_i \sigma_i^{(\alpha)}$. We conclude that, for what concerns the "random" states, the $\pm h_0$ field distribution is equivalent to an uniform field $h_i = h_0$.

Now we would like to compute the *typical* number $\mathcal{N}_{typ}(\epsilon)$ of states having a given energy density $E^{(\alpha)}/N = \epsilon$. This is equal to the typical number of states having magnetization

$m^{(\alpha)} = -\epsilon/h_0$. This is a very simple problem. Define the function

$$\mathcal{H}(x) = -\frac{1+x}{2}\log(1+x) - \frac{1-x}{2}\log(1-x)\,. \tag{4.8}$$

Then $\mathcal{N}_{typ}(\epsilon) \sim \exp\{NR\log 2 + N\mathcal{H}(\epsilon/h_0)\}$, when $|\epsilon| < \epsilon_c$, and $\mathcal{N}_{typ}(\epsilon) = 0$ otherwise. The critical energy $\epsilon_c = h_0\widehat{\epsilon}(R)$ is the positive solution of $R\log 2 + \mathcal{H}(\epsilon/h_0) = 0$. The entropy density of the system $s(\epsilon) = \log\mathcal{N}_{typ}(\epsilon)/N$ is depicted in Fig. 2. Since $s'(-\epsilon_c) > 0$ the (sub)system of the random codewords undergoes a freezing phase transition at the critical temperature $\beta_c = s'(-\epsilon_c)$. This phase transition is analogous to the one of the REM [21]: it separates an high–temperature paramagnetic phase from a low–temperature frozen one.

Let us now consider the ordered state $\alpha = 0$, whose energy is given by $E^{(0)} = -\sum_i h_i$. In this case we can apply the central limit theorem. For $N \to \infty$ the energy density of the state $\alpha = 0$ is $\epsilon^{(0)} = -(1-2p)h_0$ with probability one. We have therefore the following picture of the energy spectrum of the model: a single ordered state at $\epsilon^{(0)} = -(1-2p)h_0$, plus a bell-shaped continuum between $-\epsilon_c(h_0)$ and $\epsilon_c(h_0)$. The ordered state is thermodynamically relevant as long as it is separated by a gap from the continuum. This happens if $p < p_c(R)$, where $p_c(R)$ is the unique solution between 0 and 1/2 of the equation

$$R\log 2 + \mathcal{H}(1-2p) = 0\,. \tag{4.9}$$

Notice that Eq. (4.9) coincide with the equation determining the capacity of the binary symmetric channel [1]. This means that, in the $k, l \to \infty$ limit, Gallager codes saturate Shannon capacity.

The free energy is easily determined from the entropy:

$$f(\beta) = \min_{\epsilon}\left\{\epsilon - \frac{1}{\beta}s(\epsilon)\right\}\,. \tag{4.10}$$

The phase diagram includes three different phases: a paramagnetic (P) and a spin-glass (SG) phases, associated with the continuum part of the energy spectrum; a ferromagnetic (F) phase, associated with the ordered state. The free energy of the paramagnetic phase is given by:

$$f_P(\beta) = -\frac{R}{\beta}\log 2 - \frac{1}{\beta}\log\cosh\beta h_0\,. \tag{4.11}$$

The paramagnetic-spin glass phase boundary is given by the zero-entropy condition $\partial f_P/\partial\beta = 0$. We obtain the curve $\beta h_0 = \operatorname{arctanh}(1-2p_c(R)) \equiv h^*(R)$. At the transition the system freezes and the free energy in the spin-glass phase is

$$f_{SG}(\beta) = f_P(\beta = h^*(R)/h_0) = -h_0(1-2p_c(R))\,. \tag{4.12}$$

The ferromagnetic free energy is nothing but the energy of the ferromagnetic state:

$$f_F(\beta) = -h_0(1-2p)\,. \tag{4.13}$$

The ferromagnetic-spin glass phase boundary has therefore the simple form $p = p_c(R)$.

For sake of clarity, let us consider the magnetic field distribution which describes a binary symmetric channel, i.e. let us fix $h_0 = h_0(p) \equiv \operatorname{arctanh}(1-2p)$, cf. Eq. (2.8). The resulting phase diagram is reported in Fig. 3. The ferromagnetic-spin glass phase boundary is at

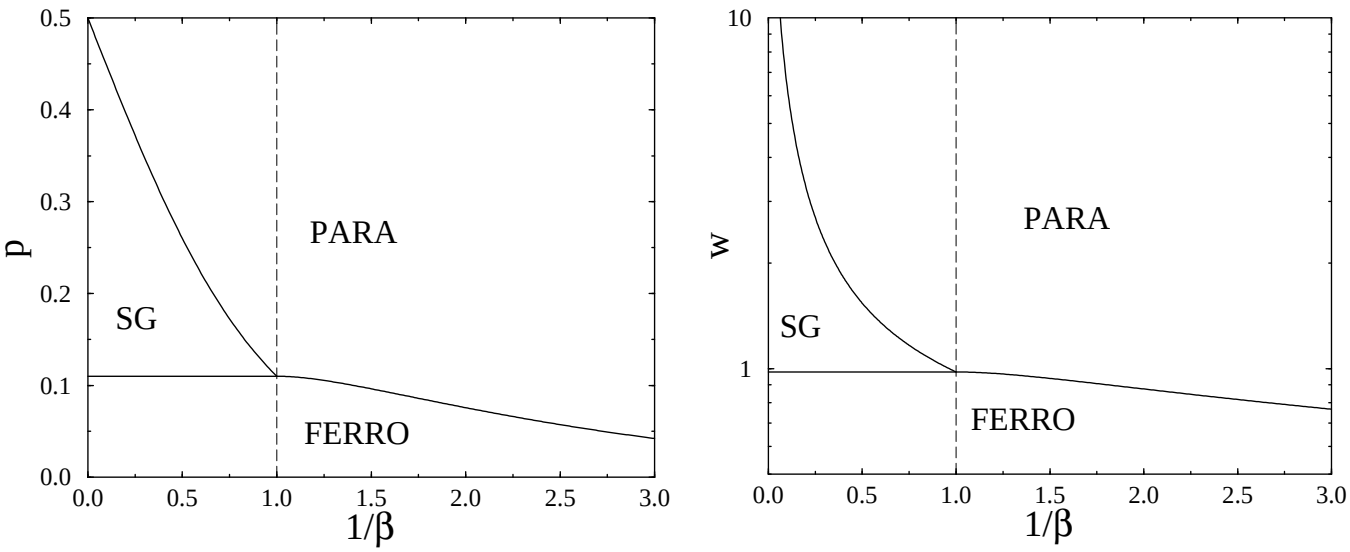

Figure 3: The phase diagram for binary (left, see Eq. (2.6)), and gaussian (right, see Eq. (2.7)) field distribution. In both cases the field distribution was chosen to satisfy Eq. (2.8).

$p = p_c(R)$. The paramagnetic-spin glass boundary is $\beta \operatorname{arctanh}(1-2p) = \operatorname{arctanh}(1-2p_c(R))$. Finally the ferromagnetic-paramagnetic phase boundary is given by

$$R \log 2 + \log \cosh \beta h_0(p) - \beta h_0(p) \tanh h_0(p) = 0 \,. \tag{4.14}$$

The triple point is at $\beta = 1$, $p = p_c(R)$, and lies on the Nishimori line.

Untill now we treated the simple case of a two-peak distribution of the magnetic fields: $p_h(h_i) = (1-p)\,\delta(h_i - h_0) + p\,\delta(h_i + h_0)$. What does it happen for a generic $p_h(h_i)$? In Appendix B it is shown that the same scenario applies with some slight modification. The free energy in the paramagnetic phase becomes

$$f_P(\beta) = -\frac{R}{\beta} \log 2 - \frac{1}{\beta} \langle \log \cosh \beta h \rangle_h \,. \tag{4.15}$$

The system undergoes a freezing transition at a critical temperature β_c determined from the condition $\partial f / \partial \beta|_{\beta_c} = 0$. For $\beta > \beta_c$, the system is in a glassy phase with free energy $f_{SG}(\beta) = f_P(\beta_c)$. Finally, the ferromagnetic phase coincides with the ordered state $\alpha = 0$, and has free energy $f_F(\beta) = -\langle h \rangle_h$.

To be specific we report in Fig. 3 the phase diagram for the gaussian distribution

$$p_h(h) = \sqrt{\frac{w^2}{2\pi}} \exp\left\{ -\frac{w^2}{2} \left[h - \frac{1}{w^2} \right]^2 \right\} \,, \tag{4.16}$$

which describes a gaussian channel with noise variance w. The triple point is located at $\beta = 1$ and $w = w_c(R)$, $w_c(R)$ being the solution of the equation below

$$R \log 2 + \langle \log \cosh h \rangle_h - \langle h \tanh h \rangle_h = 0 \,. \tag{4.17}$$

It is easy to show that the solution $R(w)$ of the above equation correspond to the capacity of a gaussian channel with constrained binary inputs [2].

5 The replica calculation

As always [22] we compute the integer moments $\langle Z^n \rangle_{h,\mathbb{C}}$ of the partition function by replicating the system n times. To the leading exponential order we get

$$\langle Z^n \rangle_{h,\mathbb{C}} \sim \int \prod_{\vec{\sigma}} d\lambda(\vec{\sigma}) d\widehat{\lambda}(\vec{\sigma})\, e^{-NS[\lambda,\widehat{\lambda}]}\,, \tag{5.1}$$

where

$$\begin{aligned} S[\lambda,\widehat{\lambda}] &= l \sum_{\vec{\sigma}} \lambda(\vec{\sigma})\widehat{\lambda}(\vec{\sigma}) - \frac{l}{k} \sum_{\vec{\sigma}_1,\ldots,\vec{\sigma}_k} \lambda(\vec{\sigma}_1)\cdot\ldots\cdot\lambda(\vec{\sigma}_k) \prod_{a=1}^{n} \delta[\sigma_1^a \ldots \sigma_k^a, +1] - \\ &\quad - \log\left\{ \sum_{\vec{\sigma}} \widehat{\lambda}(\vec{\sigma})^l \langle e^{\beta h \sum_a \sigma^a} \rangle_h \right\} - l + \frac{l}{k}\,, \end{aligned} \tag{5.2}$$

and $\vec{\sigma} = (\sigma^1, \ldots, \sigma^n)$ is the replicated spin variable. The calculations which lead to Eq. (5.2) are completely analogous to the ones of Refs. [17,19]. To be self-contained we shall sketch them in Appendix C. The free energy $f(\beta)$ is obtained by taking the saddle point of the integral (5.1) (let say $\lambda = \lambda_n^*$, $\widehat{\lambda} = \widehat{\lambda}_n^*$) and evaluating the $n \to 0$ limit: $\beta f(\beta) = \lim_{n\to 0} \partial_n S[\lambda_n^*, \widehat{\lambda}_n^*]$. The saddle point equations are

$$\widehat{\lambda}(\vec{\sigma}) = \sum_{\vec{\sigma}_1,\ldots,\vec{\sigma}_{k-1}} \lambda(\vec{\sigma}_1)\cdot\ldots\cdot\lambda(\vec{\sigma}_{k-1}) \prod_{a=1}^{n} \delta[\sigma^a \sigma_1^a \ldots \sigma_{k-1}^a, +1]\,, \tag{5.3}$$

$$\lambda(\vec{\sigma}) = \frac{\widehat{\lambda}(\vec{\sigma})^{l-1} \langle e^{\beta h \sum_a \sigma^a} \rangle_h}{\sum_{\vec{\sigma}} \widehat{\lambda}(\vec{\sigma})^l \langle e^{\beta h \sum_a \sigma^a} \rangle_h}\,. \tag{5.4}$$

The above equations are satisfied by the totally ordered solution $\lambda_0(\vec{\sigma}) = \widehat{\lambda}_0(\vec{\sigma}) = \delta_{\vec{\sigma},\vec{\sigma}_0}$, where $\vec{\sigma}_0 = (+1, \ldots, +1)$. The corresponding free energy is $f_F(\beta) = -\langle h \rangle_h$. Such a solution is is possible because of the infinite-strength ferromagnetic interactions in our model (2.3). Physically it is related to the configuration $\{\sigma_i = +1\}_{i=1,\ldots,N}$, which satisfies all the constraints[3].

5.1 Stability of the ferromagnetic phase

In the ferromagnetic solution found above (as in the ferromagnetic phase found in Sec. 4) the system is completely ordered (i.e. the magnetization is $m = 1$). This correspond to no-error communication in the coding language. Knowing the boundaries of the ferromagnetic phase is therefore of great practical relevance. Here we shall investigate the issue of local stability. The calculation is similar (although much simpler) to the one carried out for turbo codes in Ref. [12].

We start by computing the replicated action (5.2) for $\lambda(\vec{\sigma})$, $\widehat{\lambda}(\vec{\sigma})$ "near" the ferromagnetic saddle point, namely $\lambda(\vec{\sigma}) = \lambda_0(\vec{\sigma}) + \delta(\vec{\sigma})$, $\widehat{\lambda}(\vec{\sigma}) = \widehat{\lambda}_0(\vec{\sigma}) + \widehat{\delta}(\vec{\sigma})$. We first consider the case $l > 2$:

$$\delta S[\lambda_0, \widehat{\lambda}_0] = l \sum_{\underline{\sigma}} \delta(\underline{\sigma})\widehat{\delta}(\underline{\sigma}) - \frac{1}{2} l(k-1) \sum_{\underline{\sigma}} \delta(\underline{\sigma})^2 + \frac{1}{2} l\, \widehat{\delta}(\underline{\sigma}_0)^2 + O(\delta^3)\,, \tag{5.5}$$

[3]Notice that, for k even, there are 2^n solutions of the type $\lambda(\vec{\sigma}) = \widehat{\lambda}(\vec{\sigma}) = \delta_{\vec{\sigma},\vec{\tau}}$. The "spurious" solutions with $\vec{\tau} \neq \vec{\sigma}_0$ are related to the $\{\sigma_i = -1\}_{i=1,\ldots,N}$ configuration. Since we took $\langle h \rangle_h > 0$, these solutions do not have thermodynamical relevance.

where $\delta S[\lambda_0, \widehat{\lambda}_0] \equiv S[\lambda_0 + \delta, \widehat{\lambda}_0 + \widehat{\delta}] - S[\lambda_0, \widehat{\lambda}_0]$. It is convenient to integrate over $\lambda(\underline{\sigma})$ using the saddle point equation (5.3), which, for $\lambda(\vec{\sigma}) = \lambda_0(\vec{\sigma}) + \delta(\vec{\sigma})$, $\widehat{\lambda}(\vec{\sigma}) = \widehat{\lambda}_0(\vec{\sigma}) + \widehat{\delta}(\vec{\sigma})$, gives $\delta(\vec{\sigma}) = \widehat{\delta}(\vec{\sigma})/(k-1) + O(\delta^2)$. We finally get

$$\delta S[\widehat{\lambda}_0] = \frac{1}{2} \sum_{\vec{\sigma}} \zeta_{\vec{\sigma}} \widehat{\delta}(\vec{\sigma})^2 + O(\delta^2)\,, \tag{5.6}$$

where $\zeta_{\vec{\sigma}_0} = lk/(k-1)$, and $\zeta_{\vec{\sigma}} = l/(k-1)$ for $\vec{\sigma} \neq \vec{\sigma}_0$. We conclude that, for $l > 2$, the ferromagnetic phase is always locally stable and its boundaries must correspond to first order phase transitions.

For $l = 2$ the situation is physically different. Equation (5.6) is still valid, with $\zeta_{\vec{\sigma}_0} = 2k/(k-1)$ and

$$\zeta_{\vec{\sigma}} = 2 \left[\frac{1}{k-1} - \frac{\langle e^{\beta h \sum_a \sigma^a} \rangle_h}{\langle e^{\beta h n} \rangle_h} \right] \tag{5.7}$$

for $\vec{\sigma} \neq \vec{\sigma}_0$. We have therefore n different eigenvalues $\zeta_{n,\omega}$, with degeneracies $\begin{pmatrix} n \\ \omega \end{pmatrix}$, where $\omega \equiv n - \sum_a \sigma^a$. The first instability occurs for $\omega = 1$. The corresponding critical line is given by $(k-1)\langle e^{-\beta_c h} \rangle_h = 1$. This local stability condition is already known [29] in the coding community, although it has been obtained by completely different methods.

Hereafter we shall focus on the case $l \geq 3$.

5.2 Replica symmetric approximation

The simplest approximation for treating the $n \to 0$ limit, consists in choosing $\lambda(\vec{\sigma})$ and $\widehat{\lambda}(\vec{\sigma})$ to be replica symmetric, i.e. to depend upon $\vec{\sigma}$ uniquely through the symmetric combination $\sum_a \sigma^a$. A commonly adopted parametrization [30] is the following

$$\lambda(\vec{\sigma}) = \int dx\, \pi(x) \frac{e^{\beta x \sum_a \sigma^a}}{(2\cosh \beta x)^n}\,, \tag{5.8}$$

and the analogous one for $\widehat{\lambda}(\vec{\sigma})$ (with a different distribution $\widehat{\pi}(y)$). The replica symmetric order parameters $\pi(x)$ and $\widehat{\pi}(y)$ have the physical meaning of probability distributions of cavity fields. In particular

$$P(H) = \int dx\, \pi(x) \int dy\, \widehat{\pi}(y)\, \delta(H - x - y)\,, \tag{5.9}$$

is the probability distribution of the effective fields $H_i \equiv (1/\beta)\mathrm{arctanh}\langle \sigma_i \rangle$.

Using the ansatz (5.8), we easily obtain the replica symmetric free energy:

$$\begin{aligned} \beta f_P[\pi, \widehat{\pi}] \;=\; & \frac{l}{k} \log 2 - \langle \log \cosh \beta h \rangle_h + l \int dx\, \pi(x) \int dy\, \widehat{\pi}(y)\, \log[1 + t_\beta(x) t_\beta(y)] - \\ & - \frac{l}{k} \int dx_1\, \pi(x_1) \ldots \int dx_k\, \pi(x_k) \log[1 + t_\beta(x_1) \ldots t_\beta(x_k)] - \\ & - \int dy_1\, \widehat{\pi}(y_1) \ldots \int dy_l\, \widehat{\pi}(y_l) \langle \log \mathbb{F}_l(h, y_1, \ldots, y_l; \beta) \rangle_h\,, \end{aligned} \tag{5.10}$$

where we defined $t_\beta(x) \equiv \tanh \beta x$ and

$$\mathbb{F}_l(y_0, y_1, \ldots, y_l; \beta) \equiv \prod_{i=0}^{l}(1 + t_\beta(y_i)) + \prod_{i=0}^{l}(1 - t_\beta(y_i)) \,. \tag{5.11}$$

The field distributions $\pi(x)$ and $\widehat{\pi}(y)$ are determined by the saddle point equations:

$$\widehat{\pi}(y) = \int dx_1\, \pi(x_1) \ldots \int dx_{k-1}\, \pi(x_{k-1})\, \delta\left[y - \frac{1}{\beta}\mathrm{arctanh}(t_\beta(x_1)\ldots t_\beta(x_{k-1}))\right] , \tag{5.12}$$

$$\pi(x) = \int dy_1\, \widehat{\pi}(y_1) \ldots \int dy_{l-1}\, \pi(y_{l-1}) \langle \delta(x - h - y_1 - \ldots - y_{l-1})\rangle_h \,. \tag{5.13}$$

The above equations can be solved either numerically or in some particular limit. In the next Section we will see that the expansion around the random codeword limit provides rather accurate results.

5.3 One step replica symmetry breaking

To go beyond replica symmetric approximation, one has to divide the n replicas into n/m subgroups of m replicas (with $1 \leq m \leq n$). The order parameters $\lambda(\vec{\sigma})$, and $\widehat{\lambda}(\vec{\sigma})$ depend upon $\vec{\sigma}$ through the n/m variables $\widehat{\sigma}^\alpha \equiv \sum_{a=m(\alpha-1)+1}^{m\alpha} \sigma^a$. As discussed clearly in Refs. [23, 31], in the $n \to 0$ limit the order parameter becomes a functional over a probability space and the calculations becomes rather cumbersome (see Refs. [31, 32] for two viable approaches).

In our case there exists a very simple solution to the saddle point equations (5.3), (5.4) incorporating one step replica symmetry breaking:

$$\lambda(\vec{\sigma}) = \sum_{\{s^\alpha\}} \int dx\, \pi_m(x) \frac{e^{\beta x \sum_{\alpha=1}^{n/m} s^\alpha}}{(2\cosh \beta x)^{n/m}} \prod_{\alpha=1}^{n/m} \prod_{a=(\alpha-1)m+1}^{\alpha m} \delta[\sigma^a, s^\alpha] \,, \tag{5.14}$$

and the analogous one for $\widehat{\lambda}(\vec{\sigma})$ (with a different distribution $\widehat{\pi}_m(y)$). It is easy to see that the above ansatz satisfies the saddle point equations as soon as $\pi_m(x)$, $\widehat{\pi}_m(y)$ are solution of the replica symmetric equations (5.12), (5.13), with the substitution $h \to mh$. The phase described by the solution (5.14) is completely analogous to the spin-glass phase found in the random codeword model. The system is frozen in a large number of "optimal" configurations (with self-overlap $q_{EA} = 1$). The overlap between two such configurations is $q_0 = \int dx\, \pi_m(x) \int dy\, \widehat{\pi}_m(y)\, t_\beta^2(x + y)$.

Such a simple scenario (and the simple solution (5.14)) is possible because the multi-spin interactions of the model (2.5) have infinite-strength. The existence of other replica-symmetry-breaking solutions is an open issue, see Sec. 8. In the next Section we will show that our ansatz gives back the RCM solution, see Sec. 4, in the $k, l \to \infty$ limit.

The free energy of the solution (5.14) is $f_{SG,m}(\beta) = f_P(\beta m)$, see Eq. (5.10), and has to be optimized over m with $0 \leq m \leq 1$. This procedure yields the spin-glass free energy $f_{SG}(\beta) = f_P(\beta_c)$, and $m = \beta_c/\beta$. The critical temperature β_c is given by the marginality condition $\partial_m f_{SG,m}(\beta)|_{m=1} = 0$, which coincides with the zero-entropy condition $\partial_\beta f_P(\beta)|_{\beta=\beta_c} = 0$.

Let us now draw some consequences of our solution (5.14) for the phase diagram of the model. Since both the spin-glass and the ferromagnetic free energies are temperature independent, the ferromagnetic-spin glass phase boundary must stay parallel to the temperature

axis. If, for instance, we consider the binary field distribution (2.6) with $h_0 = \mathrm{arctanh}(1-2p)$, this boundary is simply given by $p = p_c(k,l)$. Moreover we notice that the energy density on the line $\beta = 1$, see Eq. (3.7), is equal to the ferromagnetic free energy. This implies that the entropy vanishes at the ferromagnetic-paramagnetic boundary for $\beta = 1$. Since the paramagnetic-spin glass boundary is determined by the zero entropy condition, this point must be the triple point. In synthesis, the main characteristics of the phase diagram depicted in Fig. 3 remain valid for finite connectivities.

6 Large k, l expansion

Here we show that the replica solution exhibited in the previous Section goes to the random codeword model solution (cf. Sec. 4) when $l, k \to \infty$ at $l/k = 1 - R$ fixed. Moreover we want to stress that this limit can be useful from a quantitative point of view. In fact, the corrections for finite k are exponentially small in k.

Notice that the free energy in the spin glass phase $f_{SG}(\beta)$ is easily obtained from the paramagnetic free energy $f_P(\beta)$. In fact we have $f_{SG}(\beta) = f_P(\beta_c)$, where the freezing temperature β_c is given by the zero-entropy condition $\partial_\beta f_P(\beta) = 0$. Moreover the ferromagnetic free energy is $f_F(\beta) = -\langle h \rangle_h$, and does not depend upon k and l. It is then sufficient to solve Eqs. (5.12), (5.13) for large k, l and evaluate Eq. (5.10) on the solution. The result is $f_P^{(exp)}(\beta)$ (exp stands for "expanded"), and allow to reconstruct the whole phase diagram as explained above.

The expansion is obtained by noticing that the product $t_\beta(x_1) \cdot \ldots \cdot t_\beta(x_{k-1})$ which appears on the right-hand side of Eq. (5.12) is exponentially small in k as long as $\pi(x)$ is supported on finite values of x. We then expand the the right-hand side of Eq. (5.13) for small values of y and plug the result in Eq. (5.12).

The calculations are straightforward. For sake of simplicity we show some consequences for the two-peak field distribution (2.6). We refer to Appendix D for the general results.

In Fig. 4 we report the modified phase diagram for $k = 6$, $l = 3$, as computed using the expansion of Appendix D (cf. Eq. (D.8)) for the paramagnetic free energy. We consider the two-peak distribution (2.6) with $h_0 = \mathrm{arctanh}(1-2p)$. The paramagnetic/spin-glass boundary is obtained by imposing the zero-entropy condition $\partial_\beta f_P^{(exp)}(\beta) = 0$. We set $f_{SG}^{(exp)}(\beta) \equiv f_P^{(exp)}(\beta_c)$. The ferromagnetic spin-glass, and ferromagnetic/paramagnetic boundaries are obtained by imposing $f_F(\beta) = f_{SG}^{(exp)}(\beta)$, and $f_F(\beta) = f_P^{(exp)}(\beta)$.

The triple point is at $\beta = 1$, $p = p_c(k,l)$. As we stressed in Sec. 3, the line $\beta = 1$ is of great practical importance, since it correspond to a widespread decoding procedure (MAP decoding). The critical noise $p_c(k,l)$ has the meaning of the threshold for no-error communication under MAP decoding. Since the ferromagnetic-spin glass phase boundary stays parallel to the temperature axis, $p_c(k,l)$ is also the threshold for any "finite-temperature" decoding [27] for $\beta \geq 1$. We get

$$p_c(k,l) = p_c^0 - \frac{1-R}{4\mathcal{H}'(1-2p_c^0)}(1-2p_c^0)^{2k} + O((1-2p_c^0)^{4k})\,, \tag{6.1}$$

where the function $\mathcal{H}(x)$ has been defined in Eq. 4.8. In the $k, l \to \infty$ limit, we recover the threshold $p_c^0 \equiv p_c(R)$ of the random codeword model, given by the solution of Eq. (4.9). The deviations from the *optimal* properties of the random-codeword model are exponentially small for large k.

Equations (5.12) and (5.13) can be solved numerically by a "population dynamics" algorithm. One represents the distributions $\pi(x)$ and $\widehat{\pi}(y)$ by two populations $\{x_i\}_{i=1,\ldots,\mathcal{L}}$ and

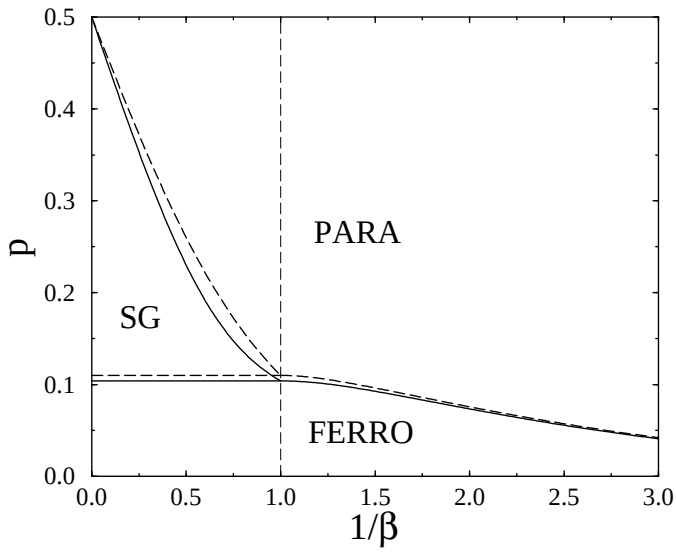

Figure 4: The phase diagram for the $(6, 3)$ code as computed from the large k, l expansion (continuous lines), and the one of the RCM (dashed lines). The vertical dashed line is the Nishimori line $\beta = 1$.

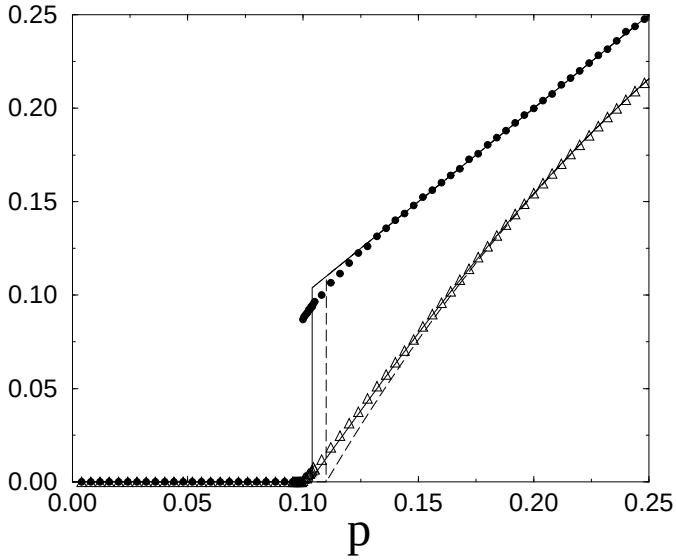

Figure 5: The error probability per bit (filled circles and upper curves), and the entropy (empty triangles and lower curves) for the $(6, 3)$ model with binary field distribution (2.6). We set $\beta = 1$ and $h_0 = \text{arctanh}(1 - 2p)$. The symbols are obtained by solving numerically the saddle point equations (5.12), (5.13). The dashed lines are the RCM results. The continuous lines are the results of the large-connectivity expansion.

$\{y_j\}_{j=1,\ldots,\mathcal{L}}$, and then iterates the equations (5.12) and (5.13). This method has been already used, for instance, in Ref. [31]. In Fig. 5, we consider once again the line $\beta = 1$ and compare the results of large k, l expansion with the numerical solution of Eqs. (5.12) and (5.13). We plot both the entropy and the average error probability per bit $\langle P_e \rangle_{h,\mathbb{C}}$, where:

$$P_e = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2}(1 - \mathrm{sign}\langle \sigma_i \rangle)\,, \tag{6.2}$$

As conclusion let us consider the problem of calculating the critical noise $p_c(k,l)$. This can be obtained either by solving numerically Eqs. (5.12) and (5.13), or from the expansion (6.1). The numerical solution yields $p_c(k,l) = 0.0997(2)$, $0.1071(2)$, $0.1091(2)$, for, respectively, $(k,l) = (6,3)$, $(8,4)$, $(10,5)$. From the expansion (6.1) we get $p_c^{exp}(k,l) \approx 0.103965$, 0.107783, 0.109195 for the same values of k and l.

7 Finite size corrections and numerical results

In this Section we compare the analytical predictions with numerical results in order to confirm the validity of the former and to investigate the nature of finite size corrections. Needless to say, the last one is a point of utmost practical importance in coding theory. Indeed it is known that the thermodynamic limit is approached exponentially fast in the ferromagnetic phase, at zero temperature [2]. We expect the same behavior to hold in the whole ferromagnetic phase.

Here we focus on the paramagnetic-spin glass phase transition. We compute the finite size corrections to the free energy of the RCM. This calculation is compared with exact enumeration calculations on small systems. Then we switch to the complete model (2.5) and compare the the numerical results with the outcome of the replica calculations, cf. Sec. 5.

7.1 The random codeword model

Let us consider, for sake of clarity, the binary distribution (2.6) with $p > p_c(R)$. This corresponds to focusing on the paramagnetic-spin glass phase transition. Under this condition the ordered state $\alpha = 0$ belongs to the continuous part of the spectrum and there is no energy gap. We shall therefore neglect this state. Its contribution is exponentially small in the thermodynamic limit.

With this assumption, we obtain the following result for the free energy density

$$f(\beta, N) = f_0(\beta) + \frac{1}{N} f_1(\beta, N) + O(1/N^2)\,, \tag{7.1}$$

The leading term has been already computed in Sec. 4. The first correction $f_1(\beta, N)$ vanishes in the paramagnetic phase and depends weakly upon N. Explicit formulae are given in Appendix E. In particular $f_1(\beta, N) \sim (1/2\beta_c) \log N$ as $N \to \infty$. The leading correction in the paramagnetic phase is exponentially small in N. In order to compute it, the ferromagnetic state cannot be neglected.

It is very easy to compute numerically the finite-N free energy for the random codeword model with binary field distribution (2.6), as long as we neglect the ordered state. All we need, for a given sample, is the energy spectrum. Let us call ν_k, with $k = 0, \ldots, N$ the number of states α, such that $E^{(\alpha)} = -h_0(N - 2k)$. The probability distribution of the spectrum $\{\nu_k\}$ is

$$P(\{\nu_k\}) = \frac{\mathcal{N}!}{\prod_{k=0}^{N} \nu_k!} \prod_{k=0}^{N} p_k^{\nu_k}\,, \tag{7.2}$$

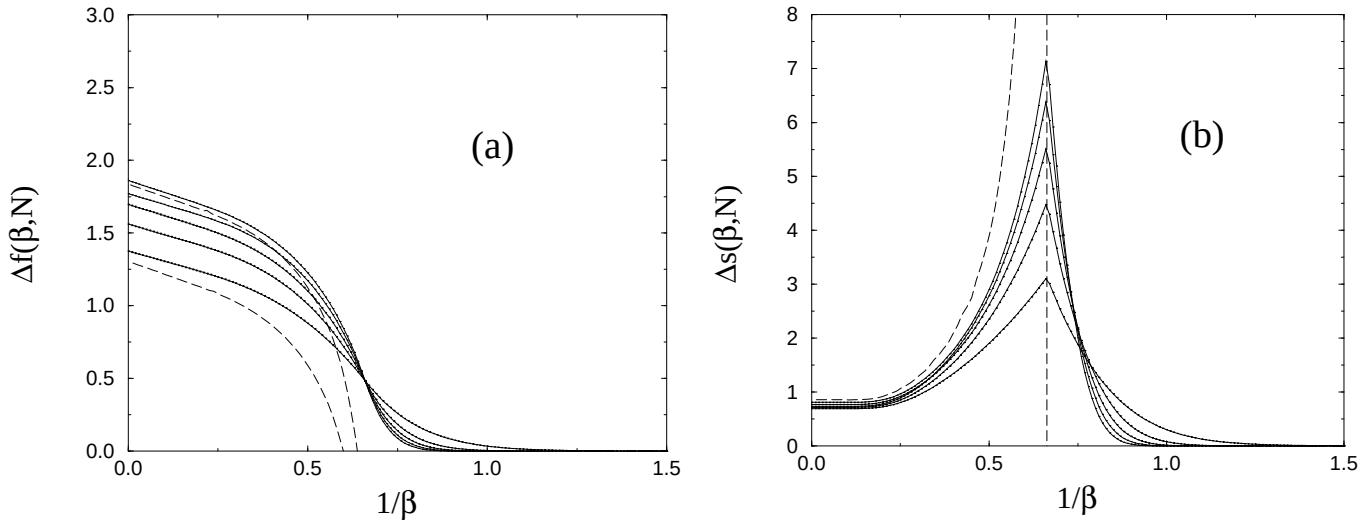

Figure 6: Finite size correction to the free energy (a) and to the entropy (b) of the RCM. The continuous lines are the results of numerical computations for $N = 40, 80, 120, 160, 200$ (error bars are not visible on this scale). The dashed lines are the analytical results for the leading finite size correction, for $N = 40, 200$ (a) and $N = 200$ (b).

where $\sum_k \nu_k = \mathcal{N} \equiv 2^{NR}$, and

$$p_k \equiv \frac{1}{2^N} \left(\begin{array}{c} N \\ k \end{array} \right) . \tag{7.3}$$

Once the $\{\nu_k\}$ have been generated with probability distribution (7.2), the partition function is given by $Z(\beta) = \sum_k \nu_k \exp\{\beta h_0(N - 2k)\}$.

We considered the RCM with rate $R = 1/2$ and binary field distribution (2.6) with $h_0 = \text{arctanh}(1 - 2p)$. The phase diagram of this model is depicted in Fig. 3. We fixed the flip probability $p = 0.2$ to be greater than the threshold $p_c(1/2) \approx 0.110025$, and computed the temperature dependence of the free energy by averaging over 10^5 realizations of the spectrum $\{\nu_k\}$.

In Fig. 6, graph (a), we plot the quantity $\Delta f(\beta, N) \equiv [f(\beta, N) - f_0(\beta)]N$, together with the theoretical prediction $f_1(\beta, N)$ for several values of N. In Fig. 6, graph (b), we consider the entropy density $s(\beta, N) \equiv \beta^2 \partial_\beta f(\beta, N)$: we plot the difference $\Delta s(\beta, N) \equiv [s(\beta, N) - s_0(\beta)]N$, for the same values of N, together with $s_1(\beta, N) \equiv \beta^2 \partial_\beta f_1(\beta, N)$ for $N = 200$ (the N dependence of $s_1(\beta, N)$ is rather weak).

Two remarks can be made by looking at Fig, 6. First, the $O(1/N^2)$ terms in Eq. (7.1) seems to be rather small. If the temperature is not too close to the critical point, the finite size corrections are well described by $f_1(\beta, N)$. Second, the curves for $\Delta f(\beta, N)$, see Fig. 6, graph (a), seem to cross at the critical point. This is expected since $\Delta f(\beta, N) \sim (1/2\beta_c) \log N$ for $\beta > \beta_c$, and $\Delta f(\beta, N) \sim e^{-\kappa N}$ for $\beta < \beta_c$. The crossing point $\beta_{N,N'}$ between the curves $\Delta f(\beta, N)$ and $\Delta f(\beta, N')$ can be used to estimate β_c. From the data of Fig. 6 we get

$$\beta_{40,80} = 1.52(1)\,, \quad \beta_{80,120} = 1.51(1)\,, \quad \beta_{120,160} = 1.51(1)\,, \quad \beta_{160,200} = 1.51(1)\,, \tag{7.4}$$

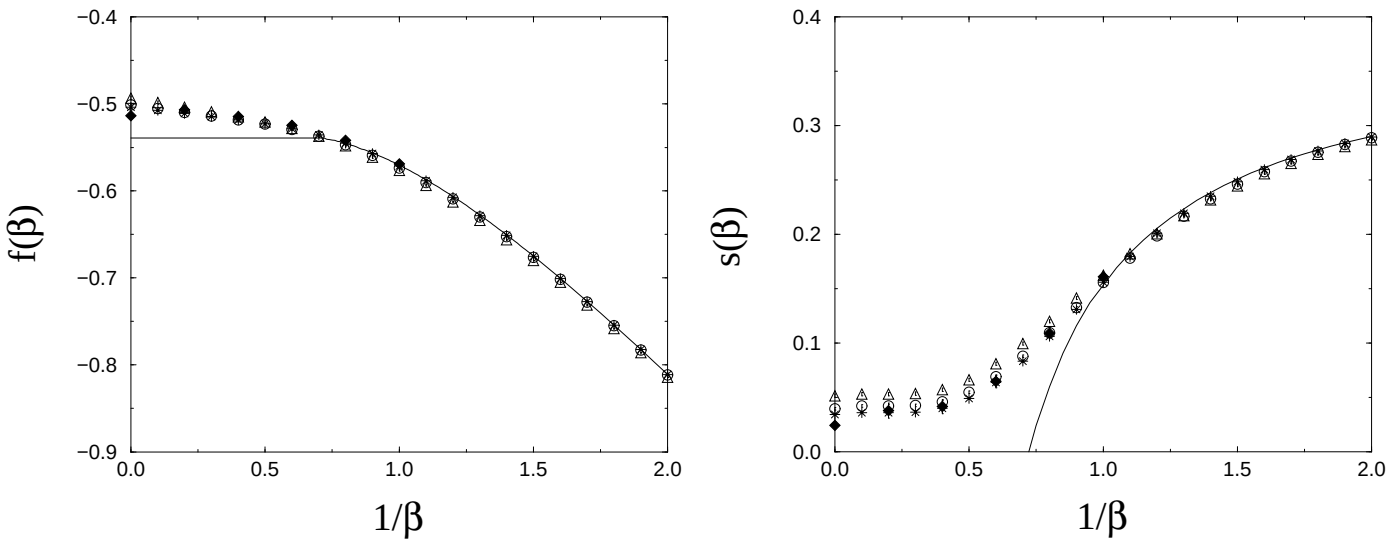

Figure 7: The free energy (left) and the entropy (right) of the $(6,3)$ model computed by exact-enumeration (symbols), and the corresponding theoretical predictions (continuous lines). The various symbols refer to different system sizes: $N = 20$ (triangles), 30 (circles), 40 (stars) and 50 (filled diamonds).

which is in good agreement with the exact result $\beta_c \approx 1.50794$.

7.2 The $(6,3)$ model

In this case we are forced to consider quite small systems since we do not know any simple form for the probability distribution of the energy spectrum. We must enumerate all the codewords (i.e. the spin configurations which satisfy the constraints in Eq. (2.5)): this takes at least $O(2^{NR})$ operations. Notice that *finding* the codewords is a simple task. It suffices to solve the linear system $\mathbb{C}\underline{x} = 0$ (mod2). A standard method (we used gaussian elimination) takes $O(N^3)$ operations [33].

As in the previous Subsection, we fixed considered the binary field distribution (2.6) with $h_0 = \text{arctanh}(1-2p)$, and $p = 0.2$. In Fig. 7 we plot the results for the free energy and the entropy densities for systems of size $N = 20, 30, 40$ (averaged over $N_{stat} = 1000$ samples) and $N = 50$ (with $N_{stat} = 20$ samples). The numerical results converge quite well to the theoretical calculation at high temperature. Below the critical temperature the convergence is very slow, as expected from the analogy with the RCM example.

The sizes considered here are too small to reach any definite conclusion on the glassy phase.

8 Discussion

The main result of this paper is the determination of the phase diagram of regular Gallager codes, see Eq. (2.5). This is depicted in Fig. 3 for the infinite connectivity limit. The phase diagram for finite connectivities has been obtained by resorting to the replica method and looks

qualitatively similar. The most important quantitative difference is the critical noise level for the ferromagnetic-spin glass phase transition. This quantity determines the performances of the corresponding code. It can be determined either by solving the mean field equations numerically, see Sec. 5, or in a large connectivity expansion, see Sec. 6. The result of the last computation is reported in Fig. 4.

The replica computation was made possible by the particularly simple one-step replica symmetry breaking solution exhibited in Eq. (5.14). We weren't able to prove that the saddle point (5.14) is either unique or the dominant one. There are however several independent indications which confirm this conclusion:

- The proposed solution is consistent with the absence of replica symmetry breaking on the $\beta = 1$ line, which has been proved in Sec. 3.
- It has been shown [19,34] that the critical noise level is the same both for zero-temperature and for temperature one decoding. This implies that the ferromagnetic-spin glass phase boundary must pass through the points $(p = p_c(k,l), 1/\beta = 0)$, and $(p = p_c(k,l), 1/\beta = 1)$, see Fig. 4 (for sake of simplicity we referred to the case of a binary field distribution). This consistent with our phase diagram.
- Our numerical results, although we restricted to fairly small systems, do not contradict our conclusions.

It can be interesting to notice that recently [35] a "factorized ansatz" has been proposed as an exact one-step replica symmetry breaking solution for some diluted spin models. The solution used in this paper is, in some sense, complementary to the one of Ref. [35].

Acknowledgments

I am grateful to B. Derrida for an illuminating discussion on the random codeword model, and to N. Sourlas for his constant support and encouragement. I thank M. Mézard and G. Parisi for their interest in the subject of this paper. This work was supported through a European Community Marie Curie Fellowship.

A Codewords in the $k, l \to \infty$ limit

In this Appendix we compute the one-codeword, and two-codeword probabilities, see Eqs. (4.1) and (4.4), for generic values of k and l. Then we show that, in the $k, l \to \infty$ limit, different codewords become statistically independent, i.e. $P_{\underline{\sigma},\underline{\tau}} \sim P_{\underline{\sigma}} P_{\underline{\tau}}$.

The one-codeword probability is, to the leading exponential order:

$$P_{\underline{\sigma}} \sim \int \prod_{\sigma} d\lambda(\sigma) d\widehat{\lambda}(\sigma) \, \exp\{N A_1(\lambda, \widehat{\lambda}; c)\} \,, \tag{A.1}$$

where

$$\begin{aligned} A_1(\lambda, \widehat{\lambda}; c) \;=\; & -l \sum_{\sigma} \lambda(\sigma) \widehat{\lambda}(\sigma) + \frac{l}{2k} \left[\left(\sum_{\sigma} \lambda(\sigma) \right)^k + \left(\sum_{\sigma} \lambda(\sigma) \sigma \right)^k \right] + \\ & + l \sum_{\sigma} c(\sigma) \log \widehat{\lambda}(\sigma) + l - \frac{l}{k} \,, \end{aligned} \tag{A.2}$$

and $c(\sigma) = (1/N)\sum_i \delta_{\sigma,\sigma_i}$ characterizes the configuration $\underline{\sigma}$. The above result can be proved by noticing that $\sum_{\underline{\sigma}} P_{\underline{\sigma}} \exp(\beta h_0 \sum_i \sigma_i) = \langle Z(h_0)\rangle_{\mathbb{C}}$, where $Z(h_0)$ is the partition function for the model (2.5) with uniform magnetic field $h_i = h_0$. The average $\langle Z(h_0)\rangle_{\mathbb{C}}$ is easily obtained from Eqs. (5.1) and (5.2) by setting $n = 1$ and $p_h(h_i) = \delta(h_i - h_0)$.

The integral (A.1) can be done through the saddle point method. Saddle point equations are more conveniently written by eliminating $\widehat{\lambda}(\underline{\sigma})$, and using the variables $\lambda_+ \equiv \sum_\sigma \lambda(\sigma)$ and $\lambda_- \equiv \sum_\sigma \lambda(\sigma)\sigma$. We get:

$$\lambda_+^k + \lambda_-^k = 2\,, \tag{A.3}$$

$$\lambda_-\lambda_+^{k-1} + \lambda_+\lambda_-^{k-1} = 2m\,, \tag{A.4}$$

where $m = \sum_\sigma c(\sigma)\sigma = (1/N)\sum_i \sigma_i$. For large k, these equations imply $\lambda_+ = 2^{1/k} + O(m^k)$, $\lambda_- = 2^{1/k}m + O(m^k)$, as soon as $-1 < m < 1$. Substituting in Eq. (A.2), we get the result anticipated in Sec. 4, see Eqs. (4.2), (4.3).

Let us now consider the two-codeword probability, cf. Eq. (4.4). Analogously to Eq. (A.1) we get:

$$P_{\underline{\sigma},\underline{\tau}} \sim \int \prod_{\sigma,\tau} d\lambda(\sigma,\tau) d\widehat{\lambda}(\sigma,\tau)\, \exp\{N A_2(\lambda,\widehat{\lambda};c)\}\,. \tag{A.5}$$

The corresponding "action" is

$$\begin{aligned} A_2(\lambda,\widehat{\lambda};c) &= -l\sum_{\sigma,\tau}\lambda(\sigma,\tau)\widehat{\lambda}(\sigma,\tau) + \frac{l}{k}\sum_{\sigma_1\ldots\sigma_k}{}' \sum_{\tau_1\ldots\tau_k}{}' \lambda(\sigma_1,\tau_1)\ldots\lambda(\sigma_k,\tau_k) + \\ &\quad + l\sum_{\sigma,\tau} c(\sigma,\tau)\log\widehat{\lambda}(\sigma,\tau) + l - \frac{l}{k}\,, \end{aligned} \tag{A.6}$$

where $c(\sigma,\tau) = (1/N)\sum_i \delta_{\sigma_i,\sigma}\delta_{\tau_i,\tau}$, and the sums $\sum'$ are restricted to $\sigma_1\cdots\sigma_k = +1$ and $\tau_1\cdots\tau_k = +1$. As before we notice that $\sum_{\underline{\sigma},\underline{\tau}} P_{\underline{\sigma},\underline{\tau}} \exp(\beta h_1 \sum_i \sigma_i + \beta h_2 \sum_i \tau_i) = \langle Z(h_1)Z(h_2)\rangle_{\mathbb{C}}$ can be obtained through a standard replica calculation, see Sec. 5 and App. C, with $n = 2$ replicas.

We now define the variables $\lambda_0 \equiv \sum_{\sigma,\tau}\lambda(\sigma,\tau)$, $\lambda_\sigma \equiv \sum_{\sigma,\tau}\lambda(\sigma,\tau)\sigma$, $\lambda_\tau \equiv \sum_{\sigma,\tau}\lambda(\sigma,\tau)\tau$, and $\lambda_{\sigma\tau} \equiv \sum_{\sigma,\tau}\lambda(\sigma,\tau)\sigma\tau$. The saddle point equations can be written in terms of these variables as follows:

$$\lambda_0^k + \lambda_\sigma^k + \lambda_\tau^k + \lambda_{\sigma\tau}^k = 4\,, \tag{A.7}$$

$$\lambda_\sigma\lambda_0^{k-1} + \lambda_0\lambda_\sigma^{k-1} + \lambda_{\sigma\tau}\lambda_\tau^{k-1} + \lambda_\tau\lambda_{\sigma\tau}^{k-1} = 4m_\sigma\,, \tag{A.8}$$

$$\lambda_\tau\lambda_0^{k-1} + \lambda_{\sigma\tau}\lambda_\sigma^{k-1} + \lambda_0\lambda_\tau^{k-1} + \lambda_\sigma\lambda_{\sigma\tau}^{k-1} = 4m_\tau\,, \tag{A.9}$$

$$\lambda_{\sigma\tau}\lambda_0^{k-1} + \lambda_\tau\lambda_\sigma^{k-1} + \lambda_\sigma\lambda_\tau^{k-1} + \lambda_0\lambda_{\sigma\tau}^{k-1} = 4q\,, \tag{A.10}$$

where $m_\sigma = \sum_{\sigma,\tau} c(\sigma,\tau)\sigma = (1/N)\sum_i \sigma_i$, $m_\tau = \sum_{\sigma,\tau} c(\sigma,\tau)\tau = (1/N)\sum_i \tau_i$, and $q = \sum_{\sigma,\tau} c(\sigma,\tau)\sigma\tau = (1/N)\sum_i \sigma_i\tau_i$. From Eqs. (A.7)-(A.10), we get, for $k \to \infty$, $\lambda_0 \simeq 4^{1/k}$, $\lambda_\sigma \simeq 4^{(1-k)/k}m_\sigma$, $\lambda_\tau \simeq 4^{(1-k)/k}m_\tau$, $\lambda_{\sigma\tau} \simeq 4^{(1-k)/k}q$, as soon as $-1 < m_\sigma, m_\tau, q < 1$. The corrections to this asymptotic behavior are of order $O(m_\sigma^k, m_\tau^k, q^k)$. Substituting this solution in Eqs. (A.5), (A.6), we get the results (4.5), (4.6).

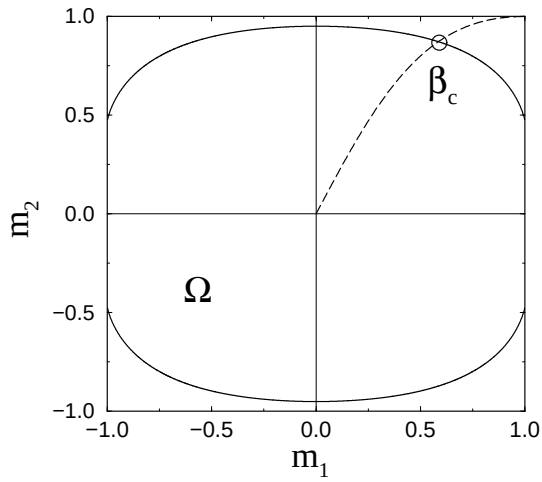

Figure 8: The RCM for $p_h(h_i) = (2/5)\,\delta(h_i - 1/2) + (3/5)\,\delta(h_i - 1)$. The continuous line encircles the region Ω (see text). The dashed line is the curve $m_1 = \tanh\beta/2$, $m_2 = \tanh\beta$, which intersect the boundary of Ω for $\beta = \beta_c$.

B The random codeword model for a generic field distribution

In this Appendix we solve[4] the RCM for a generic field distribution $p_h(h_i)$. The strategy is to start from a discrete distribution

$$p_h(h_i) = \sum_{q=1}^{\mathcal{M}} p_q\,\delta(h_i - h^{(q)})\,, \tag{B.1}$$

and then approximate a generic $p_h(h_i)$ by letting $\mathcal{M} \to \infty$.

Let us consider the distribution (B.1). In the typical sample there will be $N_1 \approx Np_1$ sites with field $h_i = h^{(1)}$ (which we can suppose, without loss of generality, to be the sites $i = 1, \ldots, N_1$), $N_2 \approx Np_2$ sites with field $h_i = h^{(2)}$ (let us say for $i = N_1 + 1, \ldots, N_1 + N_2$), and so on. For a given spin configuration $\underline{\sigma}$, we define the partial magnetization $m_q(\underline{\sigma})$ as the magnetization of the sites whose magnetic field is $h^{(q)}$. With the labeling of the sites chosen above we get

$$m_q(\underline{\sigma}) \equiv \frac{1}{N_q} \sum_{i=\mathcal{N}_{q-1}+1}^{\mathcal{N}_q} \sigma_i\,, \tag{B.2}$$

where $\mathcal{N}_q = N_1 + \ldots + N_q$. We call $\{m_q(\underline{\sigma})\}$ the *magnetization profile* of the configuration $\underline{\sigma}$.

We now consider the 2^{NR} states $\alpha = 1, \ldots, 2^{NR}$. To each of them it is associated a random codeword $\underline{\sigma}^{(\alpha)}$, where the $\sigma_i^{(\alpha)}$ are quenched variables drawn with flat probability distribution. We ask ourselves what is the typical number $\mathcal{N}_{typ}(\{m_q\})$ of states α having a

[4] I am deeply indebted with B. Derrida who explained to me how to treat this general case.

given magnetization profile $m_q(\underline{\sigma}^{(\alpha)}) = m_q$. The answer is quite easy. Define the function $\mathcal{G}(\{m_q\})$ as follows

$$\mathcal{G}(\{m_q\}) = R\log 2 + \sum_{q=1}^{M} p_q \mathcal{H}(m_q)\,, \tag{B.3}$$

where $\mathcal{H}(x)$ is given in Eq. (4.8). The typical number $\mathcal{N}_{typ}(\{m_q\})$ is obtained from $\mathcal{G}(\{m_q\})$ through the usual construction: $\mathcal{N}_{typ}(\{m_q\}) \sim \exp[N\mathcal{G}(\{m_q\})]$ if $\mathcal{G}(\{m_q\}) > 0$ and $\mathcal{N}_{typ}(\{m_q\}) = 0$ otherwise. The convex region $\Omega \equiv \{\{m_q\}|\mathcal{G}(\{m_q\}) > 0\}$ is depicted in Fig. 8 for the case $\mathcal{M} = 2$.

The energy of a state α can be written in terms of its magnetization profile: $E^{(\alpha)} = -N\sum_q p_q h^{(q)} m_q(\underline{\sigma}^{(\alpha)})$. The free energy density can therefore computed from $\mathcal{N}_{typ}(\{m_q\})$ as follows:

$$f(\beta) = \min_{\{m_q\}} \left\{ -\frac{1}{\beta}\widehat{\mathcal{G}}(\{m_q\}) - \sum_{q=1}^{M} p_q h_q m_q \right\}\,, \tag{B.4}$$

where $\widehat{\mathcal{G}}(\{m_q\}) \equiv (1/N)\log \mathcal{N}_{typ}(\{m_q\})$ (i.e. $\widehat{\mathcal{G}}(\{m_q\}) = \mathcal{G}(\{m_q\})$ inside Ω, and $\widehat{\mathcal{G}}(\{m_q\}) = -\infty$ outside).

If the expression (B.3) is used in Eq. (B.4), one gets the saddle point condition $m_q = \tanh\beta h_q$. This describes a curve in the $\{m_q\}$ space which start at $m_q = 0$ for $\beta = 0$, and ends at $m_q = \operatorname{sign} h_q$ for $\beta = \infty$. The corresponding free energy reads

$$f_P(\beta) = -\frac{R}{\beta}\log 2 - \frac{1}{\beta}\sum_{q=1}^{M} p_q \log\cosh\beta h_q\,. \tag{B.5}$$

At some critical temperature $\beta = \beta_c$ the curve $m_q = \tanh\beta h_q$ crosses the boundary of Ω. The saddle point $m_q = \tanh\beta h_q$ is no longer valid for $\beta > \beta_c$. The critical temperature can be computed from the zero entropy condition $\partial_\beta f_P|_{\beta=\beta_c} = 0$. For $\beta > \beta_c$ the entropy vanishes and the free energy is frozen to its value at the critical point: $f_{SG}(\beta) = f_P(\beta_c)$. As in Sec. 4, we must include in our analysis the ordered state $\alpha = 0$ whose free energy is $f_F(\beta) = -\langle h\rangle_h$.

The solution for a continuous field distribution $p_h(h_i)$ follows from the above results by taking the $\mathcal{M} \to \infty$ limit in Eq. (B.5). This yields Eq. (4.15). Alternatively we could have started with a continuous magnetization profile $m(h)$ from the very beginning of this Appendix.

C The derivation of Eq. (5.2)

We start by writing down the partition function of the model (2.5):

$$Z(\beta) = \sum_{\underline{\sigma}} \prod_{j=1}^{M} \delta[\sigma^{\omega_j}, +1]\, e^{\sum_i h_i\sigma_i}\,. \tag{C.1}$$

We rewrite the constraint term (i.e. the product of Kronecker delta functions) by introducing the quenched variables $D_\omega = 0, 1$, where $\omega = (i_1^\omega, \ldots, {}_k^\omega)$ runs over the k-plets of site indices.

The variables D_ω are defined by setting $D_\omega = 1$ if $\omega = \omega_j$ for some $j = 1, \ldots, M$ and $D_\omega = 0$ otherwise. With this definition we can write the replicated partition function as follows

$$\langle Z^n \rangle = \frac{1}{\mathcal{N}} \sum_{\{D\}} \sum_{\{\vec{\sigma}\}} \prod_{i=1}^{N} \left\langle e^{\beta h \sum_a \sigma_i^a} \right\rangle_h \prod_\omega \{1 - D_\omega + D_\omega \delta_n[\vec{\sigma}^\omega]\} \,, \tag{C.2}$$

where $\vec{\sigma}^\omega \equiv (\prod_{r=1}^k \sigma_{i_r^\omega}^1, \ldots, \prod_{r=1}^k \sigma_{i_r^\omega}^n)$, $\delta_n[\vec{\sigma}] \equiv \prod_{a=1}^n \delta[\sigma^a, +1]$, and $\mathcal{N}$ is a normalization constant (to be computed later).

According to our choice of the *ensemble* of check matrices, we must impose $\sum_{\omega \ni i} D_\omega = l$, for any $i = 1, \ldots, N$. This can be done by using the identity

$$\delta\left[\sum_{\omega \ni i} D_\omega, l\right] = \oint \frac{dz_i}{2\pi i} \frac{1}{z_i^{l+1}} z_i^{\sum_{\omega \ni i} D_\omega} \,, \tag{C.3}$$

where the integration path encircles the origin in the complex z_i plane. We get

$$\langle Z^n \rangle = \frac{1}{\mathcal{N}'} \sum_{\{\vec{\sigma}\}} \prod_{i=1}^{N} \oint \frac{dz_i}{2\pi i} \frac{1}{z_i^{l+1}} \left\langle e^{\beta h \sum_a \sigma_i^a} \right\rangle_h \prod_\omega \sum_{D_\omega = 0}^{1} w(D_\omega) \{1 - D_\omega + D_\omega \delta_n[\vec{\sigma}^\omega]\} \, z_\omega^{D_\omega} \,, \tag{C.4}$$

where $z_\omega \equiv \prod_{i \in \omega} z_i$. The weights $w(D_\omega)$ have been introduced for later convenience, and correspond to a rescaling of the $\{z_i\}$. Their contribution can be readsorbed by the normalization constant $\mathcal{N}'$. We set $w(1) = l(k-1)!/N^{k-1}$ and $w(0) = 1 - w(1)$. Now we can sum over the D_ω, obtaining

$$\begin{aligned} \langle Z^n \rangle \;=\; & \frac{1}{\mathcal{N}''} \sum_{\{\vec{\sigma}\}} \prod_{i=1}^{N} \oint \frac{dz_i}{2\pi i} \frac{1}{z_i^{l+1}} \left\langle e^{\beta h \sum_a \sigma_i^a} \right\rangle_h \cdot \\ & \cdot \exp\left\{ \frac{Nl}{k} \sum_{\vec{\sigma}_1, \ldots, \vec{\sigma}_k} c_z(\vec{\sigma}_1) \ldots c_z(\vec{\sigma}_k) \prod_{a=1}^{n} \delta[\sigma_1^a \ldots \sigma_k^a, +1] \right\} \,, \end{aligned} \tag{C.5}$$

where $c_z(\vec{\sigma}) \equiv (1/N) \sum_i z_i \delta_{\vec{\sigma}, \vec{\sigma}_i}$. Finally we introduce the order parameter $\lambda(\vec{\sigma})$ and its complex conjugate $\widehat{\lambda}(\vec{\sigma})$, by using the following identity

$$\begin{aligned} \exp\{N\mathcal{F}[c]\} \;=\; & \int \prod_{\vec{\sigma}} \frac{Nl}{\pi} d\lambda(\vec{\sigma}) d\widehat{\lambda}(\vec{\sigma}) \, \exp\left\{ -Nl \sum_{\vec{\sigma}} \lambda(\vec{\sigma}) \widehat{\lambda}(\vec{\sigma}) + \right. \\ & \left. + N\mathcal{F}[\lambda] + Nl \sum_{\vec{\sigma}} \widehat{\lambda}(\vec{\sigma}) c_z(\vec{\sigma}) \right\} \,. \end{aligned} \tag{C.6}$$

The use of the above identity allows to integrate over the $\{z_i\}$, obtaining Eqs. (5.1) and (5.2). The overall normalization constant can be fixed by requiring $\langle Z^n \rangle \sim 2^{Nn(1-l/k)}$ for $h_i = 0$.

D Large k, l expansion: general formulae

Let us define $t_p \equiv \langle \tanh \beta h \rangle_h$. We assume formally $t_p = O(t^p)$ where t is "small" and expand in t^k to the order t^{3k}. All the observables can be expressed in terms of the order parameters

$\pi(x)$ and $\widehat{\pi}(y)$. The solutions of Eqs. (5.12), (5.13) admit an expansion of the form

$$\pi(x) = p_h(x) + \sum_{m=1}^{\infty} \pi_m \beta^{-m} p_h^{(m)}(x) \;\; ; \qquad \widehat{\pi}(y) = \delta(y) + \sum_{n=1}^{\infty} \widehat{\pi}_n \beta^{-n} \delta^{(n)}(y)\,, \tag{D.1}$$

where $p_h^{(m)}(x) \equiv \partial_x^m p_h(x)$ and $\delta^{(n)}(y) = \partial_y^n \delta(y)$. Moreover one gets $\pi_m, \widehat{\pi}_m = O(t^{mk})$. The results for the first few coefficients are listed below:

$$\begin{aligned}
\pi_1 &= -(l-1)t_1^{k-1} - (k-1)(l-1)^2(1-t_2)t_1^{2k-3} - \qquad (D.2)\\
&\quad -\frac{1}{3}(l-1)t_3^{k-1} - \frac{1}{2}(k-1)(k-2)(l-1)^3(1-t_2)^2 t_1^{3k-5} - (k-1)^2(l-1)^3(1-t_2)^2 t_1^{3k-5} + \\
&\quad +(k-1)(l-1)^2(t_1-t_3)t_2^{k-1}t_1^{k-2} + (k-1)(l-1)^2(l-2)(t_1-t_3)t_1^{3k-4} + O(t^{4k})\,,\\
\pi_2 &= \frac{1}{2}(l-1)t_2^{k-1} + \frac{1}{2}(l-1)(l-2)t_1^{2k-2} + \qquad (D.3)\\
&\quad +(k-1)(l-1)^2(t_1-t_3)t_2^{k-2}t_1^{k-1} + (k-1)(l-1)^2(l-2)(1-t_2)t_1^{3k-4} + O(t^{4k})\,,\\
\pi_3 &= -\frac{1}{6}(l-1)t_3^{k-1} - \frac{1}{2}(l-1)(l-2)t_2^{k-1}t_1^{k-1} - \frac{1}{6}(l-1)(l-2)(l-3)t_1^{3k-3} + O(t^{4k})\,, \qquad (D.4)
\end{aligned}$$

$$\begin{aligned}
\widehat{\pi}_1 &= -t_1^{k-1} - (k-1)(l-1)(1-t_2)t_1^{2k-3} - \qquad (D.5)\\
&\quad -\frac{1}{2}(k-1)(k-2)(l-1)^2(1-t_2)^2 t_1^{3k-5} - (k-1)^2(l-1)^2(1-t_2)^2 t_1^{3k-5} + \\
&\quad +(k-1)(l-1)(t_1-t_3)t_2^{k-1}t_1^{k-2} + (k-1)(l-1)(l-2)(t_! - t_3)t_1^{3k-4} - \frac{1}{3}t_3^{k-1} + O(t^{4k})\,,\\
\widehat{\pi}_2 &= \frac{1}{2}t_2^{k-1} + (k-1)(l-1)(t_1-t_3)t_2^{k-2}t_1^{k-1} + O(t^{4k}) \qquad (D.6)\\
\widehat{\pi}_3 &= -\frac{1}{6}t_3^{k-1} + O(t^{4k})\,. \qquad (D.7)
\end{aligned}$$

The result for the paramagnetic free energy is

$$\begin{aligned}
\beta f_P(\beta) &= -R\log 2 - \langle \log\cosh\beta h\rangle_h - \frac{l}{k}t_1^k - \frac{1}{2}l(l-1)(1-t_2)t_1^{2k-2} + \frac{1}{2}\frac{l}{k}t_2^k - \\
&\quad -\frac{1}{2}(k-1)l(l-1)^2(1-t_2)^2 t_1^{3k-4} + \frac{1}{3}l(l-1)(l-2)(t_1-t_3)t_1^{3k-3} + \qquad (D.8)\\
&\quad +l(l-1)(t_1-t_3)t_1^{k-1}t_2^{k-1} - \frac{1}{3}\frac{l}{k}t_3^k + O(t^{4k})\,.
\end{aligned}$$

E Finite size corrections for the random codeword model

Let us consider the binary field distribution (2.6) with $h_0 = 1$. The results for a generic value of h_0 are obtained after a trivial rescaling of energies and temperatures: $f(\beta, h_0; N) = h_0 f(\beta h_0, 1; N)$.

As explained in Sec. 7, the finite size corrections at the paramagnetic-spin glass phase transition can be studied by neglecting the ordered state. This introduces exponentially small errors. The calculation of the free energy can be done along the lines of Ref. [21], Appendix B, which starts from the identity:

$$\langle \log Z\rangle = \int_0^\infty \frac{dt}{t}\left(e^{-t} - e^{-tZ}\right)\,. \tag{E.1}$$

We limit ourselves to quoting the outcome of the calculation. For $\beta < \beta_c$, we get $f(\beta, N) = f_P(\beta) + O(e^{-\kappa N})$[5]. For $\beta > \beta_c$ we get Eq. (7.1), with

$$f_0(\beta) = -\widehat{\epsilon}(R)\,, \quad f_1(\beta, N) = \int_0^\infty d\phi\, \rho(\phi)\, e^{-\phi} + \gamma/\beta\,, \tag{E.2}$$

$\gamma \approx 0.577216$ being the Euler constant. The function $\rho(\phi)$ is defined as the (unique) solution of

$$\beta_c \rho + \log \Psi(-N\widehat{\epsilon} + \rho) = \log(\phi) + \frac{1}{2} \log \left[\frac{\pi}{2} N (1 - \widehat{\epsilon}^2) \right]\,, \tag{E.3}$$

where $-\widehat{\epsilon}(R)$ is the ground state energy density in the thermodynamic limit, see Sec. 4. The function $\Psi(x)$ is defined as follows

$$\Psi(x) = \sum_{q=-\infty}^{+\infty} e^{-\beta_c(2q+x)} \left[1 - \exp\left(-e^{\beta(2q+x)} \right) \right]\,. \tag{E.4}$$

Notice that $\Psi(x+2) = \Psi(x)$. The $\log \Psi$ term in Eq. (E.3) gives therefore an oscillating N dependence to $f_1(\beta, N)$. Moreover, since $\Psi(-N\widehat{\epsilon} + \rho)$ remains finite for any N and ρ, $f_1(\beta, N) \sim (1/2\beta_c) \log N$ as $N \to \infty$. Finally we remark that the sum in Eq. (E.4) diverges as $\beta \downarrow \beta_c$. This gives the singularity of the free energy corrections at the critical point: $f_1(\beta, N) \sim (1/\beta_c) \log(1 - \beta_c/\beta)$.

References

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, (Wiley, New York, 1991).

[2] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*, (McGraw-Hill, New York, 1979).

[3] C. E. Shannon, Bell Syst. Tech. J. **27**, 379-423, 623-656 (1948).

[4] S.-Y. Chung, G. D. Forney, Jr., T. J. Richardson and R. Urbanke, *On the design of low-density parity-check codes within* 0.0045 *dB from the Shannon limit*, IEEE Comm. Letters, to appear.

[5] C. Berrou, A. Glavieux, and P. Thitimajshima. Proc. 1993 Int. Conf. Comm. 1064-1070.

[6] D. J. .C. MacKay, IEEE Trans. Inform. Theory **45**, 399-431 (1999).

[7] R. G. Gallager. *Low Density Parity Check Codes*, Research Monograph Series Vol. 21 (MIT, Cambridge, MA., 1963).

[8] N. Sourlas. Nature **339**, 693-694 (1989).

[9] N. Sourlas, *Statistical Mechanics of Neural Networks* Lecture Notes in Physics **368**, edited by L. Garrido (Springer Verlag, 1990).

[10] N. Sourlas, *From Statistical Physics to Statistical Inference and Back,* edited by P. Grassberger and J.-P. Nadal (Kluwer Academic, 1994), p. 195.

[5] Obviously the ordered state cannot be longer neglected in computing κ

[11] A. Montanari and N. Sourlas, Eur. Phys. J. B **18**, 107-119 (2000).

[12] A. Montanari, Eur. Phys. J. B **18**, 121-136 (2000).

[13] I. Kanter and D. Saad, Phys. Rev. Lett. **83**, 2660-2663 (1999).

[14] I. Kanter and D. Saad, Phys. Rev. E. **61**, 2137-2140 (1999).

[15] Y. Kabashima, T. Murayama and D. Saad, Phys. Rev. Lett. **84**, 1355-1358 (2000).

[16] I. Kanter and D. Saad, Jour. Phys. A. **33**, 1675-1681 (2000).

[17] R. Vicente, D. Saad and Y. Kabashima, Phys. Rev. E. **60**, 5352-5366 (1999).

[18] R. Vicente, D. Saad and Y. Kabashima. Europhys. Lett. **51**, 698-704 (2000?).

[19] Y. Kabashima, N. Sazuka, K. Nakamura and D. Saad, *Tighter Decoding Reliability Bound for Gallager's Error-Correcting Code*, `cond-mat/0010173`.

[20] H. Nishimori. J. Phys. C **13**, 4071-4076 (1980).

[21] B. Derrida. Phys. Rev. B **24**, 2613-2626 (1981).

[22] M. Mezard, G. Parisi and M. A. Virasoro, *Spin Glass theory and Beyond.* (World Scientific, Singapore, 1987).

[23] R. Monasson, J. Phys. A **31** (1998) 513-529.

[24] R. M. Tanner, IEEE Trans. Infor. Theory, **27**, 533-547 (1981).

[25] H. Nishimori. Prog. Theor. Phys. **66**, 1169-1181 (1981).

[26] H. Nishimori and D. Sherrington, *Absence of Replica Symmetry Breaking in a Region of the Phase Diagram of the Ising Spin Glass*, `cond-mat/0008139`.

[27] P. Ruján, Phys.Rev.Lett. **70**, 2968-2971 (1993).

[28] N. Sourlas. Europhys.Lett. **25**, 159-164 (1994).

[29] T. Richardson and R. Urbanke, *The Capacity of Low-Density Parity Check Codes under Message-Passing Decoding*, IEEE Trans. Inform. Theory, to appear.

[30] K. Y. M. Wong and D. Sherrington, J. Phys. A **21** L459-L466 (1988).

[31] M. Mézard and G. Parisi, *The Bethe lattice spin glass revisited*, `cond-mat/0009418`, to appear in Eur. Phys. J. B.

[32] G. Biroli, R. Monasson, M. Weigt, Eur. Phys. J. B **14**, 551-568 (2000).

[33] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes*, (Cambridge University Press, Cambridge, 1986).

[34] D. J. .C. MacKay, *On thresholds of codes*, available at `http://wol.ra.phy.cam.ac.uk/mackay/abstracts/theorems`.

[35] S. Franz, M. Leone, F. Ricci-Tersenghi and R. Zecchina, *Exact solutions for diluted spin glasses and optimization problems*, `cond-mar/0103328`.