Statistical-mechanical approach to image processing

2002 J. Phys. A: Math. Gen. 35 R81

(http://iopscience.iop.org/0305-4470/35/37/201)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) R81–R150 PII: S0305-4470(02)33137-8

TOPICAL REVIEW

Statistical-mechanical approach to image processing

Kazuyuki Tanaka

Department of Computer and Mathematical Sciences, Graduate School of Information Sciences, Tohoku University, Aramaki-aza-aoba 04, Aoba-ku, Sendai 980-8579, Japan

E-mail: kazu@statp.is.tohoku.ac.jp

Received 27 July 2002
Published 4 September 2002
Online at stacks.iop.org/JPhysA/35/R81

Abstract

The basic frameworks and techniques of the Bayesian approach to image restoration are reviewed from the statistical-mechanical point of view. First, a few basic notions in digital image processing are explained to convince the reader that statistical mechanics has a close formal similarity to this problem. Second, the basic formulation of the statistical estimation from the observed degraded image by using the Bayes formula is demonstrated. The relationship between Bayesian statistics and statistical mechanics is also listed. Particularly, it is explained that some correlation inequalities on the Nishimori line of the random spin model also play an important role in Bayesian image restoration. Third, the framework of Bayesian image restoration for binary images by means of the Ising model is reviewed. Some practical algorithms for binary image restoration are given by employing the mean-field and the Bethe approximations. Finally, Bayesian image restoration for a grey-level image using the Gaussian model is reviewed, and the Gaussian model is extended to a more practical probabilistic model by introducing the line state to treat the effects of edges. The line state is also extended to quantized values.

PACS numbers: 02.50.−r, 05.20.−y, 05.30.−d, 05.50.+q, 75.10.Nr, 89.70.+c, 03.67.Lx

1. Introduction

The design of filters for digital image processing is one of the major subjects of investigation in information processing. An important motivation is the recent massive introduction of image-processing devices into our daily life with not only personal computers but also digital cameras and mobile phones with image-transmission capabilities. In many image processing systems, the estimation of the original data from the given observed data is a task of primary importance. In the conventional approach to image processing, systems are designed to achieve this goal as accurately as possible. It is first clarified how the system generates the

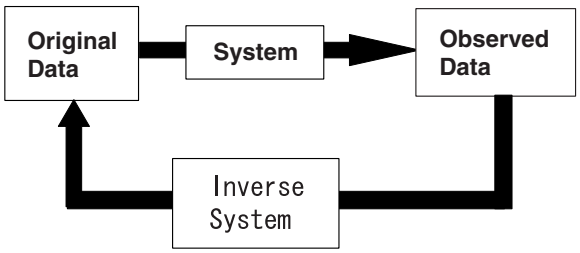

Figure 1. The data-generating system and its inverse system.

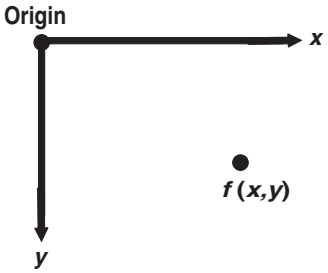

Figure 2. Convention of axis labels used in the present review and a two-dimensional light intensity function $f(x, y)$ to represent a digital image.

observed data from the original data and then the inverse system is constructed as shown in figure 1. Although a system constructed in this way behaves accurately under the same circumstances as the assumed data-generating system, it may not yield the desirable behaviour in other circumstances. In realistic cases where many users use the same system with different input data, robust systems that can process a wide variety of data with uniform reliability are highly desirable. The method of image processing using the ideas of statistical mechanics is expected to provide a way towards such a goal.

Let us explain a few basic notions in digital image processing to convince the reader that statistical mechanics has a close formal similarity to this problem. A digital image is defined on the set of points arranged on a square lattice. At each point, the intensity of light is represented as an integer number or a real number in the digital image data. A monochrome digital image is then expressed as a two-dimensional light intensity function $f_{x,y}$, where x and y denote spatial coordinates and the value of $f_{x,y}$ is proportional to the brightness of the image at the point (x, y) (see figure 2). The image can then be regarded as a matrix whose row and column indices identify a point in the image, and the corresponding matrix element value specifies the intensity of light at that point. The elements of such a digital array are called pixels. An instance of a digital image of size 7×7 with eight grey levels from 0 to 7 is shown in figure 3.

In conventional digital image processing, various linear and nonlinear filters have been designed to suit various purposes. The function of a linear filter is to take the sum of the product of the mask coefficients and the intensities of the pixels. In the 3×3 spatial linear filter, for example, the output $\hat{f}_{x,y}$ at each pixel (x, y) is given as follows:

$$\hat{f}_{x,y} = \frac{1}{9} \sum_{x'=x-1}^{x+1} \sum_{y'=y-1}^{y+1} w_{x-x',y-y'} f_{x',y'}. \tag{1}$$

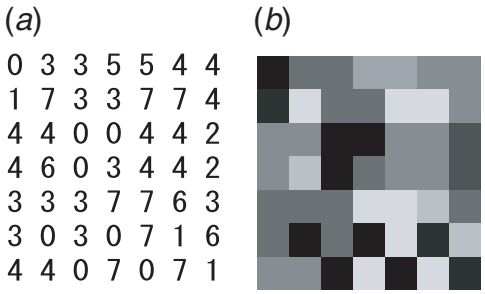

Figure 3. Digital image. (*a*) A 7×7 array with eight grey levels from 0 to 7. (*b*) The corresponding digital image.

Thus the 3×3 spatial linear filter is specified by the mask $\boldsymbol{w}$:

$$\boldsymbol{w} = \begin{pmatrix} w_{-1,-1} & w_{-1,0} & w_{-1,1} \\ w_{0,-1} & w_{0,0} & w_{0,1} \\ w_{1,-1} & w_{1,0} & w_{1,1} \end{pmatrix}. \tag{2}$$

Common examples include the low-pass spatial filter and the high-pass spatial filter with the following masks:

(i) Low-pass spatial filter: $\boldsymbol{w} = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

(ii) High-pass spatial filter: $\boldsymbol{w} = \frac{1}{9}\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$.

Another type of filter is the median filter. In the 3×3 median filter, the output $\hat{f}(x, y)$ is given as the median of the input $\{f_{x',y'} | x' = x - 1, x, x + 1, y' = y - 1, y, y + 1\}$ in the neighbourhood of the pixel (x, y). The median filter is a typical nonlinear filter.

These filters have been designed on heuristic bases. The low-pass filter and the median filter are constructed for noise reduction. Particularly, the median filter can eliminate isolated intensity spikes. On the other hand, the high-pass filter was proposed for the purpose of sharpening the image and is applied to medical imaging, object detection and other purposes. In order to achieve such robustness in conventional image processing, many of the filters are designed usually under the assumption that digital images are generated and degraded only in a specific system, and control parameters in the filters are adjusted so as to give the best performance for the special test data (or supervised data) using statistical methods [1–3]. The performance is then strongly dependent on the types of system and test/supervising data. Let us explain this problem in an explicit example.

Image restoration is the problem where the original image $\{f_{x,y}\}$ is to be estimated from the given degraded image $\{g_{x,y}\}$ as shown in figure 4. The degraded image $\{g_{x,y}\}$ is generated by adding or multiplying some noise

$$g_{x,y} = \sum_{x'} \sum_{y'} A_{x-x',y-y'} f_{x',y'} + n_{x,y} \tag{3}$$

where $n_{x,y}$ is an additive noise and $A_{x-x',y-y'}$ is a blurring noise. To obtain the restored image $\{\hat{f}_{x,y}\}$ by a filter

$$\hat{f}_{x,y} = \sum_{x'=-1}^{1} \sum_{y'=-1}^{1} w_{x-x',y-y'} g_{x',y'} \tag{4}$$

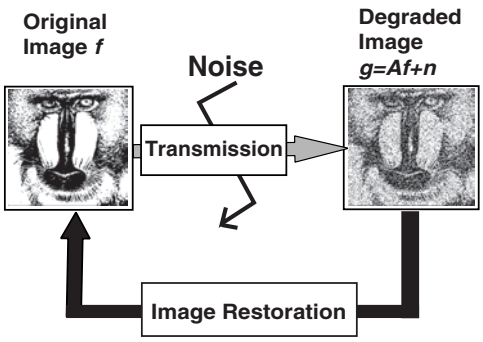

Figure 4. Image degradation and image restoration.

the coefficients $\{w_{x,y}|x = -1, 0, 1, y = -1, 0, 1\}$ have to be determined. In order to determine the coefficients $w_{x,y}$, we prepare some test/supervising images $\{f_{x,y}\}$ and their corresponding degraded images $\{g_{x,y}\}$ generated by equation (3). For these samples, the coefficients $w_{x,y}$ are determined so as to minimize the sample average of $\sum_{x'}\sum_{y'}(\hat{f}_{x,y} - f_{x,y})^2$. This method is clearly optimized to the specific type of noise in the test data. To design more robust systems that can handle very general classes of problems facing end users in practical situations, it is necessary to construct a theory on a more firm and systematic basis.

Many familiar filters for smoothing in image processing are designed on the basis of the property that the intensity $f_{x,y}$ at the pixel (x, y) takes a similar value as those in the same neighbourhood. This property should look analogous to the well-known ferromagnetic property of spin systems. To push the analogy further, it is useful to note that the system for processing images by smoothing can be translated into a classical ferromagnetic spin system by replacing a pixel with a lattice site and by assigning each grey level to a spin state. If we focus on binary image processing, we have only two grey levels, 0 and 1, as the values of the intensity function $f_{x,y}$. By introducing the replacement $s_{x,y} = 2f_{x,y} - 1$, all the possible states at each pixel are replaced by ± 1. In this case, many statistical physicists will expect that the corresponding smoothing filters may be designed using the ideas of physics, particularly, of the Ising model. Possibly, some statistical physicists may even hope that they can achieve a big breakthrough in image processing from their own standpoints, which is of course too much to expect. In practical image processing, we have to consider not only fundamental principles but also some detailed data structures. Moreover, the final goal may sometimes not be image processing but pattern recognition, robot vision, artificial intelligence, neural computation or other engineering applications. The criteria of performance inevitably depend on the situations of practical applications. Nevertheless, some ideas and techniques of physics may be useful to develop a new and universal theoretical foundation of image processing. In particular, statistical mechanics is based on probabilistic theory, and it is very effective to formulate the problem of image processing in terms of probability distribution. In this sense, statistical mechanics provides a natural framework to bridge physics and image processing. More precisely, Bayesian statistics is closely related to statistical mechanics as well as to the foundation of systematic information processing in general.

This idea has been implemented successfully for many years. Derin *et al* [4, 5] first formulated a framework for a probabilistic method of image processing based on Bayesian statistics under the assumptions on the *a priori* information for the original images and degradation process from the original image. Geman and Geman [6] and Jeng and Woods [7, 8] have constructed a probabilistic model with both intensity field and line field (that handles edges) by means of the Bayes formula. The probabilistic method was extended to

forms applicable not only to image restoration but also to edge detection, segmentation, image compression and motion detection [9, 10]. The practical implementation of probabilistic models thus constructed from Bayesian statistics involves massive calculations, and the estimation of the restored image often needs an exponential order of computational complexity except some special cases. It is useful in this respect to note that the probabilistic models for image processing can be regarded as classical spin systems with finite-range interactions, for which approximate methods to treat large-scale problems have been studied for years. Statistical-mechanical methods are applicable to the above-mentioned estimation of the restored image in probabilistic models of image processing. In the pioneering works of Geman and Geman [6] and Jeng and Woods [8], simulated annealing was employed for this purpose. The mean-field approximation was also applied to image processing by many computer scientists [16–19]. Both of these methods reduce the computational complexity from exponential to power behaviour.

Bayesian image analysis is also an interesting subject of investigation in statistics. Some interpretations of probabilistic models for image processing from the standpoint of statistics are given in [20]. Besag [21] formulated a framework for a probabilistic computational method based on Bayesian statistics and proposed a statistical scheme to determine the parameters in the model, which are called *hyperparameters*. The detailed mathematical structure of the his scheme was investigated and extended in some ways by Qian and Titterington [22, 23]. Maximum likelihood estimation is one of the conventional techniques to estimate parameters or hyperparameters from given data in statistics. The scheme of Besag is based on a maximum likelihood estimation and was realized as a probabilistic image processing algorithm by Lakshmanan and Derin [24], Iba [25], Zhang *et al* [26, 27] and Zhou *et al* [28]. Pryce and Bruce [29] and Tanaka [30, 31] investigated a scheme for *hyperparameter* determination by using the mean-field approximation from the statistical-mechanical point of view. In these statistical frameworks, the hyperparameters are determined so as to maximize evidence which is expressed in terms of the free energy of the probabilistic model.

Recently, several statistical physicists also suggested that some techniques and concepts in spin glass theory are applicable to the probabilistic computational method, because the probabilistic models are very similar to classical spin systems, particularly, the Ising model and Potts model [32]. Nishimori and Wong [33] studied the performance of a probabilistic method of image restoration by means of an infinite-range Ising model. They applied the replica method to the calculation. From the standpoint of image processing, the new development in their research is to show that the performance of the probabilistic computational method can be given only by the analytical calculation without doing any numerical experiments. In the conventional statistical method, we have to spend a lot of time doing numerical experiments to estimate the performance with high accuracy. Inoue and Tanaka [34] calculated the statistical average of evidence in Bayesian image restoration based on the infinite-range Ising model by using the replica method and analysed the statistical behaviour in the iteration process of the algorithms as the maximization of evidence.

This review is composed of two basic topics in the Bayesian approach to image restoration, and the basic frameworks and techniques are reviewed. The image restoration using the Bayes formula is a basis of probabilistic image processing and can be extended to other image processing, for example image segmentation, edge detection, image compression and motion detection [9, 10]. One topic is binary image restoration and the other one is grey-level image restoration. In section 2, we demonstrate the basic formulation of the statistical estimation from the observed data by using the Bayes formula and list the relationship between Bayesian statistics and statistical mechanics. Particularly, we explain that some correlation inequalities on the Nishimori line of the random spin model [35–38] also play an important role in Bayesian

image restoration. In section 3, we review the framework of Bayesian image restoration for binary images by means of the Ising model and give some practical algorithms by employing the mean-field and the Bethe approximations [11]. In section 4, we explain Bayesian image restoration for a grey-level image by using the Gaussian model. The Gaussian model is one of the solvable models in statistical mechanics and practical algorithms of image restoration can be derived by means of the exact closed expressions of some statistical quantities of the Gaussian model. In section 5, we introduce the line state in the Gaussian model to treat the effects of edges and demonstrate a few results of the numerical experiments for grey-level image restoration. The line state is extended to quantized values there. Section 6 is devoted to summaries and conclusions.

2. Bayesian statistics in probabilistic image processing

In this section, we summarize a general framework of image restoration by means of Bayesian statistics and some properties related to the spin glass theory, particularly, the Nishimori line in the Ising model with random interactions and random external fields.

Bayesian statistics is based on the Bayes formula. The Bayes formula can often be seen in elementary textbooks of probability and statistics [12–15]. Let us consider two events A and B. The probability that event A occurs is denoted by $\Pr\{A\}$. The conditional probability that event A occurs when event B occurs is denoted by $\Pr\{A|B\}$. The joint probability $\Pr\{A, B\}$ for events A and B is given as follows:

$$\Pr\{A, B\} = \Pr\{A|B\}\Pr\{B\} = \Pr\{B|A\}\Pr\{A\}. \tag{5}$$

From this equality, we obtain the Bayes formula

$$\Pr\{B|A\} = \frac{\Pr\{A|B\}\Pr\{B\}}{\Pr\{A\}} = \frac{\Pr\{A|B\}\Pr\{B\}}{\sum_B \Pr\{A|B\}\Pr\{B\}} \tag{6}$$

where $\sum_B$ takes the summation over all the possible events for B. The second equality is given by using $\Pr\{A\} = \sum_B \Pr\{A|B\}\Pr\{B\}$. The Bayes formula gives the conditional probability $\Pr\{B|A\}$ when the conditional probability $\Pr\{A|B\}$ and the probability $\Pr\{B\}$ are known. In equation (6), event A is given as the observed data and event B corresponds to the original information to estimate. The probability $\Pr\{B\}$ is referred to as the *a priori* probability. The conditional probability $\Pr\{A|B\}$ expresses a process generating the data from the original information. The conditional probability $\Pr\{B|A\}$ is referred to as the *a posteriori* probability. Thus the Bayes formula can be applied to the estimation of the original information from the given data.

In probabilistic image restoration by Bayesian statistics, original and degraded images are generated by the following procedures:

(i) the original image is generated according to the *a priori* process;
(ii) the degraded image is generated from the original image obtained through the previous step in the degradation process.

These procedures are shown in figure 5. These processes can be specified by probabilities $\Pr\{B\}$ for (i) and $\Pr\{A|B\}$ for (ii) or $\boldsymbol{G}$ for A and $\boldsymbol{F}$ for B in figure 5. The relationship to construct the inverse process of the process in figure 5 is given by the Bayes formula as a conditional probability. The inverse process is shown in figure 6. The conditional probability for the inverse process is referred to as the *a posteriori* probability in Bayesian statistics.

In the Bayesian approach to image processing, the Markov random field plays an important role. The Markov random field is the set of random variables in which the state of each pixel is

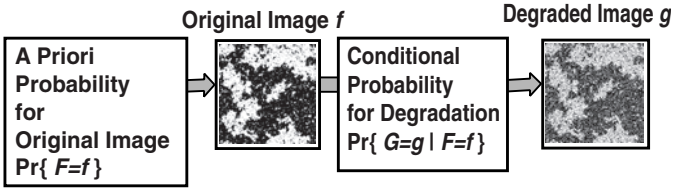

Figure 5. Prior and degradation processes generating the original and degraded images, respectively. The sets of random variables representing the original and the degraded images are denoted by $\boldsymbol{F} = \{F_{x,y}|(x, y) \in \Omega\}$ and $\boldsymbol{G} = \{G_{x,y}|(x, y) \in \Omega\}$, respectively.

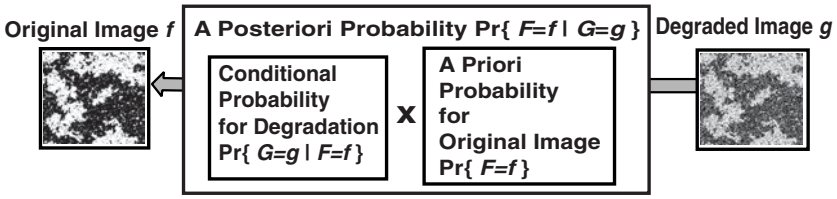

Figure 6. Inverse process of the process in figure 5.

dependent only on the configuration of its neighbourhood pixels. The two-dimensional Ising model with nearest-neighbour interactions is a typical example of a Markov random field. Generally, a Markov random field can be regarded as a two-dimensional classical spin system with interactions of finite range. Moreover, the conditional probability for the degradation process can be expressed as a term of external fields in the classical spin system, and the observed data are represented as the coefficients of the external fields.

Nishimori and Wong [33] proposed that the performance of an infinite-range model of image restoration can be estimated statistically by introducing an idea from spin glass theory. Such a viewpoint did not exist in the field of conventional image processing. Computer scientists and system engineers perform a large number of tests for all the considerable cases to check the performance of systems they designed consuming large amount of computational resources. Nishimori and Wong's proposal suggested that the concepts and techniques in spin glass theory may provide a performance check without test experiments. Similar ideas have also been applied to the performance checks in error-correcting codes [39] and mobile communications [40].

In the present section, we first give a general framework of Bayesian image restoration and define the Markov random field explicitly by using equations. We explain the equivalence between the Markov random field and Gibbs canonical distribution, the relationship between the Kullback–Leibler divergence and the free energy, and the hyperparameter determination based on the maximum likelihood estimation. Moreover, we give a framework of performance check in the probabilistic method of image restoration and mention that the framework has a close formal similarity to the spin glass theory.

2.1. Bayesian image processing

We consider an image on the square lattice $\Omega = \{(x, y)|x = 1, 2, \ldots, L_x, y = 1, 2, \ldots, L_y\}$. The lattice is assumed to consist of $|\Omega|$ pixels and to satisfy the periodic boundary conditions, so that the lattice is on a torus. The configurations of true original and given degraded images are represented by $\boldsymbol{f} = \{f_{x,y}|(x, y) \in \Omega\}$ and $\boldsymbol{g} = \{g_{x,y}|(x, y) \in \Omega\}$, respectively. In a binary

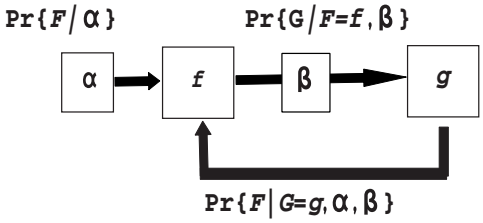

Figure 7. Image restoration by means of the Bayes formula. Here α and β are hyperparameters.

image, the set of all possible states at each pixel is denoted by $\{\pm 1\}$. The state +1 means *black* and the state -1 means *white*[1]. In recovering the original image $\boldsymbol{f}$ from the given degraded image $\boldsymbol{g}$, we use some *a priori* properties of the original image $\boldsymbol{f}$. We express the set of random variables representing the original and the degraded images by $\boldsymbol{F} = \{F_{x,y}|(x, y) \in \Omega\}$ and $\boldsymbol{G} = \{G_{x,y}|(x, y) \in \Omega\}$, respectively. In the Bayesian approach to image processing, the set of random variables is often called the *random field*.

In the present review, the basic framework of probabilistic image processing is given by means of the Bayes formula. We have illustrated the concept of the Bayes formula at the beginning of the present section by using figures 5 and 6. Now we explain the explicit form of the Bayes formula by means of the notations of the random fields $\boldsymbol{F}$ and $\boldsymbol{G}$ in the present subsection. The probability that the original image is $\boldsymbol{f}$, $\Pr\{\boldsymbol{F} = \boldsymbol{f}\}$, is called the *a priori* probability of an original image. In the Bayes formula, the *a posteriori* probability $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}\}$, that the original image is $\boldsymbol{f}$ when the given degraded image is $\boldsymbol{g}$, is expressed as

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}\} = \frac{\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\}\Pr\{\boldsymbol{F} = \boldsymbol{f}\}}{\sum_{\boldsymbol{z}} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{z}\}\Pr\{\boldsymbol{F} = \boldsymbol{z}\}} \tag{7}$$

where the summation $\sum_{\boldsymbol{z}}$ is taken over all possible configurations of images $\boldsymbol{z} = \{z_{x,y}|(x, y) \in \Omega\}$. The probability $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\}$ is the conditional probability that the degraded image is $\boldsymbol{g}$ when the original image is $\boldsymbol{f}$ and denotes the degradation process that produces the degraded image from the original image (see figure 7).

We have some criteria for obtaining the estimate of the original image $\boldsymbol{f}$, $\hat{\boldsymbol{f}} = \{\hat{f}_{x,y}|(x, y) \in \Omega\}$, as follows:

(i) Maximum *a posteriori* (MAP) estimation

$$\hat{\boldsymbol{f}} = \arg\max_{\boldsymbol{z}} \Pr\{\boldsymbol{F} = \boldsymbol{z}|\boldsymbol{G} = \boldsymbol{g}\}. \tag{8}$$

(ii) Maximum posterior marginal (MPM) estimation

$$\hat{f}_{x,y} = \arg\max_{\zeta} \Pr\{F_{x,y} = \zeta|\boldsymbol{G} = \boldsymbol{g}\}. \tag{9}$$

(iii) Thresholded posterior mean (TPM) estimation

$$\hat{f}_{x,y} = \arg\min_{\zeta} \left(\zeta - \sum_{\boldsymbol{z}} z_{x,y} \Pr\{\boldsymbol{F} = \boldsymbol{z}|\boldsymbol{G} = \boldsymbol{g}\}\right)^2. \tag{10}$$

[1] In a practical digital image, the set of all possible states at each pixel is denoted by $\{0, 1, 2, \ldots, 255\}$. The state 0 means *black* and the state 255 means *white*.

Here the notation $\arg\max_z a(z)$ means a maximizer z of the function $a(z)$, and $\Pr\{F_{x,y} = \zeta | \boldsymbol{G} = \boldsymbol{g}\}$ is the posterior marginal probability defined by

$$\Pr\{F_{x,y} = f_{x,y} | \boldsymbol{G} = \boldsymbol{g}\} \equiv \sum_{\boldsymbol{z}} \Pr\{\boldsymbol{F} = \boldsymbol{z} | \boldsymbol{G} = \boldsymbol{g}\} \delta_{z_{x,y}, f_{x,y}} \tag{11}$$

or

$$\Pr\{F_{x,y} = f_{x,y} | \boldsymbol{G} = \boldsymbol{g}\} \equiv \sum_{\boldsymbol{f}_{\Omega \setminus (x,y)}} \Pr\{\boldsymbol{F} = \boldsymbol{f} | \boldsymbol{G} = \boldsymbol{g}\} \tag{12}$$

where $\delta_{a,b}$ is the Kronecker delta. We remark that the MPM estimation and the TPM estimation are equivalent to each other in binary images.

2.2. Gibbs canonical distributions

For a specified energy function $U(\boldsymbol{f})$ and a fixed value of temperature $T(>0)$, the Gibbs canonical distribution

$$\rho(\boldsymbol{f}) = \frac{\exp\left(-\frac{1}{T} U(\boldsymbol{f})\right)}{\sum_z \exp\left(-\frac{1}{T} U(\boldsymbol{z})\right)} \qquad T > 0 \tag{13}$$

satisfies the minimization of the free energy as follows,

$$\rho(\boldsymbol{f}) = \arg\min_{\phi} \mathcal{F}[\phi] \tag{14}$$

where

$$\mathcal{F}[\phi] \equiv \sum_{\boldsymbol{z}} \phi(\boldsymbol{z}) \left(U(\boldsymbol{z}) + T \sum_{\boldsymbol{z}} \ln \phi(\boldsymbol{z}) \right). \tag{15}$$

Here the notation $\arg\min_z a(z)$ means a minimizer z of the function $a(z)$. The quantities $\mathcal{F}[\rho]$ and $\mathcal{S}[\rho] \equiv -\sum_z \rho(z) \ln \rho(z)$ are the free energy and the entropy in the statistical-mechanical model with the energy function $U(\boldsymbol{f})$, respectively. The free energy $\mathcal{F}[\rho]$ is expressed in terms of the energy function $U(\boldsymbol{f})$:

$$F \equiv \mathcal{F}[\rho] = -T \ln \left(\sum_{\boldsymbol{z}} \exp\left(-\frac{1}{T} U(\boldsymbol{z}) \right) \right). \tag{16}$$

The Kullback–Leibler divergence $\mathcal{D}_{\mathrm{KL}}[\phi \| \rho]$ can be written in terms of $\mathcal{F}[\phi]$:

$$\mathcal{D}_{\mathrm{KL}}[\phi \| \rho] \equiv \sum_{\boldsymbol{z}} \phi(\boldsymbol{z}) \ln \left(\frac{\phi(\boldsymbol{z})}{\rho(\boldsymbol{z})} \right) = \mathcal{F}[\phi] + \sum_{\boldsymbol{z}} \exp\left(-\frac{1}{T} U(\boldsymbol{z}) \right). \tag{17}$$

The Kullback–Leibler divergence $\mathcal{D}_{\mathrm{KL}}[\phi \| \rho]$ is equal to zero if and only if $\phi(\boldsymbol{f}) = \rho(\boldsymbol{f})$. The minimization of the free energy corresponds to the minimization of the Kullback–Leibler divergence.

2.3. Markov random fields

In the Bayesian approach to image processing, the random field in the *a priori* probability is often assumed to be a Markov random field [4, 5]. A Markov random field is a set of random variables in which the state of a pixel (x, y) is dependent only on the configuration of its neighbourhood of the pixel (x, y). In the present paper, for simplicity, we consider only the set of the nearest-neighbour pixels as the neighbourhood in the Markov random field.

Now we explain the definition of the Markov random field as an explicit expression. The set of nearest-neighbour pixels of the pixel (x, y) is denoted by $\mathbf{c}_{x,y} \equiv \{(x \pm 1, y), (x, y \pm 1)\}$.

In this case, the Markov random field is defined as the set of random variables satisfying the following relation,

$$\Pr\{F_{x,y} = f_{x,y} | \boldsymbol{F}_{\Omega\setminus(x,y)} = \boldsymbol{f}_{\Omega\setminus(x,y)}\} = \Pr\{F_{x,y} = f_{x,y} | F_{x',y'} = f_{x',y'}, (x', y') \in \mathbf{c}_{x,y}\} \tag{18}$$

for every pixel $(x, y) \in \Omega$. Here, $\Pr\{F_{x,y} = f_{x,y} | \boldsymbol{F}_{\Omega\setminus(x,y)} = \boldsymbol{f}_{\Omega\setminus(x,y)}\}$ is defined by

$$\Pr\{F_{x,y} = f_{x,y} | \boldsymbol{F}_{\Omega\setminus(x,y)} = \boldsymbol{f}_{\Omega\setminus(x,y)}\} \equiv \frac{\Pr\{\boldsymbol{F} = \boldsymbol{f}\}}{\Pr\{\boldsymbol{F}_{\Omega\setminus(x,y)} = \boldsymbol{f}_{\Omega\setminus(x,y)}\}} \tag{19}$$

and $\boldsymbol{F}_{\Omega\setminus(x,y)} \equiv \{F_{x',y'} | (x', y') \in \Omega\setminus(x, y)\}$.

It can be shown that the *a priori* probability with a Markov random field is reduced to a Gibbs canonical distribution only with finite-range interactions. We can prove this fact as follows: first we rewrite a probability that satisfies equation (18) in the following form:

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}\} = \exp(-U(\boldsymbol{f})) \qquad U(\boldsymbol{f}) \equiv -\ln(\Pr\{\boldsymbol{F} = \boldsymbol{f}\}). \tag{20}$$

By using equations (18) and (19) iteratively, the energy function $U(\boldsymbol{f})$ in equation (20) can be transformed as follows:

$$\begin{aligned} U(\boldsymbol{f}) = & -\ln(\Pr\{F_{x,y} = f_{x,y} | \boldsymbol{F}_{\Omega\setminus(x,y)} = \boldsymbol{f}_{\Omega\setminus(x,y)}\}) - \ln(\Pr\{\boldsymbol{F}_{\Omega\setminus(x,y)} = \boldsymbol{f}_{\Omega\setminus(x,y)}\}) \\ = & -\ln(\Pr\{F_{x,y} = f_{x,y} | F_{x',y'} = f_{x',y'}, (x', y') \in \mathbf{c}_{x,y}\}) \\ & - \ln(\Pr\{\boldsymbol{F}_{\Omega\setminus(x,y)} = \boldsymbol{f}_{\Omega\setminus(x,y)}\}) \\ = & -\sum_{(x,y)\in\Omega} \ln(\Pr\{F_{x,y} = f_{x,y} | F_{x',y'} = f_{x',y'}, (x', y') \in \mathbf{c}_{x,y}\}). \end{aligned} \tag{21}$$

We then obtain the following equation,

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}\} = \exp\left(-\sum_{(x,y)\in\Omega} U_{x,y}(f_{x,y} | f_{x',y'}, (x', y') \in \mathbf{c}_{x,y})\right) \tag{22}$$

where

$$U_{x,y}(f_{x,y} | f_{x',y'}, (x', y') \in \mathbf{c}_{x,y}) \equiv \ln(\Pr\{F_{x,y} = f_{x,y} | F_{x',y'} = f_{x',y'}, (x', y') \in \mathbf{c}_{x,y}\}). \tag{23}$$

Equation (22) means that the probability $\Pr\{\boldsymbol{F} = \boldsymbol{f}\}$ that is a Markov random field has the form of a Gibbs canonical distribution with finite-range interactions.

If the neighbourhood pixels $\mathbf{c}_{x,y}$ are specified by

$$\mathbf{c}_{x,y} \equiv \{(x - 1, y), (x + 1, y), (x, y - 1), (x, y + 1)\} \tag{24}$$

equation (18) can be written as follows:

$$\begin{aligned} \Pr\{F_{x,y} = f_{x,y} | \boldsymbol{F}_{\Omega\setminus(x,y)} = \boldsymbol{f}_{\Omega\setminus(x,y)}\} = & \Pr\{F_{x,y} = f_{x,y} | F_{x-1,y} = f_{x-1,y}, F_{x+1,y} \\ & = f_{x+1,y}, F_{x,y-1} = f_{x,y-1}, F_{x,y+1} = f_{x,y+1}\}. \end{aligned} \tag{25}$$

More general discussions about the definition of Markov random fields and the equivalence between the Markov random fields and Gibbs distributions for any lattice with general graph structure are given in [5].

2.4. Hyperparameter, maximum likelihood estimation and evidence framework

The probabilities $\Pr\{\boldsymbol{F} = \boldsymbol{f}\}$ and $\Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{f}\}$ have some model parameters that are referred to as hyperparameters. The hyperparameters of the *a priori* probability and the degradation process are denoted by α and β, respectively (see figure 7). Now, we express the degradation process and the *a priori* probability including the hyperparameters α and β in terms of the notations $\Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{f}, \beta\}$ and $\Pr\{\boldsymbol{F} = \boldsymbol{f} | \alpha\}$, respectively.

We explain the role of the hyperparameters α and β in a simple example. For $L_x = 2$ and $L_y = 1$, we consider a system Ω consisting of two pixels. The *a priori* probability distribution and the degradation process are assumed to be

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\} = \frac{\exp(\alpha f_{1,1} f_{1,2})}{4\cosh(\alpha)} \tag{26}$$

$$\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \beta\} = \frac{\exp(\beta g_{1,1} f_{1,1} + \beta g_{1,2} f_{1,2})}{4\cosh(\beta)\cosh(\beta)}. \tag{27}$$

By substituting equations (26) and (27) into equation (7), the *a posteriori* probability distribution is derived as follows:

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \beta\} = \frac{\exp(\beta g_{1,1} f_{1,1} + \beta g_{1,2} f_{1,2} + \alpha f_{1,1} f_{1,2})}{\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\exp(\beta g_{1,1}\zeta + \beta g_{1,2}\zeta' + \alpha\zeta\zeta')}. \tag{28}$$

In the MAP estimation, the restored image $\hat{\boldsymbol{f}} = \{\hat{f}_{1,1}, \hat{f}_{1,2}\}$ is given by

$$(\hat{f}_{1,1}, \hat{f}_{1,2}) = \arg\min_{\zeta=\pm1,\zeta'=\pm1}(-\beta g_{1,1}\zeta - \beta g_{1,2}\zeta' - \alpha\zeta\zeta'). \tag{29}$$

For $\alpha \gg \beta$, the restored image $\hat{\boldsymbol{f}}$ satisfies $\hat{f}_{1,1} = \hat{f}_{1,2}$. For $\alpha \ll \beta$, the restored image $\hat{\boldsymbol{f}}$ is obtained as $\hat{f}_{1,1} = g_{1,1}$ and $\hat{f}_{1,2} = g_{1,2}$. Hence, it is important how to choose the values of hyperparameters which should be determined only from the given degraded image without using the original image.

By means of the most basic framework of maximum likelihood estimation, we demonstrate how to estimate the hyperparameters when the original image is not known. Since we never know the original image $\boldsymbol{f}$ in practical applications, this scheme can be regarded as a design of optimal probabilistic model of the original image when we know the answer. In this situation, the hyperparameters α and β are determined so as to maximize quantities $\Pr\{\boldsymbol{F} = \boldsymbol{f}\}$ and $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\}$ as follows,

$$(\hat{\alpha}, \hat{\beta}) = \arg\max_{(\alpha,\beta)} \Pr\{\boldsymbol{F} = \boldsymbol{f}, \boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\} \tag{30}$$

in the original standpoint of maximum likelihood estimation. In this framework, the joint probability $\Pr\{\boldsymbol{F} = \boldsymbol{f}, \boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\} \equiv \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \beta\}\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\}$ is regarded as the likelihood of the hyperparameters α and β when the original image $\boldsymbol{f}$ and the degraded image $\boldsymbol{g}$ are given. Equation (30) is reduced to the following simultaneous equations:

$$\hat{\alpha} = \arg\max_{\alpha} \Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\} \tag{31}$$

$$\hat{\beta} = \arg\max_{\beta} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \beta\}. \tag{32}$$

This hyperparameter determination scheme was proposed by Besag [21].

By generalizing the above-mentioned hyperparameter estimation method, we explain how to determine the hyperparameters when we have no direct information on the original image $\boldsymbol{f}$. By combining equation (8) with equation (30), Lakshmanan and Derin [24] implemented the following simultaneous scheme,

$$(\hat{\alpha}, \hat{\beta}) = \arg\max_{(\alpha,\beta)} \Pr\{\boldsymbol{F} = \hat{\boldsymbol{f}}, \boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\} \tag{33}$$

$$\hat{\boldsymbol{f}} = \arg\max_{\boldsymbol{z}} \Pr\{\boldsymbol{F} = \boldsymbol{z}|\boldsymbol{G} = \boldsymbol{g}, \hat{\alpha}, \hat{\beta}\} \tag{34}$$

in the context of a segmentation problem. We note that equation (31) gives the scheme to determine the hyperparameters in the *a priori* probability from the original image whereas equations (33) and (34) do not refer to the original image directly.

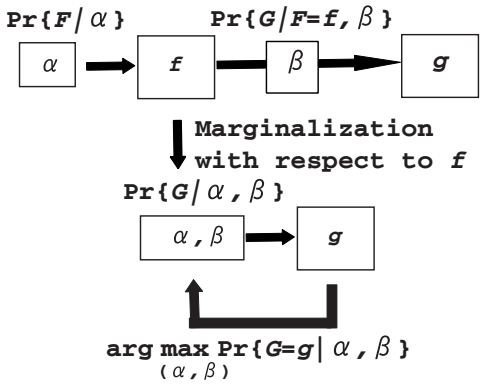

Figure 8. Evidence framework of image restoration by means of the Bayes formula.

The framework of Lakshmanan and Derin was extended to more generalized one that is referred to as a maximization of marginal likelihood [25–29]. In the maximization of marginal likelihood, the hyperparameters α and β are determined only from a given degraded image $\boldsymbol{g}$. From the standpoint of statistics, the hyperparameters α and β are determined so as to maximize a quantity $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}$ as follows,

$$(\hat{\alpha}, \hat{\beta}) = \arg\max_{(\alpha,\beta)} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\} \tag{35}$$

where

$$\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\} \equiv \sum_{\boldsymbol{z}} \Pr\{\boldsymbol{F} = \boldsymbol{z}, \boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\} = \sum_{\boldsymbol{z}} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{z}, \beta\} \Pr\{\boldsymbol{F} = \boldsymbol{z}|\alpha\}. \tag{36}$$

This situation is shown in figure 8. In this framework, the probability $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}$ is given by marginalizing the joint probability $\Pr\{\boldsymbol{F} = \boldsymbol{f}, \boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}$ over all the possible original images $\boldsymbol{f}$ and can be regarded as a marginal likelihood for α and β when the degraded image $\boldsymbol{g}$ is given, which is referred to as type II likelihood [41, 42], Akaike–Bayes information criteria (ABIC) [43] or evidence [29, 44–47] in statistics. Hereafter, $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}$ will be referred to as evidence. For example, when we consider the conditional probability $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \beta\}$ in equation (27) and the *a priori* probability $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\}$ in equation (26), the evidence $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}$ can be written for a binary image as follows:

$$\begin{aligned}\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\} &= \sum_{z_{1,1}=\pm 1} \sum_{z_{1,2}=\pm 1} \frac{\exp(\beta g_{1,1} z_{1,1} + \beta g_{1,2} z_{1,2} + \alpha z_{1,1} z_{1,2}))}{16 \cosh(\alpha) \cosh(\beta) \cosh(\beta)} \\ &= \frac{\exp(\alpha) \cosh(\beta g_{1,1} + \beta g_{1,2}) + \exp(-\alpha) \cosh(\beta g_{1,1} - \beta g_{1,2})}{8 \cosh(\alpha) \cosh(\beta) \cosh(\beta)}.\end{aligned} \tag{37}$$

It should be clear now that the hyperparameters α and β can be determined only from knowledge of the degraded image.

2.5. Configuration average and statistical performance

The probabilistic model for image restoration corresponds to a correlated random field model with the random external fields $\boldsymbol{g}$ generated by $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha^*, \beta^*\} = \sum_{\boldsymbol{z}} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{z}, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{z}|\alpha^*\}$ from the standpoint of spin glass theory. Here, we remark that the values of hyperparameters α and β, at which the observed degraded image $\boldsymbol{g}$ has been generated from

the *a priori* probability and the degradation process, are denoted by α^* and β^* so that α^* and β^* are the true values of the hyperparameters α and β in equation (36), respectively. In spin glass theory, we consider the random average of quantities $A(\boldsymbol{f})$ and $B(\boldsymbol{f}, \boldsymbol{g})$ defined by

$$\langle A(\boldsymbol{F})|\boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\rangle \equiv \sum_{\boldsymbol{f}} A(\boldsymbol{f}) \Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\} \tag{38}$$

and

$$\langle B(\boldsymbol{F}, \boldsymbol{G})|\alpha^*, \beta^*\rangle \equiv \sum_{\boldsymbol{f}} \sum_{\boldsymbol{g}} B(\boldsymbol{f}, \boldsymbol{g}) \Pr\{\boldsymbol{F} = \boldsymbol{f}, \boldsymbol{G} = \boldsymbol{g}|\alpha^*, \beta^*\}. \tag{39}$$

Here $\langle A(\boldsymbol{F})|\boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\rangle$ means the conditional expectation value of $A(\boldsymbol{f})$ due to the *a posteriori* probability $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\}$ when the degraded image $\boldsymbol{g}$ and the values of hyperparameters, α^* and β^*, are given, and $\langle B(\boldsymbol{F}, \boldsymbol{G})|\alpha^*, \beta^*\rangle$ is the conditional expectation value of $B(\boldsymbol{f}, \boldsymbol{g})$ according to the joint probability $\Pr\{\boldsymbol{F}, \boldsymbol{G}|\alpha^*, \beta^*\}$. $\boldsymbol{F}$ may be identified with spin variables and $\boldsymbol{G}$ with random fields.

Let us introduce the following quantity as a measure of performance of image restoration,

$$\begin{aligned}\mathcal{M}_{x,y}(\alpha, \beta) &\equiv \sum_{\boldsymbol{f}} \sum_{\boldsymbol{g}} (f_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha, \beta))^2 \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \alpha^*, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha^*\} \\ &= \langle (F_{x,y} - h_{x,y}(\boldsymbol{G}, \alpha, \beta))^2|\alpha^*, \beta^*\rangle \end{aligned} \tag{40}$$

where $h_{x,y}(\boldsymbol{g}, \alpha, \beta)$ is the estimation at pixel (x, y) for the original image $\boldsymbol{f}$ by the *a posteriori* probability for arbitrary values of hyperparameters α and β. The value of $h_{x,y}(\boldsymbol{g}, \alpha, \beta)$ has been obtained from MAP estimation (8), MPM estimation (9), TPM estimation (10) or other appropriate methods. We have another useful quantity defined by

$$\mathcal{L}(\alpha, \beta) = \sum_{\boldsymbol{g}} \ln(\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}) \Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha^*, \beta^*\} = \langle \ln(\Pr\{\boldsymbol{G}|\alpha, \beta\})|\alpha^*, \beta^*\rangle. \tag{41}$$

This quantity is the statistical average of the logarithm of evidence $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}$ with respect to the degraded image $\boldsymbol{g}$ that is produced by the degradation process $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \beta^*\}$ when the original image $\boldsymbol{f}$ is produced by the *a priori* probability $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha^*\}$. For the statistical quantity $\mathcal{L}(\alpha, \beta)$, we have the following rigorous inequality,

$$\mathcal{L}(\alpha, \beta) \leqslant \mathcal{L}(\alpha^*, \beta^*) \tag{42}$$

because of

$$\begin{aligned}\mathcal{L}(\alpha, \beta) - \mathcal{L}(\alpha^*, \beta^*) &= \sum_{\boldsymbol{g}} \left(\ln \left(\frac{\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}}{\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha^*, \beta^*\}} \right) \right) \Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha^*, \beta^*\} \\ &\leqslant \ln \left(\sum_{\boldsymbol{g}} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\} \right) = 0 \end{aligned} \tag{43}$$

due to Jensen's inequality. This rigorous inequality means that the maximization of log-evidence can give us the original values of hyperparameters α and β.

If we restrict ourselves to binary images, the MPM estimation (9) and the TPM estimation (10) are reduced to the same formula,

$$h_{x,y}(\boldsymbol{g}, \alpha, \beta) = \operatorname{sign}(\langle F_{x,y}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \beta\rangle). \tag{44}$$

In Bayesian statistics, we have to choose an appropriate criterion to determine the estimate of the original image. Which criterion is best depends on the degradation process. Nishimori and Wong [33] derived a rigorous inequality

$$\mathcal{M}_{x,y}(\alpha, \beta) \geqslant \mathcal{M}_{x,y}(\alpha^*, \beta^*) \tag{45}$$

in the Bayesian approach with the MPM estimation for binary image restoration. By using $|a| = a^2/|a|$ and $\mathrm{sign}(a) = a/|a|$ for any non-zero real number, inequality (45) can be proved as follows:

$$\begin{aligned}
\mathcal{M}_{x,y}(\alpha,\beta) &= 2 - 2\langle F_{x,y}\,\mathrm{sign}(\langle F_{x,y}|\boldsymbol{G}=\boldsymbol{g},\alpha,\beta\rangle)|\alpha^*,\beta^*\rangle \\
&\geqslant 2 - 2\sum_{\boldsymbol{g}}\left|\sum_{\boldsymbol{f}} f_{x,y}\Pr\{\boldsymbol{G}=\boldsymbol{g}|\boldsymbol{F}=\boldsymbol{f},\beta^*\}\Pr\{\boldsymbol{F}=\boldsymbol{f}|\alpha^*\}\right| \\
&= 2 - 2\sum_{\boldsymbol{g}}\left(\sum_{\boldsymbol{f}} f_{x,y}\Pr\{\boldsymbol{G}=\boldsymbol{g}|\boldsymbol{F}=\boldsymbol{f},\beta^*\}P_x\{\boldsymbol{F}=\boldsymbol{f}|\alpha^*\}\right) \\
&\quad\times\left(\frac{\langle F_{x,y}|\boldsymbol{G}=\boldsymbol{g},\alpha^*,\beta^*\rangle}{|\langle F_{x,y}|\boldsymbol{G}=\boldsymbol{g},\alpha^*,\beta^*\rangle|}\right) \\
&= 2 - 2\sum_{\boldsymbol{g}}\left(\sum_{\boldsymbol{f}} f_{x,y}\Pr\{\boldsymbol{G}=\boldsymbol{g}|\boldsymbol{F}=\boldsymbol{f},\beta^*\}P_x\{\boldsymbol{F}=\boldsymbol{f}|\alpha^*\}\right) \\
&\quad\times\mathrm{sign}(\langle F_{x,y}|\boldsymbol{G}=\boldsymbol{g},\alpha^*,\beta^*\rangle) \\
&= 2 - 2\langle F_{x,y}\,\mathrm{sign}(h_{x,y}(\boldsymbol{g},\alpha^*,\beta^*))|\alpha^*,\beta^*\rangle \\
&= M_{x,y}(\alpha^*,\beta^*). \qquad (46)
\end{aligned}$$

The mathematical structure in this inequality is similar to that in the correlation inequalities at the Nishimori point in the $\pm J$ model in spin glass theory [35–38]. Hence, we can regard the point $(\alpha,\beta) = (\alpha^*,\beta^*)$ as corresponding to the Nishimori line in spin glass theory. In the present review, we sometimes refer to the point $(\alpha,\beta) = (\alpha^*,\beta^*)$ as the Nishimori point in the Bayesian image restoration. The quantity $\mathcal{M}_{x,y}(\alpha,\beta)$ corresponds to a statistical average with respect to random external fields and random interactions in the spin glass model.

It should be remarked that the probabilistic models for Bayesian image restoration have no gauge invariance, whereas gauge invariance plays a very important role in the argument related to the Nishimori line in the $\pm J$ model. Nevertheless, we have some correlation inequalities that can be verified without using the gauge invariance in the $\pm J$ model. Inequality (45), which is derived without using the gauge invariance, shows that the best restored image is obtained at the Nishimori point $(\alpha,\beta) = (\alpha^*,\beta^*)$. The relationship between Bayesian statistics and the Nishimori line in classical spin systems with random interactions and random external fields was discussed in more detail from the standpoint of probabilistic information processing in [48–50].

We have demonstrated that the MPM estimation gives us the best restored image. However, many computer scientists generally use the MAP estimation for data classification or learning from data by means of Bayes statistics. This standpoint is based on the property that the image from MAP estimation is the most probable one in the following sense [56]. We denote the random fields of the original image, $\boldsymbol{F}$, in the *a priori* probability distribution $\Pr\{\boldsymbol{F}|\alpha^*\}$ and the *a posteriori* probability distribution $\Pr\{\boldsymbol{F}|\boldsymbol{G}=\boldsymbol{g},\alpha^*\}$ by the notation $\boldsymbol{F}^{\mathrm{prior}}$ and $\boldsymbol{F}^{\mathrm{posterior}}$, respectively, and introduce the following probability to understand this fact:

$$\begin{aligned}
\Pr\{\boldsymbol{F}^{\mathrm{prior}}=\boldsymbol{F}^{\mathrm{posterior}}\} \equiv \sum_{\boldsymbol{f}}\sum_{\boldsymbol{g}}&\Pr\{\boldsymbol{F}^{\mathrm{posterior}}=\boldsymbol{f}|\boldsymbol{G}=\boldsymbol{g},\alpha^*,\beta^*\} \\
&\times\Pr\{\boldsymbol{G}=\boldsymbol{g}|\boldsymbol{F}^{\mathrm{prior}}=\boldsymbol{f},\beta^*\}\Pr\{\boldsymbol{F}^{\mathrm{prior}}=\boldsymbol{f}|\alpha^*\}. \qquad (47)
\end{aligned}$$

We remark that $\Pr\{\boldsymbol{F}^{\mathrm{prior}}=\boldsymbol{F}^{\mathrm{posterior}}\}$ means the probability of the event '$\boldsymbol{F}^{\mathrm{posterior}}=\boldsymbol{F}^{\mathrm{prior}}$', that is, an image generated from the *a posteriori* probability $\Pr\{\boldsymbol{F}^{\mathrm{posterior}}|\boldsymbol{G}=\boldsymbol{g},\alpha^*,\beta^*\}$ is

equal to the original image generated from the *a priori* probability $\Pr\{\boldsymbol{F}^{\rm prior}|\alpha^*\}$. The above statement can be expressed in terms of the following inequality [56],

$$\Pr\{\boldsymbol{F}^{\rm prior} = \boldsymbol{F}^{\rm posterior}\} \leqslant \Pr\{\boldsymbol{F}^{\rm prior} = \hat{\boldsymbol{f}}^{\rm MAP}\} \tag{48}$$

where

$$\begin{aligned}\hat{\boldsymbol{f}}^{\rm MAP} &\equiv \arg\max_{\boldsymbol{z}} \Pr\{\boldsymbol{F}^{\rm posterior} = \boldsymbol{z}|\boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\} \\ &= \arg\max_{\boldsymbol{z}} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F}^{\rm posterior} = \boldsymbol{z}, \beta^*\} \Pr\{\boldsymbol{F}^{\rm posterior} = \boldsymbol{z}|\alpha^*\}. \end{aligned} \tag{49}$$

The proof is given as follows:

$$\begin{aligned}\Pr\{\boldsymbol{F}^{\rm prior} = \boldsymbol{F}^{\rm posterior}\} &\leqslant \sum_{\boldsymbol{f}}\sum_{\boldsymbol{g}} \Pr\{\boldsymbol{F}^{\rm posterior} = \hat{\boldsymbol{f}}^{\rm MAP}|\boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\} \\ &\quad \times \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F}^{\rm prior} = \boldsymbol{f}, \beta^*\} \Pr\{\boldsymbol{F}^{\rm prior} = \boldsymbol{f}|\alpha^*\} \\ &\leqslant \sum_{\boldsymbol{f}}\sum_{\boldsymbol{g}} \Pr\{\boldsymbol{F}^{\rm posterior} = \hat{\boldsymbol{f}}^{\rm MAP}|\boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\} \\ &\quad \times \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F}^{\rm prior} = \hat{\boldsymbol{f}}^{\rm MAP}, \beta^*\} \Pr\{\boldsymbol{F}^{\rm prior} = \hat{\boldsymbol{f}}^{\rm MAP}|\alpha^*\} \\ &= \Pr\{\boldsymbol{F}^{\rm prior} = \hat{\boldsymbol{f}}^{\rm MAP}\}. \end{aligned} \tag{50}$$

Inequality (48) suggests that the MAP estimation may give the best restored image in the sense that it maximizes $\Pr\{\boldsymbol{F}^{\rm prior} = \boldsymbol{F}^{\rm posterior}\}$. The probability $\Pr\{\boldsymbol{F}^{\rm prior} = \boldsymbol{F}^{\rm posterior}\}$ measures whether or not the image obtained by the MAP estimation agrees with the original image at all pixels. The quantity $\mathcal{M}_{x,y}(\alpha, \beta)$, on the other hand, is a measure of agreement at a single pixel. Both equations (45) and (48) are mathematically rigorous inequalities. Although computer scientists often adopt the MAP estimation due to the criterion (48), they should resort to simulated annealing to implement the MAP estimation because almost no models of Bayesian image processing have been solved exactly. On the other hand, statistical physicists have much experience in constructing practical algorithms which can be used to implement the MPM estimation (9) and the TPM estimation (10) approximately. Moreover, statistical-mechanical techniques for spin glass theory give us explicit formulae of statistical performance $\mathcal{M}_{x,y}(\alpha, \beta)$ for the MPM estimation and the TPM estimation in some types of models. It is therefore reasonable to consider that the MPM estimation and the TPM estimation are more practical schemes than the MAP estimation if we combine Bayesian image analysis with statistical-mechanical techniques.

In the present subsection, we have chosen the quantity $D(\hat{\boldsymbol{f}}, \boldsymbol{f}) \equiv \sum_{(x,y)\in\Omega}(\hat{f}_{x,y} - f_{x,y})^2$ as the distance between the restored image $\hat{\boldsymbol{f}}$ and the original image $\boldsymbol{f}$. We can ask a slightly different question: are there any choices for the estimator $h_{x,y}(\boldsymbol{g}, \alpha, \beta)$ other than the right-hand side of equation (44) that maximize the average distance $\langle D(\boldsymbol{F}, \boldsymbol{h}(\boldsymbol{G}, \alpha^*, \beta^*))|\alpha^*, \beta^*\rangle$? Iba has answered this question [50]. We demonstrate his careful arguments by limiting ourselves to the case of binary image restoration with pixel values ± 1. First, we introduce two statistical quantities defined by

$$\begin{aligned}&\langle (F_{x,y} - h_{x,y}(\boldsymbol{G}, \alpha^*, \beta^*))^2|\alpha^*, \beta^*\rangle \\ &\qquad \equiv \sum_{\boldsymbol{f}}\sum_{\boldsymbol{g}} (f_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 \Pr\{\boldsymbol{F} = \boldsymbol{f}, \boldsymbol{G} = \boldsymbol{g}|\alpha^*, \beta^*\} \end{aligned} \tag{51}$$

$$\begin{aligned}&\langle (F_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2|\boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\rangle \\ &\qquad \equiv \sum_{z} (z_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 \Pr\{\boldsymbol{F} = \boldsymbol{z}|\boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\}. \end{aligned} \tag{52}$$

The following equality for these statistical quantities is valid:

$$\langle (F_{x,y} - h_{x,y}(\boldsymbol{G}, \alpha^*, \beta^*))^2 | \alpha^*, \beta^* \rangle = \sum_{\boldsymbol{f}} \sum_{\boldsymbol{g}} \langle (F_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^* \rangle \Pr\{\boldsymbol{F} = \boldsymbol{f}, \boldsymbol{G} = \boldsymbol{g} | \alpha^*, \beta^*\} \tag{53}$$

The detailed proof is given as follows:

$$\begin{aligned}
&\sum_{\boldsymbol{f}} \sum_{\boldsymbol{g}} \langle (F_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^* \rangle \Pr\{\boldsymbol{F} = \boldsymbol{f}, \boldsymbol{G} = \boldsymbol{g} | \alpha^*, \beta^*\} \\
&\quad = \sum_{\boldsymbol{f}} \sum_{\boldsymbol{g}} \langle (F_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^* \rangle \\
&\qquad \times \Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{f}, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{f} | \alpha^*\} \\
&\quad = \sum_{\boldsymbol{f}} \sum_{\boldsymbol{g}} \sum_{\boldsymbol{z}} (z_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 \Pr\{\boldsymbol{F} = \boldsymbol{z} | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\} \\
&\qquad \times \Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{f}, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{f} | \alpha^*\} \\
&\quad = \sum_{\boldsymbol{f}} \sum_{\boldsymbol{g}} \sum_{\boldsymbol{z}} (z_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 \\
&\qquad \times \left(\frac{\Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{z}, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{z} | \alpha^*\}}{\sum_{\boldsymbol{z}'} \Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{z}', \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{z}' | \alpha^*\}} \right) \\
&\qquad \times \Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{f}, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{f} | \alpha^*\} \\
&\quad = \sum_{\boldsymbol{g}} \sum_{\boldsymbol{z}} (z_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 \\
&\qquad \times \left(\frac{\Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{z}, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{z} | \alpha^*\}}{\sum_{\boldsymbol{z}'} \Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{z}', \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{z}' | \alpha^*\}} \right) \\
&\qquad \times \sum_{\boldsymbol{f}} \Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{f}, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{f} | \alpha^*\} \\
&\quad = \sum_{\boldsymbol{f}} \sum_{\boldsymbol{g}} (f_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 \Pr\{\boldsymbol{G} = \boldsymbol{g} | \boldsymbol{F} = \boldsymbol{f}, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{f} | \alpha^*\} \\
&\quad = \langle (F_{x,y} - h_{x,y}(\boldsymbol{G}, \alpha^*, \beta^*))^2 | \alpha^*, \beta^* \rangle.
\end{aligned} \tag{54}$$

If $|h_{x,y}(\boldsymbol{g}, \alpha, \beta)| = 1$, the first factor on the right-hand side of equation (53) can be rewritten as follows:

$$\begin{aligned}
&\langle (F_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^* \rangle \\
&\quad = \sum_{\boldsymbol{z}} (z_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 \Pr\{\boldsymbol{F} = \boldsymbol{z} | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^*\} \\
&\quad = 2 - 2h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*) \langle F_{x,y} | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^* \rangle.
\end{aligned} \tag{55}$$

If we want to determine the optimal estimator $h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*)$ so as to minimize the quantity $\langle (F_{x,y} - h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*))^2 | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^* \rangle$, the optimal estimator $h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*)$ should be given as follows,

$$h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*) = \begin{cases} +1 & \langle F_{x,y} | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^* \rangle \geqslant 0 \\ -1 & \langle F_{x,y} | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^* \rangle \leqslant 0 \end{cases} \tag{56}$$

so that

$$h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*) = \operatorname{sign}(\langle F_{x,y} | \boldsymbol{G} = \boldsymbol{g}, \alpha^*, \beta^* \rangle). \tag{57}$$

This is a justification for the statement that $h_{x,y}(\boldsymbol{g}, \alpha^*, \beta^*)$ in equation (44) is the unique choice to minimize the statistical quantity $\langle D(\boldsymbol{F}, \boldsymbol{h}(\boldsymbol{G}, \alpha^*, \beta^*))|\alpha^*, \beta^*\rangle$. Iba has proved the statement in the more general case and has pointed out that this property is one of the important aspects of the Nishimori line for probabilistic models [50]. The point is that gauge invariance, which plays an important role in the original argument of Nishimori [48, 49], did not show up here.

Before closing this subsection, we give another interesting aspect of the statistical performance of the MAP estimation, the MPM estimation and the TPM estimation from the statistical-mechanical point of view. Let us introduce a new probability defined by

$$\rho(\boldsymbol{f}|\boldsymbol{G}, \alpha, \beta, T) \equiv \frac{\exp\left(-\frac{1}{T}H(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta)\right)}{\sum_{\boldsymbol{z}} \exp\left(-\frac{1}{T}H(\boldsymbol{z}|\boldsymbol{g}, \alpha, \beta)\right)} \tag{58}$$

where

$$H(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta) \equiv -\ln(\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \beta\}) \tag{59}$$

and $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \beta\}$ is the *a posteriori* probability. We denote the restored image $\hat{\boldsymbol{f}}$ obtained by means of the MAP estimation (8) by the notation $\hat{\boldsymbol{f}}^{\mathrm{MAP}}$. For any positive value of T, the same restored image $\hat{\boldsymbol{f}}^{\mathrm{MAP}}$ is given in terms of the probability $\rho(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta, T)$ by the following prescription:

$$\hat{\boldsymbol{f}}^{\mathrm{MAP}} = \arg\max_{\boldsymbol{f}} \rho(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta, T). \tag{60}$$

For the probability $\rho(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta)$, we define the marginal probability for $f_{x,y}$:

$$\rho_{x,y}(f_{x,y}|\boldsymbol{g}, \alpha^*, \beta^*, T) \equiv \sum_{\boldsymbol{f}_{\Omega\setminus(x,y)}} \rho(\boldsymbol{f}|\boldsymbol{g}, \alpha^*, \beta^*). \tag{61}$$

Between the *a posteriori* probability $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \beta\}$, the new probability $\rho(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta)$ and their marginal probabilities, we have the following four relationships:

$$\rho(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta, 1) = \Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \beta\} \tag{62}$$

$$\rho_{x,y}(f_{x,y}|\boldsymbol{g}, \alpha, \beta, 1) = \Pr\{F_{x,y} = f_{x,y}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \beta\}. \tag{63}$$

If the MAP estimation (8) has a unique solution, the following relationship should also be valid

$$\lim_{T\to+0} \rho(\boldsymbol{f}|\boldsymbol{g}, \alpha^*, \beta^*, T) = \prod_{(x,y)\in\Omega} \delta_{f_{x,y}, \hat{f}^{\mathrm{MAP}}_{x,y}} \tag{64}$$

and then we have

$$\lim_{T\to+0} \rho_{x,y}(f_{x,y}|\boldsymbol{g}, \alpha^*, \beta^*, T) = \delta_{f_{x,y}, \hat{f}^{\mathrm{MAP}}_{x,y}}. \tag{65}$$

Equation (65) means that we obtain the restored image $\hat{\boldsymbol{f}}^{\mathrm{MAP}}$ in the MAP estimation by using the marginal probability $\rho_{x,y}(f_{x,y}|\boldsymbol{g}, \alpha^*, \beta^*, T)$ in the limit $T \to +0$.

Now, we restrict ourselves to binary image restoration and consider the inequality (48) again. Instead of $\mathcal{M}_{x,y}(\alpha, \beta)$, we introduce the following configuration average,

$$\begin{aligned}\tilde{\mathcal{M}}_{x,y}(\alpha, \beta, T) &= \sum_{\boldsymbol{f}}\sum_{\boldsymbol{g}} (f_{x,y} - \tilde{h}_{x,y}(\boldsymbol{g}, \alpha, \beta, T))^2 \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \alpha^*, \beta^*\} \Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha^*\} \\ &= \langle (F_{x,y} - \tilde{h}_{x,y}(\boldsymbol{G}, \alpha, \beta, T))^2 |\alpha^*, \beta^*\rangle \end{aligned} \tag{66}$$

where

$$\tilde{h}_{x,y}(\boldsymbol{g}, \alpha, \beta, T) \equiv \operatorname{sign}\left(\sum_{z} z_{x,y}\rho(\boldsymbol{z}|\boldsymbol{g}, \alpha, \beta, T)\right). \tag{67}$$

Because we have $h_{x,y}(\boldsymbol{g}, \alpha, \beta) = \tilde{h}_{x,y}(\boldsymbol{g}, \alpha, \beta, 1)$, it follows that $\tilde{\mathcal{M}}_{x,y}(\alpha, \beta, 1) = \mathcal{M}_{x,y}(\alpha, \beta)$ is valid. In a similar argument to the proof of equation (48), we can show that

$$\tilde{\mathcal{M}}_{x,y}(\alpha, \beta, T) \geqslant \mathcal{M}_{x,y}(\alpha^*, \beta^*). \tag{68}$$

Inequality (68) means that if we restrict ourselves to binary image restoration and if the MAP estimation gives a unique solution $\hat{\boldsymbol{f}}^{\mathrm{MAP}}$, the MPM estimation can give us a better restored image than the MAP estimation. It must be remembered that the above statement is valid only in the case of binary image restoration. Another important point is that it is generally difficult to prove uniqueness of the minimum configuration that validates equation (64). Without solving this problem, we cannot reach the final answer to the question of the best method among the MAP, MPM or TPM estimates.

2.6. Baysian image restoration and infinite-range Ising model

Some statistical physicists may note that, if the *a priori* probability is assumed to be the infinite-range Ising model, the configuration average $\mathcal{M}_{x,y}(\alpha, \beta) = \langle (F_{x,y} - h_{x,y}(\boldsymbol{G}, \alpha, \beta))^2 | \alpha^*, \beta^* \rangle$ by the joint probability $\Pr\{\boldsymbol{F} = \boldsymbol{f}, \boldsymbol{G} = \boldsymbol{g}|\alpha^*, \beta^*\}$ can be calculated analytically with the help of the replica method. Actually, such pioneering work has been done by Nishimori and Wong [33]. They calculated the statistical difference $\mathcal{M}_{x,y}(\alpha, \beta)$ using the replica method under the assumption that the *a priori* probability is a spin-$\frac{1}{2}$ Ising model with infinite-range interactions defined by

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}\} = \Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\} \equiv \frac{\exp\left(-\frac{1}{2|\Omega|}\alpha \sum_{(x,y)\in\Omega}\sum_{(x',y')\in\Omega}(f_{x,y} - f_{x',y'})^2\right)}{\sum_{z}\exp\left(-\frac{1}{2|\Omega|}\alpha \sum_{(x,y)\in\Omega}\sum_{(x',y')\in\Omega}(z_{x,y} - z_{x',y'})^2\right)} \tag{69}$$

and the degradation process is assumed to be given by the following conditional probability

$$\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\} = \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \sigma\} \equiv \left(\frac{\beta}{2\pi}\right)^{\frac{|\Omega|}{2}} \exp\left(-\frac{\beta}{2}\sum_{(x,y)\in\Omega}(f_{x,y} - g_{x,y})^2\right) \tag{70}$$

where $f_{x,y} = \pm 1$ and $g_{x,y}$ takes any real number. By substituting equations (69) and (70) into equation (40) and using the replica method, we obtain the configuration average $\mathcal{M}(\alpha, \beta)$ in the thermodynamic limit $|\Omega| \to +\infty$:

$$\begin{aligned}\mathcal{M}(\alpha, \beta) &= \lim_{|\Omega|\to+\infty}\frac{1}{|\Omega|}\sum_{(x,y)\in\Omega}\mathcal{M}_{x,y}(\alpha, \beta) \\ &= 2 - \frac{1}{\cosh(\alpha^* m_0)}\sum_{\zeta=\pm 1}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\exp\left(-\frac{u}{2} + \alpha^* m_0 \zeta\right) \\ &\quad \times \zeta \operatorname{sign}\left(\alpha m + \beta\zeta + \frac{\beta}{\sqrt{\beta^*}}u\right)\mathrm{d}u\end{aligned} \tag{71}$$

where m_0 and m are determined from the following simultaneous equations:

$$m_0 = \tanh(\alpha^* m_0) \tag{72}$$

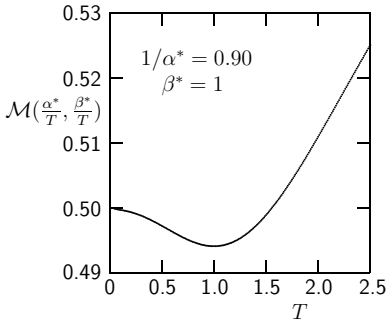

Figure 9. T-dependence of $\mathcal{M}(\frac{\alpha^*}{T}, \frac{\beta^*}{T})$ obtained from equations (71)–(73). The value $\mathcal{M}(\frac{\alpha^*}{T}, \frac{\beta^*}{T})$ at $T = 1$ corresponds to that at the Nishimori point and is the minimum value of $\mathcal{M}(\frac{\alpha^*}{T}, \frac{\beta^*}{T})$ with respect to T. This figure has been drawn using the program written by Professor J Inoue of the Graduate School of Engineering, Hokkaido University.

$$m = \frac{1}{2\cosh(\alpha^* m_0)} \sum_{\zeta=\pm 1} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{u}{2} + \alpha^* m_0 \zeta\right) \tanh\left(\alpha m + \beta\zeta + \frac{\beta}{\sqrt{\beta^*}} u\right) \mathrm{d}u. \quad (73)$$

The detailed derivation is given in [32, 33] which we omit in the present review. For the *a priori* probability (69) and the degradation process (70), the quantity $\tilde{\mathcal{M}}(\alpha, \beta, T)$ defined by equation (66) can be expressed in terms of $\mathcal{M}(\alpha, \beta)$ as follows:

$$\tilde{\mathcal{M}}(\alpha, \beta, T) = \mathcal{M}\left(\frac{\alpha}{T}, \frac{\beta}{T}\right). \quad (74)$$

As mentioned in section 2.5, $\mathcal{M}(\alpha^*, \beta^*) = \tilde{\mathcal{M}}(\alpha^*, \beta^*, 1)$ means the statistical distance between the original image $\boldsymbol{f}$ and the corresponding restored image $\hat{\boldsymbol{f}} = \boldsymbol{h}(\boldsymbol{g}, \alpha^*, \beta^*)$ given in equation (44) at the Nishimori point $(\alpha, \beta) = (\alpha^*, \beta^*)$. Moreover, if the MAP estimation (8) has a unique solution, $\mathcal{M}\left(\frac{\alpha^*}{T}, \frac{\beta^*}{T}\right) = \tilde{\mathcal{M}}(\alpha^*, \beta^*, T)$ converges to the statistical distance between the original image $\boldsymbol{f}$ and the corresponding MAP restored image $\hat{\boldsymbol{f}}^{\mathrm{MAP}}$ defined by equation (8) in the limit $T \to +0$. Now, we show the T-dependence of $\mathcal{M}\left(\frac{\alpha^*}{T}, \frac{\beta^*}{T}\right)$ obtained from equations (71)–(73) in figure 9. It is seen that $\mathcal{M}\left(\frac{\alpha^*}{T}, \frac{\beta^*}{T}\right)$ takes the minimum value $\mathcal{M}(\alpha^*, \beta^*)$ at $T = 1$ and

$$\lim_{T \to +0} \mathcal{M}\left(\frac{\alpha^*}{T}, \frac{\beta^*}{T}\right) > \mathcal{M}(\alpha^*, \beta^*). \quad (75)$$

We remark again that $\lim_{T \to +0} \mathcal{M}\left(\frac{\alpha^*}{T}, \frac{\beta^*}{T}\right)$ corresponds to the statistical distance for the MAP estimation if the MAP estimation (8) has a unique solution. This result suggests that the MPM estimation and the TPM estimation give better restored images than the MAP estimation in binary image restoration.

With the same assumption on the *a priori* probability and the degradation process, Inoue and Tanaka [34] investigated the statistical behaviour of the hyperparameter estimation in evidence framework using the replica method. The analytical calculations of the statistical average $\mathcal{M}(\alpha, \beta)$ give us the performance of the system for Bayesian image restoration without heavy numerical experiments. Some statistical physicists investigated the statistical performance of classical and quantum spin systems constructed for the image restoration by using the replica method [51–55]. For practical applications in binary image restoration, we have to investigate the statistical performance of systems where the *a priori* probability is assumed to be the two-dimensional Ising model with nearest-neighbour interactions.

2.7. Summary

In this section, we have explained the general framework of image restoration by means of Bayesian statistics. Using the Bayes formula under the assumption that the *a priori* probability is a Markov random field, the *a posteriori* probability is expressed in terms of the Gibbs canonical distribution. The latter distribution satisfies the minimization condition of the free energy, which corresponds to the Kullback–Leibler divergence in information theory. In statistics, hyperparameters in the *a posteriori* probability are determined so as to maximize the evidence which is expressed in terms of the free energy of the *a posteriori* probability. The statistical averages of the evidence and the overlap between original image and the corresponding restored image give us the statistical performance of probabilistic image processing systems analytically. The statistical average corresponds to the random average in the spin glass theory. Thus, we see that the probabilistic image processing by Bayesian statistics is closely related to statistical mechanics. In the next section, we clarify a close relationship between probabilistic image processing and statistical mechanics by introducing the framework of the Bayesian approach to binary image restoration.

3. The Bayesian approach to binary image restoration

In this section, we explain the Bayesian approach to binary image restoration in more detail. We clarify a relationship between probabilistic image processing and statistical mechanics through the framework of Bayesian statistics.

By denoting the white state and the black state at each pixel by -1 and $+1$, the Ising model with nearest-neighbour interactions and non-uniform external fields on a finite-size square lattice can be associated with binary image restoration. Then, the external field at each pixel corresponds to the degree of degradation of the given degraded image. The nearest-neighbour interactions represent the *a priori* knowledge for the original image.

In order to obtain the restored image by means of the *a posteriori* probability, we have to look for the most probable configuration in the MAP estimation or calculate the average of the random variable at each pixel in the MPM estimation. The most probable configuration and the average of the random variable at each pixel in the *a posteriori* probability correspond to the ground state configuration and the magnetization of each lattice site, respectively, in the corresponding Ising model. Usually, any image consists of at least a few hundred thousand pixels and each pixel has two possible states, white and black, in the binary image. Therefore, it is hard to obtain the most probable configuration and the magnetization at each lattice site of the corresponding Ising model since it needs unusually long computation. In order to proceed further, computer scientists did not follow the path to obtain the exact solution for some problems with such computational complexity but instead proposed ingenious methods to obtain approximate solutions with high accuracy. Statistical mechanics has been an important source of techniques for calculating the magnetization and the other statistical quantities for massive probabilistic models including the Ising model, although its main purpose is to investigate critical phenomena in the thermodynamic limit. Clearly, these techniques can be applied to the calculations of the most probable configuration and the magnetization at each lattice site for the Ising model constructed for the purpose of binary image restoration.

In this section, we explain some image restoration algorithms constructed using the mean-field and Bethe approximations [11], both of which are familiar to us in statistical mechanics. In the mean-field and Bethe approximations, the probability distribution of the massive probabilistic model is approximately expressed in a factorizable form of one-body or two-body distributions that are referred to as marginal probability distributions in probability

theory and statistics. The one-body and two-body distributions are determined so as to minimize an approximate free energy. The mean-field approximation treats only one-body distributions and the Bethe approximation also treats two-body distributions for the nearest-neighbour pairs of lattice sites. The Bethe approximation takes account of the fluctuations of the nearest-neighbour pairs of lattice sites, whereas the mean-field approximation ignores such fluctuations. In the present section, we give two approximate schemes for determining the hyperparameter by applying the mean-field and Bethe approximations to the evidence framework.

3.1. Bayes formula and evidence framework

All possible states at each pixel (x, y) in the original image $\boldsymbol{f}$ and the degraded image $\boldsymbol{g}$ are restricted to $\{\pm 1\}$. A degradation process $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\}$ and an *a priori* probability distribution $\Pr\{\boldsymbol{F} = \boldsymbol{f}\}$ are assumed to be as follows.

3.1.1. Degradation process in binary images. We have a given degraded image $\boldsymbol{g}$ which is obtained from the original image $\boldsymbol{f}$ by changing the state of each pixel to another state by the same probability $p = \frac{1}{1+\exp(2\beta)}$, independently of the other pixels. Here p is assumed to be less than $\frac{1}{2}$ so that β is always positive. The conditional probability distribution $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\}$ is given by

$$\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\} = \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \beta\} \equiv \frac{\exp\left(-\frac{1}{2}\beta\sum_{(x,y)\in\Omega}(f_{x,y} - g_{x,y})^2\right)}{(1+\exp(-2\beta))^{|\Omega|}}. \tag{76}$$

The summation $\sum_{(x,y)\in\Omega}$ and the product $\prod_{(x,y)\in\Omega}$ are taken over all the pixels (x, y).

3.1.2. A priori probability distribution in binary images. The *a priori* probability distribution, that the original image is $\boldsymbol{f}$, is given by

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}\} = \Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\} \equiv \frac{\exp\left(-\frac{1}{2}\alpha\sum_{(x,y)\in\Omega}((f_{x,y} - f_{x+1,y})^2 + (f_{x,y} - f_{x,y+1})^2)\right)}{\sum_{\boldsymbol{z}}\exp\left(-\frac{1}{2}\alpha\sum_{(x,y)\in\Omega}((z_{x,y} - z_{x+1,y})^2 + (z_{x,y} - z_{x,y+1})^2)\right)}. \tag{77}$$

The summation is defined by

$$\sum_{\boldsymbol{z}} \equiv \prod_{(x,y)\in\Omega}\sum_{z_{x,y}=\pm 1}. \tag{78}$$

By substituting equations (76) and (77) into equation (7), the *a posteriori* probability distribution is given as

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}\} = \Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \beta\} = \frac{\exp(-E(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta))}{\sum_{\boldsymbol{z}}\exp(-E(\boldsymbol{z}|\boldsymbol{g}, \alpha, \beta))} \tag{79}$$

where

$$E(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta) \equiv \frac{1}{2}\beta\sum_{(x,y)\in\Omega}(f_{x,y} - g_{x,y})^2 + \frac{1}{2}\alpha\sum_{(x,y)\in\Omega}((f_{x,y} - f_{x+1,y})^2 + (f_{x,y} - f_{x,y+1})^2). \tag{80}$$

The *a posteriori* probability distribution can be rewritten into a spin-$\frac{1}{2}$ Ising model with nearest-neighbour interactions and non-uniform external fields on a square lattice as follows:

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \beta\} = \frac{\exp(-H(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta))}{\sum_{\boldsymbol{z}}\exp(-H(\boldsymbol{z}|\boldsymbol{g}, \alpha, \beta))} \tag{81}$$

where

$$H(\boldsymbol{f}|\boldsymbol{g},\alpha,\beta) \equiv -\beta \sum_{(x,y)\in\Omega} g_{x,y} f_{x,y} - \alpha \sum_{(x,y)\in\Omega} (f_{x,y} f_{x+1,y} + f_{x,y} f_{x,y+1}). \tag{82}$$

In the framework of binary image restoration given in equations (79), the hyperparameters α and β are determined by

$$(\hat{\alpha},\hat{\beta}) = \arg\max_{(\alpha,\beta)} \Pr\{\boldsymbol{G}=\boldsymbol{g}|\alpha,\beta\}. \tag{83}$$

The logarithm of evidence $\Pr\{\boldsymbol{G}=\boldsymbol{g}|\alpha,\beta\}$ can be rewritten as follows:

$$\begin{aligned}
-\ln(\Pr\{\boldsymbol{G}=\boldsymbol{g}|\alpha,\beta\}) &= -\ln\left(\sum_{z}\exp\left(-E(\boldsymbol{z}|\boldsymbol{g},\alpha,\beta)\right)\right) \\
&\quad + \ln\left(\sum_{z}\exp\left(-E(\boldsymbol{z}|\boldsymbol{g},\alpha,\beta=0)\right)\right) + |\Omega|\ln(1+\exp(-2\beta)) \\
&= -\ln\left(\sum_{z}\exp\left(-H(\boldsymbol{z}|\boldsymbol{g},\alpha,\beta)\right)\right) \\
&\quad + \ln\left(\sum_{z}\exp\left(-H(\boldsymbol{z}|\boldsymbol{g},\alpha,\beta=0)\right)\right) + |\Omega|\ln(2\cosh(\beta)).
\end{aligned} \tag{84}$$

We remark that the first and second terms on the right-hand side correspond to the free energies of the statistical-mechanical models with energy functions $E(\boldsymbol{z}|\boldsymbol{g},\alpha,\beta)$ and $E(\boldsymbol{z}|\boldsymbol{g},\alpha,\beta=0)$, respectively. By differentiating the right-hand side of equation (84) with respect to β and α, we can derive the extremum condition of the logarithm of evidence as follows,

$$\sum_{(x,y)\in\Omega}\sum_{z} g_{x,y} z_{x,y} \Pr\{\boldsymbol{F}=\boldsymbol{z}|\boldsymbol{G}=\boldsymbol{g},\alpha,\beta\} = \tanh(\beta) \tag{85}$$

and

$$\begin{aligned}
&\sum_{(x,y)\in\Omega}\sum_{z}(z_{x,y}z_{x+1,y} + z_{x,y}z_{x,y+1}) \Pr\{\boldsymbol{F}=\boldsymbol{z}|\boldsymbol{G}=\boldsymbol{g},\alpha,\beta\} \\
&\qquad = \sum_{(x,y)\in\Omega}\sum_{z}(z_{x,y}z_{x+1,y} + z_{x,y}z_{x,y+1}) \Pr\{\boldsymbol{F}=\boldsymbol{z}|\alpha\}.
\end{aligned} \tag{86}$$

Now, we introduce the three marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)$, $\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ and $\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$:

$$\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta) \equiv \sum_{z}\rho(\boldsymbol{z}|\boldsymbol{G}=\boldsymbol{g},\alpha,\beta)\delta_{z_{x,y},\zeta} \qquad \zeta=\pm1 \tag{87}$$

$$\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta) \equiv \sum_{z}\rho(\boldsymbol{z}|\boldsymbol{G}=\boldsymbol{g},\alpha,\beta)\delta_{z_{x,y},\zeta}\delta_{z_{x+1,y},\zeta'} \qquad \zeta,\zeta'=\pm1 \tag{88}$$

$$\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta) \equiv \sum_{z}\rho(\boldsymbol{z}|\boldsymbol{G}=\boldsymbol{g},\alpha,\beta)\delta_{z_{x,y},\zeta}\delta_{z_{x,y+1},\zeta'} \qquad \zeta,\zeta'=\pm1. \tag{89}$$

By substituting equations (87)–(89) into equations (86) and (85), we can rewrite the extremum condition of the logarithm of evidence as follows:

$$\sum_{(x,y)\in\Omega}\sum_{\zeta=\pm1} g_{x,y}\zeta\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta) = \tanh(\beta) \tag{90}$$

$$\sum_{(x,y)\in\Omega}\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\zeta\zeta'\left(\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)+\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\right)$$
$$=\sum_{(x,y)\in\Omega}\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\zeta\zeta'\left(\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)+\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)\right). \tag{91}$$

From these equations, we see that the marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)$, $\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ and $\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ have to be calculated for any values of α and β, numerically. However, it is hard to calculate the marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)$, $\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ and $\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$, exactly, in terms of the definitions (87)–(89). As a key to calculating the marginal probability distributions, we explain the mean-field and Bethe approximations for the energy function (80) in sections 3.2 and 3.3.

3.2. Mean-field approximation

In this subsection, we derive the deterministic equations of the MPM estimate and the hyperparameter estimation in the mean-field approximation for the Markov random field model based on the variational principle of the free energy with respect to marginal probability distributions [17].

In the mean-field approximation, the probability distribution $\rho(\boldsymbol{f}|\boldsymbol{g},\alpha,\beta)$ is approximately expressed in terms of the marginal probability distributions

$$\rho(\boldsymbol{f}|\boldsymbol{g},\alpha,\beta)\simeq\prod_{(x,y)\in\Omega}\rho_{x,y}(f_{x,y}|\boldsymbol{g},\alpha,\beta). \tag{92}$$

By using equation (92), the approximate form of the free energy $\mathcal{F}[\rho]\equiv\sum_{\boldsymbol{z}}\rho(\boldsymbol{z}|\boldsymbol{g})(H(\boldsymbol{z}|\boldsymbol{g})+\ln\rho(\boldsymbol{z}|\boldsymbol{g}))$ for $T=1$ can be written as

$$\mathcal{F}[\rho]\simeq\mathcal{F}_{\mathrm{MF}}[\{\rho_{x,y}\}]\equiv\sum_{(x,y)\in\Omega}\sum_{\zeta=\pm1}\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)\Bigg(-\beta g_{x,y}\zeta-\alpha\sum_{\zeta'=\pm1}\zeta\zeta'\rho_{x+1,y}(\zeta'|\boldsymbol{g},\alpha,\beta)$$
$$-\alpha\sum_{\zeta'=\pm1}\zeta\zeta'\rho_{x,y+1}(\zeta'|\boldsymbol{g},\alpha,\beta)+\ln\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)\Bigg). \tag{93}$$

We note that the marginal probability distribution should satisfy the following normalization condition:

$$\sum_{\zeta=\pm1}\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)=1. \tag{94}$$

To ensure the normalization condition, we introduce the Lagrange multiplier $\nu_{x,y}$:

$$\mathcal{L}_{\mathrm{MF}}[\{\rho_{x,y}\}]=\mathcal{F}_{\mathrm{MF}}[\{\rho_{x,y}\}]+\sum_{(x,y)\in\Omega}\nu_{x,y}\left(\sum_{\zeta=\pm1}\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)-1\right). \tag{95}$$

By taking the first variation of $\mathcal{L}_{\mathrm{MF}}[\{\rho_{x,y}\}]$ with respect to $\rho_{x,y}(\xi)$,

$$\frac{\partial}{\partial\rho_{x,y}(\xi)}\mathcal{L}_{\mathrm{MF}}[\{\rho_{x,y}\}]=0\qquad\xi=\pm1 \tag{96}$$

we derive the following extremum condition:

$$\rho_{x,y}(\xi|\boldsymbol{g},\alpha,\beta)=\exp\left(-1-\nu_{x,y}+\beta g_{x,y}\xi+\alpha\xi\sum_{\zeta=\pm1}\zeta\sum_{(x',y')\in\mathbf{c}_{x,y}}\rho_{x',y'}(\zeta|\boldsymbol{g},\alpha,\beta)\right). \tag{97}$$

Here $\mathbf{c}_{x,y}$ is the set of neighbouring pixels of the pixel (x, y):

$$\mathbf{c}_{x,y} \equiv \{(x+1, y), (x-1, y), (x, y+1), (x, y-1)\}. \tag{98}$$

We can summarize the simultaneous deterministic equations with respect to the marginal probability distributions $\rho_{x,y}(\xi|\boldsymbol{g}, \alpha, \beta)$ as follows,

$$\rho_{x,y}(\xi|\boldsymbol{g}, \alpha, \beta) = \frac{\exp(-H(\xi|\boldsymbol{g}, \alpha, \beta))}{\sum_{\zeta=\pm1}\exp(-H(\zeta|\boldsymbol{g}, \alpha, \beta))} \qquad \zeta = \pm 1 \tag{99}$$

where

$$H_{x,y}(\xi|\boldsymbol{g}, \alpha, \beta) \equiv -\beta g_{x,y}\xi - \alpha\xi \sum_{\zeta=\pm1}\zeta \sum_{(x',y')\in\mathbf{c}_{x,y}} \rho_{x',y'}(\zeta|\boldsymbol{g}, \alpha, \beta). \tag{100}$$

The marginal probability distribution $\rho_{x,y}(\xi|\boldsymbol{g}, \alpha, \beta)$ can be expressed in the orthogonal series as

$$\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta) = \tfrac{1}{2}(1 + m_{x,y}(\boldsymbol{g}, \alpha, \beta)\zeta) \tag{101}$$

where

$$m_{x,y}(\boldsymbol{g}, \alpha, \beta) \equiv \sum_{\zeta=\pm1}\zeta\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta). \tag{102}$$

By substituting equation (101) into equations (99) and (100), the simultaneous deterministic equations with respect to the marginal probability distributions $\rho_{x,y}(\xi|\boldsymbol{g}, \alpha, \beta)$ can be rewritten as the equations for the moments $m_{x,y}(\boldsymbol{g}, \alpha, \beta)$:

$$m_{x,y}(\boldsymbol{g}, \alpha, \beta) = \tanh\left(\beta g_{x,y} + \alpha \sum_{(x',y')\in\mathbf{c}_{x,y}} m_{x',y'}(\boldsymbol{g}, \alpha, \beta)\right). \tag{103}$$

The approximate form of the evidence in the mean-field approximation is given as follows:

$$\begin{aligned}
\ln(\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}) &= -F(\boldsymbol{g}, \alpha, \beta) + F(\boldsymbol{g}, \alpha, \beta = 0) + |\Omega|\ln(\exp(\beta) + \exp(-\beta)) \\
&\simeq -\sum_{(x,y)\in\Omega}\left(-\beta\sum_{\zeta=\pm1} g_{x,y}\zeta\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta) - \alpha\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\zeta\zeta'\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta)\right. \\
&\quad \times(\rho_{x+1,y}(\zeta'|\boldsymbol{g}, \alpha, \beta) + \rho_{x,y+1}(\zeta'|\boldsymbol{g}, \alpha, \beta)) \\
&\quad \left. + \sum_{\zeta=\pm1}\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta)\ln\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta)\right) \\
&\quad + \sum_{(x,y)\in\Omega}\left(-\alpha\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\zeta\zeta'\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta, = 0)(\rho_{x+1,y}(\zeta'|\boldsymbol{g}, \alpha, \beta = 0)\right. \\
&\quad \left. + \rho_{x,y+1}(\zeta'|\boldsymbol{g}, \alpha, \beta = 0)) + \sum_{\zeta=\pm1}\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta = 0)\ln\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta = 0)\right) \\
&\quad + |\Omega|\ln(2\cosh(\beta)).
\end{aligned} \tag{104}$$

Since the marginal probability distributions $\rho_{x,y}^{x+1,y}(\zeta, \zeta'|\boldsymbol{g}, \alpha, \beta)$ and $\rho_{x,y}^{x,y+1}(\zeta, \zeta'|\boldsymbol{g}, \alpha, \beta)$ are approximately expressed in terms of the marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta)$, $\rho_{x+1,y}(\zeta|\boldsymbol{g}, \alpha, \beta)$ and $\rho_{x,y+1}(\zeta|\boldsymbol{g}, \alpha, \beta)$ as follows,

$$\rho_{x,y}^{x+1,y}(\zeta, \zeta'|\boldsymbol{g}, \alpha, \beta) \simeq \rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta)\rho_{x+1,y}(\zeta'|\boldsymbol{g}, \alpha, \beta) \tag{105}$$

and

$$\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta) \simeq \rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)\rho_{x,y+1}(\zeta'|\boldsymbol{g},\alpha,\beta) \tag{106}$$

the extremum conditions (90) and (91) can be approximately written as

$$\sum_{(x,y)\in\Omega} g_{x,y} m_{x,y}(\boldsymbol{g},\alpha,\beta) = \tanh(\beta) \tag{107}$$

$$\sum_{(x,y)\in\Omega} m_{x,y}(\boldsymbol{g},\alpha,\beta)(m_{x+1,y}(\boldsymbol{g},\alpha,\beta) + m_{x,y+1}(\boldsymbol{g},\alpha,\beta))$$
$$= \sum_{(x,y)\in\Omega} m_{x,y}(\boldsymbol{g},\alpha,\beta=0)(m_{x+1,y}(\boldsymbol{g},\alpha,\beta=0) + m_{x,y+1}(\boldsymbol{g},\alpha,\beta=0)). \tag{108}$$

In the case of $\beta = 0$, since both the energy function $H(\boldsymbol{f}|\boldsymbol{g},\alpha,\beta=0)$ and the moment $m_{x,y}(\boldsymbol{g},\alpha,\beta=0)$ are independent of the degraded image $\boldsymbol{g}$, the moment $m_{x,y}(\boldsymbol{g},\alpha,\beta)$ is not dependent on the site (x, y) and can be replaced by a notation $m(\alpha)$:

$$m_{x,y}(\boldsymbol{g},\alpha,\beta=0) = m(\alpha) \qquad (x,y)\in\Omega. \tag{109}$$

The mean-field equation (103) can then be reduced to

$$m(\alpha) = \tanh(4\alpha m(\alpha)). \tag{110}$$

By using this equation, the deterministic equations (107) and (108) are rewritten as

$$\beta = \operatorname{arctanh}\left(\sum_{(x,y)\in\Omega} g_{x,y} m_{x,y}(\boldsymbol{g},\alpha,\beta)\right) \tag{111}$$

$$\alpha = \frac{1}{4m(\alpha)}\operatorname{arctanh}\left(\sqrt{\frac{1}{2}\sum_{(x,y)\in\Omega} m_{x,y}(\boldsymbol{g},\alpha,\beta)(m_{x+1,y}(\boldsymbol{g},\alpha,\beta) + m_{x,y+1}(\boldsymbol{g},\alpha,\beta))}\right). \tag{112}$$

3.3. Bethe approximation

In this subsection, we derive the deterministic equations of the Bethe approximation for the Markov random field model based on the variational principle of the free energy with respect to marginal probability distributions [30, 31, 57].

In the Bethe approximation, the entropy $S \equiv -\sum_{\boldsymbol{z}} \rho(\boldsymbol{z}|\boldsymbol{g})\ln\rho(\boldsymbol{z}|\boldsymbol{g})$ is approximately expressed in terms of the marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)$, $\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ and $\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ as follows,

$$\mathcal{S}[\rho] \simeq \sum_{(x,y)\in\Omega} \mathcal{S}[\rho_{x,y}] + \sum_{(x,y)\in\Omega} \left(\mathcal{S}\left[\rho_{x,y}^{x+1,y}\right] - \mathcal{S}[\rho_{x,y}] - \mathcal{S}[\rho_{x+1,y}]\right)$$
$$+ \sum_{(x,y)\in\Omega} \left(\mathcal{S}\left[\rho_{x,y}^{x,y+1}\right] - \mathcal{S}[\rho_{x,y}] - \mathcal{S}[\rho_{x,y+1}]\right)$$
$$= -3\sum_{(x,y)\in\Omega} \mathcal{S}[\rho_{x,y}] + \sum_{(x,y)\in\Omega} \mathcal{S}\left[\rho_{x,y}^{x+1,y}\right] + \sum_{(x,y)\in\Omega} \mathcal{S}\left[\rho_{x,y}^{x,y+1}\right] \tag{113}$$

where

$$\mathcal{S}\left[\rho_{x,y}^{x+1,y}\right] \equiv -\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1} \rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\ln\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta) \tag{114}$$

and

$$\mathcal{S}\left[\rho_{x,y}^{x,y+1}\right] \equiv -\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\ln\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta). \tag{115}$$

Now we explain the above approximate form of the entropy in the Bethe approximation. All the random variables of a probability distribution $P(\boldsymbol{f})$ are independent of each other and the entropy $S \equiv -\sum_{\boldsymbol{z}} P(\boldsymbol{z})\ln(P(\boldsymbol{z}))$ is expressed as

$$S = \sum_{(x,y)\in\Omega} S_{x,y} \tag{116}$$

where $S_{x,y}$ is defined for a pixel (x, y) by

$$S_{x,y} \equiv -\sum_{\zeta=\pm1} P_{x,y}(\zeta)\ln P_{x,y}(\zeta) \tag{117}$$

$$P_{x,y}(\zeta) \equiv \sum_{\boldsymbol{z}} P(\boldsymbol{z})\delta_{z_{x,y},\zeta} \qquad \zeta = \pm1. \tag{118}$$

If we have an interaction between the nearest-neighbour pairs of pixels (x_1, y_1) and (x_1+1, y_1) and the other pixels are independent of each other in a probability distribution $P(\boldsymbol{f})$, the entropy is written as,

$$S = \sum_{(x,y)\in\Omega\setminus\{(x_1,y_1),(x_1+1,y_1)\}} S_{x,y} + S_{x_1,y_1}^{x_1+1,y_1} = \sum_{(x,y)\in\Omega} S_{x,y} + \left(S_{x_1,y_1}^{x_1+1,y_1} - S_{x_1,y_1} - S_{x_1+1,y_1}\right) \tag{119}$$

where $S_{x,y}^{x',y'}$ is defined for pairs of pixels (x, y) and (x', y') by

$$S_{x,y}^{x',y'} \equiv -\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1} P_{x,y}^{x',y'}(\zeta,\zeta')\ln P_{x,y}^{x',y'}(\zeta,\zeta') \tag{120}$$

$$P_{x,y}^{x',y'}(\zeta,\zeta') \equiv \sum_{\boldsymbol{z}} P(\boldsymbol{z})\delta_{z_{x,y},\zeta}\delta_{z_{x',y'},\zeta'} \qquad \zeta,\zeta' = \pm1. \tag{121}$$

If we have an interaction between the nearest-neighbour pairs of pixels (x_1, y_1), (x_1+1, y_1), (x_1+1, y_1+1) and (x_1, y_1+1) and the other pixels are independent of each other in a probability distribution $P(\boldsymbol{f})$, the entropy is written as

$$\begin{aligned} S &= \sum_{(x,y)\in\Omega\setminus\{(x_1,y_1),(x_1+1,y_1),(x_1+1,y_1+1),(x_1,y_1+1)\}} S_{x,y} + S_{x_1,y_1}^{(4)} \\ &= \sum_{(x,y)\in\Omega} S_{x,y} + \left(S_{x_1,y_1}^{x_1+1,y_1} - S_{x_1,y_1} - S_{x_1+1,y_1}\right) + \left(S_{x_1+1,y_1}^{x_1+1,y_1+1} - S_{x_1+1,y_1} - S_{x_1+1,y_1+1}\right) \\ &\quad + \left(S_{x_1,y_1+1}^{x_1+1,y_1+1} - S_{x_1,y_1+1} - S_{x_1+1,y_1+1}\right) + \left(S_{x_1,y_1}^{x_1,y_1+1} - S_{x_1,y_1} - S_{x_1,y_1+1}\right) \\ &\quad + \Big(S_{x_1,y_1}^{(4)} - S_{x_1,y_1}^{x_1,y_1+1} - S_{x_1,y_1+1}^{x_1+1,y_1+1} - S_{x_1+1,y_1}^{x_1+1,y_1+1} - S_{x_1,y_1}^{x_1+1,y_1} \\ &\quad + 2S_{x_1,y_1} + 2S_{x_1,y_1+1} + 2S_{x_1+1,y_1+1} + 2S_{x_1+1,y_1}\Big) \end{aligned} \tag{122}$$

where $S_{x_1,y_1}^{(4)}$ is defined for a square plaquette $\{(x_1, y_1), (x_1+1, y_1), (x_1+1, y_1+1), (x_1, y_1+1)\}$ by

$$S_{x,y}^{(4)} \equiv -\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\sum_{\zeta''=\pm1}\sum_{\zeta'''=\pm1} P_{x,y}^{(4)}(\zeta,\zeta',\zeta'',\zeta''')\ln P_{x,y}^{(4)}(\zeta,\zeta',\zeta'',\zeta''') \tag{123}$$

$$P_{x,y}^{(4)}(\zeta,\zeta',\zeta'',\zeta''') \equiv \sum_{\boldsymbol{z}} P(\boldsymbol{z})\delta_{z_{x,y},\zeta}\delta_{z_{x+1,y},\zeta'}\delta_{z_{x+1,y+1},\zeta''}\delta_{z_{x,y+1},\zeta'''} \qquad \zeta,\zeta',\zeta'',\zeta''' = \pm1. \tag{124}$$

If we have interactions between all the nearest-neighbour pairs of pixels, the entropy can be represented as follows,

$$S = \sum_{(x,y)\in\Omega} S_{x,y} + \sum_{(x,y)\in\Omega} \left(S_{x,y}^{x+1,y} - S_{x,y} - S_{x+1,y}\right) + \sum_{(x,y)\in\Omega} \left(S_{x,y}^{x,y+1} - S_{x,y} - S_{x,y+1}\right) + \bar{S} \tag{125}$$

where $\bar{S}$ is a correction term including $\sum_{(x,y)\in\Omega} S_{x,y}^{(4)}$ and the other entropies for many-body marginal probability distributions. In the Bethe approximation, we set $\bar{S} = 0$ in equation (125). This is the physical meaning of equation (113). Setting $\bar{S} = 0$ in equation (125) corresponds to the assumption that the probability distribution $P(\boldsymbol{f})$ is approximately expressed in the following factorizable form:

$$P(\boldsymbol{f}) \simeq \left(\prod_{(x,y)\in\Omega} P_{x,y}(f_{x,y})\right)\left(\prod_{(x,y)\in\Omega} \frac{P_{x,y}^{x+1,y}(f_{x,y}, f_{x+1,y})}{P_{x,y}(f_{x,y})P_{x+1,y}(f_{x+1,y})}\right) \times \left(\prod_{(x,y)\in\Omega} \frac{P_{x,y}^{x,y+1}(f_{x,y}, f_{x,y+1})}{P_{x,y}(f_{x,y})P_{x,y+1}(f_{x,y+1})}\right). \tag{126}$$

By substituting equation (113) into the expression of the free energy $\mathcal{F}[\rho] \equiv \sum_z \rho(\boldsymbol{z}|\boldsymbol{g})(H(\boldsymbol{z}|\boldsymbol{g}) + \ln \rho(\boldsymbol{z}|\boldsymbol{g}))$ for $T = 1$, we obtain the approximate form of the free energy $\mathcal{F}_{\text{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1}\right\}\right]$ as

$$\begin{aligned}
\mathcal{F}[\rho] &\simeq \mathcal{F}_{\text{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1}\right\}\right] \\
&\equiv -\sum_{(x,y)\in\Omega}\sum_{\zeta=\pm1} \left(\beta g_{x,y}\zeta\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)\right. \\
&\qquad \left. + \alpha\zeta\zeta'\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta) + \alpha\zeta\zeta'\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\right) \\
&\qquad - 3\sum_{(x,y)\in\Omega}\sum_{\zeta=\pm1} \rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)\ln\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta) \\
&\qquad + \sum_{(x,y)\in\Omega}\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1} \rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\ln\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta) \\
&\qquad + \sum_{(x,y)\in\Omega}\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1} \rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\ln\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta).
\end{aligned} \tag{127}$$

The marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)$, $\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ and $\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ satisfy the normalization conditions,

$$\begin{aligned}
\sum_{\zeta=\pm1}\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta) &= \sum_{\zeta=\pm1}\sum_{\xi=\pm1}\rho_{x,y}^{x+1,y}(\xi,\zeta|\boldsymbol{g},\alpha,\beta) \\
&= \sum_{\zeta=\pm1}\sum_{\xi=\pm1}\rho_{x-1,y}^{x,y}(\zeta,\xi|\boldsymbol{g},\alpha,\beta) = 1
\end{aligned} \tag{128}$$

and the reducibility conditions,

$$\begin{aligned}
\rho_{x,y}(\xi|\boldsymbol{g},\alpha,\beta) &= \sum_{\zeta=\pm1}\rho_{x,y}^{x+1,y}(\xi,\zeta|\boldsymbol{g},\alpha,\beta) = \sum_{\zeta=\pm1}\rho_{x-1,y}^{x,y}(\zeta,\xi|\boldsymbol{g},\alpha,\beta) \\
&= \sum_{\zeta=\pm1}\rho_{x,y}^{x,y+1}(\xi,\zeta|\boldsymbol{g},\alpha,\beta) = \sum_{\zeta=\pm1}\rho_{x,y-1}^{x,y}(\zeta,\xi|\boldsymbol{g},\alpha,\beta).
\end{aligned} \tag{129}$$

To ensure the normalization condition, we introduce the Lagrange multipliers $\left\{\nu_{x,y}, \nu_{x,y}^{x+1,y}, \nu_{x,y}^{x,y+1}\middle|(x,y)\in\Omega\right\}$ and $\left\{\Psi_{x,y}^{x+1,y}(\xi), \Psi_{x,y}^{x,y+1}(\xi)\middle|(x,y)\in\Omega\right\}$:

$$\mathcal{L}_{\text{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1}\right\}\right] = \mathcal{F}_{\text{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1}\right\}\right]$$
$$+ \sum_{(x,y)\in\Omega} \nu_{x,y}\left(\sum_{\zeta=\pm1}\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta) - 1\right)$$
$$+ \sum_{(x,y)\in\Omega} \nu_{x,y}^{x+1,y}\left(\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta) - 1\right)$$
$$+ \sum_{(x,y)\in\Omega} \nu_{x,y}^{x,y+1}\left(\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta) - 1\right)$$
$$+ \sum_{(x,y)\in\Omega}\sum_{(x',y')\in\mathbf{c}_{x,y}}\sum_{\zeta=\pm1}\Psi_{x,y}^{x',y'}(\zeta)\left(\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta) - \sum_{\zeta'=\pm1}\rho_{x,y}^{x',y'}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\right). \tag{130}$$

By taking the first variations of $\mathcal{L}_{\text{Bethe}}\left[\left\{\rho_{x,y}, \rho_{x,y}^{x+1,y}, \rho_{x,y}^{x,y+1}\right\}\right]$ with respect to $\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)$, $\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ and $\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$, we derive the expressions of the marginal probability distributions in terms of the Lagrange multipliers as follows,

$$\rho_{x,y}(\xi|\boldsymbol{g},\alpha,\beta) = \frac{\exp(\Psi_{x,y}(\xi))}{\sum_{\zeta=\pm1}\exp(\Psi_{x,y}(\zeta))} \tag{131}$$

$$\rho_{x,y}^{x+1,y}(\xi,\xi'|\boldsymbol{g},\alpha,\beta) = \frac{\exp\left(\Psi_{x,y}^{x+1,y}(\xi) + \Psi_{x+1,y}^{x,y}(\xi') + \alpha\xi\xi'\right)}{\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\exp\left(\Psi_{x,y}^{x+1,y}(\zeta) + \Psi_{x+1,y}^{x,y}(\zeta') + \alpha\zeta\zeta'\right)} \tag{132}$$

$$\rho_{x,y}^{x,y+1}(\xi,\xi'|\boldsymbol{g},\alpha,\beta) = \frac{\exp\left(\Psi_{x,y}^{x,y+1}(\xi) + \Psi_{x,y+1}^{x,y}(\xi') + \alpha\xi\xi'\right)}{\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\exp\left(\Psi_{x,y}^{x,y+1}(\zeta) + \Psi_{x,y+1}^{x,y}(\zeta') + \alpha\zeta\zeta'\right)} \tag{133}$$

where $\Psi_{x,y}(\xi)$ is defined by the following equation:

$$-3\Psi_{x,y}(\xi) + \sum_{(x',y')\in\mathbf{c}_{x,y}}\Psi_{x,y}^{x',y'}(\xi) = \beta g_{x,y}\xi \qquad \xi = \pm1. \tag{134}$$

By substituting equations (131)–(134) into the reducibility (129), we have

$$\exp\left(-\frac{1}{3}\beta g_{x,y}\xi + \frac{1}{3}\sum_{(x'',y'')\in\mathbf{c}_{x,y}}\Psi_{x,y}^{x'',y''}(\xi)\right) = \left(\sum_{\zeta=\pm1}\exp\left(-\frac{1}{3}g_{x,y}\zeta + \frac{1}{3}\sum_{(x'',y'')\in\mathbf{c}_{x,y}}\Psi_{x,y}^{x'',y''}(\zeta)\right)\right)$$
$$\times\left(\sum_{\zeta=\pm1}\exp\left(\alpha\xi\zeta + \Psi_{x,y}^{x',y'}(\xi) + \Psi_{x',y'}^{x,y}(\zeta)\right)\right)$$
$$\times\left(\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\exp\left(\alpha\zeta\zeta' + \Psi_{x,y}^{x',y'}(\zeta) + \Psi_{x',y'}^{x,y}(\zeta')\right)\right)^{-1}$$
$$(x,y)\in\Omega \quad (x',y')\in\mathbf{c}_{x,y} \quad \xi = \pm1. \tag{135}$$

Now, we introduce the following replacement,

$$\Psi_{x,y}^{x',y'}(\xi) = \sum_{(x'',y'')\in\mathbf{c}_{x,y}\setminus(x',y')}\Phi_{x,y}^{x'',y''}(\xi) + \beta g_{x,y}\xi \tag{136}$$

and rewrite the deterministic equations (135) as the following expressions:

$$\exp\left(\Phi_{x,y}^{x',y'}(\xi)\right) = \left(\sum_{\zeta=\pm1}\exp\left(\beta g_{x,y}\zeta + \sum_{(x'',y'')\in\mathbf{c}_{x,y}}\Phi_{x,y}^{x'',y''}(\zeta)\right)\right)$$
$$\times\left(\sum_{\zeta=\pm1}\exp\left(\beta g_{x',y'}\zeta + \alpha\xi\zeta + \sum_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\Phi_{x',y'}^{x'',y''}(\zeta)\right)\right)$$
$$\times\left(\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\exp\left(\beta g_{x,y}\zeta + \beta g_{x',y'}\zeta' + \alpha\zeta\zeta' + \sum_{(x'',y'')\in\mathbf{c}_{x,y}\setminus(x',y')}\Phi_{x,y}^{x'',y''}(\zeta)\right.\right.$$
$$\left.\left. + \sum_{(x'',y'')\in\mathbf{c}_{x',y'}\setminus(x,y)}\Phi_{x',y'}^{x'',y''}(\zeta')\right)\right)^{-1} \qquad (x',y')\in\mathbf{c}_{x,y} \quad \xi=\pm1. \tag{137}$$

We introduce the quantity

$$\lambda_{x,y}^{x',y'}(\boldsymbol{g},\alpha,\beta) \equiv \tfrac{1}{2}\left(\Phi_{x,y}^{x',y'}(+1) - \Phi_{x,y}^{x',y'}(-1)\right) \tag{138}$$

as an effective field from (x', y') to (x, y). Because the quantity $\Phi_{x,y}^{x',y'}(\xi)$ can be expressed by means of the orthonormal expansion in terms of an orthonormal set of polynomials[2] $\{1, \xi\}$

$$\Phi_{x,y}^{x',y'}(\xi) = \tfrac{1}{2}\left(\Phi_{x,y}^{x',y'}(+1) + \Phi_{x,y}^{x',y'}(-1)\right) + \tfrac{1}{2}\left(\Phi_{x,y}^{x',y'}(+1) - \Phi_{x,y}^{x',y'}(-1)\right)\xi \tag{139}$$

the simultaneous deterministic equations for effective fields are obtained as follows:

$$\lambda_{x\pm1,y}^{x,y}(\boldsymbol{g},\alpha,\beta) = \operatorname{arctanh}\left(\tanh(\alpha)\tanh\left(\beta g_{x,y} + \sum_{(x',y')\in\mathbf{c}_{x,y}\setminus(x\pm1,y)}\lambda_{x,y}^{x',y'}(\boldsymbol{g},\alpha,\beta)\right)\right) \tag{140}$$

$$\lambda_{x,y\pm1}^{x,y}(\boldsymbol{g},\alpha,\beta) = \operatorname{arctanh}\left(\tanh(\alpha)\tanh\left(\beta g_{x,y} + \sum_{(x',y')\in\mathbf{c}_{x,y}\setminus(x,y\pm1)}\lambda_{x,y}^{x',y'}(\boldsymbol{g},\alpha,\beta)\right)\right). \tag{141}$$

The expressions for the marginal probability distributions are given as follows:

$$\rho_{x,y}(\xi) = \frac{\exp(-H_{x,y}(\xi))}{\sum_{\zeta=\pm1}\exp(-H_{x,y}(\zeta))} \tag{142}$$

$$\rho_{x,y}^{x+1,y}(\xi,\xi') = \frac{\exp\left(-H_{x,y}^{x+1,y}(\xi,\xi')\right)}{\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\exp\left(-H_{x,y}^{x+1,y}(\zeta,\zeta')\right)} \tag{143}$$

$$\rho_{x,y}^{x,y+1}(\xi,\xi') = \frac{\exp\left(-H_{x,y}^{x,y+1}(\xi,\xi')\right)}{\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\exp\left(-H_{x,y}^{x,y+1}(\zeta,\zeta')\right)} \tag{144}$$

where

$$H_{x,y}(\xi) \equiv -\left(\beta g_{x,y} + \sum_{(x',y')\in\mathbf{c}_{x,y}}\lambda_{x,y}^{x',y'}(\boldsymbol{g},\alpha,\beta)\right)\xi \tag{145}$$

[2] Since it is valid that $\sum_{\xi=\pm1}(1\times1) = \sum_{\xi=\pm1}(\xi\times\xi) = 2$ and $\sum_{\xi=\pm1}(1\times\xi) = \sum_{\xi=\pm1}(\xi\times1) = 0$, the set of polynomials $\{1, \xi\}$ can be regarded as an orthonormal set for $\xi = \pm1$.

$$H_{x,y}^{x+1,y}(\xi,\xi') \equiv -\alpha\xi\xi' - \left(\beta g_{x,y} + \sum_{(x',y')\in \mathbf{c}_{x,y}\setminus(x+1,y)} \lambda_{x,y}^{x',y'}(\boldsymbol{g},\alpha,\beta)\right)\xi - \left(\beta g_{x+1,y} + \sum_{(x',y')\in \mathbf{c}_{x+1,y}\setminus(x,y)} \lambda_{x+1,y}^{x',y'}(\boldsymbol{g},\alpha,\beta)\right)\xi' \tag{146}$$

$$H_{x,y}^{x,y+1}(\xi,\xi') \equiv -\alpha\xi\xi' - \left(\beta g_{x,y} + \sum_{(x',y')\in \mathbf{c}_{x,y}\setminus(x,y+1)} \lambda_{x,y}^{x',y'}(\boldsymbol{g},\alpha,\beta)\right)\xi - \left(\beta g_{x,y+1} + \sum_{(x',y')\in \mathbf{c}_{x,y+1}\setminus(x,y)} \lambda_{x,y+1}^{x',y'}(\boldsymbol{g},\alpha,\beta)\right)\xi'. \tag{147}$$

According to equations (84) and (127), the approximate expression for the logarithm of evidence in the Bethe approximation is given in terms of the posterior marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)$, $\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$ and $\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)$, and the prior marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta=0)$, $\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)$ and $\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)$:

$$\begin{aligned}
\ln(\Pr\{\boldsymbol{G}=\boldsymbol{g}|\alpha,\beta\}) \simeq -\sum_{(x,y)\in\Omega}\Bigg(&-\beta\sum_{\zeta=\pm1} g_{x,y}\zeta\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)\\
&-\alpha\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\zeta\zeta'\left(\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)+\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\right)\\
&-3\sum_{\zeta=\pm1}\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)\ln\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta)\\
&+\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\ln\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\\
&+\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\ln\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta)\Bigg)\\
&+\sum_{(x,y)\in\Omega}\Bigg(-\alpha\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\zeta\zeta'\left(\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)\right.\\
&\left.+\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)\right)-3\sum_{\zeta=\pm1}\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta=0)\ln\rho_{x,y}(\zeta|\boldsymbol{g},\alpha,\beta=0)\\
&+\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)\ln\rho_{x,y}^{x+1,y}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)\\
&+\sum_{\zeta=\pm1}\sum_{\zeta'=\pm1}\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)\ln\rho_{x,y}^{x,y+1}(\zeta,\zeta'|\boldsymbol{g},\alpha,\beta=0)\Bigg)\\
&+|\Omega|\ln(2\cosh(\beta)).
\end{aligned} \tag{148}$$

In the case of $\beta = 0$, since both the energy function $H(\boldsymbol{f}|\boldsymbol{g},\alpha,\beta=0)$ and the effective fields $\lambda_{x,y}^{x',y'}(\boldsymbol{g},\alpha,\beta=0)$ are independent of the degraded image $\boldsymbol{g}$, the effective fields

$\lambda_{x,y}^{x',y'}(\boldsymbol{g}, \alpha, \beta = 0)$ are not dependent on the site (x, y) and can be replaced by notation $\lambda(\alpha)$:

$$\lambda_{x,y}^{x',y'}(\boldsymbol{g}, \alpha, \beta = 0) = \lambda(\alpha). \tag{149}$$

The simultaneous deterministic equations (140) and (141) can be reduced to

$$\lambda(\alpha) = \operatorname{arctanh}(\tanh(\alpha)\tanh(4\lambda(\alpha))). \tag{150}$$

By using this equation, the deterministic equations (90) and (91) are rewritten as

$$\beta = \operatorname{arctanh}\left(\sum_{(x,y)\in\Omega} g_{x,y} \tanh\left(\beta g_{x,y} + \sum_{(x',y')\in \mathbf{c}_{x,y}} \lambda_{x,y}^{x-1,y}(\boldsymbol{g}, \alpha, \beta)\right)\right) \tag{151}$$

$$\begin{aligned}
&\frac{1}{|\Omega|}\sum_{(x,y)\in\Omega} \operatorname{arctanh}\left\{\exp\left(-2\alpha - 2\operatorname{arctanh}\left\{\tanh\left(\beta g_{x,y} + \sum_{(x',y')\in \mathbf{c}_{x,y}\setminus(x+1,y)} \lambda_{x,y}^{x',y'}(\boldsymbol{g}, \alpha, \beta)\right)\right.\right.\right. \\
&\qquad \left.\left.\left. \times \tanh\left(\beta g_{x+1,y} + \sum_{(x',y')\in \mathbf{c}_{x,y}\setminus(x,y)} \lambda_{x+1,y}^{x',y'}(\boldsymbol{g}, \alpha, \beta)\right)\right\}\right)\right\} \\
&\qquad + \frac{1}{|\Omega|}\sum_{(x,y)\in\Omega} \operatorname{arctanh}\left\{\exp\left(-2\alpha\right.\right. \\
&\qquad \left. - 2\operatorname{arctanh}\left\{\tanh\left(\beta g_{x,y} + \sum_{(x',y')\in \mathbf{c}_{x,y}\setminus(x,y+1)} \lambda_{x,y}^{x',y'}(\boldsymbol{g}, \alpha, \beta)\right)\right.\right. \\
&\qquad \left.\left.\left. \times \tanh\left(\beta g_{x,y+1} + \sum_{(x',y')\in \mathbf{c}_{x,y+1}\setminus(x,y)} \lambda_{x,y+1}^{x',y'}(\boldsymbol{g}, \alpha, \beta)\right)\right\}\right)\right\} \\
&= \frac{\exp(2\alpha)\cosh(6\lambda(\alpha)) - 1}{\exp(2\alpha)\cosh(6\lambda(\alpha)) + 1}.
\end{aligned} \tag{152}$$

3.4. Statistical-mechanical iterative algorithms

In this section, we give statistical-mechanical iterative algorithms by means of the mean-field and the Bethe approximations in the evidence framework and the MPM estimation derived above.

First, we explain the iterative algorithm as a numerical calculation scheme to solve the nonlinear equation that has the following form,

$$\lambda = \mathcal{G}(\lambda) \tag{153}$$

where λ is referred to as the fixed point of the given function $\mathcal{G}(\xi)$. In the general numerical recipe, the procedure to solve the fixed point equation (153) is given as follows:

Basic iterative algorithm to solve the nonlinear equation $\lambda = \mathcal{G}(\lambda)$

Step 1: Set $a(0)$ as an initial value and $t \leftarrow 0$.
Step 2: Update $r \leftarrow r + 1$ and

$$a(r) \leftarrow \mathcal{G}(a(r-1)). \tag{154}$$

Step 3: Update $R \leftarrow r$ and $\lambda \leftarrow a(R)$. Stop if it is satisfied that

$$|a(r) - a(r-1)| < \varepsilon \tag{155}$$

and go to *step 2* otherwise.

Here, ε should be set as a sufficiently small positive number. Usually, it is enough to set $\varepsilon = 10^{-6}$. The iterative scheme is one of the typical techniques to solve a nonlinear equation and is also often referred to as a fixed point iteration or a fixed point method [58, 59]. Because the mean-field deterministic equations (103) and (110) and the Bethe deterministic equations (140), (141) and (150) have the forms of the simultaneous fixed point equations, the iterative algorithm to solve them can be constructed by extending the above iterative procedure.

In the mean-field approximation, the algorithm is given by using equations (103) and (110) as follows:

Algorithm 3.1. Statistical-mechanical image restoration algorithm in the mean-field approximation

Step 1: Set $r \leftarrow 0$ as an initial value.
Step 2: Update $r \leftarrow r + 1$ and

$$a_{x,y}(r) \leftarrow \tanh\left(\beta g_{x,y} + \alpha \sum_{(x',y') \in \mathbf{c}_{x,y}} a_{x',y'}(r-1)\right) \qquad (x, y) \in \Omega. \tag{156}$$

Step 3: Update $R \leftarrow r$ and $m_{x,y}(\boldsymbol{g}, \alpha, \beta) \leftarrow a_{x,y}(R)$. Stop if it is satisfied that

$$\sum_{(x,y)\in\Omega} |a_{x,y}(r) - a_{x,y}(r-1)| < \varepsilon \tag{157}$$

and go to *step 2* otherwise.
Step 4: Calculate the marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta)$, by substituting $m_{x,y}(\boldsymbol{g}, \alpha, \beta)$ into equation (101).

We can calculate the evidence $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}$ and both sides of the extremum conditions (111) and (112) by substituting the values of marginal probability distributions obtained by using the above approximations for a set of fixed values of α and β in equation (104). By running the above algorithm for various values of α and β, we obtain the estimates $\hat{\alpha}$ and $\hat{\beta}$ so as to satisfy the deterministic equations (111) and (112).

In the Bethe approximation, the algorithm is given by using equations (140), (141) and (150) as follows:

Algorithm 3.2. Statistical-mechanical image restoration algorithm in Bethe approximation

Step 1: Set $r \leftarrow 0$ as an initial value.
Step 2: Update $r \leftarrow r + 1$, and

$$a^{x,y}_{x\pm1,y}(r) \leftarrow \operatorname{arctanh}\left(\tanh(\alpha)\tanh\left(\beta g_{x,y} + \sum_{(x',y')\in\mathbf{c}_{x,y}\setminus(x\pm1,y)} a^{x',y'}_{x,y}(r-1)\right)\right) \tag{158}$$

$$a^{x,y}_{x,y\pm1}(r) \leftarrow \operatorname{arctanh}\left(\tanh(\alpha)\tanh\left(\beta g_{x,y} + \sum_{(x',y')\in\mathbf{c}_{x,y}\setminus(x,y\pm1)} + a^{x',y'}_{x,y}(r-1)\right)\right). \tag{159}$$

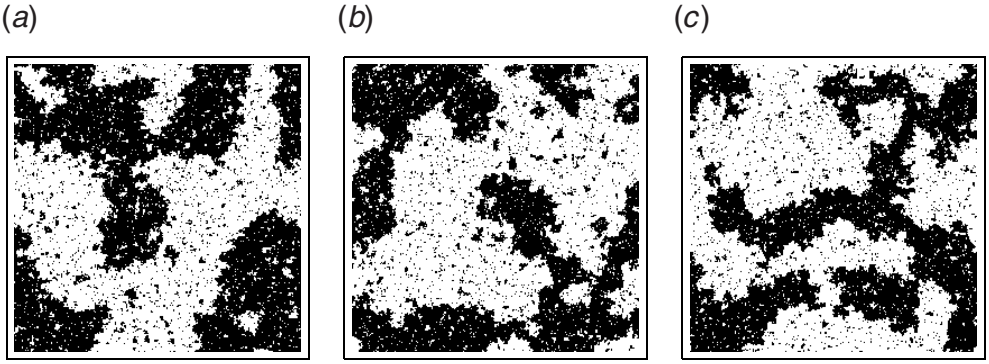

Figure 10. Binary images $\boldsymbol{f}^*$ generated by the *a priori* probability distribution (77) for $\alpha = 2.15^{-1}$. (*a*) Sample 1; (*b*) sample 2; (*c*) sample 3.

Step 3: Update $R \leftarrow r$. Set $\lambda_{x,y}^{x+1,y}(\boldsymbol{g}, \alpha, \beta) \leftarrow a_{x,y}^{x+1,y}(R)$ and $\lambda_{x,y}^{x,y+1}(\boldsymbol{g}, \alpha, \beta) \leftarrow a_{x,y}^{x,y+1}(R)$ and go to *step 4* if it is satisfied that

$$\frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} \left(\left| a_{x,y}^{x+1,y}(r) - a_{x,y}^{x+1,y}(r-1) \right| + \left| a_{x,y}^{x,y+1}(r) - a_{x,y}^{x,y+1}(r-1) \right| \right) < \varepsilon \tag{160}$$

and go to *step 2* otherwise.

Step 4: Calculate the marginal probability distributions $\rho_{x,y}(\zeta|\boldsymbol{g}, \alpha, \beta)$, $\rho_{x,y}^{x+1,y}(\zeta, \zeta'|\boldsymbol{g}, \alpha, \beta)$ and $\rho_{x,y}^{x+1,y}(\zeta, \zeta'|\boldsymbol{g}, \alpha, \beta)$, by substituting $\lambda_{x,y}^{x+1,y}(\boldsymbol{g}, \alpha, \beta)$ and $\lambda_{x,y}^{x,y+1}(\boldsymbol{g}, \alpha, \beta)$ into equations (142)–(147).

We can calculate the evidence $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \beta\}$ and both sides of the extremum conditions (151) and (152) by substituting the values of marginal probability distributions obtained by using the above approximations for a set of fixed values of α and β in equation (148). By running the above algorithm for various values of α and β, the estimates $\hat{\alpha}$ and $\hat{\beta}$ are obtained so as to satisfy the deterministic equations (151) and (152).

3.5. Numerical experiments

In this section, some numerical experiments are given. We assume that the original image is generated by the *a priori* probability distribution (77). First, three binary images in figure 10, which are generated by Monte Carlo simulations in the *a priori* probability distribution (77) for $\alpha = 2.15^{-1}$, are adopted as original images. Second, two binary images in figure 11, which are obtained from the standard images 'home' and 'mandrill,' are adopted as original images. In this subsection, the original image $\boldsymbol{f}$ and the observable degraded image $\boldsymbol{g}$ are denoted by $\boldsymbol{f}^*$ and $\boldsymbol{g}^*$.

We have a given degraded image $\boldsymbol{g}^*$ which is obtained from the original image $\boldsymbol{f}^*$ by changing the state of each pixel to another state by the same probability p, independently of the other pixels. Here p is assumed to be less than $\frac{1}{2}$. This degradation process is called a binary symmetric channel. The conditional probability distribution $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\}$ is given by

$$\begin{aligned} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\} &= \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, p\} \\ &\equiv \prod_{(x,y)\in\Omega} ((1 - \delta_{f_{x,y},g_{x,y}})p + (1-p)\delta_{f_{x,y},g_{x,y}}). \end{aligned} \tag{161}$$

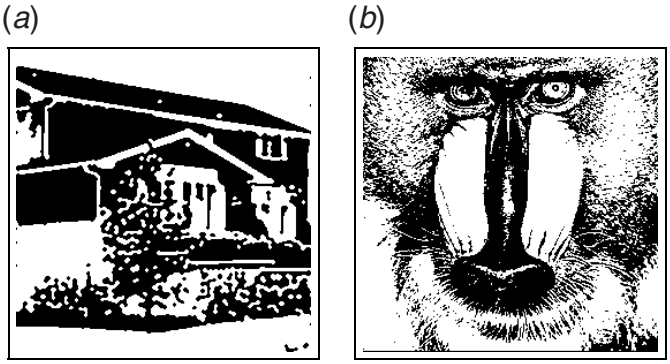

Figure 11. Artificial binary images $\boldsymbol{f}^*$ generated from the 256-valued standard images 'home' and 'mandrill' by using a thresholded processing. (*a*) Home; (*b*) mandrill.

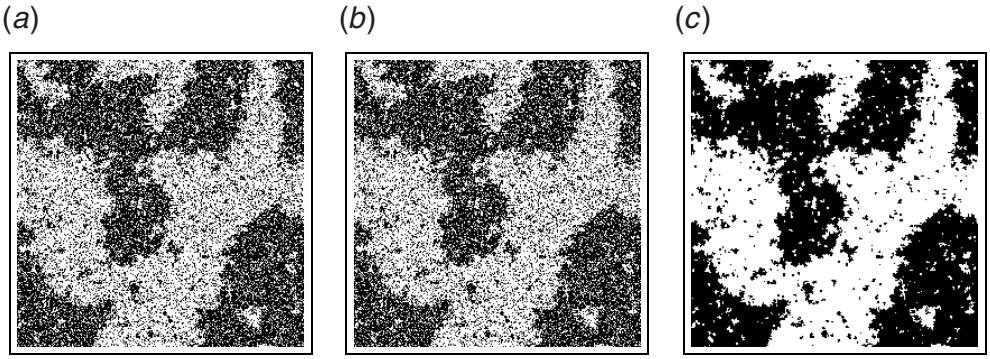

Figure 12. Restorations of the original image $\boldsymbol{f}^*$ given in figure 10(*a*). (*a*) Degraded image $\boldsymbol{g}^*$ ($p^* = 0.2$); (*b*) restored image $\hat{\boldsymbol{f}}$ by the mean-field approximation; (*c*) restored image $\hat{\boldsymbol{f}}$ by the Bethe approximation.

The conditional probability distribution can be rewritten as equation (76) by setting

$$\beta \equiv \frac{1}{2}\ln\left(\frac{1-p}{p}\right). \tag{162}$$

By setting $p = 0.2$, the degraded images $\boldsymbol{g}^*$ are produced from the original images $\boldsymbol{f}^*$ in the degradation process given in equation (161). The degraded images $\boldsymbol{g}^*$ produced by setting $p = 0.2$ from the original images $\boldsymbol{f}^*$ in figures 10(*a*)–(*c*) are shown in figures 12(*a*), 13(*a*) and 14(*a*), respectively. The true values of the hyperparameters α, p and $\beta = \frac{1}{2}\ln\left(\frac{1-p}{p}\right)$ are denoted by α^*, p^* and $\beta^* = \frac{1}{2}\ln\left(\frac{1-p^*}{p^*}\right)$, respectively. The restored images $\hat{\boldsymbol{f}}$ obtained by applying the statistical-mechanical iterative algorithms in the mean-field and the Bethe approximations to these degraded images $\boldsymbol{g}^*$ are shown in figures 12, 13 and 14.

To evaluate the restoration performance quantitatively, ten original images $\boldsymbol{f}^*$ are generated by Monte Carlo simulations in the *a priori* probability distribution (77) for $\alpha = 2.15^{-1}$. We produce a degraded image $\boldsymbol{g}^*$ from each original image $\boldsymbol{f}^*$ by means of the degradation process (76) for $p^* = 0.2$. By applying the iterative algorithms of the mean-field approximation and the Bethe approximation to each degraded image $\boldsymbol{g}^*$, we obtain the estimates of the hyperparameters $\hat{p}$ and $\hat{\alpha}$ and the restored image $\hat{\boldsymbol{f}}$ for each degraded image $\boldsymbol{g}^*$.

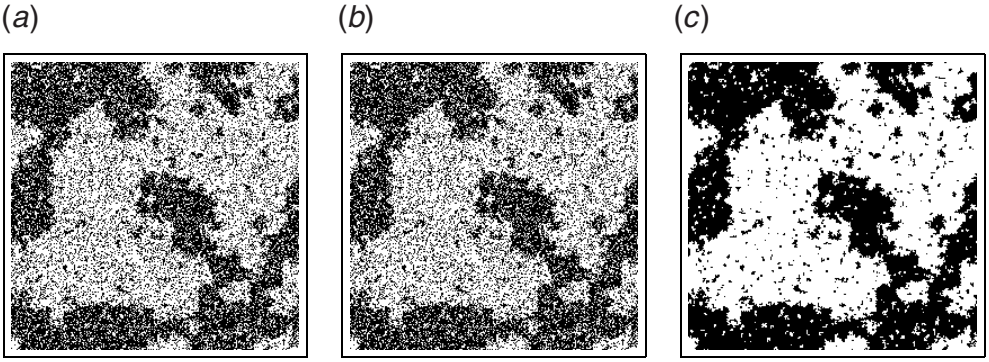

Figure 13. Restorations of the original image $\boldsymbol{f}^*$ given in figure 10(*b*). (*a*) Degraded image $\boldsymbol{g}^*$ ($p^* = 0.2$); (*b*) restored image $\hat{\boldsymbol{f}}$ by the mean-field approximation; (*c*) restored image $\hat{\boldsymbol{f}}$ by the Bethe approximation.

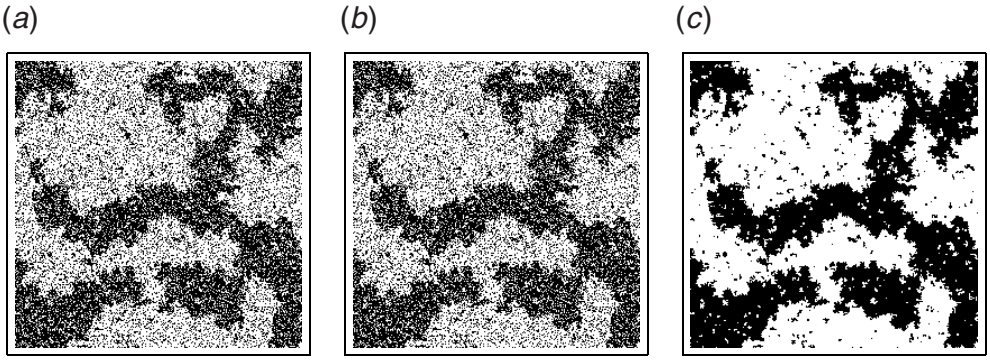

Figure 14. Restorations of the original image $\boldsymbol{f}^*$ given in figure 10(*c*). (*a*) Degraded image $\boldsymbol{g}^*$ ($p^* = 0.2$); (*b*) restored image $\hat{\boldsymbol{f}}$ by the mean-field approximation; (*c*) restored image $\hat{\boldsymbol{f}}$ by the Bethe approximation.

From these ten degraded images and the corresponding restored images $\hat{\boldsymbol{f}}$, we calculate the confidence intervals in the confidence coefficient 95% of the estimates of hyperparameters, $\hat{p}$ and $\hat{\alpha}$, and the values of the mean square error $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$,

$$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) \equiv \frac{1}{|\Omega|}\|\boldsymbol{f}^* - \hat{\boldsymbol{f}}\|^2 \tag{163}$$

and the improvement of signal to noise ratio, Δ_{SNR}:

$$\Delta_{\mathrm{SNR}} \equiv 10\log_{10}\left(\frac{\|\boldsymbol{f}^* - \boldsymbol{g}^*\|^2}{\|\boldsymbol{f}^* - \hat{\boldsymbol{f}}\|^2}\right) \quad (\mathrm{dB}). \tag{164}$$

These confidence intervals are given in table 1. We remark that, in binary image restoration, $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})/2$ is equal to the Hamming distance between the images $\boldsymbol{f}^*$ and $\hat{\boldsymbol{f}}$, $\frac{1}{|\Omega|}\sum_{(x,y)\in\Omega}\delta_{f^*_{x,y},\hat{f}_{x,y}}$. These results show that very reliable estimated values of the hyperparameters α and β can be obtained in the Bethe approximation. In the mean-field approximation, on the other hand, the estimated values are far from the original values. This result clearly shows that the mean-field approximation is not appropriate in the present situation. Similar results can be obtained by applying the Metropolis Monte Carlo simulations instead of the Bethe approximation [25].

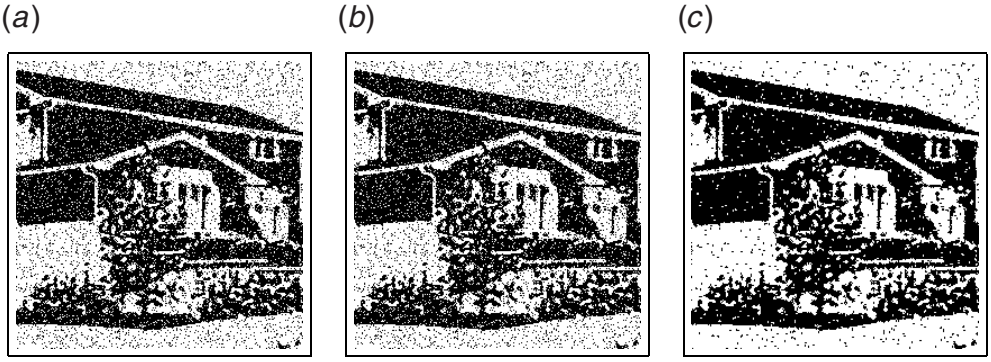

Figure 15. Binary image restoration of the original image $\boldsymbol{f}^*$ given in figure 14(*a*). (*a*) Degraded image $\boldsymbol{g}^*$ ($p^* = 0.1$); (*b*) restored image $\hat{\boldsymbol{f}}$ by the mean-field approximation; (*c*) restored image $\hat{\boldsymbol{f}}$ by the Bethe approximation.

Table 1. The confidence interval in the confidence coefficients 95% of the estimates of hyperparameters, $\hat{p}$ and $\hat{\alpha}$, and the values of $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} obtained for some degraded images $\boldsymbol{g}^*$, which are produced for $p^* = 0.2$ from ten original images $\boldsymbol{f}^*$. The ten original images $\boldsymbol{f}^*$ are generated by Monte Carlo simulations in the *a priori* probability distribution (77) for $\alpha = 2.15^{-1}$. The hyperparameters are estimated by applying the mean-field approximation and the Bethe approximation to the maximum marginal likelihood (MML) estimation.

	Mean-field approximation	Bethe approximation
$\hat{p}$	$[0.054\,98 \pm 0.001\,00]$	$[0.194\,54 \pm 0.001\,33]$
$\hat{\beta}$	$[1.573\,44 \pm 0.010\,45]$	$[0.861\,15 \pm 0.004\,60]$
$\hat{\alpha}$	$[0.268\,59 \pm 0.000\,01]$	$[0.459\,03 \pm 0.001\,66]$
$1/\hat{\alpha}$	$[4.274\,02 \pm 0.000\,11]$	$[2.689\,62 \pm 0.009\,68]$
$d(\boldsymbol{f}^*, \boldsymbol{g}^*)/2$	$[0.214\,43 \pm 0.000\,62]$	$[0.214\,43 \pm 0.000\,62]$
$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})/2$	$[0.214\,43 \pm 0.000\,62]$	$[0.092\,22 \pm 0.001\,64]$
Δ_{SNR} (dB)	$[0 \pm 0]$	$[3.929\,01 \pm 0.084\,05]$

We also performed numerical experiments for artificial binary images in figure 11 as original images. The artificial binary images $\boldsymbol{f}^*$ are generated from the 256-valued standard images 'home' and 'mandrill' by using a thresholded processing[3]. The image restoration by means of the iterative algorithms of the mean-field approximation and the Bethe approximation for the cases of $p^* = 0.1$ and $p^* = 0.2$ is shown in figures 15, 16, 17 and 18. We give in tables 2 and 3 the estimates of hyperparameters, $\hat{p}$ and $\hat{\alpha}$, and the values of the mean square error $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and the improvement of signal to noise ratio, Δ_{SNR} (dB).

In the present framework for binary image restoration, it has been shown that we can obtain satisfactory results by the Bethe approximation if the images are produced by the assumed *a priori* probability distribution. The restored images obtained by means of the mean-field approximation are not satisfactory in quality. Although we have to calculate the free energies and the correlation functions between the nearest-neighbour pairs of pixels in the *a priori* and the *a posteriori* probability distributions in the estimation of hyperparameters in the evidence framework, the accuracy in the correlation functions of the mean-field approximation is not sufficient. This is one of the reasons why the results of the mean-field approximation are not satisfactory. On the other hand, although we also obtain good

[3] The standard images 'home' and 'mandrill' are shown in figure 19 in the next section.

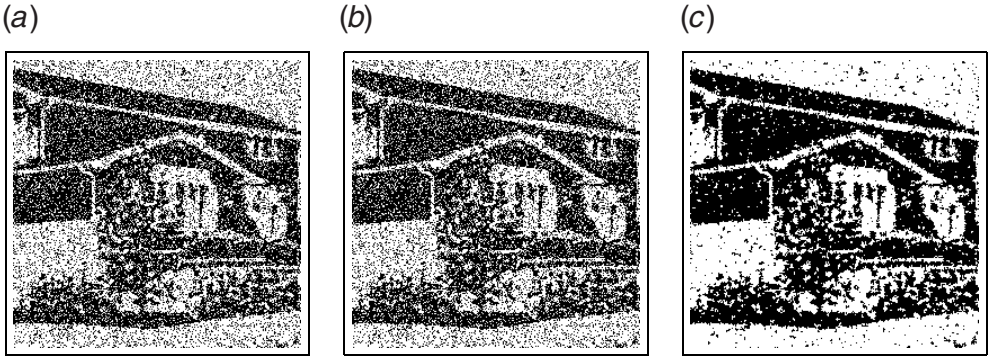

Figure 16. Binary image restoration of the original image $\boldsymbol{f}^*$ given in figure 14(*a*). (*a*) Degraded image $\boldsymbol{g}^*$ ($p^* = 0.2$); (*b*) restored image $\hat{\boldsymbol{f}}$ by the mean-field approximation; (*c*) restored image $\hat{\boldsymbol{f}}$ by the Bethe approximation.

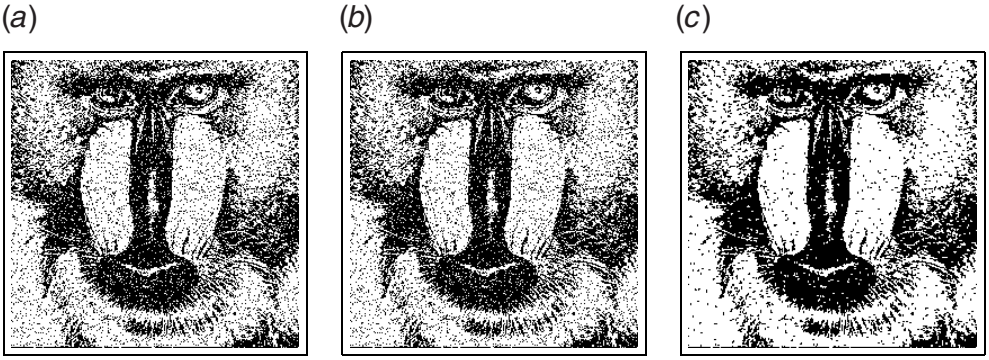

Figure 17. Binary image restoration of the original image $\boldsymbol{f}^*$ given in figure 11(*b*). (*a*) Degraded image $\boldsymbol{g}^*$ ($p^* = 0.1$); (*b*) restored image $\hat{\boldsymbol{f}}$ by the mean-field approximation; (*c*) restored image $\hat{\boldsymbol{f}}$ by the Bethe approximation.

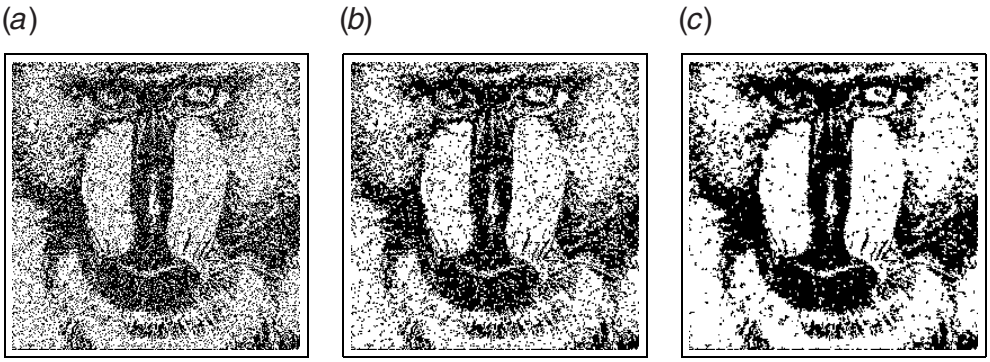

Figure 18. Binary image restoration of the original image $\boldsymbol{f}^*$ given in figure 11(*b*). (*a*) Degraded image $\boldsymbol{g}^*$ ($p^* = 0.2$). (*b*) restored image $\hat{\boldsymbol{f}}$ by the mean-field approximation; (*c*) restored image $\hat{\boldsymbol{f}}$ by the Bethe approximation.

results for image restoration in the practical binary images by the Bethe approximation, some noise remains in the results. The main reason is that the practical binary images are not

Table 2. The estimates of hyperparameters, $\hat{p}$ and $\hat{\alpha}$, and the values of $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} obtained for some degraded images $\boldsymbol{g}^*$, which are produced for various values of p^* from the original image $\boldsymbol{f}^*$ given in figure 14(*a*) (home). The hyperparameters are estimated by applying the mean-field approximation and the Bethe approximation to the MML estimation.

p^*	Approximation	$\hat{p}$	$\hat{\beta}$	$\hat{\alpha}$	$1/\hat{\alpha}$	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})/2$	Δ_{SNR} (dB)
0.1	Mean-field	0.009 96	2.299 39	0.252 37	3.962 45	0.099 53	0
	Bethe	0.048 57	1.487 48	0.410 97	2.433 24	0.047 61	3.202 83
0.2	Mean-field	0.076 93	1.242 44	0.251 35	3.978 47	0.197 91	0
	Bethe	0.132 90	0.937 79	0.392 16	2.549 95	0.083 19	3.763 91

Table 3. The estimates of hyperparameters, $\hat{p}$, $\hat{\beta}$, $\hat{\alpha}$, $1/\hat{\alpha}$, and the values of $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} obtained for some degraded images $\boldsymbol{g}^*$, which are produced for $p^* = 0.1$ and $p^* = 0.2$ from the original image $\boldsymbol{f}^*$ given in figure 14(*b*) (mandrill). The hyperparameters are estimated by applying the mean-field approximation and the Bethe approximation to the maximization of marginal likelihood.

p^*	Approximation	$\hat{p}$	$\hat{\beta}$	$\hat{\alpha}$	$1/\hat{\alpha}$	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})/2$	Δ_{SNR} (dB)
0.1	Mean-field	0.035 08	1.657 22	0.251 524	3.975 76	0.099 03	0
	Bethe	0.079 02	1.207 85	0.289 81	2.569 34	0.070 72	1.462 00
0.2	Mean-field	0.100 90	1.093 65	0.250 96	3.984 72	0.159 65	0.947 88
	Bethe	0.171 84	0.786 31	0.382 06	2.617 38	0.118 64	2.237 37

produced by the assumed *a priori* probability distribution, although the *a priori* probability distribution (77) is a fair approximation of the original images. It is necessary to improve the *a priori* probability distribution so as to be applicable to the practical image with complicated structures.

3.6. Concluding remarks

In this section, we have explained the basic framework of image restoration of binary images by means of the Bayes formula and evidence framework. The algorithms have been constructed by applying the mean-field approximation and the Bethe approximation. Some numerical experiments show that the Bethe approximation can give good results in image restoration, while the results of the mean-field approximation are not satisfactory. Nevertheless, it is obvious that our assumptions for the *a priori* probability distribution for original images are not good enough for practical images even if we restrict the original images to binary ones. From the standpoint of statistical mechanics, some extensions have been made [60–62]. Moreover, the framework can also be applied to image segmentations by extending the model system to the Potts model [24, 26].

Usually, most computer scientists and statisticians formulate Bayesian image restoration as a MAP estimation given in equation (8). In the MAP estimation, we have to search for the most probable configuration of the massive probabilistic model, a process of exponential order of complexity. Although such a problem can be treated by simulated annealing with the heat bath or the Metropolis Monte Carlo simulations, it takes a very long time for relaxation. Recently, some authors adopted the MPM estimation in equation (9) instead of the MAP estimation.

In the design of a probabilistic information processing system, the performance estimation is very important. In the image restoration, $\mathcal{M}_{x,y}(\alpha, \beta)$ in equation (40) corresponds to a statistical measure in the performance estimation. However, it has not been possible

to calculate exactly the statistical average such as $\mathcal{M}_{x,y}(\alpha, \beta)$ for the system where the *a priori* probability distribution is assumed to be the Ising model with interactions only between nearest-neighbour pairs of pixels. We can nevertheless calculate the statistical performance not as an exact value but as an approximate value by the Bethe approximation. Such investigations have been carried out for the Ising model with random bond interactions between nearest-neighbour pairs of pixels [63, 67]. Other statistical-mechanical methods will also be useful for the estimation of the statistical performance in Markov random field models. This is one of the future problems.

The Bethe approximation has already been applied to other probabilistic information processing problems. Kabashima and Saad [68] suggested that a belief propagation algorithm in error-correcting codes has a very close relationship with the Bethe approximation[4]. The formulae of Kabashima and Saad are derived by means of the following replacements,

$$\lambda^{x,y}_{x\pm1,y}(\boldsymbol{g}, \alpha, \beta) \equiv \tanh\left(\beta g_{x,y} + \sum_{(x',y')\in \mathbf{c}_{x,y}\setminus(x\pm1,y)} \lambda^{x',y'}_{x,y}(\boldsymbol{g}, \alpha, \beta)\right) \tag{165}$$

$$\lambda^{x,y}_{x,y\pm1}(\boldsymbol{g}, \alpha, \beta) \equiv \tanh\left(\beta g_{x,y} + \sum_{(x',y')\in \mathbf{c}_{x,y}\setminus(x,y\pm1)} \lambda^{x',y'}_{x,y}(\boldsymbol{g}, \alpha, \beta)\right) \tag{166}$$

in equations (140) and (141). Belief propagation is one of the familiar techniques in the probabilistic inference of intelligent information processing [71, 72]. Yedidia *et al* [73] and Kappen *et al* [74] suggested that the belief propagation algorithm can be derived from the variational principle of an approximate free energy in the Bethe approximation and proposed a method for the construction of a generalized belief propagation based on the cluster variation method [75, 76]. Their simultaneous recursion formulae in the belief propagation are equivalent to equation (137).

As another advanced mean-field method for probabilistic image restoration, some cluster-type mean-field approximations were proposed [77, 78]. The cluster-type mean-field approximations can be regarded as an extension of the mean-field approximation or the Bethe approximation from the standpoint of the effective field theory [79].

In this section, we have explained that not only the statistical-mechanical framework but also the statistical-mechanical approximations are applicable to probabilistic image processing. These situations are similar in grey-level image restoration. In the next section, we introduce a grey level image processing by means of a familiar statistical-mechanical model with continuous degree of freedom at each pixel.

4. Grey-level image processing and Gaussian model

In this section, we discuss grey-level image processing. In practical digital images, each pixel usually has 256 grey levels, 0, 1, 2, ..., 255. Such grey levels can be regarded as continuous degrees of freedom. Actually, in many conventional digital image filters, the degree of freedom at each pixel is treated as a continuous variable, which we adopt here.

One of the basic degradation processes in grey-level images is an additive white Gaussian noise with average 0 and variance σ^2 ($\sigma > 0$). Generally, when the probability distribution of the random variable A is given by

$$\Pr\{A = a\} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(a-\mu)^2\right) \qquad -\infty < a < +\infty \tag{167}$$

[4] They referred to the recursion formulae of effective fields in the Bethe approximation as Touless, Anderson and Palmer (TAP) equation [69, 70].

it is denoted by the notation $A \sim \mathcal{N}[\mu, \sigma^2]$ in probability theory and statistics. By using this notation, the degradation process from the original image $\boldsymbol{f}$ to the degraded image $\boldsymbol{g}$ by additive white Gaussian noise with average 0 and variance σ^2 can be expressed as

$$G_{x,y} - F_{x,y} \sim \mathcal{N}[0, \sigma^2] \tag{168}$$

where $F_{x,y}$ and $G_{x,y}$ are random variables at each pixel (x, y) of the original image and the degraded image, respectively.

One of the most famous linear filters is the following constrained least mean square filter [1, 80],

$$\hat{\boldsymbol{f}} = \arg \min_{\boldsymbol{z}:\|\boldsymbol{z}-\boldsymbol{g}\|^2=|\Omega|\,\sigma^2} \left(\sum_{(x,y)\in\Omega} ((z_{x,y} - z_{x+1,y})^2 + (z_{x,y} - z_{x,y+1})^2) \right) \tag{169}$$

where

$$\|\boldsymbol{z} - \boldsymbol{g}\|^2 \equiv \sum_{(x,y)\in\Omega} (z_{x,y} - g_{x,y})^2. \tag{170}$$

Here the degradation process is assumed to be additive white Gaussian noise $\mathcal{N}[0, \sigma^2]$. The algorithm for obtaining the optimal solution $\hat{\boldsymbol{f}}$ was constructed by introducing a Lagrange multiplier γ to ensure the constraint $\|\boldsymbol{z} - \boldsymbol{g}\|^2 = |\Omega|\,\sigma^2$,

$$\hat{\boldsymbol{z}}(\gamma) = \arg\min_{\boldsymbol{z}} H(\boldsymbol{z}|\gamma) \tag{171}$$

$$H(\boldsymbol{z}|\gamma) \equiv \sum_{(x,y)\in\Omega} ((z_{x,y} - z_{x+1,y})^2 + (z_{x,y} - z_{x,y+1})^2) + \gamma\|\boldsymbol{z} - \boldsymbol{g}\|^2 \tag{172}$$

and the Lagrange multiplier γ should be determined so as to satisfy the constraint $\|\hat{\boldsymbol{z}}(\gamma) - \boldsymbol{g}\|^2 = |\Omega|\sigma^2$. The restored image $\hat{\boldsymbol{z}}(\gamma)$ is obtained from the stationarity condition for $H(\boldsymbol{z}|\gamma)$ with respect to $z_{x,y}$. Some statistical physicists may note that $H(\boldsymbol{z}|\gamma)$ is the Hamiltonian of the Gaussian model [81, 82].

The Gaussian model can be solved analytically and closed expressions of some correlation functions can be calculated exactly by means of the Gaussian integral formula and the discrete Fourier transform. The Gaussian model of equation (172) expresses a smoothing effect in the first and second terms and takes account of the information from observed data as the third term. Of course, these effects are not sufficient to treat practical real-world images, because many real-world images usually have not only smooth regions but also edges between two different smooth regions and sometimes include textures. Nevertheless, as a first step towards practical image restoration, it is important to understand the basic behaviour of the image restoration scheme by means of the solvable probabilistic model from the statistical-mechanical standpoint. As mentioned in section 1, it is possible to regard the framework of the conventional filter theory in image processing as a problem of classical spin systems.

In this section, we express the basic scheme of probabilistic image restoration by means of the Gaussian model and derive formulae to calculate the statistical performance.

4.1. Bayesian image processing by means of the Gaussian model

In the framework of image restoration for grey-level images, we consider the set $\mathbb{R}$ consisting of whole real numbers and assume that the state at each pixel takes any finite real number. In this case, $\boldsymbol{F}$ and $\boldsymbol{G}$ are not discrete but continuous random variables. A degradation process $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\}$ and an *a priori* probability distribution $\Pr\{\boldsymbol{F} = \boldsymbol{f}\}$ are assumed to be as follows.

4.1.1. Degradation process in grey-level images. We adopt the additive Gaussian noise $\mathcal{N}[0, \sigma^2]$ so that the conditional probability density function $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\}$ is assumed to be

$$\Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}\} = \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \sigma\} \equiv \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{|\Omega|}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{(x,y)\in\Omega}(f_{x,y} - g_{x,y})^2\right). \tag{173}$$

4.1.2. A priori probability distribution in grey-level images. The *a priori* probability density function, that the original image is $\boldsymbol{f}$, is given by

$$\begin{aligned}\Pr\{\boldsymbol{F} = \boldsymbol{f}\} &= \Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\} \\ &\equiv \frac{\exp\left(-\frac{1}{2}\alpha\sum_{(x,y)\in\Omega}((f_{x,y} - f_{x+1,y})^2 + (f_{x,y} - f_{x,y+1})^2)\right)}{\int_{\mathbb{R}^{|\Omega|}}\exp\left(-\frac{1}{2}\alpha\sum_{(x,y)\in\Omega}((z_{x,y} - z_{x+1,y})^2 + (z_{x,y} - z_{x,y+1})^2)\right)\mathrm{d}\boldsymbol{z}} \\ &= \frac{\exp\left(-\frac{1}{2}\alpha\boldsymbol{f}\boldsymbol{C}\boldsymbol{f}^{\mathrm{T}}\right)}{\int_{\mathbb{R}^{|\Omega|}}\exp\left(-\frac{1}{2}\alpha\boldsymbol{f}\boldsymbol{C}\boldsymbol{f}^{\mathrm{T}}\right)\mathrm{d}\boldsymbol{z}}.\end{aligned} \tag{174}$$

Here $\boldsymbol{C}$ is an $|\Omega| \times |\Omega|$ matrix whose $(x, y|x', y')$-element is defined by

$$\langle x, y|\boldsymbol{C}|x', y'\rangle \equiv \delta_{x,x'}\delta_{y,y'} - \tfrac{1}{4}\delta_{x,x'+1}\delta_{y,y'} - \tfrac{1}{4}\delta_{x,x'-1}\delta_{y,y'} - \tfrac{1}{4}\delta_{x,x'}\delta_{y,y'+1} - \tfrac{1}{4}\delta_{x,x'}\delta_{y,y'-1} \tag{175}$$

and the integral is defined by

$$\int_{\mathbb{R}^{|\Omega|}}\mathrm{d}\boldsymbol{z} \equiv \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\prod_{(x,y)\in\Omega}\mathrm{d}z_{x,y}. \tag{176}$$

By substituting equations (173) and (174) into equation (7), the *a posteriori* probability distribution is given as

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}\} = \Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g}, \alpha, \sigma\} = \frac{\exp(-E(\boldsymbol{f}|\boldsymbol{g}, \alpha, \sigma))}{\int_{\mathbb{R}^{|\Omega|}}\exp(-E(\boldsymbol{z}|\boldsymbol{g}, \alpha, \sigma))\,\mathrm{d}\boldsymbol{z}} \tag{177}$$

where

$$\begin{aligned}E(\boldsymbol{f}|\boldsymbol{g}, \alpha, \sigma) &\equiv \frac{1}{2\sigma^2}\sum_{(x,y)\in\Omega}(f_{x,y} - g_{x,y})^2 + \frac{1}{2}\alpha\sum_{(x,y)\in\Omega}((f_{x,y} - f_{x+1,y})^2 + (f_{x,y} - f_{x,y+1})^2) \\ &= \frac{1}{2\sigma^2}\|\boldsymbol{f} - \boldsymbol{g}\|^2 + \frac{1}{2}\alpha\boldsymbol{f}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{f} \\ &= \frac{1}{2\sigma^2}(\boldsymbol{f} - (\boldsymbol{I} + \alpha\sigma^2\boldsymbol{C})^{-1}\boldsymbol{g})^{\mathrm{T}}(\boldsymbol{I} + \alpha\sigma^2\boldsymbol{C})(\boldsymbol{f} - (\boldsymbol{I} + \alpha\sigma^2\boldsymbol{C})^{-1}\boldsymbol{g}) \\ &\quad + \frac{1}{2}\alpha\boldsymbol{g}^{\mathrm{T}}\boldsymbol{C}(\boldsymbol{I} + \alpha\sigma^2\boldsymbol{C})^{-1}\boldsymbol{g}.\end{aligned} \tag{178}$$

The *a priori* probability distribution (174) can be regarded as a conditional autoregressive model [83, 84] in the signal processing:

$$f_{x,y} - f_{x+1,y} \sim \mathcal{N}\left[0, \frac{1}{\alpha}\right] \qquad f_{x,y} - f_{x,y+1} \sim \mathcal{N}\left[0, \frac{1}{\alpha}\right]. \tag{179}$$

This probabilistic model is equivalent to the Gaussian model in statistical physics [81, 82].

In the framework of image restoration of grey-level images given in equation (177), the hyperparameters α and σ are determined by

$$(\hat{\alpha}, \hat{\sigma}) = \arg\max_{(\alpha,\sigma)} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \sigma\} \tag{180}$$

where

$$\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \sigma\} \equiv \int_{\mathbb{R}^{|\Omega|}} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \sigma\} \Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\} \, \mathrm{d}\boldsymbol{z}. \tag{181}$$

The maximum marginal likelihood estimates of α and σ are denoted by $\hat{\alpha}$ and $\hat{\sigma}$, respectively. For the obtained estimates $\hat{\alpha}$ and $\hat{\sigma}$, the restored image $\hat{\boldsymbol{f}} \equiv \{\hat{f}_{x,y}\}(\boldsymbol{h}(\boldsymbol{g}, \hat{\alpha}, \hat{\sigma}) \equiv \{h_{x,y}(\boldsymbol{g}, \hat{\alpha}, \hat{\sigma})\})$ is determined by

$$\hat{f}_{x,y} = h_{x,y}(\boldsymbol{g}, \hat{\alpha}, \hat{\sigma}) \equiv \int z_{x,y} \Pr\{\boldsymbol{F} = \boldsymbol{z}|\boldsymbol{G} = \boldsymbol{g}, \hat{\alpha}, \hat{\sigma}\} \, \mathrm{d}\boldsymbol{z}. \tag{182}$$

4.2. Discrete Fourier transform and exact expression of evidence

By introducing the unitary matrix $\boldsymbol{U}$ defined by

$$\langle x, y|\boldsymbol{U}|p, q\rangle \equiv \frac{1}{\sqrt{|\Omega|}} \exp\left(-\mathrm{i}\,\frac{2\pi px}{L_x} - \mathrm{i}\,\frac{2\pi qy}{L_y}\right) \tag{183}$$

we can diagonalize the matrix $\boldsymbol{C}$ as follows,

$$\langle p, q|\boldsymbol{U}^{-1}\boldsymbol{C}\boldsymbol{U}|p', q'\rangle = \delta_{p,p'}\delta_{q,q'}\lambda(p, q) \tag{184}$$

where

$$\lambda(p, q) \equiv 1 - \frac{1}{2}\cos\left(\frac{2\pi p}{L_x}\right) - \frac{1}{2}\cos\left(\frac{2\pi q}{L_y}\right). \tag{185}$$

By using equation (184), the partition functions of the *a priori* and the *a posteriori* probability density functions in equations (174) and (177) can be expressed in the following form:

$$\begin{aligned} Z_{\mathrm{prior}}(\alpha) &\equiv \int_{\mathbb{R}^{|\Omega|}} \exp\left(-\frac{1}{2}\alpha \sum_{(x,y)\in\Omega} ((z_{x,y} - z_{x+1,y})^2 + (z_{x,y} - z_{x,y+1})^2)\right) \mathrm{d}\boldsymbol{z} \\ &= (2\pi)^{\frac{|\Omega|}{2}} \{\det(\alpha\boldsymbol{C})\}^{-\frac{1}{2}} = (2\pi)^{\frac{|\Omega|}{2}} \left(\prod_{p=0}^{L_x-1}\prod_{q=0}^{L_y-1} \frac{1}{\alpha\lambda(p, q)}\right)^{\frac{1}{2}} \end{aligned} \tag{186}$$

$$\begin{aligned} Z_{\mathrm{posterior}}(\alpha, \sigma) &\equiv \int_{\mathbb{R}^{|\Omega|}} \exp(-E(\boldsymbol{z}|\boldsymbol{g}, \alpha, \sigma)) \, \mathrm{d}\boldsymbol{z} \\ &= (2\pi\sigma^2)^{\frac{|\Omega|}{2}} \{\det(\boldsymbol{I} + \alpha\sigma^2\boldsymbol{C})\}^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}\alpha\boldsymbol{g}^{\mathrm{T}}\boldsymbol{C}(\boldsymbol{I} + \alpha\sigma^2\boldsymbol{C})^{-1}\boldsymbol{g}\right) \\ &= (2\pi)^{\frac{|\Omega|}{2}} \left(\prod_{p=0}^{L_x-1}\prod_{q=0}^{L_y-1} \frac{\sigma^2}{1 + \alpha\sigma^2\lambda(p, q)}\right)^{\frac{1}{2}} \\ &\quad \times \exp\left(-\frac{1}{2}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1} |G(p, q)|^2 \frac{\alpha\lambda(p, q)}{1 + \alpha\sigma^2\lambda(p, q)}\right) \end{aligned} \tag{187}$$

where

$$G(p, q) \equiv \frac{1}{\sqrt{|\Omega|}} \sum_{(x,y)\in\Omega} g_{x,y} \exp\left(-\mathrm{i}\,\frac{2\pi px}{L_x} - \mathrm{i}\,\frac{2\pi qy}{L_y}\right). \tag{188}$$

By substituting equations (186) and (187) into equation (181), the logarithm of marginal likelihood $\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \sigma\}$ is obtained in the following form:

$$\begin{aligned}\ln(\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \sigma\}) &= \ln(Z_{\text{posterior}}(\alpha, \sigma)) - \ln(Z_{\text{prior}}(\alpha, \sigma)) - \ln(\sqrt{2\pi}\sigma) \\ &= -\frac{|\Omega|}{2}\ln(2\pi) - \frac{1}{2}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\ln(1+\alpha\sigma^2\lambda(p,q)) + \frac{|\Omega|}{2}\ln(\alpha) \\ &\quad + \frac{1}{2}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\ln(\lambda(p,q)) - \frac{1}{2}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}|G(p,q)|^2\frac{\alpha\lambda(p,q)}{1+\alpha\sigma^2\lambda(p,q)}. \end{aligned} \tag{189}$$

By replacing the summation $\frac{1}{L_y}\sum_{q=0}^{L_y-1}$ by the integral $\frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\theta$ and using the integral formula,

$$\frac{1}{2\pi}\int_0^{2\pi}\ln(2(a-\cos(\theta)))\,\mathrm{d}\theta = \operatorname{arccosh}(|a|) \tag{190}$$

we can rewrite the fourth term of equation (189):

$$\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\ln(\lambda(p,q)) = -2\ln(2) + \sum_{p=0}^{L_x-1}\operatorname{arccosh}\left(2-\cos\left(\frac{2\pi p}{L_x}\right)\right). \tag{191}$$

The extremum conditions of $\rho\{\boldsymbol{G} = \boldsymbol{g}^*|\alpha, \sigma\}$ at $\alpha = \hat{\alpha}$ and $\sigma = \hat{\sigma}$ can be reduced to the following simultaneous equations:

$$\frac{1}{\hat{\alpha}} = \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\frac{\hat{\sigma}^2\lambda(p,q)}{1+\hat{\alpha}\hat{\sigma}^2\lambda(p,q)} + \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}|G(p,q)|^2\frac{\lambda(p,q)}{(1+\hat{\alpha}\hat{\sigma}^2\lambda(p,q))^2} \tag{192}$$

$$\hat{\sigma}^2 = \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\frac{\hat{\sigma}^2}{1+\hat{\alpha}\hat{\sigma}^2\lambda(p,q)} + \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}|G(p,q)|^2\frac{\hat{\alpha}^2\hat{\sigma}^4\lambda(p,q)^2}{(1+\hat{\alpha}\hat{\sigma}^2\lambda(p,q))^2}. \tag{193}$$

The restored image $\hat{\boldsymbol{f}}$ in equation (182) can be expressed explicitly as

$$\hat{\boldsymbol{f}} = \boldsymbol{h}(\boldsymbol{g}, \hat{\alpha}, \hat{\sigma}) = (\boldsymbol{I} + \hat{\alpha}\hat{\sigma}^2\boldsymbol{C})^{-1}\boldsymbol{g} \tag{194}$$

and

$$\begin{aligned}\hat{f}_{x,y} = h_{x,y}(\boldsymbol{g}, \hat{\alpha}, \hat{\sigma}) &= \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\frac{1}{1+\hat{\alpha}\hat{\sigma}^2\lambda(p,q)}\left(\cos\left(\frac{2\pi px}{L_x}+\frac{2\pi qy}{L_y}\right)\operatorname{Re}(G(p,q))\right. \\ &\quad \left. + \sin\left(\frac{2\pi px}{L_x}+\frac{2\pi qy}{L_y}\right)\operatorname{Im}(G(p,q))\right). \end{aligned} \tag{195}$$

In this way, the restoration process reduces to this simple arithmetic computation.

4.3. Statistical performance

In this section, we derive some statistical properties of the present model. We calculate the statistical averages of the mean square error $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and the log-evidence $\ln(\Pr\{\boldsymbol{g}|\alpha, \beta\})$. The results are useful for estimating the statistical performance of the Bayesian approach to image restoration analytically.

The value σ^* is for hyperparameter σ when the degraded image $\boldsymbol{g}$ is generated from the given original image $\boldsymbol{f}$ by the degradation process (173). The value α^* is for hyperparameter α when the original image $\boldsymbol{f}$ is generated by the *a priori* probability density function (174).

To treat practical grey-level real-world images, we assume that they are generated by the *a priori* probability density function (174), though such an assumption is not valid exactly.

In the maximum likelihood estimation, the value α^* is determined so as to maximize the likelihood $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\}$ with respect to the hyperparameters α and σ. The logarithm of likelihood $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\}$ can be expressed as

$$\ln(\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\}) = \frac{|\Omega|}{2}\ln(2\pi) - \frac{|\Omega|}{2}\ln(\alpha) - \frac{1}{2}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\ln(\lambda(p,q)) + \frac{1}{2}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}|F(p,q)|^2\alpha\lambda(p,q) \tag{196}$$

where

$$F(p,q) \equiv \frac{1}{\sqrt{|\Omega|}}\sum_{(x,y)\in\Omega} f_{x,y}\exp\left(-\mathrm{i}\frac{2\pi px}{L_x} - \mathrm{i}\frac{2\pi qy}{L_y}\right). \tag{197}$$

The conditions for an extremum of $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\}$ at $\alpha = \alpha^*$ are reduced to the following equations:

$$\frac{1}{\alpha^*} = \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\lambda(p,q) + \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}|F(p,q)|^2\lambda(p,q). \tag{198}$$

Now, as a measure of restoration performance, we define the statistical average of the difference between the original and corresponding restored images by

$$\mathcal{M}(\alpha,\sigma) \equiv \int\int \|\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{g},\alpha,\sigma)\|^2 \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{z},\sigma^*\}\Pr\{\boldsymbol{F} = \boldsymbol{z}|\alpha^*\}\,\mathrm{d}\boldsymbol{z}\,\mathrm{d}\boldsymbol{g} \tag{199}$$

where $\boldsymbol{h}(\boldsymbol{g},\alpha,\beta)$ is defined by equation (182) and is obtained as equation (194). By substituting equations (173), (174) and (195) into the right-hand side of equation (199) and using equation (178), we obtain the expression of the quantity $\mathcal{M}(\alpha,\sigma)$ as follows:

$$\mathcal{M}(\alpha,\sigma) = \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\frac{\sigma^{*2}}{1+\alpha^*\sigma^{*2}\lambda(p,q)} + \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\left(\frac{\sigma^{*2}}{(1+\alpha^*\sigma^{*2}\lambda(p,q))^2} - \frac{\sigma^2}{(1+\alpha\sigma^2\lambda(p,q))^2}\right)^2. \tag{200}$$

It is obvious that the statistical overlap $\mathcal{M}(\alpha,\sigma)$ gives the maximum value only at $(\alpha,\sigma) = (\alpha^*,\sigma^*)$. The result (200) was first obtained by Nishimori [85].

We consider a statistical average $\mathcal{L}(\alpha,\sigma)$ of the logarithm of marginal likelihood $\ln(\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha,\sigma\})$ with respect to the degradation process (173) and the *a priori* probability distribution (174):

$$\mathcal{L}(\alpha,\sigma) \equiv \int \ln\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha,\sigma\}\left(\int \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{z},\sigma^*\}\Pr\{\boldsymbol{F} = \boldsymbol{z}|\alpha^*\}\,\mathrm{d}\boldsymbol{z}\right)\mathrm{d}\boldsymbol{g}. \tag{201}$$

By substituting equations (173) and (174), and the third expression of equation (189) into the right-hand side of (201) and by using equation (178), we obtain the expression of the quantity $\mathcal{L}(\alpha,\sigma)$ as follows:

$$\mathcal{L}(\alpha,\sigma) = -\frac{|\Omega|}{2}\ln(2\pi) + \frac{1}{2}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\ln\left(\frac{\alpha\lambda(p,q)}{1+\alpha\sigma^2\lambda(p,q)}\right) - \frac{1}{2}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\left(\frac{\alpha\lambda(p,q)}{\alpha^*\lambda(p,q)}\right)\left(\frac{1+\alpha^*\sigma^{*2}\lambda(p,q)}{1+\alpha\sigma^2\lambda(p,q)}\right). \tag{202}$$

By differentiating the quantity $\mathcal{L}(\alpha,\sigma)$ with respect to the hyperparameters α and σ, the conditions for extrema of $\mathcal{L}(\alpha,\sigma)$ at $\alpha=\hat{\alpha}$ and $\sigma=\hat{\sigma}$ are reduced to the following simultaneous equations:

$$\frac{1}{\hat{\alpha}}=\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\frac{\hat{\sigma}^2\lambda(p,q)}{1+\hat{\alpha}\hat{\sigma}^2\lambda(p,q)}-\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\left(\frac{\lambda(p,q)(1+\alpha^*\sigma^{*2}\lambda(p,q))}{\alpha^*\lambda(p,q)(1+\hat{\alpha}\hat{\sigma}^2\lambda(p,q))^2}\right) \tag{203}$$

$$\hat{\sigma}^2=\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\left(\frac{\hat{\sigma}^2}{1+\hat{\alpha}\hat{\sigma}^2\lambda(p,q)}\right)-\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\left(\frac{\hat{\alpha}^2\hat{\sigma}^4\lambda(p,q)^2(1+\alpha^*\sigma^{*2}\lambda(p,q))}{\alpha^*\lambda(p,q)(1+\hat{\alpha}\hat{\sigma}^2\lambda(p,q))^2}\right). \tag{204}$$

It is obvious that equations (203) and (204) are satisfied by $\hat{\alpha}=\alpha^*$ and $\hat{\sigma}=\sigma^*$.

4.4. Statistical-mechanical iterative algorithms

Relations derived in the previous subsection are used here to give the algorithm for the estimation of the hyperparameters $(\hat{\sigma},\hat{\alpha})$ and the restored image $\hat{\boldsymbol{f}}$ for a given degraded image $\boldsymbol{g}$. Also, we construct the algorithms to estimate the statistical performance $\mathcal{M}(\sigma^*,\alpha^*)$ and the statistical behaviour in the iteration process for hyperparameter estimations for an original image $\boldsymbol{f}$.

First, we construct an algorithm for obtaining the estimates of hyperparameters, $\hat{\alpha}$ and $\hat{\sigma}$, and the restored image $\hat{\boldsymbol{f}}$ when a degraded image $\boldsymbol{g}$ is given. From the extremum conditions (192) and (193) of the marginal likelihood, $\Pr\{\boldsymbol{G}=\boldsymbol{g}|\alpha,\sigma\}$, and expression (195) of the restored image $\hat{f}_{x,y}$, we can construct the following recursion formulae:

Algorithm 4.1. Statistical-mechanical image restoration algorithm by the Gaussian model

Step 1: Calculate the discrete Fourier transform $G(p,q)$ of the given degraded image $\boldsymbol{g}$ by equation (188). Set $r\leftarrow 0$, $a(0)\leftarrow 1$ and $b(0)\leftarrow 1$ as the initial values.
Step 2: Update $r\leftarrow r+1$, and

$$a(r)\leftarrow\left(\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\frac{b(r-1)\,\lambda\,(p,q)}{1+a(r-1)\,b\,(r-1)\,\lambda\,(p,q)}+\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}|G(p,q)|^2\frac{\lambda(p,q)}{(1+a(r-1)\,b\,(r-1)\,\lambda\,(p,q))^2}\right)^{-1} \tag{205}$$

$$b(r)\leftarrow\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\frac{b(r-1)}{1+a(r-1)\,b\,(r-1)\,\lambda\,(p,q)}+\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}|G(p,q)|^2\frac{a(r-1)^2b(r-1)^2\lambda(p,q)^2}{(1+a(r-1)\,b\,(r-1)\lambda(p,q))^2}. \tag{206}$$

Step 3: Update $\hat{\sigma}\leftarrow\sqrt{b(r)}$, $\hat{\alpha}\leftarrow a(r)$ and $R\leftarrow r$. Go to *step 4* if it is satisfied that

$$\left|\frac{a(r)-a(r-1)}{a(r-1)}\right|+\left|\frac{b(r)-b(r-1)}{b(r-1)}\right|<\varepsilon \tag{207}$$

and go to *step 2* otherwise.

Step 4: For the obtained estimates $\hat{\alpha}$ and $\hat{\sigma}$, the restored image $\hat{\boldsymbol{f}} \equiv \{\hat{f}_{x,y}\}$ is determined by

$$\hat{f}_{x,y} \leftarrow \arg \min_{n=0,1,\ldots,255} \left(n - \int z_{x,y} \Pr\{\boldsymbol{F} = \boldsymbol{z} | \boldsymbol{G} = \boldsymbol{g}, \hat{\alpha}, \hat{\sigma}\} \, \mathrm{d}\boldsymbol{z} \right)^2. \tag{208}$$

In algorithm 4.1, R denotes the final iteration number when $a(r)$ and $b(r)$ converge. The estimation of the restored image $\hat{\boldsymbol{f}}$ in equation (208) is called the TPM estimation [20].

Second, we give an algorithm for estimating the statistical performance with respect to all possible degraded images $\boldsymbol{g}$ generated by obeying the degradation process (173) for a fixed value $\sigma = \sigma^*$ from a given original image $\boldsymbol{f}$. From equations (198) and (200), the statistical quantity $M(\alpha^*, \sigma^*)$ for the given original image $\boldsymbol{f}$ and an additive white Gaussian noise $\mathcal{N}[0, \sigma^{*2}]$ are obtained in the following procedure:

Algorithm 4.2. Statistical-performance estimation algorithm in image restoration by the Gaussian model

Step 1: Calculate the discrete Fourier transform $F(p, q)$ of the given standard image $\boldsymbol{f}$ by equation (197).
Step 2: Determine the value of α^* by

$$\alpha^* \leftarrow \left(\frac{1}{|\Omega|} \sum_{p=0}^{L_x-1} \sum_{q=0}^{L_y-1} \lambda(p, q) + \frac{1}{|\Omega|} \sum_{p=0}^{L_x-1} \sum_{q=0}^{L_y-1} |F(p, q)|^2 \lambda(p, q) \right)^{-1}. \tag{209}$$

Step 3: Calculate the quantity $M(\alpha^*, \sigma^*)$ as follows:

$$M(\alpha^*, \sigma^*) \leftarrow \frac{1}{|\Omega|} \sum_{p=0}^{L_x-1} \sum_{q=0}^{L_y-1} \frac{\sigma^{*2}}{1 + \alpha^* \sigma^{*2} \lambda(p, q)}. \tag{210}$$

Finally, we give an algorithm for estimating how the hyperparameters α and σ converge to the true values α^* and σ^* statistically with respect to all possible degraded images $\boldsymbol{g}$ generated by obeying the degradation process (173) from the given original image $\boldsymbol{f}$. From the extremum conditions (203) and (204) of the statistical average of the logarithm of marginal likelihood, $\mathcal{L}(\alpha, \sigma)$, we can construct an algorithm, which gives us a statistical estimation of how $a(r)$ and $b(r)$ converge to the estimates $\hat{\alpha}$ and $\hat{\sigma}^2$ in the iteration process of equations (205) and (206) in the maximization of marginal likelihood for a given original image $\boldsymbol{f}$, as follows:

Algorithm 4.3. Statistical behaviour estimation algorithm in hyperparameter determination

Step 1: Calculate the discrete Fourier transform $F(p, q)$ of the given standard image $\boldsymbol{f}$ by means of equation (197).
Step 2: Determine the values of α^* by

$$\alpha^* \leftarrow \left(\frac{1}{|\Omega|} \sum_{p=0}^{L_x-1} \sum_{q=0}^{L_y-1} \lambda(p, q) + \frac{1}{|\Omega|} \sum_{p=0}^{L_x-1} \sum_{q=0}^{L_y-1} |F(p, q)|^2 \lambda(p, q) \right)^{-1}. \tag{211}$$

Step 3: Set $r \leftarrow 0$, $c(0) \leftarrow 1$ and $d(0) \leftarrow 1$.
Step 4: Update $r \leftarrow r + 1$, and

$$c(r) \leftarrow \left(\frac{1}{|\Omega|} \sum_{p=0}^{L_x-1} \sum_{q=0}^{L_y-1} \frac{d(r-1)\,\lambda\,(p, q)}{1 + c(r-1) d(r-1)\,\lambda\,(p, q)} - \frac{1}{|\Omega|} \sum_{p=0}^{L_x-1} \sum_{q=0}^{L_y-1} \left(\frac{\lambda(p, q)}{\alpha^* \lambda(p, q)} \right) \left(\frac{(1 + \alpha^* \sigma^{*2} \lambda(p, q))}{(1 + c(r-1) d(r-1)\,\lambda\,(p, q))^2} \right) \right)^{-1} \tag{212}$$

Figure 19. Standard images $\boldsymbol{f}^*$. (*a*) 'Home'; (*b*) 'mandrill'.

$$d(r) \leftarrow \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\left(\frac{d(r-1)}{1+c(r-1)d(r-1)\,\lambda\,(p,q)}\right) - \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\left(\frac{c(r-1)^2 d(r-1)^2\lambda(p,q)^2}{\alpha^*\lambda(p,q)}\right) \times\left(\frac{(1+\alpha^*\sigma^{*2}\lambda(p,q))}{(1+c(r-1)d(r-1)\,\lambda\,(p,q))^2}\right). \tag{213}$$

Step 5: Update $\hat{\sigma} \leftarrow \sqrt{d(r)}$, $\hat{\alpha} \leftarrow c(r)$ and $R \leftarrow r$. Stop if it is satisfied that

$$\left|\frac{c(r)-c(r-1)}{c(r-1)}\right| + \left|\frac{d(r)-d(r-1)}{d(r-1)}\right| < \varepsilon \tag{214}$$

and go to *step 4* otherwise.

We remark that, in the above algorithm, $c(r)$ and $d(r)$ correspond to statistical averages of $a(r)$ and $b(r)$ with respect to random fields $\boldsymbol{F}$ and $\boldsymbol{G}$ in the probability $\Pr\{\boldsymbol{F}, \boldsymbol{G}|\alpha^*, \sigma^*\} \equiv \Pr\{\boldsymbol{G}|\boldsymbol{F}, \sigma^*\}\Pr\{\boldsymbol{F}|\alpha^*\}$.

In algorithms 4.2 and 4.3, we give a little more detailed explanation. In the algorithms, the value of α^*, which is maximizing the likelihood $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha\}$ in equations (196) and (197) for a given original image $\boldsymbol{f}$, is regarded as a true value of the hyperparameter α for the given original image $\boldsymbol{f}$ although the given original image $\boldsymbol{f}$ is not generated by obeying the *a priori* probability density function $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha^*\}$ in equation (174). In algorithm 4.3, the statistical behaviour of the convergence process of the hyperparameters α and σ is given as a series $\{c(r), d(r)|r = 1, 2, 3, \ldots, R\}$.

4.5. Numerical experiments

In this subsection, we give some numerical experiments of image restoration by the maximum marginal likelihood estimation and the Gaussian model for the standard images given in figure 19.

We have followed algorithm 4.1 given in subsection 4.4. The degraded images $\boldsymbol{g}^*$ are generated from the original images $\boldsymbol{f}^*$ in figure 19 by setting $\sigma^* = 10, 20, 30, 40$ and

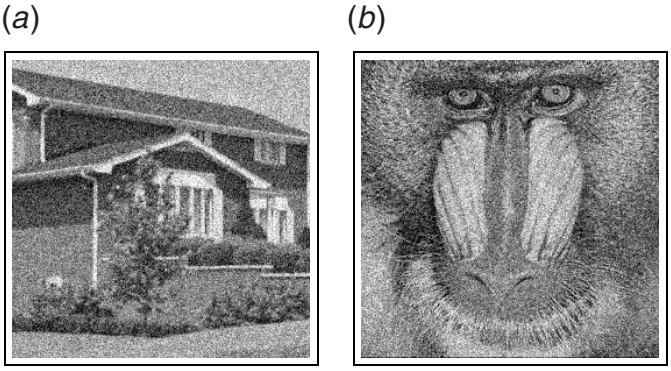

Figure 20. Degraded images $\boldsymbol{g}^*$ for $\sigma^* = 30$. (*a*) 'Home'; (*b*) 'mandrill'.

Table 4. Values of $\hat{\sigma}$, $\hat{\alpha}$, $d(\boldsymbol{f}^*, \boldsymbol{g}^*)$, $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} (dB) for the original image $\boldsymbol{f}^*$ given in figure 19(*a*).

σ^*	$\hat{\sigma}$	$\hat{\alpha}$	$d(\boldsymbol{f}^*, \boldsymbol{g}^*)$	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$	Δ_{SNR} (dB)
10	0.332	0.004 70	92.44	92.44	0
20	9.753	0.002 46	382.43	260.75	1.663 28
30	22.352	0.002 13	860.22	280.98	4.859 34
40	32.965	0.001 99	1495.63	304.54	6.911 80
50	42.508	0.001 94	2247.05	333.48	8.285 43

Table 5. Values of $\hat{\sigma}$, $\hat{\alpha}$, $d(\boldsymbol{f}^*, \boldsymbol{g}^*)$, $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} (dB) for the original image $\boldsymbol{f}^*$ given in figure 19(*b*).

σ^*	$\hat{\sigma}$	$\hat{\alpha}$	$d(\boldsymbol{f}^*, \boldsymbol{g}^*)$	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$	Δ_{SNR} (dB)
10	0.813	0.002 92	92.44	92.44	0
20	15.996	0.002 62	383.60	190.81	3.032 77
30	27.803	0.002 83	867.74	255.82	5.304 55
40	37.970	0.002 96	1521.80	314.36	6.849 30
50	47.149	0.003 08	2299.78	363.05	8.017 20

50 for the additive white Gaussian noise $\mathcal{N}[0, \sigma^{*2}]$ given in equation (173). By applying algorithm 4.1 to each degraded image $\boldsymbol{g}^*$, the corresponding restored image $\hat{\boldsymbol{f}}$ is obtained. For the restored images $\hat{\boldsymbol{f}}$, the estimates of hyperparameters $\hat{\alpha}$ and $\hat{\sigma}$ and the values of $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ in equation (163) and Δ_{SNR} in equation (164) are given in tables 4 and 5. For the original images in figure 19, the degraded images $\boldsymbol{g}^*$ and the corresponding restored images $\hat{\boldsymbol{f}}$ in $\sigma^* = 30$ are given in figures 20 and 21, respectively. The results show that some noise still remains in the restored image of 'home' but the restored image of 'mandrill' is satisfactory.

We next explain the statistical performance estimation for the original images given in figure 19. In the Gaussian model, the value of α which maximizes the *a priori* probability density function (174) is adopted as the value of α^*. In each case, the standard images 'home' and 'mandrill' given in figure 19 are set as the original image $\boldsymbol{f}^*$ in the *a priori* probability density function (174). The values of α^* and the statistical quantities $\mathcal{M}(\sigma^*, \alpha^*)$ obtained

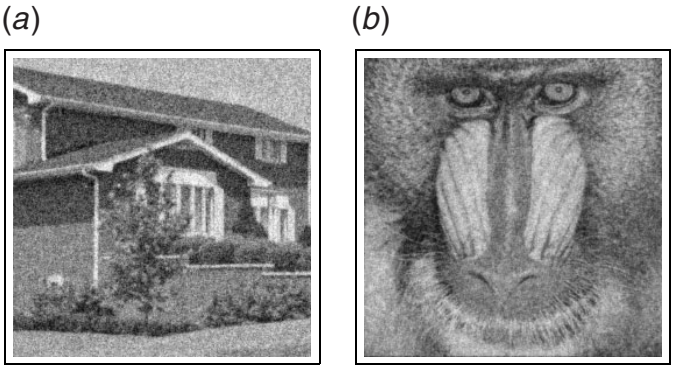

Figure 21. Restored images $\hat{\boldsymbol{f}}$ obtained by the Gaussian model for $\sigma^* = 30$. (*a*) 'Home'; (*b*) 'mandrill'.

Table 6. Values of $\mathcal{M}(\sigma^*, \alpha^*)$ when the original image $\boldsymbol{f}^*$ is set to the image given in figure 19(*a*).

σ^*	α^*	$\mathcal{M}(\sigma^*, \alpha^*)$
10	0.008 24	58.00
20	0.008 24	114.84
30	0.008 24	150.24
40	0.008 24	174.81
50	0.008 24	193.44

Table 7. Values of $\mathcal{M}(\sigma^*, \alpha^*)$ when the original image $\boldsymbol{f}^*$ is set to the image given in figure 19(*b*).

σ^*	α^*	$\mathcal{M}(\sigma^*, \alpha^*)$
10	0.004 01	72.90
20	0.004 01	172.39
30	0.004 01	244.03
40	0.004 01	295.67
50	0.004 01	335.14

from algorithm 4.2 for various values of σ^* are given in tables 6 and 7, although these standard images are not generated from the *a priori* probability density function. By using algorithm 4.3, we obtain the statistical behaviour of the iteration processes (205) and (206) in the algorithms of the maximization of evidence for standard images 'home' and 'mandrill'. The dynamical behaviour is given in figures 22 and 23. These values for statistical performance in 'mandrill' are close to the corresponding numerical experiments given in table 5. On the other hand, the values for statistical performance in 'home' are not so close to the corresponding numerical experiments given in table 4. These results show that the standard image 'mandrill' satisfies the assumption that the *a priori* probability density function is given in equation (174), while the standard image 'home' does not. These facts can also be understood from comparison of the values of α^* in tables 4 and 5 with those of $\hat{\alpha}$ in tables 6 and 7.

4.6. Constrained least mean square filter

As mentioned in the beginning of the present section, we have the constrained least mean square filter as one of the conventional linear filters in image and signal processing. When physicists try to investigate image processing from the standpoint of statistical mechanics, it is

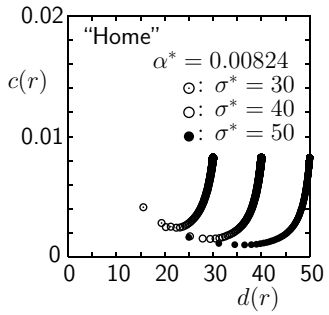

Figure 22. The statistical behaviour of the iterative process in equations (205) and (206), which is obtained from the recursion formulae (212) and (213) for the image given in figure 19(*a*).

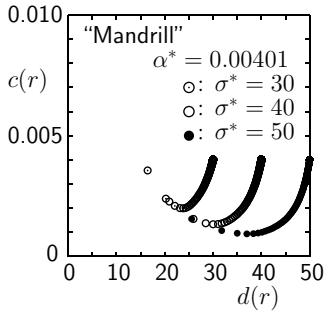

Figure 23. The statistical behaviour of the iterative process in equations (205) and (206), which is obtained from the recursion formulae (212) and (213) for the image given in figure 19(*b*).

important to understand the conventional filters. We cannot review all the conventional filters due to limitations of space in the present paper. The present author hopes that the readers see some textbooks on conventional image and signal processing [1–3]. In this subsection, we will explain a little more detailed formulation and give some numerical experiments in the constrained least mean square filter because the filter has a close relationship to the Gaussian model.

The energy functions $H(\boldsymbol{f}|\boldsymbol{g}, \gamma)$ in equation (172) and $E(\boldsymbol{f}|\boldsymbol{g}, \sigma, \alpha)$ in equation (178) are equivalent to each other except for the constant factor $\frac{1}{2}\alpha$. The present calculation by means of the discrete Fourier transform can also be applied to the constrained least mean square filter (169)–(172). Now we compare the evidence framework in the Gaussian model with the constrained least mean square filter. By a similar argument to the derivation of equation (195), the constrained least mean square filter (169)–(172) can be reduced to the following form,

$$\hat{f}_{x,y} = \frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}\frac{1}{1+\gamma\lambda(p,q)}\left(\cos\left(\frac{2\pi px}{L_x}+\frac{2\pi qy}{L_y}\right)\mathrm{Re}\,G(p,q) + \sin\left(\frac{2\pi px}{L_x}+\frac{2\pi qy}{L_y}\right)\mathrm{Im}\,G(p,q)\right) \tag{215}$$

where the Lagrange multiplier γ is determined so as to satisfy the constraint $\|\hat{\boldsymbol{z}}(\gamma) - \boldsymbol{g}\|^2 = |\Omega|\sigma^2$, which is reduced to

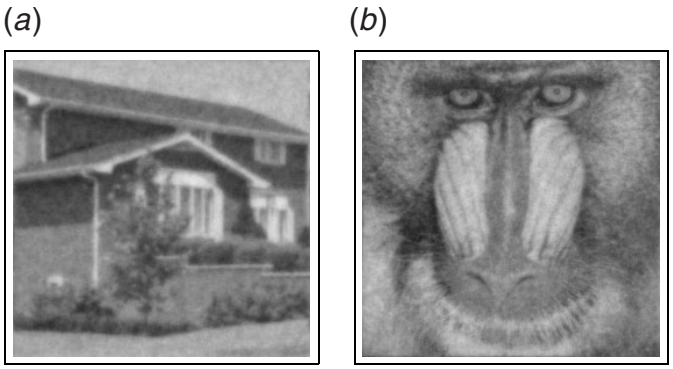

Figure 24. Restored images $\hat{\boldsymbol{f}}$ obtained by means of the constrained least mean square filter for $\sigma^* = 30$. (*a*) 'Home' (*b*) 'mandrill'.

Table 8. Values of $\hat{\gamma}$, $d(\boldsymbol{f}^*, \boldsymbol{g}^*)$, $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} (dB) obtained by using the constrained least mean square filter (169)–(172) for the original image $\boldsymbol{f}^*$ given in figure 19(*a*).

σ^*	$\hat{\gamma}$	$d(\boldsymbol{f}^*, \boldsymbol{g})$	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$	Δ_{SNR} (dB)
10	2.447 00	92.44	72.53	1.053 06
20	6.516 85	382.43	162.94	3.705 08
30	12.483 92	860.22	251.13	5.347 10
40	23.087 83	1495.63	355.02	6.245 74
50	49.954 38	2247.05	512.03	6.423 21

Table 9. Values of $\hat{\gamma}$, $d(\boldsymbol{f}^*, \boldsymbol{g})$, $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} (dB) by using the constrained least mean square filter (169)–(172) for the original image $\boldsymbol{f}^*$ given in figure 19(*b*).

σ^*	$\hat{\gamma}$	$d(\boldsymbol{f}^*, \boldsymbol{g}^*)$	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$	Δ_{SNR} (dB)
10	1.292 39	92.44	95.56	−0.143 96
20	3.985 65	383.59	209.79	2.620 94
30	8.465 61	867.74	294.77	4.689 09
40	16.445 90	1521.80	366.23	6.186 06
50	37.675 63	2299.78	450.14	7.083 36

$$\hat{\gamma} = \frac{\sigma^{*2}}{\frac{1}{|\Omega|}\sum_{p=0}^{L_x-1}\sum_{q=0}^{L_y-1}|G(p,q)|^2\frac{\lambda(p,q)^2}{(1+\gamma\lambda(p,q))^2}}. \tag{216}$$

We show in tables 8 and 9 the values of $\hat{\gamma}$, $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} for the restored images $\hat{\boldsymbol{f}}$ obtained by applying the constrained least mean square filter (215), (216) to the degraded images $\boldsymbol{g}^*$ which are generated from the original images $\boldsymbol{f}^*$ given in figure 19 for various values of σ^*. For the degraded images $\boldsymbol{g}^*$ in figure 20, the restored images $\hat{\boldsymbol{f}}$ obtained by using the constrained least mean square filter (215), (216) are shown in figure 24. Although we use the value of σ^* in the image restoration, the quality of the restored image $\hat{\boldsymbol{f}}$ is seen not to be sufficient and too blurred.

4.7. Concluding remarks

In this section, we have explained grey-level image restoration by the Gaussian model. The fact that the Gaussian model is applicable to the Bayesian approach to image restoration was pointed out by Nishimori [85]. The Gaussian model can be extended from various points of view. In the present subsection, we summarize some interesting extensions based on the Gaussian model.

From the standpoint of image processing, the Gaussian model is equivalent to a method called the conditional autoregressive model in which

$$f_{x,y} - f_{x+1,y} \sim \mathcal{N}\left[0, \frac{1}{\alpha}\right] \qquad f_{x,y} - f_{x,y+1} \sim \mathcal{N}\left[0, \frac{1}{\alpha}\right]. \tag{217}$$

Molina *et al* [83, 84] investigated the determination of hyperparameters in a simultaneous autoregressive model given in the form

$$f_{x,y} - \frac{1}{4}(f_{x+1,y} + f_{x-1,y} + f_{x,y+1} + f_{x,y-1}) \sim \mathcal{N}\left[0, \frac{1}{\alpha}\right]. \tag{218}$$

The model can be treated analytically by analogous arguments to the present section. The present author and Inoue extended these autoregressive models to a generalized version of the grey-level image restoration scheme by means of solvable random field models and investigated the hyperparameter determination in the maximization of evidence by calculating the exact closed expression of evidence using the Gaussian integral formula and discrete Fourier transform [86]. They proposed the following *a priori* probability density function:

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}\} = \Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha, \nu\} \equiv \frac{\exp\left(-\frac{1}{2}\alpha \boldsymbol{f}\boldsymbol{C}^{\nu}\boldsymbol{f}^{\mathrm{T}}\right)}{\int_{\mathbb{R}^{|\Omega|}} \exp\left(-\frac{1}{2}\alpha \boldsymbol{f}\boldsymbol{C}^{\nu}\boldsymbol{f}^{\mathrm{T}}\right) \mathrm{d}\boldsymbol{z}} \tag{219}$$

instead of equation (174). In this model, not only α but also ν are hyperparameters. The hyperparameters α, ν and σ are determined by

$$(\hat{\alpha}, \hat{\nu}, \hat{\sigma}) = \arg \max_{(\alpha,\nu,\sigma)} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \nu, \sigma\} \tag{220}$$

where

$$\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \sigma\} \equiv \int_{\mathbb{R}^{|\Omega|}} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\boldsymbol{F} = \boldsymbol{f}, \sigma\} \Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha, \nu\} \mathrm{d}\boldsymbol{z}. \tag{221}$$

We remark that the case of $\nu = 1$ corresponds to that in equations (217) and (219), and that the case of $\nu = 2$ corresponds to that in equation (218). They insisted that the picture quality of the restored image can be improved by introducing the new hyperparameter ν which takes a real number. The framework of their reference has been extended to the solvable probabilistic model in colour image restoration [87].

In the above statements, we have treated the *a priori* probability density function with only the spatially uniform hyperparameters α. Some physicists, particularly the researchers of spin glass theory, may be interested in the one with spatially non-uniform hyperparameters $\boldsymbol{\alpha} \equiv \left\{\alpha_{x,y}^{x+1,y}, \alpha_{x,y}^{x,y+1} | (x, y) \in \Omega\right\}$. Many grey-level images include not only smooth areas but also edges. If the smoothness is regarded as important, the edge areas are also smoothed out. If the edges are preserved, the noise cannot be erased sufficiently. In order to resolve this conflict, we have to introduce a strategy to change the value of the hyperparameter α in each area. One idea is to introduce adaptive smoothing. In order to achieve this goal, an inhomogeneous probabilistic model is introduced:

$$\Pr\{\boldsymbol{F}=\boldsymbol{f}\}=\Pr\{\boldsymbol{F}=\boldsymbol{f}|\boldsymbol{\alpha}\}$$
$$\equiv \frac{\exp\left(-\frac{1}{2}\sum_{(x,y)\in\Omega}\left(\alpha_{x,y}^{x+1,y}(f_{x,y}-f_{x+1,y})^2+\alpha_{x,y}^{x,y+1}(f_{x,y}-f_{x,y+1})^2\right)\right)}{\int_{\mathbb{R}^{|\Omega|}}\exp\left(-\frac{1}{2}\sum_{(x,y)\in\Omega}\left(\alpha_{x,y}^{x+1,y}(z_{x,y}-z_{x+1,y})^2+\alpha_{x,y}^{x,y+1}(z_{x,y}-z_{x,y+1})^2\right)\right)\mathrm{d}\boldsymbol{z}} \tag{222}$$

instead of equation (219). Molina *et al* [84], Aykroyd [89] and Dunmur and Titterington [88] proposed that the spatial variations of $\alpha_{x,y}^{x+1,y}$ and $\alpha_{x,y}^{x,y+1}$ are also regarded as random variables and introduced probability density functions $\mathcal{P}\left(\alpha_{x,y}^{x+1,y}\right)$ and $\mathcal{P}\left(\alpha_{x,y}^{x,y+1}\right)$ for $\alpha_{x,y}^{x+1,y}$ and $\alpha_{x,y}^{x,y+1}$, which are referred to as *hyperpriors*. The joint probability density function $\Pr\{\boldsymbol{F}=\boldsymbol{f},\boldsymbol{G}=\boldsymbol{g},\boldsymbol{\alpha}|\sigma\}$ and the *a posteriori* probability density function $\Pr\{\boldsymbol{F}=\boldsymbol{f},\boldsymbol{\alpha}|\boldsymbol{G}=\boldsymbol{g},\sigma\}$ are defined by

$$\Pr\{\boldsymbol{F}=\boldsymbol{f},\boldsymbol{G}=\boldsymbol{g},\boldsymbol{\alpha}|\sigma\}$$
$$\equiv \Pr\{\boldsymbol{G}=\boldsymbol{g}|\boldsymbol{F}=\boldsymbol{f},\sigma\}\Pr\{\boldsymbol{F}=\boldsymbol{f}|\boldsymbol{\alpha}\}\left(\prod_{(x,y)\in\Omega}\mathcal{P}\left(\alpha_{x,y}^{x+1,y}\right)\mathcal{P}\left(\alpha_{x,y}^{x,y+1}\right)\right) \tag{223}$$

and

$$\Pr\{\boldsymbol{F}=\boldsymbol{f},\boldsymbol{\alpha}|\boldsymbol{G}=\boldsymbol{g},\sigma\}\equiv\frac{\Pr\{\boldsymbol{F}=\boldsymbol{f},\boldsymbol{G}=\boldsymbol{g},\boldsymbol{\alpha}|\sigma\}}{\int\Pr\{\boldsymbol{F}=\boldsymbol{f},\boldsymbol{G}=\boldsymbol{g},\boldsymbol{\alpha}|\sigma\}\,\mathrm{d}\boldsymbol{f}\,\mathrm{d}\boldsymbol{\alpha}} \tag{224}$$

respectively. The hyperparameters $\boldsymbol{\alpha}$ and σ and the restored image $\hat{\boldsymbol{f}}$ are determined from these probability density functions. A stream of these investigations in the *hyperprior* have been dealt with mainly by statisticians [90]. Currently, the mainstream approach to achieve adaptive smoothing in practical Bayesian image restoration is to introduce a line field and the investigations of the framework based on the *hyperprior* are restricted only to statistical interests. However, it is obvious that this framework is close to spin glass theory and the author believes that many researchers in the field of spin glasses will begin to investigate it.

5. Coupled Markov random field model

In section 4, we elucidated the probabilistic computational method when the *a priori* probability density function is assumed to be given as the Gaussian model in equation (174). The idea behind this *a priori* probability is smoothing of neighbouring pixel values. If the hyperparameters are adjusted to stress spatial smoothness, the restored images will be blurred too much. If, on the other hand, the hyperparameters are adjusted to take account mostly of the observed data $\boldsymbol{g}$, many isolated noises remain. Particularly, in the grey-level image restoration, this is the limit of the simple Gaussian model as an *a priori* probability density function, and it is difficult to treat images with many edges. In order to improve this strong smoothing effect of the Gaussian model, Geman and Geman [6] and Jeng and Woods [8] introduced a method with an edge state at each nearest-neighbour pair of pixels.

In a system consisting of two pixels at (1, 1) and (1, 2), we consider the following *a priori* probability distribution with an edge state l:

$$\Pr\{F_{1,1}=f_{1,1},F_{1,2}=f_{1,2},L_{1,1}^{1,2}=l|\alpha,\gamma\}$$
$$=\frac{\sum_{l=0}^{1}\exp\left(-\frac{1}{2}\alpha(1-l)((f_{1,1}-f_{1,2})^2-\gamma^2)\right)}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sum_{l=0}^{1}\exp\left(-\frac{1}{2}\alpha(1-l)((z_{1,1}-z_{1,2})^2-\gamma^2)\right)\mathrm{d}z_{1,1}\,\mathrm{d}z_{1,2}} \tag{225}$$

instead of equation (174). If $|f_{1,1}-f_{1,2}|$ is larger than γ, the edge state l should favour existence ($l=1$). If $|f_{1,1}-f_{1,2}|$ is smaller than γ on the other hand, it should favour the 'no

edge state' ($l = 0$). In order to clarify the relationship between the classical spin system and the probabilistic model for image restoration, we introduce a temperature $T\,(>0)$ and consider the partial summation with respect to the edge state l as follows,

$$\begin{aligned}-T\ln\left\{\sum_{l=0}^{1}\exp\left(-\frac{1}{2T}\alpha(1-l)((f_{x,y}-f_{x',y'})^2-\gamma^2)\right)\right\}\\ =\ -T\ln\left\{1+\exp\left(-\frac{1}{2T}\alpha((f_{x,y}-f_{x',y'})^2-\gamma^2)\right)\right\}\\ \to\ \psi(f_{x,y}-f_{x',y'})\qquad T\to+0\end{aligned}\tag{226}$$

where

$$\psi(r)\equiv\begin{cases}\alpha(r^2-\gamma^2) & |r|\leqslant\gamma\\ 0 & |r|\geqslant\gamma.\end{cases}\tag{227}$$

This equality means that the hyperparameter γ corresponds to a control parameter for the boundary between smoothness and flatness.

In this section, we introduce a Markov random field model which has an edge state at each nearest-neighbour pair of pixels. The model with an edge state is called a coupled Markov random field model or a compound Gauss–Markov random field model. We summarize a conventional coupled Markov random field model and extend it to a quantized one.

5.1. Intensity field and line field

We extend the framework of grey-level image restoration in the previous section to a coupled Markov random field model with an edge state. The coupled Markov random field model introduced in the present subsection is the most basic one. A set of random variables for edge states is called a line field. The random field $\boldsymbol{F}$ for intensity at each pixel is called the intensity field. We introduce an edge state at each nearest-neighbour pair of pixels. All possible edge states are the 'no edge' state and the 'edge' state. The random variable for an edge at the nearest-neighbour pairs of pixels, (x, y) and $(x + 1, y)$, is denoted by $L_{x,y}^{x+1,y}$ and the one at the pairs of pixels, (x, y) and $(x, y + 1)$ by $L_{x,y}^{x,y+1}$. We assign 1 and 0 to the 'edge' state and the 'no edge' state, respectively. The random field for an edge is denoted by $\boldsymbol{L}\equiv\left\{L_{x,y}^{x+1,y},L_{x,y}^{x,y+1}|(x,y)\in\Omega\right\}$. The random field $\boldsymbol{L}$ is a line field. In this section, we consider the additive white Gaussian model (173) as a degradation process. The *a priori* probability density function, the original image being $\boldsymbol{f}$, is assumed to be,

$$\Pr\{\boldsymbol{F}=\boldsymbol{f}\}=\Pr\{\boldsymbol{F}=\boldsymbol{f}|\alpha,\gamma\}\equiv\frac{\sum_{l}\exp\left(-U(\boldsymbol{f},\boldsymbol{l}|\alpha,\gamma)\right)}{\int_{\mathbb{R}^{|\Omega|}}\sum_{l}\exp\left(-U(\boldsymbol{z},\boldsymbol{l}|\alpha,\gamma)\right)\mathrm{d}\boldsymbol{z}}\tag{228}$$

where

$$\begin{aligned}U(\boldsymbol{f},\boldsymbol{l}|\alpha,\gamma)\equiv\frac{1}{2}\alpha\sum_{(x,y)\in\Omega}\left(\left(1-l_{x,y}^{x+1,y}\right)((f_{x,y}-f_{x+1,y})^2-\gamma^2)\right.\\ \left.+\left(1-l_{x,y}^{x,y+1}\right)((f_{x,y}-f_{x,y+1})^2-\gamma^2)\right).\end{aligned}\tag{229}$$

By substituting equations (173) and (228) into equation (7), the *a posteriori* probability density function is given as

$$\Pr\{\boldsymbol{F}=\boldsymbol{f}|\boldsymbol{G}=\boldsymbol{g}\}=\Pr\{\boldsymbol{F}=\boldsymbol{f}|\boldsymbol{G}=\boldsymbol{g},\alpha,\gamma,\sigma\}=\sum_{\boldsymbol{l}}\rho(\boldsymbol{f},\boldsymbol{l}|\boldsymbol{g},\alpha,\gamma,\sigma)\tag{230}$$

where

$$\rho(\boldsymbol{f},\boldsymbol{l})=\frac{\exp(-E(\boldsymbol{f},\boldsymbol{l}|\boldsymbol{g},\alpha,\gamma,\sigma))}{\int_{\mathbb{R}^{|\Omega|}}\sum_{\boldsymbol{l}}\exp(-E(\boldsymbol{z},\boldsymbol{l}|\boldsymbol{g},\alpha,\gamma,\sigma))\,\mathrm{d}\boldsymbol{z}}\tag{231}$$

$$E(\boldsymbol{f},\boldsymbol{l}|\boldsymbol{g},\alpha,\gamma,\sigma) \equiv \frac{1}{2\sigma^2}\sum_{(x,y)\in\Omega}(f_{x,y}-g_{x,y})^2+\frac{1}{2}\alpha\sum_{(x,y)\in\Omega}\left(\left(1-l_{x,y}^{x+1,y}\right)\right.$$
$$\left.\times\left((f_{x,y}-f_{x+1,y})^2-\gamma^2\right)+\left(1-l_{x,y}^{x,y+1}\right)\left((f_{x,y}-f_{x,y+1})^2-\gamma^2\right)\right). \tag{232}$$

The Gibbs canonical distribution $\rho(\boldsymbol{f},\boldsymbol{l})$ satisfies the variational principle of minimization of the free energy functional $\mathcal{F}[\rho]$,

$$\mathcal{F}[\rho] \equiv \int_{\mathbb{R}^{|\Omega|}}\sum_{\boldsymbol{l}}(E(\boldsymbol{f},\boldsymbol{l}|\boldsymbol{g},\alpha,\gamma,\sigma)+\ln(\rho(\boldsymbol{f},\boldsymbol{l})))\rho(\boldsymbol{f},\boldsymbol{l})\,\mathrm{d}\boldsymbol{f} \tag{233}$$

under the normalization condition $\int_{\mathbb{R}^{|\Omega|}}\sum_{\boldsymbol{l}}\rho(\boldsymbol{f},\boldsymbol{l})\,\mathrm{d}\boldsymbol{f}=1$.

To facilitate analytical treatments, we introduce the marginal probability density functions $\rho_{x,y}(\zeta)$, $\rho_{x,y}^{x+1,y}(l)$ and $\rho_{x,y}^{x,y+1}(l)$ defined by

$$\rho_{x,y}(\zeta) \equiv \int_{\mathbb{R}^{|\Omega|}}\sum_{\boldsymbol{l}}\rho(\boldsymbol{f},\boldsymbol{l})\,\delta(\zeta-f_{x,y})\,\mathrm{d}\boldsymbol{f} \qquad \zeta\in\mathbb{R} \tag{234}$$

$$\rho_{x,y}^{x',y'}(l) \equiv \int_{\mathbb{R}^{|\Omega|}}\sum_{\boldsymbol{l}}\rho(\boldsymbol{f},\boldsymbol{l})\delta_{l,l_{x,y}^{x',y'}}\,\mathrm{d}\boldsymbol{f} \qquad (x',y')\in\mathbf{c}_{x,y} \quad l=0,1 \tag{235}$$

where $\delta(a)$ is the Dirac delta function. In the mean-field approximation, the probability distribution $\rho(\boldsymbol{f},\boldsymbol{l})$ is approximately expressed in terms of the marginal probability distributions:

$$\rho(\boldsymbol{f},\boldsymbol{l}) \simeq \left(\prod_{(x,y)\in\Omega}\rho_{x,y}(f_{x,y})\right)\left(\prod_{(x,y)\in\Omega}\rho_{x,y}^{x+1,y}\left(l_{x,y}^{x+1,y}\right)\right)\left(\prod_{(x,y)\in\Omega}\rho_{x,y}^{x,y+1}\left(l_{x,y}^{x,y+1}\right)\right). \tag{236}$$

By substituting equation (236) into equation (233), the free energy $\mathcal{F}[\rho]$ can be expressed as $\mathcal{F}\left[\left\{\rho_{x,y},\rho_{x,y}^{x+1,y},\rho_{x,y}^{x,y+1}\right\}\right]$ only in terms of marginal probability density functions $\rho_{x,y}(f_{x,y})$, $\rho_{x,y}^{x+1,y}\left(l_{x,y}^{x+1,y}\right)$ and $\rho_{x,y}^{x,y+1}\left(l_{x,y}^{x,y+1}\right)$. By taking the first variation of the approximate free energy $\mathcal{F}\left[\left\{\rho_{x,y},\rho_{x,y}^{x+1,y},\rho_{x,y}^{x,y+1}\right\}\right]$ with respect to $\rho_{x,y}(f_{x,y})$, $\rho_{x,y}^{x+1,y}\left(l_{x,y}^{x+1,y}\right)$ and $\rho_{x,y}^{x,y+1}\left(l_{x,y}^{x,y+1}\right)$, we obtain the simultaneous recursion formulae for $\rho_{x,y}(f_{x,y})$ as follows:

$$\rho_{x,y}(f_{x,y}) = \frac{1}{\sqrt{2\pi}\sigma_{x,y}}\exp\left(-\frac{(f_{x,y}-\mu_{x,y})^2}{2{\sigma_{x,y}}^2}\right) \tag{237}$$

$$\mu_{x,y} = \frac{\frac{1}{2\sigma^2}g_{x,y}+\alpha\sum_{(x',y')\in\mathbf{c}_{x,y}}\left(1-\lambda_{x,y}^{x',y'}\right)\mu_{x',y'}}{\frac{1}{2\sigma^2}+\alpha\sum_{(x',y')\in\mathbf{c}_{x,y}}\left(1-\lambda_{x,y}^{x',y'}\right)} \tag{238}$$

$$\frac{1}{2\sigma_{x,y}^2} = \frac{1}{2\sigma^2}+\alpha\sum_{(x',y')\in\mathbf{c}_{x,y}}\left(1-\lambda_{x,y}^{x',y'}\right) \tag{239}$$

$$\lambda_{x,y}^{x',y'} = \lambda_{x',y'}^{x,y} = \frac{1}{1+\exp\left[-\alpha\left\{\left(\sigma_{x,y}^2+\sigma_{x',y'}^2+(\mu_{x,y}-\mu_{x',y'})^2\right)-\gamma^2\right\}\right]}$$
$$(x',y')\in\mathbf{c}_{x,y}. \tag{240}$$

In these equations, $\mu_{x,y}$, $\sigma_{x,y}$ and $\lambda_{x,y}^{x',y'}$ mean the averages of $F_{x,y}$, $(F_{x,y}-\mu_{x,y})^2$ and $L_{x,y}$:

$$\mu_{x,y} \equiv \int\sum_{\boldsymbol{l}}f_{x,y}\,\rho(\boldsymbol{f},\boldsymbol{l})\,\mathrm{d}\boldsymbol{f} = \int f_{x,y}\,\rho_{x,y}(f_{x,y})\,\mathrm{d}f_{x,y} \tag{241}$$

$$\sigma_{x,y}^2 \equiv \int_{\mathbb{R}} \sum_{\boldsymbol{l}} (f_{x,y} - \sigma_{x,y})^2 \, \rho(\boldsymbol{f}, \boldsymbol{l}) \, \mathrm{d}\boldsymbol{f} = \int (f_{x,y} - \mu_{x,y})^2 \, \rho_{x,y}(f_{x,y}) \, \mathrm{d}f_{x,y} \tag{242}$$

$$\lambda_{x,y}^{x',y'} \equiv \int_{\mathbb{R}} \sum_{\boldsymbol{l}} l_{x,y}^{x',y'} \, \rho(\boldsymbol{f}, \boldsymbol{l}) \, \mathrm{d}\boldsymbol{f} = \sum_{l_{x,y}^{x',y'}=0,1} l_{x,y}^{x',y'} \, \rho_{x,y}^{x',y'} \left(l_{x,y}^{x',y'} \right). \tag{243}$$

By using the TPM estimation (10), the restored image $\hat{\boldsymbol{f}}$ is obtained by

$$\hat{f}_{x,y} = \arg\min_{\zeta \in \mathbb{R}} (\zeta - \mu_{x,y})^2. \tag{244}$$

Here, we can determine the hyperparameters σ, α and γ by employing the evidence framework (35). The evidence can be expressed in terms of the energy functions $E(\boldsymbol{f}, \boldsymbol{l}|\boldsymbol{g}, \alpha, \gamma, \sigma)$ and $U(\boldsymbol{f}, \boldsymbol{l}|\alpha, \gamma)$

$$\begin{aligned} \Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \gamma, \sigma\} = & \int_{\mathbb{R}^{|\Omega|}} \sum_{\boldsymbol{l}} \exp\left(-E(\boldsymbol{f}, \boldsymbol{l}|\boldsymbol{g}, \alpha, \gamma, \sigma)\right) \mathrm{d}\boldsymbol{f} \\ & - \int_{\mathbb{R}^{|\Omega|}} \sum_{\boldsymbol{l}} \exp\left(-U(\boldsymbol{f}, \boldsymbol{l}|\alpha, \gamma)\right) \mathrm{d}\boldsymbol{f} - \sqrt{2\pi}\sigma \end{aligned} \tag{245}$$

and can be calculated by applying the mean-field approximation to the *a priori* probability density function and the *a posteriori* probability density function, respectively. The present probabilistic model was first proposed by Geiger and Girosi [16]. They employed a saddle point approximation in order to calculate the restored image although they referred to the saddle point approximation as a mean-field approximation. We have to remark that, in general, the saddle point approximation is different from the mean-field approximation from the standpoint of statistical mechanics. The framework of the conventional mean-field approximation given in the present subsection was given explicitly by Zhang [19].

5.2. Quantized line field

Application of quantum effects to information processing is an interesting subject in the interdisciplinary field of computer science and statistical mechanics. The present author and Horiguchi [91] introduced a quantum Markov random field model which has quantum fluctuations instead of thermal fluctuations as a new type of fluctuation and proposed an iterative algorithm for image restoration from the standpoint of quantum statistical mechanics. The present author proposed a coupled Markov random field model with quantized line fields which can take states expressed as a superposition of edge states and no edge states [92]. It is an interesting problem not only for statistical physicists but also for computer scientists. In this subsection, we summarize the formulation.

A quantized line field is introduced as the following *a priori* probability density function $\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha, \gamma, h\}$:

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}|\alpha, \gamma, h\} = \frac{\mathrm{Tr}[\exp(-\boldsymbol{U}(\boldsymbol{f}|\alpha, \gamma, h))]}{\int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}[\exp(-\boldsymbol{U}(\boldsymbol{f}|\alpha, \gamma, h))] \, \mathrm{d}\boldsymbol{f}} \tag{246}$$

$$\begin{aligned} \boldsymbol{U}(\boldsymbol{f}|\alpha, \gamma, h) \equiv & \sum_{(x,y) \in \Omega} \left(-h\boldsymbol{K}_{x,y}^{x+1,y} + \alpha \left(\boldsymbol{I} - \boldsymbol{L}_{x,y}^{x+1,y} \right) \left((f_{x,y} - f_{x+1,y})^2 - \gamma^2 \right) \right) \\ & + \sum_{(x,y) \in \Omega} \left(-h\boldsymbol{K}_{x,y}^{x,y+1} + \alpha \left(\boldsymbol{I} - \boldsymbol{L}_{x,y}^{x,y+1} \right) \left((f_{x,y} - f_{x,y+1})^2 - \gamma^2 \right) \right) \end{aligned} \tag{247}$$

where

$$\boldsymbol{L}_{x,y}^{x+1,y} \equiv \prod_{(x',y')\in\Omega} \otimes \left((1-\delta_{x,x'}\delta_{y,y'}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \delta_{x,x'}\delta_{y,y'} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right) \tag{248}$$

$$\boldsymbol{L}_{x,y}^{x,y+1} \equiv \prod_{(x',y')\in\Omega} \otimes \left((1-\delta_{x,x'}\delta_{y,y'}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \delta_{x,x'}\delta_{y,y'} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right) \tag{249}$$

$$\boldsymbol{K}_{x,y}^{x+1,y} \equiv \prod_{(x',y')\in\Omega} \otimes \left((1-\delta_{x,x'}\delta_{y,y'}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \delta_{x,x'}\delta_{y,y'} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right) \tag{250}$$

$$\boldsymbol{K}_{x,y}^{x,y+1} \equiv \prod_{(x',y')\in\Omega} \otimes \left((1-\delta_{x,x'}\delta_{y,y'}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \delta_{x,x'}\delta_{y,y'} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \right) \tag{251}$$

$$\boldsymbol{I} \equiv \prod_{(x',y')\in\Omega} \otimes \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{252}$$

Here, $\boldsymbol{A}\otimes\boldsymbol{B}$ means the direct product of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. The function $\exp(\boldsymbol{A})$ for any matrix $\boldsymbol{A}$ is defined by

$$\exp(\boldsymbol{A}) \equiv \sum_{n=0}^{+\infty} \frac{1}{n!}\boldsymbol{A}^n. \tag{253}$$

If we set $h = 0$, all of the nondiagonal elements of the energy matrix $\boldsymbol{U}(\boldsymbol{f}|\alpha,\gamma,h)$ are zero and its eigenvalue is $U(\boldsymbol{f},\boldsymbol{l}|\alpha,\gamma)$ given in equation (229), and the corresponding eigenvector can be given by the following equation using $2^{2|\Omega|}$-dimensional vectors

$$|\boldsymbol{l}\rangle \equiv \prod_{(x',y')\in\Omega} \otimes \left(\left|l_{x,y}^{x+1,y}\right\rangle \otimes \left|l_{x,y}^{x,y+1}\right\rangle \right) \tag{254}$$

where

$$|1\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |0\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{255}$$

The matrices $\boldsymbol{L}_{x,y}^{x+1,y}$ and $\boldsymbol{L}_{x,y}^{x,y+1}$ can be expressed in terms of the Pauli spin matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and 2×2 unit matrix. The matrices $\boldsymbol{K}_{x,y}^{x+1,y}$ and $\boldsymbol{K}_{x,y}^{x,y+1}$ can be expressed in terms of the Pauli spin matrix $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and 2×2 unit matrix. By substituting equations (173) and (246) into equation (7), the *a posteriori* probability distribution is derived as follows:

$$\Pr\{\boldsymbol{F} = \boldsymbol{f}|\boldsymbol{G} = \boldsymbol{g},\alpha,\gamma,\sigma\} = \mathrm{Tr}(\boldsymbol{\rho}(\boldsymbol{f})) \tag{256}$$

$$\boldsymbol{\rho}(\boldsymbol{f}) \equiv \frac{\exp(-\boldsymbol{E}(\boldsymbol{f}|\boldsymbol{g},\alpha,\gamma,\sigma))}{\int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}[\exp(-\boldsymbol{E}(\boldsymbol{f}|\boldsymbol{g},\alpha,\gamma,\sigma))]\,\mathrm{d}\boldsymbol{f}} \tag{257}$$

$$\boldsymbol{E}(\boldsymbol{f}|\boldsymbol{g},\alpha,\gamma,\sigma) \equiv \sum_{(x,y)\in\Omega} \frac{(f_{x,y}-g_{x,y})^2}{2\sigma^2}\boldsymbol{I} + \boldsymbol{U}(\boldsymbol{f}|\alpha,\gamma,h). \tag{258}$$

The Gibbs canonical distribution $\boldsymbol{\rho}(\boldsymbol{f})$ satisfies the variational principle of minimization of the free energy functional $\mathcal{F}[\boldsymbol{\rho}]$,

$$\mathcal{F}[\boldsymbol{\rho}] \equiv \int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}\{(\boldsymbol{E}(\boldsymbol{f}|\boldsymbol{g},\alpha,\gamma,\sigma) + \ln(\boldsymbol{\rho}(\boldsymbol{f})))\boldsymbol{\rho}(\boldsymbol{f})\}\,\mathrm{d}\boldsymbol{f} \tag{259}$$

under the normalization condition $\int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}(\boldsymbol{\rho}(\boldsymbol{f}))\,\mathrm{d}\boldsymbol{f} = 1$. We introduce the marginal probability density function $\rho_{x,y}(\zeta)$ and the marginal probability density matrices $\boldsymbol{\rho}_{x,y}^{x+1,y}$ and $\boldsymbol{\rho}_{x,y}^{x,y+1}$ defined by

$$\rho_{x,y}(\zeta) \equiv \int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}(\boldsymbol{\rho}(\boldsymbol{f}))\delta(\zeta - f_{x,y})\,\mathrm{d}\boldsymbol{f} \qquad \zeta \in \mathbb{R} \tag{260}$$

$$\boldsymbol{\rho}_{x,y}^{x',y'} = \boldsymbol{\rho}_{x',y'}^{x,y} \equiv \frac{1}{2}\left(\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \left(2\lambda_{x,y}^{x',y'} - 1\right)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} + \tau_{x,y}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\right) \qquad (x',y') \in \mathbf{c}_{x,y} \tag{261}$$

where

$$\lambda_{x,y}^{x',y'} = \lambda_{x',y'}^{x,y} \equiv \int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}\left(\boldsymbol{L}_{x,y}^{x',y'}\boldsymbol{\rho}(\boldsymbol{f})\right)\mathrm{d}\boldsymbol{f} \tag{262}$$

$$\tau_{x,y}^{x',y'} = \tau_{x',y'}^{x,y} \equiv \int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}\left(\boldsymbol{K}_{x,y}^{x',y'}\boldsymbol{\rho}(\boldsymbol{f})\right)\mathrm{d}\boldsymbol{f}. \tag{263}$$

In the mean-field approximation, the probability density matrix $\boldsymbol{\rho}(\boldsymbol{f})$ is approximately expressed in terms of the marginal probability density function and the marginal probability density matrix:

$$\boldsymbol{\rho}(\boldsymbol{f}) \simeq \left(\prod_{(x,y)\in\Omega} \rho_{x,y}(f_{x,y})\right)\left(\prod_{(x,y)\in\Omega} \bigotimes \left(\boldsymbol{\rho}_{x,y}^{x+1,y} \bigotimes \boldsymbol{\rho}_{x,y}^{x,y+1}\right)\right). \tag{264}$$

By substituting equation (264) into equation (259), the free energy $\mathcal{F}[\boldsymbol{\rho}]$ can be expressed as $\mathcal{F}\left[\left\{\rho_{x,y}, \boldsymbol{\rho}_{x,y}^{x+1,y}, \boldsymbol{\rho}_{x,y}^{x,y+1}\right\}\right]$ in terms of $\rho_{x,y}(f_{x,y})$, $\boldsymbol{\rho}_{x,y}^{x+1,y}$ and $\boldsymbol{\rho}_{x,y}^{x,y+1}$. By taking the first variation of the approximate free energy $\mathcal{F}\left[\left\{\rho_{x,y}, \boldsymbol{\rho}_{x,y}^{x+1,y}, \boldsymbol{\rho}_{x,y}^{x,y+1}\right\}\right]$ with respect to $\rho_{x,y}(f_{x,y})$, $\boldsymbol{\rho}_{x,y}^{x+1,y}$ and $\boldsymbol{\rho}_{x,y}^{x,y+1}$, we obtain the simultaneous recursion formulae for $\rho_{x,y}(f_{x,y})$ as follows:

$$\rho_{x,y}(f_{x,y}) = \frac{1}{\sqrt{2\pi}\sigma_{x,y}} \exp\left(-\frac{(f_{x,y}-\mu_{x,y})^2}{2{\sigma_{x,y}}^2}\right) \tag{265}$$

$$\mu_{x,y} = \frac{\frac{1}{2\sigma^2}g_{x,y} + \alpha\sum_{(x',y')\in\mathbf{c}_{x,y}}\left(1-\lambda_{x,y}^{x',y'}\right)\mu_{x',y'}}{\frac{1}{2\sigma^2} + \alpha\sum_{(x',y')\in\mathbf{c}_{x,y}}\left(1-\lambda_{x,y}^{x',y'}\right)} \tag{266}$$

$$\frac{1}{2{\sigma_{x,y}}^2} = \frac{1}{2\sigma^2} + \alpha\sum_{(x',y')\in\mathbf{c}_{x,y}}\left(1-\lambda_{x,y}^{x',y'}\right) \tag{267}$$

$$\lambda_{x,y}^{x',y'} = \lambda_{x',y'}^{x,y} = \frac{\mathrm{Tr}\left(\boldsymbol{L}_{x,y}^{x',y'}\exp\left(-\boldsymbol{E}_{x,y}^{x',y'}\right)\right)}{\mathrm{Tr}\left(\exp\left(-\boldsymbol{E}_{x,y}^{x',y'}\right)\right)} \qquad (x',y') \in \mathbf{c}_{x,y} \tag{268}$$

Table 10. Values of $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} (dB) in the restored image $\hat{\boldsymbol{f}}$ obtained for the degraded image $\boldsymbol{g}^*$ given in figure 20(*a*) by the conventional coupled Markov random field model. We set $\sigma = 30$. For each fixed value of γ, the hyperparameters α are determined so as to maximize the evidence which is calculated by the mean-field approximation.

γ	α	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$	Δ_{SNR} (dB)
20	0.001 9069	427.11	3.0407
30	0.001 2585	259.24	5.2091
40	0.001 1022	204.24	6.2447
50	0.001 0518	191.10	6.5336
60	0.001 2570	211.31	6.0970
70	0.001 4596	225.07	5.8230

where

$$\boldsymbol{E}_{x,y}^{x',y'} \equiv \begin{pmatrix} 0 & -h \\ -h & \alpha\Big(\sigma_{x,y}^2 + \sigma_{x',y'}^2 + (\mu_{x,y} - \mu_{x',y'})^2 - \gamma^2\Big) \end{pmatrix}. \tag{269}$$

In these equations, $\mu_{x,y}$ and $\sigma_{x,y}$ mean the average of $F_{x,y}$ and $(F_{x,y} - \mu_{x,y})^2$:

$$\mu_{x,y} \equiv \int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}(f_{x,y}\boldsymbol{\rho}(\boldsymbol{f}))\,\mathrm{d}\boldsymbol{f} = \int_{\mathbb{R}} f_{x,y}\,\rho_{x,y}(f_{x,y})\,\mathrm{d}f_{x,y} \tag{270}$$

$$\sigma_{x,y}^2 \equiv \int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}((f_{x,y} - \mu_{x,y})^2\boldsymbol{\rho}(\boldsymbol{f}))\,\mathrm{d}\boldsymbol{f} = \int_{\mathbb{R}} (f_{x,y} - \mu_{x,y})^2\,\rho_{x,y}(f_{x,y})\,\mathrm{d}f_{x,y}. \tag{271}$$

The evidence can be expressed in terms of the energy matrices $\boldsymbol{E}(\boldsymbol{f}|\boldsymbol{g}, \alpha, \gamma, \sigma)$ and $\boldsymbol{U}(\boldsymbol{f}|\alpha, \gamma)$

$$\begin{aligned}\Pr\{\boldsymbol{G} = \boldsymbol{g}|\alpha, \gamma, \sigma\} = &\int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}(\exp(-\boldsymbol{E}(\boldsymbol{f}|\boldsymbol{g}, \alpha, \gamma, \sigma))\,\mathrm{d}\boldsymbol{f} \\ &- \int_{\mathbb{R}^{|\Omega|}} \mathrm{Tr}(\exp(-\boldsymbol{U}(\boldsymbol{f}|\alpha, \gamma))\,\mathrm{d}\boldsymbol{f} - \sqrt{2\pi}\sigma\end{aligned} \tag{272}$$

and can be calculated by applying the mean-field approximation to the *a priori* probability density matrix and the *a posteriori* probability density matrix, respectively.

5.3. Numerical experiments

In this section, we give some numerical experiments of image restoration by both conventional and quantized coupled Markov random field models. The numerical experiments are carried out for the degraded images $\boldsymbol{g}^*$ in figure 20. The degraded image is generated from the original image $\boldsymbol{f}^*$ by using the additive white Gaussian noise $\mathcal{N}[0, 30^2]$. For each fixed value of γ, we set $\sigma = 30$ and determine the hyperparameters α so as to maximize the evidence which is calculated by the mean-field approximation. We give in tables 10, 11, 12 and 13 the values of α, $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} (dB) in the restored image $\hat{\boldsymbol{f}}$ obtained for the degraded image $\boldsymbol{g}^*$ given in figure 20 by the conventional coupled Markov random field model and the quantized coupled Markov random field model. The restored images $\hat{\boldsymbol{f}}$ are shown in figures 25 and 26. From these results, we see that the quantized line field improves the quality of the restored images.

5.4. Concluding remarks

In this section, we have explained the formulation of the coupled Markov random field models which have a line field. In the conventional coupled Markov random field model, the edge state

Table 11. Values of $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} (dB) in the restored image $\hat{\boldsymbol{f}}$ obtained for the degraded image $\boldsymbol{g}^*$ given in figure 20(*b*) by the conventional coupled Markov random field model. We set $\sigma = 30$.

γ	α	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$	Δ_{SNR} (dB)
20	0.001 8513	517.27	2.2467
30	0.001 2812	370.75	3.6931
40	0.001 0894	316.56	4.3794
50	0.001 0456	301.82	4.5864
60	0.001 1251	304.62	4.5462
70	0.001 3375	312.96	4.4290

Table 12. Values of $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} (dB) in the restored image $\hat{\boldsymbol{f}}$ obtained for the degraded image $\boldsymbol{g}^*$ given in figure 20(*a*) by the quantized coupled Markov random field model. We set $\sigma = 30$ and $h = 1.5$.

γ	α	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$	Δ_{SNR} (dB)
20	0.001 9555	164.87	7.1746
30	0.001 1505	155.17	7.4380
40	0.000 7568	165.54	7.1572
50	0.000 5437	182.84	6.7255
60	0.000 4151	203.23	6.2663
70	0.000 3304	225.16	5.8212

Table 13. Values of $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$ and Δ_{SNR} (dB) in the restored image $\hat{\boldsymbol{f}}$ obtained for the degraded image $\boldsymbol{g}^*$ given in figure 20(*b*) by the quantized coupled Markov random field model. We set $\sigma = 30$ and $h = 1.5$.

γ	α	$d(\boldsymbol{f}^*, \hat{\boldsymbol{f}})$	Δ_{SNR} (dB)
20	0.002 0402	288.30	4.7855
30	0.001 2304	257.39	5.2779
40	0.000 8027	245.67	5.4804
50	0.000 5678	248.18	5.4362
60	0.000 4283	258.36	5.2617
70	0.000 3381	272.60	5.0286

has two possibilities, the 'edge' state and 'no edge' state. The present author has extended it to the coupled Markov random field model with quantized line field and has shown that the quality in restored images can be improved.

The present author also applied the cluster variation method to the conventional coupled Markov random field model and showed that the quality of the restored image can be improved [93]. In the mean-field approximation, the *a posteriori* probability distribution is approximately expressed in terms of the product of the marginal probability distributions of each pixel and each edge as shown in equation (236) and it is difficult to treat the correlations between the nearest-neighbour pairs of pixels. In the cluster variation method, the *a posteriori* probability distribution is assumed to be approximately expressed in the following

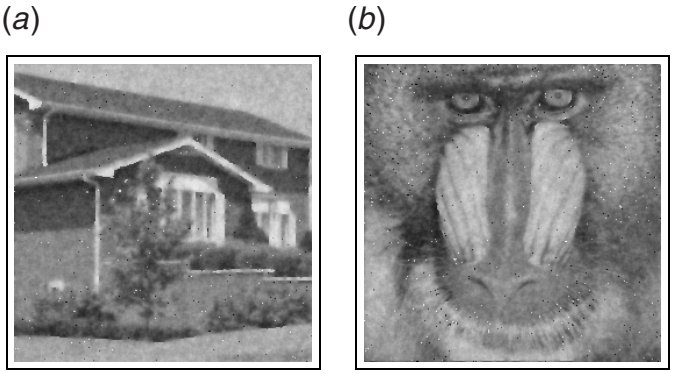

Figure 25. Restored images $\hat{\boldsymbol{f}}$ obtained by the conventional coupled Markov random field model. The degraded images $\boldsymbol{g}^*$ are given in figure 20. (*a*) 'Home' ($\sigma = 30, \gamma = 50, \alpha = 0.001\,0518$); (*b*) 'mandrill' ($\sigma = 30, \gamma = 50, \alpha = 0.001\,0456$).

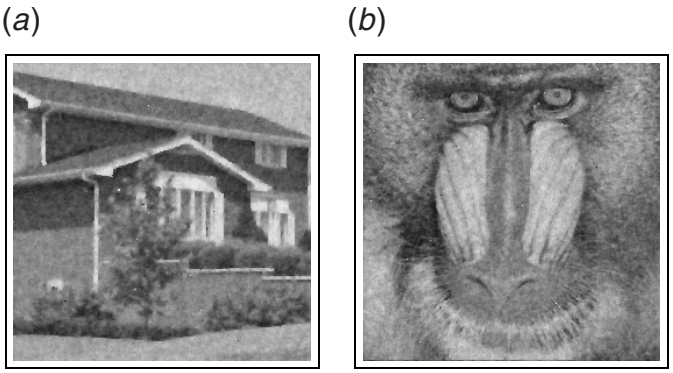

Figure 26. Restored images $\hat{\boldsymbol{f}}$ obtained by the quantized coupled Markov random field model. The degraded images $\boldsymbol{g}^*$ are given in figure 20. (*a*) 'Home' ($\sigma = 30, \gamma = 30, h = 1.5, \alpha = 0.001\,1505$); (*b*) 'mandrill' ($\sigma = 30, \gamma = 40, h = 1.5, \alpha = 0.000\,8027$).

factorizable form:

$$\rho(\boldsymbol{f}, \boldsymbol{l}) \simeq \left(\prod_{(x,y)\in\Omega} \rho_{x,y}(f_{x,y}) \right) \left(\prod_{(x,y)\in\Omega} \frac{\rho_{x,y}^{x+1,y}\left(f_{x,y}, l_{x,y}^{x+1,y}, f_{x+1,y}\right)}{\rho(f_{x,y})\rho(f_{x+1,y})} \right)$$
$$\times \left(\prod_{(x,y)\in\Omega} \frac{\rho_{x,y}^{x,y+1}\left(f_{x,y}, l_{x,y}^{x,y+1}, f_{x,y+1}\right)}{\rho(f_{x,y})\rho(f_{x,y+1})} \right) \tag{273}$$

where $\rho_{x,y}(f_{x,y})$ is given in equation (234) and $\rho_{x,y}^{x',y'}(\zeta, l, \zeta')$ is defined as follows:

$$\rho_{x,y}^{x',y'}(\zeta, l, \zeta') \equiv \int_{\mathbb{R}^{|\Omega|}} \sum_{\boldsymbol{l}} \rho(\boldsymbol{f}, \boldsymbol{l})\, \delta(f_{x,y} - \zeta) \delta_{l, l_{x,y}^{x',y'}} \delta(f_{x',y'} - \zeta')\, \mathrm{d}\boldsymbol{f}$$
$$(x', y') \in \mathbf{c}_{x,y} \quad l = 0, 1 \quad \zeta, \zeta' \in \mathbb{R}. \tag{274}$$

In equation (273), a decoupling of neighbour pixels is unnecessary in calculating the correlation between the nearest-neighbour pairs of pixels. The cluster variation method can improve the

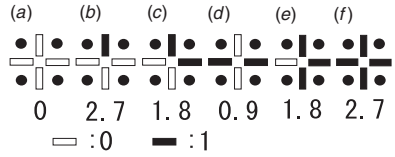

Figure 27. Interactions $V(\boldsymbol{l})$ in the line field of the coupled Markov random field model. All solid circles mean pixels. The open rectangles and the solid rectangles correspond to the 'no edge state' and the 'edge state,' respectively.

image quality of image restoration by the coupled Markov random field model. In the present section, we have demonstrated the coupled Markov random field model with quantized line fields in order to excite physicists' curiosity, but the conventional coupled Markov random field model may also improve the quality of the restored image by adopting more advanced mean-field approximations such as cluster variation method.

In practical applications, the coupled Markov random field model with interactions in the line field is employed. Its energy function is given as

$$E(\boldsymbol{f}, \boldsymbol{l}|\boldsymbol{g}, \alpha, \omega, \sigma) \equiv \frac{1}{2\sigma^2} \sum_{(x,y)\in\Omega} (f_{x,y} - g_{x,y})^2 + \frac{1}{2}\alpha \sum_{(x,y)\in\Omega} \left(\left(1 - l_{x,y}^{x+1,y}\right) (f_{x,y} - f_{x+1,y})^2 + \left(1 - l_{x,y}^{x,y+1}\right) (f_{x,y} - f_{x,y+1})^2 \right) + \omega V(\boldsymbol{l}). \tag{275}$$

Here $V(\boldsymbol{l})$ denotes interactions in the line field $\boldsymbol{l}$ and is illustrated in figure 27. The coupled Markov random field model was proposed by Geman and Geman [6] and Jeng and Woods [8], and the algorithm by the mean-field approximation was derived by Zhang [19]. The interaction term $V(\boldsymbol{l})$ in the line field $\boldsymbol{l}$ has a very complicated structure. Figure 27(*a*) means that the edge configuration associated with the nearest-neighbour square plaquette has no edges. Figure 27(*d*) means that the line passes straight through the nearest-neighbour square plaquette. Figures 27(*b*) and (*c*) correspond to the line terminating and turning at the nearest-neighbour square plaquette, respectively. The value of $V(\boldsymbol{l})$ corresponds to the energy value for each configuration of the edges in the nearest-neighbour square plaquette and is determined *ad hoc*. These values have been given in [6, 8]. If we focus only on the edge configuration, the model (275) corresponds to a two-dimensional Ising model with four-body interactions. In statistical mechanics, many two-dimensional Ising models with multi-body interactions have been investigated. It is interesting to study the critical phenomena of the model (275) from the standpoint of statistical mechanics and to clarify the relationship between the edge structure in grey-level images and critical phenomena.

In section 5.2, we have reviewed the coupled Markov random field model with quantized line states as one of the extensions of the conventional Markov random field model (see figure 27). On the other hand, some physicists [98, 99] also proposed a coupled Markov random field model with continuous line fields from the standpoint of the plane rotator model in statistical mechanics and investigated how the model gives better quality of restored image. However, the coupled Markov random field model with continuous line fields has not been investigated sufficiently yet and many open problems still remain, for example the estimation of hyperparameters.

Before closing this section, we compare the coupled Markov random field model with some conventional filters in image processing. In the conventional image processing, the low-pass filter and median filter are applied to the reduction of noise. In figures 29 and 30, we show the restored images $\hat{\boldsymbol{f}}$ obtained by applying the 3×3 low-pass filter and the

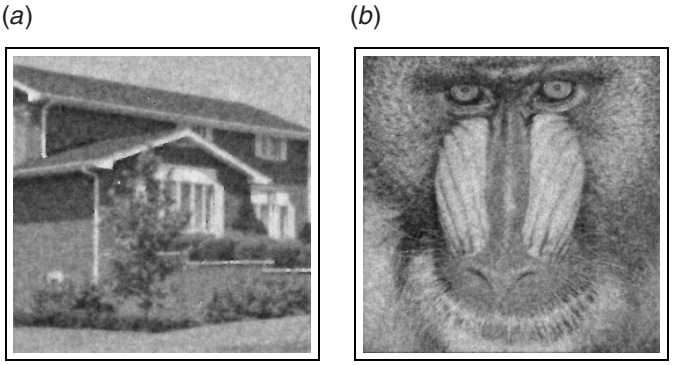

Figure 28. Restored images $\hat{\boldsymbol{f}}$ obtained by means of the conventional coupled Markov random field model given in equation (275). The degraded images $\boldsymbol{g}^*$ are given in figure 20. (*a*) 'Home' ($\sigma = 30$, $\alpha = 0.000\,603\,26$, $\omega = 1.75$, $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 171.58$); (*b*) 'mandrill' ($\sigma = 30$, $\alpha = 0.000\,403\,85$, $\omega = 1.75$, $d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 250.23$).

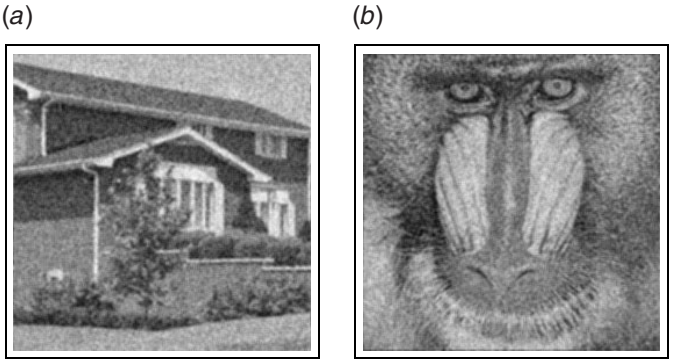

Figure 29. Restored images $\hat{\boldsymbol{f}}$ obtained by means of the 3×3 low-pass filter given in equation (4). The degraded images $\boldsymbol{g}^*$ are given in figure 20. (*a*) 'Home' ($d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 173.18$); (*b*) 'mandrill' ($d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 312.10$).

3×3 median filter to the degraded images $\boldsymbol{g}^*$ given in figure 20. The 3×3 low-pass filter and the median filter have already been explained briefly in section 1 and are given as follows:

(i) 3×3 low-pass filter: $\hat{f}_{x,y} = \frac{1}{9}\sum_{x'=x-1}^{x+1}\sum_{y'=y-1}^{y+1} g_{x',y'}$.

(ii) 5×5 low-pass filter: $\hat{f}_{x,y} = \frac{1}{25}\sum_{x'=x-2}^{x+2}\sum_{y'=y-2}^{y+2} g_{x',y'}$.

(iii) 3×3 median filter: $\hat{f}_{x,y} = \mathrm{med}\{g_{x',y'} | x' = x-1, x, x+1, y' = y-1, y, y+1\}$.

(iv) 5×5 median filter: $\hat{f}_{x,y} = \mathrm{med}\{g_{x',y'} | x-2 \leqslant x' \leqslant x+2, y-2 \leqslant y' \leqslant y+2\}$.

Here, the symbols 3×3 and 5×5 refer to the size of window in the filters. It is obvious that these methods cannot erase the noise to a satisfactory level, if we take a 3×3 low-pass filter and a 3×3 median filter. On the other hand, in the 5×5 low-pass filter and the 5×5 median filter, although the noise can be erased sufficiently, the restored image has been obtained as a blurred image (see figures 31 and 32). It is obvious that the results obtained by the coupled Markov random field model are better than those with conventional filters.

(*a*) (*b*)

Figure 30. Restored images $\hat{\boldsymbol{f}}$ obtained by means of the 3×3 median filter. The degraded images $\boldsymbol{g}^*$ are given in figure 20. (*a*) 'Home' ($d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 216.93$); (*b*) 'mandrill' ($d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 369.07$).

(*a*) (*b*)

Figure 31. Restored images $\hat{\boldsymbol{f}}$ obtained by means of the 5×5 low-pass filter given in equation (4). The degraded images $\boldsymbol{g}^*$ are given in figure 20. (*a*) 'Home' ($d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 235.12$); (*b*) 'mandrill' ($d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 380.78$).

(*a*) (*b*)

Figure 32. Restored images $\hat{\boldsymbol{f}}$ obtained by means of the 5×5 median filter. The degraded images $\boldsymbol{g}^*$ are given in figure 20. (*a*) 'Home' ($d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 222.81$); (*b*) 'mandrill' ($d(\boldsymbol{f}^*, \hat{\boldsymbol{f}}) = 400.09$).

6. Summary and discussions

In the present review, we discussed probabilistic image restoration techniques using Bayesian statistics and statistical mechanics. We explained some detailed derivations of the recursion formulae to determine the restored image and the hyperparameters and gave the explicit algorithms. Our main topics are binary image restoration for the degraded images generated by the binary symmetric channel and grey-level image restoration under additive white Gaussian noise.

In binary image restoration, the relationship with the Ising model has been clarified and the algorithms to determine the restored image and the hyperparameters are derived from the mean-field and Bethe approximations. It is shown that the results in the Bethe approximation are absolutely better than those in the mean-field approximation when we have to estimate the hyperparameters from the given degraded image by means of the evidence framework, which is a familiar technique to determine the hyperparameters in statistics. However, it should be remarked that such a difference between the mean-field approximation and the Bethe approximation appears only if we determine hyperparameters from the given degraded image by the maximization of evidence. In other words, the accuracy of evidence is strongly dependent on the approximation. If the hyperparameters have already been estimated with high accuracy by a different method or if the original values of the hyperparameters are known, all we have to do is to estimate only the restored image. In such a situation, although the Bethe approximation can supply us a better restored image than the mean-field approximation, the difference in the restored images is not as large as when the hyperparameters are determined in the maximization of the evidence. Clearly, the most important factor in Bayesian image restoration is the selection of the *a priori* probability. If we have chosen an *a priori* probability that does not fit our treated images at all, it is difficult to obtain a restored image of high quality however accurate the approximation we adopt. On the other hand, not much different image is given depending on the choice of hyperparameter values. This means that probabilistic image processing by Bayesian statistics has robustness against errors in the hyperparameter estimation. Although robustness for different choices of *a priori* probabilities is indeed a more interesting problem than that for hyperparameters within the same *a priori* probability, such an investigation has not yet been carried out very extensively.

As for grey-level image restoration, we first demonstrated the determination of the restored image and hyperparameters by a solvable probabilistic model. The solvable probabilistic model corresponds to the Gaussian model in statistical mechanics and the exact expressions of some statistical quantities are derived by the discrete Fourier transform. However, the Gaussian model is not sufficient to erase the noise in the degraded image. If we choose a large value of the hyperparameter for smoothing, some edges are erased in addition to noise and the restored image becomes blurred.

In order to improve this disadvantage of the probabilistic image restoration scheme by the Gaussian model, the line field is sometimes employed. With the line field introduced, it is hard to treat the probabilistic model and the coupled Markov random field model exactly. In the present review, we introduced the algorithm constructed by the mean-field approximation. We also demonstrated the coupled Markov random field model with the quantized line field. Although the intensity value at each pixel is observable in the degraded image, the edge state between the nearest-neighbour pairs of pixels is unobservable and subjective. From physical interests, we illustrated the scheme of image restoration by the quantized line field although we do not have any *a priori* reason to introduce the quantum-mechanical state as an edge state. The quantized line field was introduced as a state expressed in terms of the superposition of the edge state and the no edge state, such that

$$|l_{x,y}\rangle = A|l_{x,y} = 1\rangle + B|l_{x,y} = 0\rangle. \tag{276}$$

As shown in some numerical experiments, such a quantized line field can lead to the improvement of quality in the restored image, mainly due to the achievement of smoothing and noise reduction without erasing the edges.

In the present review, we demonstrated the hyperparameter determination by maximization of the evidence. As for filters, we explained the constrained least mean square filter (169)–(172), a typical basic linear filter [1, 80]. By extending the constrained least mean square filter, other frameworks for hyperparameter determination in the Markov random field model have also been proposed [94–96]. The frameworks are generally referred to as constrained optimization. In particular, Geman *et al* [94] introduced some forbidden local patterns in the line image $\boldsymbol{l} = \left\{l_{x,y}^{x+1,y}, l_{x,y}^{x,y+1} | (x, y) \in \Omega\right\}$ determined by the line variables, $l_{x,y}^{x+1,y}$ and $l_{x,y}^{x,y+1}$, and imposed the constraints that the total number $\mathcal{K}(\boldsymbol{l})$ of those forbidden local patterns should be equal to zero in the line image as follows:

$$\hat{\boldsymbol{l}} = \arg \min_{\boldsymbol{l}:\mathcal{K}(\boldsymbol{l})=0} \sum_{(x,y)\in\Omega} \mathcal{D}(\boldsymbol{l}|\boldsymbol{g}) \tag{277}$$

where $\mathcal{D}(\boldsymbol{l}|\boldsymbol{g})$ is the distance between the line image $\boldsymbol{l}$ and the given observed image $\boldsymbol{g} = \{g_{x,y} | (x, y) \in \Omega\}$. Selections of the definitions in $\mathcal{K}(\boldsymbol{l})$ and $\mathcal{D}(\boldsymbol{l}|\boldsymbol{g})$ are dependent on what type of edge detection we desire. By introducing the Lagrange multiplier γ to ensure the constraint $\mathcal{K}(\boldsymbol{l}) = 0$, we can derive the following energy function

$$H(\boldsymbol{l}|\boldsymbol{g}) = \mathcal{D}(\boldsymbol{l}|\boldsymbol{g}) + \gamma\mathcal{K}(\boldsymbol{l}) \tag{278}$$

where the second term can be regarded as a penalty term. Their framework can also be regarded as an extension of the standard regularization theory proposed by Poggio *et al* [97] and corresponds to searching for the ground state configuration of an energy function with the penalty term to exclude forbidden patterns. They introduced the Gibbs distribution in order to adopt an annealing procedure to search for the ground state configuration. Morita and Tanaka [96] gave a similar formulation for the binary image restoration as follows:

$$\hat{\boldsymbol{f}} = \arg \min_{\boldsymbol{z}:\mathcal{J}(\boldsymbol{z})=\mathcal{J}(\boldsymbol{f})} \sum_{(x,y)\in\Omega} d(\boldsymbol{f}, \boldsymbol{g}) \tag{279}$$

where $d(\boldsymbol{f}, \boldsymbol{g})$ is given by equation (163) and $\mathcal{J}(\boldsymbol{z})$ is defined as

$$\mathcal{J}(\boldsymbol{z}) \equiv \sum_{(x,y)\in\Omega} ((z_{x,y} - z_{x+1,y})^2 + (z_{x,y} - z_{x,y+1})^2). \tag{280}$$

The original image is denoted by $\boldsymbol{f}$ and $\mathcal{J}(\boldsymbol{f})$ is the number of nearest-neighbour pairs of pixels with different grades of each other in the binary image. By introducing the Lagrange multiplier γ to ensure the constraint condition $\mathcal{J}(\boldsymbol{z}) = \mathcal{J}(\boldsymbol{f})$, we can derive the following energy function:

$$H(\boldsymbol{z}|\boldsymbol{g}, \gamma) = d(\boldsymbol{z}, \boldsymbol{g}) + \gamma(\mathcal{J}(\boldsymbol{z}) - \mathcal{J}(\boldsymbol{f})). \tag{281}$$

Here, we have to estimate the number $\mathcal{J}(\boldsymbol{f})$ of nearest-neighbour pairs of pixels with different grades of each other in the original image $\boldsymbol{f}$ only from the given degraded image $\boldsymbol{g}$. We also gave the estimation framework in [96]. For binary image restoration, $H(\boldsymbol{z}|\boldsymbol{g}, \gamma = \alpha/\beta)$ is equivalent to $E(\boldsymbol{z}|\boldsymbol{g}, \alpha, \beta)/\beta$ defined by equation (80) except for a constant factor $-\alpha\mathcal{J}(\boldsymbol{f})/\beta$.

We can calculate the configuration as well as thermal averages for some quantities exactly in the Gaussian model as demonstrated in section 4. The configuration average is possible if we assume all the ideal generation processes of original images from the *a priori* probability

density function and all the ideal degradation processes. In the present review, we demonstrated the calculation of the configuration average of the mean square error between original images and their corresponding restored images using the *a priori* probability density function and the conditional probability density function of the degradation process with the original values of the hyperparameters. By using the results, we can estimate the statistical performance when the probabilistic image restoration scheme by the Gaussian model is applied to degraded images generated through the ideal original-image generating process and the ideal degradation process. Moreover, by calculating the configuration average of the evidence, we can obtain the statistical behaviour of the iteration process in the hyperparameter determination algorithm based on the maximization of the evidence. From these results obtained in the configuration average procedure, we know some important statistical performance of the proposed scheme without heavy numerical experiments.

It is well known that the configuration average is an important procedure in spin glass theory. Now it may occur to statistical physicists that the general framework for the design of the statistical performance estimation system for any probabilistic image processing systems may be formulated by results in spin glass theory. Such systematic investigations have not yet been done in practical image processing systems. In spin glass theory, many methods have been proposed to calculate the configuration average in the random classical spin systems with finite-range interactions including the two-dimensional $\pm J$ model based on the Bethe approximation [63–67]. I believe that those methods are applicable to the statistical performance estimation of probabilistic image processing systems although we have to note that the probabilistic model to treat image restoration is a correlated random-field model, which is a little different from spin glasses. This is a challenging problem.

Recently, many statistical physicists have participated in the investigation of probabilistic image processing [32]. Renormalization techniques have also been applied to probabilistic image restoration [100]. However, most of these investigations are directed towards applications of statistical-mechanical techniques to probabilistic image processing. I believe that the most important thing we, the physicists, have to aim at is to investigate the deep origin or nature of probabilistic image processing. Evidently, interactions of the classical spin systems play an important role in probabilistic image processing. Besides, most image processing systems including conventional filters work well in large-scale systems. This fact means that the cooperative phenomena in statistical mechanics play a very important role.

Although we have elucidated the probabilistic approaches to image restoration in the present review, we also have the problem of image segmentation as the other basic issue in probabilistic image processing. Before closing this review, we briefly summarize image segmentation as an example of applications of probabilistic models and statistical-mechanical techniques to other image processing problems. Image segmentation is the problem of classifing a given grey-level image into a few different areas. Image segmentation is often applied to object recognition as a pre-processing step and is regarded as an important information processing procedure for industrial and medical applications. Now we denote the observed image by $\boldsymbol{g} = \{g_{x,y}|(x, y) \in \Omega\}$ and consider classifying the observed image $\boldsymbol{g}$ to K different kinds of levels, $k = 0, 1, 2, \ldots, K-1$. In the segmented image, the set of all possible states at each pixel is denoted by $\{0, 1, 2, \ldots, K-1\}$. A most typical *a posteriori* probability distribution for the segmented image $\boldsymbol{f} = \{f_{x,y}|(x, y) \in \Omega\}$ is given as follows,

$$\rho(\boldsymbol{f}|\boldsymbol{g}, \alpha, \beta, \{\gamma_k\}) = \frac{\exp\left(\beta\delta_{f_{x,y},\phi(g_{x,y})} + \alpha\delta_{f_{x,y},f_{x+1,y}} + \alpha\delta_{f_{x,y},f_{x,y+1}}\right)}{\sum_z \exp\left(\beta\delta_{z_{x,y},\phi(g_{x,y})} + \alpha\delta_{z_{x,y},z_{x+1,y}} + \alpha\delta_{z_{x,y},z_{x,y+1}}\right)} \tag{282}$$

where

$$\phi(g_{x,y}) \equiv \arg \min_{k=0,1,2,\ldots,K-1} (\gamma_k - g_{x,y})^2. \tag{283}$$

This probabilistic model is just a Potts model with non-uniform external fields and nearest-neighbour interactions. Using this *a posteriori* probability distribution and the mean-field approximation, the probabilistic image segmentation algorithm can be constructed. Zhang [26, 27] investigated the scheme including hyperparameter estimation by maximization of the evidence. He employed the expectation-maximization procedure to maximize the evidence and combined the expectation-maximization algorithm with the mean-field approximation. More complicated probabilistic models for image segmentation were also proposed [24, 94, 101]. Moreover, probabilistic approaches based on Bayesian statistics and statistical mechanics were also applied to object detection [102, 103], motion detection [104], decoding of compressed images [105, 106] and colour image processing [87, 107, 108].

Acknowledgments

The author is grateful to Professor T Horiguchi of the Graduate School of Information Science, Tohoku University, Professor H Nishimori of the Department of Physics, Tokyo Institute of Technology, Professor J Inoue of the Graduate School of Engineering, Hokkaido University for valuable discussions. This work was partly supported by the grants-in-aid (no 13 680 384) for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Rosenfield A and Kak A C 1976 *Digital Picture Processing* (New York: Academic)
[2] Gonzales R C and Woods R E 1992 *Digital Image Processing* (Reading, MA: Addison-Wesley)
[3] Pitas I and Venetsanopoulos A N 1990 *Nonlinear Digital Filters* (Dordrecht: Kluwer)
[4] Derin H, Elliott H, Cristi R and Geman D 1984 *IEEE Trans. Pattern Anal. Mach. Intell.* **6** 707
[5] Geman D 1990 *Random Fields and Inverse Problems in Imaging (Lecture Notes in Mathematics vol 1427)* (Berlin: Springer) p 113
[6] Geman S and Geman D 1984 *IEEE Trans. Pattern Anal. Mach. Intell.* **6** 721
[7] Jeng F C and Woods J W 1990 *IEEE Trans. Inf. Theory* **36** 94
[8] Jeng F C and Woods J W 1991 *IEEE Trans. Signal Process.* **39** 683
[9] Chellappa R and Jain A (ed) 1993 *Markov Random Fields: Theory and Applications* (New York: Academic)
[10] Li S Z 1995 *Markov Random Field Modeling in Computer Vision* (Berlin: Springer)
[11] Domb C 1960 *Adv. Phys.* **9** 245
[12] Feller W 1950 *An Introduction to Probability Theory and Its Applications* (New York: Wiley)
[13] Lindley D V 1965 *Introduction to Probability and Statistics from a Bayes View-Point* (Cambridge: Cambridge University Press)
[14] Papoulis A 1984 *Probability and Random Variables, and Stochastic Processes* (New York: McGraw-Hill)
[15] Stirzakre D 1999 *Probability and Random Variables* (Cambridge: Cambridge University Press)
[16] Geiger D and Girosi F 1991 *IEEE Trans. Pattern Anal. Mach. Intell.* **13** 401
[17] Bilbro G L, Snyder W E, Garnier S J and Gault J W 1992 *IEEE Trans. Neural Networks* **3** 131
[18] Zerubia J and Chellappa R 1993 *IEEE Trans. Neural Networks* **4** 703
[19] Zhang J 1996 *IEEE Trans. Image Process.* **5** 1208
[20] Marroquin J, Mitter S and Poggio T 1987 *J. Am. Stat. Assoc.* **82** 76
[21] Besag J 1986 *J. R. Stat. Soc. Ser.* B **48** 259
[22] Qian W and Titterington D M 1989 *J. Appl. Stat.* **16** 267
[23] Qian W and Titterington D M 1992 *J. Stat. Comput. Simul.* **40** 55
[24] Lakshmanan S and Derin H 1989 *IEEE Trans. Pattern Anal. Mach. Intell.* **11** 799
[25] Iba Y 1991 *Proc. Inst. Stat. Math.* **39** 1 (in Japanese)
[26] Zhang J 1992 *IEEE Trans. Image Process.* **40** 2570

[27] Zhang J, Modestino J W and Langan D A 1994 *IEEE Trans. Image Process.* **3** 404
[28] Zhou Z, Leahy R M and Qi J 1997 *IEEE Trans. Image Process.* **6** 844
[29] Pryce J M and Bruce A D 1995 *J. Phys. A: Math. Gen.* **28** 511
[30] Tanaka K 2000 *IEICE Trans.* A **J83** 1148
(Engl. transl. *Electron. Commun. Japan* **85** 50)
[31] Tanaka K 2001 *Trans. Japan. Soc. Artif. Intell.* **16** 246
[32] Nishimori H 2001 *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford: Oxford University Press)
[33] Nishimori H and Wong K Y M 1999 *Phys. Rev.* E **60** 132
[34] Inoue J and Tanaka K 2002 *Phys. Rev.* E **65** 016125
[35] Nishimori H 1980 *J. Phys. C: Solid State Phys.* **13** 4071
[36] Morita T and Horiguchi T 1980 *Phys. Lett.* A **76** 424
[37] Nishimori H 1981 *Prog. Theor. Phys.* **66** 1169
[38] Horiguchi T 1981 *Phys. Lett.* A **81** 530
[39] Kabashima Y and Saad D 1999 *Europhys. Lett.* **45** 97
[40] Tanaka T 2001 *Europhys. Lett.* **54** 540
[41] Good I J 1965 *The Estimation of Probabilities* (Cambridge, MA: MIT Press)
[42] Berger J O 1985 *Statistical Decision Theory and Bayesian Analysis* 2nd edn (Berlin: Springer)
[43] Akaike H 1980 *Bayesian Statistics* ed J M Bernardo, M H Degroot, D V Lindley and A F M Smith (Valencia: Valencia University Press) pp 143, 185
[44] MacKay D J C 1992 *Neural Comput.* **4** 415
[45] MacKay D J C 1992 *Neural Comput.* **4** 448
[46] Bruce A D and Saad D 1994 *J. Phys. A: Math. Gen.* **27** 3355
[47] Marion G and Saad D 1996 *J. Phys. A: Math. Gen.* **29** 5387
[48] Nishimori H 1993 *J. Phys. Soc. Japan* **62** 2973
[49] Nishimori H 1994 *Physica* A **205** 1
[50] Iba Y 1999 *J. Phys. A: Math. Gen.* **32** 3875
[51] Carlucci D M and Inoue J 1999 *Phys. Rev.* E **60** 2547
[52] Inoue J 2001 *Phys. Rev.* E **63** 046114
[53] Inoue J and Carlucci D M 2001 *Phys. Rev.* E **64** 036121
[54] Tadaki T and Inoue J 2002 *Phys. Rev.* E **65** 016101
[55] Saika Y and Nishimori H 2002 *J. Phys. Soc. Japan* **71** 1052
[56] Takimoto E 2002 Private communication
[57] Tanaka K and Morita T 1995 *Phys. Lett.* A **203** 122
[58] Fröberg C-E 1985 *Numerical Mathematics: Theory and Computer Applications* (New York: Benjamin–Cummings)
[59] Schwartz H R 1989 *Numerical Analysis: A Comprehensive Introduction* (New York: Wiley)
[60] Horiguchi T, Honda Y and Miya M 1997 *Phys. Lett.* A **227** 319
[61] Tanaka K, Ichioka M and Morita T 1996 *Proc. 13th Int. Conf. on Pattern Recognition* vol 2, Track B (IEEE Computer Society Press) p 381
[62] Tanaka K, Ichioka M and Morita T 1997 *IEICE Trans.* A **J80** 260 (in Japanese)
[63] Morita T 1979 *Physica* A **98** 566
[64] Katsura S and Fujiki S 1979 *J. Phys. C: Solid State Phys.* **12** 1087
[65] Katsura S and Fujiki S 1980 *J. Phys. C: Solid State Phys.* **13** 4711
[66] Fujiki S and Katsura S 1980 *J. Phys. C: Solid State Phys.* **13** 4723
[67] Horiguchi T 1981 *Physica* A **107** 360
[68] Kabashima Y and Saad D 1998 *Europhys. Lett.* **44** 668
[69] Touless D J, Anderson P W and Palmer R G 1977 *Phil. Mag.* **35** 593
[70] Morita T and Horiguchi T 1976 *Solid State Commun.* **19** 833
[71] Pearl J 1988 *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (San Mateo, CA: Morgan Kaufmann)
[72] Lauritzen S L 1995 *Graphical Models* (Oxford: Oxford Science)
[73] Yedidia J S, Freeman W T and Weiss Y 2001 *Advances in Neural Information Processing Systems 13* ed T Leen, T Dietterich and V Tresp (Cambridge, MA: MIT Press) 689
[74] Kappen H J W and Wiegerinck W 2002 *Advances in Neural Information Processing System 14* ed T G Dietterich, S Becker and Z Ghahramani (Cambridge, MA: MIT Press) at press
[75] Kikuchi R 1951 *Phys. Rev.* **81** 988
[76] Morita T 1972 *J. Math. Phys.* **13** 115
[77] Wu C-h and Doerschuk P C 1995 *IEEE Trans. Pattern Anal. Mach. Intell.* **17** 275

[78] Wu C-h and Doerschuk P C 1995 *IEEE Trans. Pattern Anal. Mach. Intell.* **17** 391
[79] Suzuki M 1986 *J. Phys. Soc. Japan* **55** 4205
[80] Hunt B R 1973 *IEEE Trans. Comput.* **22** 805
[81] Parisi G 1988 *Statistical Field Theory* (Reading, MA: Addison-Wesley)
[82] Thompson C J 1988 *Classical Equilibrium Statistical Mechanics* (Oxford: Oxford University Press)
[83] Molina R 1994 *IEEE Trans. Pattern Anal. Mach. Intell.* **16** 1122
[84] Molina R, Katsaggelos A K and Mateos J 1999 *IEEE Trans. Image Process.* **8** 231
[85] Nishimori H 2000 *BUSSEI KENKYU* **73** 850 (in Japanese)
[86] Tanaka K and Inoue J 2002 *IEICE Trans. Inf. Syst.* D **E85** 546
[87] Tanaka K and Horiguchi T 2002 *Phys. Rev.* E **65** 046142
[88] Dunmur A P and Titterington D M 1997 *IEEE Trans. Pattern Anal. Mach. Intell.* **19** 1296
[89] Aykroyd R G 1998 *IEEE Trans. Pattern Anal. Mach. Intell.* **20** 533
[90] Mardia K V, Gill C A and Aykroyd R G (ed) 1997 *Proc. Art and Science of Bayesian Image Analysis* (Leeds: Leeds University Press)
[91] Tanaka K and Horiguchi T 1997 *IEICE Trans.* A **J80** 2117
(Engl. transl. 2000 *Electron. Commun. Japan* **83** 84)
[92] Tanaka K 2001 *IEICE Trans.* D-ll **J84** 737 (in Japanese)
[93] Tanaka K 2001 *Trans. Japan. Soc. Artif. Intell.* **16** 259 (in Japanese)
[94] Geman D, Geman S, Graffigne C and Dong P 1990 *IEEE Trans. Pattern Anal. Mach. Intell.* **12** 609
[95] Geman D and Reynolds G 1992 *IEEE Trans. Pattern Anal. Mach. Intell.* **14** 367
[96] Morita T and Tanaka K 1996 *Physica* A **223** 244
[97] Poggio T, Torre V and Koch C 1985 *Nature* **317** 314
[98] Okada M, Doya K, Yoshioka T and Kawato M 1999 *IEICE Technical Report* NC98-184 239 (in Japanese)
[99] Tanaka K, Furusato D and Horiguchi T 2002 *IEICE Trans.* D-ll **J85** 805 (in Japanese)
[100] Gidas B 1989 *IEEE Trans. Pattern Anal. Mach. Intell.* **11** 164
[101] Derin H and Elliott H 1987 *IEEE Trans. Pattern Anal. Mach. Intell.* **9** 39
[102] Nadabar S G and Jain A K 1996 *IEEE Trans. Pattern Anal. Mach. Intell.* **18** 326
[103] Li S Z 2000 *IEEE Tarns. Image Process.* **9** 273
[104] Zhang J and Hanauer G G 1995 *IEEE Trans. Image Process.* **4** 19
[105] Özcelik T, Brailean J C and Katsaggelos A K 1995 *Proc. IEEE* **83** 304
[106] Mateos J M, Katsaggelos A K and Molina R 2000 *IEEE Trans. Image Process.* **9** 1200
[107] Panjwani D K and Healey G 1995 *IEEE Trans. Pattern Anal. Mach. Intell.* **17** 939
[108] Bennett J and Khotanzad A 1999 *IEEE Trans. Pattern Anal. Mach. Intell.* **21** 537