

## Error-correcting codes and image restoration with multiple stages of dynamics

K. Y. Michael Wong

*Department of Physics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

Hidetoshi Nishimori

*Department of Physics, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo 152-8551, Japan*

(Received 7 March 2000)

We consider the problems of error-correcting codes and image restoration with multiple stages of dynamics. Information extracted from the former stage can be used selectively to improve the performance of the latter one. Analytic results were derived for the mean-field systems using the cavity method. We find that it has the advantage of being tolerant to uncertainties in hyperparameter estimation, as confirmed by simulations.

PACS number(s): 05.50.+q, 75.10.Nr, 89.70.+c

### I. INTRODUCTION

The corruption of signals by noise is a common problem encountered in information processing. To retrieve signals from messages corrupted during the transmission through noisy channels, various error-correcting codes have been proposed [1]. In particular, the error-correction mechanism of a class of parity-checking codes can be considered as a search for thermodynamically stable states of a Hamiltonian constructed in terms of the message bits [2]. These codes have been demonstrated to saturate the Shannon information bound in the limit that each encoded bit checks the parity of an infinitely large number of message bits [2,3]. While in practice, each encoded bit can only check the parity of a finite number of message bits, these codes still maintain a very low bit error probability.

The need to retrieve signals from corrupted messages is also inherent in image restoration [4]. Although parity-checking bits may not be explicitly introduced for the task, prior knowledge about the images plays a similar role. For example, the smoothness of real-world images provides a mechanism for checking the pixel values in comparison with those of their neighbors. A corresponding Hamiltonian, consisting of a ferromagnetic bias to reflect the smoothening tendency, can be constructed in terms of the image pixels. Modern techniques of image restoration based on Markov random fields correspond to the search for thermodynamically stable states of the Hamiltonian system, using methods such as simulated annealing [4].

In a recent paper, we showed that the problems of error-correcting codes and image restoration can be formulated in a unified framework [5]. In both tasks, the choice of the so-called hyperparameters is an important factor in determining their performances. Hyperparameters refer to the coefficients of the various interactions appearing in the Hamiltonian of the tasks. In error correction, they determine the statistical significance given to the parity-checking terms and the received bits. Similarly in image restoration, they determine the statistical weights given to the prior knowledge and the received data. It was shown, by the use of inequalities, that the optimal choice of the hyperparameters correspond to the maximum posterior marginal method, where there is a match between the source and model priors. The choice of

these values correspond to the Nishimori point in the space of hyperparameters [6]. This is equivalent to a thermodynamic process at finite temperature, and the task performance is better than the maximum *a posteriori* probability (MAP) method, where the values of the hyperparameters are taken to infinity, equivalent to a zero temperature process. Furthermore, from the analytic solution of the infinite-range model and the Monte Carlo simulation of finite-dimensional models, it was shown that an inappropriate choice of the hyperparameters can lead to a rapid degradation of the tasks.

In fact, hyperparameter estimation has been the subject of many previous studies [7], a recently popular one using the “evidence framework” [8]. However, if the prior models the source poorly, no hyperparameters can be reliable [9]. Even if they can be estimated accurately through steady-state statistical measurements, they may fluctuate when interfered with by bursty noise sources in communication channels. Hence it is important to devise decoding or restoration procedures which are robust against the uncertainties in hyperparameter estimation.

In this paper we propose the technique of selective freezing as a method to increase the tolerance to uncertainties in hyperparameter estimation. The technique has been studied for pattern reconstruction in neural networks, where it led to an improvement in the retrieval precision, a widening of the basin of attraction, and a boost in the storage capacity [10]. The idea is best illustrated for Ising bits or pixels with binary states  $\pm 1$ , though it can be easily generalized to other cases. In a finite temperature thermodynamic process, the Ising variables keep moving under thermal agitation. Some of them have smaller thermal fluctuations than the others, implying that they are more certain to stay in one state than the other. This stability implies that they have a higher probability to stay in the correct state for error-correction or image restoration tasks, even when the hyperparameters are not optimally tuned. It may thus be interesting to separate the thermodynamic process into two stages. In the first stage we select those relatively stable bits or pixels whose time-averaged states have a magnitude exceeding a certain threshold. In the second stage we subsequently fix (or freeze) them in the most probable thermodynamic states (for Ising variables this corresponds to the sign of the time-averaged state). Thus these selectively frozen bits or pixels are able to pro-

vide a more robust assistance to the less stable bits or pixels in their search for the most probable states. The selective freezing procedure reduces to the usual finite-temperature decoding or restoration process if all bits or pixels are frozen (since nothing happens in the second stage), or no bits or pixels are frozen (since the second stage is merely a continuation of the equilibration process of the first stage).

The two-stage thermodynamic process can be studied analytically in the mean-field model, which provides a qualitative guide to the behavior of more realistic cases of lower dimensions. However, it is necessary to make a remark about the theoretical approach. That is, as far as we have tried, the analytical solution has been inaccessible by the more conventional replica method. Rather, we have to use the cavity method to obtain the equations for the order parameters. In particular, the cavity method leads to the appearance of a term called the trans-susceptibility, which correctly describes the effects of the thermodynamics of the first stage on that of the second.

The paper is organized as follows. In Sec. II we briefly review the formulation of error-correcting codes and image restoration in a unified framework. In Secs. III and IV, we consider the mean-field model for error-correcting codes and image restoration, respectively. We derive the equations for the order parameters of the two-stage thermodynamics using the cavity method, and present numerical results illustrating the robustness of selective freezing against uncertainties in hyperparameter estimation. We further demonstrate that even when the noise model changes without the receiver/restoration agent realizing the change (i.e., it makes a wrong estimation of the prior), the task performance is still robust. For the more realistic cases of lower dimensions, simulation results illustrate the relevance of the infinite-range model in providing qualitative guidance. The conclusion is given in Sec. V.

## II. FORMULATION

Consider an information source which generates data represented by a set of Ising spins  $\{\xi_i\}$ , where  $\xi_i = \pm 1$  and  $i = 1, \dots, N$ . The data are generated according to the source prior to  $P_s(\{\xi_i\})$ . For error-correcting codes transmitting unbiased messages, all sequences are equally probable, and  $P_s(\{\xi\}) = 2^{-N}$ . For images with smooth structures, the prior consists of ferromagnetic Boltzmann factors, which increase the tendencies of the neighboring spins to stay at the same spin states, that is,

$$P_s(\{\xi\}) = \frac{1}{Z(\beta_s)} \exp\left(\frac{\beta_s}{z} \sum_{\langle ij \rangle} \xi_i \xi_j\right). \quad (1)$$

Here  $\langle ij \rangle$  represents pairs of neighboring spins,  $z$  is the valency of each site, and the partition function  $Z(\beta_s)$  is given by

$$Z(\beta_s) = \text{Tr}_\xi \exp\left(\frac{\beta_s}{z} \sum_{\langle ij \rangle} \xi_i \xi_j\right). \quad (2)$$

The data are coded by constructing the codewords, which are the products of  $p$  spins  $J_{i_1 \dots i_p}^0 = \xi_{i_1} \dots \xi_{i_p}$  for appropriately

chosen sets of indices  $\{i_1, \dots, i_p\}$ , the choice of which determines the type of code. Each spin may appear in a number of  $p$ -spin codewords; the number of times of appearance is called the valency  $z_p$ . The Sourlas code [2] is equivalent to the infinite-range model, in which all possible codewords of  $p$  spins are chosen from  $N$  spins. On the other hand, the Kabashima-Saad code [3] consists of combinations in which each spin appears in a random preselection of  $z_p$  codewords. For conventional image restoration, codewords with only  $p = 1$  are transmitted, corresponding to the pixels in the image; the inclusion of terms with  $p > 1$ , and their positive effects on restoring the original image, were also discussed in Ref. [5]. For simplicity, we restrict ourselves to the case of a single nonvanishing value of  $p$  with  $p \geq 2$ , and  $p = 1$ .

When the signal is transmitted through a noisy channel, the output consists of the sets  $\{J_{i_1 \dots i_p}\}$  and  $\{\tau_i\}$ , which are the corrupted versions of  $\{J_{i_1 \dots i_p}^0\}$  and  $\{\xi_i\}$  respectively. In the binary symmetric channel, the outputs  $J_{i_1 \dots i_p}$  are equal to  $\mp J_{i_1 \dots i_p}^0$  with probabilities  $p_J$  and  $1 - p_J$ , respectively, and  $\tau_i$  equal to  $\mp \xi_i$  with probabilities  $p_\tau$  and  $1 - p_\tau$  respectively. Thus

$$P_{\text{out}}(\{J\}, \{\tau\} | \{\xi\}) \propto \exp\left(\beta_J \sum J_{i_1 \dots i_p} \xi_{i_1} \dots \xi_{i_p} + \beta_\tau \sum \tau_i \xi_i\right), \quad (3)$$

where

$$\beta_J = \frac{1}{2} \ln \frac{1 - p_J}{p_J} \quad \text{and} \quad \beta_\tau = \frac{1}{2} \ln \frac{1 - p_\tau}{p_\tau}. \quad (4)$$

The first summation in the exponent of Eq. (3) extends over an appropriate set of the indices  $(i_1, \dots, i_p)$ .

The Gaussian channel is defined by, for a given sequence  $\{\xi_i\}$ ,

$$P_{\text{out}}(\{J\}, \{\tau\} | \{\xi\}) \propto \exp\left(-\frac{1}{2J^2} \sum (J_{i_1 \dots i_p} - J_0 \xi_{i_1} \dots \xi_{i_p})^2 - \frac{1}{2\tau^2} \sum (\tau_i - a \xi_i)^2\right). \quad (5)$$

$J_0$  and  $a$  are the strengths of the signals to be fed into the channel, and  $J^2$  and  $\tau^2$  are the variances of the noise. We note that by letting  $\beta_J$  and  $\beta_\tau$  to be  $J_0/J^2$  and  $a/\tau^2$  respectively, the input-dependent terms of Eq. (5) reduce to those of Eq. (3), which therefore can be regarded as the noise model for both binary symmetric and Gaussian channels.

According to Bayesian statistics, the posterior probability that the source sequence is  $\{\sigma\}$ , given the outputs  $\{J\}$  and  $\{\tau\}$ , takes the form

$$P(\{\sigma\}|\{J\},\{\tau\}) \propto P_{\text{out}}(\{J\},\{\tau\}|\{\sigma\})P_s(\{\sigma\}). \quad (6)$$

Using Eqs. (3) and (1), we have

$$P(\{\sigma\}|\{J\},\{\tau\}) \propto \exp\left(\beta_J \sum J_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p} + \beta_\tau \sum \tau_i \sigma_i + \frac{\beta_s}{z} \sum_{\langle ij \rangle} \sigma_i \sigma_j\right). \quad (7)$$

It often happens that the receiver at the end of the noisy channel does not have precise information on  $\beta_J$ ,  $\beta_\tau$ , or  $\beta_s$ . One then has to estimate these parameters. If the receiver estimates  $\beta$ ,  $h$  and  $\beta_m$  for  $\beta_J$ ,  $\beta_\tau$  and  $\beta_s$ , respectively, then the mean of the posterior distribution of  $\sigma_i$  is equal to the thermal average,

$$\langle \sigma_i \rangle = \frac{\text{Tr} \sigma_i e^{-H\{\sigma\}}}{\text{Tr} e^{-H\{\sigma\}}}, \quad (8)$$

where the Hamiltonian is given by

$$\langle \tilde{\sigma}_i \rangle = \frac{\text{Tr} \tilde{\sigma}_i \prod_j [\Theta(\langle \sigma_j \rangle^2 - \theta^2) \delta_{\tilde{\sigma}_j, \text{sgn}(\sigma_j)} + \Theta(\theta^2 - \langle \sigma_j \rangle^2)] e^{-\tilde{H}\{\tilde{\sigma}\}}}{\text{Tr} \prod_j [\Theta(\langle \sigma_j \rangle^2 - \theta^2) \delta_{\tilde{\sigma}_j, \text{sgn}(\sigma_j)} + \Theta(\theta^2 - \langle \sigma_j \rangle^2)] e^{-\tilde{H}\{\tilde{\sigma}\}}}, \quad (10)$$

where  $\Theta$  is the step function, and  $\tilde{H}\{\tilde{\sigma}\}$  is the Hamiltonian for the second stage, and has the same form as Eq. (9) in the first stage. To increase the flexibility in the process, the parameters  $\beta$ ,  $h$ , and  $\beta_m$  can be replaced by  $\tilde{\beta}$ ,  $\tilde{h}$ , and  $\tilde{\beta}_m$  respectively in the second stage. One then regards  $\text{sgn}\langle \tilde{\sigma}_i \rangle$  as the  $i$ th spin of the decoding/restoration process.

The most important quantity in selective freezing is the overlap of the decoded/restored bit  $\text{sgn}\langle \tilde{\sigma}_i \rangle$  and the original bit  $\xi_i$  averaged over the output probability and the spin distribution. This is given by

$$M_{\text{sf}} = \sum_{\xi} \prod \int dJ \prod \int d\tau P_s(\{\xi\}) \times P_{\text{out}}(\{J\},\{\tau\}|\{\xi\}) \xi_i \text{sgn}\langle \tilde{\sigma}_i \rangle. \quad (11)$$

Following Appendix A of Ref. [5], we can prove the inequality

$$M_{\text{sf}} \leq M(\beta = \beta_J, h = \beta_\tau, \beta_m = \beta_s), \quad (12)$$

where the right hand side is the overlap of the *single-stage* dynamics when the model parameters  $\beta$ ,  $h$ , and  $\beta_m$  match the source parameters  $\beta_J$ ,  $\beta_\tau$ , and  $\beta_s$ , respectively. Hence selective freezing cannot outperform the single-stage process

$$H\{\sigma\} = -\beta \sum J_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p} - h \sum \tau_i \sigma_i - \frac{\beta_m}{z} \sum_{\langle ij \rangle} \sigma_i \sigma_j. \quad (9)$$

One then regards  $\text{sgn}\langle \sigma_i \rangle$  as the  $i$ th bit of the decoded/restored information.

To reduce the sensitivity of the decoding/restoration process to the uncertainties in parameter estimation, we propose a two-stage process of selective freezing instead of the one-stage thermodynamic process implied by Eq. (8). In the first stage the spins evolve thermodynamically as prescribed in Eq. (8), and the thermal averages  $\langle \sigma_i \rangle$  of the spins are monitored. We may relate  $\langle \sigma_i \rangle$  to an effective field  $H_i$  by  $\langle \sigma_i \rangle = \tanh H_i$ . Spins with larger magnitudes of  $\langle \sigma_i \rangle$  correspond to larger magnitudes of  $H_i$ . They are more likely to agree with the correct message or image bit, and are less likely to change signs even when the hyperparameters vary. Their relative stability can be used to assist the less stable spins to boost their robustness against hyperparameter uncertainties. Hence we select those spins with  $|\langle \sigma_i \rangle|$  exceeding a given threshold  $\theta$ , and freeze them in the second stage of the thermodynamics. The average of the spin  $\tilde{\sigma}_i$  in the second stage is then given by

if the hyperparameters can be estimated precisely. However, we remark that the purpose of selective freezing is rather to provide a relatively stable performance when the hyperparameters cannot be estimated precisely. This cannot be revealed from the inequality, but will be confirmed by the analytic and simulation results in Secs. III and IV.

### III. INFINITE-RANGE MODEL FOR ERROR-CORRECTING CODES

Let us now suppose that the output of the transmission channel consists of only the set of  $p$ -spin interactions  $\{J_{i_1 \dots i_p}\}$ . Hamiltonian (9) then becomes

$$H\{\sigma\} = -\beta \sum_{i_1 < \dots < i_p} J_{i_1 \dots i_p} \sigma_{i_1} \dots \sigma_{i_p}, \quad (13)$$

where we have set  $\beta_m = 0$  for the case that all messages are equally probable.

Analytical solutions for the overlap are in general unavailable. We therefore consider the infinite-range model in which the exchange interactions are present for all possible pairs of sites in the Hamiltonian of Eq. (13).

To consider the transition between error-free and errored regimes, we are interested in the noise model in which

$J_{i_1 \dots i_p}$  is Gaussian with mean  $p!j_0\xi_{i_1} \dots \xi_{i_p}/N^{p-1}$  and variance  $p!J^2/2N^{p-1}$ . Since all messages are equally probable, we can apply a gauge transformation  $\sigma_i \rightarrow \sigma_i \xi_i$  and  $J_{i_1 \dots i_p} \rightarrow J_{i_1 \dots i_p} \xi_{i_1} \dots \xi_{i_p}$  to Eq. (13), and arrive at an equivalent  $p$ -spin model with a ferromagnetic bias, where

$$P(J_{i_1 \dots i_p}) = \left( \frac{N^{p-1}}{\pi J^2 p!} \right)^{1/2} \exp \left[ - \frac{N^{p-1}}{J^2 p!} \times \left( J_{i_1 \dots i_p} - \frac{p!}{N^{p-1}} j_0 \right)^2 \right]. \quad (14)$$

The Nishimori point for this model is located at  $\beta = 2j_0/J^2$ .

The infinite-range model is exactly solvable using mean-field theoretical techniques for disordered systems such as the replica or cavity method [11]. Here we use the cavity method because of its more transparent physical interpretation, and some obstacles encountered in the use of the replica method.

The cavity method uses a self-consistency argument to consider what happens when a spin is added or removed from the system. The central quantity in this method is the *cavity field*, which is the local field of a spin when it is added to the system, assuming that the exchange couplings act only one way from the system to the new spin (but not from the spin back to the system). Since the exchange couplings feeding the new spin have no correlations with the system, the cavity field becomes a Gaussian variable in the limit of large valency.

### A. Average spin in the first stage

We start with the so-called ‘‘clustering property’’ for mean-field systems [11],

$$\langle \sigma_{i_1} \dots \sigma_{i_p} \rangle = \langle \sigma_{i_1} \rangle \dots \langle \sigma_{i_p} \rangle, \quad (15)$$

where  $\langle \rangle$  represents thermodynamic averages. As shown in Appendix A, the clustering property enables us to express the thermal averages of a spin in terms of the cavity field, say, for spin 1,

$$\langle \sigma_1 \rangle = \tanh \beta h_1,$$

$$h_1 = \sum_{1 < j_2 < \dots < j_p} J_{1j_2 \dots j_p} \langle \sigma_{j_2} \rangle^{\setminus 1} \dots \langle \sigma_{j_p} \rangle^{\setminus 1}, \quad (16)$$

where the superscript  $\setminus 1$  denotes the thermal averages for a Hamiltonian in which  $\sigma_1$  and the associated exchange interactions are absent, but otherwise identical to Eq. (13). Thus  $h_1$  is the cavity field obeying a Gaussian distribution, whose mean and variance are  $pj_0m^{p-1}$  and  $pJ^2q^{p-1}/2$  respectively, where  $m$  and  $q$  are the magnetization and Edwards-Anderson order parameter, respectively, given by

$$m \equiv \frac{1}{N} \sum_i \langle \sigma_i \rangle \quad \text{and} \quad q \equiv \frac{1}{N} \sum_i \langle \sigma_i \rangle^2. \quad (17)$$

It is convenient to write

$$\beta h_i = \hat{m} + \sqrt{\hat{q}} u_i, \quad (18)$$

where

$$\hat{m} = p\beta j_0 m^{p-1} \quad \text{and} \quad \hat{q} = \frac{p}{2} \beta^2 J^2 q^{p-1}, \quad (19)$$

and  $u_i$  is a Gaussian variable with mean 0 and variance 1.

### B. Order parameters in the first stage

Self-consistently, applying the cavity argument to all terms in Eq. (17), we can obtain self-consistent equations for  $m$  and  $q$ :

$$m = \int Du \tanh G, \quad (20)$$

$$q = \int Du \tanh^2 G, \quad (21)$$

where  $Du \equiv du e^{-u^2/2}/\sqrt{2\pi}$  is the Gaussian measure and  $G = \hat{m} + \sqrt{\hat{q}}u$ . The overlap for the one-stage decoding process is given by

$$M \equiv \frac{1}{N} \sum_i \text{sgn} \langle \sigma_i \rangle = \text{erf} \frac{\hat{m}}{\sqrt{2\hat{q}}}. \quad (22)$$

Now we consider selective freezing. If we introduce a freezing threshold  $\theta$  so that all spins with  $\langle \sigma_i \rangle^2 > \theta^2$  are frozen, then the freezing fraction  $f$  is given by

$$f \equiv \frac{1}{N} \sum_i \Theta(\langle \sigma_i \rangle^2 - \theta^2) = 1 - \frac{1}{2} \text{erf} \frac{u_+}{\sqrt{2}} + \frac{1}{2} \text{erf} \frac{u_-}{\sqrt{2}}, \quad (23)$$

where  $u_{\pm} = (\pm u_0 - \hat{m})/\sqrt{\hat{q}}$  with  $\tanh u_0 = \theta$ .

### C. Average spin in the second stage

Assuming that the spin  $\tilde{\sigma}_1$  is dynamic in the second stage, we can write

$$H\{\tilde{\sigma}\} \approx H\{\tilde{\sigma}\}^{\setminus 1} - \tilde{\beta} \sum_{1 < j_1 \dots < j_{p-1}} \tilde{\sigma}_1 J_{1j_1 \dots j_{p-1}} \prod_{s=1}^{p-1} [\tilde{\sigma}_{j_s} \Theta(\theta^2 - \langle \sigma_{j_s} \rangle^2) + \text{sgn} \langle \sigma_{j_s} \rangle \Theta(\langle \sigma_{j_s} \rangle^2 - \theta^2)], \quad (24)$$

where  $H\{\tilde{\sigma}\}^{\setminus 1}$  is the Hamiltonian when spin 1 is completely removed from the system in both stages of the thermodynamic process. Removing spin 1 may cause the thermal averages of other spins to adjust slightly in the first stage. Hence some dynamic spins (with  $\langle \sigma_k \rangle^2 < \theta^2$ ) may become frozen ones (with  $\langle \sigma_k \rangle^2 > \theta^2$ ), and vice versa, so that, strictly speaking, further terms should be considered in Eq. (24) to account for these secondary effects. For example, if spin  $k$  is induced to switch from dynamic to frozen (or vice versa) on removal of spin 1, then the Taylor expansion of  $H\{\tilde{\sigma}\}$  implies that an extra term

$$\begin{aligned}
& -\tilde{\beta}(\text{sgn}\langle\sigma_k\rangle^{\setminus 1}-\tilde{\sigma}_k)[\delta(\langle\sigma_k\rangle^{\setminus 1}-\theta)-\delta(\langle\sigma_k\rangle^{\setminus 1}+\theta)](\langle\sigma_k\rangle \\
& -\langle\sigma_k\rangle^{\setminus 1})\sum_{1<j_1,\dots,j_{p-1}\neq k}J_{kj_1\dots j_{p-1}}\prod_{s=1}^{p-1}\{\tilde{\sigma}_{j_s}\Theta \\
& \times[\theta^2-(\langle\sigma_{j_s}\rangle^{\setminus 1})^2]+\text{sgn}\langle\sigma_{j_s}\rangle^{\setminus 1}\Theta[(\langle\sigma_{j_s}\rangle^{\setminus 1})^2-\theta^2]\} \\
& \quad \quad \quad (25)
\end{aligned}$$

should be incorporated into Eq. (24). Here we have neglected these terms for clarity. Nevertheless, justification *a posteriori* can be provided for their deletion.

Using a cavity argument similar to Appendix A, we can show that

$$\begin{aligned}
\langle\tilde{\sigma}_1\rangle &= \tanh\tilde{\beta}\left\{\sum_{1<j_1,\dots,j_{p-1}}J_{1j_1\dots j_{p-1}}\prod_{s=1}^{p-1}[\langle\tilde{\sigma}_{j_s}\rangle^{\setminus 1}\right. \\
& \quad \left.\times\Theta(\theta^2-\langle\sigma_{j_s}\rangle^2)+\text{sgn}\langle\sigma_{j_s}\rangle^{\setminus 1}\Theta(\langle\sigma_{j_s}\rangle^2-\theta^2)\right\}. \\
& \quad \quad \quad (26)
\end{aligned}$$

However, the effective field on the right hand side of Eq. (26) is still not a cavity field because  $\langle\sigma_{j_s}\rangle$ , which is used in the step functions to decide whether the spin  $j_s$  is dynamic or frozen in the second stage, is different from  $\langle\sigma_{j_s}\rangle^{\setminus 1}$ . Hence it may have correlations with spin 1. Taylor expansion of  $\langle\sigma_{j_s}\rangle$  about  $\langle\sigma_{j_s}\rangle^{\setminus 1}$  yields

$$\begin{aligned}
\langle\tilde{\sigma}_1\rangle &= \tanh\tilde{\beta}\left\{\tilde{h}_1+\sum_{1j\neq j_1,\dots,j_{p-2}}J_{1j_1\dots j_{p-2}}\right. \\
& \quad \times\prod_{s=1}^{p-2}[\langle\tilde{\sigma}_{j_s}\rangle^{\setminus 1}\Theta[\theta^2-(\langle\sigma_{j_s}\rangle^{\setminus 1})^2]+\text{sgn}\langle\sigma_{j_s}\rangle^{\setminus 1}\Theta \\
& \quad \times[(\langle\sigma_{j_s}\rangle^{\setminus 1})^2-\theta^2]][\text{sgn}\langle\sigma_j\rangle^{\setminus 1}-\langle\tilde{\sigma}_j\rangle^{\setminus 1}][\delta(\langle\sigma_j\rangle^{\setminus 1} \\
& \quad \left.-\theta)-\delta(\langle\sigma_j\rangle^{\setminus 1}+\theta)](\langle\sigma_j\rangle-\langle\sigma_j\rangle^{\setminus 1})\right\}, \\
& \quad \quad \quad (27)
\end{aligned}$$

where  $\tilde{h}_1$  is the generic cavity field which is now completely uncorrelated with spin 1. It is given by

$$\begin{aligned}
\tilde{h}_1 &= \sum_{1<j_1,\dots,j_{p-1}}J_{1j_1\dots j_{p-1}}\prod_{s=1}^{p-1}\{\langle\tilde{\sigma}_{j_s}\rangle^{\setminus 1}\Theta[\theta^2-(\langle\sigma_{j_s}\rangle^{\setminus 1})^2] \\
& \quad +\text{sgn}\langle\sigma_{j_s}\rangle^{\setminus 1}\Theta[(\langle\sigma_{j_s}\rangle^{\setminus 1})^2-\theta^2]\}. \\
& \quad \quad \quad (28)
\end{aligned}$$

To evaluate the difference  $\langle\sigma_j\rangle-\langle\sigma_j\rangle^{\setminus 1}$  appearing in Eq. (27), we have to apply the cavity method a second time, by comparing the changes when both spins 1 and  $j$  are removed. This is done in Appendix B, and the result is

$$\langle\sigma_j\rangle-\langle\sigma_j\rangle^{\setminus 1}=(\beta\text{sech}^2\beta h_j^{\setminus 1})(h_{j1}\tanh\beta h_1^{\setminus j}), \quad (29)$$

where

$$h_{1j}=h_{j1}=\sum_{1j\neq k_1,\dots,k_{p-2}}J_{1jk_1\dots k_{p-2}}\langle\sigma_{k_1}\rangle^{\setminus 1j}\dots\langle\sigma_{k_{p-2}}\rangle^{\setminus 1j}. \quad (30)$$

When Eqs. (29)–(30) are substituted into Eq. (27), the significant contribution comes from the terms which pair up  $J_{1jj_1\dots j_{p-2}}$  and  $J_{1jk_1\dots k_{p-2}}$ . The various terms appearing in the summation over  $j\neq j_1<\dots<j_{p-2}$  involve thermal averages in the absence of spins 1 or  $j$ . We assume that the effects of removing a spin is negligible (which can be shown to be equivalent to the replica symmetric approximation in the replica method [12]). Then, replacing the components of the terms by their mean values, and counting that  $N^{p-2}/(p-2)!$  terms appearing in the summation over  $j_1<\dots<j_{p-2}$ , we arrive at

$$\begin{aligned}
\langle\tilde{\sigma}_1\rangle &= \tanh\tilde{\beta}\left\{\tilde{h}_1+\frac{p}{2}(p-1)J^2\frac{1}{N}\sum_j[\delta(\langle\sigma_j\rangle-\theta)-\delta(\langle\sigma_j\rangle\right. \\
& \quad \left.+\theta)][\text{sgn}\langle\sigma_j\rangle-\langle\tilde{\sigma}_j\rangle](\beta\text{sech}^2\beta h_j)(r^{p-2}\tanh\beta h_1)\right\}, \\
& \quad \quad \quad (31)
\end{aligned}$$

where  $r$  is the order parameter describing the spin correlations of the two thermodynamic stages:

$$r\equiv\frac{1}{N}\sum_i\langle\sigma_i\rangle\{\langle\tilde{\sigma}_i\rangle\Theta[\theta^2-\langle\sigma_i\rangle^2]+\text{sgn}\langle\sigma_i\rangle\Theta[\langle\sigma_i\rangle^2-\theta^2]\}. \quad (32)$$

Equation (31) can be simplified by introducing the transusceptibility  $\chi_{tr}$ , which describes the response of a spin in the second stage to variations of the cavity field in the first stage, namely,

$$\chi_{tr}\equiv\frac{1}{N}\sum_i\frac{\partial\langle\tilde{\sigma}_i\rangle}{\partial h_i}. \quad (33)$$

Since  $\langle\tilde{\sigma}_i\rangle$  equals  $\text{sgn}h_i$  for  $\tanh^2\beta h_i>\theta^2$ , and  $\tanh\beta\tilde{h}_i$  otherwise, we obtain

$$\begin{aligned}
\chi_{tr} &= \frac{1}{N}\sum_i[\delta(\langle\sigma_i\rangle-\theta)-\delta(\langle\sigma_i\rangle+\theta)][\text{sgn}\langle\sigma_i\rangle \\
& \quad -\langle\tilde{\sigma}_i\rangle]\beta\text{sech}^2\beta h_i. \\
& \quad \quad \quad (34)
\end{aligned}$$

Equation (31) can thus be simplified to

$$\langle\tilde{\sigma}_1\rangle=\tanh\tilde{\beta}\left\{\tilde{h}_1+\frac{p}{2}(p-1)J^2r^{p-2}\chi_{tr}\tanh\beta h_1\right\}. \quad (35)$$

#### D. Order parameters in the second stage

The cavity field  $\tilde{h}_1$  in the second stage is a Gaussian variable. Its mean and variance are  $pj_0\tilde{m}^{p-1}$  and  $pJ^2\tilde{q}^{p-1}/2$ , respectively, where  $\tilde{m}$  and  $\tilde{q}$  are the magnetization and Edwards-Anderson order parameter, respectively, given by

$$\tilde{m} \equiv \frac{1}{N} \sum_i [\Theta(\theta^2 - \langle \sigma_i \rangle^2) \langle \tilde{\sigma}_i \rangle + \Theta(\langle \sigma_i \rangle^2 - \theta^2) \text{sgn} \langle \sigma_i \rangle], \quad (36)$$

$$\tilde{q} \equiv \frac{1}{N} \sum_i [\Theta(\theta^2 - \langle \sigma_i \rangle^2) \langle \tilde{\sigma}_i \rangle^2 + \Theta(\langle \sigma_i \rangle^2 - \theta^2)]. \quad (37)$$

Furthermore, the covariance between  $h_1$  and  $\tilde{h}_1$  is  $pJ^2 r^{p-1}/2$ , where  $r$  is given in Eq. (32).

Algebraic manipulations can be simplified if we write, for  $i=1$ ,

$$\beta h_i = \hat{m} + \sqrt{\hat{q}} u_i, \quad (38)$$

$$\tilde{\beta} \tilde{h}_i = \hat{m} + \sqrt{\hat{q}} (\eta u_i + \sqrt{1 - \eta^2} v_i), \quad (39)$$

where  $u_i$  and  $v_i$  are independent Gaussian variables with mean 0 and variance 1,  $\hat{m}$  and  $\hat{q}$  are given in Eq. (19), and

$$\hat{m} = p \tilde{\beta} j_0 \tilde{m}^{p-1} \quad \text{and} \quad \hat{q} = \frac{p}{2} \tilde{\beta}^2 J^2 \tilde{q}^{p-1}, \quad (40)$$

$$\hat{r} = \frac{p}{2} \beta \tilde{\beta} J^2 r^{p-1} \quad \text{and} \quad \eta = \frac{\hat{r}}{\sqrt{\hat{q} \hat{q}}}. \quad (41)$$

Self-consistently applying the same cavity argument to all terms in Eqs. (36), (37), (32), and (34), and performing the Gaussian average over  $u_i$  and  $v_i$ , we arrive at the following self-consistent equations for  $\tilde{m}$ ,  $\tilde{q}$ ,  $r$  and  $\chi_{tr}$ :

$$\tilde{m} = -\frac{1}{2} \text{erf} \frac{u_+}{\sqrt{2}} - \frac{1}{2} \text{erf} \frac{u_-}{\sqrt{2}} + \int_{u_-}^{u_+} Du \int Dv \tanh L, \quad (42)$$

$$\tilde{q} = 1 - \frac{1}{2} \text{erf} \frac{u_+}{\sqrt{2}} + \frac{1}{2} \text{erf} \frac{u_-}{\sqrt{2}} + \int_{u_-}^{u_+} Du \int Dv \tanh^2 L, \quad (43)$$

$$r = \left( \int_{-\infty}^{u_-} + \int_{u_+}^{\infty} \right) Du |\tanh G| + \int_{u_-}^{u_+} Du \int Dv \tanh G \tanh L, \quad (44)$$

$$\chi_{tr} = \frac{\exp(-u_+^2/2)}{J \sqrt{\pi p q^{p-1}}} \int Dv (1 - \tanh L_v^{(+)} ) + \frac{\exp(-u_-^2/2)}{J \sqrt{\pi p q^{p-1}}} \int Dv (1 + \tanh L_v^{(-)}), \quad (45)$$

where

$$L = \hat{m} + \sqrt{\hat{q}} (\eta u + \sqrt{1 - \eta^2} v) + \frac{p}{2} (p-1) \tilde{\beta} J^2 r^{p-2} \chi_{tr} \tanh G, \quad (46)$$

$$L_v^{(\pm)} = \hat{m} + \sqrt{\hat{q}} (\eta u_{\pm} + \sqrt{1 - \eta^2} v) \pm \frac{p}{2} (p-1) \tilde{\beta} J^2 r^{p-2} \chi_{tr} \theta. \quad (47)$$

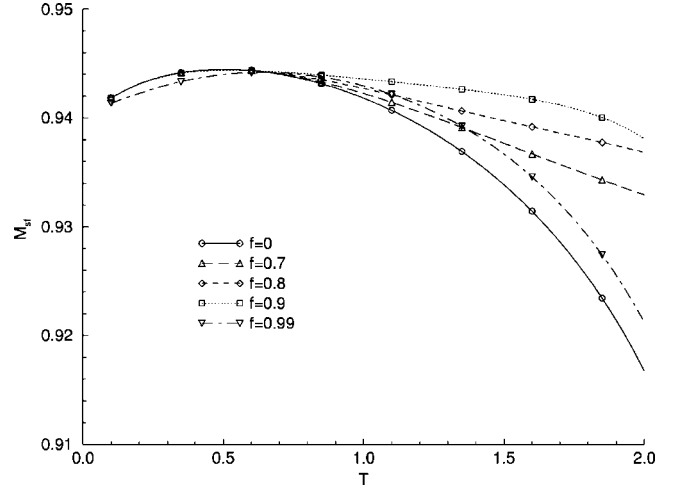


FIG. 1. The overlap  $M_{\text{sf}}$  as a function of the decoding temperature  $T$  for  $p=2$  and  $j_0=J=1$  for various given values of freezing fraction  $f$ . In this and the following figures,  $f=0$  corresponds to one-stage decoding/restoration.

Eqs. (20), (21) and (42)–(45) for the order parameters  $m$ ,  $q$ ,  $\tilde{m}$ ,  $\tilde{q}$ ,  $r$ , and  $\chi_{tr}$  form a close set of equations. The performance of selective freezing is measured by

$$M_{\text{sf}} \equiv \frac{1}{N} \sum_i [\Theta(\theta^2 - \langle \sigma_i \rangle^2) \text{sgn} \langle \tilde{\sigma}_i \rangle + \Theta(\langle \sigma_i \rangle^2 - \theta^2) \text{sgn} \langle \sigma_i \rangle]. \quad (48)$$

From the above parameters,  $M_{\text{sf}}$  can be derived as

$$M_{\text{sf}} = -\frac{1}{2} \text{erf} \frac{u_+}{\sqrt{2}} - \frac{1}{2} \text{erf} \frac{u_-}{\sqrt{2}} + \int_{u_-}^{u_+} Du \text{erf} \frac{L_u}{\sqrt{2 \tilde{q} (1 - \eta^2)}}, \quad (49)$$

where  $L_u = \hat{m} + \sqrt{\hat{q}} \eta u + [p(p-1)/2] \tilde{\beta} J^2 r^{p-2} \chi_{tr} \tanh G$ .

We have also tried to derive the above equations using the replica method. However, in the nearest results that we could find, terms involving the trans-susceptibility are absent, which we believe to be unphysical. Therefore, the replica approach to the order parameter equations remain an open question.

We show an example of the case  $p=2$  and  $j_0=J=1$  in Fig. 1, where the overlap  $M_{\text{sf}}$  is plotted as a function of the decoding temperature  $T (= \beta^{-1} = \tilde{\beta}^{-1})$  for various given values of freezing fraction  $f$ . When  $f=0$  (no spins frozen) and  $f=1$  (all spins frozen), the dynamics is equivalent to one with single stage, and the overlap reaches its maximum at the Nishimori point  $T = J^2/2j_0$ , as expected. We observe that the tolerance against variations in  $T$  is enhanced by selective freezing for certain values of  $f$ .

It is therefore interesting to consider the appropriate values of  $f$  for the best overlap at a given decoding temperature. Figures 2(a)–2(f) show that at high temperatures such as in Figs. 2(a)–2(c), there is a single maximum and its position is fairly independent of temperature, lying around  $f=0.9$  in the present case. At intermediate temperatures such as in Figs. 2(d)–2(e), there appear two maxima and as temperature

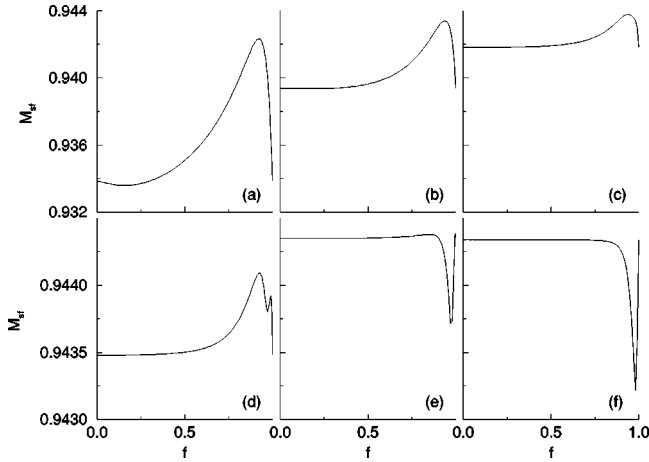


FIG. 2. The overlap  $M_{sf}$  as a function of the freezing fraction  $f$  at temperatures  $T =$  (a) 1.5, (b) 1.2, (c) 1.0, (d) 0.8, (e) 0.6, and (f) 0.4 for  $p=2$  and  $j_0=J=1$ .

changes, there is a discontinuous jump in the maximum position. Figure 2(f) shows that when the temperature is lower than the Nishimori point ( $T_N=0.5$ ), the overlap cannot be improved by selective freezing.

Figure 3 compares the overlap of the one-stage dynamics with that of the best of selective freezing. It shows that when the decoding temperature is misdetermined to be higher than its optimal value at the Nishimori point, selective freezing can provide a fairly robust performance. Furthermore, the choice of the freezing fraction for such robust performance appears to be quite independent of the temperature. The solid line in Fig. 4 locates the position for the best overlap and, as observed from Figs. 2(a)–2(f), lies in the vicinity of  $f \approx 0.9$  for a large range of temperature. The unshaded region in the same figure also indicates that selective freezing leads to an improvement in the overlap over a wide range of the parameter space.

We have also studied the dependence of the overlap on varying the freezing threshold  $\theta$  rather than the freezing fraction  $f$ . However, Fig. 5 shows that the optimal value of  $\theta$  has a much larger dependence on the temperature. This is due to

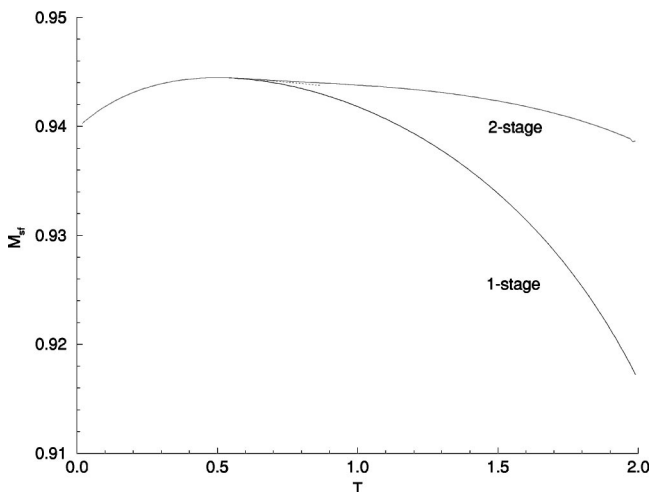


FIG. 3. The temperature dependence of the best overlap of selective freezing compared with the overlap of the one-stage dynamics for  $p=2$  and  $j_0=J=1$ .

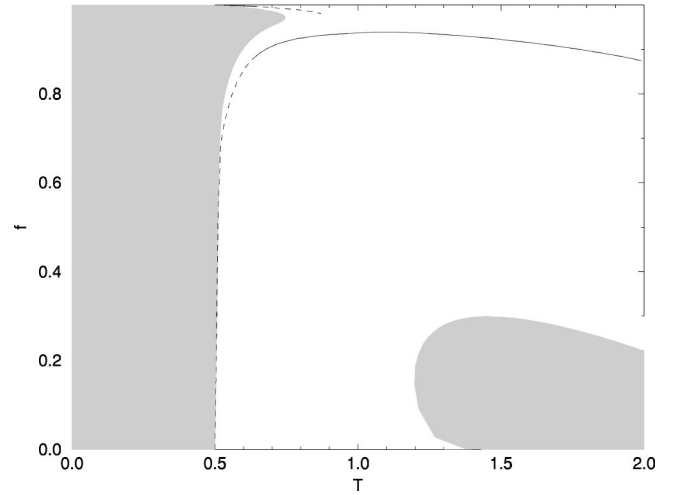


FIG. 4. The freezing fraction  $f$  for the best overlap as a function of temperature  $T$  for  $p=2$  and  $j_0=J=1$ . In this and the following figure, the solid line is the global maximum, the dashed line is the local maximum, and the shaded region has  $M_{sf} < M$ .

the sensitive dependence of the thermal averages of the spins on temperature. At high temperatures, most spins are thermally agitated, and the freezing threshold has to be set to a very low value in order to freeze a given fraction of spins. On the other hand, at low temperatures, most spins are relatively stable, and the freezing threshold has to be set to a very high value in order to keep a given fraction of spins dynamic in the second stage. We conclude that the freezing fraction is a better controlling parameter for the decoding performance.

The advantages of selective freezing are confirmed by Monte Carlo simulations shown in Fig. 6. For one-stage dynamics, the overlap is maximum at the Nishimori point ( $T_N=0.5$ ), as expected. However, it deteriorates rather rapidly when the decoding temperature increases. In contrast, selective freezing maintains a more steady performance, especially when  $f=0.9$ .

#### IV. MEAN-FIELD MODEL FOR IMAGE RESTORATION

In conventional image restoration problems, a given degraded image consists of the set of pixels  $\{\tau_i\}$ , but not the set

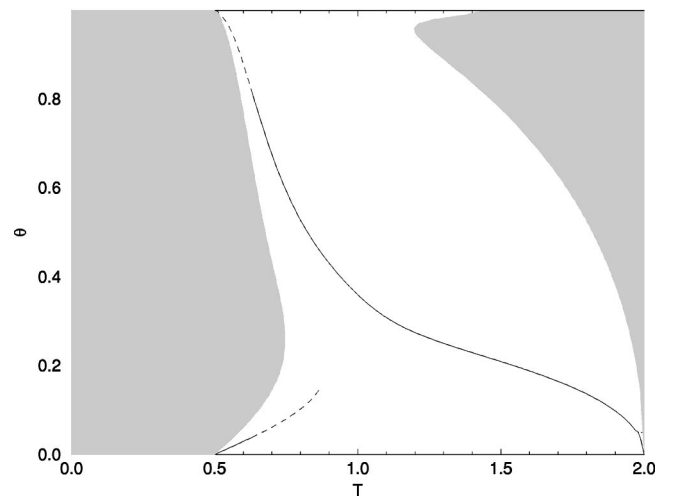


FIG. 5. The freezing threshold  $\theta$  for the best overlap as a function of temperature  $T$  for  $p=2$  and  $j_0=J=1$ .

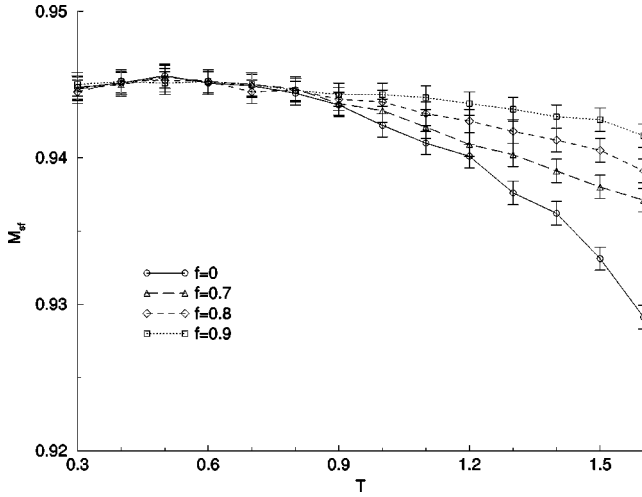


FIG. 6. Results of Monte Carlo simulations for the overlaps of selective freezing compared with those of the one-stage dynamics for  $p=2$  and  $j_0=J=1$ , corresponding to Fig. 1. The simulation parameters are  $N=1000$ , with an initial overlap of 0.9 and 200 samples. Each stage consists of 500 Monte Carlo steps per node for equilibration, and 1000 Monte Carlo steps per node for averaging.

of exchange interactions  $\{J_{i_1, \dots, i_p}\}$ . On the other hand, effective restoration requires the introduction of a model prior distribution of the pixels for smooth images. In this case the Hamiltonian corresponds to that of a random field Ising model:

$$H\{\sigma\} = -h \sum_i \tau_i \sigma_i - \frac{\beta_m}{z} \sum_{\langle ij \rangle} \sigma_i \sigma_j. \quad (50)$$

In mean-field systems, each pixel  $i$  has an extensive valency. The pixels  $\tau_i$  are the degraded versions of the source pixels  $\xi_i$ , corrupted by noise which, for convenience, is assumed to be Gaussian with mean  $a\xi_i$  and variance  $\tau^2$ , i.e.,

$$P(\tau_i | \xi_i) = \frac{\exp\left[-\frac{1}{2\tau^2}(\tau_i - a\xi_i)^2\right]}{\sqrt{2\pi\tau^2}}. \quad (51)$$

In turn, the source pixels satisfy the prior distribution in Eq. (1). Applying the cavity argument for mean-field systems, the prior distribution becomes factorizable,

$$P(\xi_i) = \frac{\exp(\beta_s m_0 \xi_i)}{2 \cosh \beta_s m_0}, \quad (52)$$

where  $m_0 = \tanh \beta_s m_0$ . The order parameter in the first stage is given by

$$\begin{aligned} m &\equiv \frac{1}{N} \sum_i \langle \sigma_i \rangle \\ &= \frac{1}{2 \cosh \beta_s m_0} \sum_{\xi=\pm 1} \exp(\beta_s m_0 \xi) \int Dx \tanh U, \end{aligned} \quad (53)$$

where  $U = \beta_m m + ha\xi + h\tau\xi$ . The overlap for the one-stage restoration process is given by

$$\begin{aligned} M &\equiv \frac{1}{N} \sum_i \xi_i \text{sgn}\langle \sigma_i \rangle \\ &= \frac{1}{2 \cosh \beta_s m_0} \sum_{\xi=\pm 1} \exp(\beta_s m_0 \xi) \xi \text{erf} \frac{\beta_m m + ha\xi}{\sqrt{2}h\tau}. \end{aligned} \quad (54)$$

Next we consider selective freezing in the second stage with a freezing threshold  $\theta$ . The freezing fraction is given by

$$\begin{aligned} f &\equiv \frac{1}{N} \sum_i \Theta(\langle \sigma_i \rangle^2 - \theta^2) = \frac{1}{2 \cosh \beta_s m_0} \sum_{\xi=\pm 1} \exp(\beta_s m_0 \xi) \\ &\times \left[ 1 - \frac{1}{2} \text{erf} \frac{u_+(\xi)}{\sqrt{2}} + \frac{1}{2} \text{erf} \frac{u_-(\xi)}{\sqrt{2}} \right], \end{aligned} \quad (55)$$

where  $u_{\pm}(\xi) = (\pm u_0 - \beta_m m - ha\xi)/h\tau$ , with  $\tanh u_0 = \theta$ . The order parameter of the second stage is given by

$$\begin{aligned} \tilde{m} &\equiv \frac{1}{N} \sum_i [\Theta(\theta^2 - \langle \sigma_i \rangle^2) \langle \tilde{\sigma}_i \rangle + \Theta(\langle \sigma_i \rangle^2 - \theta^2) \text{sgn}\langle \sigma_i \rangle] \\ &= \frac{1}{2 \cosh \beta_s m_0} \sum_{\xi=\pm 1} \exp(\beta_s m_0 \xi) \left[ -\frac{1}{2} \text{erf} \frac{u_+(\xi)}{\sqrt{2}} \right. \\ &\quad \left. - \frac{1}{2} \text{erf} \frac{u_-(\xi)}{\sqrt{2}} + \int_{u_-(\xi)}^{u_+(\xi)} Dx \tanh L \right], \end{aligned} \quad (56)$$

where  $L = \beta_m \tilde{m} + ha\xi + h\tau\xi$ . The overlap for selective freezing is given by

$$\begin{aligned} M_{\text{st}} &\equiv \frac{1}{N} \sum_i \xi_i [\Theta(\theta^2 - \langle \sigma_i \rangle^2) \text{sgn}\langle \tilde{\sigma}_i \rangle \\ &\quad + \Theta(\langle \sigma_i \rangle^2 - \theta^2) \text{sgn}\langle \sigma_i \rangle] \\ &= \frac{1}{2 \cosh \beta_s m_0} \sum_{\xi=\pm 1} \exp(\beta_s m_0 \xi) \xi \text{erf} \frac{g(\beta_m \tilde{m}) + ha\xi}{\sqrt{2}h\tau}, \end{aligned} \quad (57)$$

where

$$g(\beta_m \tilde{m}) = \begin{cases} \beta_m m - u_0, & \beta_m \tilde{m} < \beta_m m - u_0 \\ \beta_m \tilde{m}, & \beta_m m - u_0 < \beta_m \tilde{m} < \beta_m m + u_0 \\ \beta_m m + u_0, & \beta_m \tilde{m} > \beta_m m + u_0. \end{cases} \quad (58)$$

We note that since the spin-glass interaction is absent in this case, there are no trans-susceptibility effects. This is unlike the case of error-correcting codes, in which  $\chi_{tr}$  is nonzero when  $J$  is nonzero.

The three cases of the function  $g(\beta_m \tilde{m})$  in Eq. (58) correspond to three situations. When  $\beta_m \tilde{m} < \beta_m m - u_0$ , all the dynamic spins in the second stage have negative thermodynamic averages, and therefore take the value  $-1$  in the two-stage restoration process. This is equivalent to a one-stage restoration process in which all spins with thermodynamic



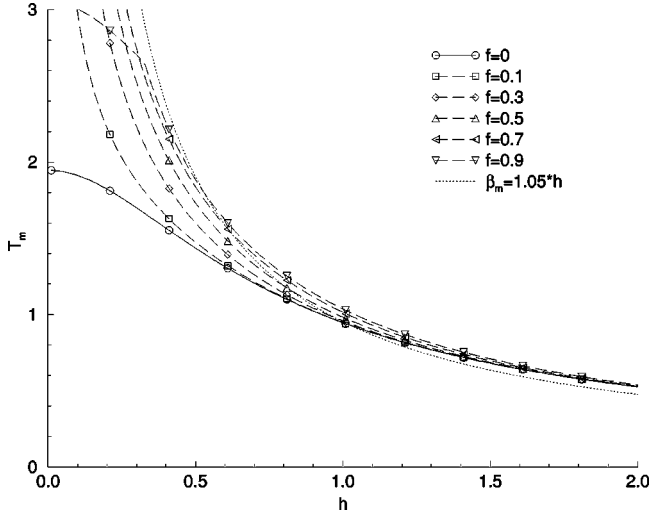


FIG. 7. The lines of optimal performance in the space of the random-field strength  $h$  and the restoration temperature  $T_m \equiv \beta_m^{-1}$  in the mean-field model of image restoration for  $a = \tau = 1$  and  $\beta_s = 1.05$ . The dotted line is the line of operation with  $\beta_m/h$  set to the optimal ratio  $\beta_s/\beta_\tau = 1.05$ . At  $f = 0.9$  the dynamic spins are completely frozen to  $+1$  to the left of the kink, but only partially so to the right.

averages above the threshold  $+\theta$  are frozen to  $+1$ , and to  $-1$  otherwise. Similarly, when  $\beta_m \tilde{m} > \beta_m m + u_0$ , all the dynamic spins in the second stage have positive thermodynamic averages. Only when  $\beta_m m - u_0 < \beta_m \tilde{m} < \beta_m m + u_0$  do we have the dynamic spins frozen to partly  $+1$  and partly  $-1$ .

We can consider the condition for the optimal performance  $M_{sf}$  of selective freezing. For a given distribution of data and noise,  $g(\beta_m \tilde{m})$  is the only adjustable parameter in Eq. (57), playing the same role as the adjustable parameter  $\beta_m m$  for one-stage dynamics in Eq. (54). In the space of  $h$  and  $\beta_m$ , the performance is optimal along the line  $h/\beta_\tau = \beta_m m/\beta_s m_0$  for one-stage dynamics [5] ( $\beta_\tau = a/\tau^2$  for Gaussian noise). Analogously, there exists a line of optimal performance defined by  $h/\beta_\tau = g(\beta_m \tilde{m})/\beta_s m_0$  for selective freezing.

An example of the lines of optimal performance is shown in Fig. 7. It is interesting to note the kinks for certain freezing fractions. They correspond to transitions of cases in which the dynamic spins are partially or completely frozen to  $\pm 1$ .

A comparison of Eqs. (54) and (57) shows that selective freezing performs as well as one-stage dynamics, but cannot outperform it. Nevertheless, selective freezing provides a rather stable performance when the hyperparameters cannot be estimated precisely. In image restoration, the usual practice is to choose a fixed ratio of  $\beta_m/h$ . Figure 8 confirms this stability along the line of operation with  $\beta_m/h$  set to the optimal ratio  $\beta_s/\beta_\tau$ . Note especially that the lines with  $f = 0.7$  and  $0.9$  attain a nearly optimal value of  $M_{sf}$  over a wide range of parameters. The kink at  $f = 0.9$  is, again, due to the appearance of the  $-1$  frozen dynamic spins (to the right of the kink).

The stable performance of selective freezing can be partly explained by the proximity of the lines of optimal perfor-

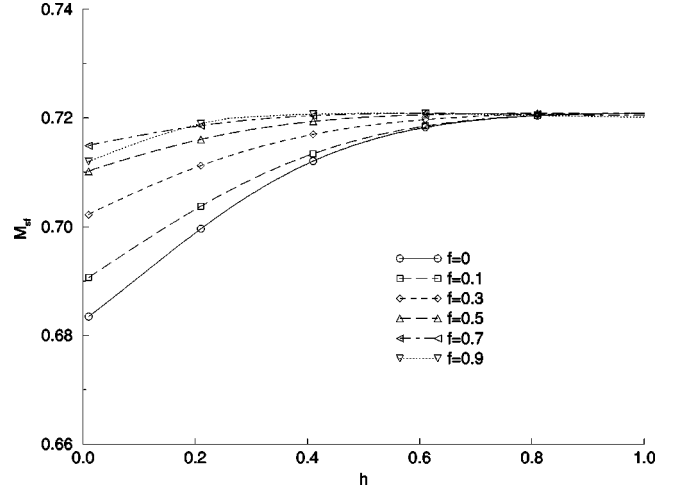


FIG. 8. The performance of selective freezing at  $a = \tau = 1$  and  $\beta_s = 1.05$ , with  $\beta_m/h$  set to the optimal ratio  $\beta_s/\beta_\tau = 1.05$  for various freezing fractions  $f$ .

mance with the line of operation which, as discussed in Ref. [5], is an important factor in hyperparameter estimation. This is illustrated by the optimal lines for small values of  $f$  near the Nishimori point  $(T_m, h) = (1.05^{-1}, 1)$  in Fig. 7.

However, the advantage of selective freezing does not only rely on the fortuitous combination of parameters. Even when the parameters are not chosen optimally, selective freezing still maintains a rather robust performance. For example, along the line of optimal performance for  $f = 0.9$  in Fig. 7, the bending at the kink only causes a modest reduction in the overlap  $M_{sf}$  in Fig. 8.

To study the robustness of the performance of selective freezing, we model a situation common in modern communication channels carrying multimedia traffic, which are often bursty in nature. Since ‘‘burstiness’’ results in intermittent interferences, we consider a noise with two Gaussian components, each with its own characteristics. A random fraction  $f_1$  of the pixels are influenced by Gaussian noise with signal strength  $a_1$  and noise variance  $\tau_1^2$ . The rest of the pixels have strength  $a_2$  and noise variance  $\tau_2^2$ . Hence the distribution of the degraded pixels are

$$P(\tau_i | \xi_i) = f_1 \frac{\exp\left[-\frac{1}{2\tau_1^2}(\tau_i - a_1 \xi_i)^2\right]}{\sqrt{2\pi\tau_1^2}} + f_2 \frac{\exp\left[-\frac{1}{2\tau_2^2}(\tau_i - a_2 \xi_i)^2\right]}{\sqrt{2\pi\tau_2^2}}, \quad (59)$$

where  $f_2 = 1 - f_1$ . The equations for the order parameters can be generalized from the single component case in a straightforward manner.

A case of interest is that the restoration agent operates on the assumption of the characteristics of the majority component of the channel, say the first component. Hence it operates at the ratio  $\beta_m/h = \beta_s \tau_1^2/a_1$ . Suppose the Gaussian noise is partly interrupted to take the characteristics of the second component, but the operation parameters cannot be adjusted

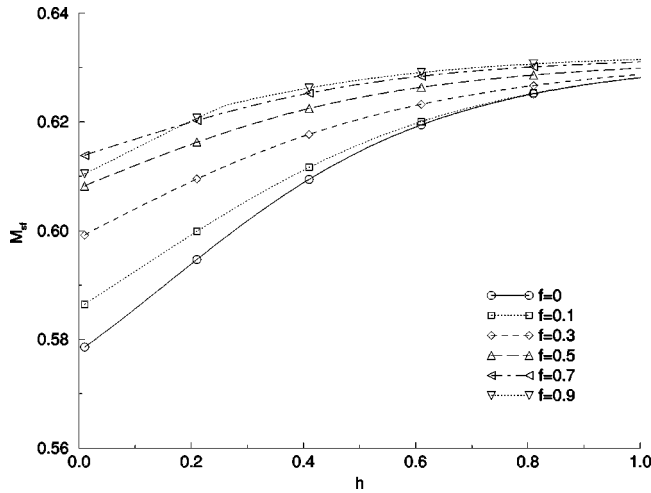


FIG. 9. The performance of selective freezing with two components of Gaussian noise at  $\beta_s = 1.05$ ,  $f_1 = 4f_2 = 0.8$ ,  $a_1 = 5a_2 = 1$ , and  $\tau_1 = \tau_2 = 1$ . The restoration agent operates by assuming the majority component, i.e.,  $\beta_m/h = \beta_s \tau_1^2/a_1$ .

soon enough; then there will be a degradation of the quality of the restored images. In the example in Fig. 9, the reduction of the overlap  $M_{sf}$  for selective freezing is much more modest than the one-stage process ( $f=0$ ).

An alternative situation is that the restoration agent is able to detect the changes in the average signal strengths and noise variance, but still operates on the assumption of a single-component Gaussian channel. Suppose that such simple statistics as  $\langle \text{sgn} \tau_i \rangle$ ,  $\langle \tau_i \rangle$  and  $\langle \tau_i^2 \rangle$  are accessible. Then the parameters  $m_0^*$ ,  $a^*$ , and  $\tau^*$  estimated by the restoration agent are obtained, for  $\tau_1 = \tau_2 = \tau$ , from the solutions of

$$m_0^* \text{erf} \frac{a^*}{\sqrt{2} \tau^*} = \langle \text{sgn} \tau_i \rangle = m_0 \left[ f_1 \text{erf} \frac{a_1}{\sqrt{2} \tau_1} + f_2 \text{erf} \frac{a_2}{\sqrt{2} \tau_2} \right], \quad (60)$$

$$m_0^* a^* = \langle \tau_i \rangle = m_0 [f_1 a_1 + f_2 a_2], \quad (61)$$

$$a^{*2} + \tau^{*2} = \langle \tau_i^2 \rangle = f_1 (a_1^2 + \tau_1^2) + f_2 (a_2^2 + \tau_2^2), \quad (62)$$

and  $\beta_s^* = \tanh^{-1} m_0^*/m_0^*$ . Using these estimated parameters, the performances in Fig. 10 improve over their counterparts based on only the majority component in Fig. 9. Still, one-stage restoration cannot avoid the performance drop when  $h$  vanishes, whereas correspondingly, selective freezing has a much more gentle drop in performance.

It is interesting to study the more realistic case of two-dimensional images, since we have so far presented analytical results for the mean-field model only. As confirmed by the results for Monte Carlo simulations in Fig. 11, the overlaps of selective freezing are much more steadier than that of the one-stage dynamics when the decoding temperature changes. This steadiness is most remarkable for a freezing fraction of  $f=0.9$ .

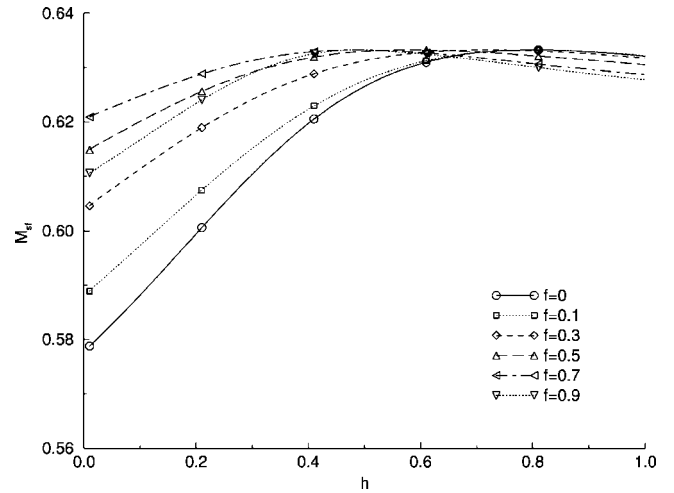


FIG. 10. Same as Fig. 9, except that the restoration agent operates with the ratio  $\beta_m/h = \beta_s^* \tau^{*2}/a^*$ , where  $\beta_s^*$ ,  $\tau^*$ , and  $a^*$  are estimated from Eqs. (60)–(62).

## V. DISCUSSION

We have introduced a multistage technique for error-correcting codes and image restoration, in which the information extracted from the former stage can be used selectively to improve the performance of the latter one. While the overlap  $M_{sf}$  of the selective freezing is bounded by the optimal performance of the one-stage dynamics derived in Ref. [5], it has the advantage of being tolerant to uncertainties in hyperparameter estimation. The performance is especially steady when the fraction of frozen spins, rather than the threshold of their thermodynamic averages, is fixed in the process. This is confirmed by both analytical and simulation results for mean-field and finite-dimensional models. As an example, we have illustrated its advantage of robustness when the noise distribution is composed of more than one Gaussian components, such as in the case of

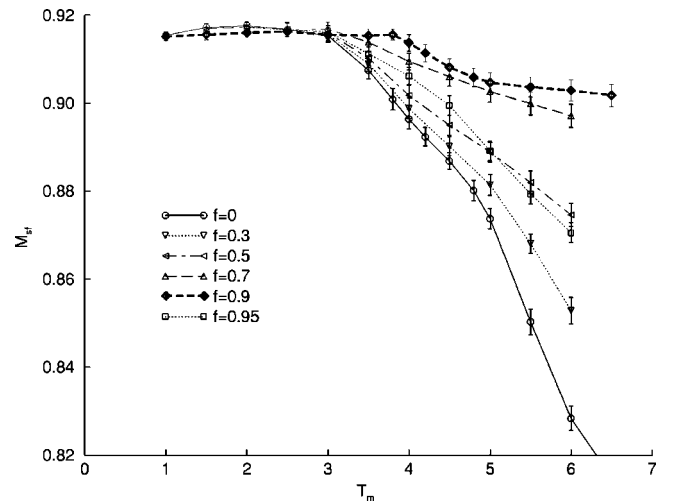


FIG. 11. Results of Monte Carlo simulations for the overlaps of selective freezing compared with those of the one-stage dynamics for two-dimensional images generated at the source prior temperature  $T_s = 2.15$ . The simulation parameters are  $N = 50 \times 50$ , with an initial overlap of 0.8 and 1000 samples. Each stage consist of 1000 Monte Carlo steps per node for averaging.

modern communication channels supporting multimedia applications.

We found that selective freezing is most useful when more than one hyperparameters have to be estimated, as illustrated by the example of image restoration, where both  $\beta_m$  and  $h$  have to be estimated. In the example of error-correcting codes discussed in Sec. III, there is only one hyperparameter  $T_m$ , and it is found that selective freezing has performance advantages only when  $T_m$  is chosen above the Nishimori point. However, more than one hyperparameter is often present in practical applications.

Selective freezing can be generalized to more than two stages, in which spins that remain relatively stable in one stage are progressively frozen in the following one. It is expected that the performance can be even more robust.

While the multistage process described here has a robust performance, it does not raise the critical temperature or the critical noise level for the existence of the ordered phase. Nor can it widen the basin of attraction for the ordered phase. Other multistage processes, proposed in Ref. [10] for neural networks, may be able to achieve this. This remains an area for further research.

We have made progress in the theoretical treatment of multistage processes using the cavity method. It allows the thermal averages of spins to be expressed in terms of the cavity fields. Since a cavity field is uncorrelated with the spin in consideration, it can in turn be expressed in terms of the means and covariances of the spin averages, thereby arriving at a set of self-consistent equations for the order parameters. In particular, there appears a trans-susceptibility term, since variations of the cavity field in the first stage are correlated with the spin average in the second stage due to the selective nature of the freezing process in the second stage. However, for the ordered phase considered in this paper, the effects of the trans-susceptibility term is not too large except near the phase boundary.

On the other hand, we have a remark about the basic assumption of the cavity method, namely, that the addition or removal of a spin causes a small change in the system describable by a perturbative approach. In fact, adding or removing a spin may cause the thermal averages of other spins to change from below to above the thresholds  $\pm\theta$  (or vice versa). This change, though often small, induces a non-negligible change of the thermal averages from fractional values to the frozen values of  $\pm 1$  (or vice versa) in the second stage. The perturbative analysis of these changes is only approximate. The situation is reminiscent of similar instabilities in other disordered systems such as the perceptron, and are equivalent to Almeida-Thouless instabilities in the replica method [13]. A full treatment of the problem would require the introduction of a rough energy landscape [13], or the replica symmetry breaking ansatz in the replica method [11]. Nevertheless, previous experiences on disordered systems showed that the corrections made by a more complete treatment may not be too large in the ordered phase. For example, corresponding analytical and simulational results in Figs. 1 and 6, respectively, are close to each other.

In practical implementations of error-correcting codes, algorithms based on belief-propagation methods, rather than Monte Carlo methods, are often employed [14]. It has recently been shown that such decoded messages converge to

the solutions of the TAP equations in the corresponding thermodynamic system [15]. Again, the performance of these algorithms are sensitive to the estimation of hyperparameters. We propose that the selective freezing procedure has the potential to make these algorithms more robust.

Incidentally, multistage dynamics has also been applied in the recently popular turbo codes [16]. Messages are coded in sequences with two possible permutations and at each iterative stage, the information derived from decoding one sequence is fed to the other in the form of external fields for each bit. The techniques developed in the present context can be used to study this iterative process.

## ACKNOWLEDGMENTS

K.Y.M.W. wishes to thank the Tokyo Institute of Technology for hospitality. H.N. is grateful to the Hong Kong University of Science and Technology for hospitality. This work was partially supported by research Grant No. HKUST6157/99P of the Research Grant Council of Hong Kong.

## APPENDIX A: THERMAL AVERAGES OF SPINS

In this appendix we derive Eq. (16) starting from the clustering property Eq. (15). For convenience we illustrate the derivation for  $p=2$ . We separate the Hamiltonian into two parts, one does not contain  $\sigma_1$  and the other does. Hence

$$H = H^{\setminus 1} - \beta \sum_{j>1} J_{1j} \sigma_1 \sigma_j. \quad (\text{A1})$$

Thus the thermal average can be written as

$$\langle \sigma_1 \rangle = \frac{\text{Tr}^{\setminus 1} e^{-H^{\setminus 1}} \text{Tr}_1 \sigma_1 \exp\left(\beta \sigma_1 \sum_j J_{1j} \sigma_j\right)}{\text{Tr}^{\setminus 1} e^{-H^{\setminus 1}} \text{Tr}_1 \exp\left(\beta \sigma_1 \sum_j J_{1j} \sigma_j\right)}. \quad (\text{A2})$$

Expanding the exponential function in the denominator and tracing over  $\sigma$ , we obtain

$$\text{Den.} = 2 \sum_{n \text{ even}} \frac{\beta^n}{n!} \sum_{j_1 \cdots j_n} J_{1j_1} \cdots J_{1j_n} \langle \sigma_{j_1} \cdots \sigma_{j_n} \rangle^{\setminus 1}. \quad (\text{A3})$$

Next, we use the clustering property to factorize the thermal average  $\langle \sigma_{j_1} \cdots \sigma_{j_n} \rangle^{\setminus 1}$ . For the coupling distribution specified by Eq. (14), only two kinds of contributions are significant in the summation over the indices  $j_1 \cdots j_n$ . In the first kind, an index  $j$  remains distinct from the rest, contributing a factor of  $J_{1j} \langle \sigma_j \rangle^{\setminus 1}$ . In the second kind, two indices become paired up. However, when  $j$  and  $k$  pair up, the thermal average  $\langle \sigma_j \sigma_k \rangle^{\setminus 1}$  becomes 1 instead of  $(\langle \sigma_j \rangle^{\setminus 1})^2$ . Hence the additional contribution due to the pairing is  $J_{1j}^2 [1 - (\langle \sigma_j \rangle^{\setminus 1})^2]$ . Other than these, the contributions due to the pairing of three or more indices are smaller by factors of  $N$ .

The denominator can now be considered as a summation over  $n$  and  $m$ , which are respectively the total number of

indices and the number of pairs of paired indices appearing in a term. The number of such terms is  $n!/m!2^m(n-2m)!$ . Hence

$$\text{Den.} = 2 \sum_{n \text{ even}} \sum_{m=0}^{n/2} \frac{\beta^n}{n!} \frac{n!}{m!2^m(n-2m)!} \times \left[ \sum_j J_{1j} \langle \sigma_j \rangle^{\setminus 1} \right]^{n-2m} \left\{ \sum_j J_{1j}^2 [1 - (\langle \sigma_j \rangle^{\setminus 1})^2] \right\}^m, \quad (\text{A4})$$

which can be simplified to

$$\text{Den.} = 2 \exp \left\{ \frac{1}{2} \beta^2 \sum_j J_{1j}^2 [1 - (\langle \sigma_j \rangle^{\setminus 1})^2] \right\} \times \cosh \left\{ \beta \sum_j J_{1j} \langle \sigma_j \rangle^{\setminus 1} \right\}. \quad (\text{A5})$$

Similarly, the numerator can be written as

$$\text{Num.} = 2 \exp \left\{ \frac{1}{2} \beta^2 \sum_j J_{1j}^2 [1 - (\langle \sigma_j \rangle^{\setminus 1})^2] \right\} \times \sinh \left\{ \beta \sum_j J_{1j} \langle \sigma_j \rangle^{\setminus 1} \right\}. \quad (\text{A6})$$

Substituting Eqs. (A5) and (A6) into Eq. (A2), we arrive at Eq. (16).

#### APPENDIX B: CHANGE IN THERMAL AVERAGES ON REMOVAL OF A SPIN

In this appendix we derive Eq. (29). For convenience we illustrate the derivation for  $p=2$ . We separate the Hamiltonian into four parts: (a) does not contain spins 1 and  $j$ , (b) contains only spins 1 and  $j$ , (c) contains spin 1 but not  $j$ , and (d) contains spin  $j$  but not 1. This yields

$$H = H^{\setminus 1j} - \beta J_{1j} \sigma_1 \sigma_j - \beta \sum_{k \neq 1j} J_{k1} \sigma_k \sigma_1 - \beta \sum_{k \neq 1j} J_{kj} \sigma_k \sigma_j. \quad (\text{B1})$$

The thermal average of  $\sigma_j$  can then be written as

$$\langle \sigma_j \rangle = \frac{\text{Tr}_{1j} \text{Tr}^{\setminus 1j} e^{-H} \sigma_j / \text{Tr}^{\setminus 1j} e^{-H^{\setminus 1j}}}{\text{Tr}_{1j} \text{Tr}^{\setminus 1j} e^{-H} / \text{Tr}^{\setminus 1j} e^{-H^{\setminus 1j}}}. \quad (\text{B2})$$

Using the mean-field technique developed in Appendix A, the denominator can be written as

$$\text{Den.} = \text{Tr}_{1j} \exp \left\{ \beta J_{1j} \sigma_1 \sigma_j + \beta \sum_{k \neq 1j} \langle \sigma_k \rangle^{\setminus 1j} (J_{k1} \sigma_1 + J_{kj} \sigma_j) + \frac{1}{2} \beta^2 \sum_{k \neq 1j} [1 - (\langle \sigma_k \rangle^{\setminus 1j})^2] \times (J_{k1} \sigma_1 + J_{kj} \sigma_j)^2 \right\}. \quad (\text{B3})$$

After collecting terms and discarding negligible ones,

$$\text{Den.} = \text{Tr}_{1j} \exp \left\{ \beta \sigma_1 \sum_{k \neq 1j} J_{1k} \langle \sigma_k \rangle^{\setminus 1j} + \beta \sigma_j \sum_{k \neq 1j} J_{jk} \langle \sigma_k \rangle^{\setminus 1j} + \beta J_{1j} \sigma_1 \sigma_j + \beta^2 (1-q) J^2 \right\}. \quad (\text{B4})$$

Together with a similar manipulation of the numerator, we obtain

$$\langle \sigma_j \rangle = \tanh \beta (h_j^{\setminus 1} + J_{j1} \tanh \beta h_1^{\setminus j}), \quad (\text{B5})$$

whose Taylor expansion yields

$$\langle \sigma_j \rangle = \langle \sigma_j \rangle^{\setminus 1} + (\beta \text{sech}^2 \beta h_j^{\setminus 1}) (J_{j1} \tanh \beta h_1^{\setminus j}), \quad (\text{B6})$$

which becomes Eq. (29) for the case  $p=2$ .

- 
- [1] R. J. McEliece, *The Theory of Information and Coding, Encyclopedia of Mathematics and its Applications* (Addison-Wesley, Reading, MA, 1977).
- [2] N. Sourlas, *Nature* (London) **339**, 693 (1989).
- [3] Y. Kabashima and D. Saad, *Europhys. Lett.* **45**, 97 (1999).
- [4] S. Geman and D. Geman, *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, 721 (1984).
- [5] H. Nishimori and K. Y. M. Wong, *Phys. Rev. E* **60**, 132 (1999).
- [6] H. Nishimori, *J. Phys. Soc. Jpn.* **62**, 2973 (1993).
- [7] Z. Zhou, R. M. Leathy, and J. Qi, *IEEE Trans. Image Process.* **6**, 844 (1997).
- [8] D. J. C. Mackay, *Neural Comput.* **4**, 415 (1992).
- [9] J. M. Pryce and A. D. Bruce, *J. Phys. A* **28**, 511 (1995).
- [10] K. Y. M. Wong, *Europhys. Lett.* **36**, 631 (1996).
- [11] M. Mézard, G. Parisi, and V. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore 1987).
- [12] K. Y. M. Wong, *Europhys. Lett.* **30**, 245 (1995).
- [13] K. Y. M. Wong, *Adv. Neural Inf. Proc. Syst.* **9**, 302 (1997).
- [14] B. J. Frey, *Graphical Models for Machine Learning and Digital Communication* (MIT Press, Cambridge, MA, 1998).
- [15] Y. Kabashima and D. Saad, *Europhys. Lett.* **44**, 668 (1998).
- [16] C. Berrou, A. Glavieux, and P. Thitimajshima, in *Proceedings of the IEEE International Conference of Communication, 1993* (IEEE, New York, 1993), p. 1064.