Image restoration using the *Q*-Ising spin glass

Jun-ichi Inoue
Complex Systems Engineering, Graduate School of Engineering, Hokkaido University, N13-W8, Kita-ku, Sapporo, 060-8628, Japan

Domenico M. Carlucci
Instituut voor Theoretische Fysica, K. U. Leuven, B-3001 Leuven, Belgium

(Received 27 June 2000; revised manuscript received 5 June 2001; published 29 August 2001)

We investigate static and dynamic properties of gray-scale image restoration by making use of the *Q*-Ising spin glass model, whose ladder symmetry allows us to take in account the distance between two spins. We thus give an explicit expression of the Hamming distance between the original and restored images as a function of the hyperparameters in the mean field limit. Finally, numerical simulations for real-world pictures are carried out to prove the efficiency of our model.

DOI: 10.1103/PhysRevE.64.036121 PACS number(s): 02.50.−r, 05.20.−y, 05.50.+q

I. INTRODUCTION

In the last decade, the problem of the image restoration (IR) has been successfully investigated by means of techniques borrowed from the field of statistical mechanics. Among them, it is certainly worth mentioning the maximum posterior marginal (MPM) and maximum *a posteriori* (MAP) estimations. From the statistical mechanical point of view, each recovered image within the MPM estimation can be regarded as the equilibrium state of ferromagnetic spin systems in the presence of random fields at finite temperature. In simple words, the reconstruction of a corrupted image is achieved by balancing the strength of a linear field, which carries the information of the degraded picture, and a ferromagnetic term that builds relatively large ''one-color'' clusters (below the transition temperature), thus suppressing the isolated pixels thought to be noise. From this point of view, the MAP estimation consists in the minimization of the same Hamiltonian at zero temperature (search for the ground state), with an appropriate scaling of the random field. The advantage of the MPM estimation over the MAP one has been pointed out by Marroquin *et al.* [1] and its performance has been investigated by several authors [2,3]. In this direction, Nishimori and Wong [4], by unifying IR problem and error-correcting code theory under a single framework, found that the optimal recovering of an image is obtained at a finite temperature (known as *Nishimori temperature* in the field of statistical mechanics). Their results, however, were restricted to the usual binary spin models (Ising), i.e., black or white images in IR jargon, and many questions about the properties of the gray-scale image restoration (GSIR) processes still remain open. A first attempt to generalize [4] to gray-level pictures has been carried out by the authors [5] by mapping the set of the pixels onto *Q*-state (chiral) Potts spins in the presence of the random fields. In that case, the symmetry of the Potts Hamiltonian (*hypertetrahedron*) reduces the problem to a two-state-*like* system, where only one bit turns out to be right, and all the others are equivalently wrong without any regard of the whole gray-level scale. Whereas this turns out to be an efficient method in the presence of white noise (each spin is flipped to any of the *Q* values with equal probability), things may be different from a transmission channel affected by Gaussian noise (the spin-flip probability distribution is a Gaussian). We thus investigate the performance (both static and dynamic) of the gray-scaled image restoration using the *Q*-Ising model [7], whose *ladder* symmetry takes into account the distance between the spin values and will allow us to say, for instance, that $Q=3$ is better than $Q=5$ if the right pixel corresponds to $Q=2$. The analytical expressions are obtained in so-called mean field limit [6], where each spin interacts with all the others. The efficiency of our model is checked by Monte Carlo simulations and iterative algorithm by using mean-field approximation. This paper is organized as follows. In Sec. II A, we introduce our model system using the infinite range model and explain the general formulation of the MPM estimation. In this paper, we focus our attention on the averaged case performance, instead of the performance for a specific choice *f* data. Therefore, in Sec. II B, we define the average of the macroscopic physical quantities over the data distribution. In Sec. II C we give an analytical expression of the Hamming distance by static calculation in the mean-field limit. In Sec. II D we derive the dynamical equations with respect to the macroscopic quantities, namely, the magnetization and the Hamming distance in terms of microscopic master equation. In Sec. III, in order to test the usefulness of the *Q*-Ising model for the GSIR, we carry out Monte Carlo simulations for real-world pictures with $Q=8$ gray-scale levels. In Sec. IV, we show an iterative algorithm based on the mean-field approximation, whose convergence is much faster than that of the Monte Carlo simulations. The last section is left for summary and discussion.

II. THE INFINITE RANGE *Q*-ISING SPIN GLASS MODEL

A. General formulation

A *Q*-gray-scale levels image is nothing but a set of pixels $\{\xi\}$ on a grid, whose values can be coded at each node as an integer variable $\xi_i \in \{1,2,\ldots,Q\}$. Without loss of generality, let our image be generated by the following prior distribution:

$$P_s(\{\xi\}) = \frac{1}{\mathcal{Z}_s} \exp\left[-\frac{\beta_s}{2} \sum_{ij} (\xi_i - \xi_j)^2 \right],$$

where $\mathcal{Z}_s$ is the usual normalization constant that is given by

$$\mathcal{Z}_s = \mathrm{Tr}_{\{\xi\}} \exp\left[-\frac{\beta_s}{2} \sum_{ij} (\xi_i - \xi_j)^2 \right].$$

In the spirit of statistical mechanics, we want to regard this picture as a snapshot of a spin system described by the Hamiltonian $\mathcal{H}_s \equiv (1/2)\Sigma_{ij}(\xi_i - \xi_j)^2$ at a specific temperature $T_s \equiv \beta_s^{-1}$. Sending our image through a noisy channel will cause the flipping of some pixels to different values. For this degrading process, we assume that each pixel ξ_i changes its state to τ_i independently. Then, the degraded pixel τ_i is given by the following conditional probability:

$$P(\tau_i|\xi_i) = \frac{1}{\sqrt{2\pi} a_\tau} \exp\left[-\frac{1}{2a_\tau^2} (\tau_i - \tau_0 \xi_i)^2 \right].$$

This means that after the transmission, the receiver observes τ_i that was violated from scaled original image $\tau_0 \xi_i$ with a standard deviation a_τ. This kind of damaging process is referred to as *Gaussian channel* (GC). Due to the independence of noisy process on each pixel, a sequence of original pixel $\{\xi\}$ is corrupted by the GC as

$$P(\{\tau\}|\{\xi\}) = \prod_i P(\tau_i|\xi_i) = \frac{1}{(\sqrt{2\pi} a_\tau)^N} \exp\left[-\frac{1}{2a_\tau^2} \sum_i (\tau_i - \tau_0 \xi_i)^2 \right].$$

In the context of Bayesian approach, the posterior distribution reads

$$\begin{aligned} P(\{\sigma\}|\{\tau\}) &= \frac{P(\{\tau\}|\{\sigma\})P(\{\sigma\})}{\mathrm{Tr}_{\{\sigma\}} P(\{\tau\}|\{\sigma\})P(\{\sigma\})} \\ &= \frac{\exp\left[-h\sum_i (\sigma_i - \tau_i)^2 - (\beta_d/2)\sum_{ij} (\sigma_i - \sigma_j)^2 \right]}{\mathrm{Tr}_{\{\sigma\}} \exp\left[-h\sum_i (\sigma_i - \tau_i)^2 - (\beta_d/2)\sum_{ij} (\sigma_i - \sigma_j)^2 \right]} \\ &\equiv \frac{\exp(-\mathcal{H}_{\mathrm{eff}})}{\mathcal{Z}_d}, \end{aligned} \tag{1}$$

where we defined the effective Hamiltonian $\mathcal{H}_{\mathrm{eff}}$ and the normalization constant $\mathcal{Z}_d$ as

$$\mathcal{H}_{\mathrm{eff}} \equiv h\sum_i (\sigma_i - \tau_i)^2 + \frac{\beta_d}{2} \sum_{ij} (\sigma_i - \sigma_j)^2 \tag{2}$$

and

$$Z_d \equiv \mathrm{Tr}_{\{\sigma\}} \exp[-\mathcal{H}_{\mathrm{eff}}],$$

respectively. The parameters h and β_d appearing in the Hamiltonian $\mathcal{H}_{\mathrm{eff}}$ [Eq. (2)] are referred to as *hyperparameters* and we cannot mention about the true values of them beforehand. This posterior distribution $P(\{\sigma\}|\{\tau\})$ is constructed in terms of a likelihood $P(\{\tau\}|\{\sigma\})$ and a prior probability $P(\{\sigma\})$ as we saw in Eq. (1). $P(\{\tau\}|\{\sigma\})$ and $P(\{\sigma\})$ are given by

$$P(\{\tau\}|\{\sigma\}) = \frac{\exp\left[-h\sum_i (\tau_i - \sigma_i)^2 \right]}{\mathrm{Tr}_{\{\tau\}} \exp\left[-h\sum_i (\tau_i - \sigma_i)^2 \right]}$$

and

$$P(\{\sigma\}) = \frac{\exp\left[-(\beta_d/2)\sum_{ij} (\sigma_i - \sigma_j)^2 \right]}{\mathrm{Tr}_{\{\sigma\}} \exp\left[-(\beta_d/2)\sum_{ij} (\sigma_i - \sigma_j)^2 \right]}.$$

The prior probability reflects our assumption on the original image that the picture should be locally smooth. As briefly mentioned in the introduction, the MAP estimation consists in maximizing the above posterior probability $P(\{\sigma\}|\{\tau\})$, that is finding the ground state $\{\sigma\}$ of the effective Hamiltonian $\mathcal{H}_{\mathrm{eff}}$ and regarding it as an estimate of true pixels.

On the other hand, in the context of the MPM estimation, we first consider the following marginal distribution:

$$P(\sigma_i|\{\sigma\}) = \sum_{\sigma \neq \sigma_i} P(\{\sigma\}|\{\tau\})$$

and then we calculate the local magnetization that is given by

$$\begin{aligned} \langle \sigma_i \rangle_{\beta_d, h} &\equiv \sum_{\sigma_i=1}^{Q} \sigma_i P(\sigma_i|\{\tau\}) = \frac{\mathrm{Tr}_{\{\sigma\}} \sigma_i \sum_{\sigma \neq \sigma_i} \exp\left[-h\sum_i (\tau_i - \sigma_i)^2 - (\beta_d/2)\sum_{ij} (\sigma_i - \sigma_j)^2 \right]}{\mathrm{Tr}_{\{\sigma\}} \exp\left[-h\sum_i (\tau_i - \sigma_i)^2 - (\beta_d/2)\sum_{ij} (\sigma_i - \sigma_j)^2 \right]} \\ &= \frac{\mathrm{Tr}_{\{\sigma\}} \sigma_i \exp\left[-h\sum_i (\tau_i - \sigma_i)^2 - (\beta_d/2)\sum_{ij} (\sigma_i - \sigma_j)^2 \right]}{\mathrm{Tr}_{\{\sigma\}} \exp\left[-h\sum_i (\tau_i - \sigma_i)^2 - (\beta_d/2)\sum_{ij} (\sigma_i - \sigma_j)^2 \right]} = \frac{\mathrm{Tr}_{\{\sigma\}} \sigma_i \exp[-\mathcal{H}_{\mathrm{eff}}]}{\mathrm{Tr}_{\{\sigma\}} \exp[-\mathcal{H}_{\mathrm{eff}}]}. \end{aligned} \tag{3}$$

Using the above expectation value, we regard the estimate of the original pixel ξ_i as $\Omega(\langle\sigma_i\rangle_{\beta_d,h})$, where function Ω is represented by a sum of step functions $\Theta(x)$ [$\Theta(x)=1$ for $x\geqslant 0$ and $\Theta(x)=0$ for $x<0$];

$$\Omega(\langle\sigma_i\rangle_{\beta_d,h})\equiv\sum_{k=1}^{Q} k\left[\Theta\left(\langle\sigma_i\rangle_{\beta_d,h}-\frac{2k-1}{2}\right)-\Theta\left(\langle\sigma_i\rangle_{\beta_d,h}-\frac{2k+1}{2}\right)\right].$$

The natural quantity measuring the quality of our restoration process, viz., the distance between the original and the recovered image, is the Hamming distance (square error)

$$D_{\mathrm{H}}(\beta_d,h)\equiv\frac{1}{2N}\sum_i\left[\xi_i-\Omega(\langle\sigma_i\rangle)_{\beta_d,h}\right]^2,$$

whose value depends upon the hyperparameters, h,β_d, appearing in the effective Hamiltonian $\mathcal{H}_{\mathrm{eff}}$. At this stage, it is important to bear in mind that the MAP estimate is recovered as the limit $\beta_d\to\infty$ (keeping their ratio constant $H\equiv h/\beta_d$) in Eq. (3). Encouraged by the results in [4] and [5], we expect that more data fed through the noisy channel improve the quality of the restored image, since the receiver will have more information about the original image. Especially, inspired by theory of error-correcting codes, in addition to the transmission of a sequence of pixels $\{\xi\}\equiv(\xi_1,\ldots,\xi_N)$ we also send the products of two original pixels, namely, $\xi_i\xi_j$, $(i,j=1,\ldots,N;i\neq j)$. Of course, we can use another kind of extra term like $(\xi_i\xi_j)^2$ or $\xi_i\xi_j\xi_k$; we use here $\xi_i\xi_j$ as a simplest form. In Sec. III, as another candidate, we test the efficiency of the extra information $(\xi_i-\xi_j)^2$, $(i,j=1,\ldots,N;i\neq j)$, namely, square distances of two pixels by Monte Carlo simulations.

For this simplest choice of extra information, each product $\xi_i\xi_j$ is also corrupted independently by the following GC:

$$P(J_{ij}|\xi_i\xi_j)=\frac{1}{\sqrt{2\pi}a_J}\exp\left[-\frac{1}{2a_J^2}(J_{ij}-J_0\xi_i\xi_j)^2\right],\quad(4)$$

namely, the degraded version of the product J_{ij} deviated from the scaled original data $J_0\xi_i\xi_j$ with width a_J. For this degrading process, we modify a likelihood $P(\{\tau\}|\{\sigma\})$ in Eq. (1) as

$$P(\{\tau\},\{J\}|\{\sigma\})=\frac{\exp\left[-(\beta_J/2)\sum_{ij}(J_{ij}-\sigma_i\sigma_j)^2-h\sum_i(\tau_i-\sigma_i)^2\right]}{\mathcal{Z}'_L}$$

with

$$\mathcal{Z}'_L\equiv\mathrm{Tr}_{\{\tau\},\{J\}}\exp\left[-\frac{\beta_J}{2}\sum_{ij}(J_{ij}-\sigma_i\sigma_j)^2-h\sum_i(\tau_i-\sigma_i)^2\right],$$

where we introduced another hyperparameter β_J. Using the same way as Eq. (1), we rewrite the posterior probability as

$$P(\{\sigma\}|\{\tau\},\{J\})=\frac{\exp[-\mathcal{H}_{\mathrm{eff}}]}{\mathrm{Tr}_{\{\sigma\}}\exp[-\mathcal{H}_{\mathrm{eff}}]}$$

with the following effective Hamiltonian

$$\mathcal{H}_{\mathrm{eff}}=\frac{\beta_J}{2}\sum_{ij}(J_{ij}-\sigma_i\sigma_j)^2-h\sum_i(\tau_i-\sigma_i)^2-\frac{\beta_d}{2}\sum_{ij}(\sigma_i-\sigma_j)^2.\quad(5)$$

B. Averaged case performance

Given the degraded version of data, namely, $\{\tau\}$ and $\{J\}$, arbitrary macroscopic physical quantity $f(\{\sigma\},\{\tau\},\{J\})$ is calculated in terms of the average over the posterior distribution $P(\{\sigma\}|\{\tau\},\{J\})$ as

$$\langle f(\{\sigma\},\{\tau\},\{J\})\rangle_{\beta_d,h}\equiv\mathrm{Tr}_{\{\sigma\}}f(\{\sigma\},\{\tau\},\{J\})P(\{\sigma\}|\{\tau\},\{J\})=\frac{\mathrm{Tr}_{\{\sigma\}}f(\{\sigma\},\{\tau\},\{J\})e^{-\mathcal{H}_{\mathrm{eff}}}}{\mathrm{Tr}_{\{\sigma\}}e^{-\mathcal{H}_{\mathrm{eff}}}}.\quad(6)$$

In the practical applications, for a given data $\{\tau\},\{J\}$, we estimate ith pixel, for example, by means of the MPM estimate as $\mathrm{sgn}(\langle\sigma_i\rangle_{\beta_d,h})$. Obviously, this estimate depends on the data $\{\tau\},\{J\}$. However, our interest here is not the performance for a specific data set $\{\tau\},\{J\}$ but the averaged performance. Therefore, when the quantity $\langle f(\{\sigma\},\{\tau\},\{J\})\rangle_{\beta_d,h}$ depends on the observed data $\{\tau\},\{J\}$, we should average them out by the distribution

$$P(\{\tau\},\{J\}|\{\xi\})=\frac{\exp\left[-\frac{1}{2a_J^2}\sum_{ij}(J_{ij}-J_0\xi_i\xi_j)^2-(1/2a_\tau^2)\sum_i(\tau_i-\xi_i)^2-\frac{\beta_s}{2}\sum_{ij}(\xi_i-\xi_j)^2\right]}{\mathrm{Tr}_{\{\tau\},\{J\},\{\xi\}}\exp\left[-\frac{1}{2a_J^2}\sum_{ij}(J_{ij}-J_0\xi_i\xi_j)^2-(1/2a_\tau^2)\sum_i(\tau_i-\xi_i)^2-\frac{\beta_s}{2}\sum_{ij}(\xi_i-\xi_j)^2\right]}$$

to obtain the averaged performance of our method. Thus, the averaged macroscopic quantity is given by

$$[\langle f(\{\sigma\},\{\tau\},\{J\})\rangle_{\beta_d,h}]_{\{\tau\},\{J\},\{\xi\}} \equiv \mathrm{Tr}_{\{\tau\},\{J\},\{\xi\}}\left[\frac{\mathrm{Tr}_{\{\sigma\}} f(\{\sigma\},\{\tau\},\{J\})e^{-\mathcal{H}_{\mathrm{eff}}}}{\mathrm{Tr}_{\{\sigma\}}e^{-\mathcal{H}_{\mathrm{eff}}}}\right]P(\{\tau\},\{J\}|\{\xi\}). \tag{7}$$

Using this definition, the performance of image restoration is measured by the following averaged Hamming distance between the original image ξ_i and the restored one, that is, $\Omega(\langle\sigma_i\rangle_{\beta_d,h})$ as

$$D_{\mathrm{H}} \equiv \mathrm{Tr}_{\{\tau\},\{J\},\{\xi\}}\left[\frac{\mathrm{Tr}_{\{\sigma\}}(1/2N)\sum_i [\xi_i-\Omega(\langle\sigma_i\rangle_{\beta_d,h})]^2 e^{-\mathcal{H}_{\mathrm{eff}}}}{\mathrm{Tr}_{\{\sigma\}} e^{-\mathcal{H}_{\mathrm{eff}}}}\right] \times P(\{\tau\},\{J\}|\{\xi\}). \tag{8}$$

In the next two subsections, we investigate the performance of image restoration in terms of this Hamming distance D_H. We focus our analysis not only on the static properties but also the dynamic properties of image restoration.

C. Static properties

In this subsection, we consider the static properties of image restoration. First of all, we should investigate the properties of original image, that is to say, the properties of the ferromagnetic Q-Ising model. However, it is quite hard to calculate the partition function or the other physical quantities for our spin system defined on two-dimensional square lattice analytically. Therefore, in this paper, we investigate the infinite range version of our model system and calculate the macroscopic physical quantities analytically. Then, the infinite range version of the prior distribution leads to

$$P_s(\{\xi\}) = \frac{1}{\mathcal{Z}_s}\exp\left[-\frac{\beta_s}{2N}\sum_{ij}(\xi_i-\xi_j)^2\right],$$

where we should notice that the argument of the exponential should be divided by N in order to take a proper thermodynamic limit. For this rather artificial model, we easily obtain the magnetization at some temperature $T_s(=\beta_s^{-1})$ as follows:

$$m_0 \equiv \frac{1}{N}\sum_i \xi_i = \frac{\mathrm{Tr}_\xi \xi \exp(2m_0\beta_s\xi-\beta_s\xi^2)}{\mathcal{Z}_s}.$$

We should keep in mind that for the infinite range Q-Ising model, the properties of the macroscopic quantities of the system are completely determined by m_0. In Fig. 1, we plot the magnetization m_0 as a function of source temperature T_s for $Q=3$ and $Q=4$. We see that for the $Q=3$ case the three states $m_0=1,\ 2$, and 3 are degenerated at $T_s=0$, while at finite temperature the middle state $m_0=2$ becomes globally stable and $m_0=1,3$ are degenerated locally stable. At high-temperature regime $T_s\to\infty$, each spin takes all the values with the same probability 1/3 and thus the corresponding magnetization is $m_0=(1+2+3)/3=2$. The transition between the ferromagnetic phase and the paramagnetic phase occurs at $T_c\sim 1.0$. In the same way as the case of $Q=3$, for $Q=4$, the four states $m_0=1,\ 2,\ 3$, and 4 are degenerated at $T_s=0$, and the middle two states $m_0=2$ and 3 become globally stable for $T_s>0$ ($m_0=1,4$ are degenerated locally stable states). The paramagnetic state is specified by the magnetization $m_0=(1+2+3+4)/4=2.5$ and the ferro-para transition occurs at $T_c\sim 1.78$. For this original image, in order to investigate the average performance of the MPM estimation, we should calculate D_H in terms of statistical mechanics of the spin system $\{\sigma\}$ with quenched disorder $\{\tau\}$, $\{J\}$, and $\{\xi\}$. For this purpose, we calculate the averaged free energy of the system described by $\mathcal{H}_{\mathrm{eff}}$ [Eq. (5)] with assistance of the replica method;

$$[\ln\mathcal{Z}]_{\{\tau\},\{J\},\{\xi\}} = \lim_{n\to 0}\frac{[\mathcal{Z}^n]_{\{\tau\},\{J\},\{\xi\}}-1}{n}$$

with

$$\mathcal{Z} = \mathrm{Tr}_{\{\sigma\}}\exp(-\beta\mathcal{H}_{\mathrm{eff}}),$$

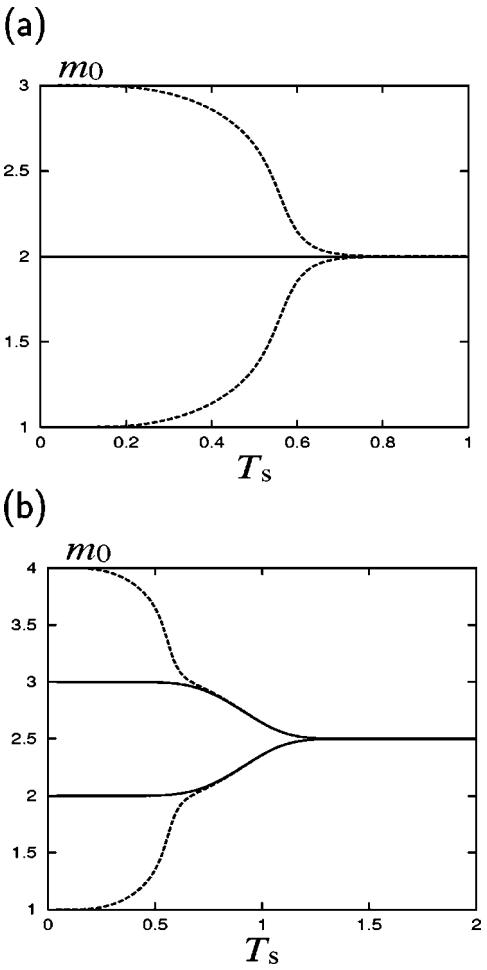

FIG. 1. The magnetization of the original image for the case of $Q=3$ (a) and $Q=4$ (b). The solid lines correspond to globally stable solutions.

which consist in replacing the quenched average of a single system with an annealed average of n replicated systems (letting $n\to 0$ at the end). Assuming a replica symmetric ansatz and by using the saddle point method, the order parameters are given by the following coupled equations:

$$m=[\langle\sigma_i^\alpha\rangle_{\beta_d,h}]_{\{\tau\},\{J\},\{\xi\}}=\mathrm{Tr}_\xi\mathcal{Q}(\xi)\int_{-\infty}^{\infty}\mathcal{D}x\,\mathcal{B}(x,\xi), \quad (9)$$

$$t=[\xi_i\langle\sigma_i^\alpha\rangle_{\beta_d,h}]_{\{\tau\},\{J\},\{\xi\}}=\mathrm{Tr}_\xi\,\xi\mathcal{Q}(\xi)\int_{-\infty}^{\infty}\mathcal{D}x\,\mathcal{B}(x,\xi), \quad (10)$$

$$q=[\langle\sigma_i^\alpha\rangle_{\beta_d,h}\langle\sigma_i^\beta\rangle_{\beta_d,h}]_{\{\tau\},\{J\},\{\tau\}}=\mathrm{Tr}_\xi\,\mathcal{Q}(\xi)\int_{-\infty}^{\infty}\mathcal{D}x[\mathcal{B}(x,\xi)]^2, \quad (11)$$

$$w=[\langle(\sigma_i^\alpha)^2\rangle_{\beta_d,h}]_{\{\tau\},\{J\},\{\xi\}}=\mathrm{Tr}_\xi\,\mathcal{Q}(\xi)\int_{-\infty}^{\infty}\mathcal{D}x\,\mathcal{C}(x,\xi), \quad (12)$$

where we should remember that the brackets $\langle\cdots\rangle_{\beta_d,h}$ and $[\cdots]_{\{\tau\},\{J\},\{\xi\}}$ are defined by Eq. (6) and Eq. (7), respectively. In the above expressions, $\mathcal{D}x\equiv(dx/\sqrt{2\pi})e^{-x^2/2}$ is the usual Gaussian measure and we define

$$\mathcal{B}(x,\xi)\equiv\frac{\mathrm{Tr}_\sigma\,\sigma\exp[U\sigma-V\sigma^2]}{\mathrm{Tr}_\sigma\exp[U\sigma-V\sigma^2]},$$

$$\mathcal{C}(x,\xi)\equiv\frac{\mathrm{Tr}_\sigma\,\sigma^2\exp[U\sigma-V\sigma^2]}{\mathrm{Tr}_\sigma\exp[U\sigma-V\sigma^2]},$$

$$\mathcal{Q}(\xi)\equiv\frac{\exp(2m_0\beta_s\xi-\beta_s\xi^2)}{\mathcal{Z}(\beta_s)},$$

with

$$U/2\equiv(\tau_0h+\beta_JJ_0t)\xi+m\beta_d+x\sqrt{(a_\tau h)^2+(a_J\beta_J)^2q},$$

$$V\equiv h+\beta_d+\beta_Jw.$$

Using the same method as the derivation of the order parameters, the averaged Hamming distance Eq. (8) is calculated and reads

$$D_{\mathrm{H}}(\beta_d,h,\beta_J)=\mathrm{Tr}_\xi\,\mathcal{Q}(\xi)\int_{-\infty}^{\infty}\mathcal{D}x[\xi-\Omega(\mathcal{B}(x,\xi))]^2.$$

It is straightforward to check that the above equations coincide with those in [4] for the $Q=2$ case. To keep things simple, we first assume that there is no glassy term in our decoding process, i.e., $\beta_J=0$. In Fig. 2, we plot the Hamming distance as a function of the decoding temperature for $Q=3$ and $T_s=0.75$. The minimum is reached at $T_d=0.75$. The same for $Q=4$ in Fig. 3. It can be shown numerically, at least for $Q=3$ and $Q=4$, that, given the original image at temperature T_s, just below the transition temperature, the optimal decoding temperature $T_d^{(\mathrm{opt})}=T_s$. The same relation turns out to be satisfied for black or white IR [4] (which corresponds to $Q=2$), differently from GSIR by the Potts model [5]. In order to compare the performance of the MPM estimate with that of the MAP one, we first investigate the scaled field $H\equiv h/\beta_d$ dependence of the MAP estimate. The MAP estimate is obtained by controlling the temperature as $T_d\to 0$ with keeping the scaled field H constant. Therefore, the Hamming distance for the MAP estimate should depend on H. In Fig. 4(a), we plot the Hamming distance of the MAP estimate as a function of H. In this figure, we set $Q=3$, $T_s=0.75$, and $a_\tau=\tau_0=1.0$. We see that the Hamming distance takes its minimum at $H_{\mathrm{opt}}=\tau_0/2a_\tau^2\beta_s=0.375$, namely,

$$D_{\mathrm{H}}(T_d=0,H)\geqslant D_{\mathrm{H}}(T_d=0,\tau_0/2a_\tau^2\beta_s).$$

This optimal value of the scaled field $H=h/\beta_d$ is obtained when we set $P(\{\tau\}|\{\xi\})=P(\{\tau\}|\{\sigma\})$ and $P_s(\{\xi\})=P_d(\{\sigma\})$, that is to say, $\beta_d=\beta_s$ and $h=\tau_0/2a_\tau^2$. In Fig. 4(b), we increase the temperature T_d keeping $H=H_{\mathrm{opt}}$ and plot the Hamming distance $D_H(T_d,H_{\mathrm{opt}})$ as a function of T_d. This figure shows that $D_{\mathrm{H}}(T_d,H_{\mathrm{opt}})$ takes its maximum at $T_d=T_s=0.75$. Therefore, we conclude that the MPM estimate achieves the lowest Hamming distance that cannot be obtained by the MAP estimation. In Fig. 4(b), we plot the $D_{\mathrm{H}}(T_d,H)$ for several values of H. From this figure, we see that as long as we choose H so as to satisfy $H\geqslant H_{\mathrm{opt}}=\tau_0/2a_\tau^2\beta_s$, the minimum value of the Hamming distance does not change.

In the limit of $T_d\to\infty$, each pixel takes $\sigma_i=1,2,3$ with the same probability 1/3, and the local magnetization leads to $\langle\sigma_i\rangle=(1+2+3)/3=2$ for all pixels i. As a result, the Hamming distance in the high-temperature limit becomes

$$D_{\mathrm{H}}(T_d\to\infty)=\frac{\sum_{\xi=1}^{3}(2-\xi)^2\exp(2m_0\beta_s\xi-\beta_s\xi^2)}{2\sum_{\xi=1}^{3}\exp(2m_0\beta_s\xi-\beta_s\xi^2)}\sim 0.1726.$$

This asymptotic behavior is checked in Fig. 4(b). We now switch on the product interaction, that is, $\beta_J\neq 0$ setting T_d and H at their optimal values. As clearly shown in Fig. 5, the performance of the restoration is dramatically improved. In this figure, the point $\beta_J=0$ corresponds to the minimum of the Hamming distance in Fig. 2.

D. Dynamics

An important and interesting problem is to determine the basin of attraction of the Hamming distance $D_{\mathrm{H}}(t)$. In fact, because of the presence of locally stable states, the final state of the restoration process is strongly dependent upon the initial condition of the dynamics. In addition, as the number of locally stable states increases with the number of the finite gray-scale levels, it becomes crucial to choose the initial state appropriately. However, as it is well known, it is difficult to treat the dynamics of spin system explicitly in finite dimension, especially, dynamics in two dimension that is the

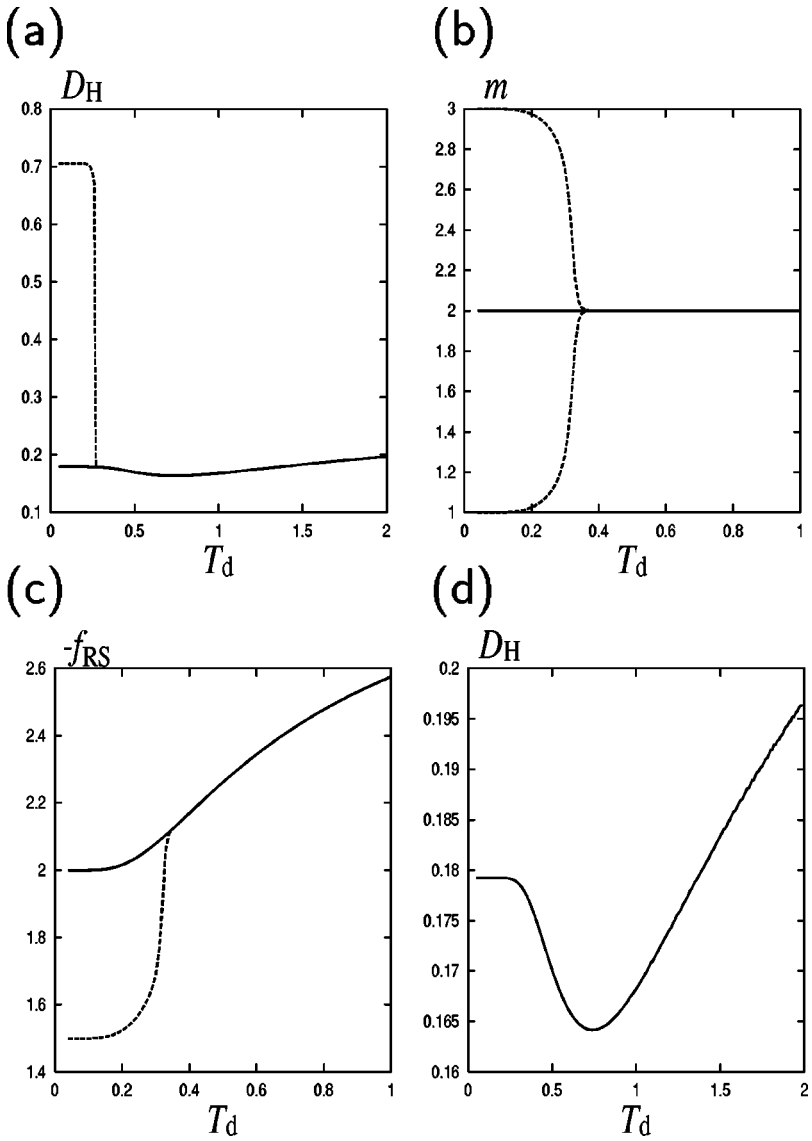

FIG. 2. The Hamming distances without glassy term ($\beta_J=0$) for the case of $Q=3$. (d) is obtained by expanding (a) around its minimum. The magnetization m and corresponding free energy $-f_{RS}$ are plotted in (b) and (c), respectively. Here, the dotted lines and the solid lines are locally stable states and globally stable states, respectively.

case of image restoration. In the previous section, we introduced the infinite range model and solved it analytically. Using this model, we derived the properties of image restoration and as we see in the next section, the results do not contradict qualitatively with the properties in two dimension. With this fact in mind, we also use the infinite range model to investigate the dynamical properties of image restoration. In the equilibrium limit $t\to\infty$, without glassy term, namely, $\beta_J=0$, the properties of image restoration in the infinite range model are completely written by magnetization m. Therefore, we assume that the dynamics of image restoration is also expressed by the time evolution of the magnetization $m(t)$. Therefore, we derive the differential equations with respect to the macroscopic variables, namely, $m(t)$ and $D_H(t)$, from the microscopic master equation. For the sake of simplicity, we restrict ourselves to the case without the glassy term. The master equation of our system leads to

$$\frac{dp_t(\{\sigma\})}{dt}=\sum_{k=1}^{N}\sum_{\sigma_{k'}=1}^{Q}[w(F_k\{\sigma\})p_t(F_k\{\sigma\})-w(\{\sigma\})p_t(\{\sigma\})]$$

with the following transition rate:

$$w(\{\sigma\})\equiv w(\sigma_k\to\sigma_{k'})=\frac{\exp[-(h+\beta_d)\sigma_{k'}^2+2(m\beta_d+h\tau_k)\sigma_{k'}]}{\sum_{\sigma_{k'}=1}^{Q}\exp[-(h+\beta_d)\sigma_{k'}^2+2(m\beta_d+h\tau_k)\sigma_{k'}]}$$

and $w(F_k\{\sigma\})\equiv w(\sigma_k\to\sigma_{k'})$, where we defined $\{\sigma\}\equiv(\sigma_1,\ldots,\sigma_k,\ldots,\sigma_N)$ and single pixel changing operator F_k that works as $F_k\{\sigma\}=\{\sigma\}'\equiv(\sigma_1,\ldots,\sigma_{k'},\ldots,\sigma_N)$.

By introducing the probability distribution of the magnetization m, viz.,

$$\mathcal{P}_t(m)\equiv\sum_{\sigma}p_t(\{\sigma\})\delta[m-m(\{\sigma\})]$$

and after some algebra we obtain the following differential equation analytically:

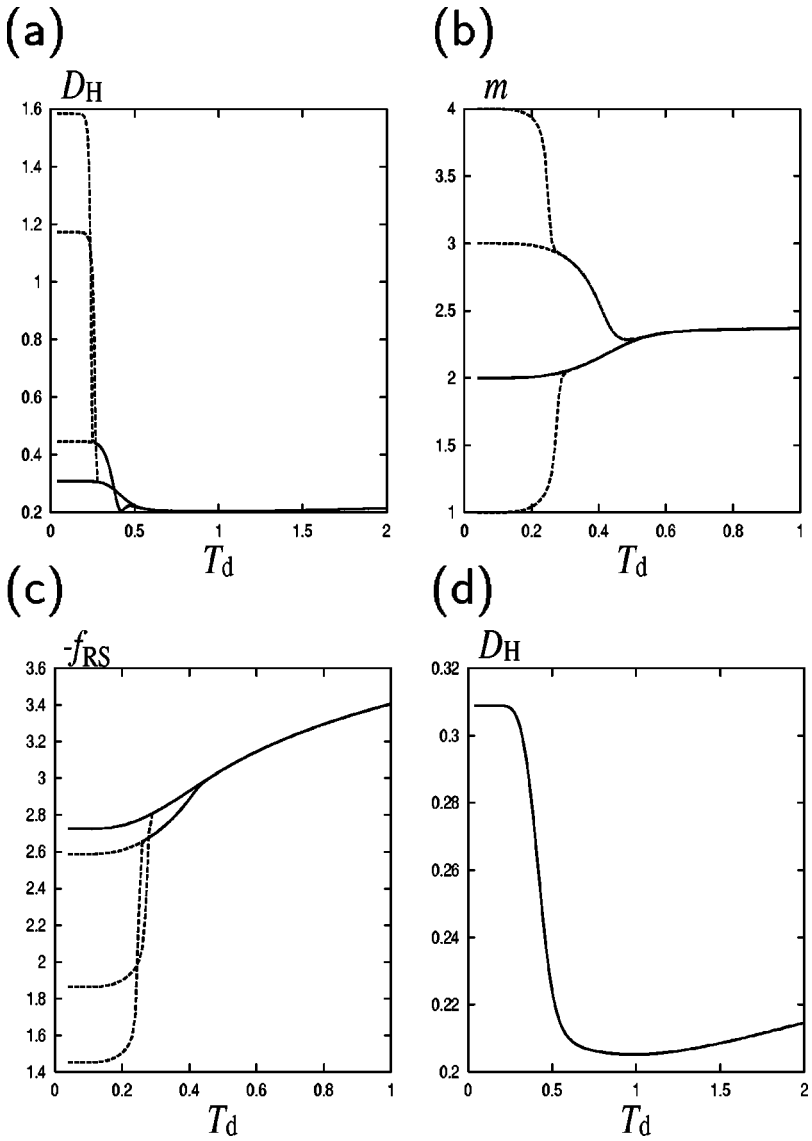

FIG. 3. The Hamming distances without glassy term ($\beta_J=0$) for the case of $Q=4$. (d) is obtained by expanding (a) around its minimum. The magnetization m and corresponding free energy $-f_{\mathrm{RS}}$ are plotted in (b) and (c), respectively.

$$\frac{dm}{dt}=-m+\sum_{\xi=1}^{Q}\mathcal{Q}(\xi)\int_{-\infty}^{\infty}Dx\,\mathcal{B}(x,\xi)\big|_{\beta_J=0}.$$

The derivation of the above differential equation is reported in Appendix A. The time evolution of the Hamming distance $D_{\mathrm{H}}(t)$ is obtained by substituting the time dependence of the magnetization $m(t)$ into $D_{\mathrm{H}}(m)$. In Fig. 6, we plot the time evolutions of the $Q=3$ Hamming distance for $T_d=T_s$ (a) and $T_d\neq T_s$ (b). From these figures, we see that if we choose the hyperparameter T_d so as to satisfy the relationship $T_d=T_s$, the Hamming distance converges to its optimal value for any initial condition. On the other hand, for $T_d\neq T_s$, there exists a threshold of the initial value of the Hamming distance $D_{\mathrm{H}}^{(c)}$ beyond which the flow $D_{\mathrm{H}}(t)$ does not converge to its optimal value. As the dynamical equation (with respect to m) is exactly the same as the time dependent Ginsburg-Landau (TDGL) equation, that is, $dm/dt=-\partial f_{\mathrm{RS}}/\partial m$, where RS is replica symmetric, the nature of the dynamics is intuitively understood as a steepest descent to a local minimum of the free energy. In fact, from Figs. 2(a)–2(c), we see that for $T_d<0.35$, there exist locally stable states. Therefore if we fail to choose the initial condition appropriately, the Hamming distance converges to the nonoptimal values. For practical situations, the corrupted image corresponds to our initial state. For the case $T_d\neq T_s$, we calculate the Hamming distance $D_{\mathrm{H}}^{(1)}$ between the original image and the noised one, which reads

$$D_{\mathrm{H}}^{(1)}\equiv\frac{1}{2N}\sum_i(\tau_i-\xi_i)^2$$

$$=\frac{\sum_{\xi=1}^{Q}[a_\tau^2+(\tau_0^2-1)\xi^2]\exp(2m_0\beta_s\xi-\beta_s\xi^2)}{2\mathcal{Z}(\beta_s)}.$$

In particular, for $a_\tau=\tau_0=1.0$, this leads to $D_{\mathrm{H}}^{(1)}=0.5$. From Fig. 5(b), we see that if we choose the corrupted image as an initial state, we destroy the observed corrupted image and the result is even worse.

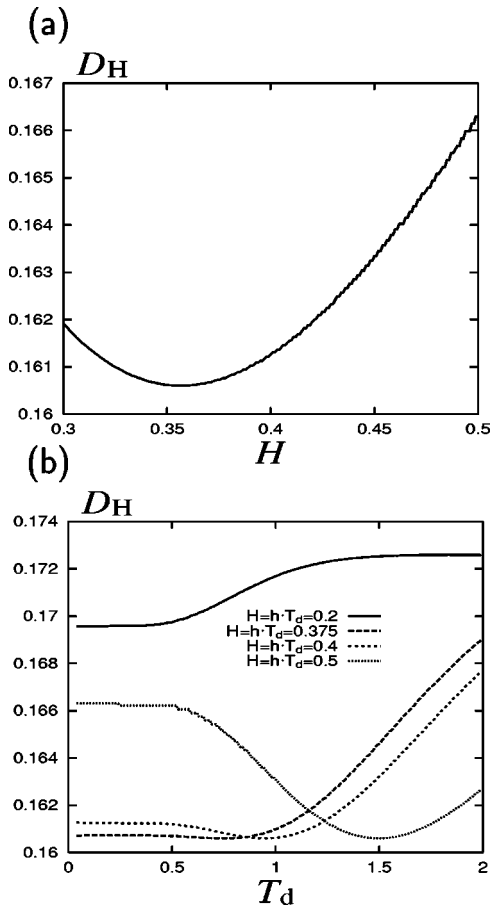

FIG. 4. The Hamming distance D_{H} of the MAP estimate for the case of $Q=3, T_s=0.75, \tau_0=a_\tau=1.0$ (a). D_{H} is plotted as a function of the scaled field $H=h/\beta_d$. We see that the minimum of D_{H} appeared at $H=H_{\mathrm{opt}}=\tau_0/2a_\tau^2\beta_s=0.375$. The Hamming distance of the MPM estimate is plotted in (b) as a function of the temperature T_d for several values of H. The figure shows that the minimum of the MPM estimate with $H=H_{\mathrm{opt}}$ is lower than that of the MAP estimate with $H=H_{\mathrm{opt}}$.

The asymptotic expressions of the magnetization and the Hamming distance in the limit of $t\to\infty$ lead to

$$m=m_*+[m(t=0)-m_*]e^{-t/t_0},$$

$$D_{\mathrm{H}}=D_{\mathrm{H}}(m_*)+\tilde{D}_{\mathrm{H}}e^{-t/t_0},$$

where m_* is a solution of the saddle point equations (9)–(12) with $\beta_J=0$ in the previous subsection. The relaxation time t_0 is given as

$$\frac{1}{t_0}=1+2\beta_d\int_{-\infty}^{\infty}\mathcal{D}x\,\mathcal{Q}(\xi)[\overline{\sigma^2}-(\overline{\sigma})^2]$$

with

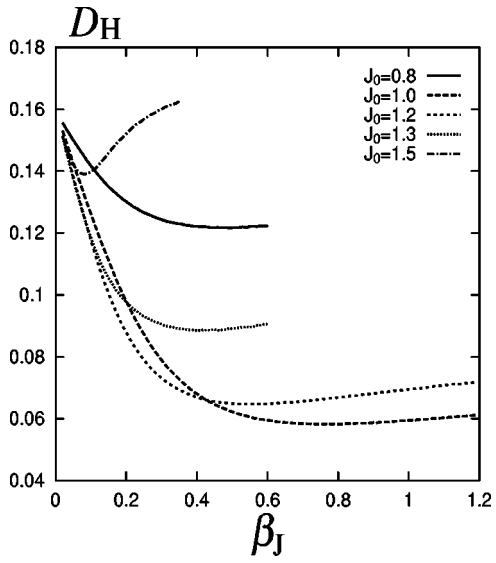

FIG. 5. The Hamming distance D_{H} as a function of the strength of the glassy term β_J for several values of J_0. We set $a_J=1.0$. The point $\beta_J=0$ corresponds to the minimum of the Hamming distance in Fig. 2.

$$\overline{(\cdots)}$$

$$\equiv\frac{\sum_{\sigma=1}^{Q}(\cdots)\exp[2(m_*\beta_d+a_\tau hx+\tau_0 h\xi)\sigma-(h+\beta_d)\sigma^2]}{\sum_{\sigma=1}^{Q}\exp[2(m_*\beta_d+a_\tau hx+\tau_0 h\xi)\sigma-(h+\beta_d)\sigma^2]},$$

and $\tilde{D}_{\mathrm{H}}$ reads

$$\tilde{D}_{\mathrm{H}}\equiv 4\beta_d[m(t=0)-m_*]\int_{-\infty}^{\infty}\mathcal{D}x[\xi-\Omega(\overline{\sigma})]\times\sum_{k=1}^{Q}k\left[\delta\left(\overline{\sigma}-\frac{2k-1}{2}\right)-\delta\left(\overline{\sigma}-\frac{2k+1}{2}\right)\right]\times[\overline{\sigma^2}-(\overline{\sigma})^2].$$

We plot the inverse relaxation time $1/t_0$ as a function of T_d for the case of $Q=3, T_s=0.75, a_\tau=\tau_0=1.0$ in Fig. 7. In this figure, we also plot the inverse relaxation time for several values of the scaled field H. We see that the inverse relaxation time $1/t_0$ takes its minimum at a finite temperature T_d. However, the inverse relaxation time $1/t_0$ never reaches zero, and the relaxation to the equilibrium state is exponential for all T_d regions.

III. MONTE CARLO SIMULATIONS

Thus far, we have worked under the assumption that all the pixels lay on an infinite-dimensional grid, an approximation that enabled us to derive exact analytical formulas. In order to test the efficiency of the Q-Ising model on the more realistic case of a two-dimensional picture, in this section, we carry out Monte Carlo simulations at finite temperature

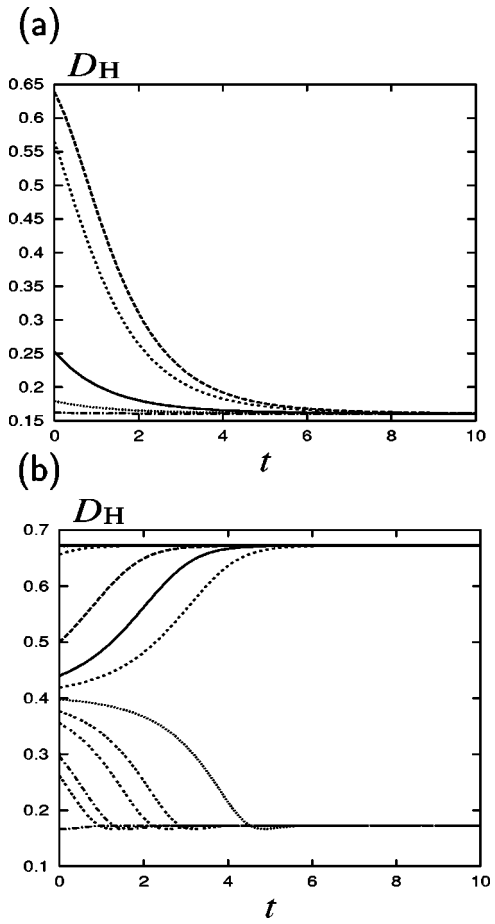

FIG. 6. The time evolutions of the Hamming distance are plotted in (a) $T_d=T_s=0.75$ and (b) $T_d=0.2\neq T_s=0.75$. We see that for the case of $T_d=T_s$, the Hamming distance converges to its optimal value for any initial state of the dynamics. On the other hand, if we set $T_d=0.2\neq T_s$, the Hamming distance converges to the wrong state that is higher than the Hamming distance between the original image and the corrupted one, that is, $D_{\mathrm{H}}^{(1)}=0.5$.

on a real-world image with $Q=8$ gray-scale level [Fig. 9(a)] corrupted by a Gaussian noise with $a_\tau=1.2$ [Fig. 9(b)]. Here the interaction in effective Hamiltonian is now restricted to the nearest neighbors spins on two-dimensional square lattice. As before, we first study the Hamming distance without the glassy term. The resulting curves averaged over 20 Monte Carlo runs are shown in Fig. 8(a) for three different values of the ratio $H=h/\beta_d$. The plots reflect indeed the mean-field behavior of Figs. 2(d) and 3(b). The corresponding restored picture at optimal values is shown in Fig. 9(c). It is evident that the ferromagnetic term succeeds in eliminating the noised pixel, i.e., isolated ones, but at the same time it also smoothens out the small true details of the original picture. For this reason, by keeping fixed $T_d^{(\mathrm{opt})}$ and H_{opt}, we switch on the glassy term, namely, $\beta_J\neq 0$. In Eq. (4) for the analysis of the infinite range model, we sent products $\xi_i\xi_j$, $(i,j=1,\ldots,N;i\neq j)$ of two pixels in original image through the channel. In this case, we send extra information about the correlation of arbitrary two pixels to receiver. However, we may send another kind of extra information. For example, it may be useful to send the square distances

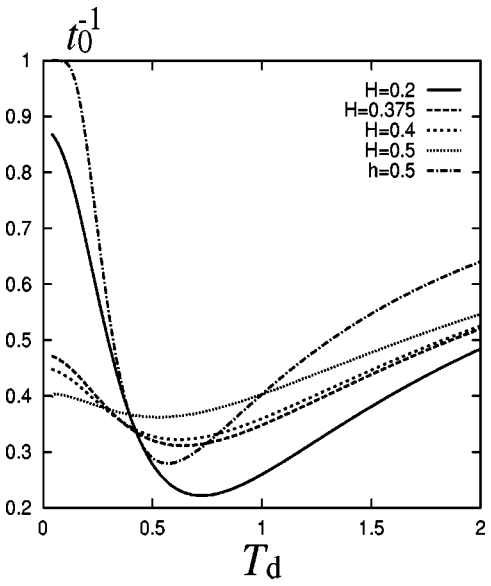

FIG. 7. The inverse relaxation time $1/t_0$ as a function of T_d for the case of $Q=3, T_s=0.75, a_\tau=\tau_0=1.0$.

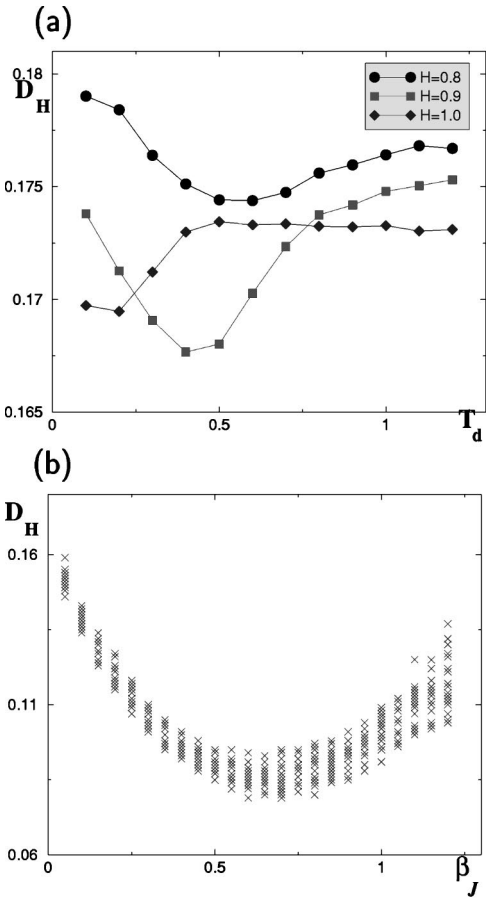

FIG. 8. The Hamming distance calculated by Monte Carlo simulation for 100×100 standard picture ''house.'' The curves averaged over 20 MCS runs are shown in [$\beta_J=0$ (a) and $\beta_J\neq0$ (b)].

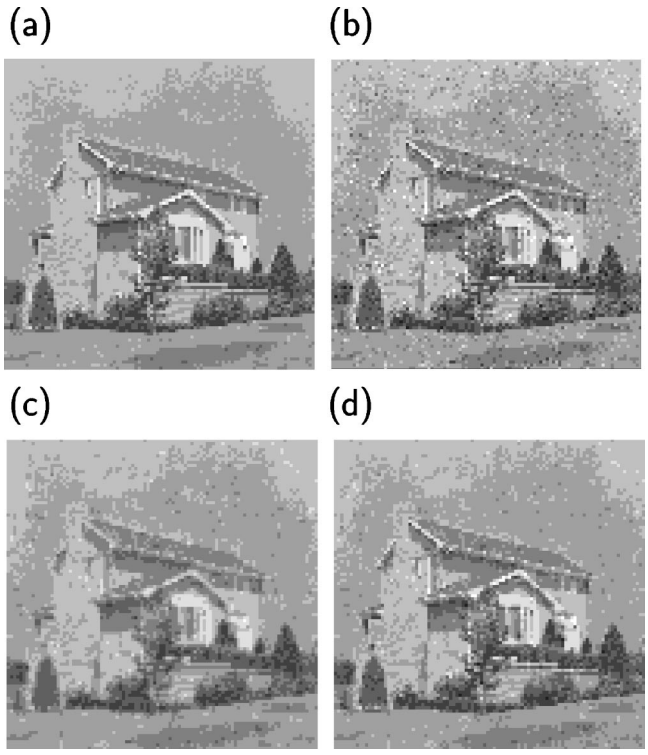

FIG. 9. The original picture (a) (''house,'' size 100×100), the corrupted picture by $a_\tau=1.2$ Gaussian noise (b), the restored pictures at $\beta_J=0$ (c) and $\beta_J\neq0$ (d) are displayed.

between two pixels, namely, $(\xi_i-\xi_j)^2$, $(i,j=1,\ldots,N;i\neq j)$ through the following noisy channel:

$$P(J_{ij}|(\xi_i-\xi_j)^2)=\frac{1}{\sqrt{2\pi}a_\tau}\exp\left[-\frac{\{J_{ij}-J_0(\xi_i-\xi_j)^2\}^2}{2a_J^2}\right].$$

For this degrading process, we use the likelihood

$$P(\{J\}|\{\sigma\})=\frac{1}{(\sqrt{2\pi})^N}\exp\left[-\frac{\beta_J}{2}\sum_{ij}\{J_{ij}-(\sigma_i-\sigma_j)^2\}^2\right]$$

to construct the posterior distribution. Here we set $a_J=J_0=1.0$ and carry out Monte Carlo simulations to check the efficiency of the above extra information. The curves in Fig. 8(b), for 20 Monte Carlo runs, show an improvement of the recovered image. The restored image at the minimum $\beta_J^{(\mathrm{opt})}$ is on the lower right corner of Fig. 9. It is evident that the extra term preserves many of the small details of the original image, for example, the white edge of the roof, which was blurred for the case of $\beta_J=0$.

IV. ITERATIVE ALGORITHM (MEAN FIELD)

The restoration by means of Monte Carlo methods is the result of a statistical process that might take long time even for powerful computers as the size of the picture increases. Therefore we apply mean-field iterative algorithm to our model. Mean-field iterative algorithm was already applied to image restoration problem by several authors in the context of the MAP estimation [3] (the Gaussian model), [8] (the Ising model using pair approximation), [9] (the Potts model using cluster zero temperature process) in the context of the MAP estimation. Therefore, these methods are referred to as *mean-field annealing*. In this paper, however, we use the mean-field iterative algorithm to obtain the MPM estimate. For this purpose, we set temperature $T=1$ during the iterations [see Eq. (B6) in Appendix B]. Recently, mean-field iterative algorithm in the context of the MPM estimation was applied to error-correcting codes problems [10] and its performance was investigated form information geometrical point of view [11].

Using the mean-field approximation with periodic boundary conditions for two dimensional square lattice of size $L_1\times L_2$, the recursion relations with respect to the local magnetization at a site (i,j), namely, m_{ij} lead to

$$m_{ij}^{(t+1)}=\frac{\mathrm{Tr}_\sigma\,\sigma e^{\omega_{ij}^{(t)}(\sigma)}}{\mathrm{Tr}_\sigma e^{\omega_{ij}^{(t)}(\sigma)}}, \tag{13}$$

$$\omega_{ij}^{(t)}(\sigma)=\{J[m_{i,j+1}^{(t)}+m_{i,j-1}^{(t)}+m_{i+1,j}^{(t)}+m_{i-1,j}^{(t)}]+2h\tau_{ij}\}\sigma-(2J+h)\sigma^2, \tag{14}$$

$$m_{i,j+L_2}=m_{ij},\quad m_{i+L_1,j}=m_{ij}, \tag{15}$$

where $\mathrm{Tr}_\sigma(\cdots)$ means the sum with respect to the gray-scale levels, namely, $\sigma=1,2,\ldots,Q$. The details of the derivation by using a variational principle are reported in Appendix B.

Then, we obtain the estimate of the pixel ξ_{ij}, namely, $\Omega(m_{ij})$ by solving the above nonlinear maps until appropriate error tolerance is satisfied. In order to investigate its performance, we introduce the following three measures:

$$D_{\mathrm{H}}\equiv\frac{1}{2L_1L_2}\sum_{i=1}^{L_1}\sum_{j=1}^{L_2}[\Omega(m_{ij})-\xi_{ij}]^2,$$

$$D_{\mathrm{H}}^{(1)}\equiv\frac{1}{2L_1L_2}\sum_{i=1}^{L_1}\sum_{j=1}^{L_2}[\Omega(m_{ij})-\tau_{ij}]^2,$$

$$D_{\mathrm{H}}^{(2)}\equiv\frac{1}{2L_1L_2}\sum_{i=1}^{L_1}\sum_{j=1}^{L_2}[\tau_{ij}-\xi_{ij}]^2,$$

where D_{H}, $D_{\mathrm{H}}^{(1)}$, and $D_{\mathrm{H}}^{(2)}$ are distances between the original image ξ_{ij} and the restored one $\Omega(m_{ij})$, the corrupted image τ_{ij} and the restored one, and the corrupted image and the original one, respectively. We choose $Q=8$, $L_1=L_2=128$, and $h=1.0$ and solve the recursion relations (13), (14), and (15) until the error

$$\varepsilon^{(t)}\equiv\frac{1}{2L_1L_2}\sum_{i=1}^{L_1}\sum_{j=1}^{L_2}|m_{ij}^{(t+1)}-m_{ij}^{(t)}|$$

becomes smaller than 10^{-5}. We list the results in Fig. 8 (''girl'') and Fig. 9 (''chair''). The original images are degraded by the Gaussian noise with a standard deviation $a_\tau=2.2$. This standard deviation gives the Hamming distance between the original image and the degraded one $D_{\mathrm{H}}^{(2)}$

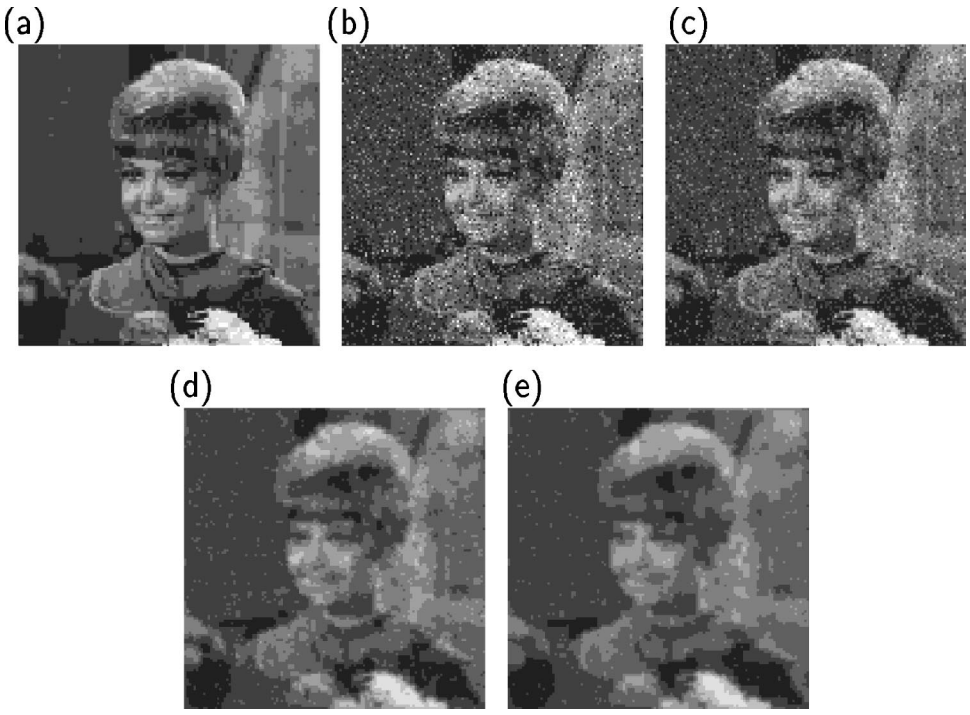

FIG. 10. The results of the iterative algorithm are displayed. The original ''girl'' picture of 128×128 (a), the degraded picture (b), the restored pictures with $J=0.2$ (c), $J=1.8$ (d), and $J=2.5$ (e) are shown.

~0.205. Obviously, the picture ''chair'' contains many more edges than the picture ''girl.'' Therefore, one of our aims of this demonstration is to check to what extent our model detects the edge parts in the real-world picture. In Fig. 10, we plot the Hamming distances D_{H} [(a)] and $D_H^{(1)}$ [(b)]. We choose the degraded image as an initial set of the pixels and investigated the J dependence of the Hamming distance. We see that the performance of the algorithm for the ''girl'' picture is much better than that of the ''chair'' picture. This is because the smoothness term in the effective Hamiltonian [Eq. (2)] is quadratic and it is hard to detect the edges in the ''chair'' picture. For both pictures, the optimal performance is achieved around the parameter $T_d\equiv1/J=1/1.8\sim0.56$. This value is not so different from the parameter that was obtained in Monte Carlo simulations [see Fig. 6(a)]. Of course, from a practical point of view, it is possible to stop the Monte Carlo simulation and not to wait for the convergence to attain the equilibrium state precisely. Then, we may regard the snapshot as the restored image if the performance is not so bad. However, in the mean-field approximation we

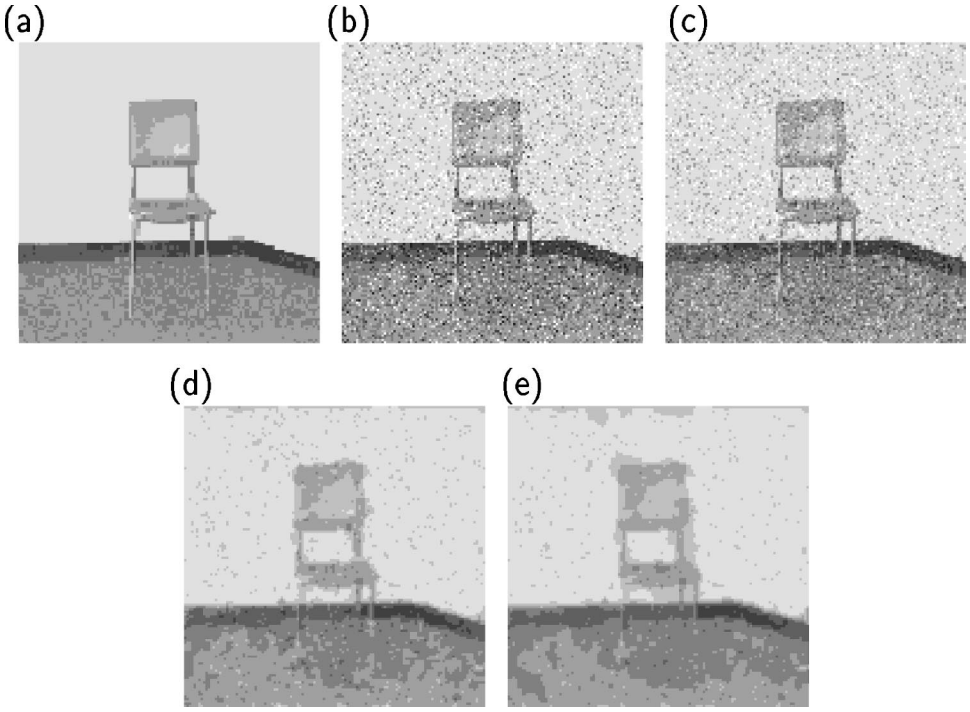

FIG. 11. The results of the iterative algorithm are displayed. The original ''chair'' picture of 128×128 (a), the degraded picture (b), the restored pictures with $J=0.2$ (c), $J=1.8$ (d), and $J=2.5$ (e) are shown.

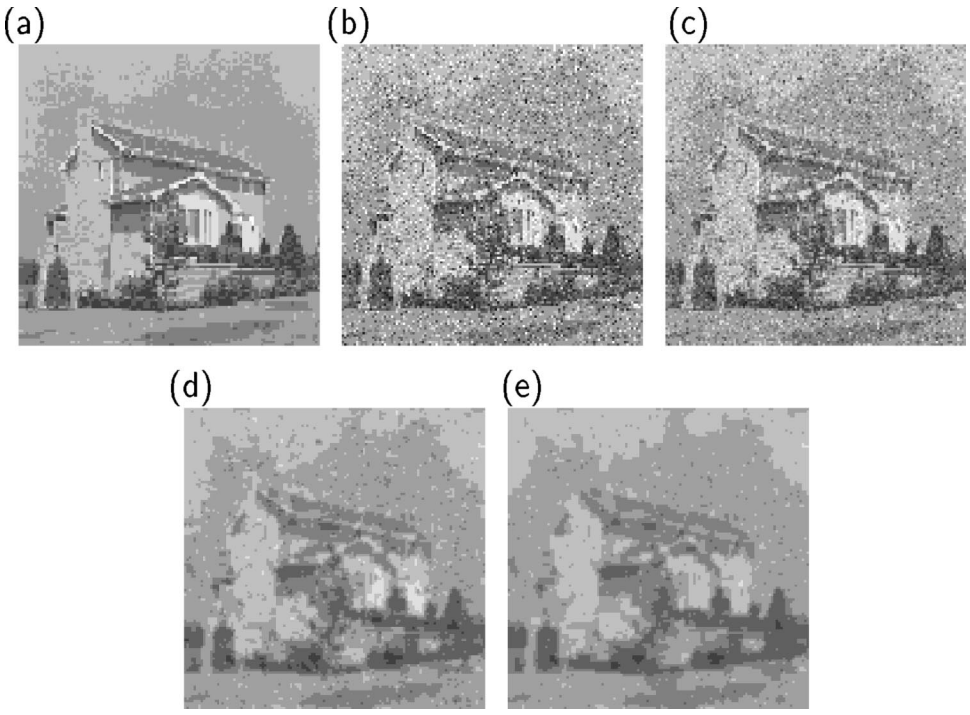

FIG. 12. The results of the iterative algorithm are displayed. The original ''house'' picture of 128×128 (a), the degraded picture (b), the restored pictures with $J=0.2$ (c), $J=1.8$ (d), and $J=2.5$ (e) are shown.

constructed here, the convergence of the iteration is guaranteed (Figs. 11–13).

V. SUMMARY AND DISCUSSION

In this paper we investigated the efficiency of the Q-Ising model for image restoration problem, when the original image is affected by Gaussian noise. By introducing the infinite range model, we gave an analytical expression for the Hamming distance, which is shown to reach its minimum at some finite temperature. We found that the optimal temperature for the GSIR using the Q-Ising model coincides with the source temperature in contrast to the chiral Potts case [5]. We also found that as in the Ising and Potts spin cases, the presence of a parity-check-*like* term greatly increases the performance of the GSIR process. Although for practical restorations of images, one would not like to smoothen out two points far away, the mean-field results provide a remarkable eye-guide for a short-range version of the effective Hamiltonian as confirmed by Monte Carlo simulation on two-dimensional pictures. From a dynamical point of view, we also obtained the time evolution of the Hamming distance analytically and found the critical initial Hamming distance beyond which the flow does not converge to its optimal value. We show the dynamical equation we obtained is exactly the same as the TDGL equation. Therefore, the destination of the dynamics is one of the locally stable states of the free energy, and if we fail to select the initial condition, the dynamics converges to the local minimum that does not give the minimum of the Hamming distance.

Recently, Skantzos and Coolen [12], reported the synchronous dynamics of one-dimensional and infinite-dimensional random field Ising models. They found that the dynamics has much more rich behavior than the sequential (Glauber) dynamics. Therefore, for our present model system, there is a

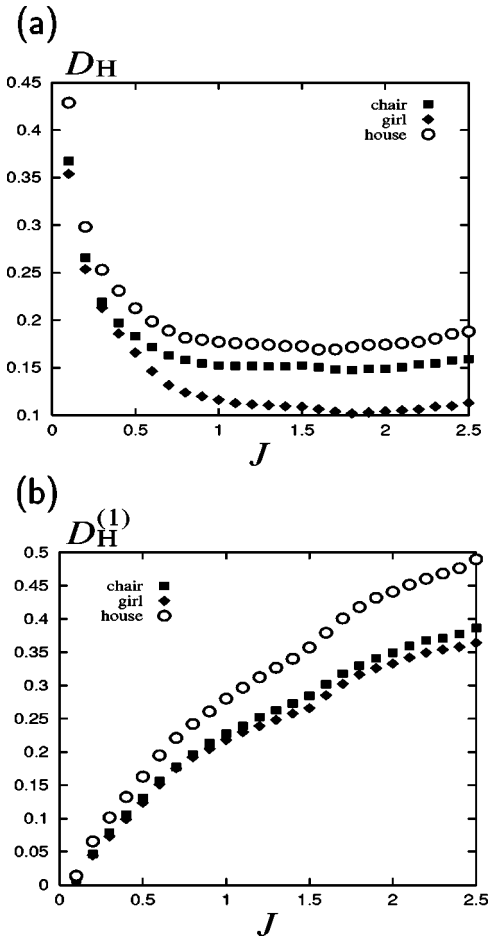

FIG. 13. The Hamming distance between the original image and the restored one D_{H} as a function of J [(a)] obtained by the iterative algorithm. The Hamming distance between the restored image and the degraded one $D_{\mathrm{H}}^{(1)}$ is shown in (b).

possibility that if we consider the synchronous dynamics instead of the sequential one, the behavior of the dynamics may be different from the results we obtained here.

Using the mean-field approximation, we also constructed the iterative algorithm that converges faster than the Monte Carlo simulation. We derived it from a variational principle of the free energy and demonstrate it for two types of the real-world pictures. From those results, we concluded that we need some extra term that would detect the edges if the picture has a lot of edges. We suppose that the glassy term we introduced in the infinite range model may play this role. This will be achieved by means of the TAP-like mean-field approximation [10].

ACKNOWLEDGMENTS

The authors acknowledge Hidetoshi Nishimori, Kazuyuki Tanaka, and Desire Bolle' for fruitful discussion and useful comments. We also thank Masato Okada for valuable comments on the dynamics of disordered spin systems. One of the authors (J.I.) was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Encouragement of Young Scientists, Grant No. 11740225, 1999–2000.

APPENDIX A: DERIVATION OF THE FLOW OF MAGNETIZATION

In this appendix, we derive the differential equation with respect to macroscopic order parameter m from microscopic master equation for the infinite range version of the Q-Ising model. For simplicity, we consider the case of no-parity check term $\beta_J=0$. For the Q-Ising model, the effective Hamiltonian is given as

$$\mathcal{H}_{\rm eff}=\frac{\beta_d}{2N}\sum_{ij}(\sigma_i-\sigma_j)^2+h\sum_i(\sigma_i-\tau_i)^2\equiv\mathcal{H}(\{\sigma\}).$$

Therefore, the energy difference due to the local spin change $\sigma_k\to\sigma_{k'}$, namely, $\Delta E\equiv\mathcal{H}(F_k\{\sigma\})-\mathcal{H}(\{\sigma\})$ is calculated in terms of the above Hamiltonians $\mathcal{H}(\{\sigma\})$ and $\mathcal{H}(F_k\{\sigma\})$ as follows:

$$\Delta E=(h+\beta_d)(\sigma_{k'}^2-\sigma_k^2)-2\beta_d m(\sigma_{k'}-\sigma_k)-2h(\sigma_{k'}-\sigma_k)\tau_k,$$

where we defined single pixel changing operator F_k as $F_k\Phi(\{\sigma\})=\Phi(\sigma_1,\ldots,\sigma_{k'},\ldots,\sigma_N)$ for arbitrary function $\Phi(\{\sigma\})$ and used the expression of the magnetization

$$m=\frac{1}{N}\sum_j\sigma_j.$$

Then, the transition rate $w(\sigma_k\to\sigma_{k'})$ is given by

$$w(\{\sigma\})\equiv w(\sigma_k\to\sigma_{k'})=\frac{e^{-\Delta E}}{\sum_{\sigma_{k'}=1}^{Q}e^{-\Delta E}}=\frac{\exp[-(h+\beta_d)\sigma_{k'}^2+2(m\beta_d+h\tau_k)\sigma_{k'}]}{\sum_{\sigma_{k'}=1}^{Q}\exp[-(h+\beta_d)\sigma_{k'}^2+2(m\beta_d+h\tau_k)\sigma_{k'}]}.$$

and $w(F_k\{\sigma\})\equiv w(\sigma_{k'}\to\sigma_k)$. For this transition rate, the master equation leads to

$$\frac{d}{dt}p_t(\{\sigma\})=\sum_{k=1}^{N}\sum_{\sigma_{k'}=1}^{Q}[w(F_k\{\sigma\})p_t(F_k\{\sigma\})-w(\{\sigma\})p_t(\{\sigma\})].$$

Here we introduce the macroscopic probability

$$\mathcal{P}_t(m)\equiv\sum_{\{\sigma\}}p_t(\{\sigma\})\delta[m-m(\{\sigma\})],$$

and consider the derivative of $\mathcal{P}_t(m)$ with respect to t, that is,

$$\begin{aligned}\frac{d}{dt}\mathcal{P}_t(m)&=\sum_{\{\sigma\}}\frac{dp_t(\{\sigma\})}{dt}\delta[m-m(\{\sigma\})]\\&=\sum_{\{\sigma\}}\left\{\sum_{k=1}^{N}\sum_{\sigma_{k'}=1}^{Q}[w(F_k\{\sigma\})p_t(F_k\{\sigma\})-w(\{\sigma\})p_t(\{\sigma\})]\right\}\delta[m-m(\{\sigma\})]\\&=\sum_{\{\sigma\}}\sum_{k=1}^{N}\sum_{\sigma_{k'}=1}^{Q}w(\{\sigma\})p_t(\{\sigma\})\left\{\delta\left[m-m(\{\sigma\})+\frac{1}{N}(\sigma_k-\sigma_{k'})\right]-\delta[m-m(\{\sigma\})]\right\}\\&=\frac{\partial}{\partial m}\left\{\sum_{\{\sigma\}}\sum_{k=1}^{N}\sum_{\sigma_{k'}=1}^{Q}w(\{\sigma\})p_t(\{\sigma\})\delta[m-m(\{\sigma\})]\frac{1}{N}(\sigma_k-\sigma_{k'})\right\}\end{aligned}$$

$$= \frac{\partial}{\partial m}\left\{ \sum_{\{\sigma\}} p_t(\{\sigma\}) \sum_{k=1}^{N} \sum_{\sigma_{k'}=1}^{Q} \left[\frac{\exp[-(h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]}{\sum_{\sigma_{k'}=1}^{Q} \exp[-(h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]} \right] \frac{1}{N}(\sigma_k - \sigma_{k'})\,\delta[m-m(\{\sigma\})] \right\}$$

$$= \frac{\partial}{\partial m} \sum_{\{\sigma\}} p_t(\{\sigma\}) \sum_{k=1}^{N} \left\{ \sum_{\sigma_{k'}=1}^{Q} \frac{\sigma_k}{N} \left[\frac{\exp[-(h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]}{\sum_{\sigma_{k'}=1}^{Q} \exp[-(h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]} \right] \right.$$

$$\left. - \sum_{\sigma_{k'}=1}^{Q} \frac{\sigma_{k'}}{N} \left[\frac{\exp[-(h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]}{\sum_{\sigma_{k'}=1}^{Q} \exp[-(h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]} \right] \right\} \delta[m-m(\{\sigma\})]$$

$$= \frac{\partial}{\partial m} \sum_{\{\sigma\}} p_t(\{\sigma\}) \left\{ \frac{1}{N} \left[\sum_{k=1}^{N} \sigma_k - \sum_{k=1}^{N} \frac{\sum_{\sigma_{k'}=1}^{Q} \sigma_{k'} \exp[(-h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]}{\sum_{\sigma_{k'}=1}^{Q} \exp[-(h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]} \right] \right\} \delta[m-m(\{\sigma\})], \quad \text{(A1)}$$

where we expanded the right-hand side of the above equation (A1) with respect to the quantity of order $1/N$. Here we should notice that

$$\frac{1}{N} \sum_{k=1}^{N} \frac{\sum_{\sigma_{k'}=1}^{Q} \sigma_{k'} \exp[-(h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]}{\sum_{\sigma_{k'}=1}^{Q} \exp[-(h+\beta_d)\sigma_{k'}^2 + 2(m\beta_d+\tau_k h)\sigma_{k'}]}$$

$$= \frac{\sum_{\xi=1}^{Q} \exp(2m_0\beta_s\xi - \beta_s\xi^2)}{\mathcal{Z}(\beta_s)} \int_{-\infty}^{\infty} \mathcal{D}x \left[\frac{\sum_{\sigma=1}^{Q} \sigma \exp[-(h+\beta_d)\sigma^2 + 2(m\beta_d + h a_\tau x + h\tau_0\xi)\sigma]}{\sum_{\sigma=1}^{Q} \exp[-(h+\beta_d)\sigma^2 + 2(m\beta_d + h a_\tau x + h\tau_0\xi)\sigma]} \right] \quad \text{(A2)}$$

should hold due to the self-averaging properties in the thermo-dynamical limit $N \to \infty$. Substituting this expression into Eq. (A1), we obtain

$$\frac{d}{dt}\mathcal{P}_t(m) = \frac{\partial}{\partial m} m \sum_{\{\sigma\}} p_t(\{\sigma\})\,\delta[m-m(\{\sigma\})] - \frac{\partial}{\partial m} \sum_{\{\sigma\}} p_t(\{\sigma\})\,\delta[m$$

$$- m(\{\sigma\})] \frac{\sum_{\xi=1}^{Q} \exp(2m_0\beta_s\xi - \beta_s\xi^2)}{\mathcal{Z}(\beta_s)} \int_{-\infty}^{\infty} \mathcal{D}x \left[\frac{\sum_{\sigma=1}^{Q} \sigma \exp[-(h+\beta_d)\sigma^2 + 2(m\beta_d + h a_\tau x + h\tau_0\xi)\sigma]}{\sum_{\sigma=1}^{q} \exp[-(h+\beta_d)\sigma^2 + 2(m\beta_d + h a_\tau x + h\tau_0\xi)\sigma]} \right]$$

$$= \frac{\partial}{\partial m} m\mathcal{P}_t(m) - \frac{\partial}{\partial m}\mathcal{P}_t(m) \frac{\sum_{\xi=1}^{Q} \exp(2m_0\beta_s\xi - \beta_s\xi^2)}{\mathcal{Z}(\beta_s)} \int_{-\infty}^{\infty} \mathcal{D}x \left[\frac{\sum_{\sigma=1}^{Q} \sigma \exp[-(h+\beta_d)\sigma^2 + 2(m\beta_d + h a_\tau x + h\tau_0\xi)\sigma]}{\sum_{\sigma=1}^{Q} \exp[-(h+\beta_d)\sigma^2 + 2(m\beta_d + h a_\tau x + h\tau_0\xi)\sigma]} \right]$$

$$= \frac{\partial}{\partial m} \left\{ \mathcal{P}_t(m) \left(m - \frac{\sum_{\xi=1}^{Q} \exp(2m_0\beta_s\xi - \beta_s\xi^2)}{\mathcal{Z}(\beta_s)} \int_{-\infty}^{\infty} \mathcal{D}x \left[\frac{\sum_{\sigma=1}^{Q} \sigma \exp[-(h+\beta_d)\sigma^2 + 2(m\beta_d + h a_\tau x + h\tau_0\xi)\sigma]}{\sum_{\sigma=1}^{Q} \exp[-(h+\beta_d)\sigma^2 + 2(m\beta_d + h a_\tau x + h\tau_0\xi)\sigma]} \right] \right) \right\}. \quad \text{(A3)}$$

Multiplying m and substituting $\mathcal{P}_t(m)=\delta[m-m(t)]$ to the left-hand side of the above Eq. (A3) and calculating integral with respect to m, we obtain

$$\int_{-\infty}^{\infty} m dm \frac{d}{dt}\delta[m-m(t)]=\frac{d}{dt}\int_{-\infty}^{\infty} m\, dm\, \delta[m-m(t)]=\frac{dm}{dt}. \tag{A4}$$

Using the same way as the left-hand side of the Eq. (A3), the right-hand side of Eq. (A3) leads to

$$\begin{aligned}
&\int_{-\infty}^{\infty} m dm \frac{\partial}{\partial m}\left\{ \delta[m-m(t)]\left(m-\frac{\sum_{\xi=1}^{Q}\exp(2m_0\beta_s\xi-\beta_s\xi^2)}{\mathcal{Z}(\beta_s)}\int_{-\infty}^{\infty}\mathcal{D}x \right.\right.\\
&\qquad\left.\left.\times\left[\frac{\sum_{\sigma=1}^{Q}\sigma\exp[-(h+\beta_d)\sigma^2+2(m\beta_d+ha_\tau x+h\tau_0\xi)\sigma]}{\sum_{\sigma=1}^{Q}\exp[-(h+\beta_d)\sigma^2+2(m\beta_d+ha_\tau x+h\tau_0\xi)\sigma]}\right]\right)\right\}\\
&=-\int_{-\infty}^{\infty} dm\,\delta[m-m(t)]\left\{ m-\frac{\sum_{\xi=1}^{Q}\exp(2m_0\beta_s\xi-\beta_s\xi^2)}{\mathcal{Z}(\beta_s)}\right.\\
&\qquad\left.\times\int_{-\infty}^{\infty}\mathcal{D}x\left[\frac{\sum_{\sigma=1}^{Q}\sigma\exp[-(h+\beta_d)\sigma^2+2(m\beta_d+ha_\tau x+h\tau_0\xi)\sigma]}{\sum_{\sigma=1}^{Q}\exp[-(h+\beta_d)\sigma^2+2(m\beta_d+ha_\tau x+h\tau_0\xi)\sigma]}\right]\right\}\\
&=-m+\frac{\sum_{\xi=1}^{Q}\exp(2m_0\beta_s\xi-\beta_s\xi^2)}{\mathcal{Z}(\beta_s)}\int_{-\infty}^{\infty}\mathcal{D}x\left[\frac{\sum_{\sigma=1}^{Q}\sigma\exp[-(h+\beta_d)\sigma^2+2(m\beta_d+ha_\tau x+h\tau_0\xi)\sigma]}{\sum_{\sigma=1}^{Q}\exp[-(h+\beta_d)\sigma^2+2(m\beta_d+ha_\tau x+h\tau_0\xi)\sigma]}\right].
\end{aligned} \tag{A5}$$

From Eqs. (A4) and (A5), we obtain the final form of the dynamical equation with respect to magnetization m as

$$\frac{dm}{dt}=-m+\frac{\sum_{\xi=1}^{Q}\exp(2m_0\beta_s\xi-\beta_s\xi^2)}{\mathcal{Z}(\beta_s)}\int_{-\infty}^{\infty}\mathcal{D}x\left[\frac{\sum_{\sigma=1}^{Q}\sigma\exp[-(h+\beta_d)\sigma^2+2(m\beta_d+ha_\tau x+h\tau_0\xi)\sigma]}{\sum_{\sigma=1}^{Q}\exp[-(h+\beta_d)\sigma^2+2(m\beta_d+ha_\tau x+h\tau_0\xi)\sigma]}\right]. \tag{A6}$$

We easily see that the above equation is exactly the same as the TDGL equation that is derived from the steepest descent of the replica symmetric free energy, that is, $-\partial f_{\mathrm{RS}}/\partial m = dm/dt$. We should also notice that in the limit of $t\to\infty$ and $dm/dt=0$, Eq. (A6) corresponds to the saddle point equation with respect to m, which was calculated by equilibrium statistical mechanics in Sec. II A.

APPENDIX B: VARIATIONAL PRINCIPLE FOR THE *Q*-ISING MODEL

In Sec. VI, we introduced the recursion relations that determine the estimate of the original image in terms of mean-field approximation [8,9]. In this appendix, we show that these recursion relations, Eqs. (13), (14), and (15), can be derived from a variational principle.

We consider the following optimization problem:

$$\min_{\rho}\{\mathcal{E}(\rho)-T\mathcal{S}(\rho)\},$$

$$\mathcal{E}(\rho)\equiv\sum_{\{\sigma\}}\mathcal{H}(\{\sigma\})\rho(\{\sigma\}),$$

$$\mathcal{S}(\rho)\equiv-\sum_{\{\sigma\}}\rho(\{\sigma\})\ln\rho(\{\sigma\}),$$

where Hamiltonian $\mathcal{H}$ is defined on the two-dimensional square lattice of size $L_1\times L_2$ $(L_1=L_2=N)$ as

$$\mathcal{H}(\{\sigma\})=\frac{J}{2}\sum_{ij}\{(\sigma_{ij}-\sigma_{i,j+1})^2+(\sigma_{ij}-\sigma_{i,j+1})^2+(\sigma_{ij}-\sigma_{i+1,j})^2+(\sigma_{ij}-\sigma_{i-1,j})^2\}+h\sum_{ij}(\tau_{ij}-\sigma_{ij})^2,$$

and $\mathcal{E}$ and $\mathcal{S}$ correspond to the energy and entropy of the system, respectively. Then, we use the mean-field approximation, that is,

$$\rho(\{\sigma\})\simeq\prod_{ij}\rho_{ij}(\sigma_{ij}).$$

We should notice that for each pixel (i,j), the following normalization condition should hold:

$$\sum_{\sigma_{ij}=1}^{Q}\rho_{ij}(\sigma_{ij})=1. \tag{B1}$$

Using the Lagrange multiplier λ_{ij}, we take into account the above normalization condition with respect to the marginal distribution, and maximize the following functional:

$$\mathcal{F}\equiv\mathcal{E}(\{\sigma\})-T\mathcal{S}(\{\sigma\})+\sum_{ij}\lambda_{ij}\left(\sum_{\sigma_{ij}=1}^{Q}\rho_{ij}(\sigma_{ij})-1\right). \tag{B2}$$

The energy $\mathcal{E}$ and the entropy $\mathcal{S}$ of the system can be written explicitly as

$$\mathcal{E}=\sum_{\sigma_{12}}\cdots\sum_{\sigma_{ij}}\cdots\sum_{\sigma_{kl}}\cdots\sum_{\sigma_{N-1N}}\mathcal{H}(\{\sigma\})\rho_{12}(\sigma_{12})\cdots\times\rho_{ij}(\sigma_{ij})\cdots\rho_{kl}(\sigma_{kl})\cdots\rho_{N-1N}(\sigma_{N-1N}),$$

$$\mathcal{S}=-\sum_{\sigma_{12}}\sum_{\sigma_{13}}\cdots\sum_{\sigma_{ij}}\cdots\sum_{\sigma_{kl}}\cdots\sum_{\sigma_{N-1N}}\prod_{ij}\rho_{ij}(\sigma_{ij})\times\sum_{ij}\ln\rho_{ij}(\sigma_{ij}).$$

The derivative of the third term of Eq. (B2) with respect to $\rho_{ij}(\sigma_{ij})$ leads to $\Sigma_{\sigma_{ij}=1}^{Q}\lambda_{ij}$, therefore, we have $(\partial\mathcal{F}/\partial\rho_{ij})=0$ as

$$\frac{\partial\mathcal{F}}{\partial\rho_{ij}(\sigma_{ij})}=\sum_{\sigma_{ij}=1}^{Q}\left\{\sum_{\sigma_{12}}\cdots\sum_{\sigma_{ij}}\cdots\sum_{\sigma_{kl}}\cdots\sum_{\sigma_{N-1N}}\mathcal{H}(\{\sigma\})\times\rho_{12}(\sigma_{12})\cdots\rho_{kl}(\sigma_{kl})\cdots\rho_{N-1N}(\sigma_{N-1N})+T\ln\rho_{ij}(\sigma_{ij})+T+T\sum_{kl\neq ij}\sum_{\sigma_{kl}=1}^{Q}\rho_{kl}(\sigma_{kl})\ln\rho_{kl}(\sigma_{kl})+\lambda_{ij}\right\}=0.$$

This leads to

$$\rho_{ij}(\sigma_{ij})=\mathcal{A}\exp\left[-\frac{1}{T}\sum_{\sigma_{12}}\cdots\sum_{\sigma_{ij}}\cdots\sum_{\sigma_{kl}}\cdots\sum_{\sigma_{N-1N}}H(\{\sigma\})\times\rho_{12}(\sigma_{12})\cdots\rho_{kl}(\sigma_{kl})\cdots\rho_{N-1N}(\sigma_{N-1N})\right]. \tag{B3}$$

Using the normalization condition (B1), we obtain the factor

$$\mathcal{A}\equiv\exp\left(-\frac{\lambda_{ij}}{T}-1-\sum_{kl\neq ij}\sum_{\sigma_{kl}=1}^{Q}\rho_{kl}(\sigma_{kl})\ln\rho_{kl}(\sigma_{kl})\right)$$

as

$$\mathcal{A}=\left(\sum_{\sigma_{ij}=1}^{Q}\exp\left[-\frac{1}{T}\sum_{\sigma_{12}}\cdots\sum_{\sigma_{ij}}\cdots\sum_{\sigma_{kl}}\cdots\sum_{\sigma_{N-1N}}\mathcal{H}(\{\sigma\})\times\rho_{12}(\sigma_{12})\cdots\rho_{kl}(\sigma_{kl})\cdots\rho_{N-1N}(\sigma_{N-1N})\right]\right)^{-1}. \tag{B4}$$

From Eqs. (B3) and (B4), the marginal distribution $\rho_{ij}(\sigma_{ij})$ reads

$$\rho_{ij}(\sigma_{ij})=\frac{\exp\left[-\frac{1}{T}\sum_{\sigma_{12}}\cdots\sum_{\sigma_{ij}}\cdots\sum_{\sigma_{kl}}\cdots\sum_{\sigma_{N-1N}}\mathcal{H}(\{\sigma\})\rho_{12}(\sigma_{12})\cdots\rho_{kl}(\sigma_{kl})\cdots\rho_{N-1N}(\sigma_{N-1N})\right]}{\sum_{\sigma_{ij}=1}^{Q}\exp\left[-\frac{1}{T}\sum_{\sigma_{12}}\cdots\sum_{\sigma_{ij}}\cdots\sum_{\sigma_{kl}}\cdots\sum_{\sigma_{N-1N}}\mathcal{H}(\{\sigma\})\rho_{12}(\sigma_{12})\cdots\rho_{kl}(\sigma_{kl})\cdots\rho_{N-1N}(\sigma_{N-1N})\right]}.$$

In order to calculate the sum $\Sigma_{\sigma_{12}}\cdots\Sigma_{\sigma_{NN-1}}(\cdots)$, we rewrite the Hamiltonian $\mathcal{H}(\{\sigma\})$ as

$$\mathcal{H}(\{\sigma\}) = -J(\sigma_{i,j+1}+\sigma_{i,j-1}+\sigma_{i+1,j}+\sigma_{i-1,j})\sigma_{ij} - 2h\tau_{ij}\sigma_{ij} + (2J+h)(\sigma_{ij})^2 - J\sum_{kl\neq ij}(\sigma_{k,l+1}+\sigma_{k,l-1}+\sigma_{k+1,l}+\sigma_{k-1,l})\sigma_{kl}$$

$$+(2J+h)\sum_{kl\neq ij}(\sigma_{kl})^2 + h\sum_{ij}(\tau_{ij})^2 - 2h\sum_{kl\neq ij}\tau_{kl}\sigma_{kl},$$

and using the relations between the local magnetization and the marginal distribution, namely, $m_{i,j+1} = \sum_{\sigma_{i,j+1}=1}^{Q}\sigma_{i,j+1}\rho(\sigma_{i,j+1})$, etc., we obtain

$$\begin{aligned} &-\frac{1}{T}\sum_{\sigma_{12}}\cdots\sum_{\sigma_{ij}}\cdots\sum_{\sigma_{kl}}\cdots\sum_{\sigma_{N-1N}}\mathcal{H}(\{\sigma\})\rho_{12}(\sigma_{12})\cdots\rho_{kl}(\sigma_{kl})\cdots\rho_{N-1N}(\sigma_{N-1N}) \\ &= \frac{J}{T}(m_{i,j+1}+m_{i,j-1}+m_{i+1,j}+m_{i-1,j})\sigma_{ij} + \frac{2h}{T}\tau_{ij}\sigma_{ij} - \frac{(2J+h)}{T}(\sigma_{ij})^2 \\ &\quad -\frac{1}{T}\sum_{\{\sigma\}\in I'}\mathcal{H}(\{\sigma\})\rho_{12}(\sigma_{12})\cdots\rho_{kl}(\sigma_{kl})\cdots\rho_{N-1N}(\sigma_{N-1N}), \end{aligned} \tag{B5}$$

where I' stands for a set of the sites except for (i,j).

Using Eq. (B5), we rewrite $\rho_{ij}(\sigma_{ij})$ as

$$\rho_{ij}(\sigma_{ij}) = \frac{\exp\left\{\frac{1}{T}[J(m_{i,j+1}+m_{i,j-1}+m_{i+1,j}+m_{i-1,j})+2h\tau_{ij}]\sigma_{ij} - \frac{(2J+h)}{T}(\sigma_{ij})^2\right\}}{\sum_{\sigma_{ij}=1}^{Q}\exp\left\{\frac{1}{T}[J(m_{i,j+1}+m_{i,j-1}+m_{i+1,j}+m_{i-1,j})+2h\tau_{ij}]\sigma_{ij} - \frac{(2J+h)}{T}(\sigma_{ij})^2\right\}},$$

where the factors

$$\exp\left[-\frac{1}{T}\sum_{\{\sigma\}\in I'}\mathcal{H}(\{\sigma\})\rho_{12}(\sigma_{12})\cdots\rho_{kl}(\sigma_{kl})\cdots\rho_{N-1N}(\sigma_{N-1N})\right]$$

appearing in both numerator and denominator of the $\rho_{ij}(\sigma_{ij})$ were canceled. As the results, we obtain m_{ij} as follows.

$$m_{ij} = \sum_{\sigma_{ij}=1}^{Q}\sigma_{ij}\rho_{ij}(\sigma_{ij}) = \frac{\sum_{\sigma_{ij}=1}^{Q}\sigma_{ij}\exp\left\{\frac{1}{T}[J(m_{i,j+1}+m_{i,j-1}+m_{i+1,j}+m_{i-1,j})+2h\tau_{ij}]\sigma_{ij} - \frac{(2J+h)}{T}(\sigma_{ij})^2\right\}}{\sum_{\sigma_{ij}=1}^{Q}\exp\left\{\frac{1}{T}[J(m_{i,j+1}+m_{i,j-1}+m_{i+1,j}+m_{i-1,j})+2h\tau_{ij}]\sigma_{ij} - \frac{(2J+h)}{T}(\sigma_{ij})^2\right\}}. \tag{B6}$$

If we set $T=1$, we can obtain the recursion relations with respect to the local magnetization m_{ij} under the periodic boundary condition as

$$m_{ij}^{(t+1)} = \frac{\sum_{\sigma=1}^{Q}\sigma e^{\omega_{ij}^{(t)}(\sigma)}}{\sum_{\sigma=1}^{Q}e^{\omega_{ij}^{(t)}(\sigma)}},$$

$$\omega_{ij}^{(t)}(\sigma) = \{J[m_{i,j+1}^{(t)}+m_{i,j-1}^{(t)}+m_{i+1,j}^{(t)}+m_{i-1,j}^{(t)}]+2h\tau_{ij}\}\sigma - (2J+h)\sigma^2,$$

$$m_{i,j+N} = m_{ij}, \quad m_{i+N,j} = m_{ij},$$

which were obtained in the previous section as Eqs. (13), (14), and (15).

[1] J. Marroquin, S. Mitter, and T. Poggio, J. Am. Stat. Assoc. **82**, 76 (1987).
[2] J. M. Pryce and A. D. Bruce, J. Phys. A **28**, 511 (1995).
[3] D. Geiger and F. Girosi, IEEE Trans. Pattern Anal. Mach. Intell. **15**, 401 (1991).
[4] H. Nishimori and K. Y. M. Wong, Phys. Rev. E **60**, 132 (1999).
[5] D. M. Carlucci and J. Inoue, Phys. Rev. E **60**, 2547 (1999).
[6] D. Bolle', H. Rieger, and G. M. Shim, J. Phys. A **27**, 3411 (1994).
[7] D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).
[8] K. Tanaka and T. Morita, Phys. Lett. A **203**, 122 (1995).
[9] T. Morita and K. Tanaka, Physica A **223**, 245 (1995).
[10] Y. Kabashima and D. Saad, Europhys. Lett. **44**, 668 (1998).
[11] T. Tanaka, Phys. Rev. E **58**, 2302 (1998).
[12] N. S. Skantzos and A. C. C. Coolen, J. Phys. A **33**, 1841 (2000).